# Research on the Path of Data Literacy Cultivation for College Students

Hao Liang[1], Wang Beibei and Shang Jun
*Shandong University of Science and Technology*
*Qingdao, China*

**Abstract.** The theoretical and practical achievements in domestic studies on data literacy are summarized by bibliometric analysis and network investigation in aspects of the connotation of data literacy, requirements of data literacy, assessment of data literacy and education of data literacy with suggestions proposed for working out the policies of data management and data literacy education, implementing the embedded data literacy education, designing the data literacy teaching contents and constructing the data culture.

## 1. Introduction

The number of various data in the public domain has increased exponentially under the background of big data, and the data has begun to have a profound impact on people's behavior and decision-making. As a standard of basic consciousness and basic ability to measure social and personal data processing, data literacy has increasingly become a key issue of the whole society, in which college students come into contact with and use data more and more frequently. It has become the main body of the provision and social data in the future, and its level of data literacy also determines the development and progress of the whole society. How to improve the data literacy of college students has become an important problem that colleges and universities are facing.

## 2. The significance of improving data literacy of College Students

Scholars at home and abroad do not have a unified definition of data literacy. Foreign scholars generally believe that data itself should be interpreted in a general way, for example, data literacy refers to the ability to master different measurement methods, process different types of data and draw correct conclusions in scientific research [1]. Calzada and other scholars believe that data literacy refers to the ability to acquire, understand, critically evaluate and manage data, and in this process, they should have the awareness and ability to abide by the moral norms [2]; Koltay believes that data literacy not only refers to the ability to obtain data and make critical evaluation and use, but also includes the ability to discuss the relationship between data lit-

---

[1]Corresponding author: Shang Jun, Shandong University of Science and Technology, Qingdao, China, Email: mandysj@163.com.

eracy and other literacy [3]. Different from foreign scholars, domestic scholars tend to think and express data literacy from the perspective of ability. For example, Huang Ruhua believes that data literacy is an extension and expansion of information literacy, mainly including data awareness, data ability and data ethics [4]; Meng Xiangbao believes that data literacy refers to the ability to use data resources to find, analyze and solve problems in the era of big data [5].

Combined with previous research results, from the perspective of talent cultivation in higher education, we can express data literacy as, which refers to people's ability to effectively discover, evaluate and use information and data. As China's higher education under the guidance of the fundamental task of moral education, how to cultivate and realize the all-round development of students is an issue of The Times. As a new era, data literacy is in urgent need of students' ability, which should be attached importance to and planned by colleges and universities.

At the same time, students should also be trained to have certain professional ethics of big data, and not to do things that violate professional ethics. On this premise, students should constantly improve their big data information processing skills [6].

## 3. The current situation of College Student's data literacy

### 3.1. The level of data awareness is low

Data consciousness refers to the process in which human brain makes judgment on data information and connects with the actual situation by analyzing the deep connotation of data information. Some college students think one-sided that data is a simple number and do not establish a preliminary data awareness[7, 8].There are also some students who can realize that data is more information-based than digital, but they can not fully analyze the essential difference between the two.Some students think that data is just interrelated digital information, and there is no essential difference between them.

### 3.2. Data processing capacity is relatively weak

Data processing mainly includes data collection, data arrangement, data expression and data analysis. From the overall analysis, today's students lack of theoretical cognition, of course, this is also a learning process, which can be improved through learning. However, through practical analysis, we found that the neglect of data collection ability makes students lack the ability to organize and select data, and can not effectively summarize the collected data. Furthermore, it seriously affects the ability of other parts of data processing, making the weak data processing ability become a common phenomenon of college students.

### 3.3. The lack of communication ability

In the time of big data demands the data processing ability of contemporary college students. First of all, college students should have the ability of data analysis and prediction. Secondly, college students should have the ability to integrate information. After analyzing and summarizing, they can share their analysis results with others to

ensure the accuracy of data and information. Finally, we should have a certain ability of official document writing, which can transform the data information into text, and accurately reflect the data relationship and information logic. But the common problem of college students is that they lack the ability to transform information into words, and can't communicate data effectively.

## 4. The influencing factors of college students' data literacy

### 4.1. Information dimension factors

In the absence of current education, college students obviously lack sufficient understanding of the types, USES and main data platforms of data, and they cannot carry out relevant data activities in a targeted way when data demands are generated. What has attracted special attention is that data ethics and data security have not been paid enough attention to, and academic misconduct in all kinds of public opinion events has aroused wide attention and intense discussion in all sectors of society, which should be paid more attention to by colleges and universities.

### 4.2. Information technology factors

With the improvement of information retrieval methods and the development of various search engines, the acquisition of data information is no longer as difficult as before, Students can obtain the data they need in a variety of ways. On the contrary, the data analysis software and platform generally have a certain threshold, in front of the skilled use level usually need to be constantly learning and practicing, Pyhton and R language, for example, to master such a new data processing programming language learning cycle, generally in two to three months above, therefore in the process of the improvement of data quality, the need of the two kinds of ability are more energy and resources.

### 4.3. Information environment factors

Data infrastructure is the premise of developing and receiving various data literacy education. Only by gaining access to various hardware and software, data literacy can be improved. At present, due to the protection and reward of intellectual property rights and labor achievements, most legitimate software and platforms are charged. These fees are generally relatively high for students. If schools fail to provide students with appropriate access to use, students will lose interest in using or even choose pirated software.

## 5. The path of cultivation of college students' data

### 5.1. Strengthen the safety awareness in order to enhance data intelligent

Protection ability in data quality education should strengthen the cultivation of student's data security first, using the data of beginning and let students form the habit

of scientific, accurate, standard, because if you don't have enough safety protection consciousness and measures, light lead to loss of data, or lead to property are violated, can be installed by local backup, the cloud remote backup, security scanning software means to ensure that data is not because the physical or network virus attacks and damage or loss, gradually develop students with data analysis, management and decision Thinking habit of big data to improve students' data intelligent protection ability.

## 5.2. Set up data analysis course to improve students' data processing ability

Colleges and universities should strengthen the training of teachers, offer courses such as data analysis and data mining to improve students' data literacy, and provide a variety of teaching projects, including thematic lectures, flipped classes, MOOC, etc. Many world-renowned institutions of higher learning pay attention to the development of data literacy education, such as Purdue University Cornell University, University of Minnesota and University of Oregon jointly carry out the data information literacy project (DIL), analyze the requirements of graduate data literacy education in different disciplines, and take effective measures to meet these requirements; Peking University, Zhejiang University and Fudan University all carry out data literacy teaching courses online and offline, covering data retrieval, data statistics, data software operation and data analysis, etc., in a flexible and free form. The development of cloud computing and big data technology has changed the speed and environment of data transmission. Therefore, the course system of data analysis should be optimized to meet the diverse data processing needs of students.

## 5.3. Create an ecological environment conducive to improving students&apos; data literacy

As the base of cultivating high-level talents, colleges and universities should strengthen the construction of digital teaching environment, build a big data resource platform, provide new data products and data analysis technology, and create a good objective environment for the improvement of students&apos; data literacy. At the same time, colleges and universities should strengthen the investment of funds, such as purchasing copyrights such as genuine software and database, and provide students with free use rights, so as to avoid the lack of data literacy caused by objective factors. In this process, teachers and students can also set up a data literacy communication team, create a sharing and collaborative communication atmosphere, take the relevant projects of data analysis as the guide, cultivate students&apos; data analysis and processing ability, decision-making ability and innovation ability in the face of practical problems, and jointly promote the improvement of data literacy of college students.

## 5.4. Make data literacy assessment for students according to the characteristics of their majors

Different majors have different emphasis on the data literacy requirements of college students. In addition to their abilities of data acquisition, analysis, selection and utilization in the general sense, students majoring in data mining and analysis, data marketing and other majors will have higher requirements on data literacy in the future. According to the professional characteristics, the data literacy evaluation standards of

college students are formulated and evaluated. First, the current situation of students' data literacy can be understood. Second, it can make students change their data concept, form scientific data thinking, and improve the understanding of data ability through evaluation. Thirdly, it can evaluate the effect of data literacy training in a certain stage in a timely and effective manner, and then adjust the training objectives and strategies in the next stage.

## 6. Conclusion

We need to learn from the successful experience of foreign countries, actively explore the content and model of data literacy education, continuously optimize the curriculum system, actively carry out data literacy education, which needs coordination and mutual assistance, unified planning, long-term cooperation, so as to achieve the long-term and stable development of data literacy education. Data literacy education is not a one-way indoctrination process, all kinds of opportunities can be created to achieve their own data literacy driven by scientific research, learning and life needs.

## Reference

[1] Love N.Taking data to new depths.Journal of Staff Developmengt, 2004(4): 183-212.

[2] Calzada PJ,Marzal MA.Incorporating data literacy into information literacy programs:Core competencies and contents.Libri, 2013, 63(2): 123-134.

[3] Kotay T. Big data big literacies. Informatio, Communication&Society, 2013, (15): 663-665.

[4] Huang Ruhua, Li Baiyang. Data Literacy Education: Expansion of Information Literacy education in the era of Big Data . Library Information Knowledge, 2016, (1): 21-29.

[5] Meng Xiangbao,Chang E,Ye Lan. Research on data literacy: origin, status quo and prospect. Journal of Chinese Library Science, 2016, 42(2): 19-126.

[6] Bai Hong Ping. Research on data ability cultivation in the era of big data. Jiangsu Science and technology information, 2017, (29): 66-67.

[7] Ma Xiaohui. Research on data literacy education of college students under the background of BIG DATA. Software Engineering, 2019, (02): 48-51.

[8] Xu Hui. Data Literacy Education for college students in the era of Big Data. Student Forum, 2017, (02): 126-127.