Git-Based Distributed Collaborative Learning: Theories, Tools and Features

Zhihong Yang[1,2(✉)]

[1] The College of Literature and Journalism, Sichuan University, Chengdu, Sichuan, China
yangjh@xbmu.edu.cn
[2] School of Journalism and Communication, Northwest Minzu University, Lanzhou, Gansu, China

Abstract. In recent years, Git-based distributed collaborative learning platforms have been increasingly used in classroom teaching, learning management and knowledge sharing in the field of education. To understand the phenomenon of Git-based distributed collaborative learning, this paper first sorts out the theoretical basis of distributed collaboration, collaborative learning and script theory, and summarizes the functions and features of distributed collaborative learning tools such as Github, Gitbook, Bookdown, and Gitbook Education. The research results show that Git-based distributed collaborative learning has positive application value in the field of education and can promote the development of online teaching at the three levels of teachers, learners and society.

Keywords: Distributed collaboration · Collaborative learning · Git · Github · Github Education

1 Introduction

With the popularization of mobile Internet devices and the wide application of cloud services, the permeability of the Internet in the world has gradually increased. According to the World Bank website data [1], the global Internet penetration rate has increased from 28.9% in 2010 to 56.7% in 2020. Information technology has profoundly changed every field and the education industry is no exception. In particular, the global spread of the COVID-19 epidemic has made school education, especially university education, more dependent on online teaching.

Internet-based teaching activities, such as Massive Open Online Courses (MOOCs), have many advantages [2], such as the ability to provide rich courses to learners on a global scale, the flexibility of learners' learning, the stronger sense of acquisition of learners and so on [3]. On the other hand, the disadvantages of online teaching have also begun to emerge. Arkorful & Abaidoo summarized many studies on the disadvantages of online teaching [4], for example, in online teaching, there is a lack of interaction between learners, teachers, and teachers and managers, and learners are required to have strong learning motivation and certain self-management abilities [5]. In terms of communication skills, due to the lack of face-to-face opportunities, learners are not sufficiently trained to

A. Volodin and I. Roumbal (Eds.): ICECIDT 2022, ASSEHR 677, pp. 124–133, 2023.
https://doi.org/10.2991/978-2-494069-02-2_15

impart what they have learned to others. What's more, many online teaching platforms require additional training, which increases the cost of time, money and management.

In view of the above shortcomings of online teaching, the industry is exploring and optimizing in different fields. In 2005, Linus Torvalds developed the epoch-making distributed version management tool Git to coordinate the development of the Linux kernel. After more than 10 years of development, Git's functions and ecosystem have grown by leaps and bounds, and it has become a leader in the field [6]. The purpose of this paper is to sort out and analyze the distributed collaborative learning platform based on Git technology from three aspects: theories, tools and features, and to clarify the theoretical basis, available tools and characteristics of Git-based distributed collaborative learning.

2 Theoretical Basis of Distributed Collaborative Learning

2.1 Distributed Collaboration

Distributed collaboration was primarily used in computer-related industries at first, referring to the fact that in the network era with enhanced mobility, people no longer rely on fixed office locations, and can collaborate remotely to achieve a certain goal at different times and locations [7]. Distributed collaboration evolved from distributed network. The so-called distributed network is a concept corresponding to a centralized network, which refers to a network connected by multiple nodes (which can be distributed in different geographical locations). Any node in such a network has at least two lines connected to it (see "Fig. 1"). Compared with the centralized network, the network security of this structure is higher, and the failure of a node will not cause the entire network to be paralyzed or blocked. The Internet is a typical distributed network structure.

It is on the infrastructure that supports distributed collaboration, such as the Internet and cloud documents, that distributed collaboration has been realized and has produced large-scale and successful cases such as GNU, Linux, Wikipedia, etc. in practice [8]. With the continuous improvement of cloud technology, distributed collaboration is developing rapidly in various industries with an irreversible trend, especially in the

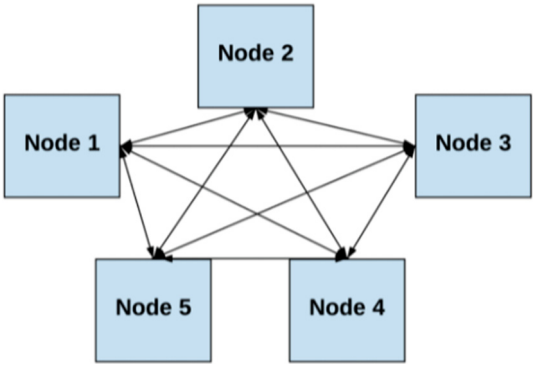

Fig. 1. A model of distributed network.

IT industry and education industry. Distributed collaboration is a new type of social collaboration that reshapes multiple modes of market behavior, cultural expression, and educational practice [8]. Yochai Benkler believes that distributed collaboration is completely decentralized and non-exclusive, not relying on management orders or market signals [9], but being based on loosely connected individuals sharing resources and outputs. High-quality distributed collaboration needs to incentivize participants to share useful information and needs to guarantee or advocate less control and more social trust in terms of institutional ethics [9].

Some researches systematically organize various distributed collaboration literature, redefine distributed collaboration as "the pursuit of a shared objective by groups that include nonproximate members, whose participation is facilitated by ICT", and propose the model distinguishes between "social" and "material" factors that contribute to distributed collaboration and differentiates "social" and "material" factors into "local" and "relational" factors. The model also incorporates the benefits of distributed collaboration to organizations [10].

Subsequent research has further revealed the complexity of distributed collaboration, finding that at least necessary safeguards in terms of flexibility, organizational culture, technical support, and management capabilities are required [11].

2.2 Collaborative Learning

Collaborative learning refers to a teaching strategy in which learners are organized in groups or teams to carry out learning activities [12]. In school teaching, there are usually three types of relationships among students: competitive relationship, mutually independent relationship, and collaborative relationship. The results of previous studies have shown that the learning performance of learners using the organizational form of collaborative learning is better than the other two forms [13, 14].

There are at least two theoretical sources of collaborative learning, one is the learning concept of zone of proximal development proposed by the former Soviet Union scholar Vygotsky. Vygotsky believed that there is a zone of proximal development between the learning tasks that the learner can and cannot complete. Through mentoring and communicating and interacting with others, learners can learn and apply knowledge content that is in the zone of proximal development. The academic concept of zone of proximal development emphasizes the importance of interaction in the context of certain types of learning content [15]. Vygotsky then developed the social constructivist learning theory, which focuses on the important influence of social and cultural mechanisms on learning in the process of learning and knowledge construction. He believed that learning is a process of cultural participation, and learners who participate in the practical activities of a learning community will better internalize the knowledge about learning [16].

Another theoretical background of collaborative learning theory is the situational learning theory proposed by Lave, J, University of California, Berkeley. Situational learning theory believes that learning is a process of social negotiation and a social and practical process involving different subjects. In this process, learners with different resources interact with learning situations, other learners, and complete learning through social interaction and collaboration [17].

Roschelle, J. of Stanford University introduced computer software, especially interaction technology, into collaborative learning in a study [18], which is considered to be one of the early pioneering research results in the field of CSCL (Computer-Supposed Collaborative Learning) [19]. On the basis of CSCL theory, many studies are exploring the factors that affect the effect of collaborative learning. Quite a few of these studies have looked in the direction of learning participants and found that collaborative learning has considerable advantages and disadvantages compared to individual learning and doesn't outperform individual learning on all metrics [20, 21].

Among the many studies on the factors that affect online collaborative learning, it is found that the information technology level of teachers and learners is a prerequisite for collaborative learning based on computer information technology [22]. Therefore, how to reduce the learning cost of information technology for teachers and learners and overcome the barriers to use should be the direction of efforts to improve the effect of collaborative learning, and the collaboration software represented by Git is also gradually recognized by the market and users because of its ease of use and powerful functions.

2.3 Script Theory

Collaborative learning is an effective teaching and learning strategy, but research has found that if the collaborative process is completely controlled by the learner, the effect of collaborative learning will not be significantly enhanced [23]. Therefore, providing learners with effective collaborative script has always been a research hotspot in the CSCL field. In 2013, Prof. Frank Fischer and his research team from the Science Learning Center, University of Munich, Germany, proposed the script theory to integrate factors that affect learners' learning outcomes and effective support modes in a CSCL environment [24].

The study found that learners construct knowledge through the interaction between teachers and other learners in the process of collaborative learning. This interactive process not only requires the individual to improve the learning effect with activities higher than his own level, but also needs the interactive support of peers higher than his own level. The "script" view in cognitive science has shifted from an early comprehensive and stable cognitive structure to a multiple and dynamic cognitive structure. The script theory of Fischer is based on the latest view of cognitive schema theory and believes that learners need to make appropriate combinations of script elements according to different situations and personal goals in the process of collaborative learning in order to achieve better learning effects [24].

Fischer defines two major elements in script theory: internal collaborative script and external collaborative script [25]. The so-called internal collaborative script refers to the existing cognitive schema in the learner's brain, which can help the learner to understand the specific meaning of the event and then take action. The internal collaborative script is dynamic and changes as the learner's experience changes. If learners are not familiar enough with the CSCL environment, there will be situations where the internal collaborative script is not detailed enough, resulting in poor learning experience and learning effects. In this case, external collaborative script should play its role. The so-called external collaborative script refers to the supporting materials and system functions provided by teachers or learning platforms, which are usually manifested in the form of artefact,

such as text materials distributed by teachers, function menus of learning platforms, certain teaching strategies, a certain role assigned for learners, etc. External collaborative script is usually set up before learning activities and is less stable and adaptable than internal collaborative script.

On the basis of dynamic memory theory, Fischer also proposed four components of script theory: play, scene, role and script. The play refers to a general understanding of collaborative learning tasks, such as group discussions or virtual collaboration. The scene refers to the knowledge about the situation that makes up the play component. For example, learners submit rebuttal materials and opinions to a certain point of view in the learning platform and the script is composed of multiple scenes. The role refers to a specific role that the learner assumes in a scene, such as the learner playing a debater with a specific position in a dialogue. The script refers to the knowledge required to form a series of activities in a specific scenario [24]. The above four types of components are not independent but have a hierarchical relationship. The uppermost layer is the play, which can be composed of several scenes; the scenes are composed of the activities of multiple roles, and the roles contain several scripts. Whether it is an internal collaborative script or an external collaborative script, it can be divided into the above four components [25].

The value of script theory is not only that it can better explain how CSCL practice stimulates learners' cognitive development, but also that it can help researchers and developers to better understand the multiple supporting roles of CSCL environment in learners' cognitive development.

3 Git-Based Distributed Collaborative Learning Tools

Although many learning management systems (LMS) have emerged around the collaborative learning education in education world and IT world, with the development of Git and Github website, distributed collaborative learning tools rooted in Git technology have gradually emerged, such as the Wiki, Issue, Pages and other services provided by the Github website, online publishing services such as Gitbook and Bookdown for the writing needs of teachers, as well as Github Education and Github Classroom which are constantly improved and added with functions by Github.

Github is a distributed version management and source code management website created using Git open source technology, with 85 million users, being the largest and most active source code management website in the world [26]. The free Wiki, Issue and Pages services provided by the Github website allow teachers to choose private-owned or public teaching plans and teaching content on the Internet. Through Issue, learners can discuss a topic with group members, teachers, and even wider users. Different from other learning management systems, when using Github for distributed collaborative learning, the collaborative scope of learning activities is not only the class members of a certain semester or stage designated by the teacher, but also all the members who have participated in previous teaching activities, or even members of the entire Internet. This system mechanism of Github makes it possible to improve the teaching effect of distributed collaborative learning. Learners can learn and collaborate more efficiently by reading past discussions. In fact, there are already many programming-related courses in colleges and universities relying on the Github website. Some studies believe that for

college teachers, Github can not only be used as a tool for organizing teaching, but also as a high-quality research data source [27].

Based on Github's large and active ecosystem, similar online publishing services such as Gitbook and Bookdown have also appeared for teachers' teaching materials writing needs. Gitbook is based on the Git distributed version management tool, supports team collaboration, and can accomplish knowledge sharing, document writing, online publishing and other purposes. In addition, Gitbook itself has open source projects hosted on Github. Since Gitbook is a commercial application, the Bookdown free open source project is derived from the Gitbook open source project and other scientific research-related open source projects (such as R package). Bookdown's functions are closer to teachers, researchers, and developers and are especially popular with practitioners in statistics and programming-related fields. Gitbook and Bookdown can not only publish content online, but also generate e-books in common formats, reducing the time cost and collaboration cost of content and format conversion for participants in teaching activities.

Github Education is an education support system launched by the Github website in 2018. There are Github Classroom, educational training, developer toolkits, open source courses and other resources that teachers and students can use for free. Based on the core functions of Github (source code management and distributed version management), Github Classroom has also made innovations different from other traditional course management platforms at the level of external collaborative script, especially strengthening the functions of automatic arrangement, collection and evaluation of students' homework, so as to enable teachers and teaching assistants to provide timely and accurate feedback on the teaching of courses of all sizes.

4 Features of Git-Based Distributed Collaborative Learning

Git technology itself emerged to solve the needs of large-scale distributed collaboration, so all distributed collaborative learning platforms and tools based on Git tools will inherit its capabilities and characteristics. Platforms such as Github, Gitbook, Bookdown, Github Education and other platforms developed around Git have attracted many institutions and users of different identities to establish a very prosperous ecosystem due to some features that other LMSs do not have. These features are summarized below:

4.1 Improving Efficiency and Promoting Career Development

From the perspective of script theory, platforms and tools such as Github, Gitbook, Bookdown, and Github Education provide more efficient external scripts suitable for large-scale collaborative learning and improve the efficiency of teachers in course management, teaching guidance, and resource updating. More importantly, Git-based distributed collaborative platforms can promote the professional development of teachers. The utility of a tool depends on the size of the population using the tool. As the user group of the Git-based distributed collaborative platforms is very large, teachers can communicate with their peers in different regions and countries on the platforms, helping teachers to build their professional reputation in the fields of peers and academic publishing.

4.2 Wider Openness Can Promote the Internalized Motivation of Learners

The learning management platforms of traditional educational institutions are usually not open to the external public and the scope of learners' collaboration can only be limited to the school or class. The Git-based distributed collaborative platforms provide multi-level and multi-dimensional collaboration methods. Their open, shared and social attributes allow learners to learn from others to solve problems on a larger scale, thereby promoting learners to develop and improve their collaboration and self-management abilities. There are a large number of real projects running on Github, Gitbook, and Bookdown platforms. It is easy for learners to find other like-minded users and they can join the development and maintenance of these real projects according to their own interests and abilities [28]. From the perspective of motivation theory [29], since enterprises and institutions recognize the contribution information of learners in Github [30], using a platform similar to Github is more conducive to the establishment and maintenance of long-term internal motivation for learners.

4.3 The Development of Open Education is Conducive to Educational Equity

Git-based distributed collaborative platforms adopt an open source strategy, which is in line with the concept of "openness and sharing" advocated by the open education movement. Judging from the trend of global education development, in the context of the rapid development of big data, cloud computing, and artificial intelligence, educational institutions, especially higher education institutions, not only pay attention to the opening of high-quality teaching resources, but also pay attention to using various tools to innovate courses and teaching models. More and more courses and researches choose Git-based distributed collaboration platforms for publishing. Compared with mature courses solely built by university teachers, most learners believe that the knowledge in the Github platform is updated in a timely manner and of higher quality [31]. Another study found that compared with 2010, the proportion of users from the third world increased significantly [32], which indicates that the high-quality educational resources published in the platform have been spilled and diffused, promoting the development of open education.

5 Conclusion

Taking the phenomenon of distributed collaborative learning based on Git as the research object, through the review of previous research and the analysis of tools such as Github, Gitbook, Bookdown and Github Education, this research believes that, in terms of theory, script theory can better explain and guide CSCL practice. In terms of tools, this research introduces the features of Github, Gitbook, Bookdown and Github Education in distributed collaborative learning. It also concludes that Git-based distributed collaborative learning has the advantages of improving efficiency and promoting professional development at the teacher level; at the learner level, it promotes the internalized motivation of learners through wider openness; at the social level, it is beneficial for the development of open education. In conclusion, Git-based distributed collaborative platforms have positive application value in course teaching, promoting personal development of

learners, and promoting open education. Educational institutions, especially colleges and universities, should draw on the strengths of existing Git-based distributed collaborative learning platforms when designing internal learning management systems.

The disadvantage of this research is that the data are mainly derived from existing research data and public data, there is no in-depth first-hand research on teachers and learners, and the validity of the conclusion of this research is affected to a certain extent.

This research is only the beginning of understanding the phenomenon of Git-based distributed collaborative learning, and in future research, it is recommended to focus on the effects of learning motivation stimulation and recommendation mechanism in Git-based distributed collaborative learning.

Authors' Contributions. This paper is independently completed by Zhihong Yang.

References

1. The World Bank Group. 2021. *Individuals using the Internet (% of population) | Data*. https://data.worldbank.org/indicator/IT.NET.USER.ZS.
2. North, S.M., R. Richardson and M.M. North. 2014. To Adapt MOOCs, or Not? That Is No Longer the Question. *Universal Journal of Educational Research* 2 (1): 69–72. https://doi.org/10.13189/ujer.2014.020108.
3. Akkoyunlu, B., and M.Y. Soylu. 2006. A study on students' views on blended learning environment. *Turkish Online Journal of Distance Education* 7 (3): 43–56.
4. Arkorful, V., and N. Abaidoo. 2015. The role of e-learning, advantages and disadvantages of its adoption in higher education. *International Journal of Instructional Technology and Distance Learning* 12 (1): 29–42.
5. Zhu, M., C. Bonk, and M. Doo. 2020. Self-directed learning in MOOCs: Exploring the relationships among motivation, self-monitoring, and self-management. *Educational Technology Research and Development* 68 (5): 2073–2093. https://doi.org/10.1007/s11423-020-09747-8.
6. Spinellis, D. 2012. Git. *IEEE Software* 29 (3): 100–101. https://doi.org/10.1109/MS.2012.61.
7. Bellotti, V., and S. Bly. 1996. Walking away from the desktop computer: Distributed collaboration and mobility in a product design team. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work*, 209–218.
8. Bollier, D. 2007. The rise of collective intelligence: Decentralized co-creation of value as a new paradigm of commerce and culture. *A Report of the Sixteenth Annual Aspen Institute Roundtable on Information Technology*. Washington, DC: The Aspen Institute.
9. Benkler, Y. 2008. *The wealth of networks*. Yale University Press.
10. O'Leary, K., R. Gleasure, P. O'Reilly, and J. Feller. 2020. Reviewing the contributing factors and benefits of distributed collaboration. *Communications of the Association for Information Systems* 47: 476–520. https://doi.org/10.17705/1cais.04722.
11. Mitchell, A. 2021. Collaboration technology affordances from virtual collaboration in the time of COVID-19 and post-pandemic strategies. *Information Technology & People*. https://doi.org/10.1108/ITP-01-2021-0003.
12. Nistor, N., L. English, and S. Wheeler. 2003. Towards the Virtual University: International On-line Learning Perspectives. IAP.
13. Kitchen, D., and D. McDougall. 1999. Collaborative learning on the Internet. *Journal of Educational Technology Systems* 27 (3): 245–258. https://doi.org/10.2190/5H41-K8VU-NRFJ-PDYK.
14. Ragoonaden, K., and P. Bordeleau. 2000. Collaborative Learning via the Internet. *Journal of Educational Technology & Society* 3 (3): 361–372.

15. Vygotsky, L.S., and M. Cole. 1978. *Mind in society: Development of higher psychological processes*. Harvard University Press.
16. Barnes, D., and F. Todd. 1995. *Communication and learning revisited: Making meaning through talk*. Taylor & Francis.
17. Lave, J., and E. Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
18. Roschelle, J. 1992. Learning by Collaborating: Convergent Conceptual Change. *Journal of the Learning Sciences* 2 (3): 235–276. https://doi.org/10.1207/s15327809jls0203_1.
19. Ma, Jing. 2013. Evolution, research fronts and focus of collaborative learning outside China: Based on mapping knowledge domain. *Open Education Research* 19 (06): 95–101. https://doi.org/10.13966/j.cnki.kfjyyj.2013.06.007.
20. Andrews, J.J., and D.N. Rapp. 2015. Benefits, costs, and challenges of collaboration for learning and memory. *Translational Issues in Psychological Science* 1 (2): 182–191. https://doi.org/10.1037/tps0000025.
21. Menekse, M., and M.T.H. Chi. 2019. The role of collaborative interactions versus individual construction on students' learning of engineering concepts. *European Journal of Engineering Education* 44 (5): 702–725. https://doi.org/10.1080/03043797.2018.1538324.
22. Shonfeld, M. 2021. Factors affecting student-teacher satisfaction with a multi-college online collaborative course. *Australasian Journal of Educational Technology*, 193–205. https://doi.org/10.14742/ajet.6073.
23. Millwood, S. 2005. The Quality of Students's Use of Evidence in Written Scientific Explanations. *Cognition & Instruction* 23 (1): 23–55. https://doi.org/10.2307/3233896.
24. Fischer, F., I. Kollar, K. Stegmann, and C. Wecker. 2013. Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist* 48 (1): 56–66. https://doi.org/10.1080/00461520.2012.748005.
25. Ouyang, Jiayu, and Qiong Wang. 2021. Shaping Learners' Interaction in CSCL Settings: Script Theory and a Review of Its Research. *Modern Distance Education Research* 33 (06): 64–72. https://doi.org/10.3969/j.issn.1009-5195.2021.06.008.
26. Github. 2021. *Build software better, together*. GitHub. https://github.com
27. AlMarzouq, M., A. AlZaidan, and J. AlDallal. 2020. Mining GitHub for research and education: Challenges and opportunities. *International Journal of Web Information Systems*.https://doi.org/10.1108/IJWIS-03-2020-0016.
28. Nielek, R., O. Jarczyk, K. Pawlak, L. Bukowski, R. Bartusiak, and A. Wierzbicki. 2016. Choose a job you love: Predicting choices of GitHub developers. In *2016 IEEE/WIC /ACM International Conference on Web Intelligence (WI)*, 200–207.https://doi.org/10.1109/WI.2016.0037.
29. Gagné, M., and E.L. Deci. 2005. Self-determination theory and work motivation. *Journal of Organizational Behavior* 26 (4): 331–362. https://doi.org/10.1002/job.322.
30. Bhasin, T., A. Murray, and M.-A. Storey. 2021. Student experiences with GitHub and stack overflow: An exploratory study. In *2021 IEEE/ACM 13th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, 81–90. https://doi.org/10.1109/CHASE52884.2021.00017.
31. Lu, Y., X. Mao, T. Wang, G. Yin, Z. Li, and W. Wang. 2019. Studying in the 'Bazaar': An Exploratory Study of Crowdsourced Learning in GitHub. *IEEE Access* 7: 58930–58944. https://doi.org/10.1109/ACCESS.2019.2915247.
32. Wachs, J., M. Nitecki, W. Schueller, and A. Polleres. 2022. The Geography of Open Source Software: Evidence from GitHub. *Technological Forecasting and Social Change* 176, 121478. https://doi.org/10.1016/j.techfore.2022.121478.