WHITE PAPER Plagiarism and the Web

A Comparison of Internet Sources for Secondary and Higher Education Students

turnitin®

Prevent Plagiarism.
Engage Students.

Table of Contents

1.0 Summary

This Turnitin study offers a look into the web sources and writing practices of secondary and higher education students in the United States. It is based on the analysis of 128 million content matches from 33 million papers (24 million from higher education and nine million from secondary) over a one-year period.

Key findings include:

- **Secondary students rely more on social networking and content-sharing sites**
  Social and content sharing sites remain the most popular category for both groups of students, though secondary students show an even greater reliance on these sources. Based on the data, it seems younger students have a more difficult time judging which sources are appropriate to use in written work.

- **Higher education students rely more on paper mills and cheat sites**
  Higher education students rely more on paper mills and cheat sites than secondary students. This may reflect the fact that college students, independent for the first time, struggle to manage their time amid an increased academic workload.

- **Wikipedia is the most popular site for both groups**
  Not surprisingly, Wikipedia reigns supreme for both secondary and higher education students with the latter showing the greatest reliance.

- **Eight sites appear among the top ten most popular web sources for both secondary and higher education students**
  Despite their differences, the two groups rely on a number of the same sources for written work. Eight of the ten top sites appear as the most popular sources for both groups of students.

- **Educators should develop specific strategies to address plagiarism**
  Creating plagiarism-proof assignments, dedicating lessons to proper sourcing and citation, and using originality checking tools in a formative manner can help educate students on originality and citation while avoiding costly and painful misconduct cases.

2.0 Introduction

If there was any doubt that plagiarism is a growing issue in this digital age it was laid to rest in a recent authoritative survey from the Pew Research Center. The survey asked 1,055 presidents of two- and four-year colleges and universities whether the problem of plagiarism has increased, decreased or stayed the same over the last 10 years. Only two percent of respondents said the problem of plagiarism has decreased while 55 percent said it has increased. Of those who believed the problem of plagiarism has increased, 89 percent believed computers and the Internet played a major role.

According to a recent Pew Center survey of College Presidents, plagiarism is a growing problem.

2011 Source: Pew Research Center

Increased

55%

Not Changed

40%

Decreased

2%
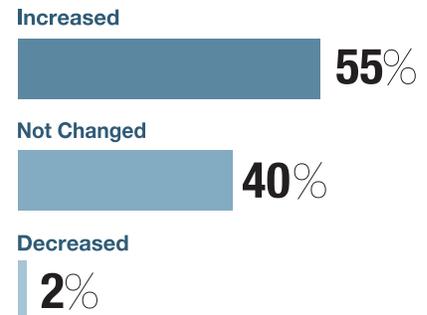
This study is the second in a series that examines the web sources that students rely on for content in their written work. The intent is to give educators greater knowledge about student sources and how to best engage and educate students about correct sourcing, citation, and attribution.

The first study, published in April 2011, found that the highest number of matches between student writing and the Turnitin Web database was from social and content sharing sites. The study also found that students relied more on legitimate educational sites (25 percent) than paper mill and cheat sites (14.8 percent).

The study provoked a good deal of discussion in the media, as well as academic and educational technology circles. Educators repeatedly requested a breakdown of student sources showing the differences between secondary and higher education students.

This study compares the web sources that secondary and higher education students use in writing assignments. By comparing the two groups, we can better understand the challenges facing educators in high schools and universities and craft more effective solutions for combating a growing problem.

3.0 Matched Content Categories

For the purposes of this study, Turnitin classified each site that appeared in a matched content report into one of the following categories.

SOCIAL NETWORKING AND CONTENT SHARING

These sites rely on a user-generated content model. The sites include social networking sites such as Facebook and MySpace, content sharing sites such as Scribd and SlideShare, and Question & Answer sites such as Yahoo! Answers and Answers.com.

HOMEWORK & ACADEMIC

The next category consists of academic, educational and homework help sites that offer legitimate educational content. Sites such as nih.gov, medlibrary.org, coursehero.com and bookrags.com are included in this category.

NEWS AND PORTALS

This category includes web sites that use a traditional publisher model such as The New York Times or Huffington Post.

PAPER MILLS AND CHEAT SITES

This category of sites share or sell papers. Sites include oppapers.com and studentoffortune.com.

ENCYCLOPEDIAS

Wikipedia, Brittanica.com and Encyclopedia.com are three of the most popular sites in this category.

OTHER

The "Other" category contains shopping sites such as Amazon.com as well as other sites that do not fall into any of the categories described above.

About This Study

This report is based on 128 million content matches from 33 million papers that were submitted to Turnitin.com between June 2010 and June 2011. Nine million papers were from secondary students. Twenty-four million came from higher education students.
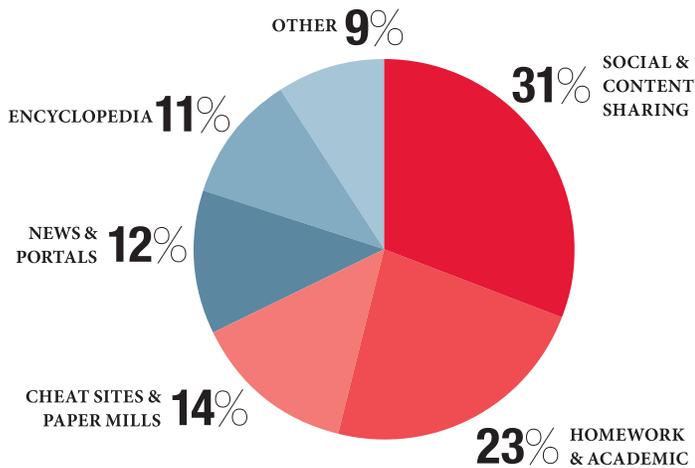
This report uses the same classification methodology as the previous report in this series. For this study we discovered a few sites that were classified incorrectly in the original study because they were actually "mirrors" of Wikipedia; meaning a personal blog that appeared to be a very popular site for student research was classified under "Social and Content Sharing" when it was, in fact, reproducing the contents of Wikipedia.

In this report, the site was categorized in the "Encyclopedia" section and removed from "Social and Content Sharing." This correction changes the earlier findings by roughly a percentage point but does not alter the analysis in a meaningful way.
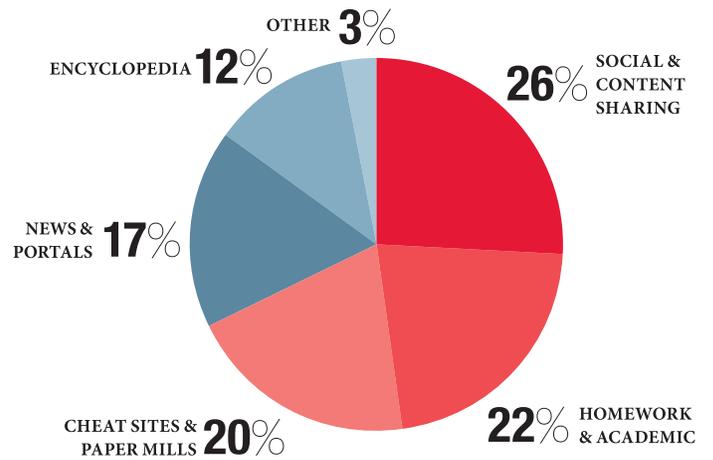
4.0 The Most Popular Web Sources for Secondary and Higher Ed Students

The graphics below show a high-level view of the most popular categories of matched content in secondary and higher education papers.

SECONDARY EDUCATION

OTHER **9**%

ENCYCLOPEDIA **11**%

NEWS & PORTALS **12**%

CHEAT SITES & PAPER MILLS **14**%

31% SOCIAL & CONTENT SHARING

23% HOMEWORK & ACADEMIC

HIGHER EDUCATION

OTHER **3**%

ENCYCLOPEDIA **12**%

NEWS & PORTALS **17**%

CHEAT SITES & PAPER MILLS **20**%

26% SOCIAL & CONTENT SHARING

22% HOMEWORK & ACADEMIC

The largest discrepancy between secondary and higher education students is in the categories of social and content sharing web sites, which are more popular among secondary students; cheat sites and paper mills which are more popular among college students; and the "other" category which includes sites such as Amazon.com. Secondary and higher education students' reliance on encyclopedias and news sites is virtually identical.

5.0 The Top Ten Most Popular Sites

While in different order, eight of the ten sites appear on the top ten lists of both secondary and higher education. The bolded sites below appear on only one of the two lists. Wikipedia is the most popular single source for both secondary and higher education students.

	Secondary		Higher Education	
1)	Wikipedia	7.99%	Wikipedia	10.74%
2)	Yahoo! Answers	7.55%	Yahoo! Answers	3.90%
3)	Answers.com	3.37%	Slideshare	3.87%
4)	eNotes	2.90%	Answers.com	3.57%
5)	Slideshare	2.38%	Oppapers.com	3.11%
6)	Scribd	2.38%	**Coursehero**	3.01%
7)	Oppapers.com	1.93%	Scribd	2.95%
8)	Amazon	1.85%	**Justanswer.com**	1.60%
9)	**Essaymania**	1.83%	eNotes	1.58%
10)	**123helpme.com**	1.71%	Amazon	1.21%

For secondary students, social and content sharing sites appear four times (Yahoo! Answers, Answers.com, SlideShare and Scribd). Three cheat sites appear in the top ten (Oppapers.com, Essaymania, 123helpme.com) as well as one academic site (eNotes), and one site categorized as other (Amazon).

For higher education students, five social sites appear (Yahoo! Answers, Answers.com, SlideShare, Scribd and Justanswer), two academic sites (Coursehero, eNotes) as well as one cheat site (Oppapers.com) and Amazon in the 'Other' category.

6.0 Key Findings

Analyzing the results led to these conclusions about the state of plagiarism and research on the Web:

- **Secondary students rely more on social networking and content sharing sites**
  By a margin of five percent, secondary students rely more on social and content sharing sites. The most obvious conclusion to draw is that younger students do not have as sound an understanding of what constitutes a proper source for written work. Instead, they are sourcing material from familiar sites.

- **Higher education students rely more on paper mills and cheat sites**
  Paper mills and cheat sites are more popular among higher education students by a margin of 5.5 percent. This may be due to the pressures students face in college. Independent for the first time, students must balance a heavy academic workload with other commitments, including social, work or extracurricular activities.

- **News and portals are most popular among higher education students**
  Students in higher education showed a great dependence on news and portal sites for information in their writing. Possible explanations for this difference could be based on behavioral patterns of the age group, the nature of the assignment, deeper engagement in current events, or a better understanding of credible sources among higher education students.

- **Wikipedia reigns supreme, especially in higher education**
  It is no surprise that Wikipedia is the leading source for matched content in both secondary and higher education. For higher education, Wikipedia appeared in three times as many matches as the second most matched source, Yahoo! Answers. In secondary, Yahoo! Answers nearly eclipsed Wikipedia in terms of popularity, once again demonstrating the popularity of content sharing sites among secondary students.

- **There are productive steps educators can take**
  Educators in secondary and higher education can help improve research, citation and attribution in student writing by implementing the approaches outlined in the next section.

7.0 Recommendations for Educators

Here are some steps educators can take to help students become better writers:

1) CREATE PLAGIARISM-PROOF ASSIGNMENTS

When designing writing assignments, choose a topic that requires an individualized approach related to personal experience or current events. Assignments are more difficult to plagiarize when students are required to submit their work at different stages of the writing process: pre-writing, drafting, revising and reflecting. The multiple drafts should reflect a significant number of progressive revisions.

2) HELP STUDENTS UNDERSTAND PROPER CITATION

Instructors should do a quick in-class activity to reinforce the difference between paraphrase, summary, and quote. Talk about your expectations for incorporating outside sources—what is acceptable and unacceptable, and explain how students should cite them. Finally, explain what plagiarism is, why it is inappropriate, and what the potential consequences are for it.

3) USE ORIGINALITY CHECKING IN A FORMATIVE MANNER

For educators using Turnitin, allowing students to view their OriginalityCheck reports for their drafts can help educate them about proper citation. Create an additional assignment that asks them to reflect and explain the Originality Report results in writing. Did they feel the report was accurate? Were there areas where they used too much unoriginal content? What will they do in their next revision to improve their paper and discover their own voice?

Finally, we encourage educators to share the information in this report with colleagues and students: https://turnitin.com/static/results/plagiarism_report.php. Only by truly understanding the growing problem of plagiarism can we actually solve it.

About iParadigms, LLC

iParadigms, LLC is the world's leading provider of web-based solutions for plagiarism prevention. The company's products include Turnitin, used by educators worldwide to check students' papers for originality, to enable web-based peer review and for digital grading of student work. iParadigms' iThenticate solution is the leader in plagiarism detection in research, publishing and many other commercial markets. The company's solutions check millions of documents each month and are used in over 100 countries. iParadigms is headquartered in Oakland, CA with an international office located in Newcastle, United Kingdom. iParadigms is backed by Warburg Pincus. For more information, please visit www.turnitin.com.