جامعة الملك سعود
كلية الآداب
مجلة الآداب
ردمد (ورقي): ٣٦٢٠-١٠١٨
ردمد (النشر الإلكتروني): ٨٣٣٩-١٦٥٨

King Saud University
College of Arts
Journal of Arts
ISSN (Paper):1018-3620
ISSN (Electronic):1658-8339

مجلة الآداب، مج (٣٤)، ع (٣)، ص ص ١٢١-١٤١، جامعة الملك سعود، الرياض (٢٠٢٢م/ ١٤٤٣هـ)
Journal of Arts, Vol. 34 (3), pp 121-141, © King Saud University, Riyadh (2022 /1443H.)

أنظمة الكشف عن الانتحال العلمي: من اللغويات إلى الأخلاقيات

يحيى الظلمي

أستاذ اللسانيات المساعد، قسم اللسانيات ودراسات الترجمة، كلية اللغات والترجمة، جامعة الملك سعود.

(قدم للنشر في ١٦/ ٥/ ١٤٤٣هـ، وقبل للنشر في ٢٨/ ٧/ ١٤٤٣هـ)

الكلمات المفتاحية: أنظمة الكشف عن الانتحال العلمي، مصادر المعلومات الأجنبية، اللغويات، المؤسسات التعليمية، الأخلاقيات.

ملخص البحث: بصرف النظر عن مقدار التعقيد الخوارزمي لأنظمة الكشف عن الانتحال العلمي، ولكون الذكاء البشري ما يزال متفوقًا على الذكاء الاصطناعي، فاحتمالية إخفاق تلك الأنظمة في التقاط الأعمال المنتحلة واردة جدًا. فالكتّاب الأذكياء لغويًا –وبالتالي مرتكبو الانتحال العلمي الذكي– يستطيعون تحويل قطعة لغوية ما إلى قطعة أخرى تبدو بمنأى عن الأولى، وتختلف عنها اختلافًا جوهريًا، في حين تظل الأفكار في الاثنتين متطابقة. وعلى نحو مماثل، فإن الانتحال العلمي بالاعتماد على مصادر المعلومات الأجنبية –وبالتالي الانتحال البين–لغوي– عادة ما يتسبب في جعل تلك الأنظمة عاجزة عن التعرف على حالات الانتحال العلمي، لا سيّما إذا كانت اللغتان متمايزتين كليًا. ولهذين السببين الأساسيين، تحاجّ هذه الورقة –باستعمال عينات من نتائج اللغويات الجنائية وبالاستناد إلى مبادئ الأخلاقيات ونظرياتها– بأن أنظمة الكشف عن الانتحال العلمي قد تثير بعض المخاوف الأخلاقية المتعلقة بالمساواة، والعدالة، والموثوقية، وأخيرًا الإثارة. على وجه التحديد، تمنح تلك الأنظمة الفضل للمنتحلين النمطيين الذين يستعملون الذكاء اللغوي، وتمنح امتيازَ الثنائية اللغوية لمرتكبي الانتحال العلمي اللغوي البين الذين يلجؤون لمصادر معلومات أجنبية، ليتمكن كلا الفريقين من مخاتلة الأنظمة عمدًا. وتعرض الورقة لبعض الجوانب التي لا بد من أخذها في الحسبان عند التعامل مع أنظمة الكشف عن الانتحال العلمي، وتختم بتوصية بإجراء دراسة مسحية مستقلة للمنتحلين الذين سبق وتعمدوا الانتحال العلمي بطريقة يعتقدون بأنها ستفلت من تلك الأنظمة ونجحوا في ذلك لدعم تلك المخاوف الأخلاقية (أو دحضها).

Plagiarism Detection Systems: From Linguistics to Ethics

Yahya Aldholmi

Assistant Professor of Linguistics, Department of Linguistics & Translation Studies, College of Languages & Translation, King Saud University.

(Received:16/ 5/1443 H, Accepted for publication 28/ 7/1443 H)

Keywords: plagiarism; plagiarism detection systems; foreign information sources; linguistics; instructional institutions; ethics.

Abstract. Regardless of how much algorithmic complexity a Plagiarism Detection System (PDS) incorporates, artificial intelligence can still be outsmarted by human intelligence, and PDSs are thus likely to fail to capture some plagiarized works. Linguistically intelligent writers (and hence intelligent plagiarists) can turn a given passage into another one that sounds substantially both remote and different from the original even though the ideas remain identical in both texts. Likewise, plagiarizing from foreign information sources (i.e., cross-language plagiarism) often renders any PDS incapable of identifying plagiarism instances, especially if the two languages are highly distant. For these two major reasons, this paper, using samples of forensic linguistics findings and several ethical principles and approaches, argues that PDSs may raise some ethical concerns related to equality, justice, universalizability, reliability, and, ultimately, altruism. Specifically, PDSs give credit to prototypical plagiarists who use linguistic intelligence and give bilingual privilege to translingual plagiarists who intentionally use foreign information sources to evade or trick the system. The paper articulates some facets that must be considered when dealing with PDSs and concludes with a recommendation that an empirical study to survey plagiarists who have intentionally tricked a PDS and escaped the consequences be conducted to illustrate (or refute) those ethical concerns.

مقدمة

قد لا يسوغ في عموم الحالات تصنيف تقنية ما على أنها أخلاقية أو غير أخلاقية بحد ذاتها (Bynum & Rogerson, 2008)، فالتقنية في الأصل محايدة، وإنما يحدد القرار الأخلاقي استعمال تلك التقنية أو توجيهها سواء من قِبل المطورين أم من قِبل المستخدمين. وبناء على ما توصلت له بعض الدراسات التي استعملت أساليب لغوية من نتائج لكشف الانتحال العلمي ومقارنتها بنتائج الكشف عنه بواسطة أنظمة آلية (plagiarism detection systems) خصوصًا حين يكون الانتحال العلمي معتمدًا على بعض مصادر المعلومات الأجنبية، فإنه يبدو أن هناك ارتباطًا بين حالات تستعمل فيها أنظمة الكشف عن الانتحال العلمي وحالات من المخاوف الأخلاقية. ويتأكد ذلك إذا علمنا أن التسمية الشائعة لهذه الأنظمة بوصفها أنظمة للكشف عن الانتحال العلمي قد تجعل مستخدمي تلك الأنظمة يعتقدون أن النتائج التي تقدمها تلك الأنظمة هي فعليًا تقريرية تثبت الانتحال العلمي أو تبرئيّة تنفيه، في حين أنها في الواقع تقدم نتائجَ لمستويات التشابه الذي قد يقتصر أحيانًا على مستويات لغوية معينة كالتشابه المعجمي أو التركيبي، وهو ما تشير إليه بعض الأعمال المتعلقة بالانتحال العلمي (على سبيل المثال، Eaton, 2021).

لقد أدرجت أنظمة الكشف عن الانتحال العلمي في الوسط الأكاديمي في وقت متأخر نسبيًا مقارنة بغيرها من الأنظمة، وذلك بغرض تعزيز النزاهة العلمية، ولمناوأة عدم الأمانة الأكاديمية (Brin, Davis, & Garcia-Molina, 1995, June). ولا شك في أن الانتحال العلمي غير مقبول أخلاقيًا، وأن اتخاذ إجراءات أكاديمية لتوعية الكتّاب عنه ومنعه والمعاقبة عليه هو أمر ضروري، غير أنه ينبغي ألا يكون الوسط الأكاديمي مدفوعًا بالجوانب المشرقة لجعل الساحة الأكاديمية أكثر صدقًا وأمانة وأخلاقيّة إلى الحد الذي يتجاهل فيه المجتمع الأكاديمي ما ترتبط به تلك الأنظمة من مخاوف أخلاقية، ويقلل من شأن ما توصل إليه البحث اللغوي حول أوجه قصور تلك الأنظمة. قد نميل إلى الثقة كثيرًا بالتقنية، ونتقبل ما تقدمه لنا تلك الخوارزميات على أنه نتيجة مسلمة، وهذا قد يقبل في حالة دون أخرى بناء على المدخلات والمخرجات، لا سيما حين يكون الفحص البشري هو الأكثر موثوقية بسبب طبيعة المدخلات، كما هو الحال مع اللغة البشرية. فهي (أي: اللغة) الأداة التي يستعملها المنتحلون حين ينتهكون الملكية الفكرية، وهي نفسها الأداة التي عن طريقها يحصل التحقيق في حالات الانتحال العلمي. إن اللغة هي ببساطة أداة الجريمة وأداة الفصل فيها، ليس فقط في حالة الانتحال العلمي، بل في كثير من الحالات التي تدرس في اللغويات الجنائية كالخلاف حول ملكية العلامات التجارية، والتشابه اللغوي بينها، والتحايل الصوتي والكتابي؛ لمحاكاة العلامات التجارية. ويأتي السؤال الجوهري: هل أظهرت التقنيات أداء كافيًا وموثوقًا به في التعامل مع لغة البشر؟ لا شك أن التقنية قد أحرزت تقدمًا كبيرًا وملموسًا في عدة جوانب متعلقة باللغة البشرية كرقمنة اللغة (digitizing)، وكالتعرف على الخطاب (speech recognition)، وفي إنتاج اللغة البشرية أو تخليقها (speech synthesis)، ولكن في الواقع ما زالت التقنية تحتاج إلى الكثير من الجهود والعقود لتقوم بتحليل اللغة البشرية بدقة خاصة في الجانب الدلالي. ولهذا، فإن السؤال الأكثر تحديدًا هنا هو: هل يمكننا الثقة بأنظمة الكشف عن الانتحال العلمي في تحليل اللغة البشرية لاتخاذ قرار حول الاشتباه في وجود انتحال علمي والفصل في كون المشتبه به قد انتحل أو لم ينتحل؟

إن الإجابة المباشرة التي قد يهرع إليها البعض هي أن هذه الأنظمة لا تستعمل من أجل الحكم على الذين يشتبه بهم في حالات الانتحال العلمي وتقرير ما إذا كانوا مذنبين أو لا، وإنما تستعمل (أو يفترض أن تستعمل) على أنها أداة مساعدة أو مؤشر أولي يلي استعماله كثير من التحقيقات البشرية، كما

ترد الإشارة إليه في بعض الدراسات (Cheong, Lau, & King, 2013). إن هذه الإجابة مقنعة وكافية لو كانت تلك الأنظمة قادرة تمامًا على اكتشاف حالات الانتحال العلمي أو حالات التشابه بصورة مطردة ومتكافئة ومتوازية. بمعنى آخر، إذا كان لكل انتحال علمي يتم اختباره بواسطة هذه الأنظمة فرص متكافئة لاكتشافه أو لعدم اكتشافه، فعلى المجتمع الأكاديمي القبول بهذه التقنية بوصفها مساعدًا على الكشف عن الانتحال العلمي طالما أن المنتحلين يتساوون في كشفهم من عدمه بصرف النظر عن كفاءتهم اللغوية ومصادرهم المعلوماتية. ومع ذلك، فقد أظهرت عدة دراسات لغوية في المقام الأول وأخرى غير لغوية (على سبيل المثال، Alzahrani, Salim, Abraham, 2012; Barrón-Cedeno, Rosso, Agirre, Labaka, 2010, August; Klopper, 2009; Mozgovoy, Kakkonen, Cosma, 2010; Sousa-Silva, 2014a) أن تلك الأنظمة تخفق في الكشف عن الانتحال العلمي في حالات كثيرة كالنصوص التي يستعمل فيها منتحلوها الذكاء اللغوي للتلاعب بها أو تلك النصوص التي اعتمد منتهكوها على مصادر معلومات أجنبية. إذا علمنا أن الاستعمال غير المناسب للمصادر إما نمطي يكون التشابه فيه بين النص الآخذ والنص المأخوذ منه مباشرًا، أو غير نمطي –وهو ما يسمى (patchwriting) "الكتابة الترقيعية"– يكون فيه صوت الكاتب المأخوذ عنه متداخلًا مع صوت الآخذ ومعتّمًا به، فإن من السهل نسبيًا اكتشاف النوع الأول. وعلى النقيض، فإن النمط الثاني يمثل تحديًا لأنظمة الكشف عن الانتحال العلمي، ويزداد الأمر تعقيدًا وصعوبة حين يعتمد الكاتب على مصادر معلومات أجنبية، فتكون لغة النص الأصل غير لغة النص الجديد، لا سيما حين تنحدر كل لغة من شجرة عائلية لغوية مختلفة، فتزيد الفجوة بين النص المصدر والنص الجديد جرّاء الاختلافات اللغوية التركيبية والدلالية. ومثل ذلك حين تكون إحدى اللغتين غير مدعومة تقنيًا مقارنة باللغة الأخرى، كما هو الحال على سبيل المثال في العربية والإنجليزية على التوالي.

ومن ثم، فإن هذه الورقة تسعى لتحاجج –من منظور أخلاقي متعمد على أدلة لغوية ومعلوماتية– بأن استعمال هذه أنظمة؛ للكشف عن الانتحال العلمي قد يكون مرتبطًا ببعض المخاوف الأخلاقية؛ وذلك لأن الأدلة كافية على أن تلك الأنظمة يتعذر عليها الكشف عن الانتحال العلمي في حالات كثيرة يقوم فيها الكاتب بخداع النظام قصدًا وينجح في ذلك. بمعنى آخر، فإن هذه الأنظمة في نسختها الحالية وبإمكانياتها الآنية أكثر قدرة على الكشف عن الانتحال العلمي لدى الطلبة والباحثين الذين لا يتمتعون بمهارات لغوية عالية أو ليس لديهم امتياز الثنائية أو التعددية اللغوية التي تمكنهم من الوصول إلى بعض المصادر المعلوماتية الأجنبية، وهاتان فئتان من المنتحلين لا تشكلان معضلة كبرى مقارنة بنقيضيهما. إن العقوبات التي تطال أولئك البسطاء الذين ينتحلون لا عن قصد ولكن لضعف تدريبهم، أو لعدم كفاءة مهاراتهم اللغوية، أو لأجل التقاليد والأعراف الكتابية المختلفة في البيئة التي يدرسون بها عن تلك التي نشأوا فيها، أو لعدم الفهم الصحيح لماهية الانتحال العلمي هي عقوبات عادة ما يُفلت منها المنتحلون البارعون الذين لديهم أدوات لغوية خاصة، ومن ضمنها الوصول لمصادر معلومات أجنبية يتعذر على أنظمة الكشف عن الانتحال العلمي كشف أصحابها. إن الأكاديميين من أساتذة ولجان تأديبية ممن يدركون كيف تعمل تلك الأنظمة يعون تمامًا أنه لا بد وأن يكون من بين الطلبة والباحثين من تعمدوا الانتحال، ثم استطاعوا أن يخاتلوا أنظمة الكشف عن الانتحال ويكسبوا الفضل –من درجات ونحوها– بما قدموه من أعمال منتحلة أو تحتوي على أجزاء منتحلة. وهذا بحد ذاته يطرح أسئلة أخلاقية متعلقة بالإنصاف، والمساواة، والموثوقية، والصدق، وهي معايير حاسمة ومهمة لدى الأخلاقيين في إصدار الأحكام الأخلاقية أو تفنيدها.

ولطالما ركزت الدراسات السابقة على مناقشة أخلاقيات المعلومات والعواقب الأخلاقية للانتحال العلمي في لغات شتى خصوصًا في الإنجليزية، ولكن قلما تناول البحث العلمي الجانب الأخلاقي لأنظمة الكشف عن الانتحال العلمي، واقتصرت بعض الدراسات اللغوية القليلة على مقارنة أداء تلك الأنظمة بالتحليل اللغوي البحت وتبيان إخفاق الأنظمة في الكشف عن الانتحال العلمي. بل ربما خشي بعض الباحثين ردة فعل المجتمع الأكاديمي لو أثار بعض التساؤلات حول أخلاقيات تلك الأنظمة على نحو جادّ؛ لأنه قد يبدو وكأنه يناوئ واحدًا من الإجراءات المتخذة لتعزيز النزاهة العلمية وتوحيد الحلول العالمية. إن هناك إشارات متناثرة في الدراسات السابقة كتلك التي نجدها عند بعض اللغويين مثل بكراري (Pecorari, 2015) إلى وجوب الحذر عند التعامل مع أنظمة الكشف عن الانتحال العلمي، وضرورة فهم ما أُعدت له وما ينبغي أن تستعمل له، إلا أن معظم الأعمال السابقة تقف عند الحد الذي تبين فيه محدودية أداء تلك الأنظمة في حالات معينة. ولذلك، فإن دراسة تجمع بين تلك الجوانب في مكان واحد لهدف واحد، ستكون تأسيسًا مهمًا للبحث العلمي اللغوي والأخلاقي المتعلق بتلك الأنظمة. ومن هنا -كما سبقت الإشارة إليه آنفًا- فإن هذه الدراسة تطرح سؤالًا حول تحقيق تلك الأنظمة لمعايير الأخلاقيات متّبعة منهج المحاججة الأخلاقية النظري الذي يضع المسلمة والنتيجة، ويستمد أدلته مما توصلت له دراسات لغوية ومعلوماتية سابقة كانت قد انتهت إلى ما يثبت إخفاق أنظمة الكشف عن الانتحال العلمي، وتحديدًا في حالات يكون فيها للمنتحلين أدوات لغوية خاصة أو تنفذ إلى مصادر معلومات أجنبية يقع عن طريقها الانتحال بالترجمة.

وتمهيدًا للشروع في مقاربة الهدف من هذه العمل، فإن الجزء التالي سيقدم مراجعة أدبية تلقي الضوء على مفهوم الانتحال العلمي بتعقيداته الدلالية والتطبيقية، وعلى جوانب من أنظمة الكشف عن الانتحال العلمي، ثم على اللغويات الجنائية والانتحال العلمي. يلي ذلك في الجزء الثاني عرض لما توصلت له الدراسات حول أنظمة الكشف عن الانتحال العلمي والأدوات اللغوية المستعملة والتحديات التي تواجهها لكشف الانتحال العلمي المعتمد على الذكاء اللغوي أو على مصادر معلومات أجنبية. وفي الجزء الأخير وبعد أن تكون الورقة قد هيّأت الأساس الذي تبنى عليه هذه المحاججة، تنتقل إلى الهدف الأساسي، وهو مناقشة المخاوف الأخلاقية اعتمادًا على المعايير الأخلاقية كما هي متناولة في نقاش أخلاقيات غيرها من الممارسات والتقنيات، واستنادًا على ما أنتجته الدراسات التي نوقشت في الجزء الذي يسبق.

مراجعة أدبية

جرت العادة أن يدرس الانتحال العلمي والملكية الفكرية في كثير من اللغات(١) من جانب معلوماتي، أي لدى متخصصي علم المعلومات، وأحيانًا من جانب اجتماعي لدى متخصصي علم الاجتماع، ولكنه منذ عهد قريب أصبح موضوعًا مشتركًا بين عدة مجالات منها مجال اللغويات (Pecorari, 2015)، وأصبح البحث اللغوي في الانتحال العلمي مؤخرًا يُربط بأنظمة الكشف عنه، والمقارنة بين الكشف الآلي والكشف البشري المبني على التحليل اللغوي المجرد. ولا شك أن ثمة علاقة منطقية بين اللغويات وأنظمة

(١) أشعر بالحاجة إلى توضيح غياب المصادر العربية عن هذا البحث لاسيما وقد اقترح أحد محكمي هذا العمل -مشكورًا- محاولة تضمين بعض المراجع العربية. لقد اطلعت على أعمال متفاوتة في العربية، ولم أجد -فيما اطلعت عليه- ما يتصل اتصالًا مباشرًا بهذا البحث. إن معظم الأدبيات الموجودة تعرّف بالانتحال العلمي ومصطلحاته وأنماطه نقلًا عن بعض المصادر الأجنبية التي استشهدتُ بها هنا أو ترجمةً لمحتواها، وليس من الناجع العدول عن المصادر الأولية إلى مراجع ثانوية أو ترجمات فقط؛ لأنها باللغة العربية. أما بعض الأدبيات الأخرى فتناقش أسباب الانتحال العلمي، وتحذر من عواقبه، وتعرض قوانين الملكية الفكرية، وتدعو لأخلاقيات البحث العلمي... وهلم جرا. وكل هذا ليس محور البحث الحالي، وسيكون من "التكثّر" استجلاب بعض تلك المراجع لهذا الغرض فحسب.

الكشف عن الانتحال العلمي؛ فتلك الأنظمة تستعمل الوحدات اللغوية كالعنصر المعجمي (lexical items) وتعتمد على تجزئة التراكيب المختلفة كحدود المكونات اللغوية (constituent boundaries) بغرض تقييم التشابه بين نصين أحدهما نص جديد والآخر نص قديم مصدري (source text) متاح للنظام. ويضاف إلى ذلك أن اللغويات الجنائية في أحد فروعها تبحث في قضايا ذات صلة كالتأليفية (authorship) سواء أكانت النصوص نصوصًا جنائية أو لا، وكذلك في قضايا تتصل بالكشف عن الانتحال العلمي وتقييمه (انظر على سبيل المثال، Coulthard, Johnson, & Wright, 2016; Gibbons, 2003)، وكل ذلك يحصل إما بالفحص اللغوي المستقل من خبراء لغويين وإما بمساعدة بعض التقنيات المتاحة كأنظمة الكشف عن الانتحال العلمي. وبما أن لهذين الجانبين علاقة بينية فإنه من الضروري قبل الانتقال إلى لبّ هذا العمل أن يُلقى الضوء على جانب من أدبياتها. ولذا، فإن هذه الورقة ستتناول -فيما يأتي أدناه- مفهوم الانتحال العلمي وأنماطه باقتضاب، ثم أنظمة الكشف عن الانتحال العلمي، ثم أخيرًا، لغويات الانتحال العلمي.

مفهوم الانتحال العلمي وأنماطه

يوصف الانتحال العلمي بأنه أخذ عمل من آخرَ ثم تمريره على أنه لآخذه (Pecorari, 2015)، وذلك العمل قد يكون فكرة أو لغة أو كلاهما. وقد يبدو الانتحال العلمي مفهومًا واضحًا في بعض الحالات، ولكنه غامض في حالات أخرى، كما يعتقد البعض بأنه مفهوم مستقر المعنى بالنسبة لهم -وإن كان ذلك مجرد اعتقاد- في حين يعتقد البعض أنه مفهوم غير مستقر. فالانتحال العلمي مفهوم يتغير بتغير الحالات والأشخاص والسياقات، وهذا يعني أنه ليس من السهل دائمًا تحديد ما إذا كانت ممارسة أحد الباحثين أو الطلبة تقع في نطاق الانتحال العلمي أو تقع خارج حدوده. بل قد يقع الانتحال فعليًا، ولكن يكون لدى أصحاب الشأن فيه تصورات مختلفة؛ فالمنتحل قد يرتكب الانتحال ولديه تصوره الخاص لما يعد انتحالًا علميًا، وما لا يعد كذلك، في حين يكون لصانعي السياسات والأساتذة تصورات مختلفة عن تلك التي لدى مرتكب الانتحال العلمي، ويتأكد ذلك حين ينتمي الطرفان لثقافات مختلفة وأنظمة تعليمية متباينة أو حتى حينما تختلف مستويات السلطة (كالطلبة مقارنة باللجان التأديبية). وهكذا فإنه يمكن أن يعرّف الانتحال العلمي ثم يفسّر التعريف بطرق متفاوتة، لا سيّما مع تقدم التقنية وإتاحة الكم الهائل من المعلومات اللذين أتاحا للانتحال العلمي التقليدي أن يتطور إلى نوع جديد يعرف علميًا باسم (digital plagiarism) "الانتحال الرقمي" (Ercegovac & Richardson, 2004) أو (cyber cheating) "الغش السيبراني" (Jones, Reid, & Bartlett, 2008). والمفارقة في هذا أن التقنية التي تساعدنا في الكشف عن الانتحال العلمي هي نفسها الوسيلة التي تساعد المنتحلين على ارتكابها (على سبيل المثال، Coulthard & Johnson, 2007)؛ لأن التقنية تقدم للكتّاب تسهيلات لاقتراض أعمال الآخرين أخلاقيًا أو انتحالها لا أخلاقيًا من مصادر يتيسر الوصول إليها على الشبكة العنكبوتية مقارنة بما كانت عليه الحال حين كانت الإفادة من نصوص المصادر مقتصرة على المواد المطبوعة (على سبيل المثال، Jones, Reid, & Bartlett, 2008).

ويمكن أن يقع الانتحال العلمي عن طريق عملية النسخ واللصق من مصادر على الشبكة العنكبوتية، أو من أي مستند رقمي أو عن طريقة استنساخ بعض المصادر المطبوعة، وهو نوع من الانتحال يدعى (verbatim) "النسخ الحرفي" (على سبيل المثال، Alzahrani, Salim, & Abraham, 2012; Mozgovoy, Kakkonen, & Cosma, 2010). النوع الآخر من الانتحال العلمي يقع حين يقضي الكاتب وقتًا لتوظيف حيل متنوعة لمواراة الانتحال العلمي، مستعملًا استراتيجية تدعى (obfuscating) "تعتيم النص" يتلاعب فيها بلغة

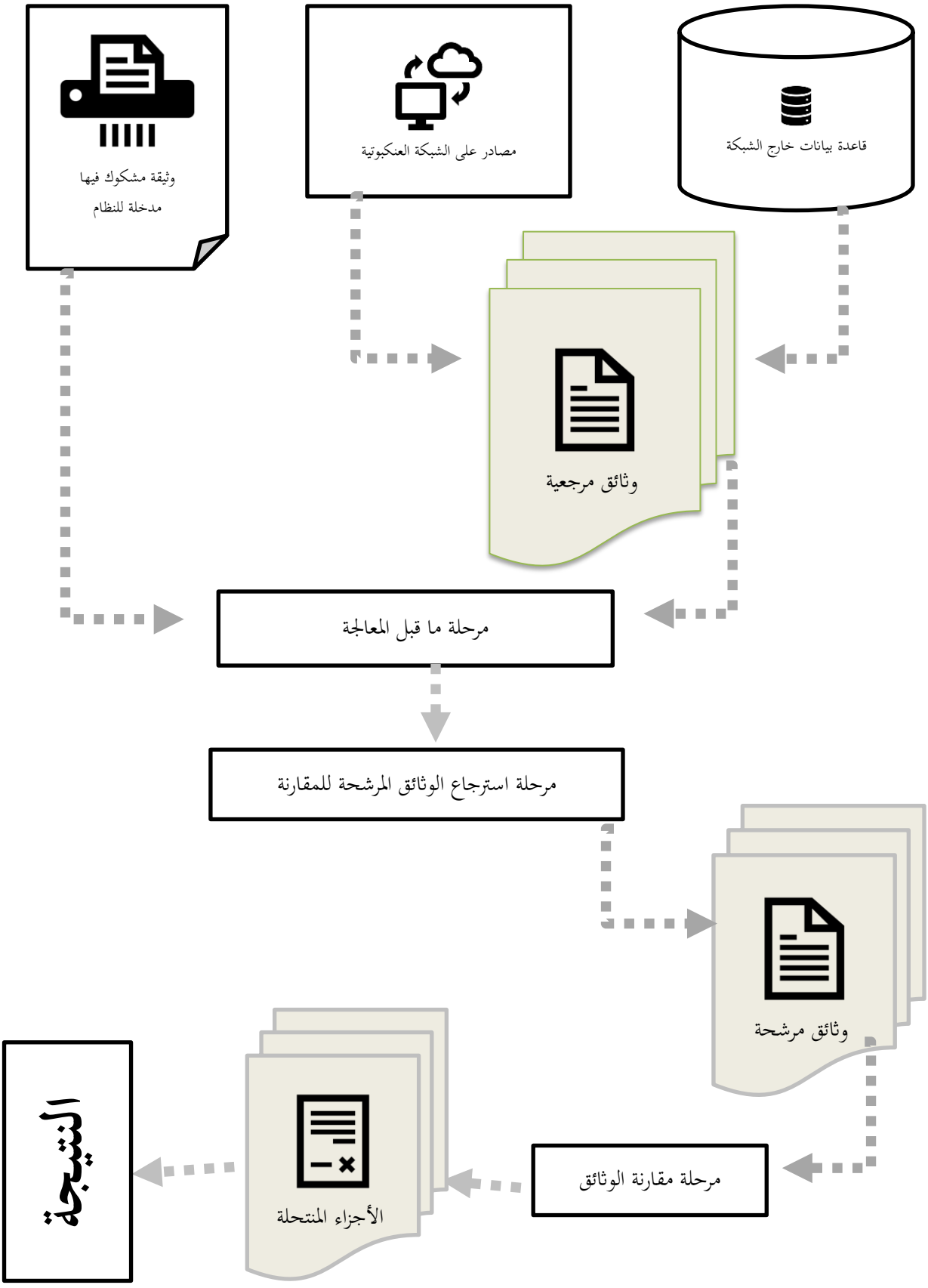
النص وأسلوبه؛ فيستبدل بعض العناصر المعجمية بأخرى، ويفكك البنية الأساسية للنص ثم يعيد هيكلتها، ومن ثم يقدّم المعلومات الدلالية ويُنزل عليها علامات الترقيم بما تقتضيه في صورتها الجديدة (Pecorari, 2015). أما النوع الثالث الذي يبدو عصيًا على الاكتشاف البشري والآلي فهو ما يطلق عليه (cross-language plagiarism) "الانتحال العلمي البين-لغوي" (Barrón-Cedeno, Rosso, Agirre, and Labaka, 2010)، ويقع بنقل محتوى نصٍّ في لغة أجنبية إلى لغة المصدر دون الإشارة إلى مؤلفه، ويتأكد ذلك حين لا يكون بين اللغتين تفاهم متبادل (mutual understanding)، ويزداد الأمر صعوبة حين تكون اللغتان منتميتين إلى عائلتين مختلفتين كعائلة اللغات الآفرو-آسيوية وعائلة اللغات الهندو-أوروبية. وقد يلجأ بعض المنتحلين إلى حِيل تقنية كإدخال النصوص في صيغة صورة داخل المستندات؛ مما يجعل الأنظمة التي تعتمد على قراءة النصوص غير قادرة على معالجة الصورة على أنها محتوى نصيّ، وكذلك استعمال بعض الروابط التي لم تعد صالحة (Mozgovoy, Kakkonen, & Cosma, 2010) إلا أن هذا النوع خارج نطاق هذا البحث.

ومع هذا التسارع التقني والضخ المعلوماتي أصبح الوسط الأكاديمي مضطرًا إلى توظيف الأدوات التقنية لمدافعة الانتحال العلمي ورفع مستوى النزاهة الأكاديمية، وبالتالي، تحسين جودة التعلم والتعليم وجودة البحث العلمي. وتعد أنظمة الكشف عن الانتحال العلمي نموذجًا مثاليًا لتوظيف التقنية في هذا المجال، وهذا ما سيتطرق له الجزء الثاني من هذه المراجعة الأدبية.

أنظمة الكشف عن الانتحال العلمي

يحال إلى أنظمة الكشف عن الانتحال العلمي أحيانًا تحت مصطلح (text-machines) "آلات النص" والمصطلحان يتناوبان في أدبيات البحث العلمي، غير أن هناك تفضيلات لدى بعض العاملين في الحقل المهني وبعض العاملين في الحقل الأكاديمي لمصطلحات دون أخرى، ولكنها في نهاية المطاف أنظمة تؤدي وظيفة واحدة أو متشابهة على الأقل. أنظمة الكشف عن الانتحال العلمي هي برمجيات تقارن الوثيقة المدخلة أو المشكوك فيها بمصادر ذات صلة بغرض تحديد مواضع التشابه ومستوياته تمهيدًا للكشف عن الانتحال وتقييمه (Lancaster & Culwin, 2005). وكثير من تلك الأنظمة هي أنظمة محددة اللغة (language-specific)، وبطبيعة الحال هي أيضًا (extrinsic) أنظمة خارجية (Alzahrani, Salim, & Abraham, 2012)، مما يجعل القصور في الطبيعة الداخلية لها (intrinsic) حين تواجه التعددية اللغوية (multilingualism) عاملًا مؤثرًا في إخفاقها في الكشف عن حالات الانتحال التي ليست ضمن لغة الوثائق المصدرية. وقد يكون الكشف عن الانتحال بالمقارنة مع بعض الوثائق المحلية خارج الشبكة العنكبوتية، وهو ما يسمى بـــــ (hermetic detection) "الكشف الحصيف"، أو مع بعض الوثائق المتاحة على الشبكة العنكبوتية، وهو ما يسمى بـــــ (web detection) "الكشف عن طريق الويب" (Mozgovoy, Kakkonen, & Cosma, 2010). كما يوضح الشكل ذي الرقم ١، يحاول النظام في الكشف عن الانتحال العلمي عن طريق الويب العثور على أوجه التشابه أو التطابق بين المستند الذي يتم التحقق منه، وبعض المصادر الأخرى المتاحة على الشبكة العنكبوتية التي يمكن للنظام الوصول لها. أما في الكشف الحصيف، فإن النظام يبحث عن أوجه التشابه في مجموعة داخلية من المصادر في قاعدة بيانات تديرها تلك المنظمة أو المؤسسة التعليمية. ومع وجود التفاوت بين النوعين، فإن ثمة بعض الأنظمة القليلة التي تتمتع بإمكانية الوصول إلى المصادر الداخلية والمصادر الخارجية على حد سواء (iParadigm, 2017).

الشكل ١. رسم يوضح العملية العامة للكشف عن الانتحال العلمي (Kanjirangat and Gupta, 2016, p.10).



وتعمل أنظمة الكشف عن الانتحال باستعمال وحدات لغوية أو كتابية مختلفة. فبعض الأنظمة مبنية على الأحرف (character-based n-gram) أو على الكلمات (word-based). فحين يجري البحث عن مستند "س" بواسطة النظام، فإن النظام يتبع تردد الوحدات الحرفية أو المعجمية والمسافات بينها (Alzahrani, Salim, & Abraham, 2012)؛ ولذا، تعرف بعض هذه الأساليب للكشف عن الانتحال العلمي بالأساليب الإحصائية، حيث يمكن لبعض هذه الأنظمة البحث عن أوجه التشابه حتى وإن كان النظام لا يميز اللغة؛ لأنها تقوم بحوسبة الوحدات فحسب (Lukashenko, Graudina, & Grundspenkis, 2007, June)، رغم أن بعض الأنظمة -وإن كانت تعمل على الوحدة الحرفية والوحدة المعجمية- قد تكون معتمدة على اللغة أيضًا (language-dependent). إن بعض الأنظمة الأخرى تعمل باستعمال السمات التركيبية أو الدلالية، فيقوم النظام بتحليل اللغة الطبيعية للمستند في نطاقين هما: أقسام الكلام، والمعلومة الدلالية (Alzahrani, Salim, & Abraham, 2012; Lukashenko, Graudina, & Grundspenkis, 2007, June) في صورة تشبه إلى حد ما التحليل البشري للغة، ولكن -دون شك- أقل كفاءة.

ومهما تكن الوحدة اللغوية أو الكتابية التي تستعملها تلك الأنظمة على اختلافها، فإن هناك طريقتين شهيرتين تدار بها عملية الكشف عن الانتحال العلمي (Weber-Wulff, 2019). تتمثل الطريقة الأولى في إجراء نظام الكشف عن الانتحال العلمي مقارنةً لعينات مختارة من الوثيقة التي يجري التحقق منها مع مصادر أخرى متاحة للنظام. وتتمثل الطريقة الثانية في استعمال النظام لوحدات معجمية مفتاحية لفحص التشابه. وبصرف النظر عن الطريقة التي يستعملها النظام لإجراء المقارنة، فإنه يظل من المحتمل أن يُرجع النظام نتيجة (false positive) إيجابية خاطئة، حيث تطوق بعض أجزاء النص بعلامة تشير إلى أنه تطابق مُقلق وهو ليس كذلك، أو أن يرجع النظام نتيجة (false negative) سلبية خاطئة حيث تُترك بعض أجزائه التي تحتوي على تشابه مقلق دون علامة (Weber-Wulff, 2019). ولهذا، فما يزال من الصعب تمكين مثل هذه الأنظمة بقدرة تضاهي قدرة الفحص البشري للغة لمعالجة النص وفهمه، وعليه، فإن الأنظمة التي صممت لهذا الغرض تخفق في التقاط حالات معينة من الانتحال العلمي، في حين أن البشر، واللغويين الجنائيين على وجه التحديد، يستطيعون الكشف عنها باستعمال بعض المبادئ والأدوات الخاصة بحقل اللغويات. وهنا، تنتقل هذه المراجعة إلى جزئها الأخير المتعلق باللغويات الجنائية والانتحال العلمي.

الانتحال العلمي واللغويات الجنائية

يميل غير اللغويين إلى تسطيح صورة التعقيد الذي تنطوي عليه اللغات البشرية، ليس إلا لأننا نولد بقدرة لغوية فطرية سرعان ما نجد أنفسنا مرغمين على الشروع في تطويرها مباشرة عقب ولادة أحدنا بوقت وجيز (Klopper, 2009). غير أن المهتمين باللغة في جوانبها الذاتية وجوانبها المتصلة بالعلوم الأخرى يدركون ذلك التعقيد بمستوياته المتعددة، ولربما كان هذا أحد الأسباب التي قادت كثيرًا من المجالات العلمية كعلوم الحاسب، والعلوم التربوية، والعلوم الاجتماعية، والعلوم الجنائية وغيرها إلى تطوير مجالات فرعية كاللغويات الحاسوبية، واللغويات التربوية (أو التعليمية)، واللغويات الاجتماعية، واللغويات الجنائية. وهي فروع تستعمل علم اللغويات لمقاربة بعض القضايا والإجابة عن بعض الأسئلة التي لا يجاب عنها دون تضافر من حقلين معرفيين. وما علم اللغويات الجنائية إلا واحد من تلك الأمثلة التي يظهر فيها العوز للخبرة اللغوية؛ لتناول بعض الحالات الجنائية التي تكون اللغة ضالعة فيها. وهو أحد الفروع الحديثة نسبيًا فعمره يناهز الأربعين عامًا، وقد كان في بدايته يتنامى ببطء، وما لبث أن تنامى سريعًا في العقدين

الماضيين (Coulthard, Johnson, & Wright, 2016). وهو -حاليًا- يتطور في اتجاهين متوازيين لا يستبعد أن يكتسب كل منهما استقلاليته مستقبلًا إذا استمر البحث العلمي في كل اتجاه على الوتيرة نفسها. اتجاه يعنى بالتحليل الصوتي الجنائي، وهو ليس ذا صلة وثيقة بهذا البحث، وكثيرًا ما يجنح إلى التنميط الكمي، واتجاه آخر يُعنى بالتحليل النصي الجنائي وكثيرًا ما يميل إلى التنميط النوعي. وعلى الرغم من أن اللغة قد تستعمل بوصفها دليلًا في المحاكم والنزاع على العلامات التجارية فيدخل هذا العمل فيها تحت الاتجاه الثاني، فإن الانتحال العلمي يموضع نفسه ضمن إطار ما تعنى به اللغويات الجنائية، كما يشير عالم اللغويات الجنائية كولثارد وزملاؤه (Coulthard, Johnson, & Wright, 2016)، وهو الجانب المتعلق بهذا العمل دون سابقه.

ويسعى علماء اللغويات الجنائية إلى تحليل النصوص لتحديد التأليفية أو ملكية النصوص، وذلك بالاستناد إلى بعض العلامات اللغوية، كتنافر الأسلوب، والاختيار المعجمي، والاختيار المشترك (co-selection)، والتفرد والتطابق، والكثافة المعلوماتية (information density)، وتردد المعجميات، ومتوسطات الأطوال من كلمات وجمل وغيرهما. وهذا الأخير، على سبيل المثال، قد استعمل في تحليل التأليف المشترك (co-authorship) في حالات سابقة (Winter & Woolls, 1996)، ويتناسب مع غرض تحليل الانتحال العلمي في نصوص الطلبة والباحثين، لا سيما حين يكون الكتّاب متحدثين غير أصليين للغة التي يكتبون بها، أو ينتمون لخلفيات ثقافية مختلفة، أو أنظمة كتابية تتباين تباينًا ملحوظًا من ناحية أطوال الوحدات اللغوية. إن من يقرأ عملًا أكاديميًا لأحد الطلبة أو الباحثين -وإن يكن هذا العمل الذي بين يديك على سبيل المثال- فإنه -دون الحاجة لإجراء أي اختبارات إحصائية- قد يلحظ نمطًا لغويًا معينًا يقوده إلى وضع بعض الفرضيات حول الكاتب، وكلما كانت لدى القارئ خبرة تحليلية لغوية كان أدنى إلى فرضيات أعمق وآكد. إن مثل هذه الأنماط اللغوية قد تقدم أدلة يستعملها اللغويون لتقديم توصيات لغوية حول تأليفية نص ما لترجيح ما إن كانت بعض الأجزاء غير المعزوة في النص منتجًا يعود للمؤلف نفسه أو اقتراضًا غير مشروع من مؤلف آخر. ومع ذلك، فإن التحليل اللغوي للنصوص المشبوهة لا ينجح إلا حين يقوم به الذكاء البشري، وقلما استطاع البحث العلمي ربط الأنماط اللغوية للانتحال العلمي بالسمات النصية المحوسبة والأتمتة بنجاح (Alzahrani, Salim, & Abraham, 2012). على سبيل المثال، حاولت بعض الدراسات التحقيق في حالة من الحالات الصعبة تتعلق بالانتحال العلمي المترجم أو البين-لغوي (Şahin, Duman, Gürses, Kaleş, & Woolls, 2019; Sousa-Silva, 2014b) وخصوصًا في بعض الأعمال الأدبية (Turell, 2004)، ولكن ظل نجاح تلك المحاولات محدودًا.

وبعد أن استعرضتُ الأدبيات المتصلة اتصالًا مباشرًا بهذا العمل، ينتقل الحديث لمناقشة كيف أن أنظمة الكشف عن الانتحال قد تخفق في أدائها، في حين تنجح اللغويات الجنائية في استعمال بعض مبادئ التحقق من التأليفية للكشف عن الانتحال العلمي. وبناء على ما يتوصل إليه التحليل والنقاش في هذا الجزء، سينتقل الجزء الأخير إلى المحاججة الأخلاقية حول الاعتماد على تلك الأنظمة في الحكم على السلوكيات الخاطئة لدى الطلبة واتخاذ العقوبات المقررة.

أنظمة الكشف عن الانتحال العلمي: الأدوات اللغوية ومصادر المعلومات الأجنبية

إن باستقراء أدبيات الانتحال العلمي نجد نوعين من الانتحال العلمي مرتبطين باللغة أو معتمدين عليها، وكلاهما يتوخى تعتيم لغة النص المنتحل باستعمال الذكاء البيلغي (intra-language intelligence) أو التقنيع البين-لغوي (cross-language masking)، ويكون الأول بالتلاعب في لغة النص المنتحل بلغة نص الكتابة، في حين يكون الثاني بتبديل لغة النص المنتحل بلغة نص الكتابة. ولأن هذين

النوعين متشابهان ومترابطان ويتعمدان تعتيم لغة النص المنتحل؛ فسنستكشفهما ونناقشهما معًا في صورة تكاملية لا منفردة تحت هذا الجزء.

لقد تنبه الباحثون إلى ضرورة وجود حقل معرفي أكثر تحديدًا يُعنى بالانتحال العلمي يفيد من أدوات اللغويات الجنائية واللغويات الحاسوبية. ففي إحدى الدراسات (Klopper, 2009) يقترح المؤلف أن تتضافر جهود الحاسوبيين والمعلوماتيين مع جهود اللغويين في اتجاه أسماه (Cyber Forensic Linguistics) "اللغويات الجنائية السيبرانية". والدافع وراء ذلك هو الجمع بين التقنيات من تخصصات مختلفة وتحسين الأنظمة الحالية التي لا تعتمد في معظمها على المعلومات الدلالية (semantic information) للكشف عن الانتحال العلمي بطريقة صحيحة. وعلى الرغم من أن كلوبر يخصص جزءًا كبيرًا من ورقته لعرض قضايا متعلقة بالحوسبة البشرية للغة بوصفها عملية عقلية، إلا أنه يتناول ضرورة الإسهام اللغوي في القضايا المتعلقة بالتأليفية، والحاجة إلى الأخذ بالأدلة اللغوية في بناء أنظمة الكشف عن الانتحال العلمي. فعلى سبيل المثال، يمكن أن تكون أنماط التردد المعجمي من علامات التأليفية أو "القلْس" النصي (regurgitation)؛ لأن المعجميات المميزة (marked lexemes)، وبالتالي المعجميات الأقل ترددًا، هي مؤشرات على البصمة اللغوية، وربما قادت إلى الكشف عن انتحال علمي. يقترح كلوبر أيضًا أن تؤخذ في الحسبان بعض النماذج اللغوية التي تخرج عن المعيارية اللغوية، كالجمع بين نمط معياري وغير معياري في موضع واحد؛ فالتباينات التأليفية في نص واحد ربما كانت دليلًا لغويًا على وجود مؤلفين أو أكثر في نص واحد، في حين يبدو النص منسوبًا لمؤلف واحد يزعم ملكيته.

إن كلوبر لا يهاجم أنظم الكشف عن الانتحال العلمي بشدة أو يتهمها بالإخفاق الكبير في الكشف عن بعض نماذج الانتحال العلمي وخصوصًا البين-لغوي منه، ولكنه يايز بجلاء بين التحليل البشري والتحليل الآلي للانتحال العلمي، ويلوّح إلى أن الأوساط الأكاديمية لديها خياران للكشف عن الانتحال العلمي، وهما: التحليل اللغوي والتحليل الخوارزمي (الآلي) للتشابه، ويبدو الأخير على أنه هو الأكثر شيوعًا لدى المؤسسات الأكاديمية والتعليمية. وهناك كثير من المؤلفات العلمية بشتى أنواعها حول التحليل الآلي على عكس التحليل اللغوي (Pecorari, 2015). وتشير كثير من الدراسات إلى أن التحليل الآلي لا ينجح ولا ينجع في كثير من الأحوال لاكتشاف الانتحال العلمي وتقييمه. ففي دراسة أخرى (Mozgovoy, Kakkonen and Cosma, 2010) على سبيل المثال، يشير الباحثون الذين قدموا مسحًا لأنواع الانتحال العلمي إلى محدودية كفاءة أنظمة الكشف عن الانتحال العلمي. فكما أشرتُ سلفًا وأشار المؤلفون أيضًا، فإن واحدة من صور الانتحال العلمي هي الانتحال بالترجمة من مصادر أجنبية (وللاستزادة حول هذا الموضوع يمكن الرجوع إلى Barrón-Cedeno, Rosso, Agirre, Labaka, 2010)؛ مما يجعل من المستبعد إن لم يكن من المستحيل أن يكون المصدر الأجنبي مُضمّنًا في قاعدة بيانات النظام أو في المصادر التي يتاح للنظام الوصول لها، بل حتى وإن يكن كذلك، فإن النظام لا يستطيع المقارنة البين-لغوية ما لم يكن النظام مزودًا بوظيفة المقارنة المتعددة اللغات (multilingual comparison functions)، وهو ما لا يشيع في تلك الأنظمة. ومما تناقشه الدراسة أعلاه "انتحال الأفكار"، وهو الانتحال الأكثر صعوبة للكشف عنه بواسطة تلك الأنظمة، ويطلق عليه المؤلفون (tough plagiarism) "الانتحال العلمي العصيّ" (ص ٥٢٦). تشير الدراسة إلى أن تلك الأنظمة تعتمد على مفاهيم لغوية كالتمثيل بشجرة البنية التركيبية (structural tree diagram)، وتكمن وراء ذلك تعقيدات كبيرة، ولكنها في الأخير متوفرة في عدد محدود من اللغات كالإنجليزية والألمانية، وعادة ما تكون تلك اللغات من شجرة لغوية واحدة. ومع أن تلك الأنظمة قد تتمكن من

الكشف عن إعادة الصياغة وإعادة بناء الهيكل التركيبي للنص الأصل، فإن درجة التشابه أي: نسبة الانتحال العلمي، ستكون أقل بكثير مما ستكون عليه حين يكون الانتحال العلمي حرفيًا أو ليس من قبيل الانتحال العلمي البين–لغوي. إن الأساس الذي تقوم عليه معظم تلك الأنظمة هو اكتشاف نسبة التشابه لا معالجة اللغة البشرية كما يعالجها الذهن البشري. وبالتالي، قد يسوغ القول: إن تلك الأنظمة لا تلتقط إلا الانتحال العلمي الذي يصدر عن الطلبة الذين لا يمتلكون المهارات اللغوية العميقة لتحويل النص المصدر إلى نص جديد يبدو مختلفًا كليًا.

إن وجود المعلومات متاحةً في كل مكان وبلغات متعددة في هذا العصر قد ضاعف فرصة أن تُعاد ترجمة نص ما في لغة ما ويُعاد استعماله بطريقة غير مشروعة في لغة أخرى (Barrón-Cedeno, Rosso, Agirre, and Labaka 2010)، وهو ما حفز بعض الباحثين إلى الاهتمام بالانتحال العلمي البين–لغوي. فعلى سبيل المثال، قام بعض الباحثين (Barrón-Cedeno, Rosso, Agirre, and Labaka 2010) بمقارنة كفاءة ثلاث طرق لاكتشاف الانتحال البين–لغوي هي: النموذج البين–لغوي القائم على الحرف (Cross-Language Character n-Gram Model [CL-CNG])، ومنهج تحليل التشابه القائم على المحاذاة البين–لغوية (Cross-Language Alignment-based Similarity Analysis [CL-ASA])، ومنهج جديد مقترح في دراستهم هو تحليل التشابه الآحاديّ اللغة والترجمة الآلية (Machine Translation and Monolingual Similarity Analysis [T+MA]). وتشير النتائج إلى أن منهج T+MA كان الأجدى من بين الثلاثة، وإلى مدى صعوبة نجاح تلك الأنظمة للانتحال العلمي البين–لغوي، وكيف أن الحلول الملموسة والجذرية لهذه المشكلة ما زالت خارج قدرات تلك الأنظمة.

ومع ما قام به المؤلفون من عمل رصين في تلك الدراسة، إلا أن متخصصي اللغويات قد يطرحون بعض التساؤلات حولها. فأول ما يؤخذ على تلك الدراسة هو اللغات التي اختيرت للمقارنة، وهي: الإنجليزية والإسبانية ولغة الباسك. لقد سوغ المؤلفون اختيارهم للغة الباسك بكونها تحتوي على عدد قليل من الباسكيين المتحدثين بها، وعلى عدد قليل من المصادر المعلوماتية على الشبكة العنكبوتية. ومع أن مسوغاتهم تبدو صحيحة من جانب، إلا أنه من المعلوم للغويين أن الإسبانية ولغة الباسك لغتان غير بعيدتين من بعض، على عكس الإنجليزية والباسك إذا ما قارناهما من وجهة نظر اللغويات التاريخية. فاللغتان تشتركان كثيرًا في الخصائص الصوتية والتركيبية والمعجمية أحيانًا؛ مما يجعل الترجمة أسهل منها حين تكون اللغتان بعيدتين كالعربية والإسبانية على سبيل المثال. وهذا يؤكد على صعوبة استعمال منهج T+MA في أزواج أخرى من اللغات خصوصًا عندما يكون نص المصدر في لغة تتمايز عن لغة التأليف، كالصينية والعربية، أو تكون إحداهما ذات مصادر معلوماتية قليلة كلغة الباسك مقارنة بلغة أخرى كالألمانية. وعندما نستحضر شأن الطلبة الأجانب أو الدوليين في أي نظام تعليمي، فإن أول ما يتبارد إلى أذهاننا من كلمتي "أجانب" و"دوليين" هو أن لديهم إمكانية الوصول إلى لغة أجنبية تختلف عن اللغة المستعملة بوصفها وسيلة للتعليم في النظام التعليمي الذي يصنفون فيه على أنهم طلبة أجانب أو دوليون. وهذا بالتالي يعني أنه ليس من المستعبد أن يلجأ بعضهم إلى مصادر أجنبية وينقلها إلى لغتهم الثانية، سواء بطريقة مشروعة أم بطريقة غير مشروعة. وهنا، يجدر التأكيد على أن العربية من اللغات القليلة التي تدعمها أنظمة الانتحال العلمي مقارنة بغيرها من اللغات الشائعة كالإنجليزية، إلا أنها أحسن حالًا من بعض اللغات الأخرى كالويلزية. فعلى سبيل المثال، تعد اللغة العربية من اللغات التي يدعمها نظام SafeAssign الذي يقوم عليه نظام إدارة التعلم المعروف "بلاك بورد" ولكن ذلك الدعم جزئي حيث يتعذر على النظام القيام بالتحليل اللغوي كإعادة الكلمات المعقدة إلى جذورها، في

حين يدعم النظام هذه الخاصية في لغات أخرى كالإنجليزية والإسبانية (Blackboard, 2021).

تقدم إحدى الدراسات لصالحة الزهراني وزملائها (Alzahrani, Salim, and Abraham, 2012) مراجعة غنية وشاملة للانتحال العلمي من منظور لغوي ومن منظور حاسوبي أو معلوماتي. وتستعمل الدراسة مصطلح (intelligent plagiarism) "الانتحال العلمي الذكي" للإشارة إلى أن المنتحل قد يقوم بالتلاعب والتعتيم للنص، حتى يبدو النص مختلفًا كليًا عن النص الأصل، وذلك بنية إخفاء المخادعة، وهي ممارسة من الصعب جدًا على أنظمة كشف الانتحال العلمي الالتفاف عليها. ويشير المؤلفون إلى نوعين من الكشف عن الانتحال العلمي: الكشف الآلي والكشف البشري، وهما ما سبقت الإشارة إليه أعلاه. وتخلص الدراسة إلى أنه على الرغم من أن أدوات كشف الانتحال العلمي قد نشأت قبل أكثر من عشرين عامًا في حينها (أي: عام ٢٠١٢م)، فإن الربط بين الأنماط اللغوية للانتحال العلمي والسمات النصية الرقمية للانتحال العلمي، ليس هو الشائع في تلك الأنظمة. وبناء على ذلك، فإن كون تلك الأنظمة تقوم على مناهج وطرق توازي الوحدة بالوحدة أو الرمز بالرمز (token-by-token) سوف يدع المجال للانتحال العلمي ليبقى غير قابل للاكتشاف الدقيق إلا بالتحليل البشري.

إن هذا الزعم ليس بدعًا لدى أولئك المؤلفين، فقد تردد ذلك الملحظ لدى باحثين آخرين. بل إن بعض الدراسات التي اختبرت التحليل اللغوي البشري أثبتت مدى نجاحه مقارنة بأنظمة الكشف عن الانتحال العلمي. فعلى سبيل المثال، وجدت إحدى الدراسات (Sousa-Silva, 2014a) أن أنظمة الانتحال العلمي التي تعتمد في العموم على المقارنة القائمة على التشابه، تخرج تقارير تنفي وجود انتحال علمي رغم أن العمل منتحل (false negative)، في حين أن التحليل اللغوي يعطي تقارير تنبئ بوجود انتحال علمي. قامت الدراسة على بيانات جُمعت من تكليفات طلابية باللغة البرتغالية، كان أصحابها قد اتهموا بممارسة الانتحال العلمي، وكان من ضمن ما ركزت عليه الدراسة الانتحال العلمي بالترجمة من اللغة الإسبانية. وبناء على عدد من الأساليب اللغوية الوصفية، عثر الباحث على قدر كبير من التلاعب بالنصوص الأصلية عند إدخالها في النصوص المشتقة التي يزعم أصحابها جِدّتها وملكيّتها. ومما توصلت إليه الدراسة هو أن أولئك المنتحلين حين يتلاعبون بالنصوص لا يعتمدون على المصدر الواحد في المرة الواحدة، وإنما يستعملون مصادر معلومات متعددة يجمعون منها المعلومات ويستقون منها المعرفة، ثم يشكلونها من تلك المشارب المختلفة في قطعة نصية واحدة ذات مغزى محدد. وهذه النقطة مهمة لجوهر هذا العمل الحالي، لما تبينه من أن أنظمة الكشف عن الانتحال قد تكون مرتبطة ببعض المخاوف الأخلاقية؛ لأنها ببساطة تكشف عن المنتحلين البسطاء، وتخفق في كشف المنتحلين القاصدين الذين يتمتعون بذكاء لغوي أو يمتلكون وصولًا لمصادر معلومات أجنبية، وسأعود لهذا بالتفصيل لاحقًا. إن ما يجعل الكشف عن الانتحال العلمي آليًا في غاية الصعوبة هو أن المنتحلين يلجؤون إلى أكثر من استراتيجية لتحوير أكثر من نص ثم مخادعة النظام (Sousa-Silva, 2014a). يعبر عن ذلك وبر ولف بقوله: "ثمة عدد هائل من النصوص المحتملة وعدد هائل آخر من الطرقة الممكنة لتحوير النصوص" (Weber-Wulff, 2019, p. 72). يضاف إلى ذلك أن كثيرًا من الأنظمة تستعمل أيضًا علامات الترقيم لتحديد أجزاء النصوص ومكوناتها؛ ولهذا، فإن التلاعب بعلامات الترقيم ومواضعها المحتملة والمتوقعة يشتت النظام، ويمكن المنتحلين القاصدين من استعمال ذلك لصالحهم.

مما سيق يتضح أن هناك نوعين من الانتحال العلمي يستعصي على تلك الأنظمة كشفه، وهما: الانتحال البيلغي الذكي والانتحال البين-لغوي (الانتحال بالترجمة). إذا كان

بعض الطلبة أو الباحثين يرتكب الانتحال العلمي عن قصد وبطريقة تجعلهم يفلتون من قدرة النظام على كشفهم، ثم يفلتون من العقوبة التي تقررها مؤسساتهم التعليمية، في حين أن طلبة أو باحثين آخرين يرتكبون الانتحال العلمي عن غير قصد أو بسبب قلة الوعي أو نقص التدريب، ثم يواجهون العقوبة الصارمة لقدرة تلك الأنظمة على كشفهم، فما مقدار الثقة التي يمكننا أن نمنحها لتلك الأنظمة التي تتجاوز المنتحلين العامدين وتعرّي المنتحلين غير العامدين؟ إن هذا يستحق النظر فيه من وجهة نظر الأخلاقيات.

أنظمة الكشف عن الانتحال العلمي والمخاوف الأخلاقية

على الرغم من تركيز البحث العلمي خصوصًا في مجال علم المعلومات ومجال التربية على عدم أخلاقية الانتحال العلمي، إلا أن القليل من الأبحاث العلمية قد ناقشت أخلاقيات أنظمة الانتحال العلمي وأخلاقيات سياسات مكافحته في المؤسسات الأكاديمية. وباستقراء الأبحاث السابقة التي تعرضت لذلك الجانب (على سبيل المثال، Canzonetta, 2019; Cheong, Lau, King, 2013; Mozgovoy, Kakkonen, Cosma, 2010)، يمكن تقسيم المخاوف الأخلاقية التي ارتبطت بأنظمة الكشف عن الانتحال العلمي تحت ثلاثة أقسام رئيسة. القسم الأول يتعلق بالمخاوف الأخلاقية المرتبطة باعتراض الطلبة على نتائج فحص التشابه بواسطة أنظمة الكشف عن الانتحال العلمي؛ وذلك لأن الطلبة في كثير من حالات الاشتباه بالانتحال العلمي يرون أن تلك الأنظمة تصنف أعمالهم تصنيفًا غير عادل بوصفها أعمالًا منتحلة جزئيًا أو كليًا، في حين يرون ذلك التصنيف لا يصف واقع أعمالهم. القسم الثاني يتمحور حول عدم ارتياح الطلبة؛ لكونهم مكرهين على إيداع أعمالهم التي يتمتعون بملكيتها الفكرية في أنظمة الكشف عن الانتحال العلمي، وتخزينها في قواعد البيانات المؤسسية، بل واستعمالها لاحقًا من قِبل المطورين لتحسين تلك الأنظمة (Pecorari, 2015)، وينظرون لذلك على أنه ضرب من ضروب انتهاك حقوق الملكية الفكرية التي تعود لهم. القسم الثالث هو أقل من سابقيه، ولكنه يظل ضمن المخاوف الأخلاقية، ويتعلق بما يُشعر به الطلبة من افتراض مسبق بأن الطلبة قد يكونون قد ارتكبوا الانتحال العلمي، أو أن لديهم استعدادًا مسبقًا لذلك، فكأن ذلك اتهام مفترض بأنهم مدانون بالانتحال العلمي ما لم يحصل العكس، وهو إثبات براءتهم.

إن تلك المخاوف مشروعة وقد نوقشت في أعمال تختلف طبيعتها عن طبيعة هذا العمل. وأما الحالة التي تناقشها هذه الورقة، فهي تلك التي يكون الطالب أو الباحث فيها منتحلًا بالفعل، ولكن أنظمة الكشف عن الانتحال العلمي لم تتمكن من التقاط ذلك بنجاح. بعبارة أخرى، إذا سلمنا بأن أنظمة الكشف عن الانتحال العلمي تخفق في كشف الانتحال العلمي أو التشابه الذي قد ينبئ بوجود الانتحال العلمي الذي قد يكشفه الفحص اللساني البشري، فإن استعمال أنظمة الكشف في تلك الحالة قد يستبطن بعض المخاوف الأخلاقية ذات الصلة؛ ولذا، ينبغي فقط استعمال تلك الأنظمة على أنها تقنية مساعدة على كشف التشابه الذي يقرر المتخصص كونه انتحالًا من عدمه. ومع أن بعض الدراسات قد قدّمت أدلة إرشادية لاستعمال أنظمة الكشف عن الانتحال العلمي (على سبيل المثال، Canzonetta, 2019)، ونبّهت بعض الدراسات الأخرى إلى أن تلك الأنظمة ليست أنظمة كاشفة للانتحال العلمي، وإنما هي أنظمة كاشفة للتشابه (على سبيل المثال، Eaton, 2021)، فإن كثيرًا منها يركز على الحالات التي يكتشفها النظام ووجوب الحذر في التعامل معها، لا على الحالات التي يخفق النظام في الكشف عنها، وهو ما يتمحور حوله هذا البحث على وجه الخصوص.

يشير المتخصصون في الأخلاقيات، وتحديدًا بعض المتخصصين في أخلاقيات الحوسبة والمعلومات (على سبيل

المثال، Bynum and Rogerson, 2008)، إلى أن البشر كثيرًا ما ينجحون في اتخاذ بعض القرارات الأخلاقية، وأن لديهم مهارات بعضها فطري وبعضها مكتسب في التعرف على ما هو أخلاقي وما هو دون ذلك، ولكن من أجل الشروع في تحليل قضية معضلة ما، يتوجب على الناظر في أي قضية أخلاقية اتخاذ موقف أخلاقي معين (ethical position). وليس من السهل تصنيف استعمال أنظمة الكشف عن الانتحال بأنها غير أخلاقية، وفي المقابل، هناك بعض المخاوف التي تتضاد وأخلاقيات المعرفة، فاتخاذ موقف بتصنيفها إما أخلاقية أو غير أخلاقية ليس كما يزعم من يرون وجوب اتخاذ موقف أخلاقي ابتداء. وهذا يعني أنه في الحالات التي تمثل معضلات أخلاقية، يجب على الناظر فيها تبيين المخاوف الأخلاقية وإثباتها، ثم ترك القرار الأخلاقي للآخرين من أفراد وجهات مؤسسية. وفيما يلي سنعرض أنظمة اكتشاف الانتحال العلمي على معايير الأخلاقيات.

إن من أهم معايير أخلاقيات تقنية ما هو مدى تحقق معيار المساواة (على سبيل المثال، Feinsod & Wagner, 2008; Kvalnes, 2015)، وهو مفهوم يؤكد على ضرورة أن يتساوى البشر دون أي نوع من التمييز ضد أحد، سواء أكان المتسبب فيه اختلافات فعلية أم متوهمة. وعندما يتعلق الأمر بأنظمة كشف الانتحال العلمي، فإن بعض الطلبة يتهيأ لهم من التمكين اللغوي سواء في لغتهم الأم أم في لغة أجنبية ما يميزهم عن أقرانهم، فيستعملونه في الانتحال العلمي "الذكي"؛ مما يساعدهم على تخطي حاجز أنظمة الكشف عن الانتحال العلمي. وأولئك الذين يرتكبون الانتحال عن قصدٍ هم الذين يفترض أن تتعامل معهم اللجان التأديبية في المؤسسات التعليمية على أنهم منتحلون، وبالرغم من ذلك، فإن أنظمة الكشف -للأسف- لا تتمكن من كشف أولئك وإنما تكشف أقرانهم ممن لا يتمتعون بأدوات لغوية قوية، وربما وقعوا في الانتحال العلمي بمحض الخطأ المعرفي كما تنص بعض الدراسات (Pecorari, 2015). فإن كانت هذه هي الحال، فإن المساواة في هذا السيناريو غير متحققة؛ وبالتالي، فإن أنظمة الكشف عن الانتحال العلمي، قد تنتج عنها معاملة لا تحقق أحد أهم معايير الأخلاقيات في بعض صورها.

ويتصل بمعيار المساواة معيار "الإنصاف" (Bynum & Rogerson, 2008)، وهو معيار آخر من معايير الأخلاقيات يستعمل في الحكم على الممارسات وحديثًا في الحكم على الابتكارات التقنية (Feinsod & Wagner, 2008; Mohapatra, 2019). فإن كان ثمة حالة ما يمارس فيها بعض أفراد المجتمع ممارسة غير أخلاقية (كالانتحال في هذا السياق) فيفلت بعضهم من العقوبة، ويقع بعضهم الآخر فيها بفضل نظام أُفترض فيه أنه يكشف التشابه، وتبعًا لذلك يدل على الانتحال العلمي على نحو متشابه لدى كل مرتكبيه، فإنه يكون قد تبين للناظر في هذا الشأن دليل على أن العدالة في صورتها التامة ليست حاضرة. ومن هنا يمكن القول إنه ليس من المضمون أن يتحقق كلا المعيارين (المساواة والإنصاف) في أنظمة الكشف عن الانتحال العلمي. ويغلب على الحال أن معظم الأساتذة الجامعيين، بل حتى ولجان التأديب لا يدركون كيف تعمل تلك الأنظمة وما يمكنها وما لا يمكنها، ومدى فعاليتها في كشف الانتحال العلمي من عدمها، بل ربما جاز لنا الذهاب إلى أبعد من ذلك بالقول إنه من غير المعتاد أن يكون هؤلاء أو بعضهم على علم بدقائق عمل تلك الأنظمة. بعبارة أكثر مباشرةً، إذا جعلنا "موثوقيتنا" مبنية على أداة يمكن بسهولة أن تفقد "موثوقيتها" في أي لحظة، فإن موثوقيتنا بممارساتنا وردّة فعلنا المبنية على نتائج تلك الأداة بدورها غير ملائمة أخلاقيًا. ويتصل هذا بمبدأ آخر ذي علاقة وهو معيار "قابلية التعميم"، وهو ما يلي نقاشه أدناه.

تعد قابلية التعميم (universalizability) أحد مبادئ الأخلاقيات (traits of moral principles) كما تشير الدراسات سواء أكانت أخلاقيات الممارسات البشرية البحتة

أو أخلاقيات التقنيات (على سبيل المثال، Fieser and Pojman, 2009). ويحتّم هذا المبدأ أن تطبق المبادئ الأخلاقية على جميع الأشخاص الذين ينظر إليهم على أنهم مشمولون بوضع مماثل، وهو أحد المعايير التي نوقشت وروجعت منذ قديم، وما يزال إلى يومنا أحد معايير الأحكام الأخلاقية. على سبيل المثال:

لا يمكننا الحكم على إجراء معين على أنه صواب لدى "س" وخطأ لدى "ص" ما لم نتمكن من العثور على اختلاف في الطبيعة أو الظروف لدى أحدهما، اختلاف يمكن أن يعدّ أساسًا منطقيًا لقبول اختلافٍ بينهما فيما يتوجب عليهما. ولذلك، فإن كنت سأحكم على إجراء ما بأنه أخلاقي حين أمارسه، فإني ضمنيًا أحكم بأنه أخلاقي أيضًا للأشخاص الآخرين طالما أن الطبيعة والظروف لا تختلف بيننا في أيٍ من جوانبها الهامة (Sidgwick, 1907, p. 209).

إن السيناريو الذي يجمع الطلبة في نظام تعليمي واحد وفي مؤسسة تعليمية واحدة وفي برنامج أكاديمي واحد وفي مقرر واحد وربما في مهمة واحدة، يفترض أن يكون محكومًا بقابلية التعميم. فإنْ تعاملنا مع الطالب أو الطالبة "س" تعاملًا مختلفًا عن الطالب أو الطالبة "ص" فإن مبدأ قابلية التعميم يصبح غير متحقق. هذا لا يعني أن هناك تقصدًا لمعاملة مختلفة ومباشرة، ولكنّ هناك عاملًا يجعل المعاملة فعليًا غير مقابلة للتعميم، وهو تضمين تلك الأنظمة في السيناريو المذكور أعلاه، في حين يعتقد الأساتذة أو اللجان أن مخرجاتها موثوقة تمامًا. وقد يصبح السيناريو أكثر سوءًا عندما يرتكب المنتحلون انتهاكًا للملكية الفكرية، ويتيقّنون لاحقًا أن أنظمة الكشف عن الانتحال العلمي يمكن أن تنطلي عليها بعض الحيل اللغوية. فأولئك يحصلون على تشجيع غير مباشر لمواصلة ما يفعلونه، بل وقد يقود غيرهم إلى تقمص الممارسة نفسها، ثم قد تتطور التصورات لاحقًا لتنحو نحوًا يعتقد فيه أولئك أن ممارساتهم هي الشيء الصحيح الذي ينبغي القيام به. فإن استطاع الطالب أو الباحث أن يوظف أدواته اللغوية؛ ليفلت من التقاطة تلك الأنظمة لأوجه التشابه مما يمكن وصفه بأنه انتحال علمي، ثم تحصّلَ على نتيجة إيجابية بناء على ما قدمه كالتقييم العالي في المقرر أو الدعم المالي للبحث، فإن التأثير لا يقتصر على انتهاك حقوق المؤلف ومخادعة الأنظمة بل قد يتعدى إلى المجتمع العلمي المحيط به من طلبة أو باحثين، حيث يشرعون في استعمال استراتيجيات مماثلة، كاللجوء إلى مصادر المعلومات الأجنبية وتفعيل الأدوات اللغوية، فتفلت النصوص من قبضة الرقيب الآلي.

ويمكن أن يكون ثمة حجة أخلاقية مضادة ومشروعة لدعم ضرورة دمج تلك الأنظمة في المؤسسات الأكاديمية والمجلات العلمية وتتمثل في الآتي. ثمة دافع لاستعمال أنظمة الكشف عن الانتحال العلمي، رغم أن الحصيلة قد تبدو مثيرة لبعض المخاوف الأخلاقية في بعض الحالات، ومن المعروف أن نظريات الأخلاقيات تختلف في النظر إلى الدافع (motive) والحصيلة (outcome) لأي ممارسة أو تقنية، وبالتالي تتغير الأحكام الأخلاقية. أي أن بعض النظريات تحكم على السلوكيات بناء على الفعل نفسه بصرف النظر عن الحصيلة في حين تحكم بعض النظريات الأخرى على الحصيلة (Fieser & Pojman, 2009). فإذا نظرنا لأنظمة الكشف عن الانتحال العلمي أمكننا تسويغ استعمالها بناء على الإجراء أو الدافع؛ لأنها إنما وجدت لتعزيز دوافع أخلاقية متعددة منها مساعدة الطلبة على أن يصبحوا أكثر وعيًا بالانتحال العلمي والتيقظ له وتجنب عواقبه، ومساعدة المؤسسات الأكاديمية على تحقيق النزاهة الأكاديمية، وغير ذلك مما يمكن أن يُساق في هذا السياق. وفي المقابل، إذا قمنا بتقييم الحصيلة لتلك الأنظمة، فإنه ما يزال بإمكاننا ربطها ببعض المخاوف الأخلاقية كعدم المساواة وعدم قابلية التعميم وبالتالي عدم الإنصاف. وقد يقدم آخرون حجة مضادة أخرى تتمثل في كون الأمر بأسره يعتمد على ما إن كان الشخص الذي يقضي بأخلاقية هذه التقنية يميل إلى تقدير الدافع أو تقدير الحصيلة، غير أن الجواب عن ذلك في وجوب شمولية النظرة والتيقظ

إلى كل الأعضاء الذين قد يتضررون من ممارسة ما وإن قلّوا؛ لأن قابلية التعميم مرتكز أساسي لتقدير المخاوف الأخلاقية. فقد يُنظر إلى أنظمة الكشف عن الانتحال على أنها تشتمل على قدر من التمييز ضد الطلبة أو الباحثين الذي يرتكبون انتحالًا علميًا غير ذكي، في حين تمنح امتيازًا لأولئك الذين قُدر لهم أن يحظوا بامتلاك أدوات لغوية تحول دون نجاح خوارزميات أنظمة كشف الانتحال العلمي التي قامت على أسس لغوية أدنى من المستويات الدلالية أو التعقيدات الثنائية اللغوية. وهذا ما قد يجعل بعض المتخصصين يميل إلى عدم وسمِ هذه الأنظمة بأنظمة الكشف عن الانتحال العلمي، ويفضل وسمها بآلات (أو أدوات) النص (text-machines/tools)؛ لأنه لا يُتوقع منها أن تكشف الانتحال العلمي، بل أن تكشف حالة تُشبه الانتحال العلمي، وعلى العنصر البشري تقرير ما إذا كانت الحالة انتحالًا علميًا أو لا، وهو ما كانت قد ألمحت له بعض الأبحاث السابقة (على سبيل المثال، Weber-Wulff, 2019) دون أن تناقشه باستفاضة.

وفي دراسة مثيرة للاهتمام عن الكتابة الأكاديمية والانتحال العلمي من منظور لغوي (Pecorari, 2015)، وهو منظور قلما ظهر في أدبيات الانتحال العلمي، يُنظر إلى الانتحال العلمي على أنه قضية بين-ثقافية؛ نظرًا لكون أطياف من الطلبة لا ينتحلون عمدًا، ولكن لأنهم ينطلقون من ثقافات متأثرة بأعراف أكاديمية وتقاليد كتابية ومعتقدات فكرية متباينة. ولسوء الحظ، فإن أولئك الطلبة هم من تتمكن أنظمة الكشف عن الانتحال العلمي من تبريزهم متلبسين بما يمكن وصفه بالانتحال العلمي في ثقافة أخرى. وفي بعض حالات الانتحال العلمي، لا يكون لدى الطلبة صورة واضحة عن دورهم بوصفهم كتّابًا، وإنما ينظرون لأنفسهم على أنهم أدوات نقل للمعلومات والمعرفة من الطرف المؤلف إلى الطرف القارئ، وليسوا منتجين لها بأنفسهم (على سبيل المثال، Abasi, Akbari & Graves, 2006: Pecorari, 2015). ولهذا السبب، لا يستبعد أن يجعل هذا التصور من دمج أعمال الآخرين في أعمالهم دون الإشارة لأصحاب الفضل (أي بالاستشهاد والاقتباس) عملًا شرعيًا في تصورهم. وإضافة إلى ذلك، فإن دور الطالب أو المتعلم في بعض الثقافات خصوصًا في الثقافات الشرقية- كالصين على سبيل المثال- دور سلبي أو غير فاعل إذا ما قورن بدوره في الثقافات الغربية؛ مما يحدو به إلى اقتراض أعمال الآخرين بطريقة تبدو أخلاقية بالنسبة له ولمجتمعه الثقافي، ولكنها ليست كذلك في الثقافات الأخرى. وهذا ما ينقل النقاش هنا إلى حالة الطلبة الأجانب أو الدوليين، مثل أولئك الذين يأتون من الشرق الأوسط؛ ليدرسوا في جامعات أجنبية تكون اللغة الإنجليزية فيها وسيلة التدريس، ويجبرون من خلال فرض سياسات القبول عليهم على تعلمها وتحقيق مستوى معين من أجل قبولهم في برامجهم المرغوبة. إن هؤلاء الطلبة ربما ارتكبوا الانتحال العلمي لأسباب مختلفة، قد يكون منها ما هو متعمد دون شك، ولكن البعض منهم يقعون في الانتحال العلمي دون قصد، ومع ذلك، يظل تصنيفهم واحدًا: منتحلون.

فإذا سلمنا أن الطلبة الدوليين الذين يكابدون المعاناة؛ لتلقي تعليمهم وكتابة أعمالهم بلغة أجنبية قد يقعون في الانتحال العلمي غير النمطي دون قصد مخادعة الأستاذ، أو اللجنة، أو القارئ، وسلمنا أيضًا بأن من بين المتحدثين الأصليين من يعي ماهية الانتحال العلمي جيدًا ولديه ثقة عالية بمهاراته الكتابية وأدواته اللغوية، ولكنه يستعمل كل ذلك عمدًا لتعتيم النصوص ومخادعة الأنظمة، فإن هذه حالة لا تتحقق فيها المعايير التي أسلفنا الحديث عنها أعلاه. ويزداد الأمر تعقيدًا إذا أخذنا في الحسبان ما يشيع عن الطلبة الدوليين من تصور أنهم كتّاب عاجزون أو منتحلون محتملون، كما يتردد في بعض المجتمعات الأكاديمية وتثبته النتائج الاستطلاعية منذ قديم (على سبيل المثال، Brogan & Brogan, 1982). وهذا بحد ذاته قضية أخلاقية أخرى؛ لأن ذلك التصور يرفع مستوى الاتهام لفئة معينة من الطلبة

يعاملون مسبقًا عليها، وعند اكتشاف أي حالة تنبئ عن تشابه قد يكون انتحالًا علميًا أو قد لا يكون، فإن معاملتهم تختلف عن معاملة نظرائهم من أولئك الذين لا يتوقع منهم الانتحال العلمي. وعلى العكس من ذلك، إذا كان المتحدث الأصلي أحادي اللغة في ثقافته الأم، ومن بين أقرانه طلبة دوليون لديهم نفوذ إلى مصادر معلومات أجنبية لا يصل إليها أحادي اللغة، فإن الطالب الأجنبي أقل عرضة لكشف انتحاله عندما يرتكب الاثنان انتحالًا علميًا. ومن هنا، لا بد من استعمال هذه الأنظمة بحذر، ولا بد من توعية من يستعملونها بجوانب نجاحها وإخفاقها؛ ليس لأنها تكشف الانتحال العلمي، ولكن لأنها في بعض سيناريوهات معقدة تخفق في الكشف عن الانتحال العلمي.

ووفقًا لبعض الأسس المستعملة في التعرف على مؤلف نص ما، فإنه ليس من المستحيل توظيف تواتر معجميات ما أو استعمال خصائص أسلوبية ما لمقاربة هوية المؤلف أو التوصل لها، بل إن البعض يذهب إلى أبعد من ذلك ويحاجّ بأن ثمة بصمة لغوية فريدة تميز مؤلفًا عن آخر (Diederich, Kindermann, Leopold, & Paass, 2003). ومع ذلك، فإن الطلبة الحاذقين لغويًا يستطيعون تمييز السمات التي تبرز مؤلفًا عن مؤلف آخر، ثم يعمدون إلى تفادي العناصر المعجمية التي قد تسترعي انتباه أنظمة كشف الانتحال العلمي وتقودها إلى الوثائق التي أخذ منها أولئك الطلبة. بل إنه من غير الواقعي أن نفترض أن كل الطلبة غير مدركين لكيفية عمل تلك الأنظمة، فبعض طلبة الحوسبة والمعلومات على معرفة جيدة بخوارزميات تلك الأنظمة التي تستعملها مؤسساتهم التعليمية ربما أكثر من معلميهم. ولهذا، يرى بعض الباحثين أن أي مؤسسة تعليمية توظف تلك الأنظمة لا بد أن تعقد ورش عمل أو تجهز أدلة مكتوبة تشرح فيها كيف تعمل تلك الأنظمة (Canzonetta, 2019). وسواء أكان الطالب واعيًا بطريقة عمل النظام من تلقاء نفسه أم؛ لأن مؤسسته التعليمية قدمت له مادة تعليمية حول النظام المستعمل لديها، وسواء أكان الطالب ذكيًا لغويًا أم تقنيًا، فإن ذلك قد يسهل عملية ارتكاب الانتحال العلمي والإفلات من العقوبة، فإن السؤال الأخلاقي حول كل ذلك يظل قائمًا، وسببه الحقيقي هو إخفاق أنظمة كشف الانتحال العلمي في تحقيق مستوى واحد من الكشف عن الانتحال في كل السيناريوهات. ومن المعلوم أن كثيرًا من المؤسسات التعليمية تستعمل نظامًا واحدًا فقط كنظام SafeAssign نظرًا للقيود المالية وبخاصة تكلفة تلك الأنظمة؛ مما يتيح الفرصة أمام بعض حالات الانتحال أن تفلت؛ لكون نظام ما يعمل بطريقة واحدة، وإلا لو كان العمل الواحد يُفحص بواسطة أنظمة مختلفة لزادت احتمالية كشف الانتحال العلمي وقلت فرصة الإفلات منها.

ولا تقف المخاوف الأخلاقية عند هذا الحد، بل إنها تتجاوزه إلى حالة أشدّ تعقيدًا يصبح الاعتقاد الشائع فيها أو يفترض فيها افتراضًا مقبولًا بأن الانتحال العلمي هو ما تُبلغ عنه أنظمة الكشف عن الانتحال، وأن أي اقتراض لا تميزه تلك الأنظمة على أنه انتحال علمي فإنه ليس بانتحال علمي. فالاستعمال المفرط لتلك الأنظمة والثقة الكاملة بها في المؤسسات التعليمية، قد يشكّلان تعريفنا وتصورنا للانتحال العلمي؛ مما يفاقم المخاوف الأخلاقية المتعلقة باستعمال تلك الأنظمة وتصورنا للانتحال العلمي. وبالرغم من ذلك، فإن البعض قد يقلل من حجم تلك المخاوف الأخلاقية؛ لأنه قلما تعمقنا في التأمل في الحالات التي يُتهم فيها الطلبة بالضلوع في الانتحال العلمي. فالدراسات التي تناقش بعض السياسات والجوانب الأخلاقية الخاصة بأنظمة الكشف عن الانتحال العلمي (على سبيل المثال، Cheong, Lau, and King, 2013) تناقشها بنبرة تبدو أكثر دعمًا لتوظيف تلك الأنظمة منعًا للانتحال العلمي وحفظًا للنزاهة الأكاديمية. يتخذ هؤلاء موقفًا واضحًا بأن التقنية محايدة، وهذا بحد ذاته صحيح في عمومه، وأن وجود عدد محدود من المعضلات الأخلاقية ينبغي ألا يحول دون استعمال أنظمة الكشف عن

الانتحال العلمي. ومن المهم الإشارة إلى أنهم يوصون بإشراك العنصر البشري في أي قرار نهائي يخص حالة اشتباه بالانتحال العلمي، وهذا صحيح كليًا وحجتهم وجيهة منطقيًا. ومع ذلك، يجب عدم اقتصار النقاش على الحالات التي يكتشفها النظام، بل يجب أن يمتد إلى الحالات التي ينجح أصحابها في التغلب على الأنظمة لتضمينها في نقاش الجانب الأخلاقي. فالسؤال الأخلاقي الناجم هنا هو: هل يصح أخلاقيًا أن يعاقب طالب قد يكون ارتكب الانتحال العلمي عن غير قصد في حين ينجو طالب آخر يجلس بجواره، ولكنه سعى عن قصد لتقديم عمل منتحل نشَله عن أحد مصادر المعلومات الأجنبية وتجاوز النظام؟ أيهما حقيق بتلك التبعات الأكاديمية الصارمة بل والتبعات الاجتماعية المؤلمة للانتحال العلمي؟ قد يزعم أحدنا أن المؤسسات التعليمية لا بد وأن تحمي نفسها باستعمال ما هو متاح لها حتى وإن كان عدم الدقة أحد سمات التقنية المتاحة حاليًا. ومع ذلك، فإن أخذنا في الحسبان أحد مبادئ الأخلاقيات وهو "الغيْريّة" على اختلاف مصطلحاتها (altruism/selfless)، فقد يقال إن على المؤسسات التعليمية الموازنة بين الذات والغير.

إلى هنا نخلص إلى أن استعمال أنظمة الكشف عن الانتحال العلمي، لا بد أن يكون مصحوبًا بالإجراءات المناسبة لتنحية بعض المخاوف الأخلاقية، ولتثقيف الأساتذة الجامعيين عن حقائق أنظمة كشف الانتحال، ولدعم الطلبة وتقرير طرق الدعم قبل تقرير سُبل العقاب. وهذا لا يعني بأي حال من الأحوال أن إثارة بعض المخاوف هو دعوة مبطنة إلى التسامح مع الانتحال العلمي، ولا إلى أن تكون المؤسسات التعليمية صديقة للممارسات الخاطئة؛ لأن هذا ببساطة مستحيل منطقيًا ومتعذر تطبيقيًا ومرفوض بدهيًّا. إنما يراد من هذه الالتفاتة تبريز بعض المخاوف الأخلاقية المدفونة في كمٍ كبير من الافتراضات الصحيحة وغير الصحيحة عن أنظمة الكشف عن الانتحال العلمي، ومناقشتها مع استحضار ما توصل إليه البحث اللغوي. إن كانت تلك الأنظمة تفتح الطريق للمنتحلين النمطيين ذوي الأدوات اللغوية العصية على أنظمة كشف الانتحال، وتغلقه أمام الكتاب البسطاء الذين لا يحاتلون تلك الأنظمة، فإن هذا يثير بعض المخاوف التي لا بد أن تقدم للطرح والمناقشة. فالمشكلة هنا ليست مقتصرة فقط على نسبة قليلة من الطلبة يفلتون من قبضة أنظمة الكشف عن الانتحال العلمي لنظل نحاول تحسين الأنظمة بناء عليها، بل إن المشكلة الحقيقية هي أن النسبة القليلة من المنتحلين قد تكون هي التي تكشف عنها تلك الأنظمة، وتفلت الأغلبية، فضلًا عن كون أولئك الذين تكشف عنهم تلك الأنظمة منتحلين غير نمطيين فحسب. وفي الوقت نفسه، هي دعوة لتطوير أنظمة أكثر كفاءة وتعقيدًا بالاستفادة مما توصلت إليه اللغويات الجنائية التي عنيت بهذا الهم البحثي، فما لم نطور أنظمة ذات ذكاء لغوي أعلى، فإن التعقيدات التقنية الأخرى لا تضمن بالضرورة كفاءة أعلى. بل إن رفع مستوى حساسية تلك الأنظمة قد يُفضي إلى تصنيف النصوص غير المنتحلة على أنها منتحلة؛ مما يؤدي إلى فشل تلك الأنظمة في وظيفتها الأساسية وتفاقم المخاوف الأخلاقية.

الخاتمة

إن اللغة سمة بشرية متطورة وواحدة من أكثر الألغاز المعرفية تعقيدًا. وعلى الرغم من إحراز تقدم ملحوظ في فهم اللغة البشرية من ناحية، وفي التطور التقني الهائل من ناحية أخرى، إلا أن القدرة البشرية الطبيعية ما تزال عصية على المجاراة، وما زال بإمكان البشر استعمال ذكائهم الطبيعي، ومنه الذكاء اللغوي لجعل بعض الأنظمة التقنية كأنظمة كشف الانتحال العلمي تعمل لصالحهم حتى فيما لا يسلك مسلكًا شرعيًا أو يتبع منهجًا أخلاقيًا. ولذلك، فأخذًا في الحسبان ما توصل إليه البحث العلمي السابق- ومعظمه من اللغويات الجنائية- من أن هذه الأنظمة تخفق في أداء وظيفتها

عندما يطبق البشر خصائصهم اللغوية الفطرية لتجاوز أنظمة صممها الإنسان بمحاكاة جانب من ذكائه، فإن هذه الأنظمة ترتبط بشيء من المخاوف الأخلاقية.

ولا يستبعد أن استعمال أنظمة الكشف عن الانتحال العلمي في المؤسسات التعليمية قد يخلق بيئة آمنة للمنتحلين الذين يعلمون كيف يجعلون تلك الأنظمة تخفق في أداء وظيفتها، في حين تساعد على كشف المنتحلين الذين لا يتمتعون بامتيازات لغوية، وبالتالي يواجهون كل العواقب الأكاديمية والاجتماعية. في مثل هذا السيناريو، يكون التعامل مع الطلبة غير متكافئ وغير عادل؛ مما يدفع ببعض الأدلة على المخاوف الأخلاقية التي تثار حول تلك الأنظمة بين يدي الباحثين والمطورين التقنيين. بعبارة أكثر بساطة، إن الجمع بين أنظمة تنجح تارة وتخفق تارة وأساتذة أو لجان غير مدركين لطبيعة تلك الأنظمة، قد يؤدي إلى ممارسات غير أخلاقية. ولهذا السبب، جاء هذا العمل لمناقشة تلك المخاوف في نمطٍ تكاملي مع النتائج البحثية السابقة خصوصًا في مجال اللغويات الجنائية. ويجدر التذكير بأن المنهج الذي استعمل في هذا العمل هو منهج نظري يقوم على جمع الأدلة، ثم المحاججة الأخلاقية كما هو معمول به في أبحاث الأخلاقيات في مجال اللغويات والمعلومات والتقنيات وغيرها، ولكنه في الوقت نفسه يؤسس قاعدة لأبحاث مستقبلية تجريبية أو ميدانية تقوم على مسح حالات الطلبة الذين واجهوا نوعًا ما من العقوبات الأكاديمية، بناء على مخرجات أنظمة الكشف عن الانتحال العلمي، ثم مقارنة تلك الحالات مع حالات طلبة آخرين يقرون بارتكاب الانتحال العلمي دون أن تطالهم العقوبة. وهناك يمكن اتخاذ قرار أخلاقي وتسويغه انطلاقًا من النتائج التي يُتوصل لها.

المصادر والمراجع

Abasi, A. R., Akbari, N., & Graves, B. (2006). Discourse appropriation, construction of identities and the complex issue of plagiarism: ESL students writing in graduate school. *Journal of Second Language Writing*, 15, 102–117. 10.1016/j.jslw.2006.05.001

Alzahrani, S. M., Salim, N., & Abraham, A. (2012). Understanding plagiarism linguistic patterns, textual features, and detection methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(2), 133-149. https://ieeexplore.ieee.org/document/5766764

Barrón-Cedeno, A., Rosso, P., Agirre, E., & Labaka, G. (2010, August). Plagiarism detection across distant language pairs. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 37- 45). Association for Computational Linguistics. https://dl.acm.org/citation.cfm?id=1873786

Blackboard (2021). *Language Support*. Retrieved 16 August 2021, from https://help.blackboard.com/SafeAssign/Student/Language_Support

Brin, S., Davis, J., & Garcia-Molina, H. (1995, June). Copy detection mechanisms for digital documents. *In ACM SIGMOD Record* (Vol. 24, No. 2, pp. 398-409). ACM. https://dl.acm.org/citation.cfm?id=223855

Brogan, K. M., & Brogan, J. D. (1983). Yet another ethical problem in technical writing (*ERIC Document Reproduction Service No. ED 229 782*). https://eric.ed.gov/?id=ED229782

Bynum, T. W., & Rogerson, S. (2008). *Computer ethics and professional responsibility*. Blackwell Publishing Ltd: Oxford, UK. https://www.wiley.com/en-us/Computer+Ethics+and+Professional+Responsibility-p-9781855548459

Canzonetta, J. (2019). Can we use plagiarism detection services responsibly? In D. Pecorari & P. Shaw (Eds.), *Student plagiarism in higher education* (pp. 74–88). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315166148-6/use-plagiarism-detection-services-responsibly-jordan-canzonetta

Cheong, C. H., Lau, T. P., & King, I. (2013). Policy and issues in deploying automated plagiarism detection systems in academic communities: A case study of Veriguide. *In IT Policy and Ethics: Concepts, Methodologies, Tools, and Applications* (pp. 1425-1450). IGI Global. https://www.igi-global.com/chapter/policy-issues-deploying-automated-plagiarism/75085

Coulthard, M., & Johnson, A. (2007). *An introduction to forensic linguistics: Language in evidence*. London and New York: Routledge. https://www.semanticscholar.org/paper/An-

https://www.tandfonline.com/doi/full/10.11120/ital.2005.04020006

Lukashenko, R., Graudina, V., & Grundspenkis, J. (2007, June). Computer-based plagiarism detection methods and tools: An overview. *In Proceedings of the 2007 international conference on Computer systems and technologies* (p. 40). ACM. https://dl.acm.org/citation.cfm?id=1330642

Mohapatra, P. K. (2019). Universalisability and Objectivity. *In An Applied Perspective on Indian Ethics* (pp. 43-48). Springer, Singapore. https://doi.org/10.1007/978-981-13-7503-3_4

Mozgovoy, M., Kakkonen, T., & Cosma, G. (2010). Automatic student plagiarism detection: Future perspectives. *Journal of Educational Computing Research*, 43(4), 511-531. https://eric.ed.gov/?id=EJ952089

Pecorari, D. (2015). *Academic writing and plagiarism: A linguistic analysis*. Bloomsbury Publishing: London. https://www.bloomsbury.com/us/academic-writing-and-plagiarism-9781441139535/

Şahin, M., Duman, D., Gürses, S., Kaleş, D., & Woolls, D. (2019). Toward an empirical methodology for identifying plagiarism in retranslation. *In Perspectives on retranslation* (pp. 166-191). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780203702819-10/toward-empirical-methodology-identifying-plagiarism-retranslation-mehmet-şahin-derya-duman-sabri-gürses-damla-kaleş-david-woolls

Sidgwick, H. (1907). *The Methods of Ethics* (7th ed.). Indianapolis: Hackett Publishing.

Sousa-Silva, R. (2014a). Investigating academic plagiarism: A forensic linguistics approach to plagiarism detection. *International Journal for Educational Integrity*, 10(1), 31-41.

ttps://www.ojs.unisa.edu.au/index.php/IJEI/article/view/932

Sousa-Silva, R. (2014b). Detecting translingual plagiarism and the backlash against translation plagiarists. *Language and Law*, 1 (1), 70-94. https://ojs.letras.up.pt/index.php/LLLD/article/view/2444

Turell, M. T. (2004). Textual kidnapping revisited: The case of plagiarism in literary translation. International *Journal of Speech Language and the Law*, 11(1), 1-26. https://doi.org/10.1558/ijsll.v11i1.1

Weber-Wulff, D. (2019). Why does plagiarism detection software not find all plagiarism? In D. Pecorari & P. Shaw (Eds.), *Student plagiarism in higher education (pp. 62–73).* Routledge.
https://www.taylorfrancis.com/books/edit/10.4324/9781315166148/student-plagiarism-higher-education-diane-pecorari-philip-shaw

Winter, E. O., & Woolls, D. (1996). *Identifying authorship in a co-written novel*. University of Birmingham report.

Introduction-to-Forensic-Linguistics%3A-Language-Coulthard-Johnson/29c329dbecde8bba59cebf83ef7c08ce4c5fb5a7

Coulthard, M., Johnson, A., & Wright, D. (2016). *An introduction to forensic linguistics: Language in evidence*. Routledge: London. https://www.crcpress.com/An-Introduction-to-Forensic-Linguistics-Language-in-Evidence/Coulthard-Johnson-Wright/p/book/9781138641716

Diederich, J., Kindermann, J., Leopold, E., & Paass, G. (2003). Authorship attribution with support vector machines. *Applied Intelligence*, 19(1-2), 109-123. https://link.springer.com/article/10.1023/A:1023824908771

Eaton, S. E. (2021). *Plagiarism in higher education: Tackling tough topics in academic integrity*. ABC-CLIO: Denver.

Ercegovac, Z., & Richardson, J. V. (2004). Academic dishonesty, plagiarism included, in the digital age: A literature review. *College & Research Libraries*, 65(4), 301-318.
https://crl.acrl.org/index.php/crl/article/view/15673

Feinsod, F. M., & Wagner, C. (2008). The ethical principle of justice: The purveyor of equality. *Annals of Long Term Care*, 16(1), 23.
https://www.hmpgloballearningnetwork.com/site/altc/article/8210

Fieser, J., & Pojman, L. P. (2009). *Ethics: Discovering right and wrong*. Belmont, WA: Wadsworth.

Gibbons, J. (2003). *Forensic linguistics: An introduction to language in the justice system*. Malden, MA: Wiley-Blackwell. https://www.wiley.com/en-us/Forensic+Linguistics%3A+An+Introduction+to+Language+in+the+Justice+System-p-9780631212478

iParadigms. TurnitIn.com. Digital assessment suite. Retrieved April 20, 2019, from http://turnitin.com

Jones, K. O., Reid, J., & Bartlett, R. (2008). Cyber cheating in an information technology age. *Digithum*, 10, 19-29.
https://digithum.uoc.edu/articles/abstract/10.7238/d.v0i10.508/

Kanjirangat, V., & Gupta, D. (2016). Study on extrinsic text plagiarism detection techniques and tools. *Journal of Engineering Science and Technology Review*, *9*(5), 9-23.

Klopper, R. (2009). The case for cyber forensic linguistics. *Alternation*, 16(1), 261-294.

Kvalnes Ø. (2015) Two Ethical Principles. *In Moral Reasoning at Work: Rethinking Ethics in Organizations. Palgrave Macmillan*, London. https://doi.org/10.1057/9781137532619_5

Lancaster, T., & Culwin, F. (2005). Classifications of plagiarism detection engines. Innovation in Teaching and Learning in Information and Computer Sciences, 4(2), 1-16.