Research Proposal: GitHub for Education

Alexey Zagalsky
University of Victoria
Victoria, BC, Canada
alexeyza@uvic.ca

Yiyun Zhao
University of Victoria
Victoria, BC, Canada
evianz@uvic.ca

Weiliang Wang
University of Victoria
Victoria, BC, Canada
weiliang@uvic.ca

INTRODUCTION

GitHub is a popular web-based hosting service for software development that utilizes the Git revision control system. It has shown to be useful in several areas that require collaboration, especially in software development and technical fields (e.g. technical writing). We propose the use of GitHub for Education, and we're not the only ones to do so.

In 2011, Greg Wilson suggested[1] that GitHub could be used for learning materials -

> *Would it be possible to create a "GitHub for education"? Right now, I think the answer is "no", because today's learning content formats make* ***merging hard****. Whatever a "GitHub for education" would look like, it would not be yet another repository of open learning materials. There are lots of those already, but almost all their content is write-once-and-upload, ... rather than* ***sharing course content in a reusable, remixable way****.*

In 2012, he elaborated[2] that

> *"GitHub for Education" isn't necessarily, "Let's put educational materials in GitHub", but rather, "Let's* ***facilitate a culture of spontaneous-but-structured collaboration*** *and improvement".*

In 2014, GitHub initiated a campaign[3] with the purpose of promoting GitHub to higher education, and while it bearly scratches the surface, it is a welcomed first step. GitHub's campaign has given stage to a discussion on the matter, and Jim Baker, a senior seveloper and a lecturer in CS at the University of Colorado, has commented on Facebook (Fig. 1)

> *"We had a great experience using GitHub to support a collaborative workflow for the 70+ students in each of the two semesters of my CS course".*

When asked for more details he replied

Figure 1. GitHub education discussion on facebook.

> *"Pull requests (PR) are the heart of the GitHub workflow, and we took advantage of PRs, including task lists so that students could report on their work in progress and get over initial humps. Any merged PR got extra credit(!). Because the course had been improved in some way - this seemed like an interesting standard for giving out extra cdedit. Consequently we mostly didn't merge PRs for labs, except for bug fixes, but we were always on the lookout for* ***better solutions than ours****. PRs were also merged for extra credit such as* ***corrections of my course notes****. Next fall we expect to have* ***autograding*** *implemented as a form of continuous integration, by running against the PRs through postcommit hooks".*

This research aims to examine whether and how GitHub can be used for educational purposes, while relying on its unique collaboration, social and awareness features. We aim to answer the following research questions:

- Is GitHub used in undergraduate courses to support learning?
- How is GitHub used to support learning?

[1] `http://software-carpentry.org/blog/2011/12/fork-merge-and-share.html`
[2] `http://software-carpentry.org/blog/2012/04/github-for-education.html`
[3] `https://education.github.com/`

CSCW'15, March 14–18, 2015, Vancouver, BC, Canada.

- What benefits and challenges are related to the usage of GitHub for education?

The contribution of this research is fourfold: (1) present up-to-date overview of related work, (2) investigate current use of GitHub *in the wild* (3) present challenges and problems related to using GitHub for education, and (4) provide recommendation and suggested practices on how GitHub could be used to support learning.

RELATED WORK

Before discussing GitHub and how its unique features might support education, we present a brief overview on technology in education, and the strong connection between Computer Supported Cooperative Work (CSCW) and education.

Technology in Education

Technology use in education began with educational films in the 1900s, and with Sidney Pressey's mechanical teaching machines [18] in the early 1920s. In the 1970s, learning material was augmented with graphics and multimedia, and in 1980s and 1990s computer based learning included the use of various technologies such as micro-worlds [16], simulations [22], and hypertext.

The early form of online learning (mid 1990s) was supported by Internet, eMail and Forums, and focused on interaction between the student and the computer. Modern form of online learning focuses on computer-mediated communications and the formation of a communities of practice [12]. It uses various technologies such as cloud computing [24], social media (e.g. Twitter [6, 19]), mobile devices [4], and MOOCs [17].

Technology had a major influence on education, and it will continue to do so (e.g. Khan Academy [25]). However, it is not trivial if and how various technologies should be used to improve learning.

Participants of the "2020 Panel", as described in the book "Technology in Education - Looking Toward 2020" [14], tried to envision the future of education and enrich people's understanding of technology's potential role in education, as well as opportunities, drawbacks and challenges. William Bossert, one of the panelists, proposes that technology should be used to encourage intellectual exchanges among students in the classroom. *Appropriately designed computer programs would make possible the solution of complex problems by a group of students, while helping "level the abilities" within a group so that less able students might take on tasks that would be viewed by others as meaningfully furthering the group activity*. Another panelist, Jim Minstrell, envisions the classroom of the future to be *full with electronic aids for teachers, helping with organizational and management tasks such as: keeping track of appointments and deadlines*. Our study suggest that the use of GitHub unique collaboration, social and awareness features might be the part of that vision.

CSCW and Education

Back in 1984, a workshop [8] organized by Paul Cashman and Irene Grief discussed the way people work and how technology could support it. The term "Computer Supported Cooperative Work" (CSCW) was first coined, while nowadays some researchers may also use the term "Computer Supported Collaborative Work". As Paul Wilson defined in his book [28] in 1991, CSCW [is] a generic term, which combines the understanding of the way people work in groups with the enabling technologies of computer networking, and associated hardware, software, services and techniques. Since then, extensive research and experimentation has been conducted to explore CSCW, for instance the investigation of awareness by Dourish & Bellotti [2], issues in the design and evaluation [7, 13], or evaluation of collaborative applications or systems [9, 20], and more.

CSCW impact extends to education as well [1], by providing mechanisms to support learning. With educational applications and systems that are designed and evaluated, we are able to see how CSCW contributes and supports education in various scenarios, such as lectures, writing assignments or lab work [1]. Moreover, findings of three research programs conducted at the University of Strathclyde show that common guidelines can be developed in the form of procedural stages to facilitate the successful implementation and usage of CSCW technology cross-sectorally in education [11].

More specifically, as opposed to other domains where distance still matters [15], the distance in education is no longer an issue with the right tools and correct environment. Internet offers various possibilities for asynchronous as well as synchronous collaboration and plays a major role in distance education [21]. Remote education systems that are based on CSCW principles provide the learner various learning methods with effective learning experiences and at a low cost [27].

Furthermore, not only students can benefit from the application of CSCW in the classroom, but also teachers could take advantages of such systems, e.g. automating assignments or monitoring students' online behaviour.

GitHub and Education

GitHub offers several unique features that greatly facilitate user's collaboration. For example: Pull Requests are a way to initiate discussion. The discussion may involve code that is visible to everyone and shows what exact changes would be merged if it is accepted, or it may invovle other content beside code (e.g. screenshots) to provide a background for discussion. When a user wishes to contribute to someone else's project, he can either *clone*[4] the project, thus creating a full clone of the project in his local environment while changes commited would directly affect the original project. This is called a Shared Repository Model[5], where Pull Requests help start code review and conversation about proposed changes before they're merged into the master branch. Alternatively, the user can *fork*[6] the entire project, thus creating a parallel project where commited changes would not directly affect the original project. This is called a Fork & Pull Model, where Pull Requests will provide a way to notify the original project maintainers about the changes you'd like them to consider.

[4] `https://help.github.com/articles/duplicating-a-repository`

[5] `http://guides.github.com/overviews/flow/`

[6] `https://help.github.com/articles/fork-a-repo`

Users can not only follow other users or projects of interest, but also broadcast their activities to their followers. Furthermore, users can discover new projects by using the *explore* feature, or share snippets by using the *Gist* feature. GitHub also supports awareness by broadcasting updates to the user's news feed. The combination of these features supports and facilitates *"a culture of spontaneous-but-structured collaboration"*[7].

GitHub is not only a prime example of CSCW, but it's also a good data source to study how people collaborate. Wu *et al.* [29] analyzed rich GitHub timeline data and found that most of the *follow* activities on GitHub are not due to developers collaborations on GitHub but their interactions outside GitHub, such as on Twitter, Hacker News, and StackOverflow. Majumder *et al*. [10] conducted an investigation on how teams are formed in a social network based on millions of software repositories spanning a period of four years and hundreds of thousands of developers from GitHub. Tsay *et al*. [26] performed a quantitative analysis of project success of over 5,000 open source software projects hosted on GitHub. They used two measures, **developer attention** and **work contribution**, for projects success, and found that projects with a high level of developer multitasking tend to receive less developer attention, but greater work contribution.

Up until recently, GitHub has mainly been used for software development, but it can also be extended to other domains that involve collaborative works - Education is one of them [5]. In this different context, both GitHub and teaching techniques might require readjustment while some of GitHub features might be repurposed. For instance, student assignments should be modified to fit GitHub's transperancy and collaboration model.

Several educators have introduced the use of GitHub into their classrooms and had a good experience.

In 2010, Luis Felipe Borjas[8] posted about the use of github organizations[9], a way to simplify management of group-owned repositories, to manage the class projects. He also applied GitHub to exams: instead of waiting until the exam deadline to upload the exams, students could just *push*[10] as many times they wished and the last *push* they made before the deadline was going to be considered as their final response. Furthermore, he strongly suggested teachers to create exams/homeworks that would translate into private *git repos* to which the students could push.

In 2011, David Humphrey stated in his blog[11] that he asked his students to use Git/GitHub and highly recommended instructors to use GitHub in classroom. He claimed that although it was a little painful to learn Git/GitHub at the beginning, the payoff would be huge.

In 2013, Griffin & Seals have used GitHub in the classroom as a version control tool while leveraging the *branch* and *merge* features [5]. When students worked on programming assignments, it was easy to *merge* it back into the original project if their version worked, also pretty handy for them to abandon that branch if it's totally off course, without destroying the original project.

Apart from these testimonies of using *"raw GitHub"* in the classroom, there are also education targeted platforms based on the GitHub model. Coursefork[12], created in March 2013, is described as *"GitHub for course creation"*[13]. It is a platform for open-sourcing and collaborating on educational material, where educators can upload course material, allow others to create copies, modify or share them with others.

These examples show that although this idea has not been promoted widely, it is being practiced *informally* and has much potential to benefit education.

METHODOLOGY

Based on our research questions, we chose to employ Qualitative methodology [3, 23] and methods. In this section we elaborate on the research questions, and study design.

Research Questions

This study strives to answer the following research questions:

Is GitHub used in undergraduate courses to support learning?
This study will examine whether GitHub is being used (formally or informally) to support learning in undergraduate courses.

How is GitHub used to support learning?
In cases where GitHub is used to support learning, we wish to investigate how it is used. Specifically, for what uses is it being used (e.g. file sharing, collaboration, time tracking). Moreover, we wish to investigate what specific features of GitHub are being used to support learning.

What benefits and challenges are related to the usage of GitHub for education?
We aim to investigate what possible benefits GitHub usage might bring to support learning, based on testimonies of the study's participants. Furthermore, we plan to examine what challenges students and lecturers face when trying to use GitHub or what problems prevent them from even trying.

Study Design

We are interested to investigate real practiotionairs (both students and lecturers), while focusing on undergraduate courses. As first phase we plan to conduct interview with two lecturers, Jim Baker and Arie Van Deursen, who had been using GitHub as a tool in their classroom and published[14] their experience. Second phase will be to form a survey questionaire based on the interview analysis from the first phase.

[7] http://software-carpentry.org/blog/2012/04/github-for-education.html
[8] http://lfborjas.com/2010/10/30/git-classroom-exams.html
[9] https://github.com/blog/674-introducing-organizations
[10] https://help.github.com/articles/pushing-to-a-remote
[11] http://vocamus.net/dave/?p=1358
[12] http://coursefork.org/
[13] http://opensource.com/education/13/9/coursefork-education-tool
[14] http://avandeursen.com/2013/12/30/teaching-software-architecture-with-github/

Third phase would be to conduct a survey with the students of these two lecturers, students who experienced the use of GitHub in their classroom. If this would not be possible, for various reasons, we will conduct a focus group based on undergraduate students from University of Victoria.

The study design allows to gain valuable insights into different group interactions, between GitHub and studnets, and between GitHub and lecturers.

EXPECTED RESULTS AND CONTRIBUTION

This research aims to examine whether and how GitHub can be used for educational purposes, while relying on its unique collaboration and social features. The contribution of this research is fourfold: (1) present up-to-date overview of related work, (2) investigate current use of GitHub *in the wild* (3) present challenges and problems related to using GitHub for education, and (4) provide recommendation and suggested practices on how GitHub could be used to support learning. Findings of this research will help examine GitHub as tool to support learning, while focusing on potential advantages and disadvantages. Furthermore, it may provide a new perspective on collaboration in education, may lead to evolvement of GitHub from a software development hub into a collaboration platform,and guide practitioners and researchers in building new services and tools to support learning.

REFERENCES

1. Dewan, P. A survey of applications of cscw including some in educational settings. *Proceedings of ED-MEDIA93* (1993), 147–152.
2. Dourish, P., and Bellotti, V. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, ACM (1992), 107–114.
3. Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering*. Springer, 2008, 285–311.
4. Gikas, J., and Grant, M. M. Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. *The Internet and Higher Education 19* (2013), 18–26.
5. Griffin, T., and Seals, S. Github in the classroom: Not just for group projects. *J. Comput. Sci. Coll. 28*, 4 (Apr. 2013), 74–74.
6. Grosseck, G., and Holotescu, C. Can we use twitter for educational activities. In *4th international scientific conference, eLearning and software for education, Bucharest, Romania* (2008).
7. Grudin, J. Why cscw applications fail: problems in the design and evaluationof organizational interfaces. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, ACM (1988), 85–93.
8. Grudin, J. Cscw: History and focus. *COMPUTER-LOS ALAMITOS- 27* (1994), 19–19.
9. Leland, M. D., Fish, R. S., and Kraut, R. E. Collaborative document production using quilt. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, ACM (1988), 206–215.
10. Majumder, A., Datta, S., and Naidu, K. Capacitated team formation problem on social networks. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, ACM (New York, NY, USA, 2012), 1005–1013.
11. McGregor, S., Thomson, A., and Ion, W. Cscw in industry and education: Transfering knowledge in engineering design. *proceedings of Concurrent Engineering, CE2000, Lyon, France* (2000).
12. McLure Wasko, M., and Faraj, S. it is what one does: Why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems 9*, 2 (2000), 155–173.
13. Neuwirth, C. M., Kaufer, D. S., Chandhok, R., and Morris, J. H. Issues in the design of computer support for co-authoring and commenting. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*, ACM (1990), 183–195.
14. Nickerson, R. S., and Zodhiates, P. P. *Technology in education: Looking toward 2020*. Routledge, 2013.
15. Olson, G. M., and Olson, J. S. Distance matters. *Human-computer interaction 15*, 2 (2000), 139–178.
16. Papert, S. Microworlds: transforming education. In *Artificial intelligence and education*, vol. 1 (1987), 79–94.
17. Pappano, L. The year of the mooc. *The New York Times 2*, 12 (2012), 2012.
18. Pressey, S. L. A machine for automatic teaching of drill material. *School & Society* (1927).
19. Reinhardt, W., Wheeler, S., and Ebner, M. All i need to know about twitter in education i learned in kindergarten. In *Key Competencies in the Knowledge Society*. Springer, 2010, 322–332.
20. Roseman, M., and Greenberg, S. Teamrooms: network places for collaboration. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, ACM (1996), 325–333.
21. Roth, J. Cscw internet tools and environments for distance education. *University of Hagen Department for Computer Science 58084* (1998).
22. Rutten, N., van Joolingen, W. R., and van der Veen, J. T. The learning effects of computer simulations in science education. *Computers & Education 58*, 1 (2012), 136–153.
23. Seaman, C. B. Qualitative methods in empirical studies of software engineering. *Software Engineering, IEEE Transactions on 25*, 4 (1999), 557–572.

24. Sultan, N. Cloud computing for education: A new dawn? *International Journal of Information Management 30*, 2 (2010), 109–116.

25. Thompson, C. How khan academy is changing the rules of education. *Wired Magazine 126* (2011).

26. Tsay, J. T., Dabbish, L., and Herbsleb, J. Social media and success in open source projects. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion*, CSCW '12, ACM (New York, NY, USA, 2012), 223–226.

27. Wang, C., Shao, M., Xia, J., and Chen, H. Research and design of remote education system based on cscw. In *Grid and Cooperative Computing*. Springer, 2004, 1038–1041.

28. Wilson, P. *Computer supported cooperative work:: An introduction*. Springer, 1991.

29. Wu, Y., Kropczynski, J., Shih, P. C., and Carroll, J. M. Exploring the ecosystem of software developers on github and other platforms. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW Companion '14, ACM (New York, NY, USA, 2014), 265–268.