

Traci Arnett Zimmerman

TWENTY YEARS OF TURNITIN: IN AN AGE OF BIG DATA, EVEN BIGGER QUESTIONS REMAIN

A funny thing happened on the way to academic integrity. Plagiarism detection software (PDS), like Turnitin, has seized control of student intellectual property. While students who use Turnitin are discouraged from copying other work, the company itself can strip mine and sell student work for profit.

--Sean Michael Morris & Jesse Stommel, "A Guide for Resisting EdTech: The Case Against Turnitin"

It was never about the loom per se. It's always about who owns the machines; it's about who benefits from one's labor, from one's craft.

--Audrey Watters, "Education Technology's Completely Over"

It is hard to believe that Turnitin has been around for 20 years. What is perhaps harder to believe is that, despite the robust challenges and tough questions it has faced during that time, Turnitin is *thriving*. Why this fact should matter to educators and students – and why, particularly, now – is the subject of this exploration.

SOME BACKGROUND AND CONTEXT

According to the information in the Turnitin [Company Background](#):

Turnitin® is the leading originality checking and plagiarism prevention service used by millions of students and faculty, and thousands of institutions worldwide.

Founded in 1996 by a group of researchers, teachers, mathematicians and computer scientists from University of California, Berkeley, iParadigms set out to create a new model for the protection of written work from misappropriation on the Internet. Ultimately, their vision was to create an instructional tool that would increase student engagement in the process of preparing written work, and help them better learn to write, think and reason.

Today, Turnitin is a key component of education worldwide. It stands at the forefront of a building awareness for best practices in teaching and learning with a service that has expanded far beyond just originality checking. Turnitin now offers a full complement of web-based services to manage a 100% paperless process for reviewing and evaluating written assignments.

It is demonstrably true that Turnitin is the “leading” plagiarism detection product¹ and, as such, it is one of the most lucrative: in fact, when iParadigms (the creators of Turnitin) was acquired by Insight Venture Partners and GIC² in 2014, it sold for \$752 million (iParadigms). According to Chris Caren,³ the chairman and CEO of iParadigms at the time of the sale, this partnership would allow Turnitin to “execute [its] vision to be the most innovative and effective technology for evaluating and improving student work” by “provid[ing] the resources and support to accelerate [its] product investment plans and expand [its] reach into international markets” (Ibid).

But what remains at issue is how Turnitin has made such profits, how they’ve built such a successful product, and how it “protects written work from misappropriation on the Internet” and “builds awareness for best practices in teaching and learning” (Turnitin “Backgrounder,” emphasis added).

CHALLENGES AND CONTROVERSIES

Despite their outward-facing success, Turnitin has faced numerous challenges over the last decades. The history of these challenges provides a scaffold upon which to build a clearer understanding of our current moment; thus, what follows here is a brief summary of some of the more formative challenges Turnitin has faced.

Questions of Accuracy

There have been multiple investigations into Turnitin’s accuracy, and the findings range from reports indicating that Turnitin incorrectly labels *too much* student writing as plagiarism (Lang, Texas Tech, 2009) to reports that Turnitin actually catches *too little* student plagiarism. (Schorn, 2007; Schorn 2015).

Susan Schorn’s (UT-Austin) studies, and Carl Straumsheim’s reporting on them for *Inside Higher Ed*, cut right to the heart of the matter: the software doesn’t work. As Schorn asserts:

We say that we are using this software in order to teach students about academic dishonesty, but we’re using software we know doesn’t work. In effect, we’re trying to teach [students] about academic honesty by lying to them. (Straumsheim).

Susan Lang (Texas Tech) agrees, indicating that while she was “not surprised” by Schorn’s findings – particularly given the “exponential explosion of information” on the Internet and the “constant shifting and reposting of content,” she is surprised that Turnitin’s software hasn’t really improved their performance over the years (Ibid). And

¹ There are a number of proprietary plagiarism detection products out there, such as SafeAssign, Plagscan, and Quetext with DeepSearch, and an equal or greater number of “free” products, such as PlagiarismChecker, Grammarly, Dupli Checker, Plagiarisma (though most of these offer “premium” plans, where you have to pay for greater access and additional features. Still, Turnitin is the leader in the plagiarism detection software (PDS) market.

² Insight Venture Partners is “a leading global private equity firm focused on high-growth investments in the technology sector,” and GIC is “Singapore’s sovereign wealth fund” (iParadigms).

³ Chris Caren is now the CEO of Turnitin: http://v1.turnitin.com/en_us/about-us/management-team

Doug Hesse (University of Denver) points out a more troubling issue when flagging “false negatives” with Turnitin: it creates a “dog-bites-man twist on the usual critique of this software...I’d rather have students slip through the cracks than stand falsely accused” (Ibid). But the bigger question here is less about who stands accused or who slips by, and more about why we continue to use this software in the first place.

One way that Turnitin attempts to resolve this tension is to claim that they don’t actually detect plagiarism. In a Turnitin [blog post](#) from 2013 titled -- “Does Turnitin Detect Plagiarism?” -- the company takes aim at what they call the “#1 misconception” about their service: “that Turnitin detects plagiarism.” To combat this “misconception,” they provide the following reasoning:

Turnitin does not detect plagiarism per se; Turnitin just finds text that matches other sources in the vast Turnitin databases and shows those matches. It is up to a human being to determine whether those text matches are a problem or not. (“Does Turnitin.”)

What’s interesting is the paradox in the first sentence: on the one hand, it proclaims that Turnitin does not detect plagiarism “per se” (that is, it doesn’t detect plagiarism “by itself” or “in itself”), but on the other hand, it asserts that Turnitin “just” finds text that matches other sources (as if that is “only” or “simply” what it does). The second sentence follows up with a more direct interpretation: Turnitin doesn’t accuse (or acquit) students of plagiarism, people do.

Still, Turnitin’s home page at turnitin.com makes clear that they do some form of plagiarism detection since the homepage banner reads: *Improve Writing. Prevent Plagiarism.*

The image shows a screenshot of the Turnitin website homepage. At the top left is the Turnitin logo. To the right are navigation links: K-12, Higher Ed, Resources, Community, Support, and a red button labeled "Contact Sales". Below the navigation is a blue banner with the text "Improve Writing. Prevent Plagiarism." and "Turnitin creates technology to improve student writing and empower original thinking." There are two buttons: "For K-12" and "For Higher Ed". A small "We're Hiring" button is in the top right corner of the banner. Below the banner is a white section with the text "Turnitin Announces Partnership with Educational Institutions to Address Contract Cheating".

Questions of Ownership and Ethics

Turnitin has also been challenged by various stakeholders on questions of ownership and ethics. There has been one major lawsuit in the United States [[Vanderhye v. iParadigms, LLC, 562 F.3d 630, 645 \(4th Cir. 2009\)](#)], where [high school students in Virginia and Arizona brought suit](#) against Turnitin for copyright infringement, but the Appellate Court upheld the District Court's ruling in favor of iParadigms, deeming the archiving of the student work as ultimately constituting fair use.

Even before the lawsuit in the United States, there were [challenges](#) to the required use of Turnitin at McGill University (Montréal, Québec) where both individual students and, later, the Canadian Federation of Students protested university requirements to submit work to Turnitin; in 2013, McGill University [announced](#) that it had ended its licensing agreement with Turnitin and, while they indicated that they would “explor[e] other software options,” they also make clear that “students [must be allowed] a reasonable alternative if they choose not to use the technological solution”(McGill).

In addition, professional organizations like the [Council for Writing Program Administrators \(CWPA\)](#), the [Conference on College Composition and Communication \(CCCC\)](#) and the [CCCC- Intellectual Property Caucus](#) all published statements that question the ethical and pedagogical soundness of Turnitin and other plagiarism detection services and offer some best practices for teaching (and practicing) academic integrity.

And when you consider that an institution-wide license for Turnitin software is somewhere “in the five figures,” it invokes much larger ethical issues. As Schorn says, The real ethical question is how you can sell a product that doesn't work to a business — the sector of higher education — that is really strapped for cash right now. We're paying instructors less, we're having larger class sizes, but we're able to find money for this policing tool that doesn't actually work. (Qtd in Straumsheim).

TURNITIN'S RESPONSES

Over the past 20 years, Turnitin has mounted numerous responses to the challenges and questions they've received. In one (rather controversial) move, Turnitin paid to send instructors to the 2009 Conference on College Composition and Communication (CCCC) to present their (less critical/more positive) research on Turnitin. Deftly covered by Scott Jaschik in *Inside Higher Ed*, the move was met with skepticism, particularly since Turnitin did not work with the National Council of Teachers of English (NCTE, the parent organization for CCCCs) and CCCCs to gain space at the conference. Instead, they emailed individual professors (who work at institutions who had adopted Turnitin) and offered to pay travel expenses for those who would be willing to present about the plagiarism detection service (Jaschik, “Buying”). Beyond the problematic, surreptitious nature of their solicitation, Turnitin was also “not forthcoming” about which panelists were (and were not) funded (Jaschik, “False Positives”). More than just a conflict of interest issue, the question of funding travel to

the conference was “especially sensitive” because “many of the people most knowledgeable about teaching composition are adjunct professors or full timers who are [not on] the tenure track and who frequently don’t have the same access as tenured professors to travel budgets and research support” (Jaschik, “Buying”).

Turnitin also put together what they labeled an “in-depth literature review “(2010) of [“Independently Published Studies on Turnitin Services”](#) made up of “35 publications [15 journal articles, 7 conference presentations and 13 other independent reports] that address the effectiveness of Turnitin services in education in reducing plagiarism and in improving the understanding and attention to academic integrity” (2). The majority of the publications (23 of 35) cited are from countries other than the United States: primarily the UK, Australia, and New Zealand, as well as Malaysia and Botswana. And, interestingly, these rather disparate studies trend toward indicating that Turnitin works best when students retain some agency (e.g., when they are able to upload, view and revise from their “originality reports”) and when faculty use transparent and “more traditional” means of educating students about the importance and role of documentation in furthering academic integrity.

TURNITIN AT 20

Despite these challenges, Turnitin remains extremely popular and is “the standard bearer for plagiarism detection for high schools and colleges” (Roll). More troubling is that the “ethical debate [seems to have] died down as the use of plagiarism detection software has increased” (Schorn, qtd in Straumsheim).

But in an age where questions of privacy, control, and use of big data are at the fore, the questions about Turnitin are also changing. Nowhere is this shift better articulated than in “A Guide for Resisting Edtech: The Case Against Turnitin,” by Sean Michael Morris and Jesse Stommel. In their essay, published in *Hybrid Pedagogy* last year, they reopen and recontextualize the debates about Turnitin over the last 20 years by seizing on the kairotic moment of 2017 and filtering it through the brilliant work that has been done on “critical digital literacies” over the last decade. Morris and Stommel also point out that the context(s) in which we understand Turnitin’s business model are changing:

The internet is increasingly a privately-owned public space. On April 3, 2017, Donald Trump signed into law a bill overturning Obama-era protections for internet users. The new law [permits Internet Service Providers \(ISPs\) to access, without permission, data about our internet use patterns](#)— from the sites we visit to the search terms we use. And this data isn’t restricted to the work we do on computers. Thanks to the [“internet of things,”](#) all our various connections can be monitored by our ISPs — from our physical location to the temperature we keep our homes to the music we ask Alexa to play for us. (In fact, Alexa processes all of our speech when it is on, even when we are not addressing it.) Every day, we participate in a digital culture owned and operated by others — designers, engineers, technologists, CEOs — who have come to understand how easily they can harvest our intellectual property, data, and the minute details of our lives. To resist this (or even to more

consciously participate in it), we need skills that allow us to “read” our world (in the Freirean sense) and to act with agency. (“A Guide”)

The “skills” that Morris and Stommel highlight are the development and deployment of critical digital literacies that prompt us to “get under the hood of edtech tools” and to accurately differentiate between “what those tools say they do” and “what they actually do” (Ibid). To guide this exploration, they provide a rubric that asks a range of specific questions – focused on everything from the ownership, politics and operation of the tool to its accessibility and design. This exercise is “not [created] to ‘take down’ or malign any specific tools or edtech companies” but to encourage users to “think in ways they haven’t” about these tools (Ibid).

But beyond just “thinking differently,” Morris and Stommel also want to inspire action and empower instructors to the kind of “ethical, activist” work of responding to (and building sites of resistance against) “those institutional demands we find unethical or pedagogically harmful,” particularly when we begin to understand that:

[s]ome platforms are not agnostic. Not all tools can be hacked to good use. Critical Digital Pedagogy demands we approach our tools and technologies always with one eyebrow raised. Some tools have good intentions squandered at the level of interface. Some tools have no good intentions at all. And when tools like these are adopted across an institution, the risks in mounting a resistance can be incredibly high, especially for contingent staff, students, and untenured faculty. (“A Guide”)

What is clear in Morris and Stommel’s piece is that Turnitin is “one of those” platforms. They support this assessment with a raft of evidence: a close analysis of Turnitin’s claims; an even closer review of Turnitin’s policies, practices and profits; and a powerful assessment of the way in which Turnitin “supplants teaching” by dehumanizing and mechanizing the ways in which we promote the understanding of authorship, ownership and academic integrity:

To an issue of academic integrity that has been the project of teaching for decades, educational technology answers with efficiency. Plug it in. Add it up. Point a finger... So, if you’re not worried about paying Turnitin to traffic your students’ intellectual property, and you’re not worried about how the company has glossed a complicated pedagogical issue to offer a simple solution, you might worry about how Turnitin reinforces the divide between teachers and students, short-circuiting the human tools we have to cross that divide. (“A Guide”)

The final section of Morris and Stommel’s essay offers a template for resistance in the form of a draft letter to be “remixed, re-imagined and revised” so that students and faculty can work to “advocate,” “educate” and build sites of resistance on their campus against Turnitin. The letter is informative, clear and “noncombative” in its language and is designed to be a living document (to be revised, reused and shared).

CODA

In response to a comment posted to “A Guide to Resisting Edtech,” which criticized the “dated” nature of their research, Stommel responded:

Agreed that many of the points here about Turnitin are dated. Sean and I have been making many of these points since at least 2007 or earlier. The problem: Turnitin's market share keeps growing. Which makes the issues here pretty evergreen in my view. ("A Guide").

Stommel's observation is (if I may re/mix up the metaphor) already "bearing fruit." Later this year, Turnitin will be debuting its newest product – "Authorship Investigation" – which will "monitor and learn the writing styles of individual students and flag content which shows considerable divergence from their previous work" (Warner). Ostensibly, this product is designed to address the problem of "contract cheating," that is "third-party produced work that is turned in for credit" (Ibid).

When I read about this new product, I thought immediately of the final entry included in Turnitin's own Literature Review of "Independently Published Studies on Turnitin Services" – a study by Wright, Owens and Nigel (2008), which offered the following advice about the use of Turnitin:

[B]oth students and teachers are trained to recognize that Turnitin should be regarded only as a tool to assist in the detection of plagiarism. It is designed to be used in tandem with (rather than entirely replace) traditional measures for assessing academic integrity (*such as identifying abrupt changes in written structure and style*). (11, emphasis added)

What's striking about this advice is that it makes clear that there are some things that should not be mechanized: "traditional [read: *human*] measures for assessing academic integrity." Ironically and poignantly, this is the focus of Turnitin's newest tool, inadvertently presaged by an "independent study" a decade ago, which warned against precisely such developments.

In addition, the argument that Turnitin is "only a tool" tends to oversimplify the ways in which it has been wielded. Morris and Stommel note this brilliantly:

While not exactly the Luddism of the 19th Century, we must ask ourselves when we're choosing edtech tools who profits and from what? Because so much of educational technology runs on the labor of students and teachers, profiting off the work they do in the course of a day, quarter, or semester, it's imperative that we understand deeply our relationship to that technology — and more importantly the relationship, or "arranged marriage," we are brokering for students. ("A Guide").

As of this writing, individual licenses for Turnitin products are not available. They can only be purchased for school districts, multiple schools or single schools. Thus, in order to understand our relationship to this technology, we have to understand how and why it is (still) being procured, who or what administrative body makes those decisions, and how we might preempt (or push back) on the arrangements that may have already been made on our behalf.

WORKS CITED

"CCCC-IP Caucus Recommendations Regarding Academic Integrity and the Use of Plagiarism Detection Services." 2006,

- <https://web.archive.org/web/20080617190408/http://ccccip.org/files/CCCC-IPpositionstatementDraft%209%2016%2006.pdf>
- CBC News. "McGill Student Wins Fight Over Anti-Cheating Website." 16 January 2004, <http://www.cbc.ca/news/canada/mcgill-student-wins-fight-over-anti-cheating-website-1.497770>
- Conference on College Composition and Communication (CCCC). "Intellectual Property-Related Motion at the CCCC Business Meeting," 16 March 2013, <http://cccc.ncte.org/cccc/committees/ip/ipreports/april-2013-report>
- "Defining and Avoiding Plagiarism: The WPA Statement on Best Practices." 2003, <http://wpacouncil.org/positions/WPAplagiarism.pdf>
- iParadigms. "Acquisition of iParadigms for \$752 Million Led by Insight Venture Partners." CISION PRNewswire, 2 June 2014, <https://www.prnewswire.com/news-releases/iparadigms-to-be-acquired-for-752-million-by-insight-venture-partners-261553291.html>
- Jaschik, Scott. "Buying Its Way Onto the Program?" *Inside Higher Ed*, 2 May 2008, <https://www.insidehighered.com/news/2008/05/02/turnitin>
- . "False Positives on Plagiarism." *Inside Higher Ed*, 13 March 2009, <https://www.insidehighered.com/news/2009/03/13/detect>
- McGill University. "Termination of Turnitin Service." 11 June 2013, <https://www.mcgill.ca/it/channels/news/termination-turnitin-service-227074>
- Morris, Sean Michael, and Jesse Stommel. "A Guide for Resisting EdTech: The Case Against Turnitin." *Hybrid Pedagogy*, 15 June 2017, <http://hybridpedagogy.org/resisting-edtech/>
- Roll, Nick. "New Salvo Against Turnitin." *Inside Higher Ed*, 19 June 2017, <https://www.insidehighered.com/news/2017/06/19/anti-turnitin-manifesto-calls-resistance-some-technology-digital-age>
- Straumsheim, Carl. "What is Detected?" *Inside Higher Ed*, 14 July 2015, <https://www.insidehighered.com/news/2015/07/14/turnitin-faces-new-questions-about-efficacy-plagiarism-detection-software>
- Turnitin. "Company Background." 2010, https://turnitin.com/static/resources/documentation/turnitin/sales/Turnitin_Company_Background.pdf
- . "Does Turnitin Detect Plagiarism?" 23 May 2013, http://turnitin.com/en_us/resources/blog/421-general/1643-does-turnitin-detect-plagiarism
- . "Understanding Account Types." https://guides.turnitin.com/Revision_Assistant/Revision_Assistant_for_Administrators/02Account_Types
- Warner, John. "Another Terrible Idea from Turnitin." *Inside Higher Ed*, 6 February 2018, <https://www.insidehighered.com/blogs/just-visiting/another-terrible-idea-turnitin>
- Watters, Audrey. "Education Technology's Completely Over." *Hackededucation.com*, 07 April 2017, <http://hackededucation.com/2017/04/07/prince>
- Zimmerman, Traci. "McLean Students File Suit Against Turnitin.com: Useful Tool or Instrument of Tyranny?" *The CCCC-IP Annual: Top Intellectual Property Developments of 2007*, <http://cccc.ncte.org/cccc/committees/ip/2007developments/mclean>

OTHER RESOURCES

- Moore, Cindy, et al. "Assessing for Learning in an Age of Comparability: Remembering the Importance of Context." *Reclaiming Accountability: Improving Writing Programs Through Accreditation and Large-Scale Assessments*, edited by Wendy Sharer et al., University Press of Colorado, Utah State University Press, 2016, pp. 17-35.
- Stommel, Jesse. "Who Controls Your Dissertation?" *Chronicle Vitae*, 7 January 2015, <https://chroniclevitae.com/news/852-who-controls-your-dissertation>
- University of Louisville, Department of English. "Policy Against the Use of Plagiarism Detection Software." March 2009, <https://louisville.edu/english/composition/policy-against-the-use-of-plagiarism-detection-software.html>
- Vie, Stephanie. "A Pedagogy of Resistance Against Plagiarism Detection Technologies." *Computers and Composition*, 30, 2013, pp. 3-15.
- Zeide, Elana. "Student Privacy Principles for the Age of Big Data: Moving Beyond FERPA and FIPPs." *Drexel Law Review*, vol. 8, 2016, pp. 339-394.