مجلة دراسات الخليج والجزيرة العربية - 51 - (198) 2025
https://doi.org/10.34120/jgaps.v51i198.3475

تقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام الذكاء الاصطناعي وفق المعايير الدولية

حمود علي العبدلي[1]✉
علي محمد أخواجه[2]

ملخص

الأهداف: هدفت الدراسة إلى تقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية وفق المعايير الدولية باستخدام نماذج الذكاء الاصطناعي لتحليل المحتوى الكمي. **المنهج:** اعتمدت الدراسة المنهج الوصفي- تحليل المحتوى لتقييم جودة، وتأثير، وأخلاقيات البحث. **النتائج:** أظهرت متوسطاً إجمالياً مقبولاً (3.00 | 0.12)، مع تفاوتات بين المجلات؛ سجلت جودة البحث أداءً جيداً (3.43 | 0.11)؛ بدعم من دقة المنهجية (3.70 | 0.77)، ووضوح النتائج (3.50 | 0.66)، لكنها عانت من ضعف الابتكار (3.20 | 0.75) ونسبة مراجع حديثة أقل من 35%؛ أما التأثير؛ فكان ضعيفاً (1.49 | 0.09)، مع ضعف الاستشهادات والمساهمة التنموية؛ في المقابل تفوقت المجلات في الأخلاقيات (4.07 | 0.14) خاصة النزاهة العلمية (4.10 | 0.86)، وحقوق الملكية الفكرية (4.20 | 0.79). ويعكس التوافق المحدود إمكانات منهجية وأخلاقية مقيدة بضعف التأثير. **الخاتمة:** لا بد من تحول إستراتيجي يُفعل الإمكانات عبر التخصصات التطبيقية وإعادة التصميم البحثي لتلبية المتطلبات المحلية والدولية. وأوصت الدراسة بتحسين التحكيم العلمي، وتشجيع النشر الدولي، وربط الأبحاث بالاحتياجات التنموية لتعزيز التأثير الأكاديمي والمجتمعي.

الكلمات المفتاحية: تقييم جودة البحث العلمي، تقنيات الذكاء الاصطناعي، المعايير الدولية، الأثر البحثي

(1) أستاذ تكنولوجيا التعليم المساعد، قسم العلوم التربوية، كلية التربية، جامعة الحديدة، الجمهورية اليمنية. haaa97@gmail.com

(2) أستاذ تكنولوجيا التعليم المساعد، قسم المناهج وطرائق التدريس، كلية التربية والعلوم، جامعة إقليم سبأ، الجمهورية اليمنية. alkhwagh2006@gmail.com

- تُسلّم البحث في: 2025/5/16، أجيز للنشر في: 2025/8/4.

تصدر عن مجلس النشر العلمي بجامعة الكويت. جميع الحقوق محفوظة للمجلة.

للاستشهاد بهذه المقالة انظر ص.242

AI-based evaluation of research quality in scientific journals of Yemeni universities according to international standards

Hamoud A. Al-abdli [1]✉
Ali M. Akhwagh [2]

Abstract

Objectives: This study evaluated the quality of research in Yemeni university journals against international standards, employing artificial intelligence (AI) models for quantitative content analysis. **Method:** The research utilized a descriptive content analysis approach to assess three core dimensions: research quality, impact, and ethics. **Results:** The findings indicated an acceptable overall mean score (3.00 | 0.12), with significant variations among journals. Research quality was rated as Good (3.43 | 0.11), bolstered by strong methodological rigor (3.70 | 0.77) and clarity of results (3.50 | 0.66). However, key weaknesses included limited innovation (3.20 | 0.75) and outdated citations (fewer than 35% of references were recent). The research impact was Weak (1.49 | 0.09), characterized by low citation rates and minimal contribution to development. In contrast, the journals performed exceptionally well in research ethics (4.07 | 0.14), particularly in scientific integrity (4.10 | 0.86) and intellectual property rights (4.20 | 0.79). Overall, the journals exhibit strong methodological and ethical potential, but this is constrained by their low impact, limiting their alignment with international benchmarks. **Conclusion:** A strategic shift is necessary to unlock this potential by focusing on applied disciplines and redesigning research frameworks to meet both local and international demands. The study recommends strengthening the peer-review process, promoting international publication, and linking research to national development needs to enhance both academic and societal impact.

Keywords: evaluation of scientific research quality, artificial intelligence techniques, international standards, research impact

(1) Assistant Professor, Department of Educational Sciences, Faculty of Education, University of Hodeidah, Yemen. haaa97@gmail.com

(2) Assistant Professor, Curriculum and Teaching Methods Department, Faculty of Education and Sciences, University of Saba Region, Yemen. alkhwagh2006@gmail.com

- Submitted: 16/5/2025, Accepted: 4/8/2025.

Published by the Academic Publication Council of Kuwait University. All rights reserved.

To cite p.242

المقدمة

يشهد الإنتاج العلمي العالمي نمواً متسارعاً؛ إذ بلغ عدد المقالات العلمية المنشورة نحو 3.3 ملايين مقال في عام 2022 مع توقّعات بارتفاع العدد إلى نحو 3.8 ملايين بحلول عام 2025؛ بناءً على معدل نمو سنوي يراوح بين 4% و5% (National Science Board, 2023)؛ وتُعد المجلات المحكّمة وسيلة لنشر هذه الأبحاث لما لها من دور محوري في نشر المعرفة وتعزيز الابتكار العلمي (Larivière & Sugimoto, 2019)؛ وتختلف جودة هذه المجلات وتأثيرها باختلاف دقة التحكيم، وأصالة الأبحاث المنشورة، ومعدلات الاستشهاد بها؛ إذ تحقق المجلات ذات معدلات الاستشهاد المرتفعة تأثيراً أكبر على المستويين العلمي والمجتمعي، في حين تظل المجلات المنخفضة الجودة محدودة الأثر؛ مما يُبرز أهمية وجود معايير تقييم دقيقة وموثوق بها (Clarivate, 2024a; Leng, 2021).

وتؤكد المعايير الدولية أهمية الاستشهادات، وشفافية التحكيم، والالتزام الأخلاقي، والتعاون الدولي؛ بوصفها مؤشرات أساسية في قياس جودة البحث العلمي (CGIAR Independent Advisory & Evaluation Service, 2023)، وتقدم Clarivate (2024a) معياراً للجودة (مثل الشفافية والمنهجية)، و4 معايير للأثر (مثل الاستشهادات)، وهي معايير تسهم في ضمان جودة النشر وتميّز المجلات. ومع التسارع المستمر في وتيرة التحول الرقمي يتزايد اعتماد الباحثين على التكنولوجيا الحديثة، ويبرز الذكاء الاصطناعي بوصفه أداة فعالة لتحليل البيانات بدقة وسرعة؛ مما يسهم في تعزيز جودة الأبحاث ومصداقيتها (Khalifa & Albadawy, 2024)، وقد بيّنت الدراسات أن هذه التقنيات تسهم في رفع إنتاجية البحث العلمي وتسريع وتيرة الاكتشافات، من خلال قدرتها على اكتشاف الأنماط وتقديم رؤى تحليلية دقيقة (Kousha & Thelwall, 2024).

كما يُظهِر التقدّم في تقنيات الذكاء الاصطناعي، خصوصاً نماذج اللغة الكبيرة (LLMs)، إمكانات لتحسين آليات تقييم المجلات العلمية من خلال تحليل النصوص، وقياس الأصالة، وتقييم المنهجية بأساليب دقيقة ومتطورة (Cheng et

al., 2023)، وتُتيح هذه النماذج استخلاص رؤى معمقة حول جودة الأبحاث؛ مما يدعم تطبيق المعايير الدولية بكفاءة وفعالية (Kousha & Thelwall, 2024)؛ ومع ذلك يبقى استخدام هذه الأدوات في العالم العربي محدوداً؛ نتيجة تحديات تقنية ونقص في الكفاءات المؤهلة؛ الأمر الذي يعوق تطبيقها على تقييم المجلات المحلية (El-Ouahi, 2023)، وهي فجوة تتجلى بوضوح في استمرار التحديات التي تعوق جودة المجلات العربية وانتشارها واندماجها في النظام العلمي العالمي.

وتنعكس هذه الفجوة بوضوح في مجموعة من التحديات البنيوية، أبرزها محدودية حضورها في قواعد البيانات العالمية؛ إذ لا تتجاوز نسبتها 0.077% من إجمالي المجلات المفهرسة في Scopus حتى عام 2023 (Benziane, 2024)، كما تُشير التقارير إلى أن حصة الدول العربية لا تتعدى 1% من الإنتاج البحثي العالمي، على الرغم من أنها تمثل 5% من سكان العالم نتيجة محدودية التمويل البحثي الذي لا يتجاوز 0.5% من الناتج المحلي الإجمالي مقارنةً بـ 2%-3% في الدول المتقدمة (الخطيب، 2021؛ UNESCO, 2021)؛ ويؤثر نقص التدريب على الكتابة العلمية والتحكيم سلباً على جودة الأبحاث؛ مما يُضعف أصالتها وابتكارها (الرازحي، 2021)، كما يحد ضعف التعاون الدولي من نشر الأبحاث في المجلات المصنفة بسبب قصور التوثيق والمنهجية (السلطان، 2023؛ El-Ouahi, 2023)، وتبقى العديد من الأبحاث التربوية بعيدة عن تقديم حلول عملية نتيجة ضعف التواصل مع المجتمع العلمي العالمي (الهبوب والفخري، 2018).

أما في اليمن؛ فتتفاقم التحديات المرتبطة بالبحث العلمي نتيجة الوضع الاقتصادي والسياسي الذي أسهم في تدهور البنية التحتية الأكاديمية والبحثية، وانخفاض معدلات الاستشهاد بالأبحاث اليمنية مقارنةً بالمجلات العالمية المرموقة (مكرد، 2010)، ويُعزى ذلك جزئياً إلى ضعف أنظمة التحكيم؛ إذ تفتقر معظم المجلات اليمنية إلى معايير دقيقة، وتعتمد -في الغالب- على مراجعات شكلية لا تستوفي المعايير الدولية (El-Ouahi, 2023)، كما تتسم العديد من الأبحاث بالضعف المنهجي، نتيجة لنقص التدريب وصعوبة الوصول إلى قواعد البيانات البحثية العالمية(Al-Munifi & Aleryani, 2021)، ويُفاقم محدودية التمويل البحثي من تردي جودة الأبحاث وتقليص فرص نشرها في المجلات

العلمية ذات التصنيف العالمي (UNESCO, 2021)، كما أن أقل من 20% من الأبحاث اليمنية تتم في إطار تعاون دولي؛ مما يُضعف فرص المجلات المحلية في بلوغ مراتب التصنيف المرموقة (El-Ouahi, 2023).

مشكلة الدراسة

في ظل هذه التحديات تبرز الحاجة إلى تقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية استناداً إلى المعايير الدولية مع توظيف أدوات الذكاء الاصطناعي؛ بهدف تشخيص نقاط القوة والضعف لتساعد في وضع إستراتيجيات واضحة لتحسين واقع النشر العلمي، وتكمن الفجوة العلمية في تركيز الدراسات السابقة على جوانب؛ مثل ضعف التمويل وقلة التدريب (الخطيب، 2021)، دون أن تقدم تحليلاً شاملاً لمخرجات البحث؛ وعلى الرغم من أن الذكاء الاصطناعي أثبت فعاليته في تحليل البيانات البحثية وتحسين جودة المخرجات العلمية من خلال تحليل النصوص، وقياس الأثر، واستخلاص مؤشرات حول الأصالة والمنهجية (Cheng et al., 2023)، فإن استخدامه لا يزال غائباً عن السياق اليمني؛ ومن هنا تهدف هذه الدراسة إلى تقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام نماذج اللغة الكبيرة (LLMs) لتحليل محتوى المنشورات في المجلات المحكمة، واستخلاص بيانات علمية دقيقة حول جودة، وتأثير، وأخلاقيات البحث، وتكمن أهمية الدراسة في سعيها إلى سد هذه الفجوة من خلال تقديم تحليل شامل وموثوق به لواقع البحوث في المجلات اليمنية؛ بما يُسهم في توفير أساس علمي لتطوير النشر الأكاديمي، ويعزز مسارات التنمية الأكاديمية والمجتمعية في اليمن والمنطقة العربية على حد سواء.

أهداف الدراسة

- بناء قائمة معايير دولية لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية.
- بناء نموذج تطبيقي لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام الذكاء الاصطناعي (LLMs).
- تحديد مستوى توافق جودة البحوث في المجلات العلمية بالجامعات اليمنية مع معايير جودة البحث الدولية.

أسئلة الدراسة

1 - ما المعايير الدولية لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية؟
2 - ما النموذج التطبيقي لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام الذكاء الاصطناعي (LLMs)؟
3 - ما درجة توافق جودة البحوث في المجلات العلمية بالجامعات اليمنية مع معايير جودة البحث الدولية؟

الإطار النظري

المحور الأول: جودة مخرجات البحث العلمي

تتجاوز جودة مخرجات البحث العلمي مجرد النشر في مجلات مرموقة لتشمل أبعاداً متعددة؛ مثل: الأصالة التي تُقدم إضافات للأدبيات، والمنهجية العلمية الدقيقة التي تضمن الموثوقية والشفافية، والأثر الأكاديمي الذي يُعزز التخصصات، والأثر المجتمعي الذي يُترجم المعرفة إلى حلول عملية، والتأثير التعليمي الذي يُنمي القدرات الأكاديمية (Leng, 2021; Moher et al., 2020)، لكـــن نظم التقييم التقليدية التي تُركّز على المؤشرات الكمية؛ مثل عدد المنشورات ومعامل التأثير تُغفل هذه الأبعاد؛ مما يُؤدي إلى تقييمات سطحيـــة تُقوِّض النزاهة البحثيـة (DORA, 2012; Hicks et al., 2015).

كما أن هذه النظم تواجه تحديات بنيوية تعوق تحقيق الجودة؛ فالتحيزات التخصصية ؛ مثل تفضيل التخصصات العالية الاستشهادات؛ مثل الطب تُنتج تقييمات غير متوازنة؛ فالأبحاث الطبية تحقق استشهادات سريعة لوفرة النشر، في حين تتطلب أبحاث الأدب أو التاريخ سنوات لإثبات التأثير؛ مما يُقلل من قيمتها في النظم الكمية (Larivière & Sugimoto, 2019)، كما أن الضغط لزيادة المنشورات يُحفز ممارسات غير أخلاقية؛ مثل تقسيم الأبحاث، أو النشر في مجلات مفترسة، أو التلاعب بالبيانات؛ مما يُضعف الأصالة والمصداقية (Moher et al., 2020; Pontika et al., 2022)، كذلك يصعب قياس الأبعاد النوعية؛ مثل الأثر التعليمي أو التأثير المجتمعي

الطويل الأمد لتعقيدها؛ مما يتطلب نهجاً هجيناً يجمع المؤشرات الكمية والتحليلات النوعية (Wilsdon et al., 2017).

ولذلك استجابت بعض المبادرات العالمية لهذه التحديات بإعادة تشكيل نهج التقييم؛ فإعلان سان فرانسيسكو (DORA) يدعو إلى التخلي عن الاعتماد الأحادي على معامل التأثير مُشجعاً تقييم المخرجات بناءً على جودتها وقيمتها العلمية (DORA, 2012)؛ وبيان ليدن دعا إلى مؤشرات متنوعة تُراعي السياقات التخصصية مع إدماج مراجعة الأقران النوعية لضمان العدالة (Hicks et al., 2015)؛ كما تركز مبادئ هونغ كونغ على تعزيز النزاهة من خلال مكافأة الممارسات المسؤولة؛ مثل الإفصاح عن البيانات والالتزام بالأخلاقيات وتجنب الاستشهادات غير المبررة (Moher et al., 2020)؛ وتُوسع هذه المبادرات نطاق المخرجات المُقيَّمة ليشمل براءات الاختراع، مجموعات البيانات، البرمجيات المفتوحة المصدر، والتقارير السياسية؛ مما يعكس تنوع مساهمات البحث (Wilsdon et al., 2017).

يُعد النشر المفتوح إستراتيجية مهمة لتعزيز جودة المعرفة وإتاحتها؛ فالأبحاث المفتوحة تحقق استشهادات أعلى مقارنة بالمنشورات التقليدية، خاصة في النشر الأخضر -الإيداع في مستودعات مفتوحة- الهجين -الجمع بين النشر التقليدي والمفتوح- (Piwowar et al., 2018)؛ ينبع هذا التأثير من زيادة الوصول العالمي؛ مما يعزز التعاون والتأثير الأكاديمي، ولكن تكاليف النشر الذهبي المرتفعة التي قد تصل إلى 3000 دولار للمقالة تشكل تحدياً للباحثين من مؤسسات ذات موارد محدودة (Pontika et al., 2022)، توفّر أدوات؛ مثل Unpaywall وoaDOI حلولاً للوصول المجاني إلى 47% من الأدبيات العلمية (Piwowar et al., 2018)، و تكشف الفجوات التخصصية عن تفـــاوت كبيـر في النشر المفتوح؛ إذ يبلغ النشر المفتوح في الصيدلة 7% مقابل 87% في الفيزياء الفلكية مثلاً، وهذا يبرز الحاجة إلى سياسات تمويل عادلة تشجع النشر المفتوح عبر التخصصات (Piwowar et al., 2018).

تفيد الدراسات أن الأبحاث المتعددة التخصصـات تُحقق تأثيراً علمياً أكبـر؛ فقد أظهرت نتائج بحوث أن هذه الأبحاث تُحقق تأثيراً متأخراً كبيراً على المدى الطويل؛

مما يعكس قدرتها على معالجة التحديات المعقدة؛ كما أظهرت دراسة Zhang et al. (2022) أن التعاون الدولـــي في الأبحاث يؤدي إلى زيادة في عدد الاستشهـادات؛ مما يؤكد أهمية بناء شبكات بحثية عالمية لتعزيز التأثير العلمي(Wang et al., 2024)؛ في المقابل يواجه إصلاح نظم تقييم البحث العلمي مقاومة مؤسسيـة نابعة من الاعتماد الراسخ على المقاييس الكمية وخشية تقويض المصداقية؛ مما يستدعـي إشراك أصحاب المصلحة، وتطوير أدوات التقييم المرنة، وتعزيز ثقافة تقدير الجودة على الكمية (Hatch & Curry, 2020).

وهذا يتطلب تدريب الباحثين على الممارسات المسؤولة؛ مثل مشاركة البيانات وتوثيق المنهجيات، وتبنّي أنظمة تقييم تُكافئ تنوّع المخرجات، وتصميم آليات شاملة تدمج بين المؤشرات الكمية والنوعية.

المحور الثاني: المعايير الدولية لتقييم جودة البحث العلمي

تُشكّل المعايير الدولية لتقييم جودة البحث العلمـي أطراً تمكّن المؤسسـات من قياس أدائها وتأثيرها عالمياً؛ إذ تجمع بين قواعد بيانات ببليومترية؛ مثل Web of Science، وScopus، وSCImago التي تُوفّر بيانات كمية عن المنشورات والاستشهادات، وأطر تقييم شاملة؛ مثل إطار التميز البحثي (Research Excellence Framework – REF) وتصنيفات QS World University Rankings التي تدمج بين المقاييس الكمية والنوعية، وتهدف هذه الأدوات إلى تقييم الأثر الأكاديمي والمجتمعي؛ إلا أنها تواجه تحديات تتعلّق بالعدالة، والشمولية، والتكيّف مع السياقات المتنوعة.

تُعد قواعد البيانات الببليومترية ركيزةً أساسية في هذه المعايير؛ فـ Web of Science، التي طورتها Clarivate Analytics، تُركّز على المجلات ذات المراجعة العالية الجودة، مقدّمة أدوات؛ مثـل Journal Citation Reports، لتحليل سمعـة المجلات بناءً على الاستشهادات ومعامل التأثير(Clarivate, 2024b, 2024c)، وتُستخدم هذه القاعدة في تقييـم الأفراد لأغراض الترقية وتصنيفـات الجامعـات، غير أن صرامتها تُقصي المنشورات غير المفهرسة دولياً؛ مما يحدّ مـن ظهـور الأبحاث المحلية (Hicks et al., 2015)، أما Scopus، التـي تديرهـا Elsevier،

فتُغطي أكثر من 25,000 مجلة إلى جانب أوراق المؤتمرات والكتب، مقدّمة مقاييس؛ مثل CiteScore لقياس متوسط الاستشهادات على مدى أربع سنـــوات، وSNIP الذي يُصحّح التحيزات التخصصية من خلال تطبيع الاستشهادات بحسب المجال العلمي (Elsevier, 2024)، وتُعتبر Scopus مفضّلة لتقييم التخصصات الناشئة والتعاون الدولي بفضل تغطيتها الواسعة، في المقابل يعتمـــد SCImago Journal Rank (SJR) على خوارزمية PageRank التي تُوزن الاستشهادات بناءً على سمعة المجلة؛ مما يُقلّل من تأثير التخصصات العالية الاستشهادات، ويُوفّر تصنيفات مجانية تُتيح الوصول إلى المؤسسات ذات الموارد المحدودة (SCImago, 2024).

تعتمد هذه القواعد على مؤشرات أداء رئيسية؛ فمعامل التأثير، المحسوب كنسبة الاستشهادات إلى عدد المقالات خلال سنتين، يُستخدم لقياس جودة المجلات، لكنه يُعاني تحيّزات تُفضّل التخصصات العالية الإنتاجية كالطب؛ إذ تصل معاملات تأثير مجلاتـه إلى 50، في حين نادراً ما تتجاوز مجلات العلوم الإنسانية 5؛ مما يؤدي إلى تقييمـــات غير عادلة (Larivière & Sugimoto, 2019)؛ أما مؤشـــر هيـــرش (h-index)، الذي يعكس إنتاجية الباحث وتأثيره من خلال عدد المقالات التي حصلت على استشهادات تساوي أو تزيد عن عددها؛ فيُفضّل الباحثين ذوي المسيرة الطويلة؛ مما يُقلّل فرص الباحثين المبتدئين (Aksnes et al., 2019) ويُعد عدد الاستشهادات مؤشراً مباشراً للتأثير، لكنه يتأثر بممارسات؛ مثل الاستشهادات الذاتية أو شبكات التعاون الأكاديمية التي قد تُضخّم التأثير دون أن تعكس الجودة الحقيقية (Bornmann & Marx, 2014)، ولضمان الإنصاف تتطلب هذه المؤشرات تطبيعاً بحسب التخصص؛ مثل استخدام الاستشهادات النسبية (Relative Citation Ratio) التي تُصحّح الفروقات التخصصية.

توفّر أطر التقييم العالمية نهجاً متعدد الأبعاد لتقييم جودة البحث العلمي؛ فإطار التميز البحثي (REF) في المملكة المتحدة يُقيّم المخرجات بنسبة 65%، والأثر المجتمعي بنسبة 20%، وبيئة البحث بنسبة 15%، مع التركيز على دراسات الحالة النوعية التي توثّق التغييرات خارج الأوساط الأكاديمية؛ مثل تحسين السياسات الصحية، أو دعم الصناعات المحلية (Stern, 2016)؛ إلا أن REF يتطلّب موارد كبيرة

لإعداد التقارير؛ مما يُشكّل تحدياً للمؤسسات ذات الإمكانيات المحدودة (Hicks et al., 2015)، أما تصنيفات QS World University Rankings؛ فتعتمد على السمعة الأكاديمية (40%)، وعدد الاستشهادات لكل عضو هيئة تدريس (20%)، ومستوى التدويل (10%) (QS, 2024)، ومع ذلك فإنها تُعاني تحيّزات لصالح الجامعات الكبرى في الدول المتقدمة؛ مثل هارفارد وأكسفورد، في حين تواجه الجامعات في الدول النامية صعوبات في الظهور بسبب قيود الموارد (QS, 2024).

تتطلب هذه المعايير تكيفاً مستمراً لمواكبة التطورات العلمية؛ فدمج مؤشرات بديلة (Altmetrics)؛ مثل التفاعل على وسائل التواصل الاجتماعي، وتنزيلات المقالات، أو التأثير على السياسات يُعزّز تقييم الأثر المجتمعي، وتُظهر تحليلات Altmetric.com أن الأبحاث ذات التفاعل الرقمي العالي تُسهم في التأثير المجتمعي أكثر من الأبحاث التقليدية (Leng, 2021)، كما تُوفر تحليلات البيانات الضخمة رؤى معمّقة حول التعاون الدولي والتأثير الطويل الأمد؛ مما يُتيح تصميم أنظمة تقويم ديناميكية تعكس التنوع الأكاديمي؛ ويتوقّف مستقبل التقييم على تحقيق التوازن بين الأدوات التكنولوجية والسياقات التخصصية لضمان معايير دولية تعكس الجودة وتُسهم في تقدم علمي عادل.

المحور الثالث: تقنيات الذكاء الاصطناعي في تقييم البحث العلمي

يُعيد الذكاء الاصطناعي تشكيل التقييم للبحث العلمي من خلال أدوات تحليلية، تُعزّز الكفاءة والدقة والعدالة، وتُتيح تقنيات معالجة اللغة الطبيعية (NLP) ونماذج اللغة الكبيرة (LLMs) تحليل النصوص العلمية بعمق؛ مما يُمكّن من تقييم الأصالة، وجودة المنهجيات، والتأثير بسرعة ودقة غير مسبوقة؛ فمثلاً تُظهر الدراسات أن نماذج؛ مثل RoBERTa المُهيّأة للنصوص العلمية تُظهر فعالية في تحديد الأصالة والأهمية في العناوين والملخصات محاكيةً التقييمات البشرية (Checco et al., 2021)، ومع ذلك فإن استخدام النصوص الكاملة يُقلّل من الدقة نظراً لتعقيدها؛ مما يتطلّب تعليمات محسّنة وبيانات إضافية؛ مثل الجداول أو الأشكال (Kousha & Thelwall, 2024)، وعلى الرغم من أن هذه النماذج تُحسّن من تحديد المساهمات العلمية فإنها لا تزال في حاجة إلى إشراف بشري لتقييم الجوانب المنهجية المعقّدة.

يُقدّم الذكاء الاصطناعي حلولاً فعّالة لتعزيز المعايير الدولية لتقييم جودة البحث العلمي؛ فخوارزميات التعلّم الآلي؛ مثل تلك المستخدمة في قاعدة Dimensions، تُحلّل المنشورات من مصادر متنوعة، بما في ذلك المستودعات المحلية وأوراق المؤتمرات؛ مما يُوسّع نطاق التقييم ليشمل الأبحاث غير المفهرسة في Web of Science أو Scopus (Herzog et al., 2020)؛ كما تُحسّن نماذج معالجة اللغة الطبيعية دقة المؤشرات من خلال تحليل سياق الاستشهادات، كاشفةً عن الاستشهادات غير المبرّرة أو التأثير المُبالغ فيه الناتج من شبكات التعاون (Waltman & van Eck, 2019)، وتُسهم أدوات؛ مثل OpenAlex، في تحليل شبكات الاستشهادات بتكلفة منخفضة؛ مما يدعم المؤسسات ذات الموارد المحدودة (Herzog et al., 2020)، وتُظهر الدراسات أن تكامل تقنيات الذكاء الاصطناعي يُعزّز دقة التقييم من خلال تحسين الشفافية وتقليل التحيّزات التخصصية (Bibri et al., 2024).

تُسرّع تقنيات الذكاء الاصطناعي عملية مراجعة الأقران من خلال أتمتة التحقق من معايير النشر، وكشف الانتحال، وتوزيع المخطوطات بناءً على مجالات الخبرة، وتُسهم أنظمة؛ مثل ScholarOne وEditorial Manager، المدعومة بخوارزميات التعلّم الآلي، في تقليل وقت المراجعة؛ مما يُعزّز كفاءة النشر العلمي (Checco et al., 2021)، كما تُستخدم نماذج؛ مثل SciBERT المُدرّبة على ملايين النصوص العلمية لتحليل جودة المنهجيات واكتشاف الانتحال؛ مما يدعم النزاهة الأكاديمية؛ إذ يمكنها رصد النصوص المعاد صياغتها بطريقة غير أخلاقية، وتقليل الاعتماد على التقديرات البشرية في اكتشاف الانتحال (Beltagy et al., 2019)؛ ومع ذلك، تظلّ هذه النماذج محدودة عند الاعتماد على الملخصات فقط؛ نظراً لافتقارها إلى التفاصيل الكافية لتقييم الدقة المنهجية؛ مما يستدعي دمج النصوص الكاملة أو البيانات التكميلية في عملية التحليل (Kousha & Thelwall, 2024).

يمتد تأثير الذكاء الاصطناعي إلى دعم الكتابة الأكاديمية؛ إذ تُحسّن أدوات؛ مثل Grammarly وWritefull، جودة النصوص من خلال تصحيح الأخطاء النحوية، وتحسين الأسلوب، واقتراح صياغات متماسكة، وتُفيد هذه الأدوات الباحثين غير الناطقين بالإنجليزية عبر تقليل الحواجز اللغوية وتعزيز جودة التواصل العلمي،

مع ذلك تُثير هذه الأدوات قضايا أخلاقية؛ مثل خطر الانتحال غير المقصود في حال عدم الإفصاح عن استخدامها، وضرورة تحديد المسؤولية عن الأخطاء أو التحيّـــزات في النصوص الناتجـــة (Lund et al., 2023)، وتُوصي الدراسات بوضع إرشادات أخلاقية تُلزم الإفصاح عن استخدام أدوات الذكاء الاصطناعي إلى جانب أنظمة تدقيق تضمن النزاهة الأكاديمية (Lund et al., 2023).

تُحدث أدوات؛ مثل ASReview وRayyan ثورة في المراجعات الأدبية المنهجية من خلال استخدام تقنيات استخراج البيانات لتحديد الأوراق ذات الصلة وتجميع البيانات؛ مما يُقلّل وقت التحليل ويزيد من الدقة عبر تقليل الأخطاء البشرية؛ وهو ما يجعلها أدوات لا غنى عنها في التخصصات ذات الأدبيات الواسعة (de Bruin et al., 2025)، كما تُوفّر قواعد البيانات المفتوحة؛ مثل OpenCitations وAltmetric.com، مؤشرات فعّالة لتقييم التأثير الأكاديمي والاجتماعي، وتُظهر تحليلات Altmetric.com أن الأبحاث ذات التفاعل الرقمي العالي؛ مثل المناقشات على منصات كـ X، تُسهم في التأثير المجتمعي بنسبة تفوق الأبحاث التقليدية (Leng, 2021).

يبرز الذكاء الاصطناعي التنبؤ بالتأثير البحثي كمجال واعد؛ إذ تُحلّل نماذج التعلّم الآلي عوامل؛ مثل عدد المؤلفين، والتعاون الدولي، وسمعة المجلة لتوقّع الاستشهادات المستقبلية، وتُظهر الشبكات العصبية العميقة أداءً متميّزاً في التنبؤ بالتأثير خلال خمس سنوات لا سيّما عند دمج الميزات النصية كالعناوين مع البيانات الببليومترية؛ مثل معامل التأثير (Abrishami & Aliakbary, 2019)، ومع ذلك تظلّ هذه النماذج محدودة في التخصصات ذات أنماط الاستشهادات المنخفضة؛ كالعلوم الإنسانية؛ بسبب ندرة البيانات التدريبية؛ مما يستدعي توافر بيانات أكثر تنوعاً (Abrishami & Aliakbary, 2019)، وتدعم هذه التقنيات النشر المفتوح من خلال تقليل تكاليف المراجعة وزيادة الكفاءة؛ إذ تُتيح أدوات؛ مثل ASReview تقييمات أولية تتوافق مع معايير التقييم الأكاديمي العالية؛ مثل REF (de Bruin et al., 2025).

يمثل وجود إطار هجين يدمج الذكاء الاصطناعي والإشراف البشري خطوة نوعية نحو أنظمة تقييم أكثر كفاءة وعدالة؛ إذ يُمكن لنماذج؛ مثل SciBERT تحليل النصوص

لتقييم الأصالة والتأثير، وتوليد تقارير آلية توضّح نقاط القوة والضعف (Beltagy et al., 2019; Checco et al., 2021)، مع احتفاظ المراجعين بدور محوري في التحقق من الدقة المنهجية التي تتطلب خبرات متخصصة (Checco et al., 2021)، ويُعزز هذا الإطار الشفافية من خلال توثيق القرارات ونشر المنهجيات على منصات مفتوحة؛ بما يُسهم في بناء الثقة ومعالجة مقاومة الأوساط الأكاديمية للأتمتة (Checco et al., 2021; Lund et al., 2023)، وينسجم هذا التوجه مع معايير دولية؛ مثل REF وPRISMA، على الرغم مما يواجهه من تحديات تقنية كاعتماده على جودة البيانات التدريبية خاصة في التخصصات الناشئة (de Bruin et al., 2025; Kousha & Thelwall, 2024)، وعلى الصعيد الأخلاقي تبرز مخاوف تتعلق بالتحيّزات الخوارزمية؛ كتمييز المؤسسات المرموقة؛ مما يستدعي تنويع البيانات وتطوير خوارزميات مفتوحة قابلة للتدقيق (Leng, 2021; Lund et al., 2023). إن تحقيق التوازن بين الأتمتة والخبرة البشرية لا يُعد تحدياً تقنياً فحسب، بل هو رؤية مستقبلية تتطلب تعاوناً بين الباحثين، والمؤسسات، ومطوّري التكنولوجيا، ومن خلال تبنّي هذه الأدوات بحذر وشفافية، تستطيع الأوساط الأكاديمية إعادة تشكيل مشهد التقييم العلمي ليصبح أكثر شمولية، ودقة، وتأثيراً.

المنهج

استُخدِم منهج تحليل المحتوى لتقييم البحوث المنشورة في المجلات اليمنية المصنفة في تصنيف أرسيف بمساعدة الذكاء الاصطناعي وفق استمارة تقييم بنيت لهذا الغرض.

مجتمع الدراسة

اختيرت المجلات اليمنية الأعلى جودة بناءً على تصنيف أرسيف العربي، وهو مرجع موثوق به ومعتمد لقياس جودة المجلات العلمية العربية وفق معايير علمية دقيقة، وهو خيار موثوق به لمستوى البحث العلمي في اليمن للأسباب الآتية:

- خضعت لتقييم مستقل يضمن الالتزام بمعايير علمية دقيقة يجعلها عينة ممثّلة وموثوقاً بها لمستوى البحث العلمي المؤسسي في اليمن.

- تحظى باعتراف أكاديمي إقليمي ودولي؛ ما يجعلها مرجعاً موثوقاً به لبحث جودة البحث العلمي في اليمن.
- تضمن تجنب المجلات ذات المعايير الضعيفة التي قد لا تحقق شروط النشر العلمي الرصين.
- توفر معياراً علمياً موثوقاً به، يُمكّن من إجراء تقييم كمّي دقيق، وإجراء مقارنات علمية، واستخلاص نتائج موضوعية قابلة للتعميم على واقع البحث العلمي في اليمن.
- توفر بيانات قابلة للقياس تتيح إجراء تحليل علمي منهجي، يمكن توظيفه في تطوير المجلات اليمنية غير المصنفة.

وبذلك يعكس اختيار هذه المجلات رؤية علمية لتقييم واقع البحث اليمني وفق معايير دولية، ويعزّز دقة نتائج الدراسة وموضوعيتها، ويعرض جدول 1 خصائص مجتمع الدراسة.

جدول 1

خصائص مجتمع الدراسة

م	عنوان الدورية	معامل "أرسيف"		الناشر	المدينة
1	الآداب للدراسات النفسية والتربوية	1.0755	مرتفع	جامعة ذمار – كلية الآداب	ذمار
2	المجلة العربية لضمان جودة التعليم الجامعي	0.7949	متوسط	جامعة العلوم والتكنولوجيا	عدن
3	مجلة الدراسات الاجتماعية	0.4872	منخفض	جامعة العلوم والتكنولوجيا	عدن
4	مجلة العلوم التربوية والدراسات الإنسانية	0.4846	منخفض	جامعة تعز – دائرة الدراسات العليا والبحث العلمي	التربة
5	مجلة الأندلس للعلوم الإنسانية والاجتماعية	0.3068	منخفض	جامعة الأندلس للعلوم والتقنية	صنعاء
6	أبحاث	0.2195	منخفض	جامعة الحديدة - كلية التربية	الحديدة
7	المجلة الدولية لتطوير التفوق	0.1852	منخفض جداً	جامعة العلوم والتكنولوجيا - مركز تطوير التفوق	عدن

تابع/ جدول 1

خصائص مجتمع الدراسة

م	عنوان الدورية	معامل "أرسيف"		الناشر	المدينة
8	الآداب	0.0702	منخفض جداً	جامعة ذمار – كلية الآداب	ذمار
9	مجلة المهرة للعلوم الإنسانية	0.1111	منخفض جداً	جامعة حضرموت – كلية التربية	الغيضة
10	مجلة الآداب للدراسات اللغوية والأدبية	0.063	منخفض جداً	جامعة ذمار – كلية الآداب	ذمار
11	مجلة جامعة حضرموت للعلوم الإنسانية	0.0227	منخفض جداً	جامعة حضرموت – الدراسات العليا	حضرموت
12	مجلة جامعة عدن للعلوم الطبيعية والتطبيقية	0.0143	منخفض جداً	جامعة عدن	عدن
13	مجلة جامعة القرآن الكريم والعلوم الإسلامية	بلا استشهادات	غير مصنّفة	جامعة القرآن الكريم والعلوم الإسلامية	صنعاء
14	مجلة الأندلس للعلوم التطبيقية	بلا استشهادات	غير مصنّفة	جامعة الأندلس للعلوم والتقنية	صنعاء
15	مجلة العلوم والتكنولوجيا	بلا استشهادات	غير مصنّفة	جامعة العلوم والتكنولوجيا – كلية الهندسة والحاسبات	عدن

استُبعدت مجلات جامعة القرآن الكريم، والأندلس، والعلوم والتكنولوجيا من مجتمع الدراسة لعدم تسجيل استشهادات في تقرير أرسيف (2024)؛ مما يُضعف موثوقية تقييمها كمياً ويدل على محدودية تأثيرها الأكاديمي؛ وبهذا اقتصر مجتمع الدراسة على 12 مجلة مصنفة ضمن أرسيف.

عينة الدراسة

اختيرت عينة متنوعة من سبع مجلات تغطي تخصصات رئيسية (الإنسانية، الاجتماعية، التربوية، والتطبيقية) لضمان تقييم شامل ومتوازن، مع مراعاة الأعلى تصنيفاً في معاملات أرسيف، والتوزيع العادل بين الجامعات، واستيفاء شروط النشر

والتحكيم والاتساق مع أهداف الدراسة؛ وُجه نموذج (LLMs) لاقتراح العينة وفق هذه المعايير، مع مراجعة بشرية للتأكد من دقتها، ويوضح جدول 2 خصائص العينة.

جدول 2

خصائص عينة الدراسة

م	عنوان الدورية	مجال التخصص	معامل "أرسيف"		الناشر	المدينة
1	الآداب للدراسات النفسية والتربوية	الدراسات النفسية والتربوية	1.0755	مرتفع	جامعة ذمار – كلية الآداب	ذمار
2	المجلة العربية لضمان جودة التعليم الجامعي	جودة التعليم	0.7949	متوسط	جامعة العلوم والتكنولوجيا	عدن
3	مجلة الدراسات الاجتماعية	العلوم الاجتماعية	0.4872	منخفض	جامعة العلوم والتكنولوجيا	عدن
4	مجلة العلوم التربوية والدراسات الإنسانية	العلوم التربوية	0.4846	منخفض	جامعة تعز – دائرة الدراسات العليا والبحث العلمي	التربة
5	أبحاث	العلوم الإنسانية	0.2195	منخفض	جامعة الحديدة – كلية التربية	الحديدة
6	مجلة الآداب للدراسات اللغوية والأدبية	الدراسات اللغوية والأدبية	0.063	منخفض جداً	جامعة ذمار – كلية الآداب	ذمار
7	مجلة جامعة عدن للعلوم الطبيعية والتطبيقية	العلوم التطبيقية	0.0143	منخفض جداً	جامعة عدن	عدن

اختير عددان من كل سنة للمجلة، ومن كل عدد اختير بحث، باستثناء مجلة واحدة اختير منها 8 أبحاث وفق المعايير الآتية:

- الحداثة الزمنية (2020–2024): لأنها تعكس التوجهات الحديثة، وتنسجم مع معايير الجودة العالمية من حيث استخدام التقنية، ودقة التوثيق، وقياس الأثر، وتمثل تطور جودة المجلات اليمنية في ظل التحديات السياسية والاقتصادية، وتتوافر في قاعدة "أرسيف" بما يسهل جمعها وتحليلها.

- التنوع المنهجي: شملت تصاميم كمية ونوعية ومختلطة.
- التغطية التخصصية: راعت تمثيل مختلف الفروع داخل كل مجال.
- سهولة الوصول: اقتصر الاختيار على الأبحاث المتاحة بالنشر المفتوح أو ضمن قواعد البيانات الأكاديمية.

أدوات الدراسة

1 - أداة التقييم ومعايير التحليل

صُمّمت أداة تقييم معيارية (Rubric) تضمنت ثلاثة مجالات رئيسية: جودة البحث العلمي، المعايير الأخلاقية، والأثر الأكاديمي والمجتمعي، احتوت على 11 معياراً و55 مؤشراً فرعياً، وفق الخطوات الآتية:

- استقصاء المرجعيات الدولية: وُجهت نماذج الذكاء الاصطناعي (ChatGPT-openAI, Google AI Studio, Grok)، لاستخلاص معايير النشر المعتمدة في قواعد؛ مثل Scopus وWeb of Science، وأطر النزاهة والشفافية من مؤسسات؛ مثل OECD وCOPE (Committee on Publication Ethics)، وإرشادات الأكاديمية الوطنية للعلوم(NAS)، والجمعية الأوروبية لأخلاقيات البحث (EASE)، واليونسكو.
- تحليل الإطار الموحد: قُورنت المعايير لتحديد العناصر المشتركة (كالأصالة والمنهجية)، والمتباينة (كالتأثير التطبيقي)، واختيرت المؤشرات الأكثر ملاءمة للسياق اليمني.
- مقارنة مع الأدبيات: رُوجع الإطار الموحد مع دراسات سابقة (Hicks et al., 2015; Moher et al., 2020; Pontika et al., 2022; Wilsdon et al., 2017) للتأكد من شموليته وموضوعيته.
- صياغة الأداة الأولية: وبناءً عليها صممت الأداة في صورتها الأولية على شكل استمارة تقييم تتضمن: المجال، المعيار، المؤشر، آلية التحليل، طريقة التقدير، وتضمنت الأداة:
  - جودة البحث العلمي وتشمل معايير: الابتكار، وضوح المشكلة، دقة المنهجية، التوثيق والمصداقية.
  - المعايير الأخلاقية وتشمل معايير: النزاهة العلمية، الشفافية، احترام الملكية الفكرية.

- الأثر ويشمل معايير: الاستشهادات، التأثير الاجتماعي والاقتصادي.

تكونت الأداة من 11 معياراً بـ 55 مؤشراً فرعياً موزعة في الجوانب الثلاثة، مع إهمال الجوانب الإدارية والتجارية والأثر الواسع في التصنيفات العالمية لتكييفه مع تحليل المحتوى البحثي في السياق اليمني؛ مما جعله دقيقاً وملائماً على الرغم من محدوديته مقارنة بالمعايير الدولية الشاملة.

صدق الأداة (المحتوى- الظاهري). خضعت الأداة لتحكيم خمسة خبراء في البحث العلمي والتقييم الأكاديمي لضمان توافقها مع المعايير الدولية، وبناءً على ملاحظاتهم عُدِّلت بعض صياغات مؤشرات الابتكار والأثر الاجتماعي والاقتصادي لتعزيز وضوحها وشفافية التطبيق، واعتمدت المؤشرات التي نالت موافقة 80% فأكثر؛ مما يعكس اتفاقاً عالياً على صدق الأداة وملاءمتها لأهداف الدراسة.

ثبات الأداة. أُعيد تحليل عينة الأبحاث نفسها بعد ثلاثة أسابيع، وحسب معامل الارتباط بيرسون (Pearson Correlation) بينهما، وبلغ 0.899؛ مما يشير إلى ثبات مرتفع جداً، كما رُفعت العينة واستمارة التحليل إلى نماذج الذكاء الاصطناعي ChatGPT، Grok، وAI Studio، طُلب منها إجراء التحليل عشر مرات وفق أوامر مخططة مسبقاً، وحساب نسبة الاتفاق بين التحليل الأول والتحليل الذي يليه، حقق ChatGPT و AI Studio نسب اتفاق راوحت بين 0.741 و0.921، في حين سجل Grok نسباً بين 0.992 و1.00، وبلغت نسبة الاتفاق بين Grok وتحليل الباحثَين 0.861، وهي نتيجة متوافقة مع ما أشار إليه Qiu et al. (2025) حول تقارب أداء البشر والذكاء الاصطناعي؛ وبطلب أمثلة من Grok للاطمئنان إلى دقة التحليل، أظهرت الأمثلة نتائج عالية الدقة، تتجاوز الأداء البشري؛ ما رجّح استخدامه لإجراء عملية تحليل محتوى الأبحاث وفق استمارة التقييم المشار إليها.

الاتساق الداخلي. استُخدمت نتائج التحليل الأول لعينة الأبحاث في حساب معامل ألفا كرونباخ، وقد بلغت القيمة الكلية 0.98؛ مما يدل على اتساق داخلي مرتفع جداً، ويعكس موثوقية الأداة، فيما راوحت القيم الفرعية بين 0.936 و0.959؛ بما يعكس تجانساً قوياً بين مكونات الأداة وصلاحيتها للتحليل.

2 - ضبط نموذج الذكاء الاصطناعي (LLMs) كأداة لتحليل الأبحاث العلمية

يُعدّ استخدام نماذج اللغة الكبيرة (LLMs) في تحليل الأبحاث العلمية نقلة نوعية في آليات التقييم الأكاديمي، لكنه يتطلب ضبطاً صارماً يضمن دقة المخرجات وموثوقيتها؛ وقد أشارت دراسات إلى أن وجود موجّه (Prompt) دقيق وواضح يُعدّ شرطاً لتوجيه سلوك النموذج وضبط استجاباته وفق منطق علمي (Tonmoy et al., 2024). ولضمان أداء النموذج وفق المعايير الأكاديمية اعتمد إطار ضبط تكراري متعدد المراحل امتدّ العمل عليه 160 ساعة، خضع خلاله النموذج لاختبارات تحقق متداخلة، هدفت إلى:

- تقليل التحيز (Bias Mitigation).
- معالجة الهلوسة الاصطناعية (Hallucination Control).
- تعزيز الصرامة المنهجية في التقييم (Rigorous Evaluation Enforcement).

وعليه؛ بُني نموذج ضبط تكراري عالي الدقة لتقييم جودة الأبحاث العلمية باستخدام الذكاء الاصطناعي التوليدي (LLMs) وفق ثلاث مراحل:

المرحلة الأولى - الضبط الأولي (Pre-Tuning). وتشمل إعداد النموذج لتقييم الأبحاث من خلال تصميم الأوامر وتحديد المعايير والتسلسل المنهجي، وتتضمن توضيح المهمة للنموذج لضمان فهم طبيعتها؛ بحيث لا يتعامل معها كنص عام بل يُحلل وفق معايير موضوعية محددة:

- تحديد الهدف بدقة ليكون الدور تحليلياً وتقييمياً وليس مجرد ملخص أو مستنتج ذاتياً.
- تزويد النموذج بمعايير تقييم صارمة مع إخضاعه لاختبارات متعددة لضمان الالتزام بها، وتعديل الأوامر تفاعلياً لمنع الاجتهاد غير العلمي.
- ضبط الأوامر والتوجيهات كأداة تقييم علمي عبر عمليات بحثية تكرارية وآليات دقيقة شملت تجارب تفاعلية، اختبارات دقيقة، وتعديلات متكررة لضمان الصرامة المنهجية وفق ما يأتي:

- التجربة الأولية، وهي: اختبار استجابة النموذج للأوامر وتحليل مدى التزامه بالمعايير وتحديد القصور.
- تحسين الأوامر، ضبط التعليمات تدريجياً لضمان استخدام استمارة التقييم كمرجع صارم ومنع الاجتهاد الذاتي.
- توظيف أدوات هندسة الأوامر (Prompt Engineering)، وفيه يستعان بنظامين متخصصين من ChatGPT-openAI لتطوير تعليمات دقيقة تحكم آلية التقييم.
- إجراء اختبارات مقارنة بين استجابات النموذج قبل ضبط الأوامر وبعده، والتحقق من دقتها مقارنة بالتحليل البشري.
- اعتماد الأوامر النهائية بعد اختبارات مكثفة لمنع التفسير غير الموضوعي وإلزام النموذج إجراءات محددة؛ مثل: تحليل وفق المعايير فقط، عرض النتائج الكمية قبل الكيفية، الالتزام بالقوالب الأكاديمية.
- ضبط الأوامر بتعليمات تمنع التفسيرات الشخصية؛ مثل: (حلل وفق المعايير دون استنتاج شخصي، لا تقدم تقييماً قبل العمليات الإحصائية، اعرض النتائج في جداول منظمة، انتقل إلى التحليل الوصفي بناءً على البيانات الرقمية).
- إضافة مرجع تحكيمي لكل مؤشر، تمثل في "آلية التحليل، وطريقة التقدير" لمنع التقييم غير المدعوم بالأدلة العلمية.

وبذلك أصبحت الأوامر النهائية أداة تقييم علمية موثوق بها، تضمن تحليل الأبحاث وفق معايير دقيقة؛ بما يعزز دقة النتائج وموثوقيتها، ويضمن هذا الإطار المنهجي منع أي تلاعب أو أخطاء تقييمية ناتجة من تفسيرات غير منهجية.

المرحلة الثانية – التقييم التفاعلي (Interactive Evaluation). وتشمل اختبار النموذج وتحسين الأوامر من خلال تجارب تفاعلية لضمان الالتزام بالمعايير الأكاديمية وتعزيز الدقة في التقييم، وتتضمن الخطوتين الآتيتين:

1 – تصميم آلية تنفيذ التقييم تدريجياً ومنظماً؛ لضمان عدم إغفال أي معيار أو القفز بين المهام بشكل غير منهجي من خلال:

- فرض تسلسل منطقي يمنع النموذج من تخطي الخطوات مع إدراج قائمة مراجعة داخلية، تفرض التحقق من اكتمال كل مرحلة قبل الانتقال إلى التالية لها وفق الترتيب:
  - استخراج بيانات البحث.
  - تقييم كل معيار بشكل مستقل وفق المؤشرات الخمسة الخاصة به.
  - حساب المتوسطات والانحرافات المعيارية.
  - تقديم تحليل كيفي داعم للتقييم الكمي.
  - إنتاج تقرير تحليلي وفق القوالب العلمية.
- استخدام قائمة مراجعة داخل الأوامر، تفرض التأكد من إكمال جميع الخطوات قبل الانتقال إلى المرحلة التالية؛ مما يعزز دقة التقييمات واتساقها.

2 - ضبط العمليات الإحصائية لضمان استناد التحليل الكمـــي إلى بيانات فعليـــة لا تقديرات وصفية من خلال:

- إلزام النموذج حساب المتوسطات والانحرافات لكل معيار ومؤشر.
- عرض جميع الحسابات في جداول قبل إصدار أي استنتاج.
- إدخال تعليمات تمنع اللغة الوصفية الغامضة؛ مثل "يبدو البحث متوافقاً"، وإلزامه تقديم بيانات كمية.
- إدراج أوامر تحقق مزدوج؛ لضمان دقة العمليات الحسابية، وفرض تحليل الفروق لمنع إعطاء تقييمات متقاربة غير منطقية.

المرحلة الثالثة - التحقق والتحسين (Validation & Refinement).

وتشمل التحقق من دقة النموذج واستقراره، وإنتاج تقارير موثوق بها، ومعالجة التحيز، والتكرار لضمان موثوقية التقييمات وتتضمن الخطوات الآتية:

1 - اختبار دقة أداء النموذج واستقراره قبل الاعتماد عليه؛ للتأكد من تقديمه نتائج متسقة عند إعادة تقييم الأبحاث نفسها؛ مما يضمن موثوقية الأداة وثباتها عبر الزمن.

- تنفيذ اختبار ثبات إعادة التطبيق (Test-Retest Reliability): إعادة تحليل الأبحاث نفسها بعد فترة، ومقارنة التقييمات بمعامل بيرسون، وأظهرت النتائج معامل قدره 0.992؛ ما يشير إلى ثبات عالٍ.
- مقارنة مخرجات النموذج مع التحليل البشري لضمان دقة الاستنتاجات، مع اختبار أدائه عند إدخال تعديلات على صياغة المعايير لمعرفة مدى تأثر استجابته بالصياغة المستخدمة.

2 - ضبط الاستجابات النهائية وإنتاج تقارير موثوق بها تتبع القوالب العلمية وتصلح للاستخدام الأكاديمي من خلال:

- إلزام النموذج بإنتاج تقارير وفق القالب:
  - مقدمة توضح البحث محل التقييم.
  - تحليل كمي يعرض المتوسطات والانحرافات لكل معيار.
  - تحليل كيفي يبرز نقاط القوة والضعف في البحث.
  - استنتاجات وتوصيات أكاديمية لتحسين جودة الأبحاث المستقبلية.
- التأكد من عدم إصدار النموذج استنتاجات قبل عرض البيانات الرقمية لكل معيار.
- مراجعة عينة من تقارير النموذج التجريبية لضمان خلوها من التحيز، وإدراج تعليمات خاصة للتعامل مع البيانات غير المكتملة أو التي تحتاج إلى تفسير إضافي.

3 - إدخال تحقق مزدوج (Double-Validation Process)؛ لمنع أي انحرافات محتملة في التقييم، والتأكد من أن كل بحث يراجع مرتين على الأقل بآلية تحقق مستقلة من خلال:

- تكرار التقييم لكل بحث بفاصل زمني، مع مقارنة النتائج لتحديد مدى الثبات.
- إلزام النموذج تقديم تقرير مراجعة ثانية (Re-evaluation Report) في حال وجود تباين في النتائج بين الجولتين.
- إضافة "آلية التدقيق العشوائي"؛ إذ تفحص عينة من الأبحاث يدوياً؛ للتحقق من دقة تقييم النموذج.

4 - ضبط استجابة النموذج؛ لمنع التحيز والتكرار وضمان أن النموذج لا يعيد إنتاج تقييمات متكررة لجميع الأبحاث؛ مما قد يـــؤدي إلى نتائج غير دقيقة أو غيـــر تمييزية، وقد اتخذت الإجراءات الآتية:

- تنويع التقييمات بناءً على تعليقات تفاعلية تعكس الفروقات الفعلية في جودة الأبحاث؛ مما يمنع إصدار تقييمات متماثلة.
- إضافة قياس تأثير القرب (Proximity Effect)؛ لمنع منح درجات متشابهة لأبحاث ذات جودة متباينة.
- تم اكتشاف بعض الأنماط المتكررة في تقييمات النموذج من خلال الاختبارات الأولية؛ ومن ثم، أدخلت آلية تفرض عليه إعادة صياغة التحليل بناءً على نقاط القوة والضعف الفعلية لكل بحث.

بهذه الخطوات ضبط نموذج (LLMs) لتقديم تقييمات دقيقة، وموضوعية، ومستقرة لجودة الأبحاث؛ مما يعزز دقة النتائج ويضمن توافقها مع المعايير الدولية لقياس جودة البحث والنشر.

معيار الحكم

صُمّم تدريج المستويات في هذه الدراسة (ممتاز: 4.5–5.0، جيد جداً: 4.0–4.4، جيد: 3.5–3.9، مقبول: 3.0–3.4، ضعيف: أقل من 3.0) بناءً على تقسيم النطاق الكلي (0–5) إلى فئات متساوية تقريباً، يمثل كل مستوى نحو 10% من النطاق، باستثناء "ضعيف" الذي يشمل ما دون المتوسط المتوقع (3.0). وتميزت هذه الدراسة بإعطاء المجالات الثلاثة –جودة البحث العلمي، والأخلاقيات، والأثر– الوزن نفسه وفق مؤشراتها، على عكس التصنيفات العالمية التي تركز بشكل أكبر على الاستشهادات كمؤشر وحيد للأثر، وتعطيها أولوية على المنهجية أو النزاهة، وهو نهج قد يغفل جوانب أساسية للدقة العلمية.

النتائج

نتائج السؤال الأول

نص السؤال الأول: ما المعايير الدولية لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية؟ للإجابة عن هذا السؤال، ووفقاً للإجراءات المتبعة، حُدّدت قائمة المعايير بثلاثة مجالات رئيسية: جودة البحث العلمي (5 معايير)، الأثر (3 معايير)، والمعايير الأخلاقية (3 معايير)، وقد خُصص لكل معيار خمسة مؤشرات؛ ليصبح المجموع الكلي 55 مؤشراً موزعة.

نتائج السؤال الثاني

نص السؤال الثاني: ما النموذج التطبيقي لتقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام الذكاء الاصطناعي (LLMs)؟ للإجابة عن السؤال اتُّبعت منهجية دقيقة تشمل ثلاث مراحل لضبط نموذج تقييم جودة الأبحاث العلمية، هي:

1 - المرحلة الأولى: الضبط الأولي (Pre-Tuning).

2 - المرحلة الثانية: التقييم التفاعلي (Interactive Evaluation).

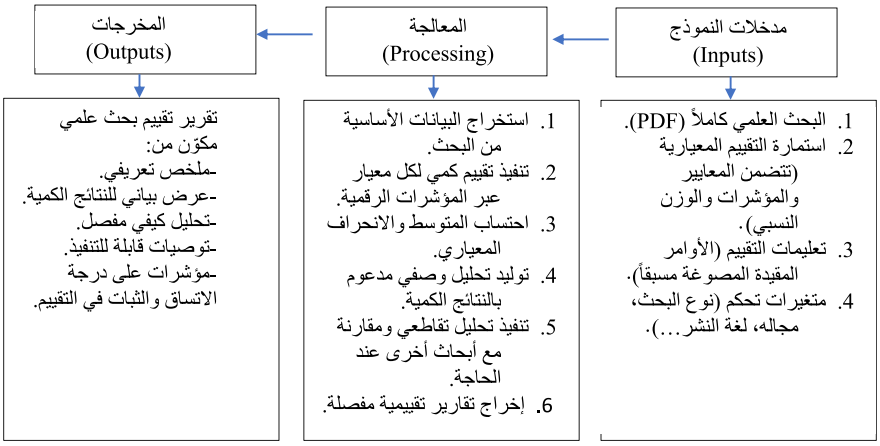
3 - المرحلة الثالثة: التحقق والتحسين (Validation & Refinement).

أسفرت هذه المراحل عن نموذج تكراري محسن لتقييم جودة الأبحاث العلمية.

التعريف الإجرائي للنموذج ومجاله: نموذج تقييمي يعتمد على الذكاء الاصطناعي (LLMs)، تم ضبطه وتحسينه باستخدام إطار صارم متعدد المراحل لتقييم جودة البحوث العلمية المنشورة في المجلات المحكمة بناءً على معايير كمية وكيفية محددة سلفاً وفق المكونات في شكل 1.

شكل 1

مكونات نموذج الضبط التكراري لتقييم جودة البحوث العلمية المقترح



ويمتاز النموذج بأنه:

- خاضع لاختبارات تحكم دقيقة لضبط التحيز والهلوسة.
- قابل لإعادة التطبيق مع ثبات في النتائج.
- موثوق به لاستخدامه في مراجعة الأبحاث ضمن المجلات أو المؤتمرات.
- يعتمد على تسلسل منطقي لا يسمح بتجاوز الخطوات التقييمية.

نتائج السؤال الثالث

نص السؤال الثالث: ما درجة توافق جودة البحوث في المجلات العلمية بالجامعات اليمنية مع معايير جودة البحث الدولية؟ وحللت المجلات المشمولة بالبحث باستخدام الإطار التقييمي المصمم لهذا الغرض، وأظهرت النتائج الواردة في الجدولين 3 و4 المتوسطات والانحرافات المعيارية لكل مجال ومعيار ومجلة.

جدول 3

المتوسطات والانحرافات المعيارية لكل مجلة على مستوى كل معيار

م	المعيار	الآداب النفسية		ضمان الجودة		الدراسات الاجتماعية		العلوم التربوية		أبحاث		الآداب اللغوية		جامعة عدن		الإجمالي		التقدير
		م	ع	م	ع	م	ع	م	ع	م	ع	م	ع	م	ع	م	ع	
1	الابتكار	3.3	0.78	3.4	0.80	3.2	0.75	3.1	0.74	3.0	0.77	3.1	0.70	3.3	0.78	3.20	0.75	مقبول
2	وضوح مشكلة البحث	3.6	0.66	3.7	0.64	3.5	0.67	3.4	0.66	3.3	0.64	3.4	0.66	3.6	0.66	3.50	0.66	جيد
3	دقة المنهجية	3.8	0.75	3.9	0.70	3.7	0.78	3.6	0.80	3.5	0.81	3.6	0.80	3.8	0.75	3.70	0.77	جيد
4	التوثيق والمصداقية	3.5	0.81	3.6	0.80	3.4	0.80	3.3	0.78	3.2	0.75	3.3	0.78	3.5	0.81	3.40	0.79	مقبول
5	وضوح النتائج	3.6	0.66	3.7	0.64	3.5	0.67	3.4	0.66	3.3	0.64	3.4	0.66	3.6	0.66	3.50	0.66	جيد
6	الأثر الأكاديمي	1.1	0.30	1.2	0.40	1.1	0.30	1.0	0.00	1.0	0.00	1.0	0.00	1.2	0.40	1.09	0.30	ضعيف
7	الأثر الاجتماعي	2.4	0.88	2.5	0.92	2.3	0.90	2.2	0.87	2.1	0.83	2.2	0.87	2.4	0.92	2.30	0.88	ضعيف
8	الأثر الاقتصادي	1.1	0.30	1.2	0.40	1.1	0.30	1.0	0.00	1.0	0.00	1.0	0.00	1.1	0.30	1.07	0.30	ضعيف
9	أصالة البحث	4.0	0.71	4.1	0.70	3.9	0.70	3.8	0.75	3.7	0.78	3.8	0.75	4.0	0.71	3.90	0.73	جيد
10	النزاهة العلمية	4.2	0.87	4.3	0.90	4.1	0.83	4.0	0.89	3.9	0.87	4.0	0.89	4.2	0.87	4.10	0.86	جيد جداً
11	حقوق الملكية الفكرية	4.3	0.78	4.4	0.80	4.2	0.75	4.1	0.83	4.0	0.77	4.1	0.83	4.3	0.78	4.20	0.79	جيد جداً
	الإجمالي	3.09	1.18	3.14	1.20	3.01	1.15	2.91	1.12	2.83	1.10	2.91	1.12	3.10	1.18	3.00	1.14	مقبول
	التقدير	مقبول		مقبول		مقبول		ضعيف		ضعيف		ضعيف		مقبول		مقبول		

أولاً: التقييم العام

تكشف البيانات من جدول 3 عن متوسط إجمالي لتقييم البحوث العلميـــة، وقد بلغ 3.00، بانحراف معياري مرتفع نسبياً (1.14)؛ مما يصنفها ضمن فئة مقبول، ويعكس هذا المتوسط تحقيق الحد الأدنى من المعايير الدولية المقبولة وفق مقياس التقييم (0-5) في هذه الدراسة؛ لكن التصنيفات العالمية -التي تركز بشدة على الاستشهادات- قد تصنفها أقرب إلى ضعيف؛ نتيجة الأثر المنخفض.

ثانياً: تحليل تفصيلي لمعايير البحث العلمي وفقاً للجدول

جودة البحث العلمي. تسجل المعايير المرتبطة بجـــودة البحوث العلمية متوسطات تراوح بين 3.20 و3.70، وهي تصنف بين "مقبول" و"جيد".

- وضوح مشكلة البحث (3.50 | 0.66): يظهر قدرة جيدة لدى الباحثين على تحديد إشكاليات واضحة ومحددة، متوافقة مع المعايير الدولية الأساسية لصياغة الأهداف، مع انحراف منخفض (0.66) يشير إلى استقرار الأداء عبر المجلات.
- الابتكار (3.20 | 0.75): يقع ضمن فئة المقبول؛ مما يشير إلى تحسين الأفكار الموجودة أكثر من تقديم ابتكارات جذرية، الانحراف المعياري (0.75) يعكس تبايناً متوسطاً بين المجلات؛ إذ تسجل بعضها (مثل "ضمان الجودة" 3.4) أداءً أفضل من أخرى (مثل "أبحاث" 3.0). مقارنة بالمعايير الدولية التي تتطلب أصالة عالية (4.5+). يظهـر هذا المعيار فجـوة واضحـة، قد ترجع إلى نقص المـوارد أو محدودية الوصول إلى قواعد بيانات حديثة تدعم التجديد البحثي.
- دقة المنهجية (3.70 | 0.77): يعكس التزاماً ملحوظاً بتصميم منهجيات علمية سليمة تدعم المصداقية الداخلية للأبحاث، لكن الوصول إلى المستوى الدولي (4.0+) يتطلب تعزيز إجراءات التحقق من الصدق والثبات، الغائبة في نسبة كبيرة من الدراسات.
- التوثيق والمصداقية (3.40 | 0.79): يشير إلى ممارسات توثيق مقبولة، لكن نسبة المراجع الحديثة (<35% في 75% من الأبحاث) تقل عن المعيار الدولي (>50%)؛ مما يضعف الارتباط بالمعرفة الراهنة.

- وضوح النتائج (3.50 | 0.66): يشير إلى قدرة جيدة على تقديم النتائج بطريقة منظمة ومفهومة، وهو ما ينسجم مع المعايير الدولية الأساسية، مع انحراف منخفض (0.66) يعكس الاتساق.

تشير هذه النتائج إلى وجـــود أساس منهجـــي متيـــن نسبياً، لكنه لا يصـــل إلى المستوى الدولي نتيجة قصور في الأدوات التحليلية المتقدمة والتوثيق الشامـــل؛ وقد يُعزى جزء كبير من التفاوت في جودة البحث العلمي إلى محدودية وصول الباحثين إلى قواعد البيانات.

التأثير الأكاديمي والاجتماعي والاقتصادي للبحث العلمي. تشمل هذه الفئة "الأثر الأكاديمي"، "الأثر الاجتماعي"، و"الأثر الاقتصادي"، وتسجل متوسطات منخفضة تراوح بين 1.07 و2.30، بتصنيف "ضعيف":

- الأثر الأكاديمـــي (1.09 | 0.30): يعكس هذا المعدل المتدنـــي ضعفاً شبـــه كامـــل في الاستشهادات والمساهمة في الحوار العلمي العالمي، مقارنة بالمعايير الدولية التي تتطلب متوسط استشهادات يراوح بين 10 و20 لكل بحث. الانحراف المنخفض (0.30) يؤكد أن هذا الضعف متجذر عبر المجلات.
- الأثر الاجتماعي (2.30 | 0.88): يشير إلى محدودية ترجمة الأبحاث إلى حلول مجتمعية، وهو ما يتعارض مع المعايير الدولية التي تؤكد التأثير العملي كمعيار أساسي للجودة.
- الأثر الاقتصادي (1.07 | 0.30): يعكس غياباً شبه تام للمساهمة في التنمية الاقتصادية أو الابتكار التطبيقي، بعيداً عن المعايير الدولية التي تربط البحث بالنمو الاقتصادي.

هذه الأرقام تكشف عن فجوة تأثيرية عميقة تعوق التوافق مع المعايير الدولية، على الرغم من وجود بنية منهجية مقبولة.

المعايير الأخلاقية والنزاهة العلمية. تشمل هذه الفئة "أصالة البحث"، و"النزاهة العلمية"، و"حقوق الملكية الفكرية"، وتسجل متوسطات مرتفعة تراوح بين 3.90 و4.20، بتصنيف "جيد" إلى "جيد جداً":

- أصالة البحث (3.90 | 0.73): يشير هذا المتوسط إلى قدرة نسبية على تقديم مساهمات معرفية جديدة، وهو قريب من "جيد جداً" (4.0)، لكنه دون المستوى الدولي (4.5+) الذي يتطلب ابتكاراً جذرياً.
- النزاهة العلمية (4.10 | 0.86): يدل هذا المتوسط على التزام عالٍ بمبادئ الأمانة العلمية، وهو يحقق تقدير "جيد جداً" ويتوافق بشكل كبير مع المعايير الدولية التي تتطلب الشفافية.
- حقوق الملكية الفكرية (4.20 | 0.79): يعكس هذا المتوسط مستوى متقدماً من الالتزام بحماية الملكية الفكرية، ويصنف "جيد جداً"، مقارباً للمعايير الدولية (4.5+).

تظهر هذه القيم أداءً متميزاً نسبياً في الجوانب الأخلاقية؛ مما يحقق توافقاً كبيراً مع المعايير الدولية في هذا المجال.

جدول 4

التقييم العام لجودة البحوث العلمية في المجلات اليمنية

م	المجلة	جودة البحث العلمي		الأثر		المعايير الأخلاقية		المجال الكلي		التقدير
		م	ع	م	ع	م	ع	م	ع	
1	المجلة العربية لضمان الجودة	3.54	0.70	1.62	0.74	4.27	0.80	3.14	1.20	مقبول
2	جامعة عدن للعلوم الطبيعية	3.56	0.71	1.57	0.73	4.17	0.79	3.10	1.18	مقبول
3	الآداب للدراسات النفسية والتربوية	3.56	0.71	1.53	0.71	4.17	0.79	3.09	1.18	مقبول
4	مجلة الدراسات الاجتماعية	3.46	0.73	1.50	0.70	4.07	0.76	3.01	1.15	مقبول
5	العلوم التربوية والدراسات	3.36	0.72	1.40	0.68	3.97	0.82	2.91	1.12	ضعيف
6	الآداب للدراسات اللغوية والأدبية	3.34	0.71	1.43	0.70	3.90	0.80	2.89	1.10	ضعيف
7	أبحاث	3.26	0.73	1.37	0.67	3.87	0.81	2.83	1.10	ضعيف
	الإجمالي	3.43	0.11	1.49	0.09	4.07	0.14	3.00	0.12	مقبول
	التقدير	جيد		ضعيف		جيد جداً		مقبول		

يُقدم جدول 4 ترتيباً للمجلات في عينة الدراسة بناءً على المتوسط الكلي للمجلة؛ إذ تراوح النتائج بين 2.83 لمجلة "أبحاث" و3.14 لـ "المجلـــة العربيـــة لضمان الجودة"، مع متوسط إجمالي يبلغ 3.00 وانحراف معياري 0.12. يُصنف هذا المتوسط البحث العلمي اليمني ضمن فئة مقبول وفق مقياس التقييم (0-5)؛ الانحراف المعياري المنخفض لا يعكس تجانساً في الأداء بين المجلات كما قد يظهر للوهلة الأولى، بل يكشف عن تعويض مضلل بين مجالات متفوقة وأخرى متدنية بشكل حاد؛ كما تؤكد ذلك قيم الانحراف المعياري الفردية لكل مجلة. يبرز هذا الواقع تحديات بنيوية عميقة تواجه البحث العلمي اليمني تتجاوز الفروقات السطحية بين المجلات؛ فعلى سبيل المثال، يعكس ضعف التمويل في مجلة "أبحاث" قصوراً في توفير البنية التحتية اللازمة للبحث،؛ مثل الوصول إلى قواعد بيانات عالمية أو أدوات تحليل حديثة، في حين تُظهر محدودية الانتشار الدولي في مجلة "العلوم التربوية والدراسات" عزلة أكاديمية تحول دون تحقيق تأثير عالمي، ويبرز نقص الربط بالاحتياجات المجتمعية في مجلة "الآداب اللغوية" فجوة بين الإنتاج البحثي والتطبيق العملي، وهي تحديات متكررة في السياق اليمني؛ نتيجة ضعف السياسات البحثية، وتدني الاستثمار في البحث التطبيقي، والاضطرابات السياسية والاقتصادية التي تعوق التواصل مع المجتمع العلمي الدولي، كما يتضح في التفصيل الآتي:

1 - جودة البحث العلمي (المتوسط العام: 3.43 | الانحراف: 0.11): تراوح المتوسطات بين 3.26 لـ "مجلة أبحاث" و3.56 لـ "جامعة عدن للعلوم الطبيعية والتطبيقية". الانحراف المعياري المنخفض يدل على أداء متقارب بين المجلات يعكس هذا التقارب وجود كفاءة أساسية في التنظيم البحثي عبر المجلات، لكنها لا ترقى إلى المعيار الدولي (4.0+)؛ على سبيل المثال مجلة "الدراسات الاجتماعية" (3.46) تحقق أداءً جيداً قد يعود إلى تركيزها على قضايا محلية ذات أولوية، في حين تعكس مجلة "أبحاث" (3.26) تأخراً يمكن ربطه بنقص الموارد أو ضعف التخطيط البحثي، يبرز هذا التباين اعتماد جودة البحث في اليمن بشكل حاسم على السياق المؤسسي والتخصصي، وغياب بنية تحتية متقدمة تلبي المعايير العالمية؛ فمجلة "جامعة عدن" تتفوق (3.56) بفضل التخصصات التطبيقية التي تتطلب منهجيات صارمة، فيما يكشف المتوسط المتدني (3.36) لـ "العلوم

التربوية والدراسات الإنسانية، الآداب للدراسات اللغوية والأدبية" عن تحديات محتملة في الموارد وضعف التخطيط الأكاديمي، هذا التباين يوحي بأن جودة البحث مرتبطة بعوامل خارجية؛ مثل الدعم المؤسسي أكثر من كونها خاصية متأصلة في البحث اليمني.

2 - الأثر (1.49 | 0.09): تراوح المتوسطات بين 1.37 لـ "مجلة أبحاث" و1.62 لـ "المجلة العربية لضمان الجودة"، وجميعها دون الحد الأدنى الدولي (2.5)؛ الانحراف المنخفض (0.09) يدل على أن ضعف الأثر مشكلة عامة، لكن الفروقات الصغيرة تعكس أن المجلات ذات التخصصات الواضحة؛ مثل "المجلة العربية لضمان الجودة" تحقق نتائج أفضل في الأثر الاقتصادي والاجتماعي من المجلات العامة؛ مثل "مجلة أبحاث". يلاحظ أن الأثر الاقتصادي والأثر الاجتماعي يمثلان نقطة ضعف بارزة؛ إذ تظهر مجلة "العلوم التربوية والدراسات الإنسانية" (1.40) قصوراً في تحويل الأبحاث إلى حلول اجتماعية أو اقتصادية، وهو ما قد يرجع إلى طبيعتها النظرية التي لا تلبي احتياجات السوق أو المجتمع المحلي، في المقابل تسجل مجلة "الآداب النفسية والتربوية" (1.53) أداءً أفضل نسبياً؛ ربما لأن مجالها النفسي يتيح تطبيقات عملية أكثر وضوحاً، أما مجلة "أبحاث" (1.37)؛ فتعكس الحد الأدنى للتأثير؛ مما يشير إلى أن ضعف التمويل والانتشار يقوض قدرتها على التأثير اقتصادياً أو اجتماعياً. هذه النتائج تكشف عن تحدّ بنيوي في اليمن؛ يتمثل في انفصال البحث عن الاحتياجات التنموية، نتيجة نقص الاستثمار في البحث التطبيقي وضعف الشراكات مع القطاعات الاقتصادية والمجتمعية.

3 - المعايير الأخلاقية (4.07 | 0.14): تراوح المتوسطات بين 3.87 لـ "مجلة أبحاث" و4.27 لـ "المجلة العربية لضمان الجودة". هذا المجال هو الأقرب إلى المعايير الدولية (4.5+)؛ "المجلة العربية لضمان الجودة" تقترب أكثر؛ ربما لأن طبيعة مجالها تتطلب التزاماً صارماً بالأخلاقيات لضمان المصداقية، تتبعها "مجلة جامعة عدن للعلوم الطبيعية والتطبيقية" (4.17)، وتحقق مجلة "الآداب اللغوية والأدبية" (3.90) أداءً جيداً؛ مما يعكس بيئة أكاديمية منظمة تدعم النزاهة؛ أما "مجلة أبحاث" (3.87)؛ فتبقى الأدنى؛ مما يشير إلى أن ضعف الموارد

قد يؤثر حتى على الجوانب الأخلاقية، التباين الطفيف (0.14) يدل على أن الالتزام الأخلاقي قاعدة عامة، ويظهر هذا المجال قوة نسبية في البحث اليمني، لكنه لا يكفي لتعويض الضعف في التأثير.

التحليل المفصل لجدول 4 يكشف أن المتوسط الإجمالي (3.00) يعكس أداءً "مقبولاً" ظاهرياً، لكنه يخفي تفاوتات حادة بين المجالات؛ إذ تتفوق بعض المجلات في جوانب محـــددة، في حين تتأخر أخرى بشكل يعكس تحديـــات بنيويـــة عميقـة في البحث اليمني تتجاوز مجرد فروق سطحية؛ فمثلاً:

- "المجلة العربية لضمان الجودة" (3.14): تتصدر الترتيب بفضل أثرها الاقتصادي والاجتماعي الأعلى (1.62)، وهو ما يرتبط بتركيزهـا على مجـــال عملـــي (ضمان الجودة التعليمية)، جـودة البحث (3.52) تدعم هذا الأداء بمنهجيات دقيقـــة، في حين تقترب المعايير الأخلاقية (4.27) من المستـــوى الدولي (4.5+)؛ مما يشير إلى التزام قوي بالنزاهة قد يعززه الدعم المؤسسي أو التدريب الأكاديمي، هذا التوازن يجعلها نموذجاً نادراً في السياق اليمني، لكنه لا يزال دون المعيار الدولي بسبب محدودية الانتشار العالمي.
- مجلة "جامعة عدن للعلوم الطبيعية والتطبيقية" (3.10): تحتل المرتبة الثانية بفضل جودة البحث العالية (3.56)، التي تعكس قدرة مؤسسية على تصميم دراسات دقيقة في العلوم التطبيقية، وهو مجال يتطلب أدوات تحليل متقدمة، المعايير الأخلاقية (4.17) تظهر التزاماً متميزاً، لكن الأثر (1.57) يكشف ضعفاً في ترجمة هذه الجودة إلى فائدة اقتصادية أو اجتماعية، هذا يعكس تحدياً بنيوياً في الجامعات اليمنية: قوة داخلية لا تتكامل مع التأثير الخارجي.
- مجلة "الآداب للدراسات النفسية والتربوية" (3.09): تُحقق أداءً جيداً بجودة بحث (3.56)؛ مما يشير إلى قدرة على معالجة قضايا نفسية وتربوية بدقة منهجية، ومعايير أخلاقية (4.17) تعكس حساسية المجال للنزاهة، لكن أثرها الاقتصادي والاجتماعي (1.53) يظل منخفضاً؛ مما قد يعود إلى ضعف الشراكات مع المؤسسات التعليمية أو المجتمعية لتطبيق النتائج، وهو تحدّ شائع في التخصصات السلوكية باليمن التي تفتقر إلى تمويل تطبيقي.

- مجلة "الدراسات الاجتماعية" حققت أداءً عاماً "مقبولاً" (3.01)، مدعوماً بنهج متقارب الأوزان بين جودة البحث العلمي (3.46) التي تعكس تركيزاً على قضايا اجتماعية ملحّة كالهجرة أو الفقر، والأخلاقيات التي تُظهر التزاماً قوياً بالنزاهة (4.07)، وضعف في الأثر (1.50) مع ذلك يشير إلى إمكانيات تطبيقية أفضل من التخصصات النظرية، وهذا الأداء يوحي بأن التخصصات الاجتماعية في اليمن قد تكون أقرب إلى التأثير المجتمعي من غيرها لكنها تعاني نقصاً في الدعم لتطوير هذا الجانب.
- مجلة "العلوم التربوية والدراسات الإنسانية" (2.91): تسجل أداءً منخفضاً في الأثر (1.40)؛ مما يعكس ضعف الربط بالواقع الاقتصادي والاجتماعي؛ ربما لأن أبحاثها تركز على النظريات التربوية دون تطبيقات عملية ملموسة، جودة البحث (3.36) تظهر كفاءة أساسية، لكنها أقل من المتوسط العام (3.43)، في حين تقترب المعايير الأخلاقية (3.97) من (4.0)؛ مما يشير إلى التزام مقبول ولكن غير متميز، هذا التراجع يعكس تحدياً في السياق اليمني، هذا الأداء يُظهر عزلة التخصصات التربوية عن احتياجات التعليم المحلي.
- مجلة "الآداب للدراسات اللغوية والأدبية" 2.89، وهو أداء دون المتوسط، جودة البحث العلمي (3.34) تتسم بكفاءة في صياغة الأسئلة وتوضيح النتائج، لكنها محدودة في الأصالة لتركيزها على تحليل النصوص الأدبية، وهي خاصية شائعة في الدراسات الأدبية، الأثر (1.43) يعاني غياب الأثر الأكاديمي والاقتصادي (1.0)؛ بسبب ندرة الاستشهادات، مع مساهمة اجتماعية محدودة (2.2) في الهوية الثقافية، الأخلاقيات (3.90) قوية بدعم من حقوق الملكية الفكرية (4.1) والنزاهة (4.0)؛ مما يعزز المصداقية، ويكشف الانحراف العالي (1.10) تفاوتاً بين الأخلاقيات القوية والأثر الضعيف؛ مما يستدعي تعزيز الابتكار والنشر العالمي.
- مجلة "أبحاث" (2.83): تتذيل الترتيب بسبب ضعفها في جميع المجالات: جودة البحث (3.26) تعكس قصوراً في التنظيم أو الموارد، والأثر (1.37) يكشف عن غياب شبه تام للتأثير الاقتصادي والاجتماعي، ربما بسبب طابعها العام الذي يفتقر إلى تركيز تخصصي، والمعايير الأخلاقية (3.87) هي الأدنى؛ مما يشير إلى أن نقص التمويل قد يؤثر حتى على الجوانب الأساسية، هذا الأداء يجعلها مثالاً للتحديات البنيوية في اليمن؛ مثل ضعف البنية التحتية البحثية والانفصال عن السياق التنموي.

الخلاصة

الاستنتاجات

1 - أظهر البحث العلمي في اليمن تحقيق الحد الأدنى من المعايير المقبولة (متوسط عام 3.00)، إلا أنه ما يـــزال يعانـــي ضعفاً في التأثير الأكاديمـــي والمجتمعـــي؛ ما يعكس غياب إستراتيجيات فعّالة لربط الإنتاج البحثي باحتياجات التنمية الوطنية والتوجهات البحثية العالمية، ويستدعي هذا الوضع إصلاحات هيكلية تعيد توجيه البحوث نحو الجودة والمواءمة مع منظومات الابتكار الدولية.

2 - يُظهر الباحثون اليمنيون التزاماً مرتفعاً بمعايير النزاهة العلمية وحقوق الملكية الفكرية (4.10–4.20)، متجاوزين عتبة Q1؛ بما يعزّز موثوقية أبحاثهم، غير أن هذا الالتزام لم يُترجم إلى تأثير علمي أو مجتمعي ملموس؛ ما يكشف عن فجوة بين جودة الممارسات وانخفاض الأثر العلمي والتنموي؛ الأمر الذي يستدعي تحويل النزاهة إلى محفز للإبداع والتأثير الفعّال.

3 - تشير نتائج تقييم المنهجية البحثية (3.20–3.70) إلى التزام الباحثين بأسس مقبولة، لكنها لا ترتقي إلى المعايير العالمية؛ مما يحدّ من دقة النتائج وشمولية المعرفة المنتجة، ويؤكـد ذلك الحاجة إلى تبنّي منهجيـــات متقدمـــة ترتكز على الابتكار والتوسع في استخدام أدوات تحليلية حديثة.

4 - يعكس ضعف الأثر الاجتماعي والاقتصادي فجوة حادة بين البحث العلمي واحتياجات المجتمع؛ إذ تظل مخرجات الأبحاث بعيدة عن التحول إلى تطبيقات عملية أو ابتكارات تدعم التنمية الاقتصادية، ويُظهر هذا الواقع محدودية القيمة التطبيقية للبحث وغياب إستراتيجيات فاعلة لربط المعرفة العلمية بسوق العمل والصناعة؛ ما يفرض إعادة التفكير في دور البحث العلمي في اليمن: أهو ترف أكاديمي؟ أم أداة تنموية تقتضي إعادة توجيه وظيفته؟

5 - يشير الانحراف المعياري المرتفع (1.14) إلى عدم تجانس في الأداء البحثي بين المجالات؛ إذ تتسم بعض الأبحاث بمستويات متقدمة من النزاهة الأخلاقية، في حين تعاني أخرى ضعفاً شديداً في التأثير الأكاديمي، هذا التناقض يعكس

خللاً بنيوياً، يتمثل في غياب تكامل البحث العلمي مع المجتمع الأكاديمي الدولي وعدم ترجمة الأبحاث إلى مخرجات ذات أثر ملموس، كما يشير إلى غياب معايير موحدة للجودة البحثية، وقصور في سياسات وزارة التعليم العالي والبحث العلمي في بناء منظومة بحثية متماسكة، لذا تبرز الحاجة الملحّة إلى إستراتيجيات وطنية واضحة تعزز جودة البحث العلمي وتضمن توافقه مع المعايير العالمية.

6 - انخفاض جودة التوثيق العلمي وحداثة المصادر يشيران إلى صعوبة الوصول إلى قواعد البيانات العلمية الحديثة؛ مما يشكل عائقاً رئيسياً أمام الابتكار؛ إذ يؤدي إلى ضعف إنتاج أفكار بحثية جديدة، ويقلل أيضاً من فرص اندماج البحث العلمي اليمني في المشهد الأكاديمي الدولي؛ وهو ما يستدعي حلولاً جذرية لتعزيز الوصول إلى قواعد البيانات.

7 - يعكس تباين التقييم بين معيار "أصالة البحث" (3.90) ومعيار "الابتكار" (3.20) توجه البحث العلمي اليمني نحو تحسين الدراسات السابقة أكثر من إحداث اختراقات معرفية جديدة، وعلى الرغم من التقدير الإيجابي للأصالة، فإن محدودية الابتكار تؤكد الحاجة الملحّة إلى تعزيز الفكر الإبداعي في الأبحاث العلمية، حتى تقترب من معايير الابتكار الدولية (4.5+) وتتمكن من إحداث تأثير علمي أكثر تميزاً وريادة.

التوصيات

بناءً على النتائج التي توصلت إليها الدراسة، وكشفت عن مواطن قوة محددة وفجوات بنيوية واضحة في منظومة البحث العلمي بالجامعات اليمنية نوصي بما يأتي:

- تسويق الجامعات ووزارة التعليم العالي هذا الالتزام الأخلاقي المرتفع كقاعدة لبناء شراكات بحثية دولية موثوق بها، كما نوصي بربط الترقيات الأكاديمية ليس بالالتزام الأخلاقي فقط بل بقدرة البحث على تحقيق أثر ملموس أيضاً؛ وذلك لتحويل هذه القوة الكامنة إلى ميزة تنافسية عالمية.

- توفير وزارة التعليم العالي والجامعات بالاستثمار اشتراكات مركزية في قواعد البيانات العالمية (مثل Scopus وWeb of Science)، هذا الإجراء سيمكّن الباحثين

من الاطلاع على أحدث التوجهات، وتحديد الفجوات البحثيـــة بدقـــة، والانتقـــال من تحسين الأفكار القائمة إلى ابتكار حلول معرفية جديدة.

- وضع إستراتيجية وطنية تعيد توجيه التمويل والحوافز البحثية نحو المشاريع التطبيقية التي تخدم أولويات التنمية في اليمن (مثل التعليم، والصحة العامـة، والأمن المائي، والطاقة المتجددة). كما يجب تشجيع المجلات العامة على التخصص التدريجي لزيادة جودتها وتأثيرها.
- تبني استمارة التقييم وتطوير نموذج الذكاء الاصطناعي (LLMs) كأداة معيارية للتحكيم في المجلات العلمية اليمنية؛ وهو ما سيضمن توحيد معايير الجودة، وتقييم الأبحاث بناءً على الأبعاد الثلاثة (الجودة، الأثر، الأخلاقيات)، وتوفير تغذية راجعة دقيقة للباحثين.
- إنشاء "وحدة لدعم النشر الدولي" في الجامعات الرئيسية لتقديم خدمات الترجمة والتحرير الأكاديمي، وتغطية تكاليف النشر في المجلات المصنفة؛ بالتوازي يجب إلزام المشاريع البحثية الممولة تضمين "خطة لقياس الأثر المجتمعي"، وتقديم تقارير توثق كيفية مساهمة البحث في حل مشكلة مجتمعية أو تطوير سياسة عامة.

المراجع

تقرير معامل التأثير والاستشهادات المرجعية العربي (أرسيف) لعام 2024. قاعدة بيانات معرفة. (2024). استرجع في 8 مارس، 2025، من https://n9.cl/kxnpy

الخطيب، برهان الدين. (2021). واقع البحث العلميّ ومتطلبات تطويره. *مجلة القرار للبحوث العلمية المحكمة*. https://n9.cl/2f9ks

الرازحي، عبدالوارث عبد سيف. (2021، مايو 25). البحث التربوي ودوره في تطوير العمليَّة التعليميَّة بجامعة الحديدة. *التنويري*. https://n9.cl/p3wjx8

السلطان، يوسف يعقوب. (2023). البحث العلمي والتطوير... الأنواع والمستلزمات والمعوقات. *مجلة العربي*،(693). https://alarabi.nccal.gov.kw/Home/Article/20120

مكرد، عائدة. (2010). *تطوير البحث العلمي بالجامعات اليمنية في ضوء الخبرات العالمية الحديثة* [ورقة مؤتمر]. مركز التطوير الأكاديمي وضمان الجودة، جامعة عدن. https://adc.aden-univ.net/5thConference_papers/research.htm

الهبوب، أحمد غالب، والفخري، نجلاء عبد الدائم. (2018). تصور مقترح لتجويد البحث العلمي في الجامعات اليمنية في ضوء متطلبات مجتمع المعرفة. *مجلة مركز التطوير الأكاديمي وضمان الجودة، 6*(11، عدد خاص)، 13–47. https://search.emarefa.net/ar/detail/BIM-1216718

Abrishami, A., & Aliakbary, S. (2019). Predicting citation counts based on deep neural network learning techniques. *Journal of Informetrics, 13*(2), 485–499. https://doi.org/10.1016/j.joi.2019.02.011

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *Sage Open, 9*(1). https://doi.org/10.1177/2158244019829575

Al-Munifi, A., & Aleryani, A. (2021). Internal efficiency of higher education system in armed conflict-affected countries: Yemen case. *International Journal of Educational Development, 86*. https://doi.org/10.1016/j.ijedudev.2021.102394

Beltagy, I., Lo, K., & Cohan, A. (2019). *SciBERT: A pretrained language model for scientific text.* In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 3615–3620). https://doi.org/10.18653/v1/D19-1371

Benziane, R. (2024). *List of journals indexed in Scopus adopting the Arabic language (Updated December 2023) [Method]*. ResearchGate. https://doi.org/10.13140/RG.2.2.11675.23849

Bibri, S. E., Krogstie, J., Kaboli, A., & Alahi, A. (2024). Smarter eco-cities and their leading-edge artificial intelligence of things solutions for environmental sustainability: A comprehensive systematic review. *Environmental Science and Ecotechnology, 19*. https://doi.org/10.1016/j.ese.2023.100330

Bornmann, L., & Marx, W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics, 98*(2), 487–509. https://doi.org/10.1007/s11192-013-1161-y

CGIAR Independent Advisory and Evaluation Service. (2023). *Applying quality of research for development frame of reference to process and performance evaluations*. https://n9.cl/r59gp

Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications, 8*(1), 25. https://doi.org/10.1057/s41599-020-00703-8

Cheng, Z., Chen, C., Li, C., & Tong, X. (2023). Evaluation of scientific research in universities based on the theories for sustainable competitive advantage. *SAGE Open, 13*(2). https://doi.org/10.1177/21582440231177048

Clarivate. (2024a). *Journal citation reports 2024: Simplifying journal evaluation*. https://n9.cl/ehbew

Clarivate. (2024b). *Journal citation reports: Simplifying journal evaluation*. https://n9.cl/gj8zm

Clarivate. (2024c). *Web of science: Trusted research intelligence*. https://clarivate.com/products/web-of-science/

de Bruin, J., Lombaers, P., Kaandorp, C., Teijema, J., van der Kuil, T., Yazan, B., Dong, A., & van de Schoot, R. (2025). ASReview LAB v.2: Open-source text screening with multiple agents and a crowd of experts. *Patterns, 6*(7). https://doi.org/10.1016/j.patter.2025.101318

DORA. (2012). *San Francisco declaration on research assessment*. American Society for Cell Biology. https://sfdora.org/read/

El-Ouahi, J. (2023). The Arabic citation index: Toward a better understanding of Arab scientific literature. *Quantitative Science Studies, 4*(3), 728–755. https://doi.org/10.1162/qss_a_00261

Elsevier. (2024). *Scopus: Comprehensive, multidisciplinary abstract and citation database*. https://www.elsevier.com/solutions/scopus

Hatch, A., & Curry, S. (2020). Research culture: Changing how we evaluate research is difficult, but not impossible. *eLife, 9*. https://doi.org/10.7554/eLife.58654

Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies, 1*(1), 387–395. https://doi.org/10.1162/qss_a_00020

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The leiden manifesto for research metrics. *Nature, 520*(7548), 429–431. https://doi.org/10.1038/520429a

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*. https://doi.org/10.1016/j.cmpbup.2024.100145

Kousha, K., & Thelwall, M. (2024). Artificial intelligence to support publishing and peer review: A summary and review. *Learned Publishing, 37*(1), 4–12. https://doi.org/10.1002/leap.1570

Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor and the politics of publication. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 369–387). Springer. https://doi.org/10.1007/978-3-030-02511-3_14

Leng, R. I. (2021). Diversity in citations to a single study: A citation context network analysis of how evidence from a prospective cohort study was cited. *Quantitative Science Studies, 2*(4), 1216–1245. https://doi.org/10.1162/qss_a_00154

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology, 74*(5), 570–581. https://doi.org/10.1002/asi.24750

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A.-M., Foeger, N., & Dirnagl, U. (2020). The Hong Kong principles for assessing researchers: Fostering research integrity. *PLoS Biology, 18*(7). https://doi.org/10.1371/journal.pbio.3000737

National Science Board. (2023). *Publications output: U.S. trends and international comparisons (NSB-2023-33)*. National Center for Science and Engineering Statistics. https://ncses.nsf.gov/pubs/nsb202333/ (Accessed: May 10, 2025)

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ, 6*. https://doi.org/10.7717/peerj.4375

Pontika, N., Klebel, T., Correia, A., Metz, I., Knoth, P., & Ross-Hellauer, T. (2022). Indicators of research quality, quantity, openness, and responsibility in institutional promotion, review, and tenure policies across seven countries. *Quantitative Science Studies, 3*(4), 888–910. https://doi.org/10.1162/qss_a_00224

Qiu, Y., Mintenig, S., Barchiesi, M., & Koelmans, A. A. (2025). Using artificial intelligence tools for data quality evaluation in the context of microplastic human health risk assessments. *Environment International, 185*. https://doi.org/10.1016/j.envint.2025.109341

QS. (2024). *QS world university rankings: Methodology*. QS Quacquarelli Symonds. https://www.topuniversities.com/qs-world-university-rankings/methodology

SCImago. (2024). *SCImago journal rank: Measuring journal prestige*. SCImago Lab. https://www.scimagojr.com

Stern, N. (2016). *Building on success and learning from experience: An independent review of the Research Excellence Framework*. UK Department for Business, Energy & Industrial Strategy. https://n9.cl/v0gjut

Tonmoy, S. M. T. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). A comprehensive survey of hallucination mitigation techniques in large language models. *arXiv*. https://arxiv.org/abs/2401.01313

UNESCO. (2021). *UNESCO science report: The race against time for smarter development*. https://unesdoc.unesco.org/ark:/48223/pf0000377433

Waltman, L., & van Eck, N. J. (2019). Field normalization of scientometric indicators. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 281–300). Springer. https://doi.org/10.1007/978-3-030-02511-3_11

Wang, J., Frietsch, R., Neuhäusler, P., & Hooi, R. (2024). International collaboration leading to high citations: Global impact or home country effect? *Journal of Informetrics, 18*(4). https://doi.org/10.1016/j.joi.2024.101565

Wilsdon, J., Bar-Ilan, J., Frodeman, R., Lex, E., Peters, I., & Wouters, P. (2017). *Next-generation metrics: Responsible metrics and evaluation for open science.* European Commission. https://ec.europa.eu/research/openscience/pdf/report.pdf

Zhang, Y., Wang, Y., Du, H., & Havlin, S. (2022). Delayed impact of interdisciplinary research. *arXiv*. https://doi.org/10.48550/arXiv.2207.04244

د. حمود علي عبده العبدلي، أستاذ تكنولوجيا التعليم المساعد، جامعة الحديدة، الجمهورية اليمنية، عميد كلية التربية السابق. دكتوراه في تكنولوجيا التعليم من كلية الدراسات العليا للتربية، جامعة القاهرة، 2012. الاهتمامات البحثية: التحـــول الرقمي في التعليم العالي لتعزيز كفاءته في دعم اقتصاد المعرفة والتنمية المستدامة؛ دمج الذكاء الاصطناعي وتحليلات التعلم في التعليم، التصميم التعليمي لتطوير بيئات تعليمية ذكية تتوافق مع فلسفة التعليم 5.0، تقييم مخرجات البحث العلمي باستخدام نماذج اللغة الكبيرة (LLMs)؛ بهدف تعزيز الأثر الاجتماعي والاقتصـــادي في العالم العربي.

haaa97@gmail.com

د. علي محمد عبدالله أخواجه، أستاذ تكنولوجيا التعليم المساعد، ورئيس قسم المناهج وطرائق التدريس، كلية التربية والعلوم بجامعة إقليم سبأ، الجمهورية اليمنية. دكتـــوراه في تكنولوجيا التعليم من كلية الدراسات العليا للتربية بجامعة القاهرة، 2017. الاهتمامـــات البحثية: التعليم الإلكتروني وتطبيقاته، استخدام الذكاء الاصطناعي في التعليم والبحث العلمي، الواقع الافتراضي والمعزز، الفصول الافتراضية، استخدام تطبيقـــات google التعليمية، أنظمـــة إدارة التعلم، المعايير التكنولوجية للمعلم والطالب والقيادات الإدارية، الاحتياجات التدريبية لأعضاء هيئة التدريس لاستخدام التقنيات الحديثة في التعليم والبحث العلمي.

alkhwagh2006@gmail.com

للاستشهاد:

العبدلي، حمود علي، وأخواجه، علي محمد. (2025). تقييم جودة البحوث في المجلات العلمية بالجامعات اليمنية باستخدام الذكاء الاصطناعي وفق المعايير الدولية. *مجلة دراسات الخليج والجزيرة العربية*، *51*(198)، 201–242. https://doi.org/10.34120/jgaps.v51i198.3475

To cite:

Al-abdli, H. A., & Akhwagh, A. M. (2025). AI-based evaluation of research quality in scientific journals of Yemeni universities according to international standards. *Journal of the Gulf and Arabian Peninsula Studies*, *51*(198), 201–242. https://doi.org/10.34120/jgaps.v51i198.3475