

Similarity Index Application to Students' Projects for Quality Control

Jumoke Soyemi^{1*} & Olugbenga Babajide Soyemi²

¹Department of Computer Science, The Federal Polytechnic, Ilaro, Nigeria.

²Department of Civil Engineering, The Federal Polytechnic, Ilaro, Nigeria

*Corresponding Author: jumoke.soyemi@federalpolyilaro.edu.ng

Abstract

Plagiarism is a global challenge presently facing most academic institutions. The wealth of resources on the internet through Information Technology provided unhindered access to those resources. The present trend among students is to download resources from the web and submit the same verbatim. The peak of this is downloading or copying someone else's project/dissertation and presenting the same as the original owner. This paper proposes the introduction of Plagiarism software in all academic institutions in Nigeria for quality control and promotion of originality among students. A well-structured questionnaire was developed and administered in The Federal Polytechnic, Ilaro among 318 participants randomly selected from Higher National Diploma (HND) II students across the five (5) schools using stratified proportional allocation technique and 212 academic staff making a total of five hundred and thirty (530) participants to gather information on the perception of introducing plagiarism software in tertiary institutions for quality assurance purposes. Also, statistical analysis was carried out on randomly selected 50 samples of past HND II projects within the Department of Computer Science. Similarity test index was conducted on the selected and digitised projects. With a similarity index benchmark of $\leq 21\%$, the results of the analysis shows 2% project with less than 21% plagiarism, 13% project with 22-30% plagiarism, 10% with 31-40% plagiarism, 15% with 41-50% and 60% of the students' project with over 50% plagiarism. The analyses from this study confirmed the level of academic fraud being practised by the students, which should be a significant concern across Academic institutions not only in Nigeria but across the globe.

Keywords: Information Technology, Internet, Plagiarism, Plagiarism software, Quality assurance

Citation

Soyemi, J. & Soyemi, O.B. (2020). Similarity Index Application to Students' Projects for Quality Control, *Journal of Women in Technical Education and Employment (JOWITED)*, The Federal Polytechnic, Ilaro Chapter, 1(1), 28-38

ARTICLE HISTORY

Received: May 7, 2020

Revised: May 29, 2020

Accepted: July 10, 2020

1. Introduction

Plagiarism is the act of presenting someone else works as ones without giving due credit to the owner of the work. Plagiarism in academics is considered as a severe offence of dishonesty and misconduct (McEvedy & Smith, 1990; Stein & Prettenhofer, 2011). This act is regarded as an academic scam or theft of intellectual property. The advent of the internet contributed tremendously to the act of plagiarism. Esra & Ahmet (2014), explored how the internet has contributed to the ever-increasing problems of plagiarism in higher education. The authors opined that would-be teachers inclined to plagiarise by copying material from the internet. Presently, the scholarly databases and the internet are quite rich and accessible to professionals (Richard, Saunders & Meek, 2016) as a means of presenting their work to the larger society for other to learn and probably take the work further. This means of communicating research work and additional contribution has its numerous advantages, even though, the abuse in the form of plagiarism needs to be seriously tackled at the academic communities where students among others are the principal culprit.

Plagiarism in academic institutions brings about loss of integrity, ethics, honesty, and downgrading institution of learning. Plagiarism has become a reoccurring challenge facing virtually all institutions of higher education globally. Resources are readily available online, which poses significant advantages and challenges in academia. Materials are carelessly copied without due reference to the original owner. This is common among both teachers and students. Assignments, term papers and projects/dissertations are downloaded and submitted without proper referencing. The inability of both teachers and students these days to do extra work has led to a sharp rise in stealing materials from all sources, particularly from the internet. They look for the easiest way out of every situation, especially with the awareness that the internet is available for any resource required for their work. This should have been to their advantage if the resources downloaded are sieved out, put to proper use, and adequately referenced.

There are various types of plagiarism. Word-for-word plagiarising is a type of plagiarism where someone else's work or write up is presented as one's idea. Plagiarising by paraphrasing is a type of plagiarism in which someone's idea is stated in one's own words without referencing the owner of the idea. A different kind of Plagiarism is Mosaic Plagiarism, here phrases and terms are taken directly from the source and then combined with one's prose. The lifted words are usually written in italics (Cavano, 2011). To curb the act of plagiarism, some software/checker tools have been developed to detect and look at similarity index between one's work and the wealth of resources available online. The software uses the internet to search through databases. Virtually, all plagiarism software works the same way because the act of detecting plagiarism is more of similarity detection or text matching that seeks similar words between two or more documents. The difference is in how each one of them carries out the check and the mode of presentation of the result. Examples of these similarity software checker are Turnitin, Viper, Grammarly, Plagscan, Whitesmoke, Article checker, Duplichecker, Plagiarismcheck, SmallSEOTools, Plagium, Plagiarism checker, Paperrrter, Dustball, Plagiarism checker plugin for WordPress, CheckForPlagiarism.net, Copyscape, Search engine reports, Plagiarisma.net, Plagtracker, DMCA scan and webconf (DigitalGYD, 2018).

The bedrock and the power of any nation lie in the quality of research conducted in their academic institutions. Researches and discoveries are meant to emerge from academia, and such findings are what serve as input to the industries and policy formulations for implementation. Consequently, if the backbone of scientific and technological discoveries is decaying, then such a nation is heading for socio-economic disaster. In the work of Carroll & Zetterling (2009), it was discovered that plagiarism had become an international issue, and many academics realise the big challenge with digitised resources on the internet, global communication and search engines on the increase. Also, in the study by Heyman (2000), 38% plagiarism was found in the works of students at Rutgers University, which supported the anxiety. Regrettably, this act is not limited to students alone; it is also becoming rampant among researchers, teachers and academics.

A project is an individual or joint study that is planned carefully and researched on by students (Thomas, 2009). It consists of solid and organised efforts motivated by a perceived opportunity when facing a problem, a need or a desire (Turner & Müller, 2003). The completion and submission of a project is a requirement for the award of a college degree and graduation in all tertiary institution of learning. At the master's level, a dissertation is required, and at the PhD level, a thesis is a requirement. At these various levels, submission of a write up is a requirement. Unfortunately, it has been observed by Debnath *et al.* (2012), that students are getting very lazy in carrying out their research and most of the submitted research works are unoriginal to qualify them for the degree so desired. (Soyemi & Isinkaye, 2017). The trend is becoming more and more every day as a result of the wealth of resources available online. Therefore, having students presenting unoriginal research work is bad enough, and the inability to find an immediate solution would impact negatively on the nation. Cryer (1997), stated that developing originality can be achieved through learning and the use of lateral thinking and creative skills to bring about creativity and enabling a period of

incubation of such skills to develop well. Therefore, originality can be achieved in research through time and effort (Blaxter, Hughes & Tight, 1998).

According to Glendinning (2014), the procedures for detecting plagiarism and academic fraud used in the E.U. were based on the reports of the 27 countries concentrating specifically on works at the bachelor and master's levels. Responses indicated that there are some forms of policies at the institutional or departmental level for academic misconduct and plagiarism in most of the institutions involved in the survey with some declining saying no effective system in place. Most participants favour safeguarding of institutional autonomy against the idea of having central commands to the higher education information (HEI) for responding to plagiarism apart from those from France, Germany, Bulgaria, and Romania. They advocated for a set of national guidelines to respond to plagiarism. The findings established that there are poorly clear policies and systems to guarantee Academic integrity in many parts of Europe.

Ehrich, Howard, Mu & Bokosmaty (2016), had in a previous study reported plagiarism as a growing problem within Australian tertiary institutions and overseas. The authors revealed students' disposition towards plagiarism as naturally non-judgmental as the policies of their institutions. Their work compared responses from Australian and Chinese institutions wherein students' were shown to have strong mentalities against plagiarism with a minority, however, displaying concerning discernments and understandings of plagiarism. The results from the administered questionnaire discovered that 90% of students in Chinese and Australian universities showed absence of consciousness to self-plagiarism as a form of any academic wrongdoing

Bruton & Childers (2016) interviewed twenty-three academics, from different departments in a public university in the southeastern United States to determine their viewpoints on student plagiarism and Turnitin. The findings showed that there are conflicting reports as to what led to student plagiarism from various faculties which include disagreements about penalties, failure to grasp basic norms of academic writing, laziness, teachers who do not adequately explain the purpose of proper referencing and many more.

From the foregoing, it can be understood that undergraduate students in Europe, the United States and other developed nations lack the understanding of plagiarism and plagiarist acts, and so commit the act freely and ignorantly

This study determines the extent of plagiarism among students of The Federal Polytechnic Ilaro, which is the case study and proposes the introduction of similarity index software in all Nigeria tertiary institutions for originality, innovations, quality assurance and escapes the torture of plagiarism

2. Methodology

2.1 Data Collection Method

Survey method and statistical analysis of one of the past sets of HND II students' projects from the Department of Computer Science were adopted in carrying out this study. The past HND II students' project were digitised and checked against plagiarism. The method of data collection was through a structured questionnaire containing awareness and perception of participants on the introduction of plagiarism software in academic institutions for quality assurance purposes.

2.2 Sample Size and Sampling Technique

This study considered the population of both HND II students, which cuts across the five (5) Schools in Federal Polytechnic, Ilaro and her academic staff in the academic session. The total number of registered HND II students was One thousand, five hundred and forty-one (1,541), while that of academic staff was four hundred and fifty (450).

Thereafter, Yamene (1967) sample size calculator was then employed to determine the appropriate sample size of three hundred and eighteen (318) scientifically from among the students' populace and Two hundred and twelve (212) from the population of academic staff. The student's sample was selected from each schools' population using Proportional Stratification sampling technique as calculated below.

Students' population size: $N = N_1 + N_2 + N_3 + N_4 + N_5 = 252 + 621 + 392 + 204 + 72 = 1541$

Where $N_1, N_2, \dots, N_5$ represents the school of: Engineering, Management, Applied Science, Environmental and Communication.

However, the aggregate sample size selected from the five schools was calculated.

$$n = \frac{N}{1 + Ne^2} = \frac{1541}{1 + 1541(0.05)^2} \cong 318$$

Where n = sample size to be selected; N = population size and e = random error term (± 0.05)

Using the stratified proportional allocation technique per stratum, we have

$$n_i = \frac{nN_i}{N} \quad i = 1, 2, \dots, 5$$

School of Engineering: $n_1 = \frac{nN_1}{N} = \frac{318(252)}{1541} = 52$; School of Management : $n_2 = \frac{nN_2}{N} = \frac{318(621)}{1541} = 128$

School of Applied Science: $n_3 = \frac{nN_3}{N} = \frac{318(392)}{1541} = 81$; School of Environmental : $n_4 = \frac{nN_4}{N} = \frac{318(204)}{1541} = 42$

School of Communication: $n_5 = \frac{nN_5}{N} = \frac{318(72)}{1541} = 15$

The required sample size for the academic staff is also computed as

$$n = \frac{N}{1 + Ne^2} = \frac{450}{1 + 450(0.05)^2} \cong 212$$

Thus, the total sample size from which a designed questionnaire was examined for this research was five hundred and thirty (530) individuals.

2.3 Method of Data Analysis

Descriptive Statistics method of data analyses was applied to scale statements and examine the order of importance through pictorial representation.

3. Results and Discussion

Graphical representation of analysis was employed to present the distribution and opinion of the sample on the issues raised in the administered questionnaire.

Figure 1 depicts the gender distribution of the two categorised participants with about 56%, 49.6% of them being male academic staff and students while 44%, 50.4% were female academic staff and students respectively. This indicates that the majority of the participants were male.

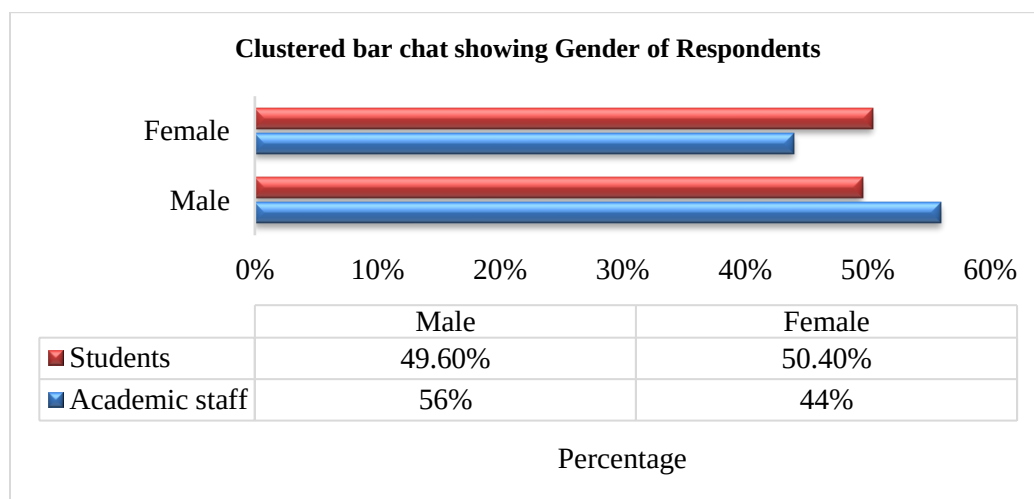


Figure 1: Distribution of the two participants by sex

It is evidenced in figure 2 that all the academic staff of the institution are aware of plagiarism, but few (20.8%) of the students said they are not aware at all while 5% of them were not even sure whether plagiarism exist or not.

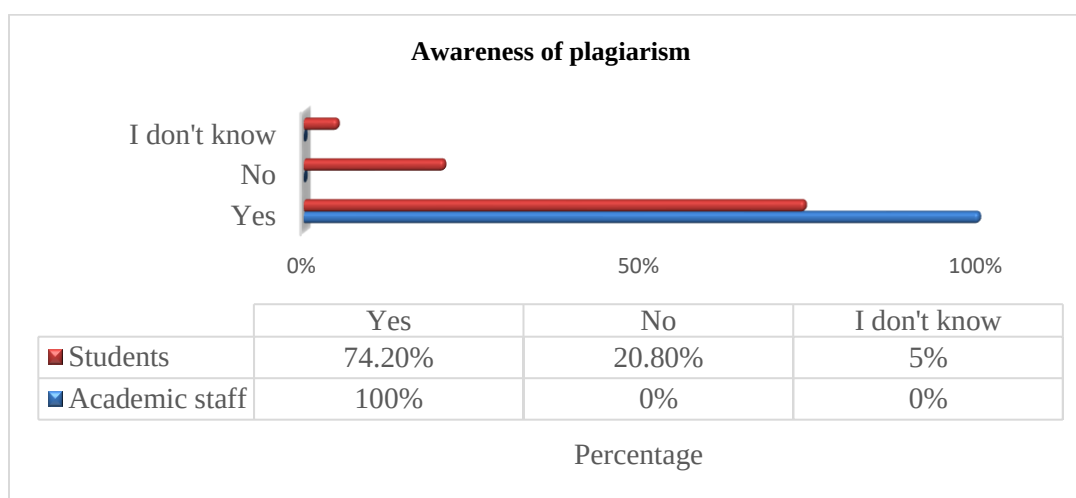


Figure 2: Distribution of participants on awareness of plagiarism

Figure 3 showed that all the randomly selected academic staff download resources from the web, 82% of the students do the same, while 18% do not. This implies that information technology has been a viable tool of use in academic institutions in the world. Since the majority of the respondents are aware of the item analysed, then the analysed item exert positive response.

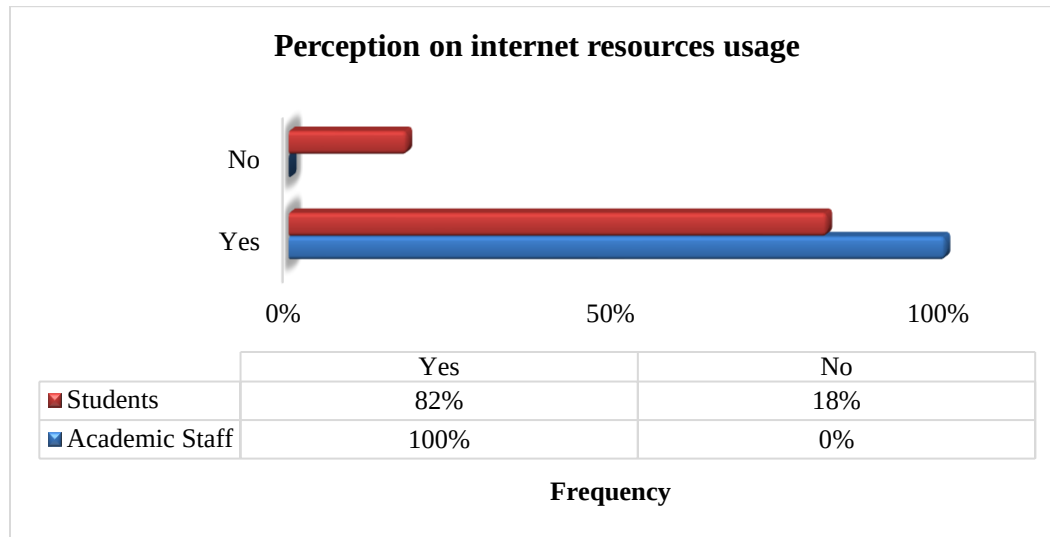


Figure 3: Distribution of participants with respect to internet usage

Figure 4 showed that majority of the randomly selected participants (both academic staff and students) downloaded publications from the internet, indicating 50% and 55%, respectively compared to project write-ups downloads (20% and 31.7%) and resource for assignments (30% and 13.3%) downloads.

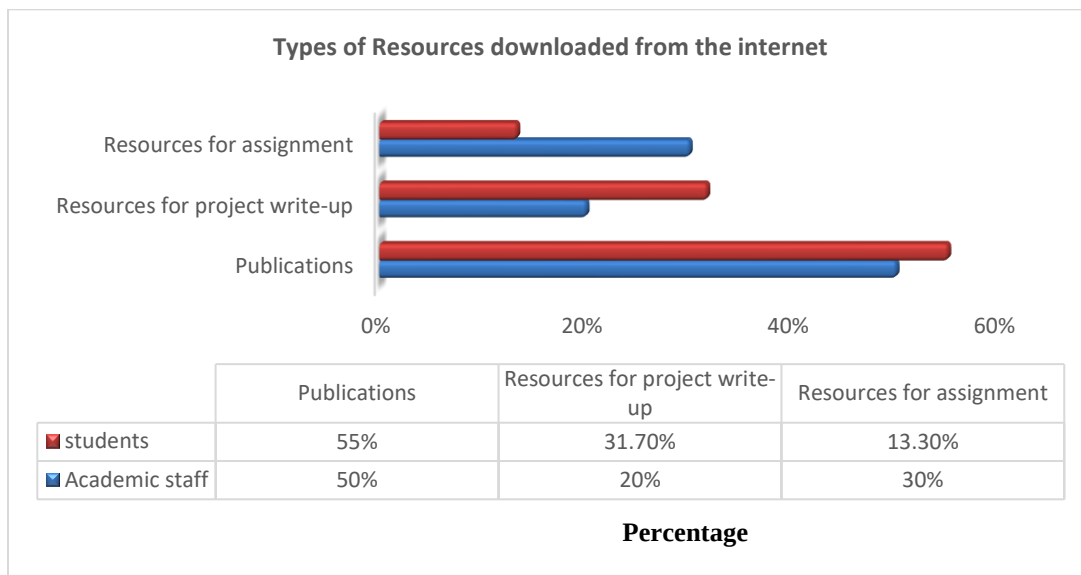


Figure 4: Distribution of participants on types of resources downloaded from the internet

Based on what the participants do with their downloaded resources, results here indicates that majority of the academic staff (94%) read and write based on their understanding and cite materials, 37.9% of the students do same, 28.3% and 33.7% of the students “copy and paste materials for use”, and “sieve materials and use” to enhance the quality of their project write-ups as depicted in Figure 5, respectively.

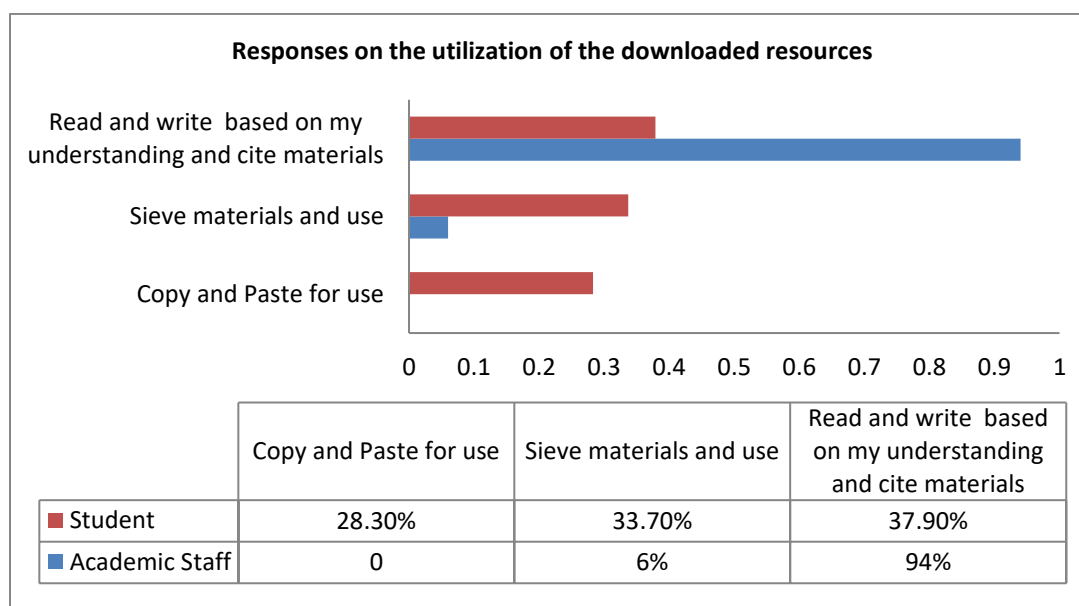


Figure 5: Distribution of participants on how they use the downloaded resources

In Figure 6, the participants (both academic staff and students) were asked if they are aware of anyone who might have been indicted on plagiarism. The results showed 62% of the academic staff who knew more than two people that plagiarised and indicted 13.8% knew no persons that plagiarised.

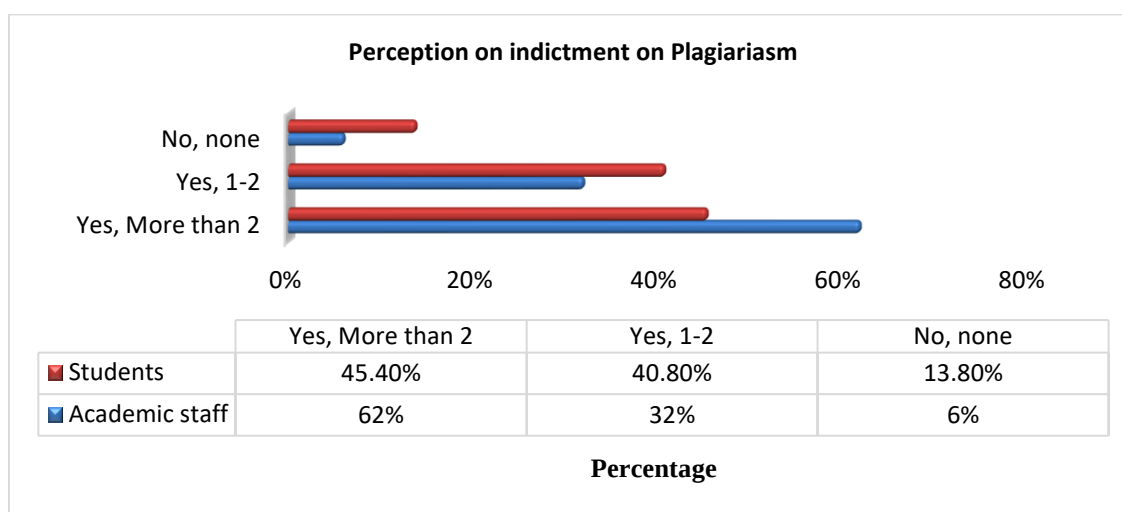


Figure 6: Distribution on the perception of participants towards indictment on plagiarism

On whether to cite resources in any write-up, 63% of the students said “Yes” 10% said “No” while 23% of them said they don’t know. But for the academic staff, all the participants (100%) said it is necessary to do so, as shown in Figure 7.

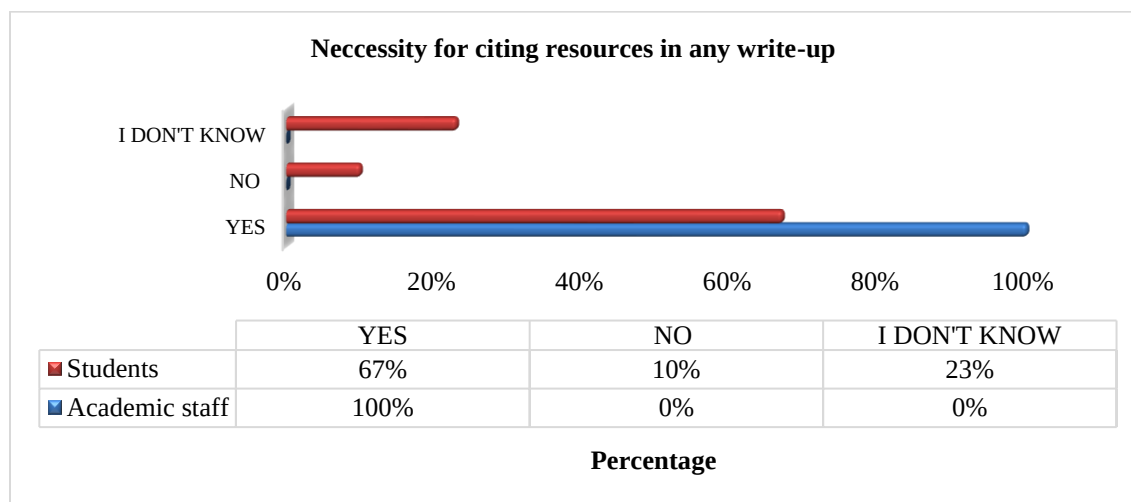


Figure 7: Distribution on the perception of participants towards necessity in citing resources in any write-up

Analysis of Figure 8 demonstrated that majority of the respondents have a positive opinion on the introduction of plagiarism software in the institution as 98% and 80% of the academic staff and students said “Yes” while 20% of the students have a negative perception on its introduction. However, 2% of the academic staff were undecided on the introduction of plagiarism software in the institution.

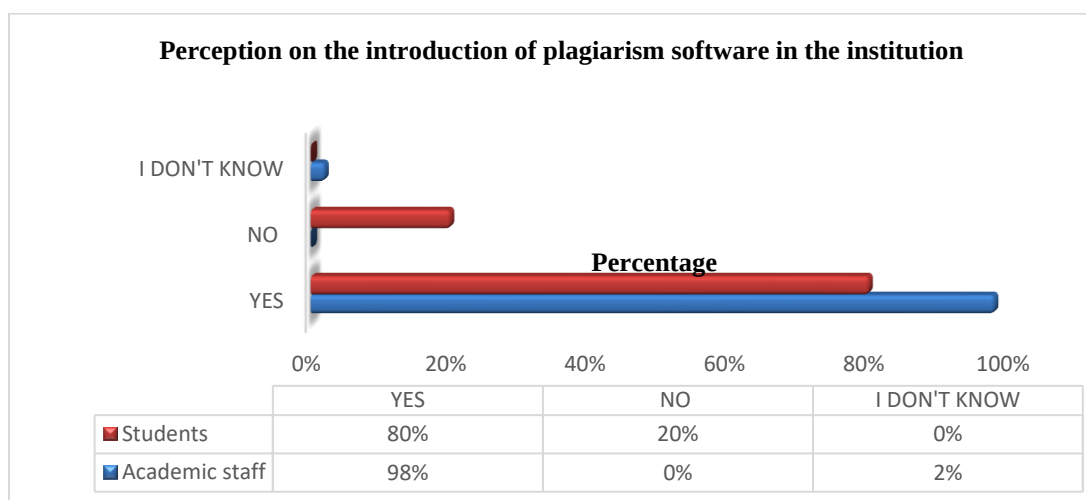


Figure 8: Distribution on the perception of participants towards necessity in citing resources

In Figure 9, the majority of the participants, both academic staff and students opined that the quality and originality of work could be attained by introducing Plagiarism software in tertiary institutions such as Federal Polytechnic, Ilaro.

Only 7% of the students have a negative perception about it, while 7% and 11% of the academic staff and students neither have a positive or negative perception of quality and originality attainment.

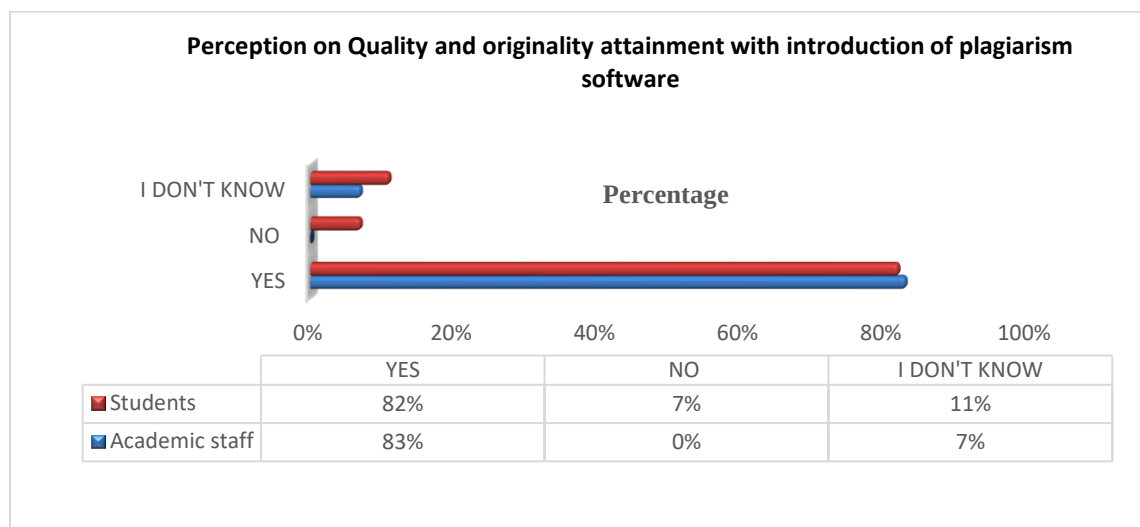


Figure 9: Distribution of participants' perception of Quality and Originality attainment

In Figure 10 on whether the concept should be introduced, evidence shows that all the academic staff (100%) are ready to take advantage of it as it will reveal the originality of their supervised students' work and create goodwill for the institution locally and internationally in the aspect of research. Meanwhile, 35.8% of the students never had a positive nor negative opinion on the introduction of the respective software.

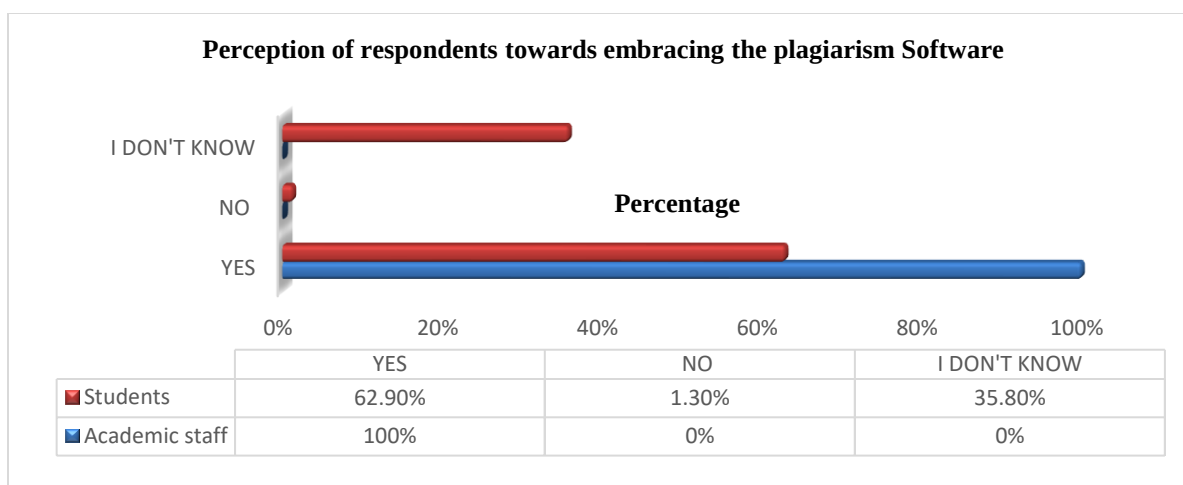


Figure 10: Distribution of participants' perception of Quality and Originality attainment as a result of plagiarism software introduction.

Figure 11 depicts the plagiarism test carried out on a past set of HND II students in the Department of Computer Science. With a similarity index benchmark of $\leq 21\%$, the results of the analysis shows 2% project with less than 21% plagiarism, 13% project with 22-30% plagiarism, 10% with 31-40% plagiarism, 15% with 41-50% and 60% of the students' project with over 50% plagiarism.

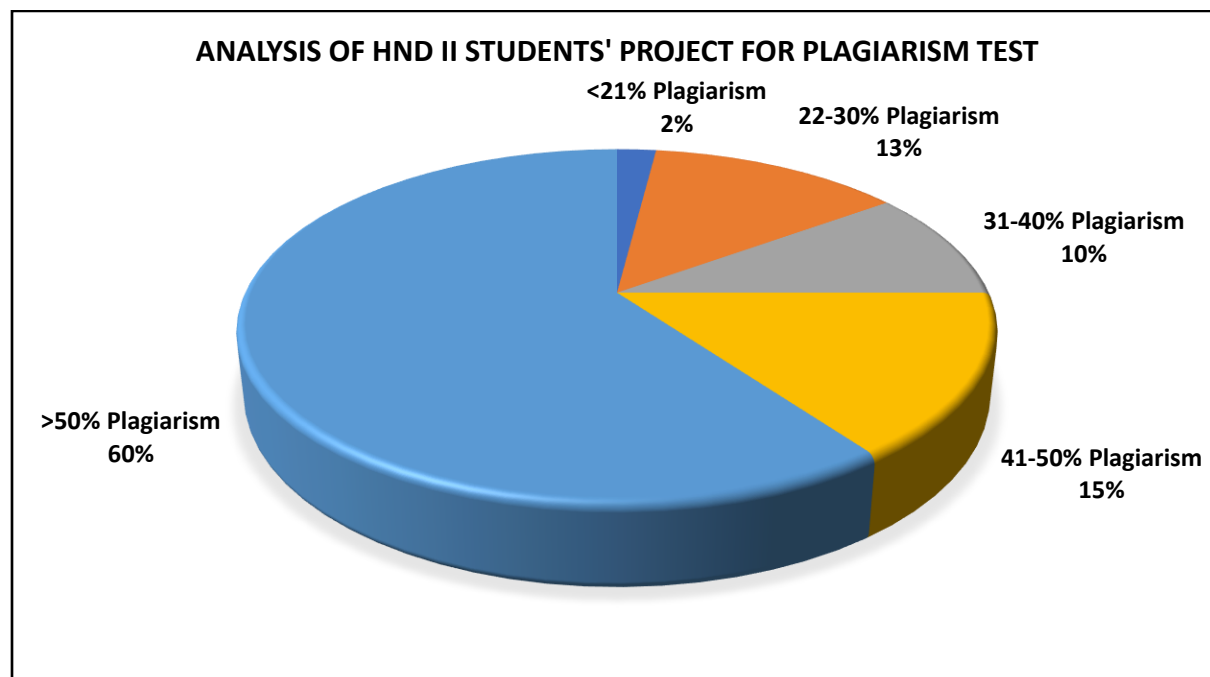


Figure 11: Analysis of Digitized Projects

4. Conclusions

The foregoing indicates that information technology via the internet is a viable tool used in academic institutions across the world because of the wealth of resources available, most of which are free of charges. Many academics, teachers and students alike take advantage of these resources to advance in academics. The students, however, have turned this benefit to a negative one in the level of copy and paste work carried out by them. The results of this study showed the level of plagiarism decadence among the students with quite a number being unaware of the crime. The results from this study confirmed the level of academic fraud being practised by the students, which should be a significant concern in the institutions, and a pointer to malpractices in the academic community. However, the majority of teachers and students are willing to embrace the introduction of the plagiarism-checking software. If introduced, will benefit both the students and their tutor in their academic work. Aside from improving the research image of the institution through quality research publications output, the academics becomes more confident and have the sense of originality. Promoting quality, originality, innovation and academic excellence can only be achieved through creative thinking and hence the proposal to introduce Plagiarism software checker in all educational institutions.

References

Blaxter, L., Hughes, C. & Tight, M. (1998). *How to research*. USA: Open University Press

-
- Bruton S. & Childers D. (2016). The ethics and politics of policing plagiarism: a qualitative study of faculty views on student plagiarism and Turnitin®, *Assessment & Evaluation in Higher Education*, 41(2), 316-330.
- Cavano, B. (2011). Plagiarism in higher education. Retrieved 12 March, 2018 from: <http://www.examiner.com/adult-deucation-in-pittsburgh/plagiarism-higher-education>
- Carroll J. & Zetterling C. (2009). *Guiding students away from plagiarism*, First edition, KTH Learning Lab and the authors.
- Cryer, P. (1997). *The research student's guide to success*. USA: Open University Press
- Debnath, D., Barthakur, N.K., Baruah, R.S. & Goswami, P.K. (2012). Problems of teaching optics in middle school: A survey in Batadraba education block of Nagaon district Assam, India. *International Journal of Scientific and Research publications*, 2(8), 1-4.
- DigitalGYD (2018). Top 20 best free online Plagiarism checker tools and websites (2018). Retrieved 12 March, 2018 from: <https://www.digitalgyd.com/top-20-best-online-plagiarism-checker-tools-free/>
- Ehrich J., Howard S.J, Mu C. & Bokosmaty S. (2016). A comparison of Chinese and Australian university students' attitudes towards plagiarism, *Studies in Higher Education*, 41(2), 231-246.
- Esra Eret & Ahmet Ok (2014). Internet plagiarism in higher education: tendencies, triggering factors and reasons among teacher candidates, *Assessment & Evaluation in Higher Education*, 39(8), 1002-1016.
- Glendinning I. (2014). Responses to student plagiarism in higher education across Europe. *International Journal for Educational Integrity*, 10(1), 4–20.
- Heyward, E. (2000). *Plagiarism and anti-plagiarism network*, NJ: Dept. of English, Rutgers University: Newark, New Jersey, USA.
- McEvedy M.R. & Smith P. (1990). *Plan write and rewrite*. Thomas Nelson Melbourne: National Library of Australia.
- Richard, L. Saunders, P & Meek, P. (2016). How to recognise and avoid Plagiarism of Tennessee at Martin, based on a website maintained by writing tutorial services, Indiana University, Bloomington Ind. University.
- Soyemi, J. & Isikaye, F.O. (2017). A web-based final year student project duplication detection system. *International Journal of Computer Application (IJCA)*, 7(1), 1-7.
- Stein, B., Lipka, N. & Prettenhofer, P. (2011). Intrinsic plagiarism analysis. *Lang Resources & Evaluation*, 45, 63–82.
- Thomas G. (2009). *How to do your research project*. Sage Publications Inc: United States
- Turner J.R. & Müller R. (2003). On the nature of the project as a temporary organisation. *International journal of project management* 21(1), 1-8.
- Yamane, T. (1967). *Statistics, an Introductory Analysis*, 2nd Ed., New York: Harper and Row.