Using GitHub Cloud Service in Training Future IT Professionals: Local Study

Olena G. Glazunova[a], Valentyna I. Korolchuk[b], Tetiana V. Voloshyna[c], Yevhenii M. Starychenko[d] and Oleksandra V. Parkhomenko[e]

National University of Life and Environmental Sciences of Ukraine, 15 Heroiv Oborony Str., Kyiv, 03041, Ukraine

Keywords: GitHub, Future IT Professionals, Collective Development.

Abstract: In today's IT industry, it is important to develop the ability of IT students to collaboratively develop software, professional and personal skills. An effective method for developing such skills in future IT specialists is to organize different types of educational projects related to different programming technologies during the execution of mini projects, group and individual project assignments, term papers, academic training within the academic disciplines. The paper summarizes the results of a pedagogical study involving 29 expert students who study Computer Science and Software Engineering and used cloud service for GitHub collaborative IT development projects. The research findings testify, the most effective characteristics of this service, according to experts, identified the possibility of collaborative development of software (i1), the convenience of bug tracking (i3) and the convenience of the code editor (i7). It offers examples and results of using GitHub cloud service in the process of executing educational projects by future IT specialists.

1 INTRODUCTION

With the development of information technology (IT), the approach to the organization of collaborative development of software products is changing. Hence, it is necessary to take into account the fact that future IT specialists should be able to adapt instantly to new situations, make appropriate decisions and quickly solve their tasks not only personally, but also while working as a team. In order for students of IT profession to continue to hold leading positions in IT industry in their professional activity, to meet the requirements of customers and employers, it is necessary to develop in them the ability to design and manage projects, to work in a team, to develop skills to use cloud services for project management and team development of software products in the process of their academic training at the university.

[a] https://orcid.org/0000-0002-0136-4936
[b] https://orcid.org/0000-0002-3145-8802
[c] https://orcid.org/0000-0001-6020-5233
[d] https://orcid.org/0000-0001-8608-5268
[e] https://orcid.org/0000-0001-8773-0185

2 THEORETICAL BACKGROUND

Cloud software for team development of software products allows users to collaborate on code, manage their versions, and more. Cloud services such as GitHub, Bitbucked, GitLab, Phabricator, Beanstalk, which were researched and described in the paper (Korolchuk, 2019). become part of the cloud-oriented scientific and educational environment of the university if used on a regular basis (Glazunova and Shyshkina, 2018). GitHub is the most popular code management platform for software development, as it enables future IT professionals to manage and collaborate on their software development training projects.

GitHub is an online Git service that hosts Git repositories and provides other features such as issue tracking. GitHub has become the prominent platform for hosting open source projects (Metz, 2015). GitHub has been embraced by the software development community as an important social platform for managing software projects and to support collaborative development (Feliciano et al., 2016). The most important benefit of using GitHub is not to support the short-term priorities of a semester-long course, but, rather, to encourage sustainable and well-documented digital development, both of student projects and the course itself (Beshero-Bondar and Parker, 2017).

Glazunova, O., Korolchuk, V., Voloshyna, T., Starychenko, Y. and Parkhomenko, O.
Using GitHub Cloud Service in Training Future IT Professionals: Local Study.
DOI: 10.5220/0012065200003431
In *Proceedings of the 2nd Myroslav I. Zhaldak Symposium on Advances in Educational Technology (AET 2021)*, pages 448-461
ISBN: 978-989-758-662-0

When using GitHub in education, one has to think about the purpose, what the goal is and then how the features in GitHub can be used to reach this goal. GitHub can be used in many different ways, but it might not be applicable in all types of courses because a certain amount of knowledge about Git is required to be able to use the features (Gunnarsson et al., 2017). Git and GitHub into data science workflows is considered best practice, and provides thoughtful advice on how to conceptualize the GitHub workflow (Fiksel et al., 2019). Other work describes a GitHub Education study that shows that using GitHub in the classroom can lead to a much improved understanding of students' project management (GitHub Education, 2020).

The most important skills employers seek in engineering are creativity, teamwork and critical thinking. Kertész (Kertész, 2015) presents the results of a collaborative learning experiment using GitHub in lab work, where the focus was on students' direct interaction with each other's learning process. Depending on how GitHub is implemented in learning programming, students may rely on GitHub for activities such as, submitting assignments, collaborating on group projects, and receiving feedback (Hsing and Gennarelli, 2019). An immediate advantage is for classes that have group projects. With GitHub Classroom, instructors can easily assign groups of students to teams and give each team their own GitHub repository within a GitHub Classroom. Students can then use Git and GitHub to collaborate on a project, just as they would in an academic or industry research project. Because teachers can see each student's commit history, it is easy to see how each student contributed to the project (Fiksel et al., 2019). A collaborative tutorial assignment on the GitHub platform was embedded in an undergraduate cybersecurity course. Students were asked to create a tutorial that would be combined with their peers' tutorials to create a course eBook. The tutorial topics were required to be in the general domain of network security. With regards to tutorial difficulty, students were told to target an audience that had completed an introductory computer networking course (Marquardson and Schuetzler, 2019).

Instead, this study was aimed at answering the following research questions: a) how to use the cloud service to collectively develop GitHub for programming training projects; b) which function-based assessment indicators affect the effectiveness of the cloud service for collective GitHub development; c) how collective GitHub development influences the formation of professional programming competence of future information technology specialists. The motivation for this study was to demonstrate the effectiveness of using the GitHub cloud service for student programming projects.

The problem of the research stems from the need to find effective information and communication technologies for the organization and implementation of various types of collaborative projects by future IT specialists. Therefore, this study was conducted to determine the characteristics of the GitHub cloud service, which affect the effectiveness of the execution of educational projects on programming in the process of training IT students and in order to form professional programming competence.

3 RESEARCH METHODOLOGY

3.1 General Background

In order to determine how effectively GitHub's cloud service enables students to carry out programmatic learning projects, such as interacting with team members, collectively working on code, sharing ideas, and reviewing each other's work during research, students' thoughts and impressions about using this cloud service were gathered. A descriptive study utilizing survey methodology was used as appropriate to achieve the objectives of the study. This allowed the researchers to gain a more detailed view of the students regarding the use of the GitHub cloud service in the execution of educational programming projects using pedagogical observation and peer review methods. The study was conducted in two stages. At the first stage, the role of experts was performed by students who evaluated the effectiveness of the GitHub cloud service for the implementation of programming training projects. The study of the first stage was conducted among the 3-rd year students of the Faculty of Information Technologies of the National University of Life and Environmental Sciences of Ukraine (NULES of Ukraine) during the second academic term of 2019 (29 students).

Twenty-nine students of IT specialities performed a collective mini-project in the process of studying one informative module in the discipline "Object Oriented programming". The study of this academic discipline was preceded by the study of "Database Organization" discipline; "Development of a program system for working with IT company management computer systems databases" was selected as the project theme. After completion of the technological practical training, an expert evaluation of the GitHub cloud service was conducted by the students in the second phase. To understand the attitude of students to the service for collaborative development, the following indicators of their evaluation from the point of view of functionality were determined: (i1) possibility of collaborative development of software; (i2) ability to manage code versions; (i3) convenience of bug tracking; (i4) ability to organize and plan teamwork; (i5) communication capability; (i6) the ability to support platforms; (i7) the convenience of the code editor; (i8) security and privacy; (i9) availability of wiki pages. In order to evaluate the GitHub collaborative development cloud service by specific indicators, a survey was developed, which consisted of 9 questions, in which the experts evaluated the importance of the indicators by assigning a ranking number.

In 2020, the second phase of the study was conducted, during which the effectiveness of the use of the GitHub cloud service was evaluated in the training of future IT professionals for the formation of professional competence. During experimental research among 3rd year students of the faculty there were two samples of students: control (96 students) and experimental groups (99 students). Student assessment was conducted during the study of the subject ”Object-Oriented Programming” for two semesters and technological practice (internship).

In the experimental study, students performed two mini-projects in academic disciplines and one group project during technological practical training. The first part of the study was to carry out the students' programming project using the resources of the e-learning course (ELC) of the academic discipline combined with the cloud service for the collective development of GitHub.

3.2 Data Analysis

The experts evaluated the significance of the developed indicators by assigning them a ranking number. The highest rated factor was assigned a rank of 1. The level of agreement of experts' opinions was determined by the coefficient of concordance. The concordance coefficient was applied to assess the degree of consistency among experts, which was calculated by the formula:

$$W = \frac{12S}{m^2(n^3-n)},$$

where

$$S = (\sum x_{ij} - \frac{\sum\sum x_{ij}}{n})^2,$$

x_{ij} – evaluations of the ranks of each object of examination, n – number of criteria evaluated, m – number of experts who evaluated the service. To calculate it, the sum of the assigned ranks and deviation squares of the rank sums from the average sum for each indicator were determined. The statistical significance of the coefficient of concordance was checked against the Pearson correlation criterion $\chi^2 = m(n-1)W$. Based on the sums obtained, the sum of the converted ranks was determined and the weight of each indicator was calculated to the formula, where $s_{ij} = x_{max} - x_{ij}$.

Student's t-test and analysis of variance were used to test the effectiveness of using the GitHub cloud service to develop professional competence. One-factor and two-factor analysis of variance with intergroup and intragroup factors – mixed-model analysis of variance (mixed-model ANOVA). Student's T-test allows to check the equality of mean values in two samples and is calculated by the formula:

$$t = \frac{|M_1 - M_2|}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}},$$

where M_1 and M_2 – mean value in control and experimental groups; σ_1 and σ_2 – standard deviation; N_1 and N_2 – sample sizes.

The variance two-factor analysis allowed to estimate the effect of two factors on different samples of objects, and the one-factor one on the influence depending on the evaluation stage.

The following conditions were considered to obtain reliable results of analysis of variance:

1. In the analyzed groups, the values of the dependent variable should be normally distributed. In this case, it is assumed that the value of the dependent variable has a normal distribution within each group, relative to the levels of factors. However, the response values do not have to have this distribution. Another weakening of the normality requirements is the normality of the distribution of model residues.
2. Homogeneity (homoskedasticity) of group variances. That is, the values of the dependent variable in each group must be statistically equal.

4 IMPLEMENTATION

One of the important types of projects in the process of IT specialists training is the projects on the collaborative development of software products, and therefore it is important to prepare students for the implementation of such projects since college times, to develop in them the necessary professional and personal skills, in particular the skills of shared software developments. When choosing cloud services for collaborative development of software products, the following issues should be taken into account: interoperability on the code, bug tracking, discussion of the code with other team members, management of versions of code and integration of additional services, availability of a repository, wiki and code editor, etc. There arises a need to integrate additional services so that the cloud services of software products collaborative development could enable us to manage projects. While implementing software development projects, the students cannot be restricted by cloud services only while organizing the teamwork; the future IT specialists also need the services, which will allow them to work together on the product code they plan to develop.

Given the rate of change in the field of IT, the number of cloud services for teamwork is constantly increasing, but assessing the functionality of such services, they can be subdivided into two categories: cloud services for project management (1) and team development of the software product (2).

Recently, educators also started using GitHub as a teaching tool for programming courses by hosting code samples and managing student tasks, and organizing teamwork (Angulo and Aktunc, 2019). The ability to use version control is a valuable skill for computer science graduates to possess. Git is a well-established, well received source version control system for the software development community and beyond (Bonakdarian, 2017; Kelleher, 2014; Haaranen and Lehtinen, 2015).

To complete the programming projects, experimental group students were asked to combine ETC resources in different academic disciplines with the GitHub service. In to complete the programming projects, students' experimental group were asked to combine ETC resources in different academic disciplines with the GitHub service. Morze and Glazunova (Morze and Glazunova, 2013) proposed the structural features, the ratio of form and content of the smart course elements and its properties: individual learning paths, content personification, the use of training elements with links to public information resources, interactive training elements, multimedia, communication and cooperation elements are substantiated. Teachers of relevant disciplines and technological practical training placed tasks of collective projects in the ELC. The ETC, posted by the teacher, contained the theoretical material (Book, Lesson resources) and course terminology (Glossary resource), lab sessions assignments (Assignment resource), and the exchange of useful resources and files (Database resource). At GitHub, student teams create their own projects, in which they can further collaborate on code writing, use the repository, perform branches, issue releases, and communicate with each other while completing study project tasks. The scheme of combining Moodle resources with the GitHub cloud service is presented in figure 1.

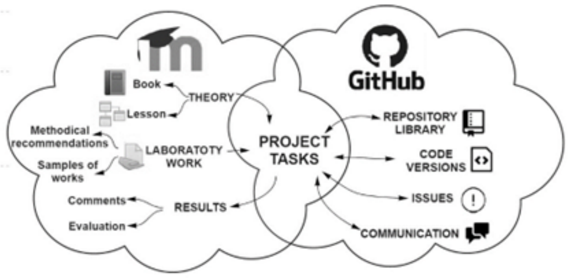


Figure 1: The scheme of combining Moodle resources with the GitHub cloud service.

For project or lab assignments, student groups can work on a public repository sharing code and ideas. It also allows cross-team communication with multiple teams working on a larger project as it happens in industry settings or teams exchanging ideas and reviewing each other's works. GitHub's cloud service provides users with a user-friendly web interface to the repository, user profile tools, change tracking, messaging, comments and online access, allowing the instructor and students to track the contributions of each team member so that students can be held accountable for their work. Thus, we can single out the following features of the GitHub cloud service (figure 2), which are important in the course of the implementation of educational projects on programming:

- programming: code editor; code versions management; bug tracking; platform support; availability of wiki pages;
- collaborative development: joint software development; teamwork planning and organization; establishing communication; security and privacy.

Precisely these features, inherent in GitHub tools, make it possible to apply different types of educational projects related to different programming technologies. This cloud service was offered to students for completing mini projects, for group and individual project assignments, term papers within academic disciplines.

In order to determine the effectiveness of the cloud service for GitHub collaborative development, 29 students were surveyed on the above-mentioned evaluation indicators after the implementation of the programming educational projects. A questionnaire was developed to ask students to assess the importance of each of the indicators:

- i1. Possibility of collaborative development of software;
- i2. Ability to manage code versions;

Figure 2: Classification of GitHub cloud service features.

- i3. Convenience of bug tracking;
- i4. Ability to organize and plan teamwork;
- i5. Communication capability;
- i6. Ability to support platforms;
- i7. Convenience of the code editor;
- i8. Security and privacy;
- i9. Availability of wiki pages.

Table 1 provides an assessment of the results of determining the effectiveness of cloud service for GitHub team development in the course of executing educational projects on programming by future IT specialists. The highest-rated indicator was assigned a rank of 1.

The concordance coefficient $W = 0.85$ indicates a high degree of convergence among experts. Pearson correlation criterion was calculated to assess the significance of the concordance coefficient $\chi^2 = m(n-1)W$. As the calculated one χ^2 (197.2) is higher than the table value (15.5) for the number of degrees of freedom $K = n - 1 = 9 - 1 = 8$ and at a given level of importance $\alpha = 0.05$, we may conclude that the obtained coefficient of concordance of 0.85 is not accidental, and therefore the results obtained are statistically significant.

Based on the obtained rank sum the weights of the indicators considered were calculated. The survey matrix was transformed into a matrix of transformed ranks according to the formula, where $s_{ij} = x_{max} - x_{ij}$, in which $x_{max} = 9$ and the weight of each indicator was calculated.

The analysis of the significance of the factors studied revealed that the following indicators were noted by the students as being the most significant ones: the possibility of collaborative development of software, the convenience of bug tracking, the convenience of the code editor, and the ability to manage code versions when completing educational projects on programming.

In the second stage of the study, students worked on tasks of various projects using the cloud service GitHub. It was proposed to carry out collective mono-projects during the study of professional disciplines or course work within such disciplines, which will allow the formation of future IT professionals professional competencies and soft skills using services for collective IT development for inverted learning (Glazunova et al., 2022).

While studying “Object Oriented Programming” academic disciplines students were asked to complete mini projects using a cloud service to collaboratively develop GitHub software. The purpose of such projects was to develop professional competencies and personal effectiveness skills in future IT specialists Students worked in teams of 4-5 people. In each team a leader was identified, he distributed the tasks among the participants of the collective project. The task of “Database Organization” organization mini projects was to design a relational model of databases for the future automated system (according to the topic chosen by students); constructing class diagrams and developing a system using a class composition.

Table 1: Assessment of the results of determining the effectiveness of cloud service for GitHub team development in the course of executing educational projects on programming by future IT specialists.

Indicator	Rank sum	S	Concordance coefficient	Pearson criterion		Sum of converted ranks	Indicator weight
				predictive	table		
i1	35	12100	0.85	197.2	15.5	226	0.22
i2	121	576				140	0.13
i3	62	6889				199	0.19
i4	147	4				114	0.11
i5	173	784				88	0.08
i6	248	10609				13	0.01
i7	96	2401				165	0.16
i8	227	6724				34	0.03
i9	196	2601				65	0.06
Σ	1305	42688				1044	1

Within the framework of mastering the ”Object Oriented Programming” academic discipline, the students were offered to implement an educational project entitled “Development of a program system for working with IT company management computer systems databases”, the objective of which was to review and analyze modern design technologies; to develop and use software standards for common computer-driven control systems; to develop software structures for the computerized management system and UML diagrams of design entities; to develop a graphical interface for computer control system software; programming and debugging using object-oriented programming techniques; testing and analysis of the performance of the developed computer management software; reporting on the performance of computer-based management systems; development of a set of standard documents to support the developed computer management software.

Within the framework of the project and technological practical training, students performed a collaborative project using the GitHub cloud service aimed at the development of their professional and personal competencies, namely: improvement of practical skills in software development and design using modern approaches and tools for flexible software development, development of teamwork skills in students, which are in demand on the modern IT labor market. The task of the team project was to develop software with web interface and relational database using HTML, CSS, JavaScript, MySQL, PHP technologies. Work on the educational project was carried out in line with the principles of Agile flexible development and Scrum methodology, which provides an incremental and iterative approach and specific roles of the participants in the development of the collaborative project. In the course of collaborative IT development using GitHub, future IT specialists kept the educational project code and necessary documentation in the public domain. In addition, a version control system was used to provide integrity and multiuser access.

5 EXPERIMENTAL RESULTS

To assess the effectiveness of the use of GitHub, two samples of students were selected: a control group (Control) without the use of a joint development service in the educational process and an experimental one using GitHub (Experiment). The comparison will be made in three stages of studying the discipline ”Programming”. Assessments were conducted in the following sequence: exam for the first semester (Examen1), results for the second semester (Examen2) and internship (EducPractic). Accordingly, the study put forward the following null hypotheses, the deviation of which will confirm the effectiveness of the use of the cloud service GitHub for the formation of professional competence: 1) the average score in the control and experimental groups does not differ; 2) the difference in the average score at different stages of evaluation is statistically insignificant; 3) the difference in the average score by groups (samples) at different stages of evaluation is statistically insignificant.

To test the first hypothesis, Student’s t-test or its non-parametric analogue, the Mann-Whitney-Wilcoxon test, will be used. To check others – analysis of variance. In particular, for the second hypothesis, one-factor analysis of variance (if necessary, non-parametric Kruskal-Wallis test) and for the third, two-way analysis of variance with intergroup and intragroup factors – mixed-model analysis of variance (mixed-model ANOVA).

Before starting the statistical verification, the power analysis should be used to determine the level of effect for the given methods that provide sample sizes. Table 2 shows the sample sizes in different sections. A significance level of 0.05 and a power of 80% were also selected for power analysis.

Table 2: The size of student samples at each stage of the study.

Evaluation stage	Group		Sum
	Control	Experimental	
Examen1	96	99	195
Examen2	96	99	195
EducPractic	96	99	195
Sum	288	297	-

To assess the possibility of neglecting the normality requirement, the effect level for Student's t-test was determined by the first hypothesis for samples of experimental and control groups, which are 288 and 297, respectively. The results of the calculations are presented in listing (figure 3).

```
pwr.t2n.test(n1=288, n2=297, sig.level=0.05,
             power=0.80)
t test power calculation
             n1 = 288
             n2 = 297
              d = 0.232
      sig.level = 0.05
          power = 0.8
```

Figure 3: Evaluation of the effect for the Student's t-test.

As figure 3 of the power analysis shows, the effect level is 0.232 (d value), which indicates the possibility of determining small effects (Cohen, 1988), accordingly, the requirement for data distribution according to the normal distribution law can be neglected.

To analyze the capacities for small groups, a calculation was made for pairwise comparisons of the obtained scores in terms of individual types of evaluation (sizes of groups 96 and 99), which is presented in listing (figure 4).

```
pwr.t2n.test(n1=96, n2=99 , sig.level=0.05,
             power=0.80)
t test power calculation
             n1 = 96
             n2 = 99
              d = 0.403
      sig.level = 0.05
          power = 0.8
    alternative = two.sided
```

Figure 4: Evaluation of the effect for pairwise comparisons of scores in terms of individual types of evaluation.

The obtained value of the effect (0.403) corresponds to the average theoretical level of the effect, therefore, when choosing a statistical procedure, it is advisable to consider the law of data distribution in the samples.

For the second hypothesis, where the test method should be a one-way analysis of variance with a group size of 195, the calculated effect is shown in listing (figure 5).

```
pwr.anova.test(k=3, n=195, sig.level=0.05,
               power=0.80)
Balanced one-way analysis of variance power
calculation
              k = 3
              n = 195
              f = 0.129
      sig.level = 0.05
          power = 0.8
```

Figure 5: Effect evaluation for analysis of variance.

The obtained value (0.129) is close to the theoretical value of the small effect (0.1) but slightly exceeds it. Therefore, when establishing the inconsistency of the distribution with the normal law, we additionally use nonparametric methods.

Since there is no statistical procedure for determining the level of effect in the power analysis for two-factor analysis with an unbalanced design (groups of different sizes), this analysis was not performed. The obtained average scores in the sections of the groups and evaluation by the results of the second stage of the experiment are presented in table 3 and in figure 6.

Table 3: The average performance of students in the control and experimental groups in terms of stages of assessment.

Evaluation stage	Group		By types of evaluation
	Control	Experimental	
Examen1	73.1	75.9	74.5
Examen2	76.2	79.9	78
EducPractic	74.3	82.2	78.2
By groups	74.6	79.3	-

According to the summary data, the difference between the overall scores in the groups is 4.7 points. At the same time, if we evaluate the pairwise differences in grades for different types of assessment, the biggest difference was in the success of students in the results of internship – 7.9 points.

Analyzing the data presented in figure 3, we see the difference in the medians, as well as the distribution of scores – the experimental group shows the best results for both general and stages (figure 7).

In this case, if the experimental group is characterized by an increase in scores in stages: Examen1 $\rightarrow$ Examen2 $\rightarrow$ EducPractic; then in the control group after the growth of average scores, the average scores of internship results are lower than for the exam in the

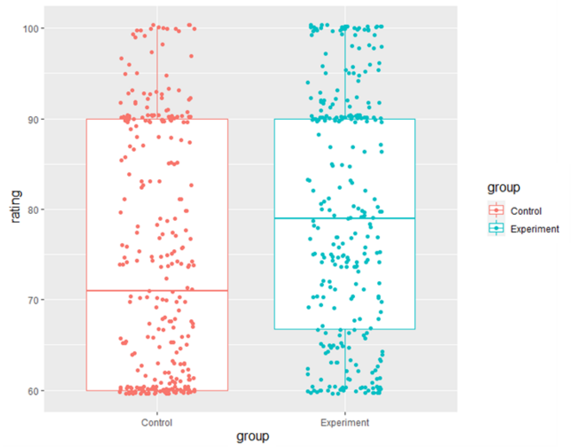


Figure 6: Average student performance based on the results of three stages of assessment.

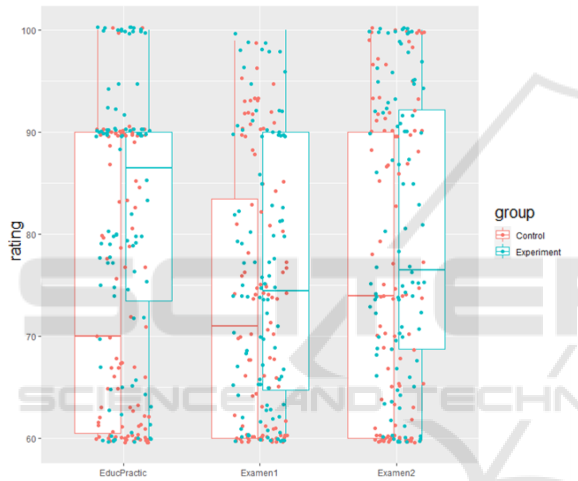


Figure 7: The success of students of control group and experimental group in terms of assessment stages.

2nd semester.

Comparing estimates by type of control, there is an increase in scores by stages (figure 8). At the same time, a significant change occurred between exams in different semesters (3.5 points).

The difference between such related stages as the exam in the 2nd semester and internship is insignificant (0.2 points), which is caused as noted above by the deterioration of grades for internship in the control group.

For the final choice of methods for estimating statistical hypotheses, the samples were tested for distribution normality using the Shapiro-Wilk test. The obtained results are presented in figure 9.

According to the data obtained from the Shapiro-Wilk tests, the null hypothesis about the normality of the distribution laws of the control (p-value $= 1.382 \cdot 10^{-15}$) and experimental groups (p-value =

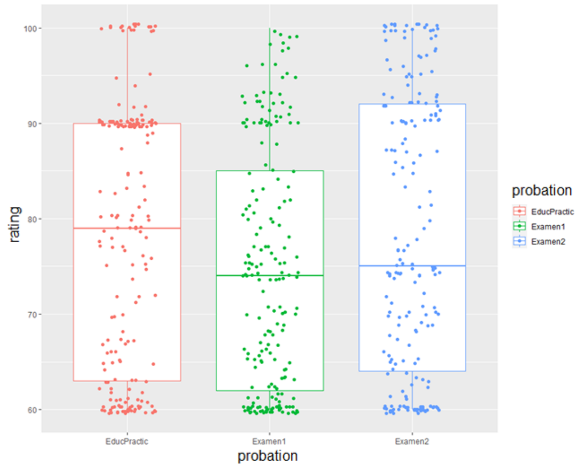


Figure 8: Student performance in terms of assessment stages.

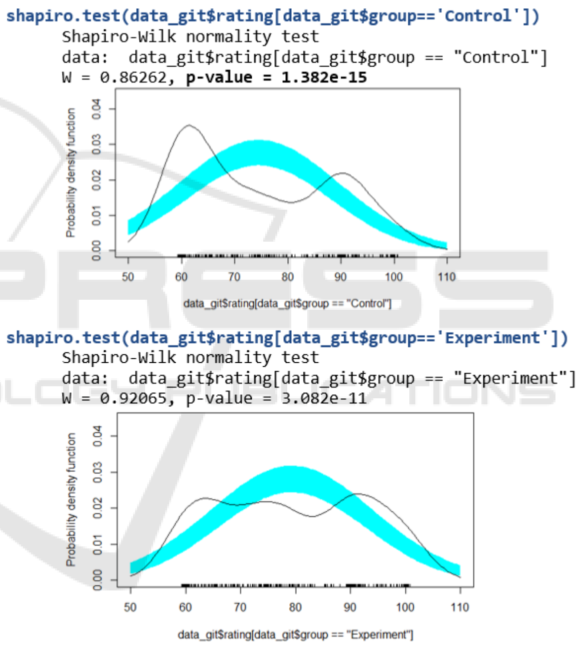


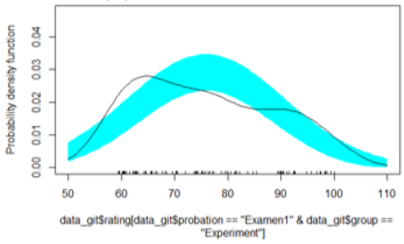
Figure 9: The results of checking the samples (control group and experimental group) for the normality of the distribution.

$3.082 \cdot 10^{-11}$) was rejected, which also illustrated the graphs of the empirical and the theoretical distribution density function.

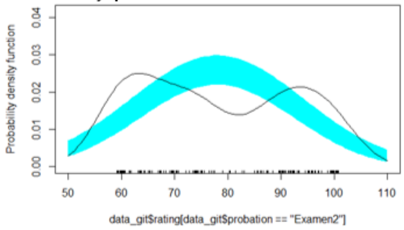
Similarly, the verification of the obtained data was carried out by types of evaluation, which is presented in figure 10.

The obtained results (p-values: $3.536 \cdot 10^{-10}$, $7.687e \cdot 10^{-11}$ and $1.741 \cdot 10^{-11}$) indicate that these samples are not subject to the normal distribution law. Because analysis of variance is used to assess the statistical significance of differences, a group variance test (Barlett test) and an emission test (Bonferroni

```
shapiro.test(data_git$rating[data_git$probation=='Examen1']
          Shapiro-Wilk normality test
     data:  data_git$rating[data_git$probation == "Examen1"]
     W = 0.89979, p-value = 3.536e-10
```




```
shapiro.test(data_git$rating[data_git$probation=='Examen2'])
          Shapiro-Wilk normality test
     data:  data_git$rating[data_git$probation == "Examen2"]
     W = 0.8889, p-value = 7.687e-11
```




```
shapiro.test(data_git$rating[data_git$
                          probation=='EducPractic'])
     Shapiro-Wilk normality test
     data:  data_git$rating[data_git$
                          probation == "EducPractic"]
     W = 0.87758, p-value = 1.741e-11
```

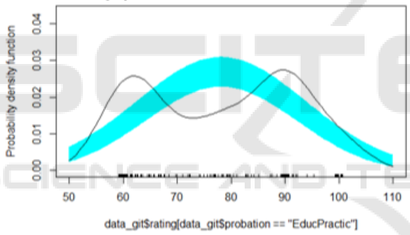


Figure 10: The results of checking the data on the normality of the distribution by type of assessment.

test) were performed.

To conduct the Barlett test, the null hypothesis was put forward that the variances between the groups are equal. The estimate of the variance homogeneity in the groups for different types of estimation is shown in listing (figure 11).

```
bartlett.test(rating ~ probation, data=data_git)
     Bartlett test of homogeneity of variances
     data:  rating by probation
     Bartlett's K-squared = 3.9919, df = 2,
                            p-value = 0.1359
```

Figure 11: Estimation of variance homogeneity for different types of estimation.

The obtained value of p-value = 0.1359 is greater than the significance level – 0.05, so there is no reason to reject the null hypothesis of the test on the equality of variances in groups. The data obtained in this way indicate that the variance data are statistically different. Similarly, the statistical equality of variances for the control and experimental groups of the third hypothesis, which is presented in listing (figure 12), was estimated.

```
bartlett.test(rating ~ interaction(group, probation),
                                   data=data_git)
        Bartlett test of homogeneity of variances
        data:  rating by interaction(group, probation)
        Bartlett's K-squared = 4.1934, df = 5,
                                   p-value = 0.5219
```

Figure 12: Estimation of dispersion homogeneity for control group and experimental group.

As in the previous case, we decide to reject the null hypothesis of equality of variances.

The Bonferroni test listings (figure 13) show no emissions for both types of analysis of variance. Accordingly, the original data meet these two requirements for analysis of variance.

```
outlierTest(aov(rating ~ probation, data = data_git))
        No Studentized residuals with Bonferroni p < 0.05
        Largest |rstudent|:
              rstudent unadjusted p-value Bonferroni p
        309 1.875526           0.061222           NA
outlierTest(aov(rating ~ group*probation,
                                  data = data_git))
        No Studentized residuals with Bonferroni p < 0.05
        Largest |rstudent|:
             rstudent unadjusted p-value Bonferroni p
        17 1.938724           0.053021           NA
```

Figure 13: Bonferroni test.

Therefore, considering the obtained estimates of the normality of the data distribution and the results of power analysis, we choose the following methods for testing hypotheses:

- Hypothesis 1: to check the equality of the average overall scores – Student's t-test; for comparisons of average scores in groups for individual types of evaluation – Mann-Whitney test, as in contrast to the overall estimates, the size of the effect is estimated as average, which does not allow to ignore the normality of the distribution.
- Hypothesis 2: despite the fact that the size of the effect is close to small, in addition to one-way analysis of variance, the results were verified through the Kruskal-Wallis test.
- Hypothesis 3: since it is impossible to perform a power analysis for two-factor analysis of variance with an unbalanced design, the requirement of normality was neglected due to the large samples in each group (Zar, 1996). Moreover, tests for the homogeneity of variances and emissions indicated the possibility of analysis of variance.

After analysis and selection of the above methods, the hypotheses put forward in the study were tested.

According to a preliminary analysis, the difference between the overall scores in the groups is 4.7 points. To assess whether this difference in scores is statistically significant (Hypothesis 1), we chose in the previous stages we chose Student's t-test. Prior to the test, a check was made for the Welch amendment to be considered for samples with different variances (figure 14).

```
var.test(rating~group,data=data_git)
F test to compare two variances
data:  rating by group
F = 1.0324, num df = 296, denom df = 287,
                          p-value = 0.7859
alternative hypothesis: true ratio of
             variances is not equal to 1
             95 percent confidence interval:
 0.8200786 1.2992479
sample estimates:
ratio of variances
           1.032437
```

Figure 14: Test for equality of variances in samples.

As you can see from the data presented in figure 14, the probability of obtaining an error of the first kind is 78.6% with a permissible 5%, to reject the null hypothesis. Therefore, the variances are statistically equal and the Welch correction is not required. The evaluation of the t-test for the overall averages in the two groups is presented in listing (figure 15).

```
t.test(rating~group,data=data_git, var.equal = TRUE)
Two Sample t-test
data:  rating by group
t = -4.2478, df = 583, p-value = 2.512e-05
alternative hypothesis: true difference in means
between group Control and group
Experiment is not equal to 0
95 percent confidence interval:
 -6.961609 -2.559435
sample estimates:
   mean in group Control mean in group Experiment
                74.55892                 79.31944
```

Figure 15: Checking the equality of average overall scores.

According to the obtained results, the actual value of the criterion $-t_f = 4.25$ exceeds the critical $-t_cr = 1.967$ for a given level of significance (0.05), which is necessary to reject the null hypothesis of equality of the two averages. Therefore, we can conclude that the difference between the mean scores between the control and experimental groups (4.7 points) is statistically significant. In this case, with a probability of 95%, this difference will be from 2.6 to 7.0 points. Accordingly, the null hypothesis is rejected.

As noted earlier, a nonparametric Mann-Whitney-Wilcoxon test, which is used for samples without a normal distribution, was calculated to confirm the difference in scores between samples at different stages of the assessment (figure 16).

```
wilcox.test(rating~group,data=data_git
[data_git$probation=='Examen1',], paired = FALSE)
Wilcoxon rank sum test with continuity correction
data:  rating by group
W = 4059.5, p-value = 0.07764
alternative hypothesis: true location
                        shift is not equal to 0

wilcox.test(rating~group,data=data_git
[data_git$probation=='Examen2',], paired = FALSE)
Wilcoxon rank sum test with continuity correction
data:  rating by group
W = 4035, p-value = 0.05729
alternative hypothesis: true location
                        shift is not equal to 0

wilcox.test(rating~group,data=data_git
[data_git$probation=='EducPractic',], paired = FALSE)
Wilcoxon rank sum test with continuity correction
data:  rating by group
W = 3206.5, p-value = 7.217e-05
alternative hypothesis: true location
                        shift is not equal to 0
```

Figure 16: Calculation of the nonparametric Mann-Whitney-Wilcoxon test.

The obtained test results indicate that a statistically significant difference between the averages in the groups is observed only for internship (7.9 points), which is indicated by the value of p-value which is less than the level of significance.

However, it should be noted that in addition to the fact that this test, like all non-parametric has less accuracy, the use of pairwise comparisons with the possibility of analysis of variance is not statistically effective. Therefore, this analysis can be considered as additional in a more efficient analysis of variance.

As noted in the previous analysis, the difference between the exams was significant in contrast to the exam (2 semester) and internship. To test the statistical significance of the differences in the mean score at different stages (Hypothesis 2), the significance of the differences was estimated (figure 17).

```
fit_probation <- aov(rating ~ probation,
                     data = data_git)
             Df Sum Sq Mean Sq F value Pr(>F)
probation     2   1701   850.6   4.555 0.0109 *
Residuals   582 108674   186.7
     ---
Signif. codes: 0 '***'0.001'**'0.01'*'0.05'.'0.1' '1
```

Figure 17: Assessing the significance of differences.

The F-test indicates that there are differences between the groups ($p < 0.05$). As noted at the stage of selection of research methods, an additional analysis was performed according to the non-parametric Kruskal-Wallis test (figure 18).

The results of the Kruskal-Wallis test confirmed the results of one-way analysis of variance.

```
fit_kruskal <- kruskal.test(rating ~ probation,
                            data = data_git)
          Kruskal-Wallis rank sum test
     data:  rating by probation
     Kruskal-Wallis chi-squared = 6.9486, df = 2,
                               p-value = 0.03098
```

Figure 18: Analysis by the nonparametric Kruskal-Wallis test.

Pairwise comparisons using the Tukey test (figure 19) and graphical data (figure 20) concluded that the difference between the exams (Examen2-Examen1 – p-value = 0.029) and the exam for 1 semester was statistically significant. Internship (Examen1-EducPractic – p-value = 0.022) and is -3.7 and 3.5, respectively. The difference between the second semester and the internship is not statistically significant (p-value = 0.994 and intersects on the graph with a vertical January) and is within the statistical error.

```
TukeyHSD(fit_probation)
Tukey multiple comparisons of means
    95% family-wise confidence level
Fit: aov(formula = rating ~ probation, data = data_git)
$probation
                          diff        lwr        upr     p adj
Examen1-EducPractic -3.6871795 -6.9389205 -0.4354385 0.0215893
Examen2-EducPractic -0.1435897 -3.3953308  3.1081513 0.9940823
Examen2-Examen1      3.5435897  0.2918487  6.7953308 0.0287904
```

Figure 19: Conducting pairwise comparisons using the Tukey test.

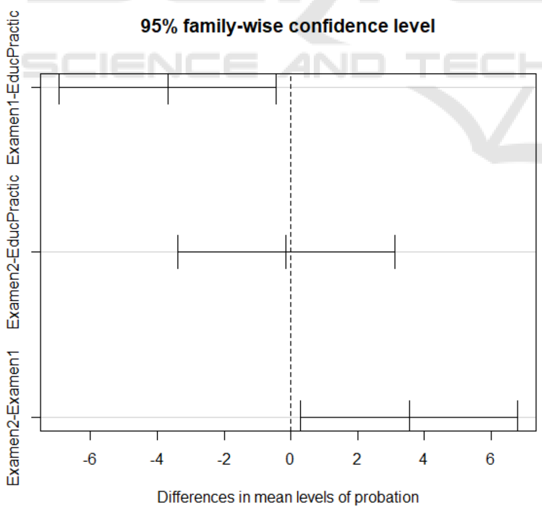


Figure 20: Graphical display of pairwise comparisons.

To check the statistical significance of the difference in the mean score by groups (figure 21) at different stages of the assessment (Hypothesis 3) was used two-factor analysis of variance with intergroup and intragroup factors – mixed analysis model (ANOVA). In our study, the intergroup factor was the distribution of groups of students relative to the control and experimental, and intragroup – the stages of evaluation.

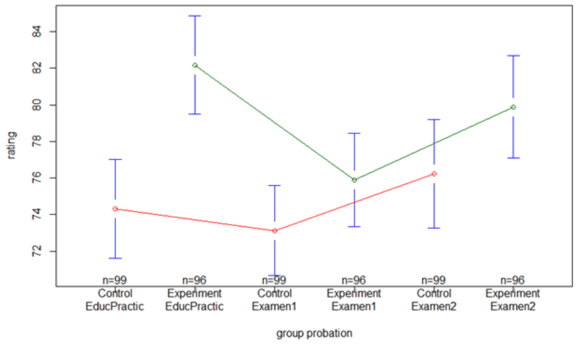


Figure 21: Plot group means and confidence intervals.

The constructed variance model for two-factor analysis of variance is presented in listing (figure 22).

```
fit <- aov(rating ~ group*probation +
      Error(id/probation), data=data_git)
      summary(fit)
Error: id
     Df Sum Sq Mean Sq F value  Pr(>F)
  group       1  3314   3314  9.247 0.00269 **
  Residuals 193  69162    358
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: id:probation
      Df Sum Sq Mean Sq F value   Pr(>F)
  probation         2  1701  850.6   9.257 0.000118 ***
  group:probation   2   728  363.8   3.959 0.019871 *
  Residuals       386  35471   91.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 22: Estimation of a two-factor variance model.

The calculated value of the Fisher criterion, based on the mean squares of the deviations within and between groups and the corresponding degrees of freedom for each of the factors is:

- For groups:

  $\nu BG = m \; - -1 = 2 - 1 = 1$

  $\nu WG = n \; - -m = 288 \; - -2 = 286$

  where m is the number of factor levels (groups), n is the number of observations (students)

  Accordingly, according to tables F of Fisher's test at a significance level of 0.05, the empirical (theoretical, critical) value is F{0.05; 1; 286} = 3.847.

- For types of assessment:

  $\nu BG = m \; - -1 = 3 - 1 = 2$

  $\nu WG = n \; - -m = 195 \; - -3 = 192$

Accordingly, according to tables F of Fisher's test at a significance level of 0.05, the empirical (theoretical, critical) value is F{0.05; 2; 192} = 3.042.

- For stages and interaction of groups and stages

  $\nu BG = (m_1 - 1) * (m_2 - 1) = (2-1) * (3-1) = 2$

  $\nu WG = (m_1 + m_2 - 1) * (n-1) = (2+3-1) + (\frac{99+96}{2} - 1) = 386$,

  where m_1, m_2 is the number of levels for intergroup (group) and intragroup (stages) factors, n is the number of observations (students) in each sample divided into groups and stages – F{0.05; 2; 386} = 3.847.

According to the results of the study, the dependence of the averages for groups and types of assessment is statistically significant, as indicated by the value of the actual F statistics, exceeding the critical values found: 9.247>3.890 and 9.257>3.042 (or according to p-value – 0.0026 and 0.00011, respectively, which is much less than the significance level of 0.05). That is, two-factor analysis of variance proved the preliminary results on the deviation of 1 and the second null hypotheses.

As for the interaction of groups and stages, is the question in assessing the statistical significance of differences by groups at different stages, the value of the obtained Fisher statistics is more than critical – 3.959>3,847 (p-value = 0.019871) indicates the presence of an additional effect from the interaction. That is, there is a significant difference in the average of individual groups by type of assessment. For the obtained model we will make multiple comparisons according to the Tukey test:

- The group comparison (figure 23) confirmed the t-test data on the statistical significance of the difference in the rating score for the control and experimental groups, which is 4.7 points.

```
emmG <- emmeans(fit, ~ group)
pairs(emmG)
contrast              estimate   SE  df t.ratio p.value
Control - Experiment     -4.76 1.57 193  -3.041  0.0027
```

Figure 23: Multiple group comparisons.

- From the obtained comparisons (figure 24) of the averages by stages (types of evaluation) for all groups, we see that the significant difference in the average evaluations corresponds to the data of one-factor analysis for the second hypothesis.

- Pairwise comparisons of the difference between the mean scores for the groups divided by both factors are significant for most comparisons (figure 25).

```
emmG <- emmeans(fit, ~ group)
pairs(emmG)
contrast               estimate    SE  df t.ratio p.value
EducPractic - Examen1     3.726 0.971 386   3.838  0.0004
EducPractic - Examen2     0.176 0.971 386   0.181  0.9820
Examen1 - Examen2        -3.550 0.971 386  -3.657  0.0008
```

Figure 24: Multiple comparisons by type of assessment.

```
emm <- emmeans(fit, ~ group*probation)
pairs(emm)
contrast                                        estimate   SE  df t.ratio p.value
Control EducPractic - Experiment EducPractic      -7.874 1.93 390  -4.089  0.0007
Control EducPractic - Control Examen1              1.172 1.36 386   0.860  0.9556
Control EducPractic - Experiment Examen1          -1.593 1.93 390  -0.827  0.9624
Control EducPractic - Control Examen2             -1.939 1.36 386  -1.423  0.7128
Control EducPractic - Experiment Examen2          -5.582 1.93 390  -2.899  0.0453
Experiment EducPractic - Control Examen1           9.046 1.93 390   4.698  0.0001
Experiment EducPractic - Experiment Examen1        6.281 1.38 386   4.540  0.0001
Experiment EducPractic - Control Examen2           5.935 1.93 390   3.082  0.0266
Experiment EducPractic - Experiment Examen2        2.292 1.38 386   1.656  0.5617
Control Examen1 - Experiment Examen1              -2.765 1.93 390  -1.436  0.7052
Control Examen1 - Control Examen2                 -3.111 1.36 386  -2.283  0.2034
Control Examen1 - Experiment Examen2              -6.754 1.93 390  -3.508  0.0067
Experiment Examen1 - Control Examen2              -0.347 1.93 390  -0.180  1.0000
Experiment Examen1 - Experiment Examen2           -3.990 1.38 386  -2.883  0.0474
Control Examen2 - Experiment Examen2              -3.643 1.93 390  -1.892  0.4087

P value adjustment: tukey method for comparing a family of 6 estimates
```

Figure 25: Multiple comparisons by groups and types of assessment.

Comparisons between the same types of evaluations between the control and experimental groups showed:

- The difference in the average scores on the exam for the first semester (Control Examen1 – Experimental Examen1) is statistically insignificant (p-value = 0.7052>0.05);
- The difference in the average scores on the exam for the second semester (Control Examen2 – Experimental Examen2) is statistically insignificant (p-value = 0.4087>0.05);
- The difference in the average scores on internship (Control EducPractic – Experimental EducPractic) is statistically significant (p-value = 0.0007<0.05) and is 7.8 points.

These results confirmed the results of the Mann-Whitney-Wilcoxon test that the effectiveness of the implementation of experimental methods affected the results of internship, during which students actively used the cloud service GitHub to implement collective projects.

Assessing separately the dynamics of differences in each of the groups by stages, we see that in the control group, each change in the average score between different types of assessment is insignificant, which indicates the absence of any dynamics in traditional learning: "Control Examen1 – Control Examen2" – p-value = 0.2034, "Control EducPractic – Control Examen2" – p-value = 0.7128, "Control EducPractic – Control Examen1" – p-value = 0.9556. In contrast, for the experimental group where such changes were statistically significant.

6 CONCLUSIONS

The educational projects on programming are an effective method for shaping the professional and personal competencies of future IT specialists. To work on educational projects, you should use modern cloud services for collaborative IT development, such as GitHub. The most effective features of this service are the possibility of collaborative development of software (i1), the convenience of bug tracking (i3), and the convenience of the code editor (i7), which are determined by the statistical processing of student peer review. Other features of this service that have also been explored include the ability to manage code versions (i2), the ability to organize and plan teamwork (i4), the ability to communicate (i5), the ability to support platforms (i6), security and privacy (i8), availability of wiki pages (i9).

The GitHub cloud service can be applied to complete mini projects, group or individual work, term papers, or during the academic training of the students. The examples of using GitHub discussed in the paper show that the specific features of this service completely satisfy the needs of students of IT profession in the implementation of the tasks of educational projects on programming. And affects the formation of professional competence of future specialists in information technology, as evidenced by the results of the study. The obtained results of the experimental group are higher by 5.93% than in the control group. Crucial to the formation of professional competence are the use of cloud service in the implementation of collective projects for software development during training.

Further research can be aimed at theoretical substantiation and development of methods of flexible training of future information technology specialists using services for joint software development in the implementation of educational projects, as the technology of agile learning is closest to real conditions in software development.

REFERENCES

Angulo, M. A. and Aktunc, O. (2019). Using GitHub as a Teaching Tool for Programming Courses. In *2018 Gulf Southwest Section Conference*, Austin, TX. https://peer.asee.org/31594.

Beshero-Bondar, E. E. and Parker, R. J. (2017). A GitHub Garage for a Digital Humanities Course. In Fee, S. B., Holland-Minkley, A. M., and Lombardi, T. E., editors, *New Directions for Computing Education: Embedding Computing Across Disciplines*, pages 259–276. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-54226-3_15.

Bonakdarian, E. (2017). Pushing Git & GitHub in Undergraduate Computer Science Classes. *J. Comput. Sci. Coll.*, 32(3):119–125.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, 2 edition. https://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf.

Feliciano, J., Storey, M.-A., and Zagalsky, A. (2016). Student Experiences Using GitHub in Software Engineering Courses: A Case Study. In *Proceedings of the 38th International Conference on Software Engineering Companion*, ICSE '16, page 422–431, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/2889160.2889195.

Fiksel, J., Jager, L. R., Hardin, J. S., and Taub, M. A. (2019). Using GitHub Classroom To Teach Statistics. *Journal of Statistics Education*, 27(2):110–119. https://doi.org/10.1080/10691898.2019.1617089.

GitHub Education (2020). 2020 GitHub Education Classroom Report: Insights from the next generation of software developers. https://education.github.com/classroom-report/2020.

Glazunova, O. G., Korolchuk, V. I., Parhomenko, O. V., Voloshyna, T. V., Morze, N. V., and Smyrnova-Trybulska, E. M. (2022). Methodology for using Cloud-oriented Environment for Flipped Learning of the Future IT Specialists. In Semerikov, S., Osadchyi, V., and Kuzminska, O., editors, *Proceedings of the 1st Symposium on Advances in Educational Technology - Volume 1: AET*, pages 445–460. INSTICC, SciTePress. https://doi.org/10.5220/0010925100003364.

Glazunova, O. G. and Shyshkina, M. (2018). The Concept, Principles of Design and Implementation of the University Cloud-based Learning and Research Environment. In Ermolayev, V., Suárez-Figueroa, M. C., Yakovyna, V., Kharchenko, V. S., Kobets, V., Kravtsov, H., Peschanenko, V. S., Prytula, Y., Nikitchenko, M. S., and Spivakovsky, A., editors, *Proceedings of the 14th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume II: Workshops, Kyiv, Ukraine, May 14-17, 2018*, volume 2104 of *CEUR Workshop Proceedings*, pages 332–347. CEUR-WS.org. https://ceur-ws.org/Vol-2104/paper_158.pdf.

Gunnarsson, S., Larsson, P., Månsson, S., Mårtensson, E., and Sönnerup, J. (2017). Enhancing Student Engagement Using GitHub as an Educational Tool. In *Introduction to Teaching and Learning in Higher Education*. Genombrottet, Lunds tekniska högskola. https://lucris.lub.lu.se/ws/files/27854342/group1_github_final.pdf.

Haaranen, L. and Lehtinen, T. (2015). Teaching Git on the Side: Version Control System as a Course Platform. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '15, page 87–92, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/2729094.2742608.

Hsing, C. and Gennarelli, V. (2019). Using GitHub in the Classroom Predicts Student Learning Outcomes and Classroom Experiences: Findings from a Survey of Students and Teachers. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, SIGCSE '19, page 672–678, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3287324.3287460.

Kelleher, J. (2014). Employing git in the classroom. In *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*, pages 1–4. https://doi.org/10.1109/WCCAIS.2014.6916568.

Kertész, C.-Z. (2015). Using GitHub in the classroom - a collaborative learning experience. In *2015 IEEE 21st International Symposium for Design and Technology in Electronic Packaging (SIITME)*, pages 381–386. https://doi.org/10.1109/SIITME.2015.7342358.

Korolchuk, V. (2019). Cloud services for collective projects preparation processes of future IT-professionals: Analysis and selection criteria. *New pedagogical thought*, 100(4):46–51. https://doi.org/10.37026/2520-6427-2019-100-4-46-51.

Marquardson, J. and Schuetzler, R. M. (2019). Teaching Tip: Learning by Teaching through Collaborative Tutorial Creation: Experience using GitHub and AsciiDoc. *Journal of Information Systems Education*, 30(1):10–18. https://jise.org/Volume30/n1/JISEv30n1p10.html.

Metz, C. (2015). How Github Conquered Google, Microsoft, and Everyone Else. https://www.wired.com/2015/03/github-conquered-google-microsoft-everyone-else/.

Morze, N. V. and Glazunova, O. G. (2013). What Should be E-Learning Course for Smart Education. In Ermolayev, V., Mayr, H. C., Nikitchenko, M. S., Spivakovsky, A., Zholtkevych, G., Zavileysky, M., Kravtsov, H., Kobets, V., and Peschanenko, V. S., editors, *Proceedings of the 9th International Conference on ICT in Education, Research and Industrial Applications: Integration, Harmonization and Knowledge Transfer, Kherson, Ukraine, June 19-22, 2013*, volume 1000 of *CEUR Workshop Proceedings*, pages 411–423. CEUR-WS.org. https://ceur-ws.org/Vol-1000/ICTERI-2013-p-411-423-MRDL.pdf.

Zar, J. H. (1996). *Biostatistical Analysis*. Prentice Hall, Upper Saddle River, 3 edition.