تحديات في تقييم جودة ومصداقية الأدب الرمادي المولَّد بواسطة الذكاء الاصطناعي*

ترجمة
رغداء بلال
raghdaa.bilal@gmail.com

المستخلص:

في مشهد متسارع التطور للتواصل العلمي، يعد الأدب الرمادي مصدراً هاماً للمعلومات. وبما أن الأدب الرمادي يُعرّف على أنه عبارة عن معلومات مسجلة لا تخضع لسيطرة الناشرين التجاريين وأنها متاحة للجميع دون عملية مراجعة الأقران التقليدية، فهو يتضمن مجموعة واسعة من المواد مثل: التقارير والأطروحات العلمية وورقات المؤتمر والأوراق البيضاء.

أدى ظهور الذكاء الاصطناعي إلى حدوث ثورة في صناعة المحتوى من خلال توفير سرعة وحجم غير مسبوقين للمعلومات. ومع ذلك، فإن هذا التقدم يطرح تحديات فريدة، لاسيما في إدارة وتقييم سلامة وجودة المحتوى المولَّد بواسطة الذكاء الاصطناعي. تتناول الدراسة هذه التحديات وتسلط الضوء على القيود المفروضة على طرق التقييم التقليدية للمخرجات المولَّدة بواسطة الذكاء الاصطناعي. كما تناقش نهضة الذكاء الاصطناعي في إنتاج الأدب الرمادي وأوجه القصور في مناهج التقييم التقليدية والقضايا الخاصة المرتبطة بالمحتوى المولَّد بواسطة الذكاء الاصطناعي.

الكلمات المفتاحية: الأدب الرمادي، الذكاء الاصطناعي، تحديات، طرق التقييم

مقدمة: فهم الأدب الرمادي

يمثل الأدب الرمادي أي مصدر بيانات مسجلة ومرجعية ومستدامة، أو مصدر معلومات للقيمة الحالية أو المستقبلية، ويكون متاحاً للجميع دون عملية مراجعة تقليدية من قبل الأقران (Savic, 2017).

---

*مقالة مترجمة عن:

Savić, D. (2025). Challenges in evaluating the Quality and Credibility of AI-generated Grey Literature. In PISA 2025 Conference*: "How Scientific Communication Practices are Evolving*". GreyNet International. *https://doi.org/10.26069/greynet-2025-000.769-gn*

تاريخ استلام البحث: 18-8-202 تاريخ إجازة البحث: 5-9-2026

مع الاعتراف بغياب وجود عملية مراجعة تقليدية من قِبل الأقران للأدبيات البيضاء العادية، يُراعي هذا التعريف جميع العناصر الرئيسية لمفهوم الأدب الرمادي، وهي: الحفظ طويل الأمد والاستدامة والاستخدام والقيمة.

وعلى الرغم من أن هذه الدراسة تتناول بشكل رئيسي مجموعات البيانات الرمادية، يوضح الشكل (1) أنواع وأشكال عديدة من الأدب الرمادي. ويُدرج موقع GreyNet الإلكتروني (greynet.org) أكثر من 150 نوعاً من المستندات، بما في ذلك قواعد البيانات ومجموعات البيانات وجداول البيانات وأوراق البيانات وبيانات الأقمار الصناعية وبيانات المنتج.

ببلوغرافيات	مخطوطات مرفوضة	منشورات منظمات غير حكومية وشركات استشارية
أوراق مناقشة	مخطوطات غير مرسلة	فيديوهات
نشرات إخبارية	ملخصات مؤتمر	مقالات ويكي
عروض تقديمية	فصول كتاب	بريد الكتروني
تقارير تقييم البرامج	مراسلات شخصية	مدونات ووسائل التواصل الاجتماعي
ملاحظات فنية	اتصالات غير رسمية	مجموعة بيانات
منشورات وكالات حكومية	بيانات تعداد سكان	تقارير لجان
تقارير وكالات تمويل	نسخ أولية	ورقات عمل
تقارير غير منشورة	مواصفات	تقارير شركات
أطروحات	براءات الاختراع	كتالوجات
وثائق السياسة	ويبينارز	خطابات
		تقارير على مواقع الانترنت

الشكل (1) أنواع الأدب الرمادي

انتشار مخرجات الذكاء الاصطناعي في إنتاج الأدب الرمادي

إن تحديد الكمية الدقيقة من الأدب الرمادي الناتج يومياً أو إجمالي الكمية المتاحة هو أمر صعب نظراً لطبيعته المتنوعة واللامركزية. فالعديد من أنواع الأدب الرمادي لا يمكن غالباً تتبعها بنفس طريقة الأدبيات المنشورة رسمياً، مما يجعل التقديرات الدقيقة صعبة. ومع ذلك يمكن تقديم بعض الرؤى العامة. على سبيل المثال، إن الإنتاج اليومي للأدب الرمادي هو أمر

جوهري، وخاصةً بالنسبة إلى عدد المؤسسات والمنظمات والأفراد الذين يصدرون التقارير والأوراق البيضاء والمستندات المشابهة. كما ارتفع معدل إنشاء هذه الأدبيات بشكل ملحوظ في السنوات الأخيرة مع ازدياد المنصات الرقمية وسهولة النشر الإلكتروني.

يُعد حجم البيانات المجمعة والمدارة هائلاً، فهو يشمل ملايين المستندات على مستوى العالم. على سبيل المثال، توفر مستودعات مثل النظام الدولي للعلوم والتكنولوجيا الزراعية (نظام أجريس AGRIS) 12 مليون سجل، والنظام الدولي للمعلومات النووية (نظام إينيس INIS) 5 مليون سجل، ونظام معلومات عن الأدب الرمادي في أوروبا (OpenGrey) 700 ألف سجل، أو قاعدة بيانات للأطروحات والرسائل العلمية (بروكويست ProQuest) 3 مليون سجل، إمكانية الوصول إلى مجموعات كبيرة، لكنها لا تزال تمثل جزءاً ضئيلاً مما هو موجود.

بالنسبة لرسائل البريد الإلكتروني، اعتباراً من عام 2025، تشير التقديرات إلى ما يقارب 376.4 مليار رسالة بريد إلكتروني مرسلة ومستلمة يومياً حول العالم. ويشهد هذا الرقم ارتفاعاً مطرداً على مدار السنوات، ومن المتوقع أن يصل إلى 408.2 مليار رسالة بريد إلكتروني يومياً بحلول عام 2027.

أما فيما يتعلق بمستخدمي البريد الإلكتروني، يستفيد حوالي 4.48 مليار شخص من خدمات البريد الإلكتروني على مستوى العالم في عام 2024. ومن المتوقع أن يرتفع هذا العدد إلى أكثر من 5 مليار بحلول عام 2027 (oberlo.com).

تبين هذه الإحصائيات، على الرغم من انتشار منصات التراسل البديلة، الأهمية الدائمة للبريد الإلكتروني كوسيلة أساسية للتواصل سواءً على الصعيد الشخصي أو المهني.

أثر الذكاء الاصطناعي على إنشاء الأدب الرمادي:

أحدث الذكاء الاصطناعي ثورة في إنتاج الأدب الرمادي من خلال قدرته على أتمتة وتحسين جوانب متنوعة لعملية إنشاء المحتوى. وتشمل الأنواع الأساسية للأدب الرمادي المولّد بواسطة الذكاء الاصطناعي ما يلي:

- **ملخصات الأبحاث:** تستطيع أدوات الذكاء الاصطناعي مثلCPT-4 (نظام المصطلحات الإجرائية الحالي) أن تختصر الدراسات المعقدة إلى ملخصات موجزة.

- **المطبوعات الأولية**: يُسرّع الذكاء الاصطناعي عملية صياغة نتائج الأبحاث الأولية، وتسهيل نشرها.
- **التقارير الفنية**: من خلال تجميع البيانات، يُنتج الذكاء الاصطناعي تقارير مفصلة حول مواضيع متنوعة، لتحسين إمكانية الوصول إلى المعلومات الفنية.
- **البحث**: أظهرت المنصات المدعومة بالذكاء الاصطناعي مثل Semantic Scholar (semanticscholar.org) قدرة الذكاء الاصطناعي على تنظيم وتقديم الأدب الرمادي بشكل أكثر فعالية، مما يمكّن الباحثين من تصفح كميات هائلة من البيانات بسهولة أكبر. كما يتيح للمستخدمين استكشاف أكثر من 200 مليون منشور أكاديمي وأكثر من 120 مليون براءة اختراع، ويتضمن ذلك العديد من المستندات غير خاضعة لمراجعة الأقران (معهد كارولينسكا،2023). وعلى الرغم من أن محرك البحث الأكاديمي BASE (محرك بحث بيليفيلد الأكاديمي)(base-search.net) لم يُصرَّح عن استخدام الذكاء الاصطناعي في وظائف البحث الأساسية، إلا أنه يضم مجموعة ضخمة أكثر من 420 مليون سجل، العديد منها أدب رمادي.

يتميز دور الذكاء الاصطناعي في إنشاء الأدب الرمادي بالعديد من المزايا:

- **السرعة:** يقلل الذكاء الاصطناعي من الوقت اللازم لإنتاج الأدب الرمادي حيث يمكنه توليد ملخصات وتقارير وبيانات وصفية بسرعة.
- **الحجم:** تتيح قابلية التوسع في الذكاء الاصطناعي معالجة مجموعات البيانات التي تساهم لزيادة كبيرة في توافر الأدب الرمادي.
- **التخصيص:** يستطيع الذكاء الاصطناعي تخصيص المحتوى لجمهور محدد والتكيف مع سياقات متنوعة، وإنتاج مستندات متعددة اللغات، مما يُوسع مجال تطبيق الأدب الرمادي.

أظهرت دراسة أجراها ماير- شونبرغر وكوكير عام 2013 قدرة البيانات الكبيرة، المدعومة بواسطة الذكاء الاصطناعي، على تغيير طريقة نشر المعلومات من خلال جعل كميات هائلة من البيانات قابلة للإدارة والتنفيذ.

التحديات التي يطرحها الأدب الرمادي المولّد بواسطة الذكاء الاصطناعي

على الرغم من مزايا الأدب الرمادي المولّد بواسطة الذكاء الاصطناعي، فإنه لا يخلو من التحديات:

- **الهلوسات وعدم دقة الحقائق:** على الرغم من تطور نماذج الذكاء الاصطناعي، فإنها لا تزال قادرة على إنتاج محتوى غير صحيح أو مزيف، مما يفرض مخاطر على مصداقية الأدب الرمادي (Flanagin et al.,2023).
- **تحديات التقييم:** تتجاوز الوتيرة السريعة لتوليد محتوى الذكاء الاصطناعي قدرات عمليات مراجعة الأقران التقليدية، مما يستلزم طرق تقييم جديدة.
- **قضايا الشفافية:** تربك الطبيعة المبهمة لخوارزميات الذكاء الاصطناعي (مشكلة "الصندوق الأسود") عملية التحقق من المصادر والمنطق الكامن وراء المحتوى المولّد بواسطة الذكاء الاصطناعي (Bender et al.,2021).

ناقش تقرير أعدّه .Bender et al (2021) مخاطر "الببغاوات العشوائية" في نماذج اللغة، حيث يولّد الذكاء الاصطناعي نصوصاً تبدو منطقية ولكنها تفتقر إلى الأساس الواقعي، مما يؤكد الحاجة إلى التمحيص في مخرجات الذكاء الاصطناعي (Bender et al.,2021).

الطرق التقليدية في تقييم الأدب الرمادي

لقد شكلت الطرق التقليدية، مثل مراجعة الأقران ومؤشرات القياس الببلوغرافي، حجر الأساس في تقييم الأدب الرمادي:

- **مراجعة الأقران:** تشمل خبراء متخصصين في المجال لتقييم جودة وملائمة المحتوى.
- **مؤشرات القياس الببلوغرافي:** تساعد مقاييس مثل تحليل الاقتباس ومعامل التأثير والمقاييس البديلة في قياس نفوذ وانتشار الأدب الرمادي.

وبالرغم من ذلك، فإن مدى إمكانية تطبيق هذه الطرق على المحتوى المُولّد بواسطة الذكاء الاصطناعي محدودة، لأنها مصممة لتقييم نصوص بشرية.

قيود طرق التقييم التقليدية للمحتوى المولّد بواسطة الذكاء الاصطناعي

تواجه طرق التقييم التقليدية العديد من القيود عند تطبيقها على المحتوى المولّد بواسطة الذكاء الاصطناعي:

- **عدم الكفاية في تقييم مخرجات الذكاء الاصطناعي:** تفشل غالباً مقاييس التقييم التي تركز على الإنسان في التقاط الفروقات الدقيقة للمحتوى المولّد بواسطة الذكاء الاصطناعي، مثل مصداقيته وأصالته.
- **سرعة توليد المحتوى:** يستطيع الذكاء الاصطناعي إنتاج محتوى بوتيرة تتخطى قدرة أنظمة المراجعة التقليدية، مما يؤدي إلى خطر نشر معلومات غير مُدققة.
- **التحيز وسوء التفسير:** قد تشمل أنظمة الذكاء الاصطناعي على تحيزات من بيانات التدريب الخاصة بها أو تسيء تفسير المفاهيم العلمية المعقدة، مما يزيد من تعقيد عملية التقييم (MIT Sloan EdTech,2024)

دعا كل من فينسنت وهيرشت إلى تطوير أطر تقييم جديدة مصممة خصيصاً لمواجهة التحديات الفريدة التي يفرضها المحتوى المولّد بالذكاء الاصطناعي.

تعد كل من SIFT وRADAR من الأطر المستخدمة لتقييم مصادر المعلومات ولاسيما في السياقات الرقمية والأكاديمية. ترمز SIFT إلى التوقف والتحقق من المصدر وإيجاد تغطية أفضل وتتبع الادعاءات والاقتباسات والوسائط إلى سياقها الأصلي. تولد هذه الطريقة التي طورها خبير محو الأمية الرقمية مايك كولفيلد على التقييم السريع للمعلومات قبل استثمار الوقت في قراءة أو مشاهدة المصدر. أما RADAR فهو اختصار للملائمة والسلطة والتاريخ والمظهر والسبب. وتساعد أداة التقييم على تقييم جودة ومصداقية وملائمة المعلومات (Mandalios,2013).

الحلول المقترحة وأفضل الممارسات

لمعالجة أوجه القصور في طرق التقييم التقليدية، تم اقتراح الحلول التالية:

- **معايير التقييم الخاصة بالذكاء الاصطناعي**: تطوير مقاييس وأدوات وأطر عمل جديدة مصممة خصيصاً لسمات المحتوى المولّد بواسطة الذكاء الاصطناعي.
- **تعزيز الشفافية**: تنفيذ تصنيف واضح يشير إلى دور الذكاء الاصطناعي في إنشاء المحتوى

لضمان أن القراء على دراية بمصادر النص.

- **أدوات الكشف عن الذكاء الاصطناعي**: استخدام البرمجيات لتحديد المحتوى المولّد بواسطة الذكاء الاصطناعي ووضع علامة عليه، وبالتالي الحفاظ على سلامة الأدب الرمادي.
- **المبادئ التوجيهية الأخلاقية**: وضع معايير شاملة للاستخدام المسؤول للذكاء الاصطناعي في التواصل العلمي.
- **تحديد واضح للمحتوى المولّد بواسطة الذكاء الاصطناعي**: استخدام عبارة مثل "أنشئت هذه المقالة جزئياً باستخدام أدوات الذكاء الاصطناعي".

يؤكد تقريراً مشـتركاً بين منظمة اليونيسـكو UNESCO ومنظمة التعاون الاقتصـادي والتنمية OECD على أهمية تعزيز التعاون متعدد التخصـصـات لتطوير مبادئ توجيهية وأطر عمل فعّالة لتقييم المحتوى المولّد بواسطة الذكاء الاصطناعي (UNESCO & OECD, 2022).

الخاتمة

يحُدث الذكاء الاصطناعي ثورة في إنتاج الأدب الرمادي من خلال توفير سرعة وحجماً وإمكانات تخصيص غير مسبوقين. وبالرغم من ذلك، فإن أساليب التقييم التقليدية غير كافية لتقييم المحتوى المولَّد بواسطة الذكاء الاصطناعي مما يستدعي إتباع أساليب جديدة. تفرض مسائل أساسية، كالهلوسات ومخاوف تتعلق بالدقة والاعتبارات الأخلاقية، تحديات كبيرة لمصداقية المحتوى العام المولّد بواسطة الذكاء الاصطناعي. وتعد الحلول المقترحة بما في ذلك أطر التقييم الخاصة بالذكاء الاصطناعي وتعزيز الشفافية ضرورية للحفاظ على النزاهة العلمية. كما يعد التعاون متعدد التخصصات بين خبراء الذكاء الاصطناعي والمتخصصين في المجال أمراً بالغ الأهمية لتطوير استراتيجيات تقييم فعّالة.

يوفر دمج الذكاء الاصطناعي في إنتاج الأدب الرمادي فرصاً غير مسبوقة للابتكار والكفاءة. ومع ذلك، فإنه يستلزم أيضاً تحولاً جذرياً في كيفية تقييمنا لهذا المحتوى وضمان سلامته. فمن خلال تبني أساليب تقييم جديدة، يمكننا تسخير إمكانات الذكاء الاصطناعي مع الحفاظ على مصداقية وفائدة الأدب الرمادي.

وبينما نمضي قدماً، من الضروري أن نبقى مرنين ويقظين حيث يتطلب التطور التكنولوجي السريع تحسيناً مستمراً لأساليب التقييم لضمان استمرار الأدب الرمادي كمصدر موثوق وقيّم في المجتمع العلمي.

المراجع:

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.

Caulfield, M. (2024). Evaluating Sources. Simon Fraser University. www.lib.sfu.ca/help/research-assistance/evaluating-sources.

Flanagin, S., Bibbins-Domingo, K, Berkwits, M. & Christiansen, M.A. (2023). Nonhuman "Authors" and Implications for the Integrity of Scientific Publication and Medical Knowledge. JAMA. 2023;329(8):637-639 doi:10.1001/jama.2023.1344.

Karolinska Institute, University Library Web Team (2023) https://kib.ki.se/en/search-evaluate/greyliterature.

Mandalios, J. (2013). RADAR: An approach for helping students evaluate Internet sources. Journal of Information Science, 39(4), 470-478.

Mayer-Schönberger, V., & Cukier, K. (2013). Big Data: A Revolution That Will Transform How We Live, Work, and Think. Eamon Dolan/Houghton Mifflin Harcourt.

MIT Sloan EdTech (2024). When AI Gets It Wrong Addressing AI Hallucinations and Bias.

https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/

Savić, D., 2017. Rethinking the Role of Grey Literature in the Fourth Industrial Revolution. 10th Conference on Grey Literature and Repositories: proceedings [online]. Prague: National Library of Technology. TGJ (TheGrey Journal) Special Winter Issue, Volume 14, 2018.

Vincent, N., & Hecht, B. (2018). Data Strikes: Evaluating the Feasibility of Social Protest against Data-Driven Technologies. Proceedings of the ACM on Human-Computer Interaction.

UNESCO & OECD. (2022). Guidelines for the Responsible Use of AI in Scientific Communication.

Challenges in Evaluating the Quality and Credibility of AI-Generated Grey Literature

Translated By
Raghdaa Bilal
raghdaa.bilal@gmail.com

Abstract:

In the rapidly evolving landscape of scientific communication, grey literature (GL) stands as an important resource. Defined as recorded information not controlled by commercial publishers and made publicly available without the traditional peer-review process, GL includes a wide range of materials such as reports, theses, conference papers, and white papers.

The advent of Artificial Intelligence (AI) has revolutionized content creation, offering unprecedented speed and volume. However, this advancement presents unique challenges, particularly in managing and evaluating the integrity and quality of AI-generated content. This paper examines these challenges, highlighting the limitations of traditional evaluation methods for AI-generated outputs. It discusses the rise of AI in GL production, the inadequacies of conventional assessment approaches, and the specific issues associated with AI-generated content. The paper concludes with proposed solutions and best practices to ensure the integrity of scientific communication in the AI-driven era.

Keywords: Grey Literature; AI; Challenges; Evaluation Methods