الذكاء الاصطناعي "التوليدي/ التدهوري" ونزاهة البحث

بقلم
يوري سيلان
ترجمة: د. سمير محمد

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

مقدمة

إن ظاهرة الذكاء الاصطناعي (AI) تنطوي في جوهرها على مفارقة. فمن جهة هو ذكاء توليدي. وقد أفادت هذه السمة التوليدية البحث المعاصر؛ إذ مكّنت الباحثين من توليد الأفكار وتعزيز فرص البحث مع توفير الوقت والتكاليف. ومن جهة أخرى يبدو أن الطبيعة التوليدية للذكاء الاصطناعي تقود على نحو يكاد يكون حتميًا إلى نقيضها فتُنتج الإنتروبيا عبر انهيار النموذج وتصبح طبيعةً تدهورية. نستكشف في هذه المقالة مدى إمكان عدّ الطبيعة التدهورية الضمنية للذكاء الاصطناعي التهديدَ الرئيس طويل الأمد لنزاهة البحث (RI)، لأن كثيرًا من المشكلات الأخرى المرتبطة بتأثير الذكاء الاصطناعي في نزاهة البحث يمكن النظر إليها بوصفها آثارًا مترتبة عليها. وفي الجزء الأول، نقدم عرضًا عامًا لتأثير الذكاء الاصطناعي في نزاهة البحث يتناول أخلاقيات الذكاء الاصطناعي واستخدامه في التعليم (AIED)، والذكاء الاصطناعي باعتباره «منطقة رمادية» أو ممارسة بحثية مشكوكًا فيها (QRP)، وإدراج مبادئ استخدامه في مدونات السلوك ومواقف الناشرين الأكاديميين والجامعات منه. وفي الجزء الثاني نبحث كيف يشكل انهيار الذكاء الاصطناعي التوليدي وتحوله إلى ذكاء اصطناعي تدهوري تهديدًا حاسمًا لنزاهة البحث في المستقبل. ونؤكد أن السبيل الوحيد لمنع الآثار الضارة للطبيعة التدهورية للذكاء الاصطناعي في نزاهة البحث هو الاحتفاظ بمجموعات البيانات الأصلية التي أنشأها البشر أساسًا لأنظمة الذكاء الاصطناعي، وإضافة مجموعات بيانات بشرية المنشأ جديدة إليها باستمرار. ومن ثم فإن أحد المبادئ الأساسية المتعلقة بتأثير الذكاء الاصطناعي في نزاهة البحث هو مسؤولية ضمان بقاء الذكاء الاصطناعي متجذرًا في الواقع الذي صنعه البشر. غير أن هذا يقودنا إلى المنظور الاجتماعي-التقني للذكاء الاصطناعي، وهو ما نتناوله في الجزء الثالث حيث نقوّم التأثير الاجتماعي والأخلاقي الأوسع للذكاء الاصطناعي التدهوري. ونشدد على حدوث تحول جوهري في متطلبات الثقة البشرية تجاه المجتمع، وندعو إلى إدارة استباقية لمخاطر الذكاء الاصطناعي تكون أكثر شمولًا فيما يتصل بنزاهة البحث.

الكلمات المفتاحية: الذكاء الاصطناعي؛ نزاهة البحث؛ انهيار النموذج؛ أخلاقيات الذكاء الاصطناعي؛ الذكاء الاصطناعي في التعليم؛ إنتروبيا الذكاء الاصطناعي؛ النظم الاجتماعية-التقنية.

1. نزاهة البحث وصعود الذكاء الاصطناعي التوليدي

تُعد نزاهة البحث (RI) كفاءة أساسية في البحث ومن ثم فهي خطوة حاسمة في العملية البحثية. غير أن نزاهة البحث تعرضت في السنوات الأخيرة للتهديد بسبب صعود الذكاء الاصطناعي التوليدي (Selan and Metljak 2023)، الذي لا يقتصر على توليد النصوص بل يمتد إلى قدرات متعددة الوسائط. ويجعل تطور هذه التقنية مخرجاتها صعبة التمييز من المعلومات التي ينشئها البشر (European Commission, 2025).

---

* (De) generative AI and research integrity, (Jurij Selan) / Humanities and Social Sciences Communications - 2026
DOI: 10.1057/s41599-026-08248-y
مجلة اتصالات العلوم الإنسانية والاجتماعية/ تاريخ الاستلام: 9 تشرين الأول/أكتوبر 2025 تاريخ القبول: 29 حزيران/يونيو 2026

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

لقد أدى التطور الهائل للذكاء الاصطناعي في السنوات الأخيرة إلى طمس الحدود الفاصلة بين الإبداع البشري والإبداع الذي تنتجه الآلة. وقد دفع ذلك إلى إعادة تقويم البحث الأكاديمي وأثار سؤالًا عما إذا كان لا يزال يمكن النظر إليه بوصفه مسعى بشريًا حصريًا (Gulumbe et al., 2024). إن السماح للذكاء الاصطناعي بالدخول إلى البحث الأكاديمي أعاد تعريف المناهج التقليدية والنماذج الأخلاقية المتعلقة بقيم التأليف والأصالة والصرامة المنهجية (Inam et al., 2024; Gulumbe et al., 2024).

ومن منظور نزاهة البحث يمكن النظر إلى الذكاء الاصطناعي التوليدي باعتباره «سلاحًا ذا حدين» (Nanda, 2021)، كما أن دمجه في المساعي الأكاديمية يطرح مفارقة كبيرة (Gulumbe et al., 2024). فهو قادر على دعم البحث الأكاديمي وتعزيزه، لكنه قادر أيضًا على حجب أصالة الإبداع البشري (Nanda, 2021). وقد يؤدي ذلك إلى تآكل الثقة بالعلم ونزاهته، لأنه يثير أسئلة حول كيفية الحصول على المعلومات وكيفية إجراء البحث وكيفية تقويم المنشورات. وعند استخدام ما يسمى بالذكاء الاصطناعي ذي «الصندوق الأسود»، الذي يوحي بالسلطة والموثوقية يكون خطر التحيز وتحريف الحقائق وتلفيق النتائج كبيرًا. لكن بما أن توقع حظر الذكاء الاصطناعي من البحث توقع غير واقعي فإن السؤال الرئيس هو: كيف يمكن للذكاء الاصطناعي أن يتعايش مع مبادئ نزاهة البحث؟ (Gulumbe et al., 2024).

وعلى الرغم من أن تهديد الذكاء الاصطناعي لنزاهة البحث كان متوقعًا منذ عدة أعوام (Nanda, 2021)، وأن أولى التشريعات في الاتحاد الأوروبي وُضعت سنة 2019، مثل «المبادئ التوجيهية الأخلاقية للذكاء الاصطناعي الجدير بالثقة» (European Commission, 2019)، فإن الأمر لم يصبح قضية جدية تستوجب النظر داخل المجتمع الأكاديمي وفي تعليم نزاهة البحث إلا في تشرين الثاني 2022، عندما أصبحت قدرة ChatGPT على توليد النصوص متاحة على نطاق واسع (European Commission, 2019; Moya et al., 2024). ومع صعود ChatGPT، ظهرت فورًا المعضلات الآتية: هل يمكن عدّ المحتوى الذي ينتجه الذكاء الاصطناعي محتوى أصيلًا؟ كيف ننسب التأليف إلى الأعمال التي ينشئها الذكاء الاصطناعي؟ ما الآثار التي يحدثها الذكاء الاصطناعي في تعليم الطلبة وفي تنمية قدراتهم على التفكير النقدي؟ (Gulumbe et al., 2024) ما تهديدات الذكاء الاصطناعي لنزاهة البحث، وكيف يمكن معالجتها؟ هل الذكاء الاصطناعي أداة تساعد الباحثين والطلبة على إنتاج بحوث أفضل، أم أداة تجعل الغش أسهل؟ (Nanda, 2021).

ولأن أدوات الذكاء الاصطناعي تستطيع توليد نصوص وصور ومحتويات أخرى تبدو من إنتاج البشر، بالاستناد إلى المعرفة المتاحة على الإنترنت، فقد عُدَّ استخدامها منذ البداية عاملًا يقوض نزاهة البحث بدرجة كبيرة. ولذلك شرعت الحكومات والجامعات والناشرون، على نحو متسارع، في تطبيق تدابير وقائية لتجنب الاستخدام غير الملائم للذكاء الاصطناعي في البحث.

2. أخلاقيات الذكاء الاصطناعي

يُعد تأثير الذكاء الاصطناعي في نزاهة البحث جزءًا من الموضوع الأوسع المتعلق بالقضايا الأخلاقية المحيطة بالذكاء الاصطناعي، وهو موضوع يتضمن تقويم تأثيره في جوانب الحياة البشرية كافة وتقاطع هذا التأثير مع

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

حقوق الإنسان الأساسية. ويمكن رد جذور المعضلات الأخلاقية جزئيًا إلى اللغة المستخدمة لوصف تقنيات الذكاء الاصطناعي. فقد ابتُكر مصطلح «الذكاء الاصطناعي» سنة 1956 (Holmes et al., 2022). وأوحى اختيار لفظ «الذكاء» بإضفاء صفات بشرية على وعي يشبه وعي الإنسان وعلى قدراته وفاعليته (Holmes et al., 2022). وقد ترسخ استخدام مصطلحات تجسيمية مثل «الذكاء» و«التعلم» و «التعرّف» لوصف الأنظمة القائمة على الآلة في سردية الذكاء الاصطناعي إلى حد أننا بدأنا نعتقد أن الذكاء الاصطناعي يمكن أن يكون مسؤولًا عن أفعاله (Holmes et al., 2022). لذلك لا بد لفهم الذكاء الاصطناعي فهمًا سليمًا من منظور أخلاقي من إدراك طابعه «التجسيمي». وما يسمى بمحو أمية الذكاء الاصطناعي لا يقتصر على المكونات التقنية بل يشمل أيضًا جوانبه البشرية ومنها آثاره في الأفراد والديناميات السياسية وديناميات القوة المؤثرة فيه (Holmes et al., 2022). وينبغي أن ينصب التركيز الرئيس على أخلاقيات الشركات التي صممت التقنية وصناع القرار المشاركين، لا على أخلاقيات التقنية ذاتها (Holmes et al., 2022).

ويبدو أن أخلاقيات الذكاء الاصطناعي هي الموضوع الرئيس الذي ينبغي للجميع تعلمه (Holmes et al., 2022). فهي تتناول قضايا متعلقة بالبيانات مثل الموافقة وخصوصية البيانات؛ وتحليل البيانات بما في ذلك الشفافية والثقة والمساءلة وسوق البيانات الشخصية مع التركيز على المسؤولية والخصوصية. وقد وجد جوبين وزملاؤه (2019) 84 مجموعة منشورة من المبادئ التوجيهية الأخلاقية للذكاء الاصطناعي، وحددت أغلبيتها النقاط الخمس الآتية: الشفافية، والعدالة والإنصاف، وعدم الإضرار، والمسؤولية، والخصوصية. وعلى نحو مماثل، حدد خان وزملاؤه (2021) المبادئ والتحديات الأساسية لأخلاقيات الذكاء الاصطناعي. وتمثلت المبادئ في الشفافية والخصوصية والمساءلة والإنصاف. أما التحديات فتمثلت في نقص المعرفة الأخلاقية، وغموض المبادئ أو إفراطها في العمومية، والتعارضات في التطبيق، واختلاف تفسيرات المبادئ، ونقص الفهم التقني، والقيود الإضافية. وتحدد منظمة التعاون الاقتصادي والتنمية (OECD, 2019) القيم والمبادئ الأساسية للذكاء الاصطناعي في: النمو الشامل والتنمية المستدامة والرفاه وحقوق الإنسان والقيم الديمقراطية بما فيها الإنصاف والخصوصية والشفافية وقابلية التفسير؛ والمتانة والأمن والسلامة؛ والمساءلة.

ولا تزال القوانين أو التشريعات المنظمة للذكاء الاصطناعي قليلة. وقد عرّفت «المبادئ التوجيهية الأخلاقية للذكاء الاصطناعي الجدير بالثقة»، الصادرة سنة 2019 باسم المفوضية الأوروبية (European Commission 2019)، الذكاء الاصطناعي الجدير بالثقة بأنه: (1) قانوني، أي يحترم جميع القوانين واللوائح السارية؛ و(2) أخلاقي، أي يحترم المبادئ والقيم الأخلاقية؛ و(3) متين. وبناء على ذلك، وضعت سبعة متطلبات لعدّ الذكاء الاصطناعي جديرًا بالثقة، هي: الفاعلية البشرية والإشراف البشري؛ والمتانة التقنية والسلامة؛ والخصوصية وحوكمة البيانات؛ والشفافية؛ والتنوع وعدم التمييز والإنصاف؛ والرفاه الاجتماعي والبيئي؛ والمساءلة.

وتُعد لائحة الذكاء الاصطناعي (قانون الذكاء الاصطناعي) التي اقترحها الاتحاد الأوروبي في نيسان 2021 أول لائحة للذكاء الاصطناعي تصدرها سلطة كبرى على مستوى العالم (Schwemer et al., 2021). وفي 1 آب 2024 دخل قانون الذكاء الاصطناعي حيز التنفيذ (European Commission, 2024)، وسيصبح نافذًا بالكامل في 2 آب 2026. ويهدف القانون إلى ضمان قدرة الأوروبيين على الثقة بالذكاء الاصطناعي من خلال تحديد مخاطره

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وتنظيمها. ويحدد النهج القائم على المخاطر أربعة مستويات: ذكاء اصطناعي ذو مخاطر غير مقبولة، وذكاء اصطناعي عالي المخاطر، وذكاء اصطناعي محدود المخاطر، وذكاء اصطناعي أدنى المخاطر (European Commission, 2024). ويحظر القانون الذكاء الاصطناعي ذا المخاطر غير المقبولة ويفرض متطلبات على الذكاء الاصطناعي عالي المخاطر. فالذكاء الاصطناعي ذو المخاطر غير المقبولة يهدد سلامة الناس وسبل عيشهم وحقوقهم، ويتعارض مع الكرامة والحرية والمساواة والديمقراطية وسيادة القانون وحقوق الإنسان الأساسية، ومنها عدم التمييز وحماية البيانات والخصوصية وحقوق الطفل. ومن أمثلة الممارسات المحظورة أنظمة التقييم الاجتماعي التي تستخدمها الحكومات، والألعاب التي تشجع السلوك الخطر. أما الذكاء الاصطناعي عالي المخاطر فيهدد حياة الناس وصحتهم، كما في الأنظمة الحرجة للسلامة ومجالات أخرى تشمل القياسات الحيوية، والبنية التحتية (النقل)، والتعليم (أنظمة القبول، ومراقبة السلوك المحظور للطلبة)، والتوظيف وإدارة العاملين، ومكونات السلامة (الجراحة بمساعدة الروبوت)، وإقامة العدل، وإدارة الهجرة ومراقبة الحدود، وإنفاذ القانون. ولا يُحظر الذكاء الاصطناعي عالي المخاطر، لكنه يخضع لمتطلبات شديدة الصرامة. ويشير الذكاء الاصطناعي محدود المخاطر إلى الأنظمة التي تفتقر، أساسًا، إلى الشفافية بشأن استخدام الذكاء الاصطناعي، من دون ارتباط واضح بفئتي المخاطر غير المقبولة أو العالية. ويعرّف قانون الذكاء الاصطناعي تأثير الذكاء الاصطناعي في نزاهة البحث، في المقام الأول، بأنه «خطر محدود»، ويفرض «التزامات بالشفافية» على استخدامه في البحث من أجل تعزيز الثقة؛ كضمان إبلاغ البشر على نحو صحيح باستخدام الذكاء الاصطناعي، وإتاحة التعرف إلى المحتوى المولد به (European Commission, 2024).

3. استخدام الذكاء الاصطناعي في التعليم

تتمثل إحدى القضايا الأخلاقية الحاسمة المتعلقة بالذكاء الاصطناعي في تأثيره في التعليم (Holmes et al., 2022). وتظهر قضايا عدة بشأن استخدام الذكاء الاصطناعي في التعليم (AIED)، مثل: كيف نضمن ألا يمس الذكاء الاصطناعي في التعليم حقوق الإنسان؟ وهل يعالج الأنواع الصحيحة من مهمات التعلم؟ وهل يسعى إلى جعل التعلم أكثر كفاءة فحسب، أم يعمل أيضًا على تحسينه بوصفه نشاطًا اجتماعيًا وإنسانيًا؟ وهل المقصود منه مساعدة المربين، أم إزاحتهم وجعلهم متقادمين؟ (Holmes et al., 2022) وكيف تتداخل المصلحة الخاصة لشركات الذكاء الاصطناعي مع الصالح العام؟ ولمن يدين نظام الذكاء الاصطناعي بالولاء، ولأي غرض يعمل؟ ومن يتخذ القرارات: الطلبة أم النظام التعليمي أم المدارس أم الشركات أم السياسيون؟ ومن يُسمح له بجمع الآثار الرقمية للطلبة ومن يستطيع الوصول إليها ومن يمكنه استخدام المعلومات ومن سيجني الفائدة منها؟ وأخيرًا بما أن الذكاء الاصطناعي يحتاج إلى موارد تقتصر على البلدان والمجتمعات الأكثر ثراءً فكيف ينطوي ذلك على التمييز؟

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وتشير هذه الأسئلة إلى أن الذكاء الاصطناعي في التعليم قد يقود إلى أداتنة التعليم وخصخصته، لأن معظم أدوات الذكاء الاصطناعي المستخدمة في المؤسسات التعليمية توفرها شركات تجارية. وقد يؤدي استخدام الذكاء الاصطناعي إلى تفاقم عدم المساواة بدلًا من الحد منها لدى الفئات المحرومة والفقراء والطلبة ذوي الإعاقة ومن لا تتوافر لهم التقنية أو الكفاءة التقنية المناسبة أو المهارات اللغوية (Biggs et al., 2018). وبما أن منتجًا تجاريًا واحدًا قد يعتمده نظام التعليم الحكومي بأكمله في دولة أو إقليم فقد يقود ذلك إلى استعمار التعليم بالذكاء الاصطناعي. ونظرًا إلى أن غالبية أدوات الذكاء الاصطناعي تُدرَّب باللغة الإنجليزية وتقوم عليها فقد تؤدي اللغة أيضًا دورًا في تعزيز هذا النوع من الاستعمار (Cotterell et al., 2020).

ومن الشواغل المهمة الأخرى الكيفية التي يتلاعب بها الذكاء الاصطناعي بالأطفال بطرائق قد لا يلاحظها الآباء والمدارس أصلًا (Holmes et al., 2022). فالمعضلة لا تتمثل فقط في أن فاعلية الأطفال واستقلالهم الذاتي تقيّدهما الشركات التجارية، بل أيضًا في أن الشركات الخاصة تُهندس نمو الأطفال داخل أنظمة مغلقة عصية على الاختراق. ولذلك لا تتعلق القضية أساسًا بحماية البيانات، بل بحماية الأطفال من أصحاب المصالح الخارجيين الذين يتدخلون في نموهم (Holmes et al., 2022).

ومن ثم يبرز سؤال عما إذا كان ينبغي تقويم هذه الأنظمة قبل إدخالها إلى الصف الدراسي وفق معايير مشابهة لتلك المطبقة على التدخلات الطبية. ومن المدهش أن الحكومات في أنحاء العالم لا تُبدي تحفظًا على السماح بدخول أنظمة الذكاء الاصطناعي التجارية إلى التعليم العام من دون وعي كامل بجميع الآثار التي قد تُحدثها في الأطفال، ومن دون مساءلة مباشرة أمام المتعلمين (Holmes et al., 2022).

الذكاء الاصطناعي بوصفه «المنطقة الرمادية» أو الممارسات البحثية المشكوك فيها

يمكن النظر إلى حالات سوء السلوك الأكاديمي التي تتضمن أدوات الذكاء الاصطناعي في إطار «متصل النزاهة الأكاديمية»، أي العلاقة بين النزاهة الأكاديمية وسوء السلوك الأكاديمي (Eaton, 2023; Moya et al., 2024). ويمكن تحديد آثار الذكاء الاصطناعي في البحث بوصفها آثارًا محددة الحدود وأخرى غير محددة الحدود (Moya et al., 2024). ووفقًا لـ Moya وزملائه فإن أبرز الآثار المحددة الحدود هي: (أ) تسهيل الذكاء الاصطناعي الغش الذي يتعذر اكتشافه؛ و(ب) عواقب التلفيقات التي يولدها الذكاء الاصطناعي؛ و(ج) تكريس التحيزات في الذكاء الاصطناعي، مثل العنصرية والتمييز الجنسي الكامنين ضمنيًا في بيانات التدريب. أما أكثر الآثار الأخلاقية غير محددة الحدود شيوعًا فهي: (أ) هل يُعد استخدام الذكاء الاصطناعي شكلًا من أشكال الانتحال؟ (ب) أين يقع حد المقبول في استخدام الذكاء الاصطناعي؟ (ج) من يكون المؤلف عند استخدام الذكاء الاصطناعي؟ و(د) هل يستطيع الطلبة التعلم باستخدام الذكاء الاصطناعي؟ (Moya et al., 2024).

وفي حين تشير الآثار الأخلاقية محددة الحدود إلى الآثار التي يوجد بشأنها قدر من الاتفاق بين أصحاب المصلحة، تفتقر الآثار غير محددة الحدود إلى جواب أو حل واضح بالأبيض والأسود. وفيما يتعلق بقضايا نزاهة البحث، يمكن عد الآثار الأخلاقية غير محددة الحدود للذكاء الاصطناعي في نزاهة البحث جزءًا من «المنطقة الرمادية» أو «الممارسات البحثية المشكوك فيها» (QRP) (Selan and Metljak, 2023; Butler et al., 2017)، أي

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

إن تأثيرها في نزاهة البحث لا يكون صوابًا أو خطأً بصورة مباشرة، بل يعتمد على السياق.

فمثلًا، يمكن أن تكون لأدوات الذكاء الاصطناعي، مثل مولدات النصوص، استخدامات مشروعة ومصرح بها (Moya et al., 2024). إذ تستطيع تقديم أفكار تكون نقطة بداية للكتابة، ومساعدة المؤلفين على تجاوز انسداد الكتابة، وإطلاق عملية الكتابة، ومساعدة الكتّاب على الشروع في مهمة ما أو تقديم تمثيلات بديلة للبيانات. كما يمكن لهذه الأدوات تبسيط المفاهيم المعقدة، واقتراح وجهات نظر أو إمكانات أو مواقف، وإلهام الإبداع، أو العمل محفزًا للنقد والتطوير فيما يتصل بصيغ البيانات البديلة.

غير أن مولدات النصوص نفسها وأدوات إعادة الصياغة والترجمة يمكن أن تُستخدم أيضًا لإخفاء الغش؛ مثل تقديم المخرجات على أنها عمل أصيل وهو ما يمكن عده كتابة بالنيابة (ghost-writing) (Currie, 2023)، أو استخدام أدوات إعادة الصياغة والترجمة لإعادة استعمال عمل الآخرين من دون نسبته إليهم كما في استخدام «العبارات المعذبة» (تغيير الكلمات باستخدام قاموس مرادفات) و«الترجمة العكسية» (ترجمة النص إلى لغة أخرى ثم إعادته إلى لغته الأولى للتحايل على برامج مطابقة النصوص) (Moya et al., 2024).

ولأن الانتحال أحد أخطر انتهاكات نزاهة البحث، بدأ تطوير أدوات مطابقة النصوص في مطلع القرن الحادي والعشرين لتسهيل كشف الانتحال (Andrade-Hidalgo et al., 2025). غير أن الذكاء الاصطناعي قوّض الأساس الذي تقوم عليه هذه الأدوات. فبسبب قدراته التوليدية يمكن إنتاج نصوص علمية كاملة تتجنب كشف الانتحال والتعرف البشري (Currie, 2023)، لكن ذلك يأتي على حساب أخطاء في المعلومات والاستشهادات معًا. واتساقًا مع التشبيه التجسيمي توصف الأخطاء التي يميل الذكاء الاصطناعي إلى ارتكابها بتشبيهات مأخوذة من الطب النفسي (Currie, 2023). فالهلوسة أو الوهم يشيران إلى استجابة معقولة تبدو صحيحة لكنها ليست كذلك؛ ويُعرف إنشاء مراجع أو صور أو حقائق وهمية لتعزيز أطروحة أو تقرير باسم الاختلاق؛ والوهم الحسي مماثل للهلوسة، لكنه خطأ قائم على التشابه، أي مزج عناصر مترابطة؛ أما الهذيان فهو استجابة معقدة أو غير منطقية تنتج عن خوارزمية مثقلة أو مشوشة. وإضافة إلى ذلك، قد يبدو الذكاء الاصطناعي وكأنه يغير مزاجه عند الاستجابة (Currie,لا 2023).

ويذكّرنا بيسون (2023) بتحذير بيك من أن سوء السلوك البحثي كان يتفاقم بالفعل في الأعوام السابقة وأن استخدام الذكاء الاصطناعي لم يفعل سوى تسريع هذا الاتجاه. فقد أتاح استخدام التقنية في السنوات الأخيرة إنشاء شبكات مثل «مصانع الأوراق البحثية»، التي تستخدم الذكاء الاصطناعي لإنتاج أوراق مزيفة يصعب للغاية على المؤسسات اكتشافها. وكثيرًا ما قبلت عمليات مراجعة المؤتمرات المهنية ملخصات مزورة، كما اجتاز جزء كبير من الأوراق التي اختلقها الذكاء الاصطناعي مراجعة الأقران (Currie, 2023). ويمكن إنشاء التزييفات العميقة وغيرها من البيانات والصور والتقارير الزائفة باستخدام خوارزميات الذكاء الاصطناعي التوليدي في الأشعة والطب النووي (Currie, 2023). وقد أدت التطورات في الذكاء الاصطناعي إلى ارتفاع سوء السلوك البحثي ارتفاعًا أُسيًا. ففي سنتي 2014-2015، فحص بيك وزملاؤه 20,000 ورقة بحثية، ووجدوا أن 782 ورقة ينبغي سحبها بسبب التلاعب بالصور (Bik et al., 2016). وإذا كان ذلك هو الوضع قبل عشرة أعوام، فيمكننا تصور مقدار ما أصبح عليه من سوء اليوم، بعد أن غدت مولدات الصور بالذكاء الاصطناعي شديدة القوة

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

ومتاحة على نطاق واسع. ويحذر اليم على نحو مماثل من أن استخدام الذكاء الاصطناعي التوليدي أصبح أحد العوامل الرئيسة المسببة لسوء السلوك البحثي منذ سنة 2022 (Alam, 2024).

ومنذ إدخال الذكاء الاصطناعي التوليدي إلى الاستخدام الواسع تناولت دراسات كثيرة في سياقات ثقافية متعددة عدد الباحثين الذين يستخدمونه، والأغراض التي يُستخدم فيها، والتحديات والمخاطر المتصورة.

فمثلًا بحث الزهراني (2024) كيفية تأثير أدوات الذكاء الاصطناعي التوليدي في باحثي التعليم العالي في المملكة العربية السعودية ووجد أن معظم الباحثين الشباب يحملون آراء إيجابية، ولديهم معرفة جيدة باستخدام الذكاء الاصطناعي التوليدي في البحث. وهم يدركون أن هذه الأدوات قد تحوّل البحث الأكاديمي عبر تقليل وقت التحليل، وتوليد الأفكار، وزيادة الإنتاجية، وتحسين الكفاءة. ومع ذلك، شدد الباحثون الشباب، على الرغم من تفاؤلهم، على أهمية التدريب والدعم والإرشاد الملائم في استخدام أدوات الذكاء الاصطناعي التوليدي، فضلًا عن الاعتبارات الأخلاقية. وأبرزوا كذلك التزامهم بالممارسات البحثية المسؤولة، وضرورة الشفافية، والحاجة إلى معالجة التحيزات المحتملة المرتبطة بهذه الأدوات.

ومن جهة أخرى وجد اندرسون وزملاؤه (2025) تصورات أكثر ترددًا بين الباحثين الدنماركيين. ووجدوا أن الباحثين الشباب أميل إلى استخدام الذكاء الاصطناعي التوليدي من الباحثين الأكبر سنًا وأن الباحثين يمكن تقسيمهم إلى ثلاث مجموعات. تفضل المجموعة الأولى استخدام الذكاء الاصطناعي التوليدي «كعامل يحمل عبء العمل» (35.2%). وتنظر هذه المجموعة إليه بوصفه أداة لتسريع العمليات أو للمساعدة في المسائل التقنية المرتبطة بالبحث، لكنها متشككة في استخدامه في المهمات الإبداعية، مثل مراجعة الأقران وتوليد الأفكار. وتفضل المجموعة الثانية استخدامه «مساعدًا لغويًا فقط» (24.0%). وهؤلاء الباحثون هم، عمومًا، الأكثر تشككًا في استخدامه في معظم مراحل العملية البحثية، باستثناء الاستعانة به في التحرير اللغوي. وتتكون المجموعة الثالثة من باحثين يفضلون استخدامه «مسرّعًا للبحث» (40.7%). ويقيّم هؤلاء الباحثون الذكاء الاصطناعي التوليدي بإيجابية كبيرة عمومًا في جميع حالات الاستخدام تقريبًا لزيادة الإنتاجية، ولا سيما في تحليل البيانات وتصميم البحث. وبناء على ذلك، خلص اندرسون وزملاؤه (2025) إلى أن الباحثين ينظرون بسلبية أكبر إلى مهمات تصميم التجارب ومراجعة الأقران، في حين ينظرون بإيجابية متكررة إلى التحرير اللغوي وتحليل البيانات. وتشدد أغلبية الأكاديميين على ضرورة الاستخدام النقدي والتأملي للذكاء الاصطناعي التوليدي، في حين يُعد توليد الصور أو تعديلها، والبيانات الاصطناعية، موضوعين مثيرين للجدل على نحو خاص.

إدراج مبادئ استخدام الذكاء الاصطناعي في مدونات السلوك

بسبب التطور السريع توجد حاجة ملحة إلى إدراج الذكاء الاصطناعي في المبادئ والأطر التنظيمية والتعليمية والأخلاقية. وفي ضوء التهديدات المحددة لنزاهة البحث، اقتُرحت قيم ومبادئ أساسية لاستخدام الذكاء الاصطناعي، وأُدرجت في مدونات سلوك متعددة خلال السنوات الأخيرة.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وقد جرى تحديث «المدونة الأوروبية لقواعد السلوك من أجل نزاهة البحث» ومراجعتها لمعالجة الفرص والتهديدات التي يطرحها الذكاء الاصطناعي (Brent, 2023; Allea, 2023). وتتضمن المدونة الجديدة إرشادات بشأن الذكاء الاصطناعي، وتعكس التطورات الحديثة في هذا المجال (Brent, 2023). ووفقًا للمدونة، يمكن للذكاء الاصطناعي أن يؤثر في نزاهة البحث بطريقتين (Brent, 2023): الأولى توليد البيانات وكتابة الأوراق البحثية، والثانية استخدامه أداة لكشف الانتحال. وفي الحالتين، تكون الشفافية هي الشاغل الأساس. فأدوات الذكاء الاصطناعي تتمتع بقدرة هائلة، ويمكن استخدامها لتحقيق فائدة كبيرة، لكن أهم متطلب هو أن يكون استخدامها شفافًا. وهذا هو الاهتمام الرئيس للمدونة المنقحة (Brent, 2023).

وأُجريت التعديلات المتعلقة بالذكاء الاصطناعي على المدونة في ثلاثة مواضع (Allea, 2023). ففيما يتصل بإجراءات البحث أو الممارسة البحثية، عُدلت الفقرة 6 الخاصة بالإبلاغ لتشمل الذكاء الاصطناعي، وصارت تنص على ما يأتي: «يُبلغ الباحثون عن نتائجهم ومناهجهم، بما في ذلك استخدام الخدمات الخارجية أو الذكاء الاصطناعي والأدوات الآلية، بطريقة تتوافق مع الأعراف المقبولة في التخصص، وتيسر التحقق أو التكرار، حيثما ينطبق ذلك» (Allea, 2023). وعلى نحو مماثل، حُدثت الفقرة 2 الخاصة بالشفافية ضمن بروتوكول المراجعة والتقويم، فأصبحت تتطلب الإفصاح عن استخدام الذكاء الاصطناعي، وتنص على ما يأتي: «يراجع الباحثون والمؤسسات والمنظمات البحثية الطلبات المقدمة للنشر أو التمويل أو التعيين أو الترقية أو المكافأة ويقوّمونها بطريقة شفافة ومبررة، ويفصحون عن استخدام الذكاء الاصطناعي والأدوات الآلية» (Allea, 2023). وأخيرًا، أضيفت، ضمن باب الانتهاكات، في قسم سوء السلوك البحثي وغيره من الممارسات غير المقبولة، فقرة إضافية عرّفت بوصفه سوء سلوك: «إخفاء استخدام الذكاء الاصطناعي أو الأدوات الآلية في إنشاء المحتوى أو صياغة المنشورات» (Allea, 2023).

وتشكل «المدونة الأوروبية المنقحة لقواعد السلوك من أجل نزاهة البحث»، إلى جانب «المبادئ التوجيهية الأخلاقية للذكاء الاصطناعي الجدير بالثقة» لسنة 2019، الأساس الذي تقوم عليه «المبادئ التوجيهية الحية للاستخدام المسؤول للذكاء الاصطناعي التوليدي في البحث»، الصادرة عن منتدى منطقة البحث الأوروبية (European Commission, 2025). وتهدف هذه المبادئ إلى توجيه استخدام الذكاء الاصطناعي في البحث في القطاعين العام والخاص. وهي تقر بأهمية الذكاء الاصطناعي لما يتيحه من فرص كثيرة، لكنها تقر أيضًا بالمخاطر التي يطرحها، وفي مقدمتها خطر التضليل واسع النطاق وغيره من الاستخدامات غير الأخلاقية ذات الأثر الاجتماعي الكبير. وتلتزم هذه المبادئ بالإعلان العالمي لحقوق الإنسان، وتتوافق مع توصية اليونسكو بشأن أخلاقيات الذكاء الاصطناعي (Kaushik, 2021)، وكذلك مع مبادئ منظمة التعاون الاقتصادي والتنمية للذكاء الاصطناعي (OECD, 2019).

وتحدد المبادئ التوجيهية أربعة مبادئ أساسية للاستخدام المسؤول للذكاء الاصطناعي في البحث (European Commission, 2025): الموثوقية، أي أن تكون المعلومات التي ينتجها الذكاء الاصطناعي قابلة للتحقق وإعادة الإنتاج، وأن يجري تجنب التحيزات؛ والأمانة، أي الإفصاح عن استخدام الذكاء الاصطناعي التوليدي؛ والاحترام، أي مراعاة الأثر البيئي للذكاء الاصطناعي، وآثاره الاجتماعية، والتحيز، والتنوع، وعدم التمييز،

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

والإنصاف، ومنع الضرر؛ والمساءلة، أي مسؤولية الباحث عن جميع المخرجات التي ينتجها الذكاء الاصطناعي.

ولكي يُستخدم الذكاء الاصطناعي التوليدي بمسؤولية، ينبغي للباحثين:

1. أن يكونوا مسؤولين عن المخرجات.

2. أن يستخدموا الذكاء الاصطناعي بشفافية.

3. أن ينتبهوا إلى الخصوصية والسرية وحقوق الملكية الفكرية.

4. أن يحترموا القانون الوطني وقانون الاتحاد الأوروبي والقانون الدولي.

5. أن يواصلوا التدريب والتعلم بشأن الاستخدام السليم للذكاء الاصطناعي.

6. أن يتجنبوا الإفراط في استخدام أدوات الذكاء الاصطناعي في الأنشطة الحساسة التي تؤثر في باحثين ومنظمات أخرى، مثل مراجعة الأقران والتقويمات وغيرها (European Commission, 2025).

مواقف الناشرين الأكاديميين والجامعات من الذكاء الاصطناعي

نظرًا إلى المعضلات المرتبطة بالذكاء الاصطناعي، طبقت مؤسسات علمية وهيئات تمويل وناشرون كثيرون سياسات لنزاهة البحث تتضمن بنودًا بشأن استخدام الذكاء الاصطناعي، وكشفه، والتأليف، والغش، والانتحال (Moya et al., 2024). وقد حظرت بعض المؤسسات استخدام الذكاء الاصطناعي في البداية، لكن هذا النهج يبدو غير واقعي مع تزايد اندماجه في الحياة اليومية. ولذلك قدمت المؤسسات إرشادات وتدريبًا لتعزيز فهم قدرات الذكاء الاصطناعي، والمساعدة في تطوير المهارات اللازمة لمستقبل تحركه تقنياته (Moya et al., 2024). وتشدد هذه الإرشادات على أهمية محو أمية الذكاء الاصطناعي وإدماجه في الاستراتيجيات التعليمية مع الحفاظ على النزاهة الأكاديمية.

وفي تحليل مفصل لإرشادات كبار الناشرين، راجع جولومبي وزملاؤه التغييرات وإدراج سياسات الذكاء الاصطناعي في الإرشادات العلمية (Gulumbe et al., 2024). ويستبعد الناشرون الذكاء الاصطناعي، عمومًا، من التأليف، مؤكدين الدور الحاسم الذي يؤديه الذكاء البشري والمساءلة في العمل الأكاديمي. كما يدعون إلى الإفصاح الصريح والصادق عن استخدام الذكاء الاصطناعي، ويشددون على ضرورة الحفاظ على أصالة البحث ونزاهته. وتؤكد هذه المعايير أيضًا ضرورة أن يضمن المؤلفون فرادة المعلومات المنتجة بدعم الذكاء الاصطناعي وصحتها. ومن خلال تقديم الذكاء الاصطناعي بوصفه أداة مساعدة، لا بديلًا من البراعة البشرية، تُظهر المعايير المنقحة التزامًا بالحفاظ على المعايير الأكاديمية. وتتناول القواعد كذلك القضايا الأخلاقية المرتبطة بميل الذكاء الاصطناعي إلى تضخيم التحيزات في البيانات، وخطر تقويض نزاهة البحث الأصيل (Gulumbe et al., 2024).

وقد توحدت معظم المؤسسات العلمية الآن حول المبادئ الآتية لاستخدام الذكاء الاصطناعي في البحث (ACM US Technology Policy Office, 2023; STM, 2021): الشفافية والمساءلة، والجودة والنزاهة، والخصوصية والأمن، والإنصاف، والتنمية المستدامة.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

3. إنتروبيا الذكاء الاصطناعي: من الذكاء الاصطناعي التوليدي إلى الذكاء الاصطناعي التدهوري

كما تقدم دُرست «طبيعة الذكاء الاصطناعي ذات الحدين» وأُدرجت في السنوات الأخيرة، ضمن مدونات السلوك والمبادئ الأساسية لمؤسسات علمية كثيرة، مثل الجامعات والناشرين.

ومع ذلك ثمة قضية تتعلق بتأثير الذكاء الاصطناعي في نزاهة البحث لم تحظ بتأكيد أو استكشاف وافيين، مع أنها قد تحمل بعضًا من أشد الآثار ضررًا في جودة البحث ونزاهته مستقبلًا. فقد أعمَتنا الإثارة التي صاحبت مشهد الذكاء الاصطناعي سريع التطور عن مفارقة قاسية: قد تشكل مخرجات الذكاء الاصطناعي نفسها تهديدات كبيرة وشواغل أخلاقية لتطوره مستقبلًا (Lynch, 2024). وهذه القضية، التي توصف بأنها «إنتروبيا الذكاء الاصطناعي»، قد تقوض جميع منجزاته (Marr, 2024).

انهيار النموذج

أُثيرت قضية إنتروبيا الذكاء الاصطناعي حديثًا في أوراق علمية ومنشورات إعلامية (Bhatia, 2024)، بحثت مخاطر استخدام مخرجات الذكاء الاصطناعي نفسه لتدريب نماذج الذكاء الاصطناعي المستقبلية (Hammond, 2024). وقد أطلق على هذه الظاهرة اسم «انهيار النموذج» (Marr, 2024). تقليديًا تُدرَّب أنظمة الذكاء الاصطناعي على مجموعات بيانات هائلة جُمعت من الإنترنت. وهذه المجموعات التي أنشأها البشر في الأصل تُظهر تعقيد الإبداع والثقافة البشريين وتنوعهما (Marr, 2024). وتُعرف البيانات التي أنشأها البشر باسم «البيانات العضوية» (Marr, 2024). ولتدريب GPT-3، احتاجت OpenAI إلى أكثر من 650 مليار كلمة انتُقيت من بيانات خام على الإنترنت يزيد حجمها مئة مرة، وقد رُشّح 98% منها واستُبعد الباقي (Brown et al., 2020). وهذه كمية هائلة من البيانات لإمداد الذكاء الاصطناعي بها (Snoswell, 2024). غير أنه مع تزايد إنتاج الذكاء الاصطناعي التوليدي للمحتوى يتنامى الميل إلى استخدام البيانات التي يولدها الذكاء الاصطناعي أي ما يسمى «البيانات الاصطناعية» (Marr, 2024)، بدلًا من البيانات العضوية البشرية مصدرًا للتدريب (Romano, 2023). ومنذ إطلاق ChatGPT سنة 2022، أخذ الناس ينشرون قدرًا متزايدًا من المحتوى المولد بالذكاء الاصطناعي على الإنترنت. وقد كتب سام ألتمان على منصة X أن OpenAI تولد نحو 100 مليار كلمة يوميًا (Altman, 2024)، وكلها تدخل قاعدة بيانات الإنترنت، وتجعل من الأصعب يومًا بعد يوم معرفة ما هو حقيقي؛ إذ لا توجد وسيلة للتمييز بين البيانات العضوية والاصطناعية (Austin, 2024). لذلك بدأ الباحثون سنة 2023 يتساءلون عما إذا كان بإمكانهم الاعتماد حصريًا على البيانات الاصطناعية لتدريب نماذج الذكاء الاصطناعي. ولو كان ذلك ممكنًا، لكان أرخص بكثير، وأقل إثارة للمشكلات الأخلاقية والقانونية من جمع البيانات البشرية (Snoswell, 2024).

والسؤال هو: ماذا يحدث عندما يُدرَّب نموذج ذكاء اصطناعي لا على البيانات العضوية وحدها بل أيضًا على بيانات اصطناعية أنتجها الذكاء الاصطناعي نفسه؟ وماذا يحدث لنموذج الذكاء الاصطناعي عبر الأجيال المتعاقبة؟ (Shumailov et al., 2024) تقوم أنظمة الذكاء الاصطناعي على الاحتمالات إذ تتنبأ بالنتيجة الأرجح وتولدها استنادًا إلى البيانات التي دُربت عليها. فإذا دُرّب نموذج جديد على مخرجات ولدها الذكاء الاصطناعي، نشأ تأثير حلقة ذاتية، أي نوع من الحلقة المفرغة، تبدأ فيها المخرجات بالتقارب نحو حلول أضيق وأعم وأكثر

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

تجانسًا، لأنها الحلول الأعلى احتمالًا (Hammond, 2024). ونتيجة لذلك، يميل النموذج إلى تجاهل العناصر غير المألوفة والمفاجئة في البيانات الأصلية التي دُرب عليها (Marr, 2024). وهذا ما يسمى «تأثير غرفة الصدى»، الذي قد يقود إلى انهيار النموذج (Marr, 2024). وعندما يتعلم الذكاء الاصطناعي من مخرجاته يتدهور فهمه للعالم لأن هذه المخرجات لا تكون مثالية أبدًا فتغدو المعلومات أكثر تجانسًا وتدهورًا مع الزمن. وفي النهاية ينهار النموذج، ويعجز عن توليد محتوى جديد ذي معنى. ويشبه ذلك تكرار نسخ نسخة؛ إذ تفقد كل نسخة جديدة بعض تفاصيل الأصل، فتنتج صورة أقل دقة وأكثر ضبابية للعالم (Marr, 2024). ويمكن بدلًا من ذلك تصور طلبة لا يقرؤون الكتاب المدرسي الفعلي أبدًا بل يتعلمون من ملاحظات طلبة آخرين؛ فيصبح فهمهم مشوهًا وجزئيًا مع مرور الوقت وينتهي إلى معرفة قاصرة ومنحازة (Infobip, 2024). ويذكّر إيتون بفيلم ملتبلسيتي من بطولة مايكل كيتون حيث يصنع بطل الفيلم نسخًا جديدة من نفسه (Eliot, 2024). ويتدهور تدريجيًا؛ يكون التغير في البداية غير ملحوظ تقريبًا لكنه يصبح أكثر وضوحًا مع كل عملية استنساخ جديدة (Marr, 2024). ومع مرور الزمن تتراكم الأخطاء وتشوه إدراك النموذج للواقع (Vue.ai, 2023).

وثمة فرق بين انهيار النموذج (model collapse) وانهيار الانماط (mode collapse) / (Lutkevich, 2023). يحدث انهيار الأنماط في شبكة توليدية خصومية (GAN) عندما تُولد عينات محدودة بصرف النظر عن المدخلات ويفشل النموذج في تغطية الأنماط المتعددة ضمن توزيع متنوع للبيانات. أما انهيار النموذج فيعني فشل النموذج في التعلم على نحو سليم وهو كامن في نماذج التعلم الآلي كلها عند استخدام بيانات تدريب اصطناعية. وينتج تدهور الذكاء الاصطناعي أيضًا من تدريبه على بيانات ولدتها نماذج أخرى، لا من تدريبه على مخرجاته وحدها. ويبين شومايلوف وزملاؤه أن تدريب نموذج ذكاء اصطناعي على نتائج نموذج ذكاء اصطناعي آخر يمكن أن يسبب بمرور الزمن انهيار النموذج (Shumailov et al., 2024).

وينتج انهيار النموذج من ثلاثة عوامل (Shumailov et al., 2024). الأول هو تراكم الأخطاء، حيث تنتقل عيوب التكرارات السابقة ويضخمها كل جيل جديد من النماذج، فتفضي إلى مخرجات تنحرف عن أنماط البيانات الأصلية. والعامل الثاني هو فقدان بيانات الذيل؛ ويحدث عندما تُستبعد الوقائع غير الشائعة من بيانات التدريب، فتُحجب بذلك مفاهيم كاملة. وأخيرًا، تعمل حلقات التغذية الراجعة على إدامة أنماط محدودة، فتنتج مخرجات متحيزة أو متكررة (Lutkevich, 2023).

ويميز شومايلوف وزملاؤه بين الانهيار المبكر والانهيار المتأخر للنموذج (Shumailov et al., 2024). يقع الانهيار المبكر عندما يفقد الذكاء الاصطناعي معلومات عن ذيول التوزيع. فتختفي الذيول أي الأحداث منخفضة الاحتمال وينكمش التوزيع. وفي النهاية تصبح المخرجات أكثر تشابهًا ولا يبقى بينها وبين البيانات الأصلية إلا شبه ضئيل (Lutkevich, 2023). أما في الانهيار المتأخر، فيدمج الذكاء الاصطناعي أنماطًا ينبغي أن تظل متميزة، وينهار النموذج في نمط واحد (Shumailov et al., 2024).

وفي الحالتين يعني هذا الانهيار انخفاض جودة المخرجات وتنوعها: فمع تزايد الأجيال تبتعد المخرجات، اعتباطيًا، عن المصدر الأصلي أكثر فأكثر وهذا فقدان للجودة؛ ومع تزايد الأجيال يقل التباين في المخرجات تدريجيًا، وهذا فقدان للتنوع (Shumailov et al., 2024). ومن دون قدر كافٍ من البيانات العضوية الجديدة

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

ستتراجع تدريجيًا جودة النماذج التوليدية المستقبلية أي دقتها أو تنوعها أي قدرتها على الاسترجاع (Alemohammad et al., 2023; Shumailov et al., 2024). وتبدأ الصور مثل الوجوه والجمل والأرقام كلها بالتشابه (Austin, 2024).

ويقدم شومايلوف وزملاؤه (Shumailov et al., 2024) و بورجي (Borji, 2024) تفسيرات رياضية تبين كيف يؤدي انهيار النموذج حتمًا إلى فقدان التنوع والإبداع في المخرجات خلال بضعة أجيال فقط. فبعد تكرارات عدة يتقارب التوزيع المنتظم في نهاية المطاف نحو توزيع طبيعي مما يؤدي إلى الانهيار؛ إذ تُسوّى بنية البيانات الأصلية نتيجة تكرار أخذ العينات وإعادة المواءمة، فتُمحى خصائصها المميزة تدريجيًا وينشأ منحنى أقرب إلى شكل الجرس (Lutkevich, 2023). فمثلًا عند استخدام نص عن عمارة العصور الوسطى مدخلًا أنتج النموذج في دورة الجيل التاسع مخرجات مكونة من كلام غير مفهوم (Lutkevich, 2023). وعلى نحو مماثل يقدم فينغر مثالًا افتراضيًا لنموذج ذكاء اصطناعي طُلب منه توليد صور للكلاب (Wenger, 2024). سيميل النموذج إلى إعادة إنتاج السلالات الأكثر شيوعًا في مجموعة التدريب مثل كلاب المسترد الذهبي. وبعد عدد كافٍ من الدورات، سيتجاهل النموذج سلالات الكلاب غير الشائعة، ولن يولد إلا صورًا لكلاب المسترد الذهبي. وفي النهاية سيعجز عن توليد محتوى ذي معنى ولن ينتج إلا بقعًا لونية بلا معنى؛ أي إنه سينهار (Wenger, 2024).

وتُعرف ظاهرة انهيار النموذج أيضًا بتسميات: «الذكاء الاصطناعي يأكل الذكاء الاصطناعي» (Vue.ai, 2023)، و«أكل الذكاء الاصطناعي للحوم جنسه» (Infobip, 2024)، و«الذكاء الاصطناعي ذاتي الافتراس» (Alemohammad et al., 2023)، و«ذكاء أوروبوروس الاصطناعي» (Vue.ai, 2023)، و«ذكاء هابسبورغ الاصطناعي» (Sadowski, 2023). وتوحي هذه التسميات بأن الذكاء الاصطناعي يتغذى على جنسه أو يلتهم نفسه أو ذيله أو ينخرط في شكل من أشكال سفاح القربى الرقمي أو التزاوج الداخلي (Snoswell, 2024)، مما يؤدي إلى انهيار ينهي «سلالته» وقدرته على «التكاثر» على نحو سليم. كما يطلق Alemohammad وزملاؤه على ذلك اسم «اضطراب الالتهام الذاتي للنموذج» (Model Autophagy Disorder: MAD)، قياسًا على مرض جنون البقر (Alemohammad et al., 2023).

تأثير الطبيعة التدهورية للذكاء الاصطناعي في نزاهة البحث

قد تبدو قضية انهيار النموذج للوهلة الأولى مشكلة تقنية متخصصة وضيقة (Marr, 2024). غير أن انهيار النموذج قادر على تقويض موثوقية أنظمة الذكاء الاصطناعي؛ ومن ثم تتجاوز القضية التحديات التقنية، وتثير شواغل أخلاقية واجتماعية كبيرة (Borji, 2024). والواقع أن المشكلات المرتبطة بنزاهة البحث، مثل هلوسات الذكاء الاصطناعي، يمكن النظر إليها بوصفها نتيجة مباشرة لإنتروبيا الذكاء الاصطناعي.

فما تأثير الذكاء الاصطناعي التدهوري في نزاهة البحث؟

أولًا، يصبح الذكاء الاصطناعي التدهوري أكثر محدودية في الإبداع والابتكار. فإدماج مخرجات الذكاء الاصطناعي في بيانات تدريبه المستقبلية ينشئ حلقة تغذية راجعة تقل معها احتمالية أن يقدم النموذج

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

مخرجات متنوعة أو جديدة تدريجيًا (Infobip, 2024). ولا تستطيع النماذج المنهارة أن تبتكر حقًا أو تدفع الحدود إلى الأمام (Infobip, 2024). ويرتبط ذلك أيضًا بـ«اختراق المكافأة». فكثير من نماذج الذكاء الاصطناعي تحركها أنظمة حوافز تهدف إلى تحسين مقاييس محددة. ومع الزمن قد يتعلم الذكاء الاصطناعي «التحايل» على النظام بإنتاج استجابات تزيد المكافأة إلى أقصى حد، لكنها تفتقر إلى الإبداع أو الأصالة. فضلًا عن ذلك، قد يحفز تدريب النموذج الذكاء الاصطناعي على اللجوء افتراضيًا إلى استجابات «آمنة» لتجنب إنتاج مخرجات مسيئة أو مثيرة للجدل، فينتج عن ذلك مخرجات باهتة ومتجانسة (Infobip, 2024).

ثانيًا، يفقد الذكاء الاصطناعي التدهوري صلته بالعالم الحقيقي، ويقود إلى تمثيل الواقع تمثيلًا خاطئًا، فتنتج هلوسات الذكاء الاصطناعي. وعندما تُدرَّب النماذج حصريًا على بيانات اصطناعية، فإنها تبدأ بنسخ الأنماط الموجودة داخل تلك البيانات، بدلًا من التعلم من دقائق البيانات المستمدة من البشر (Infobip, 2024). ويكون هذا الذكاء الاصطناعي التدهوري أقل قدرة على التعامل مع تحديات العالم الحقيقي التي تتطلب فهمًا دقيقًا وحلولًا قابلة للتكيف. كما أنه يشوه الواقع ويلفقه، ويقدم نتائج خاطئة من حيث الوقائع أو مضللة أو عديمة المعنى (Infobip, 2024). ونتيجة مباشرة للفقدان المستمر لجودة مخرجات الذكاء الاصطناعي وتنوعها، يفقد الذكاء الاصطناعي صلته بمعرفة العالم الحقيقي، ويغدو أكثر عرضة للهلوسة.

ثالثًا، يكرس الذكاء الاصطناعي التدهوري التحيزات والصور النمطية الموجودة ضمنيًا في مجموعات البيانات الأصلية، ويقلص التنوع الاجتماعي-الثقافي (Infobip, 2024). فعندما يواصل الذكاء الاصطناعي التعلم من بياناته، يضخم التحيزات والصور النمطية الموجودة في بيانات التدريب الأصلية، ويعزز أعرافًا اجتماعية ضارة قد لا تكون موجودة في المادة المصدرية إلا على نحو خفي (Romano, 2023). والأحداث منخفضة الاحتمال، التي كثيرًا ما تتضمن فئات مهمشة أو أوضاعًا غير مألوفة، معرضة على نحو خاص لأن «تنساها» خوارزميات الذكاء الاصطناعي (Marr, 2024). ونتيجة لذلك، يصبح الذكاء الاصطناعي أقل قدرة على فهم مطالب المجتمعات المختلفة والاستجابة لها، ويعزز التحيزات وأوجه اللامساواة القائمة (Marr, 2024). ويفقد التنوع الاجتماعي-الثقافي، إلى حد المحو الثقافي لبعض الجماعات (Snoswell, 2024). ويبعث هذا على قلق خاص في مجالات مثل توليد الأخبار، والمحتوى التعليمي، والبحث العلمي، حيث تكون الدقة والموثوقية في غاية الأهمية (Romano, 2023; Marr, 2024).

ويقود ذلك كله إلى أخطر نتيجة بالنسبة إلى نزاهة البحث: فقدان الأصالة والثقة بالذكاء الاصطناعي (التوليدي/التدهوري). فمع ازدياد انتشار مخرجات الذكاء الاصطناعي المهلوسة والمتحيزة والنمطية، ومع ازدياد صعوبة التمييز بين المحتوى الذي يولده الذكاء الاصطناعي والمحتوى الذي ينشئه البشر، نواجه حتمًا فقدان الأصالة والثقة بالذكاء الاصطناعي (Romano, 2023). وبما أنه لا توجد وسيلة للتمييز بين البيانات العضوية والاصطناعية، لا يمكننا أن نثق بما هو حقيقي وما ليس حقيقيًا، وما هو مبتكر وما ليس مبتكرًا، وما هو متحيز وما ليس متحيزًا.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

إضعاف القدرات والخرف الاصطناعي

يمكن تلخيص الأثر العام للذكاء الاصطناعي التدهوري وفقًا لـ فاولر، بأنه أثر «إضعاف القدرات العقلية» (dumbing down) (Fowler, 2023). وكما في لعبة «الهاتف»، يزداد تشوه نموذج الذكاء الاصطناعي المدرب على بيانات أنشأها بنفسه مع كل تكرار (Sweenor, 2023). وفي حلقة التغذية الراجعة هذه، يضخم الذكاء الاصطناعي الأخطاء والتحيزات باستمرار بمرور الزمن. ولذلك يتدهور الذكاء المحاكى للذكاء الاصطناعي؛ والأهم أن هذا التدهور قد يؤثر أيضًا في ذكاء مستخدميه إذا بالغوا في الاعتماد عليه. فالذكاء الاصطناعي التوليدي لا يصبح تدهوريًا هو نفسه فحسب، بل قد يدفع الباحثين أيضًا إلى التدهور و«الانهيار».

ولكن كيف يمكن حدوث أثر «إضعاف القدرات العقلية» هذا إذا كان الذكاء الاصطناعي يحاكي، في الظاهر، القدرات الإدراكية البشرية، وقد وُصف طويلًا بمصطلحات تجسيمية مثل «الذكاء» و«الإبداع»؟ ينطوي التشبيه التجسيمي على عيب حاسم. ففي الواقع، يفتقر الذكاء الاصطناعي إلى سمة مميزة أساسية للإدراك البشري، وهي التي تجعل الإنسان ذكيًا ومبدعًا حقًا. فعلى خلاف البشر، تركز أنظمة الذكاء الاصطناعي على العناصر الشائعة، وتهمل أو تنسى النادر وغير المألوف وغير المحتمل. وعلى الرغم من أن التعميمات جانب مهم من الإدراك البشري أيضًا، فإن الإبداع البشري ينبع من غير المحتمل والاستثنائي (Shumailov et al., 2023). والتركيز على غير المألوف والنادر هو ما يميز الإبداع والذكاء البشريين. في المقابل، تبالغ نماذج الذكاء الاصطناعي في تقدير الأحداث المحتملة، وتقلل من تقدير الأحداث غير المحتملة (Shumailov et al., 2024). وعبر الأجيال المتعاقبة، تدمر الأحداث المحتملة مجموعة البيانات، فيؤدي ذلك إلى إزالة الذيول، أي العناصر غير المحتملة لكنها حاسمة في مجموعة البيانات، والتي تسهم في دقة النموذج وتباين مخرجاته. وتميل نماذج الذكاء الاصطناعي إلى توليد متواليات تزداد احتمالًا من البيانات الأولية عبر الأجيال المتعاقبة، وتبدأ بإدخال متواليات غير المحتملة، أي أخطائها (Shumailov et al., 2024). ومع الزمن تتراكم الأخطاء ويُساء تفسيرها بشدة (Lutkevich, 2023). ونتيجة لذلك، يعجز الذكاء الاصطناعي عن إنتاج مخرجات تعكس حقًا ثراء التجربة البشرية وتعقيدها (Marr, 2024). ويكون مصير طبيعته التوليدية أن تصبح تدهورية. وتبدأ الفنون والعلوم وغيرها من الوسائط الإبداعية بتقليد نجاحات الصناعة على نحو نمطي، وتغدو متشابهة في مظهرها وإحساسها (Vue.ai, 2023).

ويصف لي أثر إضعاف القدرات العقلية للذكاء الاصطناعي أيضًا بأنه «خرف الذكاء الاصطناعي» (Goovaerts, 2024). فكما يحدث للبشر، حيث يجلب التقدم في العمر غالبًا اعتلالات عقلية، منها خرف ألزهايمر، يقع أمر مشابه لنماذج الذكاء الاصطناعي. إن عملية شيخوخة هذه النماذج، التي تُدرّب ويُعاد تدريبها بأعداد متزايدة من معلمات الإدخال، تقود فعليًا إلى اكتساب النموذج معرفة مفرطة في التعميم، إلى درجة يعجز معها عن استكشاف أي موضوع واحد بعمق.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

الحد من الذكاء الاصطناعي التدهوري في البحث

تبدو العملية التدهورية شاملة وحتمية في أي نظام ذكاء اصطناعي (Shumailov et al., 2024). ولذلك يبرز سؤال عما إذا كان ثمة علاج قابل للتطبيق لمنعها أو الحد منها (Borji, 2024). فهل ستواصل نماذج الذكاء الاصطناعي المستقبلية التحسن، أم أن مصيرها التدهور حتى تبلغ انهيارًا كليًا للنموذج؟ (Dohmatob et al., 2024) يتمثل تحدي نزاهة البحث في المستقبل في ضمان أن تعكس نماذج الذكاء الاصطناعي تعقيدات العالم الحقيقي، وأن تولد محتوى إبداعيًا، وأن تبعث الثقة بوصفها أدوات بحث موثوقة (Infobip, 2024).

وقد اقتُرحت إجراءات عدة لمنع انهيار النموذج، مثل تنويع بيانات التدريب، والإشراف البشري، وأنظمة المكافأة البديلة، والرصد الاستباقي؛ وهي إجراءات تعتمد جميعها، في نهاية المطاف، على «تنظيم البيانات» الفعال (Infobip, 2024). ويعني تنظيم البيانات أنه بدلًا من الاعتماد على مجموعات بيانات ضخمة ضعيفة الترشيح، ينبغي بذل جهد بشري لتنظيم بيانات الإدخال وانتقاء أمثلة تدريب متنوعة وتمثيلية. ومن دون هذا التنظيم، سيكون مصير الذكاء الاصطناعي التوليدي تضخيم التحيزات القائمة وفقدان التنوع في اللغة والصور (Hammond, 2024). ومن الضروري ضمان تدريب الذكاء الاصطناعي على بيانات بشرية بأعلى جودة. غير أن هذا أصبح تحديًا يكاد يستحيل تجاوزه (Marr, 2024). فالذكاء الاصطناعي يحتاج إلى البيانات البشرية، لكن العالم الرقمي يغمره على نحو متزايد سيل من البيانات الاصطناعية، مما يجعل التمييز بين المادة العضوية والاصطناعية أشد صعوبة، ويعقد تنظيم بيانات بشرية غير ملوثة لتدريب النماذج المستقبلية (Marr, 2024). ويبدو ترشيح البيانات الاصطناعية شبه مستحيل، وقد تصبح هذه الجهود أكثر مشقة على المدى الطويل، إذ يزداد التمييز بين النوعين صعوبة كل يوم (Snoswell, 2024).

ولا توجد، في الوقت الحاضر طريقة موثوقة للتمييز بين البيانات العضوية والاصطناعية أو لتتبع بيانات النماذج اللغوية الكبيرة على نطاق واسع (Lutkevich, 2023). لذلك يبرز السؤال: كم بقي من الوقت قبل أن تتلوث نماذج الذكاء الاصطناعي تلوثًا كاملًا (Marr, 2024)؟ وفي المستقبل القريب ستسارع شركات الذكاء الاصطناعي إلى اقتناء البيانات البشرية حيثما أمكن (Lutkevich, 2023). وإذا نظرنا إلى البيانات بوصفها سلعة نفيسة شبيهة بالنفط: «لا بيانات، لا ذكاء اصطناعي توليدي» (Eliot, 2024)، وإلى البحث عن البيانات بوصفه شبيهًا بملاحقة النفط، ظهر السؤال: متى ستنفد البيانات كلها؟ (Villalobos et al., 2022) ومتى نبلغ النقطة التي تكون عندها جميع البيانات العضوية المتاحة قد خضعت للمسح؟ (Eliot, 2024) ويرى بعضهم أن البيانات عالية الجودة قد تنفد في وقت مبكر لا يتجاوز سنة 2026، وأن مخزون البيانات منخفضة الجودة سينفد بحلول سنة 2030 (Villalobos et al., 2022). ومن المؤكد أن ذلك سيعوق الاتجاه الحالي نحو نماذج ذكاء اصطناعي متزايدة الحجم تقوم على مجموعات بيانات هائلة.

وعلى الرغم من أن هذا قد يبدو رأيًا متطرفًا إلى حد ما، إذ من غير المرجح أن يتوقف البشر عن إنتاج بيانات بشرية المنشأ (Eliot, 2024)، فإن المشكلة هي أننا نشهد بالفعل اتجاهًا يقل فيه عدد من يبدعون من دون مساعدة الذكاء الاصطناعي. ولذلك، ومع أن البشر سيواصلون دائمًا إنتاج بيانات جديدة، فإن هذه البيانات ستكون، على نحو لا مفر منه، مدعومة بالذكاء الاصطناعي بدرجة متزايدة.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

ويرى شومايلوف وزملاؤه (Shumailov et al., 2024) أن تدريب الذكاء الاصطناعي باستخدام البيانات الاصطناعية ليس مستحيلًا إذا أُدير بصورة سليمة (Nature Publishing Group, 2024). وتتمثل إحدى طرائق معالجة ذلك في وضع علامة مائية على المحتوى المولد بالذكاء الاصطناعي أو وسمه بطريقة ما، على النحو الوارد في التشريع الحكومي الأسترالي الحديث (Snoswell, 2024). والعلامات المائية إشارات مضمّنة غير مرئية، لكنها سهلة الكشف، وتتيح التعرف إلى المحتوى المولد بالذكاء الاصطناعي وإزالته من مجموعات بيانات التدريب (Wenger, 2024). غير أن العلامات المائية ليست جوابًا شاملًا، لأنها سهلة الحذف، كما أن تبادل معلوماتها يتطلب تنسيقًا واسعًا بين شركات الذكاء الاصطناعي.

وإلى أن تُحل قضية تنظيم البيانات تؤكد ظاهرة انهيار النموذج ميزة السبق في بناء نماذج الذكاء الاصطناعي التوليدي (Shumailov et al., 2024; Vue.ai, 2023). وبفضل ميزة السبق يُرجح أن تكون النماذج الأولى التي دُربت بالكامل على بيانات أنشأها البشر، أدق النماذج وأكثرها موثوقية مما سنحصل عليه على الإطلاق. أما النماذج المستقبلية فستصبح، حتمًا، تدهورية مع اعتمادها المتزايد على المحتوى المولد بالذكاء الاصطناعي في التدريب (Marr,2024). والنماذج المدربة على مجموعات بيانات مأخوذة من الإنترنت قبل تلوثه ببيانات الذكاء الاصطناعي تمثل العالم الحقيقي تمثيلًا أفضل؛ ولذلك سيكون من المثير للاهتمام أن نرى ما سيحدث للنماذج المستقبلية التي لا يمكن تدريبها على بيانات إنترنت نقية كهذه. **فهل أفضل نماذج الذكاء الاصطناعي هي، بالفعل، النماذج الموجودة بين أيدينا اليوم؟** (Wenger, 2024).

وسيكون تنظيم البيانات حاسمًا للحد من مخاطر الذكاء الاصطناعي التدهوري، وهو ما يبدو، من بعض الوجوه، مناقضًا لمفهوم الطبيعة التوليدية والمستقلة للذكاء الاصطناعي وفلسفتها (Hammond, 2024). وسيكون الإسهام البشري ضروريًا لاتخاذ قرارات مستنيرة بشأن البيانات التي ينبغي استخدامها في تدريب الذكاء الاصطناعي. وهذا يبين ضرورة تبني نهج أكثر تمركزًا حول الإنسان في تدريب الذكاء الاصطناعي، يؤكد أهمية الإشراف البشري والمساءلة (Hammond, 2024). **وكما يقول هاموند:** في حين تعتمد أنظمة الذكاء الاصطناعي الحالية اعتمادًا كبيرًا على الحجم الهائل للتحسن والتوسع، يبين الذكاء البشري أن الجودة، لا الكمية، قادرة على إنتاج عمليات تعلم أكثر فاعلية. ومن ثم، من الضروري إعادة النظر في مناهج تدريب الذكاء الاصطناعي، بحيث يصبح الذكاء التنظيمي جزءًا جوهريًا من تطويره. ويكمن التحدي الحقيقي، والفرصة الحقيقية أيضًا، في إنشاء أنظمة ذكية لا تنتج محتوى عالي الجودة فحسب، بل تكون كذلك قادرة على التمييز في شأن البيانات التي تتعلم منها. ومن خلال تعزيز هذا المستوى من الاستقلال والحكمة، يمكننا التعامل على نحو أفضل مع مخاطر الذكاء الاصطناعي ومكاسبه، وضمان أن تُثري هذه الأنظمة عالمنا بدلًا من أن تقلل تنوعه وعمقه (Hammond, 2024).

وبناء على ذلك، فإن صون نزاهة البحث ينطوي على الحد من إمكان انهيار النموذج وآثاره. ولا يمكن تحقيق ذلك إلا بالاحتفاظ بمجموعات البيانات الأصلية التي أنتجها البشر أساسًا لأنظمة الذكاء الاصطناعي، وبإدخال بيانات بشرية أصلية جديدة باستمرار. ويتيح هذا للمستخدمين التمييز بين المحتوى الذي يولده الذكاء الاصطناعي والمحتوى الذي ينشئه البشر (Romano, 2023).

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

4. الذكاء الاصطناعي التدهوري والمنظور الاجتماعي-التقني

يتمثل أحد المبادئ الحاسمة المتعلقة بتأثير الذكاء الاصطناعي في نزاهة البحث في مسؤولية ضمان بقائه متجذرًا في الواقع القائم على البشر (Sweenor, 2023). ويثبت تاريخ التقنية أنها لا تعمل أبدًا بالطريقة المقصودة تمامًا، وليست مجرد أداة محايدة لتحقيق غايات إيجابية (Williamson, 2023). وقد تطور الحماس الأولي للذكاء الاصطناعي، القائم على شعار «إذا كنا نستطيع، فلنفعل»، إلى نهج أكثر تبصرًا بشأن منافعه الفعلية: «مجرد قدرتنا على فعل شيء لا يعني أنه ينبغي لنا أن نفعله» (Currie, 2023). غير أن هذا يُدخلنا أكثر في المنظور الاجتماعي-التقني للذكاء الاصطناعي.

ووفقًا لكودينا وفان دي بويل (2024)، تقوم فكرتان محوريتان في هذا المنظور:

أولًا: التقنيات مكونات ضمن أنظمة أكبر، مثل نظام الطاقة أو نظام النقل، تعمل على مستويات متعددة: عالمية ووطنية وحضرية وغيرها.

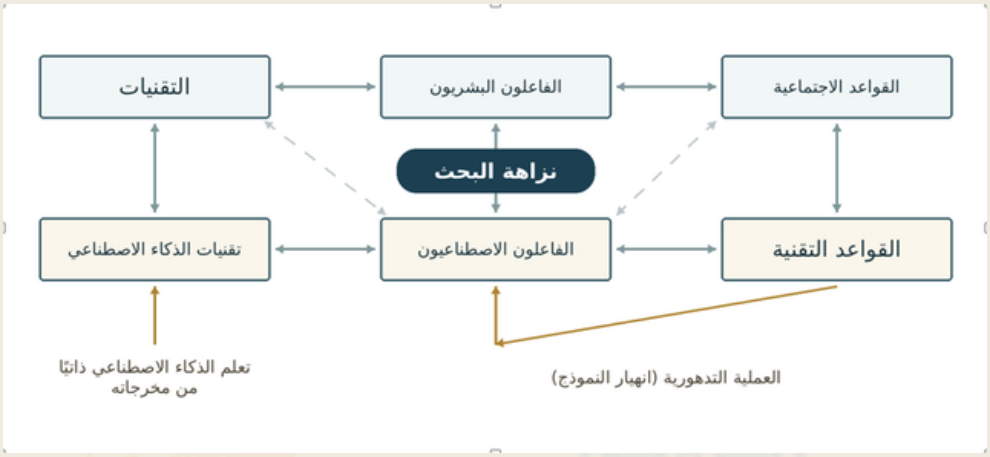
ثانيًا: العناصر البشرية أو الاجتماعية مكونات أساسية في النظام الاجتماعي-التقني. وبالمثل يرى كريستيانيني وزملاؤه (2023) أن أنظمة الذكاء الاصطناعي آلات اجتماعية تتكون من عناصر مادية وبشرية تتفاعل بطريقة منظمة. ولذلك كما يشدد دالك (2024)، ينبغي للتعريف الاجتماعي-التقني للذكاء الاصطناعي أن «ينص على غرض بشري واضح للذكاء الاصطناعي، فيضع التقنية ضمن نطاق المجتمعات المتمركزة حول الإنسان»، وأن «يسمح بمرونة التوقعات التي يحملها المستخدمون بشأن مقدار الاستقلال الذي تمارسه تقنية الذكاء الاصطناعي أثناء خدمتها غرضها».

أنظمة الذكاء الاصطناعي بوصفها نظمًا اجتماعية-تقنية

يرى كروس وزملاؤه (2006) أن النظم الاجتماعية-التقنية التقليدية تتكون من ثلاثة عناصر: التقنيات والفاعلون البشريون والمؤسسات الاجتماعية أي القواعد والقيود الاجتماعية. ولأن التقنية التقليدية لا تتعلم ذاتيًا، فإن هذه النظم لا تتكيف إلا من خلال الفاعلين البشريين. غير أن تقنية الذكاء الاصطناعي تتسم بالاستقلال والتفاعل والتكيف الذاتي؛ ولذلك، فإن أنظمة الذكاء الاصطناعي، بوصفها نظمًا اجتماعية-تقنية، تضم، وفقًا لفان دي بويل (2020)، عنصرين إضافيين (الشكل 1). وبما أنها تستطيع التعلم من بيئتها وتكييف نفسها فهي تشتمل على فاعلين اصطناعيين غير بشريين وغير قصديين. ويخضع هؤلاء الفاعلون لقواعد تقنية تنظم التفاعلات بينهم، بطريقة تشبه تنظيم القواعد الاجتماعية للتفاعلات بين الفاعلين البشريين.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

الشكل 1. نظام الذكاء الاصطناعي بوصفه نظامًا اجتماعيًا-تقنيًا ونزاهة البحث



ملاحظة: يمكن النظر إلى نزاهة البحث في النظم الاجتماعية-التقنية للذكاء الاصطناعي بوصفها حصيلة التفاعلات بين الفاعلين البشريين والفاعلين الاصطناعيين والقواعد الاجتماعية والتقنية. وقد تعطل العمليات التدهورية في تقنيات الذكاء الاصطناعي عناصر النظم الاجتماعية-التقنية، فتؤثر في الفاعلين الاصطناعيين والبشريين معًا، وتقوض المؤسسات والقواعد الاجتماعية، وتفضي إلى قضايا أخلاقية وقضايا تتعلق بنزاهة البحث.

غير أنه إذا أدخل الذكاء الاصطناعي مكونات جديدة كليًا إلى النظم الاجتماعية-التقنية، فلن يكون ذلك مجرد تغيير طفيف في المجتمع، بل قد يحول النظم الاجتماعية-التقنية القائمة تحولًا جوهريًا أشد من أي ابتكار تقني سابق (Kudina and van de Poel, 2024). وقد يكون للذكاء الاصطناعي أثر مقلق في المجتمع، ويمكن أن يغير بعمق ممارسات اجتماعية مثل التفاعل البشري والإبداع، ومؤسسات مثل النظم التعليمية (Kudina and van de Poel, 2024).

ويقدم النهج الاجتماعي-التقني للذكاء الاصطناعي منظورًا أكثر تكاملًا لقضاياه، ويؤكد أهمية ثلاثة جوانب تؤثر فيه: المؤسسات الاجتماعية، والثقافة، والحوكمة (Kudina and van de Poel, 2024). فالمؤسسات الاجتماعية تحدد وتنظم كيفية تفاعل تقنية الذكاء الاصطناعي والبشر. وتنعكس الثقافة، بوصفها بناءً من المعتقدات والتوقعات الاجتماعية، في مجموعات البيانات التي يُدرّب عليها الذكاء الاصطناعي. غير أن الذكاء الاصطناعي، لكونه متكيفًا ذاتيًا، لا يبني على البنى الثقافية فحسب، بل يتحداها أيضًا. ولا تشير الحوكمة إلى المبادئ الأخلاقية لاستخدام الذكاء الاصطناعي وحدها، بل تبرز كذلك جوانب السياسة الأوسع. ويرى النهج الاجتماعي-التقني أن معالجة القدرة الاضطرابية للذكاء الاصطناعي تتطلب أكثر من الاعتبارات الأخلاقية؛ إنها تتطلب السياسة، أي طيفًا من القرارات التقنية والاجتماعية والاقتصادية والحوكمية المنسقة (Kudina and van de Poel, 2024).

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وضمن المنظور الاجتماعي-التقني، لا يعود النظر إلى القضايا الأخلاقية والمعنوية المتعلقة بالذكاء الاصطناعي بوصفها شؤونًا بشرية خالصة، بل بوصفها شأنًا هجينًا يتضمن تفاعلات البشر مع بيئتهم الاجتماعية-الثقافية والمادية الأوسع. وفي إطار النهج الاجتماعي-التقني، ينبغي تناول الأخلاق بوصفها خاصية دينامية تتجسد ضمن نظام معقد من أبعاد الإنسان والتقنية والعالم (Kudina and van de Poel, 2024).

ويترتب على أطروحة الطبيعة المتشابكة للأخلاق مع البيئة الاجتماعية-المادية أنه، على الرغم من بقاء الناس فاعلين أخلاقيين ينظرون في الأوضاع ذات الإشكالات الأخلاقية ويتخذون قرارات بشأن أنظمة الذكاء الاصطناعي، فإن أنظمة الذكاء الاصطناعي ذاتها ليست منفصلة عن عملية اتخاذ القرار الأخلاقي لدى الناس. (Kudina and van de Poel, 2024) والعلاقة بين القيم والتقنية تفاعلية لا أحادية الاتجاه، لأن للقيم طبيعة مزدوجة: فهي، من جهة، تقدم توجيهًا أخلاقيًا في الأوضاع الاجتماعية-التقنية الإشكالية؛ ومن جهة أخرى، هي أيضًا نتاج للممارسات الاجتماعية-التقنية، التي قد تتحدى القيم القائمة، وتراجعها، وتقترح قيمًا جديدة (Van den Poel and Kudina, 2022).

متطلبات الثقة وأنظمة الذكاء الاصطناعي (التوليدية/التدهورية)

يبرز، ضمن المنظور الاجتماعي-التقني، السؤال الآتي بشأن الذكاء الاصطناعي التدهوري ونزاهة البحث: كيف يمكن النظر إلى قضايا نزاهة البحث والطبيعة التدهورية للذكاء الاصطناعي من منظور اجتماعي-تقني، وما الذي قد يعنيه النهج الاجتماعي-التقني بشأن تأثير الطبيعة التدهورية للذكاء الاصطناعي في نزاهة البحث؟

تمثل الطبيعة التدهورية للذكاء الاصطناعي من منظور اجتماعي-تقني، أثرًا اضطرابيًا واضحًا في النظم الاجتماعية-التقنية وفي المجتمع عمومًا. ولأن أنظمة الذكاء الاصطناعي تتضمن فاعلين اصطناعيين يحاكون الذكاء والأخلاق، فإن انهيار نموذج تقنية الذكاء الاصطناعي يؤثر فيهم، أي يؤدي إلى تدهورهم، وينتج، كما تقدم، «خرف الذكاء الاصطناعي» وأثر إضعاف قدرات هؤلاء الفاعلين. وعلى الرغم من أن ذكاء هؤلاء الفاعلين الاصطناعيين وأخلاقهم ليسا إلا محاكاة، فإن تدهورهم قد يؤثر أيضًا في الذكاء والأخلاق الحقيقيين للفاعلين البشريين، لأن الفاعلين الاصطناعيين والبشريين يتفاعلون داخل النظام الاجتماعي-التقني للذكاء الاصطناعي (انظر الشكل 1). ولذلك قد يدفع انهيار الفاعلين الاصطناعيين الفاعلين البشريين أيضًا إلى التدهور المعرفي والأخلاقي و«الانهيار»، فتنتج قضايا أخلاقية وأخرى متعلقة بنزاهة البحث.

غير أن الباحثين لن «يتدهوروا» إلا إذا وثقوا ثقة عمياء بمخرجات الذكاء الاصطناعي. ويبدو هذا غير مرجح بالنظر إلى الوعي النقدي والتدريب على نزاهة البحث الذي يتلقاه الباحثون. فالفاعلون الاصطناعيون والبشريون، من باحثين وطلبة وغيرهم، يتفاعلون ديناميًا داخل النظم الاجتماعية-التقنية للذكاء الاصطناعي؛ والفاعلون البشريون، مثل الباحثين، ليسوا مستخدمين سلبيين قد تتدهور معرفتهم وأخلاقهم بسبب عطل تقني في الذكاء الاصطناعي، بل هم أفراد أذكياء وواعون أخلاقيًا، ومدربون على مبادئ نزاهة البحث التي تهيئهم للتعامل بمسؤولية مع تقنية الذكاء الاصطناعي.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

ويشير هذا، من منظور اجتماعي-تقني، إلى أنه على الرغم من احتمال تدهور الذكاء الاصطناعي بمرور الزمن، فإن ذلك لن يؤدي إلى تراجع الباحثين؛ بل سيدفعهم إلى البحث عن بدائل للذكاء الاصطناعي التوليدي، أو حتى العودة إلى مصادر أكثر موثوقية، ولا سيما بالنظر إلى عدم جدارة المعلومات المولدة بالذكاء الاصطناعي بالثقة.

غير أن ذلك قد يشير إلى التحول الأهم والأكثر جوهرية في مواقف الفاعلين البشريين داخل النظم الاجتماعية-التقنية للذكاء الاصطناعي إزاء ما يسمى «متطلبات الثقة». وكما يؤكد روك وزملاؤه (2025)، تُعد متطلبات الثقة حاسمة لحسن عمل النظم الاجتماعية-التقنية، وتتحقق عندما تُستوفى معايير الدقة والأمن والخصوصية، مما يعزز تحقق الثقة. ففي النظم الاجتماعية-التقنية التقليدية، كانت الثقة قيمة أساسية في التفاعلات بين الفاعلين البشريين والتقنيات والمؤسسات الاجتماعية. وقد اعتاد الناس الوثوق بوسائل الإعلام والاتصال والمعلومات التي توفرها التقنية والمجتمع. غير أن هذا تغير تغيرًا كبيرًا مع الذكاء الاصطناعي التوليدي. وتعكس العبارة المأثورة من قصيدة ألكسندر بوب لسنة 1709، «إن قدرًا قليلًا من المعرفة شيء خطير»، أخطار الثقة الزائفة بقدرات الذكاء الاصطناعي أيضًا (Currie, 2023). ففي البداية بدا، مع الذكاء الاصطناعي التوليدي، أن المعرفة والإبداع اللامحدودين أصبحا في متناولنا. لكن ذلك اتضح أنه مجرد وهم، ضخمّه «تأثير دانينغ-كروغر» (Currie, 2023)، إذ يمنحنا الذكاء الاصطناعي ثقة زائفة تدفعنا إلى المبالغة في تقدير قدراتنا ومعارفنا. وقد مكّننا من أن نصبح شبه خبراء في أي مجال، غير أن المعلومات ليست معرفة بعد كما ذكّرنا ألبرت أينشتاين ذات مرة (Currie, 2023).

ولذلك ينطوي الذكاء الاصطناعي على خطر «أن يقود الأعمى أعمى آخر»، وهو لا يستوفي متطلبات الثقة في حالات كثيرة. ويدفع ذلك الناس إلى عدم الثقة به، ومن ثم عدم الثقة بالمجتمع الذي يتغلغل فيه. وهكذا قلب الذكاء الاصطناعي تمامًا الثقة التي صاحبت النظم الاجتماعية-التقنية تقليديًا. ففي أنظمة الذكاء الاصطناعي، لا يمكن للمرء أن يثق بأي شيء يراه أو يسمعه أو يقرأه أو يُنقل إليه. وبذلك نقلت الطبيعة التدهورية للذكاء الاصطناعي إدراك الواقع الاجتماعي من الثقة إلى انعدام الثقة.

وإلى ماذا يمكن أن يقود ذلك لاحقًا؟ ربما إلى نوع من رد الفعل المضاد للعولمة الرقمية. فكما يشير Currie، لا توفر المعرفة الحقيقية إلا المصادر الأولية، أما المصادر الثانوية، ومنها الذكاء الاصطناعي، فلا تقدم، في أفضل الأحوال، إلا معلومات، وفي أسوأها معلومات مضللة (Currie, 2023). وبما أننا نعجز عن الثقة بأي شيء يتجاوز المصادر الأولية وتجاربنا المادية المباشرة وتفاعلاتنا الاجتماعية، فقد يبدأ المجتمع بإعادة توجيه نفسه نحو المصادر الأولية والمجتمعات المحلية والخبرة العملية المباشرة. وهذا يؤثر بالفعل في مؤسسات مثل التعليم. فلأن الأساتذة لا يستطيعون الوثوق بالأعمال التي ينتجها الطلبة في منازلهم، بدأ النظام التعليمي يبتعد عن الواجبات والمقالات المنزلية، ويتجه نحو كتابة المقالات داخل الصف في دفاتر الامتحان، والامتحانات الشفوية، وساعات الحضور المكتبي الإلزامية، وغيرها من الواجبات التي تطلب من الطلبة إظهار معرفتهم في الوقت الفعلي (Shirky, 2025). إن التعليم يعود ببساطة إلى طرائق تقليدية، مثل التحدث والاستماع والقراءة، ثبت أنها أسس للثقافة الأكاديمية منذ نشأتها.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وقد ينشأ وضع مشابه في استخدام الذكاء الاصطناعي في البحث. فما جدوى استخدام الذكاء الاصطناعي إذا كان معروفًا بإنتاج نتائج زائفة، أو الهلوسة، أو تمثيل الجماعات تمثيلًا خاطئًا، أي إذا كان يتدهور؟ عند استخدام الذكاء الاصطناعي في البحث، يلزم قدر كبير من التدقيق اللغوي والتحقق من الحقائق، وهذا يخلق مفارقة. فنحن نسعى إلى استخدامه للحصول على مساعدة سريعة، لكن هذه الميزة تتضاءل لأن التحقق من مخرجاته يحتاج إلى وقت طويل. وعلى الرغم من إمكان أن يكون الذكاء الاصطناعي مفيدًا، فإن فائدته محدودة بدرجة كبيرة بسبب ميله إلى الخطأ وضرورة التحقق الدقيق منه. فهل يعني ذلك، في النهاية، أن الاستخدام السليم للذكاء الاصطناعي، وفقًا لجميع معايير نزاهة البحث، يتطلب في الواقع وقتًا والتزامًا أكبر من عدم استخدام أدواته أصلًا، **ومن ثم لا يستحق العناء؟**

وقد تقود هذه الاعتبارات إلى وضع يبدو متناقضًا، وتزيد إبراز أهمية نزاهة البحث داخل المجتمع البحثي. فهل سيشجعنا الذكاء الاصطناعي حقًا على قضاء وقت أطول في إنجاز عمل أفضل، أم أن الحقيقة القاسية هي أنه لن يفعل سوى دفع المستخدمين إلى إنتاج مزيد ومزيد من الأوراق المزيفة، بدءًا من تقارير تلاميذ المدارس ووصولًا إلى الأوراق العلمية والروايات وغيرها (Vaughan-Nichols, 2025)؟ ولذلك يرى بعضهم أن توقع الاستخدام المسؤول للذكاء الاصطناعي وهم، ويجادلون بأن عبارة «مستخدم مسؤول للذكاء الاصطناعي» تناقض لفظي (Vaughan-Nichols, 2025). غير أن الباحثين، إذا تعاملوا مع نزاهة البحث بجدية، سيترددون تلقائيًا في الاعتماد على الذكاء الاصطناعي. ولأنه غير جدير بالثقة، يتضح أن النهج الأكثر أمانًا هو تجنب الإفراط في الاعتماد عليه أو الإفراط في استخدامه. وإذا أراد شخص أن يكون باحثًا «جيدًا»، فسوف يلتزم بمعايير نزاهة البحث التي قد تتعارض مع استخدام الذكاء الاصطناعي، لأن عدم الالتزام بها قد يؤدي إلى عواقب اجتماعية سلبية لا تقتصر على التبعات القانونية، بل تشمل الضرر بالسمعة أيضًا. ولذلك، كلما بدا أن الذكاء الاصطناعي التدهوري يهدد نزاهة البحث بدرجة أكبر، زاد احتمال أن يقلل الباحثون استخدامهم له، مثل استخدامه في كتابة الأوراق، لأنه غير جدير بالثقة، ولأن المراجعين، في مراجعات الأوراق والمؤتمرات وغيرها، سيرفضون ببساطة الأوراق التي ينتجها الذكاء الاصطناعي لانخفاض جودتها. ومع أن هذا قد لا يكون واقع الحال بالضرورة في الحاضر، فإنه سيكون كذلك بالتأكيد إذا تدهور الذكاء الاصطناعي.

الإدارة الاستباقية لمخاطر الذكاء الاصطناعي التدهوري ونزاهة البحث

يتيح النهج الاجتماعي-التقني فهمًا أفضل لكيفية عمل أنظمة الذكاء الاصطناعي، وللكيفية التي يمكن أن تحول بها المجتمع ومؤسسات حاسمة مثل الديمقراطية (Kudina and van de Poel, 2024). ويرى بعضهم أن لوائح الذكاء الاصطناعي، شأنها شأن لوائح المنشطات في الرياضة، تصبح متقادمة بحلول الوقت الذي تتحول فيه إلى سياسة رسمية، لأنها تعجز عن استباق الأضرار المستقبلية (Bockting et al., 2023). ويبدو المخالفون دائمًا متقدمين بخطوة على جهات الملاحقة. غير أن لوائح مثل قانون الذكاء الاصطناعي تُحدث بالفعل أثرًا اجتماعيًا-تقنيًا في ممارسات الشركات، كما يتضح من حملات الضغط المكثفة التي تمارسها الشركات ضدها (Kroet, 2025).

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وتُعد القواعد أو اللوائح الاجتماعية، مثل قوانين الذكاء الاصطناعي، جزءًا من إدارة المخاطر، ولها مزايا وعيوب. وعلى الرغم من أن لوائح الذكاء الاصطناعي قد لا تكشف الآثار السلبية المحتملة كلها، فإنها تعالج كثيرًا من المخاطر الاجتماعية-التقنية التي يُرجح أن تكون مهمة في المستقبل. غير أن كيسليش وزملاءه (2025)، ومن منظور اجتماعي-تقني، ينتقدون أساليب التقويم القائمة والنهج التنظيمية لإدارة مخاطر الذكاء الاصطناعي، مثل قانون الذكاء الاصطناعي، ويحددون أحد عشر قصورًا في هذه الإدارة: غياب المساءلة الديمقراطية، ونزع الطابع السياسي عن التقويم، وغياب الشمول، وتحيز الخبراء، ومحدودية الإرشاد في قرارات الموازنة بين البدائل، واستبعاد «المجاهيل المجهولة»، والتركيز على إصلاح المخاطر، والتركيز على الامتثال، والتركيز على القياس الكمي، وغياب مراعاة السياق، والتركيز على التقنية بوصفها المصدر الرئيس للمخاطر. وبدلًا من ذلك، يدعون «الباحثين وصناع السياسات وصناعة التقنية إلى الانخراط في نهج استباقية تتجاوز الامتثال وتمارين قوائم التحقق، وتقر بالمسؤولية التي يتحملونها عن آثار تتجاوز ما هو معروف» (Kieslich et al., 2025).

وتُعد إدارة المخاطر أو تقويمها استباقيًا أمرًا حاسمًا من منظور اجتماعي-تقني (Kieslich et al., 2025)، لأنها قد تمنع فقدان الثقة في المجتمع. وتركز النهج التقليدية لإدارة المخاطر على الجوانب الكمية؛ فتعالج «المعلومات المعروفة» (مخاطر لها احتمالات إحصائية وآثار قابلة للقياس) و«المجاهيل المعروفة» (أوجه عدم يقين يتعذر قياسها لأن آثار تقنية معينة غير معروفة)، لكنها تستبعد «المجاهيل المجهولة». فمثلًا، لا يُلزم قانون الذكاء الاصطناعي المطورين إلا بالتعامل مع مخاطر «الاستخدام المقصود» لأنظمة الذكاء الاصطناعي (Kieslich et al., 2025). ومن ثم لا يكون تقويم المخاطر هذا شاملًا، ولا يعالج العوامل النظامية الأوسع، ولا يستطيع حساب الفروق الفردية أو الجماعية، ولا يعالج تعارض القيم. في المقابل، يشدد النهج الاستباقي على التعامل النوعي مع «المجاهيل المجهولة»، وهي المخاطر التي تنشأ عندما لا نكون واعين أصلًا بأن أشياء أو أنشطة معينة قد تترك آثارًا ضارة. ولمراعاة هذه «المجاهيل المجهولة»، نحتاج إلى نهج أكثر شمولًا وأساليب بديلة ذات طبيعة نوعية لا تعتمد حصرًا على القياس الكمي.

وفيما يتعلق بنزاهة البحث، يتمثل السؤال في الكيفية التي يمكن أن تسترشد بها الإدارة الاستباقية للمخاطر في الاستخدام المسؤول للذكاء الاصطناعي في البحث. ولأن نهج إدارة المخاطر الاستباقية ينتقل من الكشف السلبي عن الاحتيال إلى الجهود الاستباقية الرامية إلى تحديد المخاطر قبل وقوعها، ولا سيما المخاطر التي تُعد «مجاهيل مجهولة»، فإنه قد يكشف كيف يمكن للذكاء الاصطناعي أن يتدخل على نحو خفي في نزاهة البحث بطرائق غير متوقعة، حتى قبل ظهور هذه القضايا. وقد ترتبط هذه «المجاهيل المجهولة»، فيما يتصل بنزاهة البحث، على نحو خاص بقضايا «المنطقة الرمادية» أو «الممارسات البحثية المشكوك فيها» التي نوقشت آنفًا. وبما أن هذه الممارسات ليست صوابًا أو خطأً على نحو واضح، بل تعتمد على السياق، فمن الصعب التنبؤ بما إذا كان استخدام الذكاء الاصطناعي سيعود بالنفع أو الضرر على البحث، ومتى يحدث ذلك. فكما تقدم، تستطيع مولدات النصوص، مثلًا، تمكين البحث من جهة، وتسهيل الاحتيال من جهة أخرى.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

وفي حالات الممارسات البحثية المشكوك فيها، يمكن استخدام الأساليب القائمة على السيناريو بفاعلية لاستكشاف المعضلات الأخلاقية في البحث. فقد بيّن انتس وزملاؤه (2009) أن أكثر البرامج التعليمية نجاحًا في نزاهة البحث كانت «قائمة على الحالات وتفاعلية وتتيح للمشاركين تعلم تطبيق مهارات اتخاذ القرار الأخلاقي في العالم الحقيقي والتدرب عليها». وتتمثل فائدة الأساليب القائمة على السيناريو في أنها تصف مستقبلات متعددة وبديلة ومعقولة، وتتناول خبرات فئة متنوعة من المستخدمين، وتشركهم، بمن فيهم غير الخبراء، في التفكير الاستباقي في تطورات مستقبلية متنوعة (Kieslich et al., 2025). وبذلك يمكنها الإقرار بأوجه عدم اليقين، أي «المجاهيل المجهولة»، الكامنة في المستقبل.

ولذلك يمكن استخدام نهج «التصور الاجتماعي-التقني القائم على السيناريو» (Scenario-based Sociotechnical Envisioning: SSE)، الذي اقترحه كيسليش وزملاؤه (2025) بوصفه أسلوبًا للحوكمة الاستباقية للمخاطر، استخدامًا فعالًا لتوجيه الاستخدام المسؤول للذكاء الاصطناعي في البحث. والجانب الحاسم الذي يميز هذا النهج عن غيره من أساليب إدارة المخاطر الاستباقية هو استخدامه السرديات أو السيناريوهات لتصور القضايا المعقدة المتعلقة باستخدام الذكاء الاصطناعي في البحث بطريقة ملموسة. ويمكن تطوير هذه السيناريوهات من خلال مشاهد قصيرة، ودراسات حالة، ولعب أدوار، وتقديمها في صورة معضلات. كما يمكن استخدام المعضلات لأغراض تعليمية في صورة ألعاب معضلات، مثل اللعبة التي طورتها جامعةإيراسموس روتردام (2020) لنزاهة البحث، أو اللعبة المخصصة للذكاء الاصطناعي التي طورها كيم وزملاؤه (2025). ولعبة المعضلة نوع من التجارب الفكرية يواجه فيها الباحثون سيناريوهات محاكية للحياة الواقعية، ويجب عليهم الاستجابة لها أخلاقيًا. وتكون النهج القائمة على السيناريو، مثل ألعاب المعضلات، مفيدة بصورة خاصة في الأوضاع المعقدة، والشديدة الارتباط بالسياق، والصعبة الاستباق أو التنبؤ، مثل الأوضاع الواقعة في «المنطقة الرمادية» لنزاهة البحث. وبهذه الطريقة تعالج أوجه القصور في أساليب التقويم والنهج التنظيمية القائمة، ولا سيما من خلال التشديد على السياق والقيم المجتمعية والعوامل البشرية، وهي مصادر مخاطر شديدة الصلة بنزاهة البحث. ومن خلال استخدام نهج مثل التصور الاجتماعي-التقني القائم على السيناريو، يمكن أن تصبح تقويمات نزاهة البحث أكثر شمولًا، وأكثر استجابة للسياق الاجتماعي-التقني الذي تُنشر فيه أنظمة الذكاء الاصطناعي، وأن تساعد على تحديد قضايا نزاهة البحث قبل ظهورها.

خاتمة

آفاق المستقبل

على الرغم من أن الذكاء الاصطناعي التوليدي أنتج مخرجات أصيلة وإبداعية على نحو لافت في العلم والفن معًا، فإن تدهوره قد يقوض إمكاناته (Vue.ai, 2023). وإذا دُربت نماذج الذكاء الاصطناعي على بيانات مولدة به، فقد نشهد تدهورًا في البحث مستقبلًا (Marr, 2024). ويمكن للذكاء الاصطناعي التدهوري أن يهدد البحث ونزاهته تهديدًا حاسمًا، لأنه يقوض الجودة، بما في ذلك الصرامة الفكرية والإبداع والاستثنائية؛ والموثوقية، مثل تزييف المعلومات؛ والإنصاف، من خلال إبراز التحيزات والصور النمطية؛ والتمثيل، من خلال تمثيل

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

الأحداث النادرة والفئات المهمشة ووجهات النظر تمثيلًا خاطئًا أو استبعادها. كما يثير شواغل أخلاقية، مثل تقويض اتخاذ القرار القائم على معلومات منخفضة الجودة ومتحيزة، وقد يترك آثارًا اقتصادية واجتماعية كبيرة، مثل خفض ثقة الجمهور بتقنيات الذكاء الاصطناعي (Sweenor, 2023).

فما الذي يمكن فعله لمنع الذكاء الاصطناعي التدهوري أو الحد منه؟

ثمة جانب تقني، يتمثل عامله الأساس في كيفية تدريب نماذج الذكاء الاصطناعي. وعلى الرغم من إغراء الاعتماد على البيانات الاصطناعية، لأنها أسهل وأرخص في الحصول عليها، يشدد ماري (2024) على وجوب مقاومة هذا الإغراء. ولضمان بقاء نماذج الذكاء الاصطناعي دقيقة وذات صلة، ينبغي أن نواصل تدريبها على خبرات بشرية أصيلة، مع احترام حقوق الأفراد الذين تُستخدم بياناتهم. كما ينبغي تحديث نماذج الذكاء الاصطناعي بانتظام من خلال تعريضها لبيانات بشرية المنشأ جديدة، ويحتاج مجتمع الذكاء الاصطناعي إلى مزيد من التعاون في مشاركة مصادر البيانات واستراتيجيات التدريب للمساعدة على منع إعادة تدوير البيانات (Marr, 2024).

ويمثل انهيار النموذج تحديًا، لكن كثيرين يرون إمكان معالجته بإبقاء الذكاء الاصطناعي متجذرًا في الواقع الذي أنشأه البشر (Marr, 2024). ومع ذلك، قد لا يكون من الضروري تجنب البيانات الاصطناعية بالكامل عند تدريب نماذج الذكاء الاصطناعي المستقبلية. فقد بيّن جيرستجراسر وزملاؤه (2024) أنه إذا جُمعت البيانات الاصطناعية إلى جانب البيانات الحقيقية، بدلًا من مجرد إحلالها محل البيانات الحقيقية، أمكن للنوعين العمل معًا؛ إذ تؤسس البيانات العضوية البيانات الاصطناعية في الواقع، وتحدد أي البيانات الاصطناعية ينبغي الاحتفاظ به وأيها ينبغي استبعاده. ويعني هذا إمكان إدارة الحجم الكبير من البيانات الاصطناعية باستخدام قدر صغير فقط من البيانات العضوية. إذ يمكن استخدام جزء صغير من البيانات العضوية لترشيح كميات هائلة من البيانات الاصطناعية أو إثرائها (Eliot, 2024). غير أن ذلك يفترض تمييزًا واضحًا بين البيانات العضوية والاصطناعية، وهو أمر قد يكون إشكاليًا؛ فالتمييز بين النوعين يكاد يكون مستحيلًا حتى اليوم، حين لا نزال واعين بالفرق، كما في قدرتنا على تحديد ما إذا كانت لوحة ما مزيفة أو أصلية لـ بيكاسو، فكيف بالمستقبل حين قد يختفي تمامًا الفرق بين البيانات الأصلية والاصطناعية؟ إن اشتراط فصل بيانات الذكاء الاصطناعي عن البيانات العضوية يثير قضايا بشأن أصول البيانات على الإنترنت، إذ لا يزال من غير الواضح كيف يمكن تتبع البيانات الاصطناعية على نطاق واسع (Shumailov et al., 2024).

وما يمنحنا الأمل هو أن ظاهرة شبيهة بانهيار النموذج عولجت بنجاح قبل وقت طويل من الانتشار الواسع لتقنيات مثل ChatGPT (Lutkevich, 2023; Shumailov et al., 2024). فقد حاولت «مزارع المحتوى» سنوات طويلة التأثير في خوارزميات البحث لتغيير كيفية ترتيبها للمعلومات. ودفع الأثر الضار لمحاولات التسميم هذه في نتائج البحث إلى مراجعة خوارزميات البحث. فمثلًا، تفضل Google المحتوى الصادر من مصادر موثوقة، مثل النطاقات التعليمية، وتخفض قيمة المعلومات التي تبدو منخفضة الجودة أو صادرة من مزارع المحتوى. وهذا يوحي بوجود أمل في تنظيم فعال لبيانات نماذج الذكاء الاصطناعي مستقبلًا أيضًا.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

غير أن الجانب التقني ليس إلا جزءًا من صورة أكبر. فإلى جانب الاعتبارات التقنية، ينبغي أيضًا معالجة الجوانب الاجتماعية-التقنية من خلال حوكمة استباقية للمخاطر، تكون شاملة ومتوافقة مع السياق الاجتماعي-التقني الذي تُنشر فيه أنظمة الذكاء الاصطناعي، وتساعد على تحديد الآثار قبل وقوعها. وبدلًا من الاقتصار على رد الفعل، تهدف النهج الاستباقية إلى تحديد الآثار المستقبلية والاستجابة لها استباقيًا من أجل الحفاظ على ثقة الجمهور بالمجتمع. وفي إطار هذا الأسلوب الاستباقي، اقترحت مجموعة من العلماء والمؤسسات الشريكة، في تشرين الأول 2023، مجموعة من «المبادئ التوجيهية الحية» لاستخدام الذكاء الاصطناعي التوليدي، وحددت المساءلة والشفافية والإشراف المستقل مبادئ أساسية لاستخدامه (Bockting et al., 2023). وبناء على هذه المبادئ، اقترحت إنشاء مجلس علمي لتقويم سلامة أنظمة الذكاء الاصطناعي وصلاحيتها وأخلاقياتها، قبل أن يضر الذكاء الاصطناعي بالعلم وثقة الجمهور.

وقد أصبحت المسؤولية في استخدام الذكاء الاصطناعي موضع تأكيد وفهم جيدين داخل المجتمع البحثي. ويدرك الأكاديميون وجوب التحقق السليم من مخرجات الذكاء الاصطناعي، وأن المنشورات عالية الجودة لا يمكن إنتاجها إلا بأدنى قدر من دعمه. ولذلك لا يتمثل السؤال الأكثر أهمية لمستقبل البحث في ما إذا كان الذكاء الاصطناعي التوليدي سيؤثر في المجتمع، بل في كيفية حدوث تأثير الذكاء الاصطناعي التدهوري، والمجالات التي يكون فيها هذا التأثير أشد بروزًا وضررًا.

ومن الواضح أن جواب هذا السؤال يعتمد بدرجة كبيرة على السياق، لأن تأثير الذكاء الاصطناعي التدهوري يختلف تبعًا لعوامل كثيرة. فهو يعتمد على مجال البحث؛ إذ تستفيد بعض العلوم من الذكاء الاصطناعي أكثر، في حين قد تتعرض علوم أخرى لضرر أكبر. فمثلًا، يكون الذكاء الاصطناعي التوليدي أكثر قابلية للتكيف في علوم الحاسوب، بينما قد يكون استخدامه أشد إشكالية في العلوم الإنسانية والاجتماعية. ويعتمد الأمر أيضًا على الثقافة الأكاديمية في بلد أو نظام معين؛ إذ يتركز ضغط النشر في بعض البيئات الأكاديمية على إنتاج عدد كبير من المنشورات، ويشجع النشر السريع، في حين يكون التركيز في بيئات أخرى أقل على الكم وأكثر على المنشورات عالية الجودة. ومع وضع ذلك في الحسبان، يطور الناشرون أيضًا ممارسات مختلفة، ونشهد تزايدًا في البحوث منخفضة ومتوسطة الجودة التي تغمر النظام. وقد يسبب ذلك مزيدًا من المشكلات في البحث في الأدبيات، لأن المنشورات منخفضة الجودة سيُستشهد بها في بحوث لاحقة، مما يضخم التأثير التدهوري للذكاء الاصطناعي في نزاهة البحث.

وخلاصة القول، لا شك في أن تقنيات الذكاء الاصطناعي تترك تأثيرًا حاسمًا، وربما اضطرابيًا وتدهوريًا، في البحث ونزاهته. غير أنه، كما يؤكد كيسليش وزملاؤه (2025)، قد يبقى لدينا وقت لإدارة هذه العملية بمسؤولية إذا انخرط الباحثون وصناع السياسات وصناعة التقنية استباقيًا في نهج توقعية، وتولوا زمام الفعل في تشكيل مستقبل الذكاء الاصطناعي. ومع ذلك، يبدو قول هذا أسهل من فعله، لأننا اعتدنا بالفعل، وربما على نحو لا رجعة فيه، على الخدر المريح الذي وفره لنا الذكاء الاصطناعي.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

الإفصاحات

المصالح المتنافسة: يصرح المؤلف بعدم وجود مصالح متنافسة.

توافر البيانات: تنطلق الدراسة من نهج نظري؛ ولذلك لا تنطبق مشاركة البيانات على هذا البحث، إذ لم تُولد بيانات ولم تُحلل.

الموافقة الأخلاقية: لا تنطبق، لعدم وجود مشاركين بشريين.

الموافقة المستنيرة: لا تنطبق، لعدم وجود مشاركين بشريين.

بيان التمويل: لم يتلق البحث أي تمويل.

إسهامات المؤلف: البحث من تأليف فردي. صاغ S .J الفكرة، وراجع الأدبيات، وطوّر النظرية، وأعد المخطوطة.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

المصادر

1. ACM US Technology Policy Office (2023). Principles for the development, deployment, and use of generative ai technologies. https://www.acm.org/binaries/content/assets/public-policy/ustpc-approved-generative-ai-principles

2. Alam, S. (2024). Trends in research integrity concerns and the evolving role of the publisher. Insights: the UKSG journal 37(13). https://doi.org/10.1629/uksg.663

3. Alemohammad, S., Casco-Rodriguez, J., Luzi, L., Humayun, A. I., Babaei, H., LeJeune, D., Siahkoohi, A., & Baraniuk, R. G. (2023). Self-Consuming Generative Models Go MAD. ArXiv. https://doi.org/10.48550/arXiv.2307.01850

4. Allea. (2023). The European Code of Conduct for Research Integrity. https://allea.org/code-of-conduct/

5. Al-Zahrani, A. M. (2024). The impact of generative AI tools on researchers and research: Implications for academia in higher education. Innovations in Education and Teaching International, 61(5), 1029–1043. https://doi.org/10.1080/14703297.2023.2271445

6. Altman, S. (2024). X. https://x.com/sama/status/1756089361609981993?lang=en

7. Andersen, J. P., Degn, L., Fishberg, R., Graversen, E. K., Horbach, S. P. J. M., Schmidt, E. K., Schneider, J. W., & Sørensen, M. P. (2025). Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions. Technology in Society, 81, 102813. https://doi.org/10.1016/j.techsoc.2025.102813

8. Andrade-Hidalgo, G., Mio-Cango, P. & Iparraguirre-Villanueva, O. (2025). Exploring the Impact of Artificial Intelligence on Research Ethics - A Systematic Review. J Acad Ethics 23, 1053–1070. https://doi.org/10.1007/s10805-024-09579-8

9. Antes, A. L., Murphy, S. T., Waples, E. P., Mumford, M. D., Brown, R. P., Connelly, S., & Devenport, L. D. (2009). A Meta-Analysis of Ethics Instruction Effectiveness in the Sciences. Ethics & Behavior, 19(5), 379–402. https://doi.org/10.1080/10508420903035380

10. Austin, D. (2024). Degenerative AI & what happens when AI trains on AI data: artificial intelligence trends. eDiscoveryToday. https://ediscoverytoday.com/2024/08/26/degenerative-ai-what-happens-when-ai-trains-on-ai-data-artificial-intelligence-trends/

11. Bhatia, A. (2024). When A.I.'s output is a threat to A.I. itself. New York Times https://www.nytimes.com/interactive/2024/08/26/upshot/ai-synthetic-data.html

12. Bik, E. M., Casadevall, A., & Fang, F. C. (2016). The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications. mBio, 7(3), e00809-16. https://doi.org/10.1128/mBio.00809-16

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

13. 1.Bisson, R. (2023, July 19). Rise of AI stokes fears research misconduct could accelerate. Research Professional News. https://www.researchprofessionalnews.com/rr-news-world-2023-7-rise-of-ai-stokes-fears-research-misconduct-could-accelerate/

14. Bockting, C. L., van Dis, E. A. M., van Rooij, R., Zuidema, W., & Bollen, J. (2023). Living guidelines for generative AI - why scientists must oversee its use. Nature 622, 693-696. https://doi.org/10.1038/d41586-023-03266-1

15. Borji, A. (2024). A note on Shumailov et al. (2024): "AI models collapse when trained on recursively generated data". ArXiv. https://doi.org/10.48550/arXiv.2410.12954

16. Böttcher, F., & Thiel, F. (2018), Evaluating research-oriented teaching: A new instrument to assess university students' research competences. Higher Education, 75(1), 91–110. https://doi.org/10.1007/s10734-017-0128-y

17. Brent, T. (2023, June 27). European research integrity code updated to reflect advances in artificial intelligence. Science Business. https://sciencebusiness.net/news/AI/european-research-integrity-code-updated-reflect-advances-artificial-intelligence

18. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language Models are Few-Shot Learners. ArXiv. https://doi.org/10.48550/arXiv.2005.14165

19. Butler, N., Delaney. H., & Spoelstra S. (2017). The gray zone: questionable research practices in the business school. Academy of Management Learning & Education, 16(1), 94–109. https://doi.org/10.5465/amle.2015.0201

20. Council of Europe (2023, April 26). EduTalks@Council of Europe – Artificial Intelligence and Academic Integrity. Council of Europe. https://www.coe.int/en/web/education/-/edutalks-council-of-europe-artificial-intelligence-and-academic-integrity

21. Cristianini, N., Scantamburlo, T. & Ladyman, J. (2023). The social turn of artificial intelligence. AI & Society 38. 89–96 (2023). https://doi.org/10.1007/s00146-021-01289-8

22. Currie, G. M. (2023). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? Seminars in Nuclear Medicine 53(5). 719-730. https://doi.org/10.1053/j.semnuclmed.2023.04.008

23. Currie, G. M. (2023). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? Seminars in Nuclear Medicine, 53(5), 719–730. https://doi.org/10.1053/j.semnuclmed.2023.04.008.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

24. Dahlke, J. (2024). A.I. go by many names: towards a sociotechnical definition of artificial intelligence. arXiv:2410.13452. https://doi.org/10.48550/arXiv.2410.13452

25. Dawson, A. G. (2023). Artificial intelligence and academic integrity. Aspen Publishing.

26. Dohmatob, E., Feng, Y., Yang, P., Charton, F., & Kempe, J. (2024). A Tale of Tails: Model Collapse as a Change of Scaling Laws. Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. https://openreview.net/forum?id=KVvku47shW

27. Eaton, S. E. (2023, March 4). Artificial intelligence and academic integrity, post- plagiarism. University World News. https://www.universityworldnews.com/post.php?story=20230228133041549

28. Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? Journal of Responsible Technology, 13, 100060. https://doi.org/10.1016/j.jrt.2023.100060

29. Eliot, L. (2024). Rethinking the doomsday clamor that Generative AI will fall apart due to catastrophic model collapse. Forbes. https://www.forbes.com/sites/lanceeliot/2024/06/30/rethinking-the-doomsday-clamor-that-generative-ai-will-fall-apart-due-to-catastrophic-model-collapse/

30. Erasmus University Rotterdam (2020). Dilemma game. https://www.eur.nl/en/about-university/policy-and-regulations/integrity/research-integrity/dilemma-game

31. European Commission (2019). Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

32. European Commission (2025). Living guidelines on the responsible use of generative AI in research. https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf

33. European Commission. (2024). AI Act. Official Journal of the European Union 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

34. Fowler, D. S. (2023). AI in Higher Education: Academic Integrity, Harmony of Insights, and Recommendations. Journal of Ethics in Higher Education, (3), 127–143. https://doi.org/10.26034/fr.jehe.2023.4657

35. Gerstgrasser, M., Schaeffer, R., Dey, A., Rafailov, R., Pai, D., Sleight, H., Hughes, J., Korbak, T., Agrawal, R., Gromov, A., Roberts, D. A., Yang, D., Donoho, D., & Koyejo, S. (2024). Is model collapse inevitable? Breaking the curse of recursion by accumulating real and synthetic data. ArXiv. https://doi.org/10.48550/arXiv.2404.01413

36. Goovaerts, D. (2024). AI degeneration: here's why aging models could be a problem. FIERCE Network. https://www.fierce-network.com/ai/ai-degeneration-heres-why-aging-models-could-be-problem

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

37. Gulumbe, B. H., Audu, S. M. & Hashim, A. M. (2024). Balancing AI and academic integrity: what are the positions of academic publishers and universities?. AI & Soc 40, 1775–1784. https://doi.org/10.1007/s00146-024-01946-8 \

38. Hammond, K. (2024). Degenerative AI: The risks of training systems on their own data. Center for Advancing Safety of Machine Intelligence (CASMI). https://casmi.northwestern.edu/news/articles/2024/degenerative-ai-the-risks-of-training-systems-on-their-own-data.html

39. Holmes, W., Persson, J., Chounta, I-A., Wasson, B., and Dimitrova, V. (2022). Artificial intelligence and education. A critical view through the lens of human rights, democracy and the rule of law. Council of Europe. https://rm.coe.int/prems-092922-gbr-2517-ai-and-education-txt-16x24-web/1680a956e3

40. Hussain, M. M. (2023). The policy efforts to address racism and discrimination in higher education institutions: the case of Canada. CEPS Journal, 13(2), 9–29. https://doi.org/10.26529/cepsj.965

41. Infobip. (2024). What is AI model collapse?. https://www.infobip.com/glossary/model-collapse

42. Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nat Mach Intell 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

43. Kaushik, A. (2021). UNESCO's Recommendation on the Ethics of AI. https://montrealethics.ai/unescos-recommendation-on-the-ethics-of-ai/

44. Khan A. A., Badshah S., Liang P., Khan B., Waseem M., Niazi M., Akbar M. A. (2021). Ethics of AI: A Systematic Literature Review of Principles and Challenges. arXiv. https://doi.org/10.48550/arXiv.2109.07906

45. Kieslich, K., Helberger, N., & Diakopoulos, N. (2025). Scenario-Based Sociotechnical Envisioning (SSE): An Approach to Enhance Systemic Risk Assessments. SocArXiv. https://doi.org/10.31235/osf.io/ertsj_v1

46. Kim, SE., Kim, K., Lee, J., Ko, Y., Wang, Y., So, HJ. (2025). Dilemmas in AI Ethics: A Digital Game for Moral Reasoning and Collective Decision-Making. In: Cristea, A.I., Walker, E., Lu, Y., Santos, O.C., Isotani, S. (eds) Artificial Intelligence in Education. AIED 2025. Lecture Notes in Computer Science, vol 15878. Springer, Cham. https://doi.org/10.1007/978-3-031-98417-4_31

47. Kroes, P., Franssen, M., Van de Poel, I., & Ottens, M. (2006). Treating socio-technical systems as engineering systems: Some conceptual problems. Systems Research and Behavioral Science, 23(6), 803–814. https://doi.org/10.1002/sres.703.

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

48. Kroet, C. (2025). Big Tech spending on Brussels lobbying hits record high, report claims. Euronews. https://www.euronews.com/next/2025/10/29/big-tech-spending-on-brussels-lobbying-hits-record-high-report-claims

49. Kudina, O., van de Poel, I. (2024). A sociotechnical system perspective on AI. Minds & Machines, (21). https://doi.org/10.1007/s11023-024-09680-2

50. Lutkevich, B. (2023). Model collapse explained: How synthetic training data breaks AI. TechTarget. https://www.techtarget.com/whatis/feature/Model-collapse-explained-How-synthetic-training-data-breaks-AI

51. Lynch, M. (2024). When A.I.'s output is a threat to A.I. itself. The Tech Advocate. https://www.thetechedvocate.org/when-a-i-s-output-is-a-threat-to-a-i-itself/

52. Marr, B. (2024). Why AI models are collapsing and what it means for the future of technology. Forbes. https://www.forbes.com/sites/bernardmarr/2024/08/19/why-ai-models-are-collapsing-and-what-it-means-for-the-future-of-technology/

53. Moya, B., Eaton, S.E., Pethrick, H., Hayden, A., Brennan, R., Wiens, J., & McDermott, B. (2024). Academic Integrity and Artificial Intelligence in Higher Education (HE) Contexts: A Rapid Scoping Review. Canadian Perspectives on Academic Integrity. https://doi.org/10.55016/ojs/cpai.v7i3.78123

54. Nanda, N. (2021, December 7). Is artificial intelligence a threat to academic integrity? Original by Turnitin. https://www.ouriginal.com/is-ai-a-threat-to-academic-integrity/

55. Nature Publishing Group (2024). Using AI to train AI: Model collapse could be coming for LLMs, say researchers. TechXplore. https://techxplore.com/news/2024-07-ai-collapse-llms.html

56. NPA Core Competencies Committee (2007-2009). The NPA postdoctoral core competencies. https://www.nationalpostdoc.org/page/CoreCompetencies

57. OECD (2019). OECD AI principles overview. https://oecd.ai/en/ai-principles

58. OECD Future of Education and Skills 2030 (2019). OECD learning compass 2030: A series of concept notes. http://www.oecd.org/education/2030-project/teaching-and-learning/learning/learning-compass-2030/OECD_Learning_Compass_2030_Concept_Note_Series.pdf

59. Olatunde Oduoye, M., Javed, B., Gupta, N., & Valentina Sih, C. M. (2023). Algorithmic Bias and Research integrity; the role of non-human authors in shaping scientific knowledge with respect to Artificial Intelligence (AI); a perspective. International Journal of Surgery. Advance online publication. https://doi.org/10.1097/JS9.0000000000000552

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

60. Romano, U. (2023). From Generative to Degenerative AI: Risks of Self-Feeding Loop of AI Learning. LinkedIn. https://www.linkedin.com/pulse/from-generative-degenerative-ai-risks-self-feeding-loop-romano-0z3pf

61. Roque, G., Nascimento, J., Souza, R., Alves, C, & Araújo, J. (2025). Trust requirements in sociotechnical systems: A systematic literature review. Information and Software Technology, 186, 107796. https://doi.org/10.1016/j.infsof.2025.107796.

62. Sadowski, J. (2023). X. https://x.com/jathansadowski/status/1625245803211272194?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1625245803211272194%7Ctwgr%5Ee5d714dc7b691b94d9611b54e9ed6ee250564e99%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Ftheconversation.com%2Fwhat-is-model-collapse-an-expert-explains-the-rumours-about-an-impending-ai-doom-236415

63. Schwemer, S. F., Tomada, L. and Pasini, T. (2021), Legal AI Systems in the EU's proposed Artificial Intelligence Act. Proceedings of the Second International Workshop on AI and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2021), held in conjunction with ICAIL 2021, June 21, 2021, Sao Paulo, Brazil. https://ssrn.com/abstract=3871099

64. Selan, J., & Metljak, M. (2023). Developing and Validating the Competency Profile for Teaching and Learning Research Integrity. Center for Educational Policy Studies Journal, 13(3), 33–74. https://doi.org/10.26529/cepsj.1618

65. Shirky, C. (2025). Students Hate Them. Universities Need Them. The Only Real Solution to the A.I. Cheating Crisis. The New York Times. https://www.nytimes.com/2025/08/26/opinion/culture/ai-chatgpt-college-cheating-medieval.html

66. Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The Curse of Recursion: Training on Generated Data Makes Models Forget. ArXiv. https://doi.org/10.48550/arXiv.2305.17493

67. Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. Nature 631, 755–759. https://doi.org/10.1038/s41586-024-07566-y

68. Snoswell, A. J. (2024). What is 'model collapse'? An expert explains the rumours about an impending AI doom. The Conversation. https://theconversation.com/what-is-model-collapse-an-expert-explains-the-rumours-about-an-impending-ai-doom-236415

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

69. STM (2021). AI ethics in scholarly communication: STM best practice principles for ethical, trustworthy and human-centric AI. https://s3.eu-west-2.amazonaws.com/stm.offloadmedia/wp-content/uploads/2024/08/10033009/2021_04_29_STM_AI_White_Paper_April2021-1.pdf

70. Sweenor, D. (2023). AI Entropy: The Vicious Circle of AI-Generated Content. Understanding and Mitigating Model Collapse. Medium. https://medium.com/data-science/ai-entropy-the-vicious-circle-of-ai-generated-content-8aad91a19d4f

71. The US National Academies of Sciences, Engineering, and Medicine (2017). Fostering integrity in research. The National Academies Press. https://doi.org/10.17226/21896

72. Trachtenberg, B. (2023). ChatGPT, artificial intelligence, and academic integrity. Office of Academic Integrity. University of Missouri. https://oai.missouri.edu/chatgpt-artificial-intelligence-and-academic-integrity/

73. Turnitin. (2023). Academic integrity in the age of AI. Turnitin. https://www.turnitin.com/resources/academic-integrity-in-the-age-of-AI

74. University of Cambridge (2023). Artificial Intelligence. University of Cambridge. https://www.plagiarism.admin.cam.ac.uk/what-academic-misconduct/artificial-intelligence

75. Van de Poel, I. (2020). Embedding values in Artificial Intelligence (AI) systems. Minds and Machines, 30(3), 385–409. https://doi.org/10.1007/s11023-020-09537-4

76. Van de Poel, I., & Kudina, O. (2022). Understanding Technology-Induced Value Change: A pragmatist proposal. Philosophy & Technology, 35(2), 40. https://doi.org/10.1007/s13347-022-00520-8.

77. Vaughan-Nichols, S. J. (2025). Some signs of AI model collapse begin to reveal themselves. The Register. https://www.theregister.com/2025/05/27/opinion_column_ai_model_collapse/

78. Villalobos, P., Sevillay, J., Heimx, L., Besirogluz, T., Hobbhahn, M., & Ho, A. (2022). Will we run out of data? An analysis of the limits of scaling datasets in machine learning. ArXiv. https://doi.org/10.48550/arXiv.2211.04325

79. Vue.ai. (2023). The degeneration of Generative AI. https://www.vue.ai/blog/ai-transformation/the-degeneration-of-generative-ai/

80. Wenger, E. (2024). AI returns gibberish when trained on generated data. Nature 631, 742-743. https://doi.org/10.1038/d41586-024-02355-z

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

81. Williamson, B. (2023). Degenerative AI in education. Code Acts in Education. https://codeactsineducation.wordpress.com/2023/06/30/degenerative-ai-in-education/

82. York University (2023, August 16). AI technology and academic integrity for instructors. York University. https://www.yorku.ca/unit/vpacad/academic-integrity/ai-technology-and-academic-integrity/

83. Zobel, J. (2023, May 19). We need to retain research integrity in the AI era. Pursuit and Research, University of Melbourne. https://pursuit.unimelb.edu.au/articles/we-need-to-retain-research-integrity-in-the-ai-era

00964- 7710122232 Alfaidcenter2011@gmail.com العراق - بغداد - الكرادة

تأسس مركز الفيض العلمي لاستطلاع الرأي والدراسات المجتمعية في بغداد بموجب شهادة التسجيل الصادرة عن الأمانة العامة لمجلس الوزراء -دائرة المنظمات غير الحكومية المرقمة (1J775330) بتاريخ ٢٠١٢/٤/٢٦، وهو مركز علمي بحثي يهتم بإجراء الاستطلاعات والدراسات الميدانية فضلا عن إعداد الأوراق البحثية والمقالات حول قضايا الحياة المجتمعية للأسرة والمواطن، والدولة بمؤسساتها المختلفة.

للتواصل

00964- 7710122232

Alfaidcenter2011@gmail.com

www.al-faidh.com

العراق - بغداد - الكرادة