

Reform and Practice of the Data Acquisition and Preprocessing Course in Digital Intelligence Education

Kun Qin^a, Yan Gong^a, Xiaoliang Meng^{a,b}, Zheng Ji^a, Gang Li^a, Yichun Xie^c, Zhipeng Gui^a, Songtao Ai^d, Wanlin Gong^b, Pengcheng Zhao^a, Longgang Xiang^d, Changhui Yu^a

^a School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China

^b Student Engineering Training and Innovation Practice Center, Wuhan University, Wuhan 430072, China

^c Institute for Geospatial Research and Education, Eastern Michigan University, Ypsilanti, MI 48197, USA

^d Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430072, China

© Higher Education Press 2025

Abstract Data acquisition and preprocessing is a core course on digital intelligence at Wuhan University that is designed to cultivate students' understanding of data sources and improve preprocessing methods. The course aims at fostering digital thinking and literacy and enhancing intelligent computing skills. This study examined digital intelligence education and reform practices integrated into the data acquisition and preprocessing course, which covered web data, social sensing data, remote sensing data, sensor network data, unmanned aerial vehicle data, and 3D data. Moreover, the study explored the development and implementation of the course's teaching platform, which was based on the open geospatial engine.

Keywords digital intelligence education, web data, social sensing data, remote sensing data, sensor network data, unmanned aerial vehicle data, 3D data

1 Introduction

Digital intelligence education is an interdisciplinary talent cultivation model that leverages Big Data and AI technologies as core enablers. Its primary objective is to cultivate and develop students' digital thinking and literacy, intelligent computing skills, and digital competencies to equip them with the necessary skills to address challenges in the digital intelligence era.

Therefore, Wuhan University has established a comprehensive program dedicated to nurturing digital talents (Meng, 2024; Wu, 2024; Zhang, 2024).

The data acquisition and preprocessing course is the first digital intelligence education course in the spatial information science area. It is a general education course that covers theories and technologies related to data acquisition and preprocessing in digital intelligence education. For students in the science and engineering disciplines, such as computer science, electronic information engineering, remote sensing, physics, and chemistry, this course is recommended as an elective course for freshmen and sophomores. For students in the liberal arts and social sciences, such as literature, art, history, and management, it is suggested as an optional course for juniors and seniors without any prerequisites.

Data acquisition and preprocessing course aims at cultivating students' understanding of data sources and improving preprocessing methods while fostering digital intelligence skills. The course officially commenced in the second semester of 2024 and represented a significant step forward in advancing digital intelligence education reform. This report introduces the course's digital intelligence education and reform practices, including the acquisition and preprocessing of various data types based on digital intelligence thinking, such as web data, social sensing data, remote sensing data, sensor network data, unmanned aerial vehicle (UAV) data, and 3D data. It also presents the development and practice of the data acquisition and preprocessing teaching platform, which is based on the Open Geospatial Engine (OGE) platform.

Received January 20, 2025; revised February 10, 2025; accepted February 14, 2025

Yan Gong (✉)

E-mail: gongyan@whu.edu.cn

2 Teaching Objectives and Course Contents of the Data Acquisition and Preprocessing Course

2.1 | Teaching Objectives

Data acquisition and preprocessing course is a general education course primarily designed for undergraduates in digital intelligence talent cultivation, while also catering to the quality education needs of non-data science postgraduates. With three objectives, including moral value shaping, knowledge acquisition, and skill development, the course aims at cultivating students' understanding of data sources and preprocessing methods, enhancing their research skills, and fostering a spirit of scientific exploration and lifelong learning.

The first teaching objective is moral value shaping. The data acquisition and preprocessing course aims at cultivating students' truth-seeking and pragmatism. Data acquisition techniques serve as a fundamental ability to understand and perceive the world. Through learning and practising different data acquisition techniques, students develop greater awareness and the ability to trace information back. Moreover, data preprocessing techniques are essential for improving data quality and forming a solid foundation for subsequent analyses and applications. Students are encouraged to adopt a quality-first mindset and a commitment to excellence (Gong et al., 2024).

The second teaching objective is knowledge acquisition. Students gain a comprehensive understanding of multi-source data and develop proficiency in the acquisition and preprocessing of multiple types of data, such as web data, social sensing data, remote sensing data, sensor networks, UAV data, and 3D data (Qin et al., 2023).

The third teaching objective is skill development. Students develop the technical expertise to acquire and preprocess multi-source data, as well as practical skills in the real world (Lin, 2022; Sheng et al., 2018).

2.2 | Course Content

The data acquisition and preprocessing course explores theories and techniques related to data acquisition and preprocessing in digital intelligence education. The curriculum covers data source analysis and the acquisition and preprocessing of various types of data. Moreover, the course contains four practical sessions, including social sensing data acquisition and preprocessing practice, remote sensing data acquisition

and preprocessing practice, sensor network data acquisition and preprocessing practice, and 3D data acquisition and preprocessing practice.

The course is 2 credits covering 36 hours, with a recommended class size of no more than 30 students. The course covers 12 academic weeks, with 3 one-hour sessions per week. The course is comprised of 7 theoretical chapters and 4 practical sessions, as well as an integrative discussion session in the final week. The first practical session is social sensing data acquisition and preprocessing. Students can complete tasks on the OGE platform remotely using a personal computer. However, the second, third, and fourth practical sessions as mentioned above require access to a remote sensing laboratory or a sensor network laboratory.

Through the theoretical teaching and practical components of this course, students are expected to master the techniques of data acquisition and preprocessing, thereby forming a solid foundation for their subsequent studies in related professional courses. The course content system regarding the relationships between teaching objectives and teaching contents is illustrated in Table 1. For example, *Chapter 1: Data source analysis* introduces multi-source data, which enables students to understand the origin of data and assess the source, which achieves the first teaching objective, namely moral value shaping. Moreover, this also helps students master knowledge about different data types, which achieves the second teaching objective, namely knowledge acquisition.

3 Exploring Digital Intelligence Education Methods

The teaching methods of data acquisition and preprocessing course have been reformed and redesigned in alignment with the principles of digital intelligence education. As a fundamental component of data science education, the curriculum covers different technologies. For example, one topic relates to AI technologies, as remote sensing data acquisition and preprocessing rely on deep learning technologies and sensing network data acquisition and preprocessing apply robot technologies.

The course begins with an analysis of data sources, enabling students to understand the origins of data, followed by the introduction of various methods of data acquisition and preprocessing. Moreover, a teaching platform based on the OGE platform is under development. By integrating digital intelligence education mindset and technologies, a comprehensive reform and exploration of teaching methods has been carried out. The main digital intelligence education methods adopted are summarized below.

Table 1 Teaching content, class hour, teaching approach, and teaching objective

Teaching content	Class hour	Teaching approach	Teaching objective
<i>Chapter 1: Data source analysis</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Chapter 2: Web data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Chapter 3: Social sensing data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Practice 1: Social sensing data acquisition and preprocessing practice</i>	3	Practical training	Objective 3
<i>Chapter 4: Remote sensing data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Practice 2: Remote sensing data acquisition and preprocessing practice</i>	3	Practical training	Objective 3
<i>Chapter 5: Sensor network data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Practice 3: Sensor network data acquisition and preprocessing practice</i>	3	Practical training	Objective 3
<i>Chapter 6: UAV data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Chapter 7: 3D data acquisition and preprocessing</i>	3	Theoretical instruction	Objective 1 Objective 2
<i>Practice 4: 3D data acquisition and preprocessing practice</i>	3	Practical training	Objective 1 Objective 2
<i>Achievement exhibition and discussion</i>	3	Theory and practice	Objective 1 Objective 2 Objective 3

Notes. UAV: unmanned aerial vehicle, Objective 1: moral values shaping, Objective 2: knowledge acquisition, Objective 3: skills development.

3.1 | Understanding Data Origins and Source Issues

Data acquisition and preprocessing is the first step in the data lifecycle. The teaching philosophy of this course is to provide high-quality data sources for digital intelligence education using advanced technologies, addressing issues at the data source, and ensuring data quality. It focuses on understanding data origins and identifying potential issues at six data sources, including web data source, remote sensing data source, social sensing data source, sensor network data source, UAV data source, and 3D data source, as well as other data sources.

First, the web data source mainly refers to web text data in this course. With the rapid development of network technology, the Internet has increasingly become a major source of data. The Internet offers an extensive and diverse repository of data, accessible through various web services, enabling users to obtain almost any data they need.

Second, the remote sensing data source mainly refers to remote sensing satellite images in this course. Remote sensing is a technology for acquiring data by capturing electromagnetic waves from target objects using remote sensors, commonly known as remote sensing cameras, without direct contact, enabling long-distance perception of the targets. With advancements in remote sensing technologies, numerous satellites

orbiting Earth generate large amounts of remote sensing data daily, thus enhancing data accessibility.

Third, the social sensing data source in this course includes socio-economic and human mobility data. The development of remote sensing science and technology has expanded the focus of traditional methods from solely observing the earth to also monitoring human society and social dynamics. Utilizing smartphones, cameras, and various devices equipped with location information, near-real-time social sensing has made it possible to collect data reflecting human activities and societal dynamics.

Fourth, the sensor network data sources are ubiquitous, forming an omnipresent sensor network that can perceive various environmental factors in real-time, such as temperature, humidity, and noise, providing real-time sensor network data.

Fifth, the UAV data source is an online source for data collected on unmanned aircraft and their sub-components. With the development of the low-altitude economy, UAVs are becoming increasingly prevalent. Equipping UAVs with sensors, spatial photography, and scanning can be conducted. Through data processing and reconstruction, the locations and attributes of spatial objects can be obtained.

Sixth, 3D data source provides techniques for students to understand the real world. By using various scanning devices to acquire point cloud data of target objects, 3D models are quickly constructed, allowing for

the creation of dynamic and three-dimensional representations of the real world.

Besides the aforementioned 6 major data sources, there are many other data sources, such as statistical data, survey data, and existing data from other systems, which are also important research objects for data science.

3.2 | Remote Sensing Advancing From Earth Observation to Social Observation

Traditional remote sensing can be understood as physical feature-based remote sensing, which involves the detection, perception, and analysis of the physical properties and geometric features of ground objects. For example, based on different reflection characteristics of electromagnetic waves from different ground objects, various types of ground cover, such as water bodies, soil, and vegetation, and these different types of ground cover can be identified. Physical feature-based remote sensing primarily focuses on earth observation. However, with advancements in remote sensing technology, its scope has expanded to include both earth and human observation. This course covers both physical feature-based remote sensing and social sensing technologies for studying human society.

The initial remote sensing is earth observation. This course introduces four aspects of knowledge regarding traditional remote sensing, including data acquisition platforms, sensors, characteristics and processing of data, and remote sensing in precision agriculture. It covers satellite orbits and multi-level remote sensing data acquisition platforms encompassing space, air, and ground, as well as various remote sensors. Through the analysis of rich remote sensing imagery, students gain an understanding of the capabilities of remote sensing technologies. Moreover, the course introduces key characteristics of remote sensing data, including objectivity, wide coverage, timeliness, and comprehensiveness.

In the practical sessions of the earth observation, students engage in an exercise on remote sensing data acquisition and preprocessing. This includes understanding the characteristics of Gaofen 1 (GF-1) remote sensing data, learning how to query and acquire GF-1 data, and using ENVI remote sensing software for practical operations. GF-1 is a high-resolution multispectral satellite launched in 2013. The practice is designed to help students grasp geometric and radiometric processing methods for GF-1 remote sensing data and gain proficiency in the complete process and operational implementation of extracting vegetation information in Wuhan using GF-1 remote sensing data.

Advanced remote sensing is social observation.

The global database of events, language, and tone (GDELT) is a news media database that monitors news in print, broadcast, and online media worldwide in real time. This data can be used to analyze and extract key information, such as individuals, locations, organizations, and event types. This course uses GDELT data as an example to introduce social sensing technologies and methods. By utilizing GDELT data, international relation networks can be constructed and analyzed, including their small-world properties, scale-free characteristics, and temporal variations (Li et al., 2020).

In the practical sessions of social observation, students engage in an exercise that includes the acquisition and preprocessing of GDELT data, construction and analysis of country interaction networks, and evaluation of temporal changes in these networks. Moreover, the programming language used for this practice is Python.

3.3 | Ubiquitous Sensor Networks and Real-Time Sensing and Processing

The sensor network is an important technical approach for acquiring data in real-time. This course covers real-time sensing and processing technologies based on sensor networks, including sensor network data acquisition technology, preprocessing technology, information service technology, and typical application cases.

In the practical session, three exercises are conducted, including radar sensor data acquisition and preprocessing, temperature and humidity sensor data acquisition and import, and Hikvision camera video stream data acquisition and visualization.

3.4 | UAV-Based 3D Modelling to Construct a Digital Twin World

The coursework on UAV-based 3D modelling comprises two main parts, including unmanned remote sensing data acquisition and preprocessing and 3D scanning data acquisition and preprocessing.

Unmanned remote sensing data acquisition and preprocessing include UAV remote sensing systems, mission planning and design, and data processing. These contents are also included in a practical session.

3D scanning data acquisition and preprocessing involve the principles of 3D point cloud scanning, data acquisition, and processing.

3.5 | Educational and Experimental Platform Building Based on OGE

OGE, developed by Wuhan University, is an open-earth engine designed to meet the growing demand for

spatiotemporal information services in the digital earth realm. This project aims to meet the needs of digital earth spatiotemporal information services, research global spatiotemporal observation data organization and management methods, construct high-performance geographic analysis and AI analysis models, propose global spatiotemporal data knowledge service technologies, develop a data-ready, analysis-ready, decision-ready and open-sharing earth engine system, and build a spatiotemporal information infrastructure, serving the construction of digital twin ties, and further providing a controllable spatiotemporal information base for natural resource management, urban governance, public services, smart transportation, disaster relief, and ecological construction.

This course leverages the abundant data and computing resources of OGE to develop a teaching and experimental platform. The platform enhances students' digital thinking and skills through real data as learning materials, real models based on algorithms, real processing with computing power, and real scenarios for solving problems. This OGE-based educational and experimental platform is under development and will be made available to the public as soon as possible.

4 Conclusions

Data acquisition and preprocessing is one of the 17 digital intelligence education courses at Wuhan University. This study explored the reform and practice of digital intelligence teaching in this course and summarized its 5 key educational approaches. The first approach is understanding data origins and source issues. The second approach is expanding from earth observation to social observation. The third approach is utilizing ubiquitous sensor networks for real-time sensing and processing. The fourth approach is employing UAV-based 3D modelling to construct a digital twin world. The fifth approach is developing an educational and experimental platform based on OGE.

Data science and AI serve as important theoretical and technological foundations for data acquisition and preprocessing. As these fields develop, more advanced and effective methods of data

acquisition and preprocessing will continue to emerge. The reforms and practices of the data acquisition and preprocessing course offer valuable insights for other digital intelligence education courses.

Future research should focus on deepening and enriching various data acquisition and preprocessing methods, completing all practical sessions based on the OGE platform, and further enhancing students' digital literacy and digital intelligence skills.

Acknowledgment Thanks to the editor and anonymous reviewers for their constructive comments and suggestions.

Conflict of Interest The authors declare that they have no conflict of interest.

Data Availability Statements The authors confirm that all data generated or analyzed during this study are included in this published article.

References

- Gong, J. Y., Qin, K., Tang, X. H., Yue, P., & Cheng, P. G. (2024). *Fundamentals of geographic information systems*. 3rd ed. Beijing: Science Press. (in Chinese).
- Li, D. R., Guo, W., Chang, X. M., & Li, X. (2020). From earth observation to human observation: Geocomputation for social science. *Journal of Geographical Sciences*, 30(2), 233–250.
- Lin, Z. Y. (2022). *Data acquisition and preprocessing*. Beijing: Posts & Telecom Press. (in Chinese).
- Meng, X. L. (2024). *The standards and sharing guidelines for the digital intelligence education practice and innovation platform at Wuhan University*. Wuhan: Wuhan University Press. (in Chinese).
- Qin, K., Lu, B. B., Chen, J. P., Li, X., Li, Y. B., & Xu, Y. Q. (2023). *Spatial data analysis*. 3rd ed. Wuhan: Wuhan University Press. (in Chinese).
- Sheng, Y. H., Zhang, K., & Yang, L. (2018). *Spatial data acquisition and management*. Beijing: Science Press. (in Chinese).
- Wu, D. (2024). *The guidelines for building digital intelligence education support system at Wuhan University*. Wuhan: Wuhan University Press. (in Chinese).
- Zhang, P. W. (2024). *The white paper on digital education of Wuhan University*. Wuhan: Wuhan University Press. (in Chinese).