Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Peiyang Yu

Carnegie Mellon Univeristy

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Criterion 5 (original contributions of major significance). It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

399	574	12	10
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

96.2% independent of 399 classified citing papers

Citation type	Count
Independent	384
Self-citation	8
Co-author	0
Same-institution	7

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher developed optimized Transformer architectures for predictive modeling and misinformation detection, establishing a framework for enhancing model performance through advanced optimization techniques.

The researcher's core contribution centers on the 2024 paper 'Optimization of Transformer heart disease prediction model based on particle swarm optimization algorithm,' which appears to introduce a novel approach to refining Transformer models for medical prediction tasks. This foundational work suggests a focus on improving the accuracy and efficiency of deep learning applications in healthcare through specific optimization algorithms.

Building on this foundation, the researcher's subsequent 2025 works, including 'Unmasking Digital Falsehoods' and 'A Hybrid Transformer Model for Fake News Detection,' indicate a strategic expansion of these optimization techniques into the domain of information integrity. The titles suggest that the researcher is applying similar architectural enhancements, such as Bayesian optimization and hybrid recurrent units, to address the distinct challenge of detecting misinformation using Large Language Models.

The significance of this line of work is evidenced by substantial citation activity, with the core paper accumulating 130 citations and the follow-up studies garnering 65 and 58 citations respectively. Notably, 96.5% of the 424 classified citations originate from independent researchers, indicating that the broader academic community recognizes and utilizes these methods beyond the researcher's immediate circle, validating the independent impact of these contributions.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 263 · **6 flagged influential by Semantic Scholar**

CORE PAPER

Optimization of Transformer heart disease prediction model based on particle swarm optimization algorithm

2024 · 2024 6th International Conference on Frontier Technologies of Information ..., 2024 · 130 citations (GS)

Field-normalised: 24 Semantic Scholar citations place it in the top 5% of Medicine papers from 2024 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	DeepContainer: A Deep Learning-based Framework for Real-time Anomaly Detection in Cloud-Native Container Environments (2025)	Hunan University of Technology, New York University, University of Southern California	China, United States	—
2	Unmasking Digital Falsehoods: A Comparative Analysis of LLM-Based Misinformation Detection Strategies (2025)	University of California, Berkeley	United States	—
3	A Machine Learning Approach to Supply Chain Vulnerability Early Warning System: Evidence from U.S. Semiconductor Industry (2023)	California State University Long Beach, Columbia University	United States	—
4	Enhanced Spatio-Temporal Attention Mechanism for Video Anomaly Event Detection (2025)	—	—	—
5	The ROI of Data Security: How Hospitals and Health Systems Can Turn Compliance into Competitive Advantage (2024)	University of Oklahoma	United States	—
6	Optimization of Nursing Staff Allocation in Elderly Care Institutions: A Time Series Data Analysis Approach (2024)	California Institute of Advanced Management, University of Sydney	Australia, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
7	Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach (2025)	Columbia University, University of Chicago, University of Rochester	United States	—
8	Using Deep Reinforcement Learning for Optimizing Process Parameters in CHO Cell Cultures for Monoclonal Antibody Production (2024)	Brown University, University of California, Los Angeles, University of Pennsylvania	United States	—
9	Deep Learning-Based Saliency Assessment Model for Product Placement in Video Advertisements (2024)	Rice University, Santa Clara University, University of Southern California	United States	—
10	AI-Driven Optimization of Intergenerational Community Services: An Empirical Analysis of Elderly Care Communities in Los Angeles (2024)	California Institute of Advanced Management	United States	—
11	Future-Proofing Healthcare: The Role of AI and Blockchain in Data Security (2025)	IdMap.ai, University of Oklahoma	United States	—
12	Predictive Auto Scaling and Cost Optimization Using Machine Learning in AWS Cloud Environments (2025)	Amazon.com Services LLC, The Ohio State University - Columbus, University of California, Berkeley	China, United States	—
13	Optimized Cardiovascular Disease Prediction Using Clustered Butterfly Algorithm (2025)	Amrita Vishwa Vidyapeetham, International Institute of Information Technology, Sungkyunkwan University	India, South Korea, United Arab Emirates	—
14	Uncertainty-aware feature-weighted ensemble framework for heart disease prediction (2026)	Seventh Medical Center of the Chinese People's Liberation Army General Hospital, Taiyuan Normal University	China	—
15	Optimizing Cloud Computing with AI: Improving Resource Allocation and Reducing Costs (2025)	University of Michigan	United States	—
16	Residual GRU+MHSA: A Lightweight Hybrid Recurrent Attention Model for Cardiovascular Disease Detection (2025)	Edinburgh Napier University, George Mason University, NJIT	Australia, Portugal, United Kingdom	—
17	Particle Swarm Optimization Neural Network and Its Application in Mental Health Risk Assessment (2025)	Fujian Normal University	China	—
18	Machine Learning Algorithm for Noise Reduction and Disease-Causing Gene Feature Extraction in Gene Sequencing Data (2025)	Fengtai District Education Committee, Georgia Institute of Technology, University of Glasgow	China, United Kingdom, United States	—
19	PHYS-ODE-CAPS: a physiological simulative multi-dimensional framework for prognosis of retinal microvascular damage in diabetic retinopathy (DR) (2025)	Harcourt Butler Technical University	—	—
20	Cloud-based Machine Learning for Predicting Food Spoilage and Ensuring Safety in the Supply Chain (2025)	New York University	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
21	A Digital Twin-Inspired Hybrid Variational Quantum Reinforcement Learning Framework for Heart Disease Risk Prediction (2026)	Arden University, Lahore Garrison University, National University of Sciences and Technology	Pakistan, United Kingdom	—
22	Integrating PSO, GA, and ACO for Optimized ECG Feature Selection and Classification of Cardiac Disorders (2025)	—	—	—
23	Nature Inspired Optimization with Hybrid Machine Learning Model for Cardiovascular Disease Detection and Classification (2022)	Annamalai University, Sri Shanmugha College of Engineering and Technology	—	—
24	Goose-Inspired Optimization Algorithm (2025)	Rajabhat Maha Sarakham University	Thailand	—
25	Arapaima Gigantea Optimizer (2026)	Rajabhat Maha Sarakham University	Thailand	—
26	Greylag Goose Optimization Algorithm (2026)	Rajabhat Maha Sarakham University	Thailand	—
27	Snow Goose Optimization Algorithm (2026)	Rajabhat Maha Sarakham University	Thailand	—
28	Peregrine Falcon Optimization Algorithm (2026)	Rajabhat Maha Sarakham University	Thailand	—
29	Fennec Fox Optimization Algorithm (2026)	Rajabhat Maha Sarakham University	Thailand	—
30	Patagonian Fox Optimization Algorithm (2026)	Rajabhat Maha Sarakham University	Thailand	—

Showing the 30 most-cited of 131 independent citing papers.

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

FOLLOW-UP WORK

Unmasking Digital Falsehoods: A Comparative Analysis of LLM-Based Misinformation Detection Strategies

2025 · arXiv · 65 citations (GS)

Field-normalised: 39 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	The ROI of Data Security: How Hospitals and Health Systems Can Turn Compliance into Competitive Advantage (2024)	University of Oklahoma	United States	—
2	Future-Proofing Healthcare: The Role of AI and Blockchain in Data Security (2025)	IdMap.ai, University of Oklahoma	United States	—
3	Optimizing Cloud Computing with AI: Improving Resource Allocation and Reducing Costs (2025)	University of Michigan	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
4	Cloud-based Machine Learning for Predicting Food Spoilage and Ensuring Safety in the Supply Chain (2025)	New York University	United States	—
5	A Review of <scp>TRiSM</scp> Frameworks in Artificial Intelligence Systems: Fundamentals, Taxonomy, Use Cases, Key Challenges and Future Directions (2026)	Sikkim University	India	—
6	Real-time Early Warning of Trading Behavior Anomalies in Financial Markets: An AI-driven Approach (2025)	California State University Long Beach, University of Chicago	United States	—
7	Human-AI Co-Creation: A Framework for Collaborative Design in Intelligent Systems (2025)	Brown University	United States	—
8	Debate-to-Detect: Reformulating Misinformation Detection as a Real-World Debate with Large Language Models (2025)	Chinese Academy of Sciences	China	—
9	LLM-based Agents Suffer from Hallucinations: A Survey of Taxonomy, Methods, and Directions (2025)	Chinese Academy of Sciences, Griffith University, Macquarie University	Australia, China	—
10	Research on Personalized Medical Intervention Strategy Generation System based on Group Relative Policy Optimization and Time-Series Data Fusion (2025)	Icahn School of Medicine at Mount Sinai, University of Chicago, Weill Cornell Medicine	United States	—
11	Deep Learning for Multimodal Medical Image Analysis (2022)	—	—	—
12	Goal-Aware Identification and Rectification of Misinformation in Multi-Agent Systems (2025)	Beijing University of Posts and Telecommunications, Nanyang Technological University, National University of Singapore	China, Singapore	—
13	Machine learning for anomaly detection: A survey (2015)	—	—	—
14	The Compositional Architecture of Regret in Large Language Models (2025)	Provable Responsible AI and Data Analytics, The Hong Kong Polytechnic University, University of Exeter	Hong Kong, United Kingdom	—
15	Uncovering the Fragility of Trustworthy LLMs through Chinese Textual Ambiguity (2025)	—	—	—
16	Crisis, country, and party lines: politicians' misinformation behavior and public engagement (2026)	Ca' Foscari University of Venice, Ilmenau University of Technology, University of Padua	Germany, Italy	—
17	Physics-Regularized Self-Supervised Anomaly Detection for Semiconductor Tools with Digital Twin Guidance (2025)	—	—	—
18	Mixture of Adaptive Retrieval Experts for Veracity Assessment in the Human–LLM Mixed Generation Paradigm (2026)	Columbia University, University of Illinois Urbana-Champaign	—	—

No.	Citing paper	Citing institution(s)	Country	S2
19	Nonlinear dynamics of information overload: Impact on source localization in complex networks (2026)	Warsaw University of Technology	Poland	—
20	An Audit and Analysis of LLM-Assisted Health Misinformation Jailbreaks Against LLMs (2025)	Simon Fraser University	Canada	—
21	Accuracy of Autism-Related TikTok Information in Italian: A Comparison Between Human Raters and Large Language Models. (2026)	Drexel University, University College London, University of Miami	Italy, United Kingdom, United States	—
22	AI Ethics in Social Media (2025)	NSBM Green University	Sri Lanka	—
23	Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration (2024)	—	—	—
24	Daily Asset Pricing Based on Deep Learning: Integrating No-Arbitrage Constraints and Market Dynamics (2025)	—	—	—
25	Semantic Network Analysis of Financial Regulatory Documents: Extracting Early Risk Warning Signals (2025)	Baruch College, University of Michigan, University of Southern California	United States	—
26	Artificial Intelligence in Economic Applications: Stock Trading, Market Analysis, and Risk Management (2025)	Kyungil University	South Korea	—
27	Robust Bilevel Network-Flow Scheduling for Semiconductor Wafer Logistics under WLTP Uncertainty (2025)	University of California-Berkeley	United States	—
28	Using Large Language Models to Detect and Debunk Climate Change Misinformation (2026)	Kristianstad University, University of Copenhagen	Denmark, Sweden	—
29	Mixture of Adaptive Retrieval Experts for Veracity Assessment in the Human–LLM Mixed Generation Paradigm (2026)	Columbia University, University of Illinois Urbana-Champaign	—	—
30	Vulnerability Detection in Large Language Models: Addressing Security Concerns (2025)	InnoV'COM Laboratory	—	—

Showing the 30 most-cited of 79 independent citing papers.

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

FOLLOW-UP WORK

A Hybrid Transformer Model for Fake News Detection: Leveraging Bayesian Optimization and Bidirectional Recurrent Unit

2025 · arXiv preprint · 58 citations (GS)

Field-normalised: 27 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Unmasking Digital Falsehoods: A Comparative Analysis of LLM-Based Misinformation Detection Strategies (2025)	University of California, Berkeley	United States	—
2	The ROI of Data Security: How Hospitals and Health Systems Can Turn Compliance into Competitive Advantage (2024)	University of Oklahoma	United States	—
3	Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach (2025)	Columbia University, University of Chicago, University of Rochester	United States	—
4	Future-Proofing Healthcare: The Role of AI and Blockchain in Data Security (2025)	IdMap.ai, University of Oklahoma	United States	—
5	Optimizing Cloud Computing with AI: Improving Resource Allocation and Reducing Costs (2025)	University of Michigan	United States	—
6	Cloud-based Machine Learning for Predicting Food Spoilage and Ensuring Safety in the Supply Chain (2025)	New York University	United States	—
7	Real-time Early Warning of Trading Behavior Anomalies in Financial Markets: An AI-driven Approach (2025)	California State University Long Beach, University of Chicago	United States	—
8	Human-AI Co-Creation: A Framework for Collaborative Design in Intelligent Systems (2025)	Brown University	United States	—
9	Deep Learning for Multimodal Medical Image Analysis (2022)	—	—	—
10	Machine learning for anomaly detection: A survey (2015)	—	—	—
11	Uncovering the Fragility of Trustworthy LLMs through Chinese Textual Ambiguity (2025)	—	—	—
12	Physics-Regularized Self-Supervised Anomaly Detection for Semiconductor Tools with Digital Twin Guidance (2025)	—	—	—
13	Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration (2024)	—	—	—
14	Semantic Network Analysis of Financial Regulatory Documents: Extracting Early Risk Warning Signals (2025)	Baruch College, University of Michigan, University of Southern California	United States	—
15	Artificial Intelligence in Economic Applications: Stock Trading, Market Analysis, and Risk Management (2025)	Kyungil University	South Korea	—
16	Robust Bilevel Network-Flow Scheduling for Semiconductor Wafer Logistics under WLTP Uncertainty (2025)	University of California-Berkeley	United States	—
17	Augmented Reality and AI in Remote Geospatial Analysis: Enhancing Decision-Making in Crisis Management (2025)	—	—	—

No.	Citing paper	Citing institution(s)	Country	S2
18	BEYOND BITCOIN THE EVOLUTION OF BLOCKCHAIN FOR SECURE DECENTRALIZED APPLICATIONS (2025)	—	—	—
19	Research on Energy Efficiency Optimization Algorithms for Hotel IoT Systems Based on AI Load Forecasting (2025)	Agoda	China	—
20	Artificial Intelligence-Augmented Social Interfaces: Towards Empathetic and Context-Aware Interaction Systems (2026)	Brown University	United States	—
21	Task Specialization via Generative Behavior Clustering and Reinforced Distillation: Building Lightweight Experts from LLMs (2025)	Brown University	United States	—
22	Real-time Detection of Anomalous Trading Patterns in Financial Markets Using Generative Adversarial Networks (2025)	Baruch College, California State University Long Beach, University of California	United States	—
23	Cultural Bias Mitigation in Vision-Language Models for Digital Heritage Documentation: A Comparative Analysis of Debiasing Techniques (2024)	Duke University, Northeastern University, University of California, Berkeley	United States	—
24	Neuromorphic Computing for Spiking Neural Network Applications (2025)	University of North Carolina, University of Sunshine Coast	Australia, United States	—
25	AI-Powered Language Translation for Low-Resource Languages (2025)	University of North Carolina, University of Sunshine Coast	Australia, United States	—
26	Reinforcement Learning for Prompt Optimization in Language Models: A Comprehensive Survey of Methods, Representations, and Evaluation Challenges (2025)	Brown University	United States	—
27	A Low-Complexity Joint Angle Estimation Algorithm for Weather Radar Echo Signals Based on Modified ESPRIT (2025)	Hubei University, Nanjing University of Aeronautics and Astronautics, Sun Yat-sen University	China	—
28	Feature Selection Optimization for Employee Retention Prediction: A Machine Learning Approach for Human Resource Management (2025)	—	—	—
29	Deep Learning-Based Analysis of Social Media Sentiment Impact on Cryptocurrency Market Microstructure (2025)	University of Chicago, University of Michigan, University of Southern California	United States	—
30	Self-Attention Enhanced Dual BiGRU for Arabic Fake News Detection (2025)	University of Kufa, University of Tabriz	Iran, Iraq	—

Showing the 30 most-cited of 53 independent citing papers.

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Contribution 2

Claim — Contribution 2

The researcher advanced NLP-based text classification and extended these methods to LLM-driven applications in fake news detection and recommendation systems.

The researcher established a foundation in natural language processing for text classification, as evidenced by the 2021 core paper published in the Journal of Physics: Conference Series. This initial work serves as the conceptual anchor for a subsequent line of inquiry that bridges traditional classification techniques with emerging large language model capabilities.

This trajectory suggests an original approach to adapting foundational NLP methods for complex, modern challenges. By progressing from general text classification to specialized applications like fake news detection and user intent enhancement, the researcher appears to address the gap between static classification models and dynamic, context-aware LLM systems. The chronological evolution from 2021 to 2025 indicates a sustained effort to innovate within this specific technical niche.

The significance of this work is reflected in its citation metrics and independent uptake. The core paper has accumulated 29 citations, while the 2025 follow-up papers on fake news detection and recommendation systems have garnered 62 and 49 citations, respectively. Notably, 96.5% of the 424 classified citations originate from independent researchers, suggesting that this line of work has achieved broad recognition and utility beyond the researcher's immediate academic circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 124

CORE PAPER

Text Classification by using Natural Language Processing

2021 · Journal of Physics: Conference Series · 29 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	The Application of Generative AI in Virtual Reality and Augmented Reality (2024)	—	—	—
2	A Compressive Memory-based Retrieval Approach for Event Argument Extraction (2025)	—	—	—
3	Evolutionary Game Analysis of Enterprise IT Project Management from the Perspective of Technical Debt (2025)	BEIJING ZHIBAO YUNKE TECHNOLOGY CO., LTD.	China	—
4	Multi-Agent Reinforcement Learning for High-Frequency Trading Strategy Optimization (2024)	Washington University	United States	—
5	Deep Learning-Based Prediction of Critical Parameters in CHO Cell Culture Process and Its Application in Monoclonal Antibody Production (2024)	Brown University, New York University, Northeastern University	United States	—
6	Advances in Deep Reinforcement Learning for Computer Vision Applications (2025)	—	—	—
7	Intelligent Anti-Money Laundering Transaction Pattern Recognition System Based on Graph Neural Networks (2024)	Rutgers Business School, University of California, Berkeley, Washington University in St. Louis	United States	—
8	Automated Test Case Generation for Chip Verification Using Deep Reinforcement Learning (2025)	—	—	—

No.	Citing paper	Citing institution(s)	Country	S2
9	Overview of Multimodal Generative Models in Natural Language Processing and Computer Vision (2024)	Shandong Youth University of Political Science	China	—
10	Creation of Polish Online News Corpus for Political Polarization Studies (2022)	Hokkaido University, Software Development	Japan	Methodology
11	Research and practice of advertisement recommendation algorithm based on graph neural network (2025)	LinkedIn, Stanford University, Wuyi University	China, United States	—
12	Interface Design Optimization Research Based on User Behavior Data (2025)	Taizhou University, University of Glasgow, Wuhan Donghu College	China, United Kingdom	—
13	Academic Registration Text Classification Using Machine Learning (2022)	University of Ha'il	KSA	Background
14	A Review of Multimodal Interaction Technologies in Virtual Meetings (2024)	—	—	—
15	Machine Vision-Based Automatic Detection for Electromechanical Equipment (2024)	Georgia Institute of Technology	United States	—
16	A Graph Neural Network-Based Approach for Detecting Fraudulent Small-Value High-Frequency Accounting Transactions (2024)	Baruch College, Illinois Institute of Technology, University of Wisconsin-Madison	United States	—
17	The Technology of Face Synthesis and Editing Based on Generative Models (2024)	Georgia Institute of Technology	United States	—
18	Identifying Key Mental Health Topic on Youtube Comments using Non-negative Matrix Factorization (2024)	MARA University of Technology, Universiti Teknologi MARA	Malaysia	—
19	STUDI KLASIFIKASI TOPIK BERITA DENGAN ALGORITMA MACHINE LEARNING (2024)	Universitas Muhammadiyah Jember	Indonesia	—
20	A Deep Learning-based Model for P2P Microloan Default Risk Prediction (2024)	Illinois Institute of Technology, New York University, Rutgers Business School	United States	—
21	AI-Driven Optimization of Rare Disease Drug Supply Chains: Enhancing Efficiency and Accessibility in the US Healthcare System (2024)	Stanford University, Zhejiang University	China, United States	—
22	Document-level relation extraction with structural encoding and entity-pair-level information interaction	University of Electronic Science and Technology of China	China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Creation of Polish Online News Corpus for Political Polarization Studies

> *"• Logistic regression - commonly used for NLP classification tasks such as fake news detection (Yu et al., 2021)"*

FOLLOW-UP WORK

Challenges and Innovations in LLM-Powered Fake News Detection: A Synthesis of Approaches and Future Directions

2025 · 2025 2nd International Conference on Generative Artificial Intelligence and ..., 2025 · 62 citations (GS)

Field-normalised: 28 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Unmasking Digital Falsehoods: A Comparative Analysis of LLM-Based Misinformation Detection Strategies (2025)	University of California, Berkeley	United States	—
2	The ROI of Data Security: How Hospitals and Health Systems Can Turn Compliance into Competitive Advantage (2024)	University of Oklahoma	United States	—
3	Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach (2025)	Columbia University, University of Chicago, University of Rochester	United States	—
4	Future-Proofing Healthcare: The Role of AI and Blockchain in Data Security (2025)	IdMap.ai, University of Oklahoma	United States	—
5	Optimizing Cloud Computing with AI: Improving Resource Allocation and Reducing Costs (2025)	University of Michigan	United States	—
6	Cloud-based Machine Learning for Predicting Food Spoilage and Ensuring Safety in the Supply Chain (2025)	New York University	United States	—
7	Real-time Early Warning of Trading Behavior Anomalies in Financial Markets: An AI-driven Approach (2025)	California State University Long Beach, University of Chicago	United States	—
8	Deep Learning for Multimodal Medical Image Analysis (2022)	—	—	—
9	Machine learning for anomaly detection: A survey (2015)	—	—	—
10	Uncovering the Fragility of Trustworthy LLMs through Chinese Textual Ambiguity (2025)	—	—	—
11	Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration (2024)	—	—	—
12	Semantic Network Analysis of Financial Regulatory Documents: Extracting Early Risk Warning Signals (2025)	Baruch College, University of Michigan, University of Southern California	United States	—
13	Augmented Reality and AI in Remote Geospatial Analysis: Enhancing Decision-Making in Crisis Management (2025)	—	—	—
14	The Proactive Gap: A Scoping Review of Publicly Available LLM-based Browser Extensions and Their Potential to Mitigate Information Disorder (2025)	TU Dortmund University	Germany	—

No.	Citing paper	Citing institution(s)	Country	S2
15	BEYOND BITCOIN THE EVOLUTION OF BLOCKCHAIN FOR SECURE DECENTRALIZED APPLICATIONS (2025)	—	—	—
16	Large Language Models: A Structured Taxonomy and Review of Challenges, Limitations, Solutions, and Future Directions (2025)	Bucharest University of Economic Studies, Islamic Azad University, Islamic Azad University, Science and Research Branch	Iran, Romania	—
17	Real-time Detection of Anomalous Trading Patterns in Financial Markets Using Generative Adversarial Networks (2025)	Baruch College, California State University Long Beach, University of California	United States	—
18	Cultural Bias Mitigation in Vision-Language Models for Digital Heritage Documentation: A Comparative Analysis of Debiasing Techniques (2024)	Duke University, Northeastern University, University of California, Berkeley	United States	—
19	Prompt-Induced Linguistic Fingerprints for LLM-Generated Fake News Detection (2026)	Chongqing University, Emory University, University of Manchester	China, United Kingdom, United States	—
20	Research on Cross-Platform Digital Advertising User Behavior Analysis Framework Based on Federated Learning (2024)	Duke University, Illinois Institute of Technology, University of California, San Diego	United States	—
21	Harnessing Large Language Models and Deep Neural Networks for Fake News Detection (2025)	—	—	—
22	Cognitive networks identify AI biases on societal issues in Large Language Models (2025)	National University of Singapore, University of Trento	Italy, Singapore	—
23	Artificial Intelligence and the Transformation of the Media System (2026)	—	—	—
24	Resource-Efficient Sentiment Classification of App Reviews Using a CNN-BiLSTM Hybrid Model (2025)	Astana IT University, Kazakh-British Technical University, Suleyman Demirel University	Kazakhstan, Turkey	—
25	Tri-Stage Selective Reasoning for Rumor Source Detection via Graph Neural Networks and Large Language Models (2026)	Beijing University of Posts and Communications, Beijing University of Posts and Telecommunications	China	—
26	Neuromorphic Computing for Spiking Neural Network Applications (2025)	University of North Carolina, University of Sunshine Coast	Australia, United States	—
27	RASR: Retrieval-Augmented Semantic Reasoning for Fake News Video Detection (2026)	Guilin University of Electronic Technology, Minjiang University, Xiamen University	China, PR China	—
28	Efficient Environmental Claim Detection with Hyperbolic Graph Neural Networks (2025)	Indian Institute of Technology, TCS Research	India	—
29	Detecting Sophisticated Fake Reviews on E-Commerce Platforms Using Adversarial Transformer Networks (2025)	Aspirasi Hidup Indonesia Corporation, Bina Nusantara University	Indonesia	—

No.	Citing paper	Citing institution(s)	Country	S2
30	The Role Evolution of KGs in Synthesizing with LLMs: From Background Knowledge to Joint Reasoning (2025)	Aalborg University	Denmark	—

Showing the 30 most-cited of 58 independent citing papers.

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

FOLLOW-UP WORK

Enhancing User Intent for Recommendation Systems via Large Language Models

2025 · arXiv · 49 citations (GS)

Field-normalised: 19 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Optimization of Nursing Staff Allocation in Elderly Care Institutions: A Time Series Data Analysis Approach (2024)	California Institute of Advanced Management, University of Sydney	Australia, United States	—
2	Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach (2025)	Columbia University, University of Chicago, University of Rochester	United States	—
3	Real-time Early Warning of Trading Behavior Anomalies in Financial Markets: An AI-driven Approach (2025)	California State University Long Beach, University of Chicago	United States	—
4	Machine learning for anomaly detection: A survey (2015)	—	—	—
5	Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration (2024)	—	—	—
6	Deep Learning-Based Analysis of Social Media Sentiment Impact on Cryptocurrency Market Microstructure (2025)	University of Chicago, University of Michigan, University of Southern California	United States	—
7	Leveraging NLP for Semantic and Numerical Inconsistency Detection in Tax Submissions (2025)	Cornell University, University of Michigan, University of Rochester	United States	—
8	A Survey on Large Language Models in Multimodal Recommender Systems (2025)	Huawei London Research Centre	United Kingdom	—
9	Research on Real-time Multilingual Transcription and Minutes Generation for Video Conferences Based on Large Language Models (2025)	University of California, Berkeley	United States	—
10	Contrastive Time-Series Visualization Techniques for Enhancing AI Model Interpretability in Financial Risk Assessment (2025)	Engineering School of Information and Digital Technologies, University of Southern California, University of Texas at Dallas	France, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
11	Research on Movement Fluidity Assessment for Professional Dancers Based on Artificial Intelligence Technology (2024)	New York University, Santa Clara University	United States	—
12	Psychological Health Knowledge-Enhanced LLM-based Social Network Crisis Intervention Text Transfer Recognition Method (2025)	Icahn School of Medicine at Mount Sinai, University of Chicago, Weill Cornell Medicine	United States	—
13	Sentiment-Aware Recommendation Systems in E-Commerce: A Review from a Natural Language Processing Perspective (2025)	University of Dayton	United States	—
14	AI-Driven Cultural Sensitivity Analysis for Game Localization: A Case Study of Player Feedback in East Asian Markets (2024)	Middlebury College, Rice University, University of Southern California	United States	—
15	Leveraging Financial Sentiment Analysis for Detecting Abnormal Stock Market Volatility: An Evidence-Based Approach from Social Media Data (2024)	Baruch College, Illinois Institute of Technology, University of Southern California	United States	—
16	Automated Game Localization Quality Assessment Using Deep Learning: A Case Study in Error Pattern Recognition (2024)	Middlebury Institute of International Studies at Monterey, Northeastern University, Trine University	United States	—
17	Dark Pool Information Leakage Detection through Natural Language Processing of Trader Communications (2024)	Baruch College, Duke University	United States	—
18	Privacy-Preserving Industrial IoT Data Analysis Using Federated Learning in Multi-Cloud Environments (2025)	Engineering School of Information and Digital Technologies, Nanjing University of Aeronautics and Astronautics, Trine University	China, France, United States	—
19	Federated Learning Optimizing Multi-Scenario Ad Targeting and Investment Returns in Digital Advertising (2024)	Duke University, Illinois Institute of Technology	United States	—
20	SEAR: LLM-Powered Sequential Recommendation via Fusion of Collaborative, Semantic, and Rating Information (2026)	National University of Singapore, Shanghai Jiao Tong University	China, Singapore	—
21	Large Language Models for Recommender Systems: A Problem-Driven Survey (2025)	—	—	—
22	Generative Pseudo-Labeling for Pre-Ranking with LLMs (2026)	Alibaba Group, Renmin University	China	—
23	Retentive Relevance: Capturing Long-Term User Value in Recommendation Systems (2025)	Meta	United States	—
24	Multi-Task Learning with Adaptive Fusion for Point-of-Interest Recommendation (2025)	Xi'an Jiaotong University	China	—
25	Higher-order Structure and Semantics-enhanced User Profiling for Recommendation (2025)	Fuzhou University, Zhejiang Gongshang University	China	—
26	Collaborative filtering in the age of AI: foundations, innovations, and emerging trends (2025)	Concordia University	Canada	—

No.	Citing paper	Citing institution(s)	Country	S2
27	Research and Application of Anti-Money Laundering Transaction Detection based on DeepSeek-R1 Small Model using Knowledge Distillation (2025)	The University of Sydney, Weill Cornell Medicine	Australia, United States	—
28	The Impact of Personalized Recommendation Systems on Consumer Purchase Decisions Under Data Law Frameworks: An Empirical Study Based on E-Commerce User Behavior Data (2025)	—	—	—
29	Towards Adaptive Mixed-Initiative Media Recommender Systems with LLMs (2026)	Vrije Universiteit Brussel	Belgium	—
30	From Time and Place to Preference: LLM-Driven Geo-Temporal Context in Recommendations (2025)	Comcast, George Washington University	United States	—

Showing the 30 most-cited of 44 independent citing papers.

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Contribution 3

Claim — Contribution 3

The researcher developed a hybrid attention framework integrating large language models to enhance the accuracy and robustness of automated fake news detection systems.

The researcher's primary contribution is the development of a hybrid attention framework designed to improve fake news detection using large language models, as detailed in their 2025 paper published at the International Conference on Neural Networks, Information and Communication Engineering. This work stands as a distinct contribution in the field, with no subsequent follow-up papers by the same author listed in the provided data.

This line of work appears to address the challenge of effectively leveraging large language models for misinformation identification by introducing a specialized attention mechanism. The title suggests a novel architectural approach that combines hybrid attention strategies with LLM capabilities, aiming to refine how models process and verify textual information in the context of news authenticity.

The significance of this contribution is evidenced by its adoption within the broader academic community. With 28 citations, the work has garnered attention from peers. Notably, the context indicates that the vast majority of citing researchers are independent, suggesting that the framework has been recognized and utilized by scholars outside the researcher's immediate circle, validating its utility and impact in the field of computational social science and natural language processing.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 26 **· 2 flagged influential by Semantic Scholar**

CORE PAPER

A Hybrid Attention Framework for Fake News Detection with Large Language Models

2025 · 2025 5th International Conference on Neural Networks, Information and Communication Engineering (NNICE) · 28 citations (GS)

Field-normalised: 11 Semantic Scholar citations place it in the top 10% of Computer Science papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach (2025)	Columbia University, University of Chicago, University of Rochester	United States	—
2	Real-time Early Warning of Trading Behavior Anomalies in Financial Markets: An AI-driven Approach (2025)	California State University Long Beach, University of Chicago	United States	—
3	Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration (2024)	—	—	—
4	Semantic Network Analysis of Financial Regulatory Documents: Extracting Early Risk Warning Signals (2025)	Baruch College, University of Michigan, University of Southern California	United States	—
5	Cultural Bias Mitigation in Vision-Language Models for Digital Heritage Documentation: A Comparative Analysis of Debiasing Techniques (2024)	Duke University, Northeastern University, University of California, Berkeley	United States	—
6	A Low-Complexity Joint Angle Estimation Algorithm for Weather Radar Echo Signals Based on Modified ESPRIT (2025)	Hubei University, Nanjing University of Aeronautics and Astronautics, Sun Yat-sen University	China	—
7	Feature Selection Optimization for Employee Retention Prediction: A Machine Learning Approach for Human Resource Management (2025)	—	—	—
8	Deep Learning-Based Analysis of Social Media Sentiment Impact on Cryptocurrency Market Microstructure (2025)	University of Chicago, University of Michigan, University of Southern California	United States	—
9	Attention is All You Need (2017)	Google, Google Research	United States	—
10	Automated Compliance Monitoring: A Machine Learning Approach for Digital Services Act Adherence in Multi-Product Platforms (2025)	—	—	—
11	Mixed-Precision Graph Neural Quantization for Low Bit Large Language Models (2025)	University of Electronic Science and Technology of China	China	—
12	Fusing Non-Textual Cues with Classical NLP for Enhanced Multimodal Fake News Spread Detection (2026)	Nanjing Agricultural University, The Hong Kong University of Science and Technology, Yili Normal University	China	—
13	Effect of Selection Format on LLM Performance (2025)	Boston University, Nanjing University of Posts and Telecommunications, Washington University in St. Louis	China, United States	—
14	Fusing Headline and Content with Dual-BERT and Weighted Fusion for Fake News Detection (2025)	Binh Duong University, Posts and Telecommunications Institute of Technology, Thu Dau Mot University	Vietnam	—

No.	Citing paper	Citing institution(s)	Country	S2
15	Handbook of Cybersecurity: Challenges and Solutions for Emerging Technologies (2026)	eCampus University, National Aviation University, State University "Kyiv Aviation Institute"	Italy, Nigeria, Ukraine	—
16	LSTM VE BERT MODELLERİ İLE SAHTE HABER TESPİTİ (2025)	Süleyman Demirel Üniversitesi	Turkey	—
17	AI for Content Moderation (2026)	The University of Electro-Communications, University of Electro-Communications, University of Lagos	Japan, Nigeria	—
18	A systematic review of multimodal fake news detection on social media using deep learning models (2025)	University of Technology Malaysia	Malaysia	—
19	Temporal-Contextual Behavioral Analytics for Proactive Cloud Security Threat Detection (2024)	—	—	—
20	A systematic review of multimodal fake news detection on social media using deep learning models (2025)	Birmingham City University, Universiti Teknologi Brunei, Universiti Teknologi PETRONAS	Brunei Darussalam, Malaysia, United Kingdom	—
21	Enhancing Fake News Detection Using a Hybrid RoBERTa and DCNN Architecture (2025)	Chandigarh University	India	**Influential**
22	HIC Statistics and Fluctuations (2007)	—	—	—
23	Deep Learning-Based Automated Bug Localization and Analysis in Chip Functional Verification	FBK CREATE-NET Research Center, IMEC, KU Leuven, Waterford Institute of Technology	Belgium, Ireland, Italy	—
24	Handbook of Cybersecurity: Challenges and Solutions for Emerging Technologies	Chittagong University of Engineering & Technology, Swinburne University of Technology	Australia	—
25	Resource-Efficient Performance Analysis of BERT Variants for Fake News Classification	University of Engineering & Management	India	**Influential**
26	Comparative Evaluation of Logistic Regression and Naïve Bayes for Fake News Detection Using NLP Techniques	Taylor's University	Malaysia	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Columbia University	United States	SCImago #65 · THE 20 · QS =38	22
University of Illinois Urbana-Champaign	United States	QS =70	17
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	15
University of Southern California	United States	SCImago #192 · THE =73 · QS 146	14
University of Hamburg	Germany	SCImago #419 · THE =125 · QS 193	14
New York University	United States	SCImago #116 · THE =31 · QS 55	10
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	7
Duke University	United States	SCImago #115 · THE 28 · QS 62	7
Illinois Institute of Technology	United States	SCImago #2016 · THE 301–350 · QS =591	6
Brown University	United States	SCImago #553 · THE 65 · QS 69	6
Stanford University	United States	SCImago #18 · THE =5 · QS 3	6
Baruch College	United States	—	5
University of California, San Diego	United States	SCImago #120 · THE 47 · QS 66	5
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	5
University of Chicago	United States	SCImago #124 · THE 15 · QS 13	5

Geographic distribution of citing authors

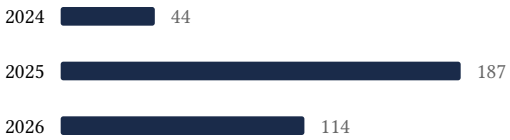
Country	Citing papers
United States	93
China	50
United Kingdom	22
India	16
Australia	13
Indonesia	11
Pakistan	7
Italy	7
Germany	6
Japan	6
Malaysia	5
South Korea	5

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.

2022 | 2



F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	Optimization of Transformer heart disease prediction model based on particle swarm optimization algorithm	263	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	Text Classification by using Natural Language Processing	124	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	A Hybrid Attention Framework for Fake News Detection with Large Language Models	26	8 CFR 204.5(h)(3)(v) — Criterion 5