

متابعة إشراف بحث الدكتوراة
تخصص تقنية المعلومات
بحث بعنوان:

تطوير آلية للتعامل مع الذكاء الاصطناعي في البيانات غير
الخاضعة للرقابة نتيجة التحيز من تقييم المخاطر إلى التخفيف والتعافي

Prepared by

اعداد الباحث : علي الشهري
اشراف الاستاذ الدكتور : محمد الخولي

2026



جامعة ميدأوشن
كلية تقنية المعلومات
مشروع خطة أطروحة للتسجيل لدرجة الدكتوراة بعنوان:

تطوير آلية للتعامل مع الذكاء الاصطناعي في البيئات غير الخاضعة للرقابة نتيجة التحيز من تقييم المخاطر إلى التخفيف والتعافي

Developing a framework for Artificial Intelligence in uncontrolled environments due to bias, From Risk Assessment to Mitigation and Recovery

01

شكر وتقدير

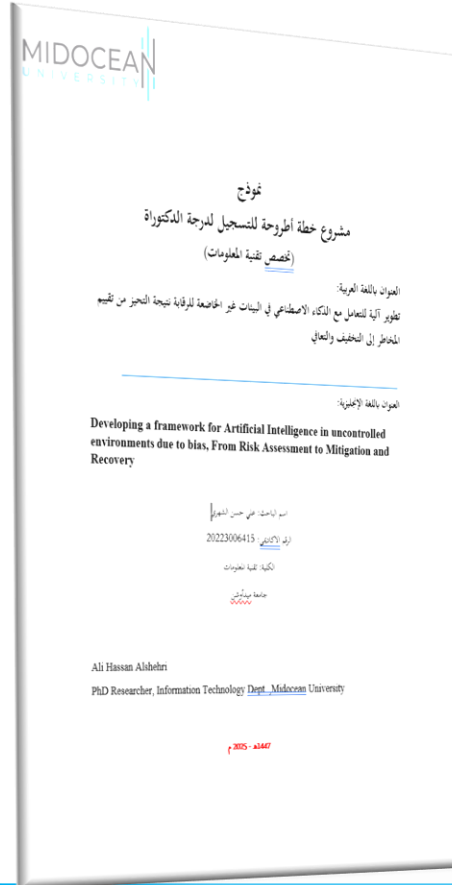


كل الشكر والتقدير للأستاذ الدكتور محمد الخولي على جهوده
ودعمه وتوجيهاته التي تدعم تطوير الباحثين لتقديم مساهمة
بحثية تثري وتضيف لجهود البحث العلمي في جامعة ميدأوشن
وكذلك لفريق الدعم بالجامعة

محتويات العرض:

ما تم عمله خلال هذا الفصل الدراسي:

١. العمل على موضوع البحث وتنفيذ التعديلات المطلوبة
٢. قراءة العديد من المراجع والدراسات والأبحاث العلمية
٣. الاستمرار في الاطلاع على التجارب الجديدة في مجال الذكاء الاصطناعي وتحيز هذه الأنظمة
٤. تنفيذ بعض التجارب العملية لمحاولة اعداد بيئة عملية للبحث العلمي
٥. بعض التحديات والمعوقات التي واجهها الباحث أثناء الفصل الدراسي



الاستمرار في تطوير خطة البحث العملي وموضوعه:

في بداية الفصل الدراسي وبعد حضور عدد من اللقاءات مع التوجيه الأكاديمي تم استكمال عدد من الفصول في البحث بما يحقق معايير الأصالة ويضيف لجهود البحث العلمي ويسهم في خدمة المعرفة وخصوصاً في مجال يعد اليوم واحد من أكثر المجالات أهمية وهو الذكاء الاصطناعي تم العمل على الموضوع البحثي وهو:

العنوان باللغة العربية:

تطوير آلية للتعامل مع الذكاء الاصطناعي في البيئات غير الخاضعة للرقابة نتيجة التحيز من تقييم المخاطر إلى التخفيف والتعافي

ب. العنوان باللغة الإنجليزية:

Developing a framework for Artificial Intelligence in
uncontrolled environments due to bias, From Risk
Assessment to Mitigation and Recovery

منهجية الدراسة:

باستخدام منهجية البحث المقارن، يجمع البحث بين التحليل النظري والدراسة المقارنة بين التعامل مع الذكاء الاصطناعي الضيق والعام في حالات الخروج عن السيطرة:

المنهج النظري

- تقييم الأطر النظرية الحالية حول مخاطر الذكاء الاصطناعي: رؤى من نظرية التحكم وإدارة المخاطر
- مراجعة الأدبيات حول آليات وتحديات التحكم في الذكاء الاصطناعي
- المنظور الأخلاقي: الأبعاد الأخلاقية والقانونية لإيقاف أنظمة الذكاء الاصطناعي عند تعطلها

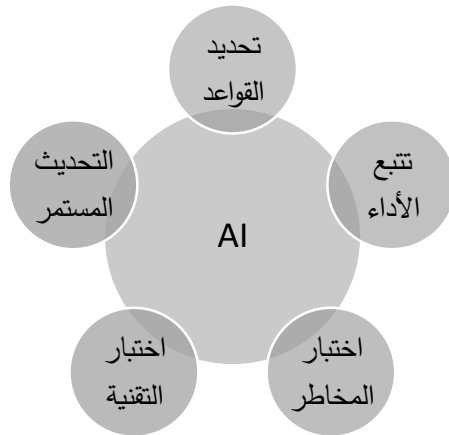
منهج دراسة الحالة:

- تكوين بيئة افتراضية: كيفية عمل أنظمة الذكاء الاصطناعي لتقييم سلوكها.
- اختبار سيناريوهات الخروج عن السيطرة: إنشاء سيناريوهات يتم فيها تجاوز الحدود التشغيلية، والتحليل المطلوب لسلوك الذكاء الاصطناعي عند استخدام آليات التوقف.
- تقييم فعالية آلية التحكم: تقييم نجاح آلية التحكم في تعطيل الأنظمة الخارجة عن السيطرة نتيجة التحيز مثل الاستجابة، وقابلية الاستغلال، وتأثير التوقف على استمرارية النظام.

فرضية البحث:

ولتحقيق أهداف البحث تم الاعتماد على الفرضيات التالية:

- يمكن للتحيز أن يؤثر على سلامة وأمن الذكاء الاصطناعي وهو ما يؤثر على مستوى الثقة بهذه الأنظمة.
- يمكن للذكاء الاصطناعي أن يتخذ قرارات متحيزة بناءً على مدخلات تؤثر على قراراته ومخرجاته ويزداد مستوى التحيز بزيادة التعليم الخاطئ للنظام
- هناك دور لكشف التحيز وتخفيفه والتعافي منه في تحسين نتائج أنشطة الذكاء الاصطناعي وكذلك تحسين معايير السلامة.



كيف سيتم تنفيذ الجانب العملي من الدراسة:

سيتم تطبيق النموذج العملي التالي مع التركيز على التحيز:

قراءة العديد من المراجع والدراسات والأبحاث العلمية:

Google Scholar

artificial intelligence safety

About 20,100 results (0.11 sec)

Articles

Any time
Since 2025
Since 2024
Since 2021
Custom range...

Sort by relevance
Sort by date

Leveraging **artificial intelligence** and advanced food processing techniques for enhanced food **safety**, quality, and security: a comprehensive review
[SB Dhal](#), D Kar - Discover Applied Sciences, 2025 - Springer
... **Artificial intelligence** is emerging as a transformative force in addressing the multifaceted challenges of food **safety**... driven technologies, such as **machine learning**, deep learning, natural ...
☆ Save 📄 Cite Cited by 16 Related articles All 3 versions

Artificial Intelligence for Patient **Safety** and Surgical Education in Neurosurgery
[T Sugiyama](#), [H Sugimori](#), [M Tang](#), [M Fujimura](#) - JMA Journal, 2025 - jstage.jst.go.jp
... Overview of our research project on **artificial intelligence** (AI) for neurosurgery aimed at ...
... (circulation model for AI enhancement...
... versions

for **artificial intelligence**
2025 - academic.oup.com
... **artificial intelligence** (AI) ... research into
... **Artificial intelligence safety** and security ...
... 8 versions

...: Fatigue monitoring and risk
... and practical ..., 2025 - books.google.com
... and **machine learning** technology to crew
... guarantee for flight **safety**, but also ...
... versions

Artificial intelligence in pharmaceutical
manufacturing: Enhancing efficiency, product quality, and **safety**
MV Rajesh, K Elumalai - Journal of Holistic Integrative Pharmacy, 2025 - Elsevier

RTIC
Journal of Technologies
Information and Communication
2025, 5(1), 33-52
e-ISSN: 2584-7965
Research Article
OPEN ACCESS

Enhancing Proactive Cyber Defense: A Theoretical Framework for AI-Driven Predictive Cyber Threat Intelligence

Kamrul Hossen¹, Farhad Hossain², Al Amin³, Yashab Saitradhar⁴, Israt Jahan Jenny¹, Shukik Mahmud^{1*}

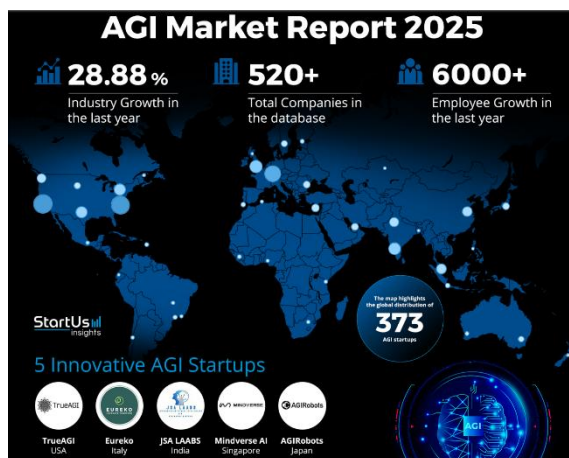
¹ Tishu University, Indiana, United States
² St. Francis College, Brooklyn, New York, United States
³ Mahabub International University, Fairfield, Iowa, United States
⁴ University of Bremen, Bremen, Germany
⁵ Japan-Rangladesh Robotics and Advanced Technology Research Center, Niphamat, Bangladesh
* Corresponding Author: shukikmahmud@gmail.com

Chadane Hossen, K. Hossain, F. Amin, A. Saitradhar, Y. Jenny, I. J. and Mahmud, S. (2025) Enhancing Proactive Cyber Defense: A Theoretical Framework for AI-Driven Predictive Cyber Threat Intelligence, *Journal of Technologies Information and Communication*, 5(1), 33-52. <https://doi.org/10.55227/jst.v5i1.24175>

ARTICLE INFO
Received: 25 Dec 2024
Accepted: 16 Mar 2025

ABSTRACT
The rapid evolution of cyber threats and the dynamic nature of the threat landscape have necessitated the development of proactive and predictive defense mechanisms. This research proposes an AI-driven framework for predictive cyber threat intelligence aimed at enhancing organizational cybersecurity by identifying and mitigating threats before they materialize. The framework integrates diverse data sources, including network logs, endpoint data, and threat intelligence feeds, to generate actionable insights using advanced machine learning algorithms such as anomaly detection, pattern recognition, and predictive analytics. A continuous feedback loop ensures the adaptability of the framework through

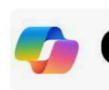
الاطلاع على التجارب الجديدة في مجال تحيز الذكاء الاصطناعي:



ChatGPT



Gemini



Microsoft Copilot



Grok



Claude



Comprehensive Bias Report

Detailed analysis of textual and numerical biases within automation contexts, offering insights and actionable recommendations.

#AutomationBias #AIethics

#BiasMitigation #Human-CentricDesign

HTML Bias Report

Share output Download File

بعض التحديات والمعوقات التي واجهها الباحث أثناء العام الدراسي:

١. صعوبة تطبيق التحيز على نماذج في الجانب الصحي والتقنيات ذاتية التشغيل والنقل بسبب عدم توفر نماذج مناسبة وما زال العمل جاري لتوفير نماذج يمكن العمل عليها للبحث.
٢. قلة المراجع فيما يتعلق بتحيز الذكاء الاصطناعي كونه ما زال حديثاً نوعاً مع الحرص على أن تكون المراجع حديثة خلال هذا العام 2026 وما زالت الشركات التي تقدم المنتجات تقدمها على دفعات وما زال الباحثين في المراحل الأولى للتعامل مع هذه التقنيات وكذلك اعداد البحث يركز على التحيز في أنظمة الذكاء الاصطناعي ورصد السلامة والأمن في هذه التقنية وإمكانية الخروج عن السيطرة من تقييم المخاطر إلى التخفيف والتعافي
٣. التسارع في ظهور تقنيات ونسخ جديدة من أنظمة الذكاء الاصطناعي يجعل البحث يسعى للتركيز على أفضل الخيارات التقنية التي يمكن رصدها لتحيز الذكاء الاصطناعي ومقارنتها بشكل علمي ما بين محدودة الكفاءة و نوعية الخدمات المقدمة.
٤. تنوع بعض آليات التعامل مع التحيز في مجال اخلاقيات الذكاء الاصطناعي وكذلك الجانب القانوني.
٥. محدودية المراجع العربية المختصة التي تنقد بشكل علمي الذكاء الاصطناعي من ناحية التحيز ولعل هذا البحث يكون إن شاء الله إضافة للمحتوى العربي في هذا المجال.



Thank You شكراً لكم