Report from the Working Group on Text-Matching Software for Detecting Plagiarism in Student Work

Presented to CLRC on 23 May 2013

Executive Summary
A working group of the Campus Law Review Committee (CLRC) examined the use of Text Matching Software (TMS) as a tool to detect plagiarism in student work. The group met over the 2012-13 academic year to study the issue and made the following recommendations:

1. The University-wide adoption of a particular TMS, either through purchase or internal development, should only occur if and when it can be demonstrated that such a system can be effective across a significant number of disciplines and that its use does not undermine other mechanisms for preventing plagiarism.

2. If a TMS is adopted or developed, other more effective measures to promote academic integrity and deter cheating and plagiarism should not be abandoned.

3. There is currently no need for a separate policy on TMS.

4. Regardless of whether or not the University adopts a particular TMS, guidelines for best practices in the use of TMS should be developed.

5. The University should, wherever possible, use existing resources to provide information on the use of TMS.

6. If a TMS is used, students must be notified in the course syllabus and provided with reasonable alternatives if they object to their materials being stored in a database.

7. Any instructor or unit who adopts a TMS is responsible for ensuring that its use complies with existing policies and laws.

Background

Rationale and History
In 2011, the report of the Academic Integrity Task Force, a group set up to advise the Office of Student Judicial Affairs and the Dean of Student on the state of academic integrity on campus, noted that a number of people had raised concerns about the use of text matching software to detect plagiarism at the University of Alberta. The groups based that recommendation on an extensive survey of students, teaching assistants, and instructors at the University of Alberta and on comments generated during focus groups. Some of the comments were observations that such resources were valuable tools in protecting academic integrity and urged their adoption. Others, however, expressed concern about the impact of the use of such software on students’ intellectual property rights and privacy and argued that the use of such software negatively affected classroom teaching and the relationship between teacher and student. The report also observed that there were already instances of units adopting such software but that there were no University policies or guidelines which governed its use. The Task Force recommended that General Faculties Council set up a working group to examine the use of
such software and to make recommendations to the community. That recommendation read as follows:

General Faculties Council should appoint a group to review electronic detection resources such as turnitin.com, identify their pedagogical strengths and pitfalls, and make recommendations to the community as to how they should be used if they are adopted. There is a growing interest in the use of electronic text-matching software and several units have already adopted some form of text matching software for use in detecting plagiarism. It seems inevitable that their use will become more pervasive in coming years and we need to be strategic in looking at how they are employed so as to minimise any unintended consequences to our students and classrooms.

Campus Law Review Committee, the GFC committee responsible for addressing student discipline issues, including plagiarism, established a working group to fulfill that recommendation and to gain a better understanding of digital plagiarism detection tools and their use at the University of Alberta.

**Principles**

- Protecting academic integrity is a vital interest and obligation of all of the members of the academic community
- We must protect the reputation of the university and of our faculty, staff, students and graduates
- Students have a right to expect that they will not be unreasonably disadvantaged in any disciplinary or pedagogical process
- Creating unnecessary burdens on instructors is a barrier to effective teaching

**Mandate**

Investigate and possibly make recommendations to Campus Law Review Committee regarding the use of text matching software for the purpose of detecting student plagiarism at the University of Alberta.

**Functions**

1. Review available tools and techniques for the use of text-matching software to detect plagiarism;
2. Identify the extent of their use at the University of Alberta;¹
3. Review benefits and pitfalls of using text-matching software tools; and
4. If appropriate, make recommendations to CLRC.

¹ Note, there is no central record of the use of TMS at the University of Alberta. While a simple Google search uses text-matching algorithms, our interest is in those TMS platforms that batch process student assignments and/or use a database for comparison. Since Individual professors or departments can subscribe to any of these services and are not required to report on their use, we could not realistically gauge the extent to which TMS systems are employed at the University.
Membership

- Chris Hackett, Office of Student Judicial Affairs, Chair
- Deborah Eerkes, Office of Student Judicial Affairs
- Jayson MacLean, Student OmbudService
- Dustin Chelem, Vice President Academic, University of Alberta Students’ Union
- Nathan Andrews, Vice President Academic, University of Alberta Graduate Students’ Association
- Elaine Geddes, Associate Dean, Faculty of Business
- Murtaza Jamaly, Undergraduate student at large
- Karon Dragon, Graduate student at large
- Andrew MacMillan, Faculty member, Department of Biochemistry
- Terry Nadasdi, Faculty member, Department of Linguistics
- Frank Peters, Faculty member, Department of Educational Policy Studies
- Walter Bischof, Faculty member, Department of Computing Science
- Gordie Mah, Information Technology Security Officer, Office of the Vice-Provost & Associate Vice-President (Information Technology)

What is Text-Matching Software?

Text-matching software is an electronic tool used for detecting plagiarism in student assignments or papers. It works by comparing electronically submitted assignments against documents posted on the Internet and/or a database of other assignments, and identifying when text from the student assignment matches text from other sources. It is often called plagiarism detection software; however, this is a misconception, as this kind of software does not distinguish between properly cited quotations and plagiarism. It is limited to identifying exact matches in text.

There is a broad variety of TMS, ranging from Google and other search engines to elaborate subscription-based services like Turnitin.com, databases built on site, and services like MOSS, which is used for comparing computer code. TMS can operate via a web browser, as a standalone program, or as a component of learning management systems. TMS can be used to check a single paper that has sparked suspicion for the instructor, a single paper in which the author wants to rule out unintentional or sloppy plagiarism in their own work, or all of the papers submitted for a class.

Furthermore, TMS can also check papers against one or more of the following:
1. Databases of previously submitted materials
2. Databases checking submitted assignments against each other
3. Databases of academic journals
4. The Internet via a search engine

Examples of TMS include:
1. Turnitin.com
2. SafeAssign
3. Wcopyfind
In addition, some schools, departments, and instructors have created their own TMS. Turnitin.com is the largest of the TMS and has contracts with a number of Canadian schools as well as institutions around the world. It compares papers against a database of previously submitted papers as well as searching the Internet. It can be licensed by an institution, a faculty, a department, or an individual instructor.

TMS systems are not without controversy, particularly when they are used to screen the work of an entire class and all of the submitted materials are retained in an off-site database, possibly outside of Canada. Some student groups and instructors have raised concerns about the implications of database supported TMS in terms of student privacy and intellectual property rights, judicial fairness, and the instructor-student relationship.

**Background information and research**

The Working Group reviewed:
- Scholarly articles on TMS
- Scholarly articles on academic integrity
- Newspaper and journal articles on TMS, related academic integrity issues, and issues surrounding the implementation of TMS at other Canadian institutions.
- Documentation on specific TMS
- University Policies at institutions that have adopted TMS

**Selected research summary**

1 Academic integrity should be approached through a variety of practices. The most effective methods also have the effect of reinforcing a positive instructor-student relationship. As one such practice, TMS might be a valuable tool in protecting academic integrity but it cannot be used alone.

2 Students plagiarise for a variety of reasons. Fifty three percent (53%) of Canadian undergraduate students reported having plagiarised at least once in an academic integrity survey at Canadian post-secondary institutions. (McCabe and Christiensen Hughes, 2006, p. 17) While we tend to focus on students who plan to plagiarise, we know that a significant number, likely a majority, plagiarise out of ignorance or panic. TMS may not deter the latter students but it will help identify them when they do. The literature on TMS as a deterrent to plagiarism is limited and does not provide clear evidence that TMS, on its own, deters cheating among the students who plagiarise because they are panicked. (Walker, 2010) Even where it is suggested that it is a deterrent, it is difficult to sort out how much of the reduction happened because the instructor engaged the students in a discussion of plagiarism and indicated that they
would follow through on suspected plagiarism and how much was solely the result of the TMS.

3 Some students who are deterred may switch to another class rather than not cheat at all. This is consistent with information provided by students in the Academic Integrity Survey.

4 Dalhousie terminated its contract with Turnitin and UBC restricted its use, both in 2011, over disagreements concerning the location of the databases used to store their students’ submissions as well as backups. The concern was about the degree of risk that students’ papers could potentially be accessed by the United States government for purposes other than academic integrity. Other people have raised concerns about the potential risk of the Government of Canada gaining access to such databases. Still others have objected to the requirement for students to submit their intellectual property to a database which may serve to benefit a for-profit company.

5 TMS systems struggle to identify paraphrased material – the more sophisticated the algorithm, the more likely such material will be identified but there are no systems that are infallible when students revise the text that is being copied. Consequently, TMS systems are most effective in detecting “cut and paste” plagiarism.

6 TMS will not detect cases where students purchase original papers from paper mills. We have seen a significant growth in the number of individuals and companies who offer to write assignments or papers for students for a fee. (Hager, 2012) These papers are often written with the expectation of an average grade, making them easier to produce on demand and less costly. Social media and such websites as kijiji.ca have made it easier for content creators and students intent on cheating to connect. The growth of these types of resources for cheaters reinforces the need to treat TMS as one part of a multi pronged approach to academic integrity. There is a concern that increased use of TMS may push students intent on plagiarising to use these services more widely.

7 Student groups and instructors at several universities have raised concerns about compelling students to contribute their intellectual property to a database owned by a third party. There have been no court decisions prohibiting it in Canada but critics have argued that this is a moral and pedagogical as well as a legal issue.

8 Institutions who have adopted TMS systems have typically created policies which govern their use. The Working Group is unaware of any institutions that have a TMS policy not related to a specific software package.

**TMS Review**

The Task Force looked at several of the most commonly used TMS in order to gain an understanding of how they worked. In addition, Walter Bischof demonstrated the use of Damocles, a TMS developed at Monash University that uses existing search engines to check for plagiarised material. Dr. Bischof corresponded with the author of Damocles who noted that
the costs of running queries in Google or some other major search engine, required for the software to function, had become prohibitive.

**Information from Ryerson and Dalhousie Universities**

The Working Group conducted a conference call with representatives from two universities that have licensed TMS. Ryerson University has had a license with Turnitin.com for 10 years and Dalhousie University has a license with Safe Assign, although it formerly had one with Turnitin.

The external participants were:

1. Bob Mann, Dalhousie University Manager of Appeals and Discipline, Dalhousie University
2. Diane Pirner, Professor of Nursing, Ryerson University
3. Donna Bell, Academic Integrity Coordinator, Ryerson University

**Key information and observations:**

1. Mann declined to discuss the reasons for terminating the Turnitin contract but indicated that they were happy with Safe Assign. He noted that they had switched to maintaining only a database of papers generated at Dalhousie so their students' papers were not being stored off site.

2. Bell indicated that Turnitin costs had gone up substantially over the years, particularly because of the inclusion of unwanted and unneeded features.

3. Both institutions indicated only a minority of instructors used TMS. Mann noted that it was used mainly in large classes in the Arts. Bell said that, despite the University's Turnitin license and the support offered for that program by the institution, some instructors used a free program called Viper.

4. Both institutions have a policy that governs how their respective TMS programs can be used. Ryerson includes a requirement that, if Turnitin is going to be used, the course syllabus must indicate that students will be required to submit their papers to Turnitin and there must be an opt out clause for students who do not wish to submit their work through the database. Bell indicated that if there was suspicion of plagiarism, an instructor could choose to run a paper through Turnitin even if there was no indication it was going to be used in the course or if a student had exercised their right to opt out. Mann said that Dalhousie had recently created a policy requiring an opt out provision in courses that use Safe Assign. Opt out requirements in both institutions were set by individual professors, and included such measures as requiring an additional oral component to an assignment, submitting an annotated bibliography with a paper or providing outlines and drafts along with the final version, among others.

5. Bell indicated that numbers of plagiarism cases did not decline over the ten years that Ryerson has been using Turnitin, but that they remained steady. She noted that it is not clear whether the number of students caught in "panic" or "ignorance" plagiarism
balanced out the number of students who withdrew from courses in order to avoid
detection, or whether the TMS has no deterrent value at all.

6 All indicated that it was vital to provide training for instructors specific to the institutional
TMS in order to interpret the results from the TMS. Despite that, both Mann and Bell
indicated that there were relatively few cases of suspected plagiarism that came forward
based solely on TMS results, rather the TMS results were used to trigger suspicions of
the paper or assignment and the instructor used her or his judgement to decide what
evidence supported a charge.

7 All indicated that the system worked best in identifying the reuse of papers previously
submitted at the same institution.

8 Pirner said that she had had many students tell her that they appreciated that Turnitin
protects them from having their material misused by other students.

9 Despite a campaign by the Students’ Union at Ryerson to encourage students to opt out
en masse, Pirner and Bell indicated that very few students have chosen to opt out of
Turnitin. Mann said that the Dalhousie Students Union had raised concerns about where
databases were stored.

Information provided by the Freedom of Information and Protection of Privacy Office

Dr. Harry Davis, Information and Privacy Officer, provided the following guidelines to the
committee for instances when instructors compelled students to run their work through a TMS
that stores copies of the students work in a database. Since these databases will inevitably
include personal information, an instructor employing a TMS must ensure that:

1 Students have clear information up front that the software will be employed.
2 Students are given clear information as to how their personal information will be used,
how it will be stored, and how long it will be retained.
3 Students in the course have a real and reasonable opportunity to opt out of having their
assignments and papers stored in the database.

Information Provided by the Office of General Counsel

Marie Strauss, Senior Legal Counsel in the Office of General Counsel met with the committee to
discuss legal implications of the use of text matching software. She noted that it was difficult to
give legal advice on a category of software, since they operate differently and therefore the legal
implications would not be the same. She indicated that there could be potential concerns about
intellectual property and privacy which an instructor might need to address.

She also noted that, in general, TMS is simply another tool which can be used to help detect
plagiarism and does not, in itself, interfere with Section 30.1.1(1)b, which promises the right “to
be presumed not to have committed an offence until his or her commission of an offence has
been established on a balance of probabilities, before an impartial and unbiased decision-
maker.”
Discussions of the Committee

The committee reviewed the above literature, interviews and information, and held discussions raising concerns from their own perspectives. The following questions guided the discussions:

1. Does the University need a policy on the use of TMS? Should it issue non-binding guidelines for the purposes of educating students and instructors? Should the University do nothing at all?
2. If a database is used by a TMS, should there be restrictions on storage?
3. How does the use of TMS affect instructors?
4. How does the use of TMS affect students?
5. How should any policy or guidelines offset any perceived problems created by the use of TMS?

Results of the committee discussions:

1. The investigation and deliberations over the use of TMS revealed a series of issues, both positive and negative, that should be addressed if either an instructor or the University as a whole were to adopt a TMS system.

2. A TMS that automates the process of searching the Internet and/or academic databases for an individual paper or assignment when a suspected case of plagiarism is identified can be a useful tool for instructors. A TMS system can help to identify sources of copied materials, potentially reducing the time that an instructor needs to invest in order to refer a case to the Faculty. At the same time, the focus of the process is still very much on the instructor using his or her judgment to determine whether or not material is plagiarised. The process is triggered by indications that suggest the paper is worth investigating for plagiarism. A TMS would identify text that was identical to other sources, which would necessitate further interpretation and investigation by the instructor.

3. The introduction of mandatory submission of papers to a TMS may bring benefits and also raise concerns that would need to be addressed:
   a. If students know that papers will be checked as a matter of course, it is reasonable to assume at least some of them will choose not to plagiarise in that course, although there is still no empirical evidence that is the case.
   b. The use of a TMS could be a deterrent to plagiarism, but that deterrent value might be limited to students who deliberately choose to plagiarise as opposed to students who plagiarise out of panic or ignorance and thus don't plan ahead. It is also likely that more of “panic” and “ignorance” plagiarists will be caught through the use of a TMS than by current systems alone. Students who are deterred from plagiarising in a course because it uses a TMS may choose not to plagiarise in that course, or they may simply transfer to another section or course in which they believe the instructor to be less vigilant.  

---

2 A common practice for students, according to student participants in the focus group as part of the 2011 University of Alberta Academic Integrity Survey.
c Over-reliance on TMS could create complacency on part of the instructor, which could lead to more opportunities for students to engage in plagiarism other than the “cut and paste” variety, or various other types of academic dishonesty.

d Those few students who believe they can beat the system, either because they have purchased an original paper through a paper mill or because they have altered the copied text sufficiently to “fool” a TMS will likely not be deterred and may or may not be detected.

e The mandatory submission of papers to a TMS may affect the instructor-student relationship by creating a climate in which students feel they have to prove their innocence. Using a TMS to check all papers in a course can be perceived as the assumption that all students are plagiarists unless the software establishes their innocence. That same perception is not generated by the use of, for example, a Google search conducted after suspicions are aroused about the use of a source. The perception of distrust that the use of TMS can imply can undermine the student-instructor relationship. If a TMS is to be used, it is critically important that it be framed as a measure to ensure the protection of honest students, rather than implying a presumption that all students will plagiarise.

f TMS can be used as a tool for teaching appropriate citation by having students review reports on drafts of their material before submission.

g Written notice that a TMS will be used should be included in the course syllabus so that students can make an informed decision about participating in the course, either by withdrawing from the course or knowing that they may have to exercise their right to opt out of the use of the TMS, where such exists. Such notice would have the added advantage of drawing attention to the TMS, which could increase its deterrence value.

h In institutions that do adopt a single TMS, there is a need to train instructors in how to use the software and how to effectively interpret the results. A TMS that provides scores, originality reports, or degrees of uniqueness only indicates how much of the material is original, not whether it has been cited correctly or not. A paper with a large number of quotations may receive a higher score than a paper that contains plagiarised material. This could produce false positives which, if not correctly understood by the instructor, could lead to students having to deal with the stress of defending themselves when called before the Faculty and can place unnecessary demands on the discipline process. If an instructor is using a TMS at the University of Alberta, the manner in which students are treated in the process can significantly affect the students’ well being and perception of fairness in the discipline process, even though instructors do not impose sanctions under the Code of Student Behaviour. It is vital that instructors treat students as innocent until they have been found to have committed an offence under the Code of Student Behaviour.

i The use of mandatory screenings could lead to a sense of complacency on the part of instructors, causing them to minimise or not use other measures to protect
academic integrity in a class, measures which the academic integrity literature would suggest have a greater deterrent value than TMS. Similarly, instructors need to have an understanding of ways that students can beat the TMS they are employing in order to guard against it. Again, the best way to counteract these problems is through training of instructors.

Increased demands for training for instructors raises the concern about sessional and other contract instructors who may not have a great deal of experience and who do not have a significant amount of time to dedicate to training. This is an ongoing issue for the University in terms of academic integrity but, in at least some instances, we may be increasing instructor workload through training demands by introducing TMS systems rather reducing it. Likewise, it may create an additional burden on University resources to provide training materials and/or sessions for all instructors.

A simple Google search employs some of the same methods as TMS. However, the complexities of using TMS increase when a database built on stored student papers is added to the mix. The major players in TMS, including Turnitin and Safe Assign, use databases which store papers previously submitted as part of the data set for future comparison.

The mandatory submission and retention of student papers means that students have to surrender intellectual property and personal information.

a. While we have no clear court decisions in Canada on the impact of the use of TMS on students' intellectual property rights, we do have a moral and ethical responsibility as an academic institution to treat such rights with respect.

b. We do have clearer direction on the use of personal information because of the Freedom of Information and Protection of Privacy Act (FOIPP). Any instructor using a database driven TMS must ensure that its use is FOIPP compliant. As such, students must be informed in the course syllabus that the software will be used, what information will be stored in the database, for what purpose(s) it will be used, where it will be stored and for how long. The information about databases must include backups or other copies of the databases.

c. There seems to be diminishing value in larger scale TMS databases. National and international databases may catch students who purchase papers from large scale paper mills that sell the paper more than once to different students. The trend in paper mills, however, seems to be towards local individuals or small companies that generate average quality papers on demand.

c. Even more common, however, is the practice of using portions of friends’ papers/assignments, or obtaining assignments submitted in previous years for the same course. The most significant value of a database system is to prevent the reuse of assignments from year to year or across sections of a course. As such, that database would be best scaled at a departmental or faculty scale.
Locally held databases would also be easier to manage the locations of the data and all backups as well as the purging of the data in accordance with the promises made to students when it was collected. Students must have a reasonable opportunity to opt out of the use of the TMS in a course which employs a database driven TMS. The working group was concerned that some of the opt out measures employed at Ryerson and Dalhousie were made intentionally onerous to discourage students from opting out, to prevent additional work for the instructor.

d Opt out provisions should not be onerous and should not require a student to prove their innocence. The opt out alternatives should reflect good practice in protecting academic integrity in classes. It would be useful to provide instructors with some possible examples of opt out alternatives.

e It is vital to be clear that opting out of the use of a TMS does not mean that students are allowed to opt out of having their work checked for plagiarism in a different fashion.

5 Further concerns and comments:

a There does not appear to be a need for a separate policy on TMS; current policies and laws, particularly the Code of Student Behaviour and FOIPP, are sufficient unless the University were to decide to adopt a single TMS.

b All participants should understand their legal and policy requirements when it comes to the use of TMS in University of Alberta courses. The University should ensure that information is readily available for faculty and students.

c TMS can be very expensive and it is important that anyone looking to adopt a software package understand those costs and explore alternatives that may be just as effective but at a lower cost.

d Many of the perceived benefits of TMS can be achieved in other ways without the expense and complexity of using a TMS. For example, the Working Group noted that changing assignments from year to year has a strong deterrent effect on reuse of previous assignments, as does engaging students in a discussion about plagiarism and its consequences.

e Given the variety of disciplines and academic cultures represented at the University of Alberta, it is highly unlikely that a single TMS could meet the needs of all instructors. Should a TMS be adopted by a department or instructor, any evaluation process should include legal and FOIPP considerations, and account for the University's information management, privacy, and security requirements.
Recommendations:

1. The University-wide adoption of a particular TMS should only occur if and when it can be demonstrated that such a system can be effective across a significant number of disciplines and that its use does not undermine other mechanisms for preventing plagiarism. Any decision to adopt a TMS system should take into account the concerns and issues identified in this report.

2. If a TMS is adopted, other more effective measures to promote academic integrity and deter cheating and plagiarism should not be abandoned. Research shows that a positive student-teacher relationship is a strong deterrent to academic dishonesty.

3. There is no need for a separate policy on TMS unless the University adopts a single platform. Otherwise, existing laws and policies cover the major issues surrounding the use of TMS.

4. Regardless of whether or not the University adopts a particular TMS, guidelines for best practices in the use of TMS, based on the findings of the Working Group, should be developed. The guidelines should draw attention to existing policies and legal concerns and be clear about acceptable and effective use of TMS. The guidelines should also include examples of reasonable alternatives to be used in the event a student elects to opt out where that is a possibility.

5. Existing teaching and academic integrity resources should be updated to include information and guidance on the use of TMS and a reference to the guidelines. Some examples of such resources might include materials provided by Faculties for instructors, the Academic Integrity Handbook for Instructors and TAs, and the TIE website.

6. If a course involves the use of TMS, students must be duly notified in the course syllabus and they must be provided with reasonable alternatives if and when they object to having their course work stored in a TMS database.

7. Any instructor or unit who adopts a TMS is responsible for ensuring that its use is in compliance with existing policies and laws.
Resources reviewed by the Working Group on the use of Text Matching Software, 2012/2013

**University of Alberta Documents:**

Code of Student Behaviour, University of Alberta.

**Research:**


**Websites:**
SafeAssign by Blackboard: [http://safeassign.com](http://safeassign.com)

Turn It In: [http://turnitin.com](http://turnitin.com)

**Other Institutions:**
University of British Columbia: [http://vpacademic.ubc.ca/integrity/turnitin-at-ubc](http://vpacademic.ubc.ca/integrity/turnitin-at-ubc)

University of Toronto: [http://www.teaching.utoronto.ca/teaching/academicintegrity/turnitin/e-resource.htm](http://www.teaching.utoronto.ca/teaching/academicintegrity/turnitin/e-resource.htm)

Dalhousie University: [http://www.dal.ca/dept/university_secretariat/academic-integrity/plagiarism-cheating/plagiarism-detection-resource.html](http://www.dal.ca/dept/university_secretariat/academic-integrity/plagiarism-cheating/plagiarism-detection-resource.html)

**Media:**


Parry, Marc. “Software catches (and also helps) young plagiarists,” the Chronicle of Higher Education, November 6, 2011.
Text Matching Software Working Group Meeting Schedule and Requests for Feedback

Meeting Schedule
28 November 2012
11 December 2012
21 January 2012
6 February 2013
25 February 2013
25 March 2013

Requests for Electronic Feedback
20 February 2013
16 April 2013
29 April 2013
2 May 2013