

Student Perspectives on Source-Code Plagiarism: Case Study of Three Programming Courses

Matija Novak

University of Zagreb / Faculty of Organization and Informatics, Varaždin, Croatia
matija.novak@foi.unizg.hr

Abstract - Plagiarism is a serious offence and students plagiarise for various reasons. But sometimes students plagiarise and don't even know that they are doing it. In this research a case study was conducted where questions, which describe different situations, were given to students. Questions are taken from an existing study which was conducted in the UK to get the perspective of academics. Similar questions were later given to students in the UK and China to get the student perspective. In this case, Croatian students from three programming courses (at three different levels) were questioned. In this paper the most interesting findings of student perspectives on plagiarism are presented together with some comparisons to the previous studies.

Keywords - plagiarism; academic offence; student perspective; programming

I. INTRODUCTION

Every teacher has to deal with plagiarism on a regular basis. There are two main aspects while dealing with plagiarism: detection and prevention. While detection is necessary, the first step in dealing with plagiarism should be prevention. There are many techniques to prevent plagiarism [1] and academic dishonesty in general. One technique is to educate students on that topic. It might seem that it is obvious what plagiarism is but research like [2] shows that this is not the case.

This is especially true when talking about source-code plagiarism which is the focus of this paper. Simple definition of plagiarism, that can be found in dictionaries, is that plagiarism is “the act of taking the writings of another person and passing them off as one’s own” [3]. While this is correct there are different aspects that are not covered, especially when talking about source-code plagiarism.

Therefore, researchers try to define source-code plagiarism more precisely. A good definition is given by Cosma and Joy in [4] which is: “*Source-code plagiarism in programming assignments can occur when a student reuses ... source-code authored by someone else and, intentionally or unintentionally, fails to acknowledge it adequately ... , thus submitting it as his/her own work. This involves obtaining ... the source-code, either with or without the permission of the original author, and reusing ... source-code produced as part of another assessment (in which academic credit was gained) without adequate acknowledgement ... The latter practice, self-plagiarism, may constitute another academic offence*”.

In this paper the goal is to find out the student perspective on plagiarism. This study is conducted in Croatia on three programming courses that are on three different levels. *The main idea is to compare whether there are any differences in perception between students in these courses.* In addition, the results are compared to existing studies which analyse academic (in UK) and student perspectives (in UK and China).

The rest of the article is structured as follows. Section 2 describes the related work and Section 3 the methodology. Results are presented in Section 4 and in Section 5 discussion is given. Section 6 concludes.

II. RELATED WORK

There are different studies [5,6,7,8] that analyse student and academic perspectives on source-code plagiarism. In 2006, Cosma and Joy [6] conducted a survey about academic perspective on source-code plagiarism. The goal was to find out what constitutes source-code plagiarism. Two issues have received controversial responses regarding source-code reuse and self-plagiarism. The definition of source-code plagiarism given in introduction was partially based on this study.

Later in 2011, Joy, Cosma, Yau and Sinclar [7] conducted a similar survey on student perspectives. Based on the results they divided the students into three groups: the ones that plagiarise deliberately, those that do so negligently and those who genuinely do not understand what plagiarism means. Most misunderstood issues were about source-code reuse from previous assignments, citation, limits of collaboration and when code was converted to a different language. When compared to the academic perception the common issue is code reuse.

In [6,7] the research was done in the UK, while in 2014 a study [8] was made which compared student perspectives between UK and Chinese students. This study showed that, in most areas, there is a common understanding and misunderstanding of what constitutes source-code plagiarism. Although, there were three topics that UK students understand and students from China don't. These were: copying code from a book without acknowledgement, paying someone, and copying and submitting another student's work without consent.

In recent years, while there are studies [9,10] that deal with student perspective on plagiarism focusing on text, there are no new studies focusing on source-code.

III. METHODOLOGY

In this study a survey was conducted on three programming courses in Croatia. Course A is a course at a first year university graduate level. Course B is a course at third year university undergraduate level. Course C is a third year professional undergraduate level. Students in all three courses had already passed some programming courses so it is expected that they know something about source-code plagiarism.

At first the idea was to use the same questions that were used in [7] or [8]. Unfortunately, the original survey and exact questions used were not available. Fortunately, the complete questions used in [6] were found in [11] which is a detailed report of the whole study. While [6] focuses on academics from the papers it is visible that the questions in [7,8] are similar to questions in [6] so that it is possible to compare some results.

A. Survey execution and layout

The survey was performed online over Learning Management System (LMS) Moodle. Access to the survey had only students that enrolled in the course. *The survey was translated to Croatian to eliminate the problem of misunderstanding the questions.*

The survey was divided into five parts. First part was just one yes/no question: *“Do you know what the term source-code plagiarism means?”*.

Second part gave the students two definitions of source-code plagiarism. One was the definition presented in the introduction. The second definition was [12]:

“Plagiarism, in programming assignments, is the act of taking a significant amount of source-code parts (up to the entire source-code) from other students or from the Internet and using it without noting its origin. A ‘significant amount’ means that the similarity between two solutions of a programming assignment is high enough that an expert (teacher, ethical board, etc.) considers specific student work as sufficiently ‘real’ plagiarism to accuse the student of plagiarism.”

These two definitions were considered most accurate on what constitutes as source-code plagiarism. If interested in other definitions there is a systematic literature review paper [13] which gives good overview of definitions about source-code plagiarism. The reason for giving students these definitions was to give them understanding of what was meant by source-code plagiarism. There was a suspicion, before starting the survey, that some students are totally ignorant of what source-code plagiarism means, which was confirmed.

One might argue that these definitions do not include open source-code, but this will depend on the course. If using open source-code is allowed then this can not be considered plagiarism. For example, in the case of the three courses students need to make the assignments themselves from scratch. Now it is not forbidden to gather ideas from others, but if they do use code from someone else they have a form to fill out where they can state which parts of the code were taken and from where.

Third part contained three questions about what source-code plagiarism can involve. *These questions had the answers Agree, Disagree, Neither Agree or Disagree (i.e. don't know).* Part four contained 15 questions (or better said statements) that describe different situations. *These questions had answers which are: Academic offence, Not an academic offence and don't know.* Here a note was given: *“When you answer the following questions, think about whether this would be considered an academic offence or not, in a graded assignment.”*

Part five contained only two questions which were:

1. What do you think is the minimum percentage that your code needs to be similar to another student so that the teacher starts suspecting you of plagiarism and looks into it in more detail? Answer from 0% to 100%.
2. Here you write any comments if you like.

B. Questions of part three and four

The questions that were given to the students were not all questions that exist in [11]. The reason for this was that some questions were not that relevant since the students do not have these parts in their assignments in these courses. For example: *“Plagiarism in programming assignments can involve documentation of computer programs.”* In courses where the study is done students do not have to write documentation. Also, some questions deal with non-graded assignments which were not the focus of this study. This reduction also had another purpose to limit the number of questions. If the survey is longer, less students want to spend the time to answer it.

The questions in part three were:

1. Plagiarism in Programming assignment can involve the Source-code of a computer program.
2. Plagiarism in Programming assignment can involve the Comments within the source code.
3. Plagiarism in Programming assignment can involve the User-interface of a computer program.

The questions in part four were:

1. A student reproduces/copies someone else source-code without making any alterations and submits it without providing any acknowledgements.
2. A student reproduces/ copies someone else source-code. Adapts the code to his/her own work and submits it without providing any acknowledgements.
3. A student pays another person (other than a student on the same module) to create part or whole of source-code and submits it as his/her own work.
4. A student pays a fellow student on the same module to create part or whole of source-code and submits it as his/her own work.

5. A student steals another student's source-code and submits it as his/her own work without making any alterations.
6. A student steals another student's source-code, edits and submits it as his/her own work
7. For a group assignment, students between different groups exchange parts of source-code with the consent of their fellow group members, and integrate the borrowed source code within their work as if it was that group's own work.
8. For a group assignment, students between different groups exchange parts of source-code without their fellow group members knowing, and integrate the borrowed source code within their work as if it was that group's own work.
9. Assume that students were not allowed to resubmit material they had originally created and submitted previously for another assignment. For a graded assignment, a student has copied parts of source-code that he had produced for another assignment without acknowledging it.
10. Two students work together for programming for a programming assignment that requires students to work individually and students submit very similar source-codes.
11. For an assignment, the student has copied source-code from a book and has intentionally not provided any acknowledgments.
12. For an assignment, the student has copied source-code from a book and has unintentionally not provided any acknowledgments.
13. Providing pretend reference (i.e. references that were made-up by the student that do not exist)
14. Providing false reference (i.e. referenced exist but not match the source-code that was copied)
15. Modifying the program output to make it seem as if the program works.

IV. RESULTS

In total 206 students filled out the survey, the response rate of the survey was as follows:

- Course A: 87.5% response rate, 70 responses out of 80 students;
- Course B: 70.4% response rate, 95 responses out of 135 students;
- Course C: 77.6% response rate, 41 response out of 53 students.

The results were first analysed together and then per course, the results per course are further presented in parentheses. Afterwards, to see if there are any differences in the perception of students between the three courses the statistical analysis was performed using chi-square test.

The hypothesis for each question was: Null hypothesis (H0) – The student perception is not dependent on the

course. Alternative hypothesis (H1) – The student perception is dependent on the course.

Regarding question 1 (in the first part of the survey) where students were asked if they know what source-code plagiarism means on average 85% answered yes (Course A - 91.4%, Course B – 86.3%, and Course C – 78%). These results are not unexpected since the students are at least three years into their study program. Based on this question it can be said that at least most of them have heard of the term source-code plagiarism. Based on the responses of other questions it is clear that they do not understand the meaning of source-code plagiarism completely.

A. Results of part three questions

The results of the survey's third part were interesting. First question in this part was that source-code plagiarism can involve source-code. It was expected that everyone would agree, but surprisingly only 71.5% agreed (Course A – 78.6%, Course B – 72.6% and Course C – 63.4%). Less than 10% answered that they don't know. First explanation was that students did not honestly fill out the survey but by analysing other questions from these respondents it does not seem so. Especially, some of the really good comments in the last question came from students that answered this question with disagreement. It might be that they do not know what source-code plagiarism is. But, since definitions were given to them before this part, the only explanation might be they did not understand the question.

Regarding comments, around 38% of students agree that it might be part of source-code plagiarism (Course A – 28.6%, Course B – 45.3%, Course C – 41.5%) and 21% don't know. Regarding user interface, around 52% (Course A – 62.9%, Course B – 56.8% and Course C – 36.6%) agreed that it can involve source-code plagiarism and 16% don't know.

When this is compared to what academics in the UK have answered [11] (comments ~68% agree, user interface ~60% agree) it is not surprising that students have lower agreement. What is interesting is that Course A, which is on a university graduate level, has the lowest agreement for comments (28.6%) and highest for user interface (62.9%). On the other hand, Course C, which is a professional undergraduate, is exactly opposite lowest for user interface (36.6%) and higher for comments (41.5%).

The explanation might be that students in higher years know that comments can be easily deleted and modified without changing the logic. Also they already know that comments are not part of the real functionality while the user interface is. Also, such students build quite complex user interfaces. Therefore, source-code plagiarism highly involves user interface and not so much comments. Students in Course C still use comments to explain code for themselves and see it as a possibility an indication of plagiarism if these comments are the same. In addition, these students build quite simple user interfaces so plagiarism is more likely to happen in comments than in user interfaces.

TABLE I. RESULTS OF THE THIRD PART

Q	Percentage (%)							
	Course A		Course B		Course C		Total	
	AO	NO	AO	NO	AO	NO	AO	NO
1	93	6	100	0	95	5	96	4
2	47	37	55	34	68	22	57	31
3	74	13	62	16	68	27	68	19
4	76	11	66	16	68	27	70	18
5	97	3	100	0	83	17	93	7
6	77	10	76	10	73	20	75	13
7	19	54	30	55	44	44	31	51
8	77	16	74	11	73	17	75	14
9	36	50	31	46	54	24	40	40
10	33	44	23	55	24	56	27	52
11	43	34	85	10	90	7	73	17
12	43	34	41	36	51	27	45	32
13	64	13	68	10	76	10	69	11
14	49	16	58	13	54	15	53	14
15	33	43	33	44	24	46	30	44

Q-question, AO- academic offence, NO – not an academic offence
 marked bold – significant differences between courses

B. Results of part four questions

This part of the survey had most questions (Q). The results (in percentages) are presented in Table 1 for answers plagiarised (academic offence - AO) and not plagiarised (not an academic offence - NO). The terms plagiarism and academic offence are in this section used interchangeably. While it is clear that they are not the same, in this study there were no answers about other academic offences. The don't know is not given because of the space but can easily be calculated since the three answers add up to 100%. Most questions do not have a significant statistical difference between groups (course). For those that have it will be explicitly stated what the χ^2 and p value are.

For Q1 (copying without modification) there is a clear understanding (96% agree) that it is AO. For Q2 (copying with modification) there is a big decrease in agreement (57% agree) that this is AO. Now the reason is explained in the last question in part five. Students could not decide how big the modification is. And this is something that corresponds to the answers of academics in [11] that it depends on the degree of adaptation. Now in the three courses the rule is that no adaptation is allowed so it makes sense that many answered that it is plagiarism.

For Q3 (paying third person) and Q4 (paying another student) the results are almost equal. Around 70% agree that it is AO. This percentage is almost the same in [11] where ~75% agree. Here the academics say that this is not plagiarism rather some other academic offence so students might have the same reasoning. Although, this was not stated in the comments so we don't know. In this question there is a significant difference in perception between students of the three courses with $\chi^2(4)=9.972, p=0.041$.

For Q5 (stealing without modification) the understanding is clear 93% say it is AO, while for Q6 (stealing with modification) this drops to 75%. Here again the explanation is similar as for the difference between Q1 and Q2 and that is the unknown factor how big the adaptation is. The academics on both questions were around 83% that it is plagiarism. What is interesting is that I Q5 there exists a difference between groups (courses) with $\chi^2(2)=20.561, p<0.001$. This is obvious because in Course 3 there is 17% that say this is NO, while in Course A 0% say NO and Course B 3% say NO. The explanation might be that students on the university level are much more ethical. Similar observation is present in Q6, but there is no significant difference between courses.

Regarding group work Q7 (code exchange with knowledge of the group) and Q8 (code exchange without knowledge of the group) there is a big difference. When other team members know that the exchange is taking place only 30% say this is AO. On the contrary when it is done without the knowledge of team members 74% say this is plagiarism. On both questions academics here agree with ~83% that it is plagiarism. From this question it can be said that students do not understand that AO is taking place even when the other side agrees to the exchange. Also, it seems that students don't understand the boundaries of what is allowed. As with Q5 there is a significant difference between groups (courses) in Q7 with $\chi^2(4)=10.5177, p=0.033$. In the two university courses there is a high percentage that it is not an academic offence. In Course A only 19% and in course B only 30% think it is AO, while in Course C 44% think it is AO. It is around 50% for all three courses that it is NO, the difference comes from the don't know answers. The reason for the difference might be that students at the university level have much more courses that have group work and on the professional level it is more individual projects.

Question Q9 goes into the domain of self-plagiarism which academics in [11] don't agree with and only 29% think this is plagiarism. Students here are more strict with 40% saying that it is plagiarism and 20% don't know.

Question 10 (two students working together while they should be working alone) has only 27% say this is plagiarism and 22% don't know, while ~68% academics say it is. This question confirms the statement made in question Q7 that there is a lack of understanding of the boundaries of what is allowed.

Questions 11-14 are about copying from books and citations. In question Q11 (copying from book without citation intentionally) 73% agree that it is AO, but there is a significant difference between courses with $\chi^2(4)=44.657, p<0.001$. The difference is mostly due to Course A where only 43% say it is AO and 23% don't know while in the other two courses it is more than 85% that it is AO.

In question 12 (copying from book without citation unintentionally) 45% say it is AO while ~68% of academics [11] say it is. Here it is shown that students are much more relaxed. In question 13 (copying from book with intentionally putting made up references) 69% think

it is AO and in Q14 (copying from book with intentionally putting wrong references) 53% think it is AO. Q13 and Q14 have lower agreement that it is AO than for Q11 but higher than Q12. When compared to Q11-Q14 it can be concluded that students think it worse when it is intentional the same as academics. Still, academics are much more strict since from their point of view how can one know if it was intentional or not. Also, referencing something (Q13 and Q14) even if it is not correct, more students are uncertain whether this is AO (larger percentage ~25% don't know answers) while when nothing is cited it is more clear to them that it is AO.

Regarding modifying the output 30% think it is AO while 25% don't know. Here academics [11] say ~20% that it is and the reasoning is that it is not plagiarism rather than some other kind of cheating. Some comments from students also state that this question is not plagiarism but rather an incorrect implementation.

C. Results of part five questions

In this last part the question was how big the similarity in percentages should be to be considered plagiarism. Answers in Course A ranged from 5% to 90% with an average of 56%. Answers in Course B ranged from 5 to 90% with an average of 59%. Answers in Course C ranged from 5% to 100% with an average of 44%.

In Figure 1 student answers are shown in graph form. The graph shows the frequency of percentages in ranges of 10 percent. From this graph it is visible that Course C is different from Course A and Course B. In Course A and Course B more students gave higher percentages while in Course C more students gave lower percentages. What is visible from this is that there is no agreement and as with academics [11] it depends.

The last question were comments and partially they were mentioned to explain the results of previous questions. While most students did not write anything there were some good observations.

Some comments did go into the direction that they think that students who study together, have the same courses and learn from the same sources will produce the same code. Such a statement is not correct size just by using the same sources for learning it will probably produce higher similarities but not the same. Also, when talking in context of the three courses where projects are relatively large the overall similarity will be smaller than maybe in courses with small assignments. Usually students see only one project, maybe two and do not have the overview of all projects in course as teachers so it is not unexpected that the opinions differ.

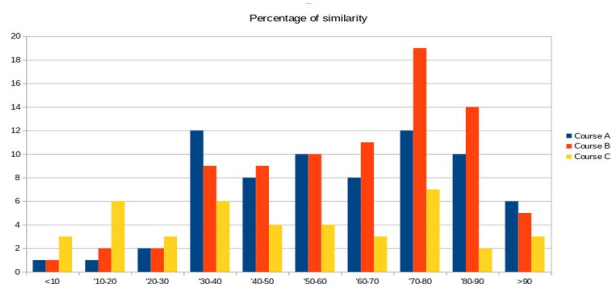


Figure 1. Frequency of similarity percentage in student responses

Other comments got into discussion that just if something is similar it is not necessarily plagiarism. While this is true and agrees with many studies, it is going into the domain of detection and this is not the focus here.

One issue that was pointed out regarding the questions itself was that it was not clear how much copying/adapting are the questions talking about. This is fair comment and such questions could be split, in the future, into multiple questions where the amount is specified.

Some comments talked about code reuse and copying from websites like stack-overflow and how they think it is not needed to be cited. Here again is the problem of how much is necessary to be cited. Another issue is what is the correct way to cite source-code.

Good suggestions in the comments were that students should be educated. Another was that if there is suspicion of plagiarism they should be asked to explain the code written. Explaining the code is a partial solution since it could be that the student has the knowledge to explain but does not know how to write it from scratch. Also, it could be that it was just to make a shortcut to do the assignment faster. Are such situations then acceptable?

V. DISCUSSION

In the previous section the results were discussed on a per question basis in this section a discussion is given based on all results. Also it compares the results to [8]. As it can be seen, questions describe different situations: copying from students, paying someone for doing their assignments, stealing, group work, self plagiarism and modifying output.

Based on the first question the assumption that there are students that do not know what is meant by source-code plagiarism is confirmed. Around 15% of students do not know what is source-code plagiarism. Although 85% say they know what source-code plagiarism means, by analysing other questions it was shown that this understanding is limited.

In the case of copying (Q1) and stealing (Q5) without modification the understanding is excellent and almost all students (above 90%) know this is considered plagiarism. In [8] such situations are understood by 80% of UK students and only 43% of Chinese students.

When modification is involved (Q2, Q4, Q6, Q8) or when talking about paying someone (Q3) there is less understanding, but still many students (approximately 60%) understand that. When looking at [8] situations of modification in the UK understanding is around 70% and in China 57%. While paying someone in the UK 80% of students understand and in China only 16%.

Copying from books (Q11) when no reference is given most students understand (around 73%). On the other hand, in situations where it was copying from books and giving false references (Q14), made up references (Q13) or unintentional not referencing (Q12) the understanding is moderate (between 45-70%). According to [8] regarding copying from books without referencing 81% of UK students understand this is plagiarism but only 26%

from China understand that. On the topic of incorrect references in the UK 11% understand and in China 10%.

The most problematic domain for students in the three courses was collaboration (Q7, Q10) where the understanding is quite low (around 30%). According to [8] the situation of collaboration on individual work is 34% in the UK and 8% in China, while group collaboration is 32% in the UK and 14% in China.

Regarding self plagiarism (Q9) understanding is moderate (around 40%). Based on [8] in the UK the understanding is around 3% and in China around 4%.

Lastly, the modification of program output is considered plagiarism by 30% and 44% do not think it is plagiarism. In [8] such a situation is not considered plagiarism and 80% of UK students and 45% of Chinese students understand.

From these results there is the *indication* that Croatian students are better in understanding self-plagiarism, copying and stealing without modification and collaboration than students from the UK and China. On the other hand they have worse understanding when copying with modification is done or copying from books then students in the UK but better than Chinese students. *These results are not completely accurate since the questions are similar but not exactly the same.* Nonetheless, a good indication is given what to expect.

Overall students in Croatia in the three courses have similar understandings of source-code plagiarism regardless of the study program. The only significant differences appear in cases of: paying someone (Q3), stealing without modification (Q5), code exchange in group work with knowledge of the group (Q7) and copying from a book without citation intentionally (Q11).

Additionally, based on questions in part three in all three courses students have similar level of understanding that source-code plagiarism can involve.

VI. FUTURE WORK

This survey gave a good first view on students' perspective about source-code plagiarism. In the future it is planned to make additional surveys on other programming courses to see if the results are consistent.

As indicated in the comments, some questions need to be more precise to get more clear responses so in future it is planned to make this improvement.

Another possibility for future work is to give students real source-code which is considered plagiarism by academics and code that is not considered plagiarism and to ask them to make a decision. This would give inside if there is an agreement.

VII. CONCLUSION

Not all students know what source-code plagiarism means and those that do know, do not understand it completely. Plagiarism is still a highly researched topic and even academics don't agree on all topics what

constitutes plagiarism so it is of no surprise that students don't know.

Based on the presented results it can be said that students from Croatia in the three programming courses have a solid understanding of what constitutes source-code plagiarism in most cases. Almost all students understand that copying and stealing is plagiarism but when it comes to modifications the understanding starts to drop. In comparison to students in the UK and China it seems that Croatian students are somewhere in between.

The most misunderstood topic is about collaboration, self-plagiarism and where can source-code plagiarism occur. It can be stated based on this research that students are not adequately educated on the topic of source-code plagiarism. Nonetheless, this is not excuse for allowing plagiarism. While special attention needs to be about what are the boundaries of collaboration on a particular course.

REFERENCES

- [1] T. Hodgkinson, H. Curtis, D. MacAlister, and G. Farrell "Student academic dishonesty: The potential for situational prevention" *J. Crim. Just. Educ.* 27(1), pp. 1–18, 2016 DOI:<https://doi.org/10.1080/10511253.2015.1064982>
- [2] Drisko, J. W. "What Is Plagiarism, How to Identify It, and How to Educate to Avoid It" *Journal of Social Work Education*, pp. 1-12, 2022
- [3] Encyclopædia Britannica Inc, "plagiarism" *Encyclopædia Britannica*, 2015. [Online]. Available: <http://www.britannica.com/topic/plagiarism>. [Accessed: 30-Jan-2023].
- [4] Cosma, G., and Joy, M. "Towards a definition of source-code plagiarism" *IEEE Transactions on Education*, 51(2), pp. 195-200, 2008
- [5] Yeo, S. "First-year university science and engineering students' understanding of plagiarism" *High Education Research & Development*, 26(2), pp. 199-216, 2006
- [6] Cosma, G., and Joy, M. "Source-Code Plagiarism: a Uk Academic Perspective" *Annual Conference of the HEA Network for Information and Computer Science* pp. 116–120, 2006
- [7] Joy, M., Cosma, G., Yau, J. Y.-K., and Sinclair, J. "Source Code Plagiarism - A Student Perspective" *IEEE Transactions on Education*, 54(1), pp. 125–132, 2011 DOI:<https://doi.org/10.1109/TE.2010.2046664>
- [8] Zhang, D., Joy, M., Cosma, G., Boyatt, R., Sinclair, J., and Yau, J. "Source-code plagiarism in universities: a comparative study of student perspectives in China and the UK" *Assessment & Evaluation in Higher Education*, 39(6), pp. 743–758, 2014 DOI:<https://doi.org/10.1080/02602938.2013.870122>
- [9] Murugathas, K., and Santharoban, S. (2022). "Students' perspectives on plagiarism detection" *VINGNANAM Research Conference*, Shri Lanka, 2002
- [10] Raisyah, N. "Understanding on Plagiarism from English Language Education Students' Perspective", Doctoral dissertation, *Universitas Islam Kalimantan MAB*, 2022
- [11] Georgina, C., and Mike, J. "Source-code plagiarism: A UK academic perspective" *Research Report CS-RR-422*, Department of Computer Science, University of Warwick, 2006, Available: https://wrap.warwick.ac.uk/61520/1/WRAP_Joy_cs-rr-422.pdf [Accessed: 30-Jan-2023]
- [12] Author, Author. "Title" Doctoral dissertation, *University of xxx. Faculty of xxx*, year
- [13] Novak, M., Joy, M., and Kermek, D. "Source-code similarity detection and detection tools used in academia: a systematic review." *ACM Transactions on Computing Education*, 19(3), pp. 1-37, 2019