Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Hailun Ding

IBM Research

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. 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

202	239	13	6
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.

88.4% independent of 95 classified citing papers

Citation type	Count
Independent	84
Self-citation	1
Co-author	9
Same-institution	1

107 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 pioneered storage-efficient logging systems using representation learning, establishing a foundation for automated attack investigation and learned provenance graph storage.

The researcher's core contribution centers on ELISE, a 2021 paper introducing a storage-efficient logging system powered by redundancy reduction and representation learning. This work serves as the foundational pillar for a subsequent line of inquiry into intelligent log management and security analysis.

This line of work appears to address the challenge of efficiently storing and analyzing massive volumes of log data. By applying representation learning to reduce redundancy, the researcher moved beyond traditional compression methods. The follow-up papers, AIRTAG (2023) and a 2023 study on learned provenance graph storage, suggest an expansion of these techniques toward automated attack investigation and specialized graph storage, indicating a coherent evolution from basic efficiency to advanced security applications.

The significance of this work is evidenced by substantial citation activity. The core paper has garnered 37 citations, while the follow-up works have received 72 and 33 citations respectively. Notably, 88.4% of the 95 classified citations originate from independent researchers, demonstrating that the broader academic community recognizes and builds upon these contributions outside the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 36 · **8 flagged influential by Semantic Scholar**

CORE PAPER

{ELISE}: A storage efficient logging system powered by redundancy reduction and representation learning

2021 · 30th USENIX Security Symposium (USENIX Security 21), 3023-3040, 2021 · 37 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	An Insider Threat Investigation Method by Graph Analysis with Log Texts	Chinese Academy of Sciences	China	—
2	" Get in Researchers; We're Measuring Reproducibility": A Reproducibility Study of Machine Learning Papers in Tier 1 Security Conferences	University of Florida	United States	—
3	Poison forensics: Traceback of data poisoning attacks in neural networks	University of Chicago	United States	Background
4	Logshrink: Effective log compression by leveraging commonality and variability of log data	Chongqing University, Sun Yat-sen University, The University of Newcastle	Australia, China	**Methodology**
5	Accurate and scalable detection and investigation of cyber persistence threats	Karlsruhe Institute of Technology, University of Virginia	Germany, United States	—
6	Rethinking tamper-evident logging: A high-performance, co-designed auditing system	Florida State University, University of Virginia	United States	—
7	Mint: Cost-Efficient Tracing with All Requests Collection via Commonality and Variability Analysis	Alibaba Group, Sun Yat-sen University	China	—
8	DEHYDRATOR: Enhancing Provenance Graph Storage via Hierarchical Encoding and Sequence Generation	Zhejiang University of Technology	China	—

No.	Citing paper	Citing institution(s)	Country	S2
9	DNSLogzip: A Novel Approach to Fast and High-Ratio Compression for DNS Logs	Beijing University of Posts and Telecommunications, China Mobile Communications Group Shandong Co., Ltd., China Telecom Co., Ltd.	China	—
10	LogFold: Compressing Logs with Structured Tokens and Hybrid Encoding	Singapore Management University, Sun Yat-sen University	China, Singapore	—

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 Logshrink: Effective log compression by leveraging commonality and variability of log data

"However, they cannot fully disclose the redundancy of log files, which are well formatted and might enable more effective compression [19, 20]."

FOLLOW-UP WORK

{AIRTAG}: Towards automated attack investigation by unsupervised learning with log texts

2023 · 32nd USENIX Security Symposium (USENIX Security 23), 373-390, 2023 · 72 citations (GS)

Field-normalised: 48 Semantic Scholar citations place it in the top 10% of Computer Science papers from 2023 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Deep learning-based intrusion detection systems: A survey	Tsinghua University	China	**Influential**
2	Self-supervised machine learning framework for online container security attack detection	Cisco, Meta Platforms, North Carolina State University	China, United States	—
3	On the Reproducibility of Provenance-based Intrusion Detection that uses Deep Learning	Lahore University of Management Sciences, SRI, University of Arizona	Canada, New Zealand, Pakistan	—
4	Log2graphs: an unsupervised framework for log anomaly detection with efficient feature extraction	Mohamed bin Zayed University of Artificial Intelligence, University of Electronic Science and Technology of China	China, United Arab Emirates	**Methodology**
5	Attack Effect Model based Malicious Behavior Detection	Nanjing University	China	**Influential**
6	Training with Only 1.0‰ Samples: Malicious Traffic Detection via Cross-Modality Feature Fusion	Purdue University, Tsinghua University	China, United States	—
7	From Alerts to Intelligence: A Novel LLM-Aided Framework for Host-based Intrusion Detection	Microsoft, University of California, Irvine, University of California, Los Angeles	United States	**Influential**
8	An End-to-End Framework for Functionality-Embedded Provenance Graph Construction and Threat Interpretation	University of Alberta	Canada	—

No.	Citing paper	Citing institution(s)	Country	S2
9	SoK: Harmonizing Attack Graphs and Intrusion Detection Systems	Delft University of Technology, Sapienza University of Rome	Italy, Netherlands	Influential
10	ProvAgent: Threat Detection Based on Identity-Behavior Binding and Multi-Agent Collaborative Attack Investigation	Chinese Academy of Sciences, Zhongguancun Laboratory	China	—
11	KnowHow: Automatically Applying High-Level CTI Knowledge for Interpretable and Accurate Provenance Analysis	Peking University, Shanghai Jiao Tong University, Southeast University	China	—
12	CLIProv: A Contrastive Log-to-Intelligence Multimodal Approach for Threat Detection and Provenance Analysis	Beijing University of Posts and Telecommunications	China	—
13	After the breach: Incident response within enterprises	UC San Diego	United States	Methodology
14	An Insider Threat Investigation Method by Graph Analysis with Log Texts	Chinese Academy of Sciences	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 Log2graphs: an unsupervised framework for log anomaly detection with efficient feature extraction

"AIRTAG combines these improvements, using BERT for log semantic extraction and introducing one-class support vector machine (OC-SVM), in the detection module [12]."

METHODOLOGY After the breach: Incident response within enterprises

"AirTag's [7] main intuition is that logs and causal graphs are different representations of the same problem, however logs are in Euclidean space, and easier for ML models to learn."

FOLLOW-UP WORK

The case for learned provenance graph storage systems

2023 · 32nd USENIX Security Symposium (USENIX Security 23), 3277-3294, 2023 · 33 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Deep learning-based intrusion detection systems: A survey	Tsinghua University	China	—
2	From Alerts to Intelligence: A Novel LLM-Aided Framework for Host-based Intrusion Detection	Microsoft, University of California, Irvine, University of California, Los Angeles	United States	—
3	Rethinking tamper-evident logging: A high-performance, co-designed auditing system	Florida State University, University of Virginia	United States	—
4	DEHYDRATOR: Enhancing Provenance Graph Storage via Hierarchical Encoding and Sequence Generation	Zhejiang University of Technology	China	Influential
5	Risk taxonomy, mitigation, and assessment benchmarks of large language model systems	Ant Group, Institute of Information Engineering, Chinese Academy of Sciences, Tsinghua University	China	Background

No.	Citing paper	Citing institution(s)	Country	S2
6	A survey on advanced persistent threat detection: a unified framework, challenges, and countermeasures	CSIRO, Nanjing University of Science and Technology, University of Wollongong	Australia, China	—
7	Generating robust counterfactual witnesses for graph neural networks	Aalborg University, Case Western Reserve University	Denmark, United States	—
8	Provenance-enabled explainable ai	Alibaba Cloud, Georgetown University	China, United States	—
9	Omnisec: LLM-driven provenance-based intrusion detection via retrieval-augmented behavior prompting	Wenzhou University of Technology, Zhejiang University of Technology	China	—
10	ProvX: Generating Counterfactual-Driven Attack Explanations for Provenance-Based Detection	Nanyang Technological University, Xidian University	China, Singapore	—
11	How Far Should We Need to Go: Evaluate Provenance-based Intrusion Detection Systems in Industrial Scenarios	Peking University, Tencent, Tsinghua University	China	—
12	LESS: Efficient Log Storage System Based on Learned Model and Minimum Attribute Tree	Tsinghua University	China	**Influential**

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 developed training-time mitigation strategies to address both injected and natural backdoors, establishing a foundational approach for enhancing model robustness against diverse adversarial threats.

The researcher's contribution centers on the 2022 paper 'Training with more confidence: Mitigating injected and natural backdoors during training.' This work appears to propose methods for detecting and neutralizing backdoor vulnerabilities directly within the training process, rather than relying solely on post-training defenses. The title suggests a focus on improving model confidence as a mechanism for security, addressing both maliciously injected triggers and naturally occurring data anomalies that may act as backdoors.

This line of work addresses a critical gap in machine learning security by targeting the training phase, where backdoors are typically implanted. By distinguishing between injected and natural backdoors, the research indicates a nuanced understanding of threat vectors. The absence of follow-up papers by the same researcher in this dataset suggests that this single publication serves as a standalone, seminal contribution that established a clear methodological framework without requiring immediate iterative extensions by the author.

The significance of this work is evidenced by its citation record, with 70 citations indicating strong uptake in the field. Notably, 88.4% of the 95 classified citing papers originate from independent researchers, suggesting that the methodology has been widely adopted and validated by the broader scientific community outside the researcher's immediate circle. This high degree of independent engagement underscores the work's impact as a reliable reference point for subsequent studies in adversarial machine learning.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 25 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Training with more confidence: Mitigating injected and natural backdoors during training

2022 · Advances in Neural Information Processing Systems 35, 36396-36410, 2022 · 70 citations (GS)

Field-normalised: 59 Semantic Scholar citations place it in the top 10% of Computer Science papers from 2022 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Narcissus: A practical clean-label backdoor attack with limited information	Sony AI, Texas A&M University-Commerce, Virginia Tech	Japan, United States	—
2	Ibd-psc: Input-level backdoor detection via parameter-oriented scaling consistency	Alibaba Group, Griffith University, Harbin Institute of Technology	Australia, China, Singapore	Background
3	Certifiably robust graph contrastive learning	The Pennsylvania State University	United States	—
4	Badexpert: Extracting backdoor functionality for accurate backdoor input detection	Princeton University, Zhejiang University	China, United States	Background
5	Sok: The last line of defense: On backdoor defense evaluation	Delft University of Technology, Ikerlan Research Center, Radboud University	Italy, Netherlands, Norway	—
6	Revisiting {Training-Inference} Trigger Intensity in Backdoor Attacks	Xi'an Jiaotong University	China	—
7	Magnitude-based neuron pruning for backdoor defens	Shanghai Jiao Tong University	China	—
8	BackdoorMBTI: A Backdoor Learning Multimodal Benchmark Tool Kit for Backdoor Defense Evaluation	Shanghai Jiao Tong University	China	—
9	Model pairing using embedding translation for backdoor attack detection on open-set classification tasks	Idiap Research Institute	Switzerland	**Methodology**
10	Reconstructive neuron pruning for backdoor defense	Fudan University, Sony AI, University of Copenhagen	China, Denmark, Japan	Background
11	Towards a proactive {ML} approach for detecting backdoor poison samples	Princeton University	United States	**Influential**
12	Black-box backdoor defense via zero-shot image purification	New Jersey Institute of Technology, University of Georgia	United States	**Methodology**
13	Backdoorbox: A python toolbox for backdoor learning	Tsinghua University, Zhejiang University	China	—
14	CLIBE: Detecting dynamic backdoors in transformer-based NLP models	Zhejiang University	China	—
15	Progressive poisoned data isolation for training-time backdoor defense	University of Macau	China, Macau	—
16	Preference poisoning attacks on reward model learning	University of Wisconsin-Madison, Washington University in St. Louis	United States	—
17	Watch out! simple horizontal class backdoor can trivially evade defense	CSIRO, Fudan University, Nanjing University of Science and Technology	Australia, China	—

No.	Citing paper	Citing institution(s)	Country	S2
18	Mask and restore: Blind backdoor defense at test time with masked autoencoder	Stony Brook University	United States	—
19	Secure transfer learning: Training clean models against backdoor in (both) pre-trained encoders and downstream datasets	Griffith University, Huazhong University of Science and Technology	Australia, China	—
20	Bridging distribution shift and ai safety: Conceptual and methodological synergies	New York University, University of California, Santa Barbara	United States	—
21	TruVRF: Towards triple-granularity verification on machine unlearning	Nanjing University of Science and Technology, Zhejiang University	China	—
22	Attacking neural networks with neural networks: Towards deep synchronization for backdoor attacks	Lehigh University, New Jersey Institute of Technology, University of Georgia	United States	—
23	Secure Transfer Learning: Training Clean Model Against Backdoor in Pre-Trained Encoder and Downstream Dataset	Deakin University, Griffith University, Huazhong University of Science and Technology	Australia, China	—
24	Rethinking pruning for backdoor mitigation: an optimization perspective	Shanghai Jiao Tong University	China	Background
25	Mithridates: Auditing and boosting backdoor resistance of machine learning pipelines	University of Massachusetts Amherst	United States	—

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 Model pairing using embedding translation for backdoor attack detection on open-set classification tasks

"With respect to defenses, which are techniques used during training to prevent the learning of backdoors while the genuine behavior is learned, [18] proposes a technique which prevents a machine learning algorithm from learning linear decision regions and filters out inputs that are potentially..."

METHODOLOGY Black-box backdoor defense via zero-shot image purification

"Some purification approaches focus on retraining the model using clean or carefully selected training data to reduce the impact of the backdoor [64, 59, 31, 27, 55]."

Contribution 3

Claim — Contribution 3

The researcher advanced the field of adversarial machine learning by critically re-evaluating and refining methodologies for reverse-engineering trojan triggers in neural networks.

CLAIM: The researcher's significant contribution centers on the 2022 paper 'Rethinking the reverse-engineering of trojan triggers,' which serves as the foundational work in this specific line of inquiry. This publication represents a focused effort to address critical aspects of model security and integrity.

ORIGINALITY: The title suggests a departure from or a critical reassessment of existing paradigms for identifying malicious triggers. By framing the work as 'rethinking' the process, the researcher appears to have identified limitations or inefficiencies

in prior approaches, offering a novel perspective or improved framework for understanding how trojan attacks can be detected and analyzed.

SIGNIFICANCE: The work has garnered substantial attention, with 80 citations indicating its relevance to the community. Notably, 88.4% of the citing papers originate from independent researchers, demonstrating that the contribution has resonated beyond the researcher's immediate circle and has been adopted by the broader scientific community as a valuable reference point.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 24 · **2 flagged influential by Semantic Scholar**

CORE PAPER

Rethinking the reverse-engineering of trojan triggers

2022 · Advances in Neural Information Processing Systems 35, 9738-9753, 2022 · 80 citations (GS)

Field-normalised: 61 Semantic Scholar citations place it in the top 10% of Computer Science papers from 2022 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Narcissus: A practical clean-label backdoor attack with limited information	Sony AI, Texas A&M University-Commerce, Virginia Tech	Japan, United States	—
2	Nearest is not dearest: Towards practical defense against quantization-conditioned backdoor attacks	Central South University, Tencent, Wuhan University	China	Background
3	Ibd-psc: Input-level backdoor detection via parameter-oriented scaling consistency	Alibaba Group, Griffith University, Harbin Institute of Technology	Australia, China, Singapore	—
4	Fedtracker: Furnishing ownership verification and traceability for federated learning model	Hong Kong University of Science and Technology, Sun Yat-sen University, WeBank	China, Hong Kong	**Methodology**
5	Certifiably robust graph contrastive learning	The Pennsylvania State University	United States	—
6	Beear: Embedding-based adversarial removal of safety backdoors in instruction-tuned language models	Georgia Tech, University of California, Berkeley, University of Chicago	United States	Background
7	Badexpert: Extracting backdoor functionality for accurate backdoor input detection	Princeton University, Zhejiang University	China, United States	Background
8	What can discriminator do? towards box-free ownership verification of generative adversarial networks	Chongqing University, Nanjing University of Aeronautics and Astronautics, Wuhan University	China	Background
9	Sok: The last line of defense: On backdoor defense evaluation	Delft University of Technology, Ikerlan Research Center, Radboud University	Italy, Netherlands, Norway	—
10	Data free backdoor attacks	The Pennsylvania State University, University of California Berkeley, University of California, Santa Barbara	United States	—
11	BAN: detecting backdoors activated by adversarial neuron noise	Delft University of Technology, Radboud University, Vrije Universiteit Amsterdam	Netherlands	**Methodology**

No.	Citing paper	Citing institution(s)	Country	S2
12	DISTIL: Data-Free Inversion of Suspicious Trojan Inputs via Latent Diffusion	École polytechnique fédérale de Lausanne (EPFL), EPFL	Switzerland	—
13	Disabling Self-Correction in Retrieval-Augmented Generation via Stealthy Retriever Poisoning	Hong Kong University of Science and Technology, The Hong Kong University of Technology and Science	Hong Kong	—
14	Towards backdoor stealthiness in model parameter space	Delft University of Technology, Radboud University	Netherlands	—
15	Made: Graph backdoor defense with masked unlearning	Peking University	China	—
16	Harmless backdoor-based client-side watermarking in federated learning	Huawei Cloud, The University of Hong Kong	China, Hong Kong	—
17	Revisiting {Training-Inference} Trigger Intensity in Backdoor Attacks	Xi'an Jiaotong University	China	—
18	Adversarially robust anti-backdoor learning	Karlsruhe Institute of Technology (KIT)	Germany	—
19	Magnitude-based neuron pruning for backdoor defens	Shanghai Jiao Tong University	China	**Methodology**
20	Impedance Leakage Vulnerability and Its Utilization in Reverse-Engineering Embedded Software	Florida International University	United States	—
21	BackdoorMBTI: A Backdoor Learning Multimodal Benchmark Tool Kit for Backdoor Defense Evaluation	Shanghai Jiao Tong University	China	—
22	Model pairing using embedding translation for backdoor attack detection on open-set classification tasks	Idiap Research Institute	Switzerland	**Methodology**
23	SecureGaze: Defending Gaze Estimation Against Backdoor Attacks	Delft University of Technology, The Pennsylvania State University	Netherlands, United States	—
24	STEP: Detecting Audio Backdoor Attacks via Stability-based Trigger Exposure Profiling	National University of Singapore, Xi'an Jiaotong University, Zhejiang University	China, Singapore	—

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 Fedtracker: Furnishing ownership verification and traceability for federated learning model

"FSS before and after the overwriting attack. and FeatureRE [56] trying to remove the watermark inside the FL model."

METHODOLOGY BAN: detecting backdoors activated by adversarial neuron noise

"To address this problem, featureRE [37] proposes a detection method using feature space triggers."

METHODOLOGY Magnitude-based neuron pruning for backdoor defens

"Some advanced detection methods also conduct reverse engineering with the backdoored model to recover the trigger pattern [31, 28, 33, 12]."

METHODOLOGY Model pairing using embedding translation for backdoor attack detection on open-set classification tasks

"Later methods mostly alleviate the requirement for training data though some methods still require access to clean data such as [10], [12]–[15], [17] which may be easier to find."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Rutgers University	United States	—	8
Zhejiang University	China	SCImago #6 · THE 39 · QS 49	8
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	6
Sun Yat-sen University	China	SCImago #40 · THE 201–250 · QS =276	5
Peking University	China	SCImago #11 · THE 13 · QS 14	5
Delft University of Technology	Netherlands	SCImago #359 · THE 57 · QS =47	5
University of Massachusetts Amherst	United States	SCImago #788 · QS =247	4
New Jersey Institute of Technology	United States	SCImago #2104 · THE 501–600 · QS 761-770	4
Shanghai Jiao Tong University	China	SCImago #10 · THE 40 · QS =47	4
Virginia Tech	United States	—	4
University of Georgia	United States	SCImago #597 · THE 351–400 · QS 525	3
Radboud University	Netherlands	QS 279	3
Fudan University	China	SCImago #46 · THE 36 · QS 30	3
New York University	United States	SCImago #116 · THE =31 · QS 55	3
Chinese Academy of Sciences	China	SCImago #2	3

Geographic distribution of citing authors

Country	Citing papers
China	48
United States	46
Australia	6
Singapore	6
Netherlands	5
Hong Kong	4
Germany	4
Italy	3
Denmark	2
Japan	2
New Zealand	2
Canada	2

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.

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	{ELISE}: A storage efficient logging system powered by redundancy reduction and representation learning	36	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Training with more confidence: Mitigating injected and natural backdoors during training	25	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	Rethinking the reverse-engineering of trojan triggers	24	8 CFR 204.5(i)(3) — Outstanding Researcher