



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Academic Integrity And Text Matching Software Tools

Dr. Deepesh Kumar Thakur, Faculty-English, School of Liberal Arts & Management- Humanities, DIT University, Dehradun

Er. Durgesh Kumar Thakur, M.Phil Scholar, B R Ambedkar Bihar University, Muzaffarpur

Abstract

The aims of this paper is to look at the impact of global studies on issues related to academic integrity in the modern digital era. These include using automated techniques to detect plagiarism online. Additionally, it discusses a number of crucial issues at the center of managing plagiarism in the digital age. It also discusses the actual challenges that educational institutions confront in juggling the demands of academic productivity, upholding the standards of truthfulness and integrity, and using methods such as evaluation or publication to produce outcomes. Digital lessons could be misused.

Key Word: *Integrity, Digital, Software tools, plagiarism, Legal Standing*

Introduction

The presence and absence of Internet paraphernalia and World Wide Web have played an important role in the dominion of academic integrity. Many have alleged Internet for assisting in academic transgression by means of plagiarism, contract cheating, collusion and cheating. Nevertheless, others uphold the contribution of Internet related technologies as a means to extend the philosophy of authorship, association and worldwide contribution to writing and as a chance to remodel traditional concept of attribution. Worldwide the cases of plagiarism and acts of unethical academic conduct have attracted great media attention. The public perception thus generated is that there is deterioration in the academic integrity in this digital age. The emerging idea of word ownership presented a problem for states. England became a pioneer in the development of official rights law. William Wordsworth and other poets advocated for the official recognition of authors as owners of their works of art. The poets contended that as a literary work is an intellectual creation, law using proprietary rights should protect it. Ownership of books that were formerly owned by bookseller or individual readers. The recognition of a person's official rights was legislated in the Statute of Anne of 1710. The act provided authors with a legal right to claim their literary works as property, which could be enforced through legislation. Another milestone was the emergence of copyright law in England through a legal decision when in 1741 the famous English poet Alexander Pope sued Edmund Curll. Some of its essential elements related to authorship are still cherished today. Furthermore, plagiarism cannot be defined or conceptualized without it. Other European nations in the eighteenth century took a different stance on the issue of legal authorship rights. They claimed authorship

based on the idea that creators have moral rights to their creations. This took away the authors' economic right to their work while still granting them credit rights as the work's creators.

The Berne Convention: Recognition of Authorship & International Obligations

The Berne Convention for the 'Protection of Literary and Artistic Works' of 1886 is the first international agreement to protect authorial rights and the creators of literary and artistic works. Although the international copyright laws have little direct influence within national borders, the member countries are likely to integrate significant portions of them into their national legislation. To protect author rights, many nations have created their own national copyright laws. However, in order to survive in the global economy, it is becoming more and more important for countries to join international organisations like the World Intellectual Property Organisation (WIPO). Consequently, 120 nations have already ratified a number of international agreements and treaties. These countries are required to write their national laws in compliance with the international agreements they have ratified. Copyright is no exception, and it has had an impact on how various countries have dealt with plagiarism.

Plagiarism and its Legal Standing

Since it developed because of copyright law and appeals to both civil and criminal legal theories, law as such does not define plagiarism. In civil law, plagiarism is associated with a violation of the moral rights of the author. On the other hand, the idea of plagiarism is related to criminal law and is defined as misappropriating and stealing another person's words. "A thief in literature; one who steals the thoughts or writings of another" is what Mallon (1989, p. 11) refers to as a plagiarist. The terms "stolen," "illicit," "theft," "kid-napped," and "misappropriated" are also well described as legal terms in the criminal law codes of nations that have adopted the English system of law. However, conceptions of authorship and thus protection from copying cannot be understood solely through the lens of legislation. For many years, literary critics and cross-cultural theorists have written about the concept of authorship and contested the legal definition of authorship. Fascinatingly, numerous colleges across the globe have incorporated some of these words into their policies regarding plagiarism and academic misconduct. In the event of a "breach," different "penalties" or "sanctions" are applied (Sutherland-Smith 2010); however, the most frequent outcomes for those found "guilty" of the "offence" of plagiarism are a fine, suspension, or removal from study. This demonstrates how deeply ingrained these phrases are in criminal law.

Concept of Authorship in the Digital Age

The character and meaning of authorship have changed thanks to the internet in many different ways. The old concepts of textuality have come into conflict with this process because of the digital age. According to numerous studies (Howard 2007; Pecorari 2013; Sutherland-Smith 2005a, b, 2008, 2013), students do not regard the Internet in the same way as print-based literature in terms of attribution. No matter how one feels about individual, academic, or institutional perspectives on technology and authorship or ownership of texts, it is clear that in the digital age, technology is changing how we think about authorship.

Authorship as expressed in Plagiarism Policies

Institutional as well as individual author credit is given top emphasis by many academic institutions since it is thought to be crucial for academic survival. They take great effort to ensure that the staff's (and students') original works are free of plagiarism. The majority of colleges claim that their academic research staff is productive, and publications serve as proof of this claim. Plagiarism is equated with "the academic

death penalty" wherever it is found (Howard 1995; Sutherland-Smith 2008). Higher education institutions create honor codes or policies against plagiarism, which are typically posted online, to control the practice. This is done in the hopes that everyone involved in the academic life of the institution—staff, students, and faculty—will read it and agree to abide by them. The question of whether plagiarism laws should be categorised under "academic misconduct" or under "academic integrity" measures has however sparked a new internal institutional discussion.

This is not always the case because many organisations still allow for the policies below misbehavior and the structure lack of attribution in the form of "breaches" or "offences." A recent extensive study done across many European countries indicated the use of phrases like "guilt," "offence," "penalties," and "misconduct" to define how plagiarism was presented by European institutions (Glendinning 2014). Additional research suggests that university regulations are still based on antiquated ideas of authorship and that punishment continues to play a large role in how judgements are made (Hartle et al. 2009; Sutherland-Smith 2010, 2013, and 2014). This isn't always the case, though; many companies still permit infractions of the rules listed below and an absence of attribution in the form of "breaches" or "offences." A recent comprehensive study carried out across several European countries found that European institutions referred to plagiarism as "guilt," "offence," "penalties," and "misconduct" (Glendinning 2014). Another study claims that university rules are still based on antiquated ideas of authorship and that punishment still has a big impact on how judgements are made (Hartle et al. 2009; Sutherland-Smith 2010, 2013, 2014). The message the university policies are sending to the world at large is therefore the rhetoric of criminal law. However, university rules usually place their regulations in the area of student "responsibilities" under the heading of academic integrity when they create these concerns using the concept of moral integrity, or ethics. The words "ethics", "ethical practices", "and intellectual nature of work", "moral responsibility and professional responsibility" are used to define the intellectual work as an ethical practice. On the other hand, complaints of academic integrity violations must result in educational value rather than just retaliation for change to be comprehensive within an institution (Park 2004; Sutherland-Smith 2014).

Attempts to Define Plagiarism

Detection of plagiarism requires a precise definition of what constitutes plagiarism. In this regard Teddi Fishman observed, "Even among academics, there is no standard or agreed upon definition of plagiarism" (2009, p. 1). Therefore, there will be instances when in spite of a university policy or other guidelines, instructors may not be sure where the line is to be drawn between slapdash scholarship and plagiarism. Teddy proposed a definition consisting of five parts (2009, p. 5) which incorporates words, thoughts, or other products that are attributable to another person, but as such are not attributed. It encompasses plagiarism related to copy-and-paste, unattributed paraphrases or patch-writings and translated works. Fishman although does not address intention directly, but remarks that if it is used for some sort of personal gain then it falls under plagiarism. However, intention is stated as part of a number of definitions of plagiarism specifically in university policies related to cases of academic misconduct. On the other hand misappropriation of intellectual property, legalistic notions such as copyright violations or stealing are the focal point of many other definitions. An author claims that plagiarism exists in a text whether or not the submitter or publisher meant for it to. Since plagiarism is evident in the text itself, the display of purpose may be a factor in determining the repercussions or a sanction. Therefore, it is much simpler to define what is not included in any definition of plagiarism, such as altered images or falsified data, ghostwritten papers, honorary authorship, and similar practices. These behaviours are not plagiarism; rather, they constitute academic misconduct. The issue presently is that it is not practical to develop software that is capable of

consistently detecting it in the absence of a specific description. Typically, the software can only detect same word placement or, to a lesser extent, text similarities. As a result, software can only be used as a tool to perhaps identify the use of words from other sources; it cannot prove plagiarism. Thus, only humans are capable of making such decisions.

Tall Claims and Assurances

The plagiarism detection software are marketed by the related companies with all sorts of claims and assured results. Catchy tag lines like “Advanced online plagiarism detection” (CatchIt First), “Originality check” (Turnitin), “Easy, quick and accurate” (Ephorus), “based on the latest research in computer linguistics” (Plag Scan), “saves time” (Urkund) are used to woo the instructors. The tall claims made raise the expectations that it can rapidly and most reliably detect all of the plagiarism in the research papers. Furthermore the reports prepared which are to be presented to the honors board or unit coordinator are expected to clearly identify the plagiarized segments and express an ultimate judgment as to the severity of the replication. These companies also offer plagiarism-free certificates, which the students can submit along with their research papers. Additionally, these businesses provide credentials that students can submit with their research papers that are free of plagiarism. The technology detects plagiarism through comparison with a source that has previously been published. Just because plagiarism was not found does not indicate its original; it could be that the system was unable to access the source because of a variety of issues.

Shortfalls and Functionality of Current Systems

The widespread consensus is that computers can solve any problem as long as the right algorithms are established for them. One of these issues is the detection of plagiarism in texts, although using algorithms to solve it only allows for partial success. The procedure of identifying plagiarism in a work often involves two steps. Finding probable sources is the first stage. The next is determining how much text is correlating between the source that was found and the document that is being examined. The Systems use one of two methods to look for probable sources in this procedure. It can either utilise its own database of likely sources or extract a text from the document being analysed and run a search on it. Several different algorithms can be used to construct and analyse these databases.

Results Analysis

Numerous studies on plagiarism advise that when examining and evaluating the findings that the systems return, one should pay close attention. A number is used to indicate the severity of the text overlap that has been discovered. However, it is crucial to understand that many systems use various and unidentified techniques and typically only examine a small percentage of the text given for analysis, thus for the same text they will return different values. If the test is performed even after a ten-minute interval, the identical system may nevertheless produce different results. It seems to reason that each time a different section of the in question text is utilized as a test sample. Most systems also generate a report in addition to allocating a number. However, these, too, confuse the situation further rather than making it easier to comprehend the results. Most commonly, these are charged with using tangled language and grammar, meaningless or confusing numbers, perplexing markup, and making no sense. Some of them are difficult to manage in a university setting, while others are unmanageable in a general setting. In addition, there may be issues with their transfer inside the institution or with the outside examiners. Additionally, there are issues with printing it and keeping it in the student's file. . Additionally, if the report is prepared later, the results may differ since the sources are no longer accessible on the internet. In order to generate documentation

for an academic integrity board, reports must include crucial details like the line number and page of the overlapped text. As a result, it becomes challenging the board members to independently investigate the claim and they must constantly rely on the software.

The structure and the explanation text in the reports present challenges throughout the interpretation process and necessitate sufficient expertise on the part of the examiner to correctly interpret the system reports. The over-reporting of plagiarism by the systems—by which we mean that they reported more sources than was appropriate and even texts that barely overlapped the relevant paper were reported as prospective sources—is another issue with interpretation. They even mark legitimately quoted text as not original. This is seen in numerous systems and is most likely the result of flawed programming.

Creation of Database and Related Issues

Many systems store the submitted papers in a database so that they can be used to compare them to papers in the future. This seems like an excellent idea can be used for term papers on related topics that are prepared every semester. However, in situations when institutions use services to identify plagiarism and students submit their papers to the system, it is technically against their copyright. A thesis is typically created without a disclosure agreement, making it illegal to upload it to any other third-party server. Some solutions let the installation of software locally on the instructor's PC in place of utilising a database run by any organisation. The algorithm then use a search engine to verify certain passages from the texts being examined. The system either displays an exact replica of the search engine results or compiles them and evaluates them before grading and showing them to the user. However, these devices have restrictions and can only be utilised for a few number of queries per day. In order to get results fast, especially when term papers or final theses are due, the company offering the plagiarism detection software must collaborate with a search engine. As a result, this kind of software is particularly unproductive because it wastes time and produces findings that are difficult to interpret. The instructors are therefore happier with manual methods of plagiarism detection.

Services of Doubtful Nature

Some businesses offer services that are questionable. A few of them provide these services "free of charge" with the express intent of gathering texts submitted for use by paper mills. Others provide money-back guarantees on the purchase of their software if the customer is dissatisfied. However, in practice, they fall short of their word; if a customer emails customer service with concerns, such concerns are ignored, and a refund of the purchase price is therefore ruled out. Even one of the businesses transferred the papers that students submitted to a stolen Turnitin account while offering plagiarism detection to students at an exorbitant price. The latter's reports were sufficiently manipulated that it appeared as though the concerned corporation had produced them. When challenged, the corporation claimed that the similarities in the results were really an accident.

Collusion

While the software must look through a limited range of texts, there is one sort of plagiarism that certain systems can identify efficiently. Each document in this case can be compared to the others to look for texts that are similar. Due to the rapidly rising number of comparisons, this is only effective for a small number of documents. Furthermore, although though it might appear unusual for such a closed to be available, there are undeniably common circumstances at universities when this is the case. Collusion is what this is known as, or, as Zauner (2014, p. 18) describes it, "die böse Schwester der Teamarbeit" (the wicked sister

of teamwork). A typical example of collusion when students have clear cut instructions to work alone but do not, or the instructions were not so clearly conveyed, and the students collectively created a text which each of them submit with no or slight alterations (Sutherland-Smith 2013). A common instance of collusion is when students are given explicit instructions to work independently but fail to do so or the instructions are not sufficiently communicated, and the students collectively produce a manuscript that each student submits with no or minor modifications (Sutherland-Smith 2013). At the HTW Berlin in 2012, a test of 18 systems that was solely focused on finding collusion in texts and programme code was conducted (Weber-Wulff et al. 2012). The results demonstrated that there is software that can reliably detect collusion. The primary factor was that the prospective sources were found in the submitted articles rather than on the public Internet. However, it should be emphasized that the techniques that were effective at spotting text collusion turned out to be ineffective at doing the same for computer programme collusion. Furthermore, the software found collusion far more quickly in words than in programme code.

Plagiarism Detection – Advance Technological Research

Advanced technical research is being done in a number of areas to improve the effectiveness and efficiency of software that detects text matching. However, even if sophisticated methods for spotting text similarities are created, the systems will not be available as products and will not provide the "confirmation test" that administrators and teachers are looking for. At the University of Weimar in Germany, researchers are working to automate intrinsic plagiarism detection, which looks for plagiarism by examining the internal structure and pointing out variations in style rather than by looking for potential sources. In actuality, it has a strong connection to the authorship identification issue. In a similar vein, a research team from the University of Constance in Germany has been investigating citation-based plagiarism detection. This approach ignores the text and just compares the references' names, rankings, and citation styles. Currently, the citations must be manually coded, which greatly limits the applicability of the technique. However, it has been discovered, for instance, that plagiarism in translation can be found if the citation styles used in the text have a significant amount of overlap (Gipp 2014). Numerous research organisations across the world are investigating strategies to identify semantic plagiarism. These very experimental methods look for texts that have a similar meaning structure to one being used in a text and attempt to map out its meaning. Alternatively, it merely recognises synonym substitution, word insertion or deletion, and word rearrangement. In addition to these, there are experimental systems that try to extract the meaning from paragraphs so they may be contrasted with others. However, there are currently no systems that are close to being made accessible for general usage.

Conclusion

Plagiarism being a composite notion has various elucidations to the practical details. The originality and ownership of author is the basis of the plagiarism policy construction of university. Anywhere when the conventional Idealistic view of authorship is embraced along with the idea that author holds legal rights of property over their creations enforceable through any court of law, form the part of academic misconduct regulations of the university. Such policies are laid in terminology from criminal law and the outcome is penalty of some kind. The objective is to deter students from taking such actions or else the punishment is the penalty for the act. However, when the morality is recognized the focus of plagiarism policies is academic integrity sphere of policy regulations. The objective is to edify students that ascribing or acclaiming authorial creation is the truthful, moral and appropriate action to take for the contribution these authors have made towards the public realm.

The academic integrity policy lays stress on enlightening rather than penalizing the contravention and transgression perception. Various researchers would contend that concept of authorship is archaic considering the cooperative writing spaces that are present and thriving in the worldwide technology field and that the fundamental lawful doctrines of the Romantic concept of authorship are outdated. Considering that the conventional perspective of authorship will constantly straggle after the altering sphere of digital textual production, now the universities should recognize the effect of such opposing thoughts especially in shaping the academic concept, strategy and compliance with academic integrity. Accepting authorship in this digital age necessitates regularly re-examining the fundamentals emphasizing the policies and processes of university academic integrity. Comprehending inter-textuality and integrating up-and-coming textual forms in learning, it requires durable pedagogical approaches so that the action policy is put into practice. The plagiarism detection systems make plentiful promises, however the drawback are complicated and profound. Therefore, software cannot be relied upon as the lone instrument for determining plagiarism, because algorithms can be formed or erroneously executed. At the most the software can only provide evidence which must be essentially evaluated by a human being to determine whether a text is a plagiarism or not. Practically, the software must not be regularly used for all student texts, but merely be employed as a supplementary device in the academic integrity apparatus of an institution.

References:

Mallon, T. (1989). *Stolen words*. San Diego: Harcourt.

Gipp, B. (2014). *Citation-based plagiarism detection: Detecting disguised and cross-language plagiarism using citation pattern analysis*. Berlin: Springer Vieweg.

Glendinning, I. (2014). *Responses to student plagiarism across higher education institutions across Europe*. *International Journal for Educational Integrity*, 10(1), 4–20.

Hartle, R. T., Kimmins, L., & Huijser, H. (2009). *Criminal intent or cognitive dissonance: How does self-plagiarism fit into academic integrity?* In *Conference proceedings, 4th Asia Pacific Conference on Educational Integrity*, Wollongong. <http://ro.uow.edu.au/apcei/09/papers/5/>

Howard, R. M. (1995). *Plagiarisms, authorships and the academic death penalty*. *College English*, 57(7), 788–806.

Park, C. (2004). *Rebels without a clause: Towards an institutional framework for dealing with plagiarism by students*. *Journal of Further and Higher Education*, 28(3), 291–306.

Pecorari, D. (2013). *Teaching to avoid plagiarism: How to promote good source use*. Maidenhead: Open University Press.

Sutherland-Smith, W. (2005b). *Pandora's Box: Academic perceptions of student plagiarism in writing*. *Journal of English for Academic Purposes*, 4(1), 83–95.

Sutherland-Smith, W. (2008). *Plagiarism, the internet and student learning: Improving academic integrity*. London: Routledge.

Weber-Wulff, D., Köhler, K., & Möller, C. (2012). *Collusion detection system test report 2012*. [Web page]. <http://plagiat.htw-berlin.de/collusion-test-2012/>. Accessed 11 Apr 2015.

Zauner, H. (2014). *Wissenschaftliches Fehlverhalten—Münsteraner Kettenplagiate*. In *Laborjournal*, 09, 17–18.

