Correlation Analysis of Developer Interaction on Stack Overflow Website towards Perceived Programming Skills among Information Technology Students in Jose Maria College

Jerwin Carreon[1], Japhet Floren[2], Jory Benedicto[3], Erica Celine Fernandez[4], Vishy Te[5], Kate Saramosing[6], Riza Mae Sinangote[5]

[1,2,3,4,5]*Student, Bachelor of Science in Information Technology Jose Maria College, Philippine*
[6]*Faculty, Jose Maria College*

Abstract- **There is massive information and addresses that can be found through the internet. Internet users visit websites like forums and blogs worldwide to seek information, communicate especially to get ideas. Stack Overflow is one of the popular sites ever visited by communities. The main objective of this site is to enable users to ask questions, collaborate and seek help commonly to a certain programming context or errors. Furthermore, programmers find it an effective and interesting collaborating site. Considering this site has been for many years on the internet. May it be a beginner trying to figure out programming, average or professionals with different logics and ways of contributions find this site very influential. This alone has shaped individuals' programming skills both negatively and positively. This study aims to produce a statistical record of correlation between developer interactions of Stack Overflow within Jose Maria College IT students and perceived programming skills. Jose Maria College is an institution that runs an IT program which means programming as one of their core subjects, students are no excuse and exposed to encountering programming issues. Often, the students are susceptible to visiting Stack Overflow to find answers for their assignments and projects. Nevertheless, this study will show the impact of the site on the programming skills of the students. Commonly visualization figures and text inputs indicating the instruments and findings.**

Keywords- **Correlational Research, Stack Overflow, Developer Interaction, Programming Skills, Jose Maria College**

I. INTRODUCTION

One of the challenges of a programmer or a student learning programming languages is finding sources through the internet. Seeking for guidance and solutions to their program errors by visiting discussion forums. Where they can ask questions and assistance to professionals. Specifically Stack Overflow, a discussion forum for programmers, is considered one of the largest contributors where they share different techniques, knowledge, and languages. By then researchers believe that the perceived skills achieved by visiting the website are correlated and inspired from the forum's sources.

Discussion's forums were conceptualized as early as 1980. The main objective was to create a platform for asynchronous communication among users. They are seen as online collaborative spaces where a diverse set of people converge and share knowledge. Due to this huge repository of knowledge contributed by people from all over the world, online forums can be effective sources of learning. They were also seen as enablers of blended and constructivist learning.

As described in a research article by International Conference on Information and Communication Technologies (ICICT, 2014). Stack Overflow is a Q&A forum for software professionals, students and programmers. Users can register for free and post their questions related to programming concepts such as programming languages, databases, XML, etc. Each question can be accompanied by a set of tags used to categorize the question with other similar questions. Tags are also used by a user to search for questions asked by the other users. Users can get reputation scores based on their contribution and interaction on the forum. Users and questions also have badges attached to them to indicate their quality.

Regarding teaching a computer programming language, many potential advantages of OL delivery have still to be realized. These include using cut-and-paste techniques to reuse existing code and the ability to compile and execute code examples while studying lecture material.

According to Tanveer Ahmed and Abshishek Srivastava's research, they particularly examine the example of Stack Overflow (SO). Though speculative, it is often regarded as one of the finest internet forums for technical debates. This assertion is supported by the literature's increasing interest in examining the platform from various angles. There is research devoted to badges, recognizing expert conduct, detecting emotional signals in developers, evaluating and suggesting tags, gamifying participation, and forecasting the status of inquiries, among other things. Some papers highlights programming skilss in design and development of information systems such in [9] and [10].

According to a study paper titled "From Asking to Answering: Increasing Your Involvement on Stack Overflow." Stack Overflow offers a platform for expert users to address queries regarding software libraries, alleviating substantial burdens on their owners and creators. Additionally, it serves as a supplement to API documentation. There are also plugins for integrated development environments (IDEs) that enable developers to interact with Stack Overflow as they code.

While not all users come to platforms like Stack Overflow to become contributors, many studies indicate that a significant proportion of users are hesitant to engage due to a perceived lack of competence. These emotions are likely to be particularly strong while transitioning from asking to answering questions. Furtado et al.'s empirical research corroborates this intuition: they discover that users who submit many queries in a short period are more likely to abandon the site in the future than they are to provide replies.

The researchers want to test and obtain knowledge towards ways of students as a developer interacting to the features of Stack Overflow. Getting information towards the level of satisfaction a student can get and reactions about the collaboration of others. Nevertheless to observe the perceived programming skills for being into Stack Overflow based on their experience and impressions. This involves within Jose Maria College IT students alone who are recognized as visitors and users of the said website.

Theoretical framework

This article is based on Linda Harasim's idea of Online Collaborative Learning (OCL), which advocates for a flexible learning environment facilitated by online collaboration, sharing, and communication to reshape formal, non-formal, and informal education for the knowledge era. This method encourages students to be active and involved in gaining information about their community. As a result, it supplants certain conventional real exercises performed daily by teachers. The primary goal is not to replace instructors, but to help them improve communication between the two sides.

Additionally, Dekeyser (2007) asserts that Skill Acquisition Theory "is that the acquisition of a wide variety of skills exhibits a remarkable similarity in development from initial representation of knowledge to eventual fluent, spontaneous, largely effortless, and highly skilled behavior, and that this set of phenomena can be explained by a set of fundamental principles common to all skills acquisition." In summary, as Speelman (2005) points out, skill acquisition may be seen of as a subset of learning, with learning defined as "the storage of information in memory about some environmental or cognitive experience." Thus, he defines skill development as a kind of learning in which "skilled actions may become routineized and even automated under certain circumstances."
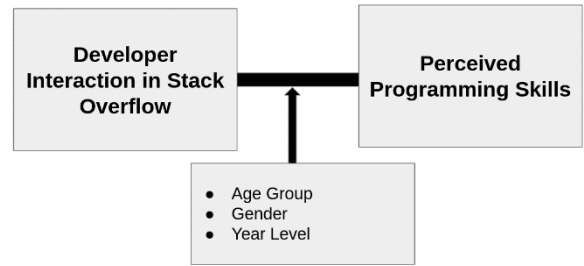
Conceptual Framework

Figure 1. Conceptual Framework of the study

Figure 1 shows Developer Interaction on Stack Overflow as the independent variable and its relationship towards perceived Programming skills to students as the dependent variable. Defining the approaches and connection involved to the study.

II. METHODOLOGY

Research Design

The design used in this study was employed research design, which aims to investigate the Correlation Analysis of Developer Interaction on Stack Overflow Website towards Perceived Programming Skills Among Information Technology Students in Jose Maria College in the school year 2020-2021.The researchers developed a quantitative questionnaire to use as the survey instrument. In this study, the researchers use correlational research as the design because it investigates the correlation between programmers who use stack overflow often and those who use it less often or not at all. Answering the research question hypothesis is the central purpose of the study. According to Fraenkel and Wallen (2009), correlational research aims to find out the relationship between two or more variables and their cause and effect [6]. Research design is the conceptual structure within which research was conducted. It is the plan of how a researcher arranges his/her research to find the answer to the question or the statement problem.

Research Locale and Sampling

The study was conducted at Jose Maria College, Davao City, Davao del Sur. In this season of the pandemic, the researchers use an online survey questionnaire. The respondents are the official college students currently enrolled in BSIT. The study was in the semester of the academic year 2020-2021.

Population refers to all the members of a particular group. It is the group of interest to which the researchers would like to generalize the result of a study. The study was conducted in the school academic year 2020-2021. The researchers decided that the respondents of this study are the BSIT Students. It's about the Correlation Analysis of Developer Interaction on Stack Overflow Website towards Perceived Programming Skills Among Information Technology Students in Jose Maria College in the school year 2020-2021 To collect the appropriate data needed for the research, the researchers will

collect information by giving survey questionnaires on the Students of BSIT at Jose Maria College from 1st year to 4th year of school year 2020-2021 via Online Survey Questionnaire.

In the researcher's sampling they used Slovin's Formula which is written as; n= N / (1+Ne2) where: n= the number of samples

N=the total population. There are 75% recognized of the population of Bachelor of Science Technology Students in Jose Maria College per year level and assigned a confidence level of 90% with a 25% margin of error, the group's sample size is as follows:

Out of the 13 1st year students there are 7 responses, out of the 35 2nd year students the group recognized 9 responses,

From the 2 4th year students the group recognized 2 responses

Simple Random sampling is the most basic form of sampling. Every member of the population has an equal treatment. This paper employed simple random sampling in deploying online survey questionnaires. Random sampling was used in the selection of the respondents.

The respondents were given ample time to answer the survey sheets, after which the researchers retrieved the instrument for safekeeping. The data that was gathered in the deployment of the instruments were statistically treated and were interpreted.

Data Collection Procedure

The researchers of this study formulated an online researcher-made questionnaire based on the indicators of the study's variables. Then, their research adviser validated the online questionnaire before the distribution through a Google survey form. After validating the online questionnaires, the letters were sent to the following respondents, which is asking permission for them to participate in an online survey. The researchers administered the questionnaire to the respondents to ensure that the questionnaires served their purpose. At the same time, the researcher discussed and explained the purpose of the study to the respondents. Right after the respondents answered the questionnaires, the researchers retrieved them immediately. The data gathered through the online survey were tallied, tabulated, analyzed, and interpreted by the researchers to answer the questions as posted in the problem statement or research question.

Research Instrument

The group used Google forms Online Survey In their Questionnaire to gather information from Bachelor of Science in Information Technology Students in Jose Maria College who are the group's research participants and will be the ones to take their research questionnaire consisting of the questions that the group have gathered from this adopted questionnaire [questionnaire source] with the hopes that all participants will answer accordingly. The researchers added a positive approach in their questions since the questions from the adopted questionnaire had a negative approach.

Ethical Considerations

The researchers will ensure that each of the respondents would be notified through online chat via messenger. The notification sent to the respondents will ask the participant's approval to participate in the said online survey. As carriers of good conduct, the researchers make sure that they would keep their promise to the participants and that their personal information is confidential and that the group will never share it with anyone for their protection. The researchers ensure that all willing participants are not harmed and that all participants have been given complete information about what is needed and what, if any, possible negative effects may result from their participation. Respect for the dignity of research participants would be the group's top priority and the protection of research participants' privacy.

III. RESULTS AND DISCUSSIONS

This study intends to find out the degree of the relationship of developer interaction in Stack Overflow website and perceived programming skills among information technology students in Jose Maria College.

Results

Based on the findings of this study it has revealed that there is no significant difference on the level of developer interaction in Stack Overflow Website when analysed according age group and year level. Meanwhile, p-value 0.036 for Gender < 0.05, then we reject the null hypothesis. There is a significant difference on the level of developer interaction in Stack Overflow Website when analysed according Gender.

Significant Difference on the level of developer interaction in Stack Overflow website When Analysed According to Age Group, Gender, and Year Level

Test Variables (ANOVA)	Level	Mean	Std. Deviation	F	Sig.	Decision
Age Group	18 - 25 years old 26 - 30 years old	3.95 3.80	.616 .566	.114	.737	Accept Ho
Gender	Male Female	4.04 3.60	.570 .644	4.641	.036	Reject Ho
Year Level	1st Year 2nd Year 3rd Year 4th Year	3.55 3.98 3.96 4.80	.619 .593 .403 .283	2.762	.053	Accept Ho

In addition, since, *p*-values for Age Group and Gender are .289 and .060 > 0.05, respectively, then we do not reject the null hypothesis. There is no significant difference in the perceived programming skills level in Stack Overflow Website when analysed according to Age Group and Gender. Meanwhile, p-value 0.030 for Year Level < 0.05, then we reject the null hypothesis. There is a significant difference on the level of perceived programming skills in Stack Overflow Website when analysed according Year Level. Constructed on the Post-hoc Test, on Multiple Comparisons' Table, it

revealed that there is a significant mean different between First Year and Second Year, and between First Year and Fourth Year.

Test Variables (ANOVA)	Level	Mean	Std. Deviation	F	Sig.	Decision
Age Group	18 - 25 years old 26 - 30 years old	3.79 3.10	.899 .000	1.151	.289	Accept Ho
Gender	Male Female	3.89 3.32	.923 .619	3.721	.060	Accept Ho
Year Level	1st Year 2nd Year 3rd Year 4th Year	3.06 3.84 3.88 4.90	.862 .882 .295 .141	3.263	.030	Reject Ho

Lastly, the table below shows the positive correlation between Developer Interaction in Stack Overflow Website and Perceived Programming Skills. Since, p-value is .000<0.05, then we reject the null hypothesis. There is a significant relationship between Developer Interaction in Stack Overflow Website and Perceived Programming Skills. This positive correlation implies that when the eloper Interaction in Stack Overflow Website increases, it can be assumed that Perceived Programming Skills also increases. On the strength of relationship between variables, with the value of r=.807, it has a Strong Positive Relationship.

Correlations Relationship between Developer Interaction in Stack Overflow Website and Perceived Programming Skills

Range of r-value	Description Interpretation
0.00 to 0.20	Negligible Positive Relationship
0.21 to 0.40	Weak Positive Relationship
0.41 to 0.60	Moderate Positive Relationship
0.61 to 0.80	Strong Positive Relationship
0.81 to 0.99	Very High/Very Strong Positive Relationship
1.00	Perfect Relationship

IV. CONCLUSION

In light of the data gathered based on these findings show that Developer Interaction on the Stack Overflow Website and Perceived Programming Skills have a substantial link. With this positive link, it can be argued that as eloper Interaction on the Stack Overflow Website increases, so does Perceived Programming Skills. It has a Strong Positive Relationship with a value of r=.807 on the strength of the relationship between variables. Therefore we conclude that there is statistically significant relationship between the Developer Interaction on Stack Overflow Website towards Perceived Programming Skills among Information Technology Students in Jose Maria College.

The result was also supported in the article entitled,"From Asking to Answering: Getting More Involved on Stack Overflow", Individuals use Stack Overflow to learn and developers seek solutions to problems they face at work. There are even integrated development environment (IDE) plugins that allow developers to interface with Stack Overflow while coding. Stack Overflow provides a forum in which power users can answer questions about software libraries, relieving their owners and developers of a significant load. It also plays a complementary role to API documentation.

REFERENCES

[1] Bachschi T., Hannak A., Lemmerich F., and Wachs J., "From Asking to Answering: Getting More Involved on Stack Overflow", arXiv e-prints, 2020.*: Getting More Involved on Stack Overflow*

[2] Amalarethinam, George & Lucia, T., "*ScienceDirect International Conference on Information and Communication Technologies (ICICT 2014) Customer Facilitated Cost-based Scheduling (CFCSC) in Cloud*".2015. 660-667.

[3] R. Dagohoy, M.V. Buladaco, J. Jacinto, J. Pintado, & L. Ibañez, "Social Media Marketing Towards Consumer Buying Behavior: A Case in Panabo City",5, 22-30, 10.47772/IJRISS, 2021, 5202.

[4] R. Dekeyser, "Skill acquisition theory", In B. VanPatten & J. Williams (Eds.), Theories in second language acquisition: An introduction (pp. 97-113). New Jersey: Lawrence Erlbaum Associates, Inc.,2007.

[5] C. P. Speelman, & K. Kirsner, "Beyond the learning curve: The construction of mind. Oxford University Press", 2005.

[6] K. Buckreus, "Connectivism: Informing Distance Education Theory", Pedagogy and Research, Morinville, Alberta, Canada, 2020.

[7] L. Harasim, *Learning Theory and Online Technologies,* RoutledgeFalmer, Taylor & Francis Group, 2012, New York.

[8] T. Ahmed, & A. Srivastava, "Understanding and evaluating the behavior of technical users. A study of developer interaction at StackOverflow", Human-centric Computing and Information Sciences, 7 (1), doi: 10. 1186/s13673-017-0091-8, 2017.

[9] C. Magallanes, M. Ortiz, M.N. Seville, E. Tuliao, S. Tejada, N.G. Eroy & M.V. Buladaco, "Analysis And Design Of A Sales And Inventory Management Information System For A Motorcycle Parts And Accessories Store", International Journal of Scientific Research and Engineering Development-– Volume 4 Issue 3, May -June 2021.

[10] [M.V. Buladaco & F.L. Ubay, "GETMOSYS: A Plant Pathologists Geotagging And Monitoring System For Infected Banana Plant", International Journal Of Scientific & Technology Research Volume 9, Issue 04, April 2020.