

IN AN EFFORT TO REDUCE ACCIDENTAL PLAGIARISM BY COMPUTER SCIENCE STUDENTS

Ed Lindoo, Ph.D.

H. Wayne Huizenga School of Business and Entrepreneurship

Nova Southeastern University, 3301 College Avenue,

Fort Lauderdale-Davie, Florida 33314-7796 | 800-541-668

772-260-1910 direct, edlind@nova.edu

ABSTRACT

Plagiarism is nothing new to academia; however, two forces are driving Computer Science majors to accidentally plagiarize on written papers. The first force is the Internet and the electronic information it provides. Because of the ease of access and the simplicity to copy/paste, many students are not always going through the regiment of analyzing, synthesizing paraphrasing, and assessing the quality and accuracy of information they copy. The second force can often be a combination of the amount of code reuse computer science instructors allow and, Computer Science instructor's not always explaining plagiarism; especially as it relates to non-Computer Science courses. As a result, it has been found within the research of this paper that Computer Science majors have a higher incidence of plagiarism as compared to other disciplines. To that end, the intent of this paper is to reinforce to Computer Science educators the importance of including a brief lecture on plagiarism at the start of each course.

INTRODUCTION

Cheating on tests and plagiarizing others works has been occurring since the dawn of higher education. During the past twenty years, one might say that this has turned into an epidemic, mostly due to electronic and online sources that simplify copying from the Internet or "cribbing from smartphones." [3] While it is true that anyone who sets out to cheat or intentionally plagiarize can easily do so, now there is also a good chance they will be caught by any one of a number of electronic tools available today. Interestingly, in February 2010, Stanford University disclosed that 23% of their plagiarism issues were from computer science majors, yet this student body made up only 6.5% of the overall University. (Marsan Network world)

If you copy text from a source, or several sources, and change only a few words here and there the result is what the Harvard Guide to Using Sources [6] calls "*mosaic plagiarism*". Even if you never intended to copy a source, you could end up committing this type of plagiarism as a result of careless note-taking and confusion as to where your source's ideas end and your own ideas begin" [3] Similarly, Rebecca Moore Howard, a professor of writing and rhetoric at New York's Syracuse University, states that "college students commonly incorporate whole

paragraphs from a source into their papers, changing only a few words, without using quotation marks”, a process she calls “patchwriting.” [3]

This is where Computer Science majors seemingly get in trouble.....more so than in any other academia. Typically Computer Science instructors focus on programming activities and while many might make mention of plagiarizing code, little is discussed about writing research papers. At the same time many of these instructors encourage code reuse. In fact Hamilton stated “Most programming courses describe code reuse as good practice” [10] Why reinvent the wheel? Find code snippets.....use them. But where exactly do we draw the line? Three lines of code...10 lines of code...10-20% of your overall program? What might become acceptable (copying and pasting a 20 line function into an 800 line program of code) in a Computer Science course is not acceptable in a management research paper. Compounding the issue for Computer Science students, computer programming textbooks and manuals rarely contain discussion of plagiarism, copyright issues or correctly referencing source code within one’s program. Yet, most College and University Computer Science disciplines require core courses that are not programming related. As Computer Science majors go back-and-forth between general academia and programming courses, confusion can easily ensue.

ONLINE MAKING IT EASY

To plagiarize, according to the Merriam-Webster dictionary, is “to steal or pass off (the ideas or words) of another as one’s own; use without crediting the source.” Merriam-Webster defines cheating as “using trickery that escapes observation”. As a result we can say there is a distinct difference between Plagiarizing and Cheating. However, we can also see that there can be varying degrees of plagiarism, and that the higher the degree the closer it is to cheating. For example, acquiring a paper from a student who turned in the paper two years prior and simply changing the submitters name on the cover page is 100% plagiarism and certainly considered cheating. Just as, obtaining a copy of the final exam (with answers) and submitting them as one’s own work, is outright cheating, not plagiarizing. Supporting this, Donald L. McCabe, a professor of management and global business at Rutgers University, found that both professors and students feel less likely to consider plagiarism cheating, especially if the plagiarized information comes in the form of copy/pasting from the Internet. Based on his interviews McCabe suggests that only one in four undergraduates consider cut-and-paste plagiarism to be serious cheating [3]

Even so, experts are split as to whether students today are less sensitive to plagiarism in an age when they constantly re-mix, copy-and-paste and re-tweet others’ creations online. “Sharing is in the DNA of the Internet,” especially on social media like Twitter and Facebook, says Urs Gasser, executive director of Harvard’s Berkman Center for Internet & Society. It is no longer clear, to the young and old just what -composes plagiarism. That said, it comes as no surprise that because of search engines and text-matching software, plagiarism seems more

common today. “It’s easier to catch”, suggests the Berkman Center's Gasser. Today a professor can confirm his/her suspicions in a matter of seconds! Just grab some text and put it in Google, or better yet, upload the entire paper to Turnitin.com[3]

In an effort to eliminate plagiarism, many schools scan student papers using plagiarism-detection software such as Turnitin which matches students’ writing against a database of published sources and previously submitted student papers from around the world. Approximately 1,500 colleges and 4,000 secondary schools use Turnitin, in addition to a number of grade schools, and about 100 college admissions offices use it to check the originality of essays on applications. (Turnitin.com) During 2011, Turnitin saw more than 60 million submissions, however there are dozens of other similar programs, such as Blackboard's SafeAssign. The bottom line is that students need to be aware that their submissions are probably being scanned.

Another option, one that has been implemented for a number of years by many schools, is to let Student’s check their papers on Turnitin before submitting them. It has long been felt that students would learn from their mistakes. However Harrington [4], states that it “is more likely to teach students how to right-click words” for synonyms “and scramble phrases to get acceptable scores on Turnitin” Further, students have learned to take advantage of Internet tools to avoid detection, such as using Google Translate, to translate a plagiarized paragraph into Spanish and then back into English so that it uses different wording from the original. What’s really alarming is that we find students in postgraduate programs lacking the critical skills that they should have to compose a proper paper [2].

HOW DO STUDENTS RESEARCH IN THE DIGITAL AGE?

Many educators know that student’s reliance on Internet-based search has replaced the more rigorous and traditional approaches to research. In a recent report the Pew Research Center found that while teachers value that which the Internet provides (99 percent agreement) for empowering student access to information, the report concluded that digital technologies are more distracting than helpful from an academic standpoint. In short, what constitutes “research” for students today has come to mean “Googling.” [7]

The Pew report shows that the ease with which information “appears” online allows students to avoid any of the questions that may surface concerning the quality and intent of information they “research.” The Pew survey revealed that only one percent of those surveyed reported as “excellent” the ability of students “to recognize bias in online content.” As for their “ability to assess the quality and accuracy of information they find online,” only three percent reported that they found students to be “excellent.”

This data supports the following insights into student research behavior, specifically:

- **Students appear to value immediacy over quality in online research,**

The ease with which “the answer” may be found online places sites such as Wikipedia, homework help sites, answer sites, and other social and content sharing sites to the top in terms of source matches.

- **Students often use cheat sites and paper mills as sources**

Less a research competency issue than a moral and ethical one, the significant number of sources that match to cheat sites and paper mills suggest that for students there is a bias towards immediate outcomes and results rather than towards concerted effort to meet assignment goals.

- **There is an over reliance on the “wisdom of the crowd”**

Students appear to demonstrate a strong appetite for crowd-sourced content in their research. Though it is not immediately evident why students seek these sources out, the strong reliance on these types of sites indicate difficulty assessing the authority and legitimacy of the content these sources present.

- **Student “research” is synonymous with “search”**

The frequent and uninhibited use of sites with limited educational value (as defined by the quality and authority of content) in student work underscores a preference for “searched,” rather than “researched” content.

- **Existing student source choices warrant a need for better search skills**

In addition to a preference for immediacy, the popularity of crowd-sourced content online indicates that a majority of students are engaging in cursory or shallow searches for content. At play may be an absence of awareness of how search engines work and how to effectively conduct searches to find appropriate content. What also appears to be absent is the use of criteria (whether internally—or externally—defined) to judge that content [7]

Similarly, Turnitin found that just over half of Internet sources in student writing came from legitimate educational resources. 57% of matches come from academic and homework sites, news and portal sites and encyclopedias. That means that 43% of Turnitin matches lead to sites that are academically suspect, including cheat sites and paper mills, shopping sites, and social and user-generated content. Drilling deeper into this, it was found that 19% of content matches came from paper mills and cheat sites so one can conclude that at a minimum, 19% of matched content comes from sources of academic disrepute. With Wikipedia being the most popular source for unoriginal content in student papers, and the ease to which students can copy/paste from electronic sources, the burden has been placed upon instructors to spend more and more time looking for plagiarism when their focus should be on teaching [10]

CONCLUSION

It has become clear that many computer science majors do not have an adequate understanding of why copy/paste procedures are invalid in general academia. It is also clear that this is typically the reason that their research papers are being exposed. Students feel they know how to use Google, however, when it comes to academic research, Google is a starting point at best. The real danger for students is the ease at which they can copy/paste electronic information. With the Internet and online plagiarism checkers flagging thousands of papers each term, teachers are being taxed to turn-in student plagiarism. While it is not clear just how many additional hours teachers now spend reviewing flagged papers, there can be no doubt that it has added time and stress into each instructors course loads, ultimately rendering less time for educating.

Therefore it is recommended that computer science instructors spend more time up front in their courses not only discussing plagiarism but to demonstrate techniques to avoid accidental plagiarism. For example, breaking the habit of copy/paste is hard so as one does this, highlight all of the copied text and change the background to some color. Use different colors for different sources and once research has been done, start paraphrasing. Different instructors, different techniques, but the important thing to remember; the more we can do upfront to prevent plagiarism, the less we will see.

REFERENCES

- [1] Callahan, D., *The Cheating Culture: Why More Americans Are Doing Wrong to Get Ahead*, Harcourt, 2004.
- [2] Fishman, T., International Center for Academic Integrity. www.academicintegrity.org, 2013.
- [3] Glazer, S., *Plagiarism and Cheating*. CQ Researcher Online. Retrieved 3/20/13 from: <http://photo.pds.org:5012/cqresearcher/document.php?id=cqresrre2013010400&action=print>, 2013.
- [4] Harrington, D., “(Moral) Hazards of Scanning for Plagiarists,” Blog: Moving Markets into the Light <http://davidharrington.com/?p=594>, 2011.
- [5] Hamilton, M., Tahaghoghi, S., and Walker, C. *Educating Students about Plagiarism Avoidance – A Computer Science Perspective*. RMIT University, 2004.
- [6] *Harvard Guide to Using Sources*, Harvard College Writing Program, Retrieved 3/26/13 from <http://isites.harvard.edu/icb/icb.do?keyword=k70847&pageid=icb.page342054>
- [7] Parker, K., Lenhart, A., and Moore, K., *The Digital Revolution and Higher Education*, Pew Research Center, Aug. 28, 2011, www.pewsocialtrends.org/2011/08/28/the-digital-revolution-and-higher-education, 2011.
- [8] Pinsky, M., Instructors Web Site. Retrieved 3/27/13 from <http://chuma.cas.usf.edu/~pinsky/teachers.htm>, 2013.

[9] Tomar, D., *The Shadow Scholar: How I Made a Living Helping College Kids Cheat*, p. 110., 2012.

[10] *The Sources in Writing – Higher Education*. White paper retrieved May 2, 2013 from www.turnitin.com

[5zz] Hamilton, M., Tahaghoghiih, S., and Walker, C. *Educating Students about Plagiarism Avoidance – A Computer Science Perspective*. RMIT University, 2004.