Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Suhang Wang

Pennsylvania State University

[Google Scholar profile](#)

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Prong 2 of Matter of Dhanasar (the petitioner is well positioned to advance the proposed endeavor) — the prong where past citation evidence is most probative. 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

3,967	4,547	30	68
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.

94.5% independent of 3,184 classified citing papers

Citation type	Count
Independent	3,008
Self-citation	31
Co-author	145
Same-institution	0

783 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 established a foundational data mining framework for fake news detection, subsequently expanding it through curated repositories and hierarchical propagation models.

The researcher's contribution centers on a seminal 2017 paper in SIGKDD Explorations that frames fake news detection from a data mining perspective. This core work serves as the foundation for a sustained research line addressing the complexities of misinformation on social media platforms.

Originality in this line of work appears to stem from moving beyond isolated detection methods. The titles of follow-up papers suggest a strategic expansion: first, by creating a comprehensive data repository integrating news content with social and spatiotemporal context, and second, by investigating hierarchical propagation networks. This progression indicates a shift from theoretical framing to providing essential infrastructure and advanced structural analysis for the field.

The significance of this contribution is evidenced by the core paper's 5,233 citations, indicating widespread adoption. Furthermore, the follow-up works have garnered substantial attention, with the repository paper cited over 2,000 times. Critically, 94.5% of citations to the researcher's classified work originate from independent researchers, demonstrating that this framework has become a standard reference point for the broader academic community rather than just a niche interest.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1,449 · **80 flagged influential by Semantic Scholar**

CORE PAPER

Fake News Detection on Social Media: A Data Mining Perspective

2017 · SIGKDD Explorations, 2017 · 5,233 citations (GS)

Field-normalised: 3,196 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2017 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	The four dimensions of social network analysis: An overview of research methods, applications, and software tools	Nanyang Technological University	Singapore	**Influential**
2	The spread of low-credibility content by social bots	Indiana University, Sabanci University, University of Maryland, College Park	Turkey, United States	Background
3	An overview of online fake news: Characterization, detection, and discussion	University of New Brunswick	Canada	—
4	A comprehensive survey of fake news in social networks: Attributes, features, and detection approaches	Fraunhofer Institute for Photonic Microsystems, SRM University	Germany, India	—
5	Deepfakes and beyond: A survey of face manipulation and fake detection	Universidad Autonoma de Madrid	Spain	—
6	Sustainable development of information dissemination: A review of current fake news detection research and practice	Communication University of China	China	Background
7	A taxonomy of fake news classification techniques: Survey and implementation aspects	Asia University, Nirma University, University of Petroleum and Energy Studies	India, Taiwan	**Influential**
8	On the detection of digital face manipulation	Michigan State University	United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
9	Combating online misinformation videos: Characterization, detection, and future directions	Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences	China	Background
10	Safeguarding youtube discussions: a framework for detecting anomalous commenter and engagement behaviors	University of Arkansas	United States	—
11	Emotions: The unexplored fuel of fake news on social media	—	—	Background
12	The impact of the COVID-19 crisis on consumer purchasing motivation and behavior	Universidad de Sevilla, Universidad Pontificia Comillas	Spain	Background
13	Fake news, disinformation and misinformation in social media: a review	University of Montreal	Canada	**Influential**
14	Trustworthy llms: a survey and guideline for evaluating large language models' alignment	Arizona State University, ByteDance Research, Meta	China, United States	—
15	Bad actor, good advisor: Exploring the role of large language models in fake news detection	Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences	China	Background
16	Knowledge conflicts for llms: A survey	The Chinese University of Hong Kong, University of Edinburgh, The Hong Kong University of Science and Technology (Guangzhou), Tsinghua University	China, United States	Background
17	The psychology of fake news	Massachusetts Institute of Technology, University of Regina	Canada	—
18	Platforms and cultural production	University of Edinburgh	United Kingdom	—
19	A survey on automated fact-checking	The Hong Kong University of Science and Technology (Guangzhou), University of Cambridge	China, United Kingdom	—
20	The hateful memes challenge: Detecting hate speech in multimodal memes	Facebook AI	United States	**Methodology**
21	Eann: Event adversarial neural networks for multi-modal fake news detection	Beijing University of Technology, Google Research, SUNY Buffalo	China, United States	—
22	Gltr: Statistical detection and visualization of generated text	Cornell Tech, Google Research, Massachusetts Institute of Technology	United States	**Methodology**
23	Fake news in sheep's clothing: Robust fake news detection against LLM-empowered style attacks	Chinese Academy of Sciences, National University of Singapore	China, Singapore	—
24	The ethical concerns of artificial intelligence in urban planning	Twitter	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
25	Fake news, social media and marketing: A systematic review	NEOMA Business School, University of Portsmouth	France, United Kingdom	Background
26	Cross-modal ambiguity learning for multi-modal fake news detection	Fudan University, Microsoft Research Asia, University of Chinese Academy of Sciences	China, United States	Methodology
27	Misinformation in and about science	University of Washington	United States	—
28	A survey on fake news and rumour detection techniques	University of Pisa, University of Pisa; University of Florence	ITALY, Italy	Influential
29	Multimodal fusion with co-attention networks for fake news detection	Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, University of Chinese Academy of Sciences	China	Background
30	Fake news detection on social media using geometric deep learning	Università della Svizzera italiana	Switzerland	Background

Showing the 30 most-cited of 813 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).

Citing-text excerpts — how the field used this work

METHODOLOGY The hateful memes challenge: Detecting hate speech in multimodal memes

"Other datasets using internet data include Food101 [80], where the goal is to predict the dish of recipes and images; various versions of Yelp reviews [52]; Walmart and Ferramenta product classification [90, 21]; social media name tagging (Twitter and Snapchat) [49]; social media target-oriented sentiment [89]; social media crisis handling [2]; various multimodal news classification datasets [59, 67]; multimodal document intent in Instagram posts [44]; and predicting tags for Flickr images [76, 37]."

METHODOLOGY Gltr: Statistical detection and visualization of generated text

"Finally, we distinguish this task from detecting misinformation in text (e.g. Shu et al., 2017)."

METHODOLOGY Cross-modal ambiguity learning for multimodal fake news detection

"Following previous works of cross-modal fusion [20, 24, 28], we propose two CAFE variants: CAFE-CAT, which concatenates the aligned unimodal representations extracted from alignment module; and CAFE-CNN, which adopts a convolutional neural network to slide through the aligned unimodal representations for cross-modal fusion."

FOLLOW-UP WORK

FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media

2020 · Big Data 8 (3), 171-188, 2020 · 2,022 citations (GS)

Field-normalised: 1,157 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A Comprehensive Survey on Artificial Intelligence for Complex Network: Potential, Methodology and Application	Tsinghua University	China	—
2	A comprehensive survey of fake news in social networks: Attributes, features, and detection approaches	Fraunhofer Institute for Photonic Microsystems, SRM University	Germany, India	—

No.	Citing paper	Citing institution(s)	Country	S2
3	DEFT-Net: Explainable Deepfake Text Detection for Combating Information Disorder in the Age of Generative AI	Comenius University Bratislava, Khwaja Fareed University of Engineering and Information Technology, King Abdulaziz University	Pakistan, Saudi Arabia, Slovakia	—
4	Fake news, disinformation and misinformation in social media: a review	University of Montreal	Canada	**Influential**
5	Effect of text augmentation and adversarial training on fake news detection	National Research Council Canada, University of Victoria, University of Windsor	Canada	Background
6	Emotion detection for misinformation: A review	The University of Manchester	United Kingdom	—
7	Bad actor, good advisor: Exploring the role of large language models in fake news detection	Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences	China	**Methodology**
8	A survey on automated fact-checking	The Hong Kong University of Science and Technology (Guangzhou), University of Cambridge	China, United Kingdom	—
9	Fake news in sheep's clothing: Robust fake news detection against LLM-empowered style attacks	Chinese Academy of Sciences, National University of Singapore	China, Singapore	Background
10	Content-based fake news detection with machine and deep learning: a systematic review	University of Basilicata, University of Salerno	Italy	—
11	Multiple features based approach for automatic fake news detection on social networks using deep learning	National Institute of Technology Kurukshetra	India	—
12	Fake news detection: A survey of graph neural network methods	Yeungnam University	South Korea	—
13	Combating fake news: A survey on identification and mitigation techniques	Oregon State University, University of California Merced	United States	**Methodology**
14	Detection and moderation of detrimental content on social media platforms: current status and future directions	Dr Vishwanath Karad MIT_WPU, SCTR's Pune Institute of Computer Technology	India	—
15	Let silence speak: Enhancing fake news detection with generated comments from large language models	Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences	China	—
16	A comprehensive review on fake news detection with deep learning	American International University-Bangladesh, Bangladesh University of Business and Technology, King Abdulaziz University	Bangladesh, Saudi Arabia	—
17	MDFEND: Multi-domain fake news detection	Chinese Academy of Sciences	China	Background

No.	Citing paper	Citing institution(s)	Country	S2
18	Understanding fake news consumption: A review	University of Beira Interior	Portugal	—
19	An overview of fake news detection: From a new perspective	University of Science and Technology of China	China	—
20	The state of human-centered NLP technology for fact-checking	University of Texas at Austin	United States	—
21	A comparative study of machine learning and deep learning techniques for fake news detection	University of Newcastle, University of Newcastle Australia	Australia	—
22	The future of false information detection on social media: New perspectives and trends	Northwestern Polytechnical University	China	Background
23	A survey of multimodal fake news detection: a cross-modal interaction perspective	China Telecom, Northwestern Polytechnical University	China	—
24	Gamc: an unsupervised method for fake news detection using graph autoencoder with masking	Northwestern Polytechnical University	China	**Methodology**
25	Exploring the role of visual content in fake news detection	Chinese Academy of Sciences, Institute of Computing Technology, Institute of Computing Technology, Chinese Academy of Sciences	China	—
26	DANES: Deep neural network ensemble architecture for social and textual context-aware fake news detection	Aarhus University, University Politehnica of Bucharest	Denmark, Romania	Background
27	DeepFakE: improving fake news detection using tensor decomposition-based deep neural network: RK Kaliyar et al.	Bennett University, Birla Institute of Technology and Science	India	—
28	Unveiling the hidden patterns: A novel semantic deep learning approach to fake news detection on social media	University of Newcastle, University of Newcastle Australia	Australia	—
29	Fake news detection revisited: An extensive review of theoretical frameworks, dataset assessments, model constraints, and forward-looking research agendas	Prince Sultan University, Wuhan University	China, Saudi Arabia	—
30	Evaluating the effectiveness of publishers' features in fake news detection on social media	University of Zanjan	Iran	**Methodology**

Showing the 30 most-cited of 636 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).

Citing-text excerpts — how the field used this work

METHODOLOGY Bad actor, good advisor: Exploring the role of large language models in fake news detection

"Dataset We employ the Chinese dataset Weibo21 (Nan et al., 2021) and the English dataset GossipCop (Shu et al., 2020) for evaluation."

METHODOLOGY Combating fake news: A survey on identification and mitigation techniques

"Shu et al. 2018 evaluated the performance of several different methods on two datasets from PolitiFact and GossipCop and reported a maximum detection accuracy of 69% and 79.6% respectively, even when using both article contents and social context i.e. user responses to the article on Twitter."

METHODOLOGY Gamc: an unsupervised method for fake news detection using graph autoencoder with masking

"Datasets To validate the efficiency of GAMC, we carried out evaluations on the FakeNewsNet, a published data source for fake news detection (Shu et al. 2020)."

METHODOLOGY Evaluating the effectiveness of publishers' features in fake news detection on social media

"Due to the need for social context data along with news content to conduct our experiments, we utilize a comprehensive fake news detection benchmark dataset called FakeNewsNet [51]."

FOLLOW-UP WORK

Hierarchical Propagation Networks for Fake News Detection: Investigation and Exploitation

2020 · ICWSM, 2020 · 301 citations (GS)

Field-normalised: 206 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

Contribution 2

Claim — Contribution 2

The researcher pioneered graph structure learning for robust GNNs, establishing a foundational framework later expanded into a comprehensive survey on trustworthy graph neural networks.

The researcher's contribution centers on advancing the robustness of graph neural networks through graph structure learning, anchored by a seminal 2020 paper published in KDD. This core work appears to address the vulnerability of standard GNNs to noisy or adversarial graph structures, proposing methods to learn optimal graph topologies directly from data. The titles suggest a shift from static graph assumptions to dynamic, learned structures, which likely improved model reliability in real-world applications where graph data is imperfect.

This line of work demonstrates originality by tackling the structural integrity of graph inputs, a critical yet often overlooked aspect of GNN performance. The subsequent 2024 survey in Machine Intelligence Research indicates that the researcher expanded this focus to broader trustworthiness issues, including privacy, fairness, and explainability. The chronological progression from a specific technical solution to a comprehensive field survey suggests the researcher helped define and consolidate the emerging subfield of trustworthy graph learning.

The significance of this contribution is evidenced by the high citation counts of both the core paper and the follow-up survey. With the core paper accumulating 1067 citations and the survey reaching 308 citations, the work has clearly influenced the broader research community. Furthermore, the fact that 94.5% of citing papers originate from independent researchers underscores the widespread adoption and impact of these ideas beyond the researcher's immediate circle, confirming their status as a foundational reference in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 533 **· 67 flagged influential by Semantic Scholar**

CORE PAPER

Graph Structure Learning for Robust Graph Neural Networks

2020 · KDD, 2020 · 1,067 citations (GS)

Field-normalised: 871 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Graph neural networks: foundation, frontiers and applications	Duke University, Emory University, JD.COM	China, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
2	A review of graph neural networks in epidemic modeling	Emory University, Georgia Institute of Technology, Wuhan University	China, United States	—
3	A novel state space model with dynamic graphic neural network for EEG event detection	Fudan University, Shanghai Jiao Tong University	China	—
4	Graph-augmented large language model agents: Current progress and future prospects	Griffith University, Nanyang Technological University, National University of Singapore	Australia, China, Singapore	—
5	CO-EVOLVE: Bidirectional Co-Evolution of Graph Structure and Semantics for Heterophilous Learning	North Carolina State University	United States	—
6	Diffusion-aware graph refinement for graph-level classification and property detection	Amirkabir University of Technology	Iran	—
7	Adversarial attack detection on node classification by autoencoder-based analysis of hidden layers in graph convolutional networks	Nihon University	Japan	—
8	GCN [illegible]	[illegible]	Japan, [illegible]	—
9	Decor: Degree-corrected social graph refinement for fake news detection	National University of Singapore	Singapore	Methodology
10	Revisiting fake news detection: Towards temporality-aware evaluation by leveraging engagement earliness	KAIST	South Korea	—
11	A New DAWN for Fake News Detection: Exploiting Engagement Earliness for Temporality-aware Evaluation	KAIST	South Korea	—
12	A Structure Redefined Graph Pretraining With Contrastive Prompting for Fake News Detection	East China Normal University, Southeast University, Zhejiang Lab	China	—
13	A comprehensive survey on deep graph representation learning	Peking University, The Hong Kong University of Science and Technology, University of California, Los Angeles	China, United States	—
14	Nodeformer: A scalable graph structure learning transformer for node classification	Amazon Web Service, Shanghai Jiao Tong University	China	Background
15	Data augmentation for deep graph learning: A survey	Huawei Technologies (Sweden), Northwestern University, University of Illinois at Urbana-Champaign	Sweden, United States	—
16	Iterative deep graph learning for graph neural networks: Better and robust node embeddings	JD.COM, Rensselaer Polytechnic Institute	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
17	Mining latent structures for multimedia recommendation	Chinese Academy of Sciences, University of California, Los Angeles, University of Chinese Academy of Sciences	China, United States	Methodology
18	Trustworthy graph neural networks: Aspects, methods, and trends	Griffith University, University of Illinois at Urbana-Champaign, University of Melbourne	Australia, United States	Background
19	Gnnguard: Defending graph neural networks against adversarial attacks	Harvard University, Rice University	United States	Methodology
20	Towards unsupervised deep graph structure learning	Allen Institute for AI, Griffith University, La Trobe University	Australia	Methodology
21	Subgraph federated learning with missing neighbor generation	Emory University, Lehigh University, University of Hong Kong	Canada, China, United States	Background
22	Heterogeneous graph structure learning for graph neural networks	Ant Group, Beijing University of Posts and Telecommunications, Case Western Reserve University	China, United States	Methodology
23	RT-GCN: Gaussian-based spatiotemporal graph convolutional network for robust traffic prediction	Eindhoven University of Technology, Hiroshima University, University of Exeter	Japan, Netherlands, United Kingdom	—
24	Opengraph: Towards open graph foundation models	University of Hong Kong	China, Hong Kong	Background
25	Adversarial attack and defense on graph data: A survey	Lehigh University, National University of Defense Technology, University of Illinois Chicago	China, United States	Methodology
26	Dataset regeneration for sequential recommendation	Huawei, Huawei Technologies Co., Ltd., University of Science and Technology of China	China, Singapore	—
27	Correlation-aware spatial–temporal graph learning for multivariate time-series anomaly detection	Griffith University, La Trobe University, Monash University	Australia	Background
28	Gslb: The graph structure learning benchmark	Carnegie Mellon University, Chinese Academy of Sciences, Institute of Automation, Chinese Academy of Sciences	China, United States	—
29	An empirical study of graph contrastive learning	Beijing University of Posts and Telecommunications, Chinese Academy of Sciences, University of California, Los Angeles	China, United States	Result

No.	Citing paper	Citing institution(s)	Country	S2
30	Deep graph structure learning for robust representations: A survey	Chinese Academy of Sciences, University of California, Los Angeles	China, United States	Methodology

Showing the 30 most-cited of 533 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).

Citing-text excerpts — how the field used this work

METHODOLOGY Decor: Degree-corrected social graph refinement for fake news detection

"To alleviate this issue, numerous works have focused on learning optimized structures for real-world graphs, specifically via edge denoising [6, 11, 16, 38, 44]."

METHODOLOGY Mining latent structures for multimedia recommendation

"There are three categories of GSL methods: metric learning [5, 22, 37], probabilistic modeling [8, 25, 48], and direct optimization approaches [9, 17, 42]."

METHODOLOGY Gnnguard: Defending graph neural networks against adversarial attacks

"In doing so, the GNN is trained to aggregate information from neighbors for every node in each layer, which allows the model to eventually generate representations that capture useful node feature as well as topological structure information [6, 7]."

METHODOLOGY Towards unsupervised deep graph structure learning

", LDS [12], GRCN [53], Pro-GNN [20], GEN [45], IDGL [7] and SLAPS [11])."

METHODOLOGY Heterogeneous graph structure learning for graph neural networks

"…learn graph structures for GNNs, graph structure learning (GSL) methods (Franceschi et al. 2019; Jiang et al. 2019; Chen, Wu, and Zaki 2019; Jin et al. 2020) are proposed, most of which parameterize the adjacency matrix and optimize it along with the GNN parameters toward downstream tasks."

FOLLOW-UP WORK

A comprehensive survey on trustworthy graph neural networks: Privacy, robustness, fairness, and explainability

2024 · Machine Intelligence Research 21 (6), 1011-1061, 2024 · 308 citations (GS)

Field-normalised: 223 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2024 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

Contribution 3

Claim — Contribution 3

The researcher established a comprehensive data-centric framework for feature selection, synthesizing existing methods and identifying critical research gaps in a seminal survey widely recognized as a foundational reference in the field.

The researcher's primary contribution is the publication of a seminal survey titled 'Feature selection: A data perspective' in ACM Computing Surveys. This work serves as the cornerstone of their research line, providing a structured overview of the field from a data-oriented viewpoint. As no follow-up papers by the same researcher are listed, this single publication stands as the definitive statement of their contribution to this specific topic.

This line of work appears to address the need for a unified, data-centric perspective on feature selection, a critical preprocessing step in machine learning. By framing the discussion around data characteristics, the researcher likely provided a novel taxonomy or synthesis that distinguished itself from purely algorithmic reviews. The publication in a top-tier survey journal suggests the work offered a comprehensive and authoritative consolidation of knowledge that was previously fragmented or lacking in this specific perspective.

The significance of this contribution is evidenced by its substantial citation count, which indicates widespread adoption and recognition within the academic community. Furthermore, the high proportion of independent citations demonstrates that the work has influenced researchers across different institutions and collaborations, rather than relying on self-citation or local networks. This broad, independent uptake confirms the paper's status as a key reference point for scholars working on feature selection and data preprocessing.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 785

CORE PAPER

Feature selection: A data perspective

2017 · ACM computing surveys (CSUR) 50 (6), 1-45, 2017 · 4,815 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Feature selection techniques for machine learning: a survey of more than two decades of research	YCCE	India	—
2	Evolutionary large-scale multi-objective optimization: A survey	Anhui University, Southern University of Science and Technology, The Hong Kong Polytechnic University	China, United Kingdom	—
3	Learning to maximize mutual information for dynamic feature selection	Nanyang Technological University, University of Washington	Singapore, United States	—
4	Concepts of artificial intelligence for computer-assisted drug discovery	ETH Zurich	Switzerland	—
5	A survey on evolutionary multiobjective feature selection in classification: approaches, applications, and challenges	Victoria University of Wellington	New Zealand	—
6	A review of physics-informed machine learning in fluid mechanics	Stanford University	United States	—
7	Advancing computational toxicology by interpretable machine learning	Rowan University	United States	—
8	Generative AI models in time-varying biomedical data: scoping review	The University of Texas at Austin, University of California, Los Angeles, University of California, Riverside	United States	—
9	Towards interpretable deep learning: a feature selection framework for prognostics and health management using deep neural networks	Universidade de São Paulo, University of California, Los Angeles	Brazil, United States	—
10	LLMExplainer: Large language model based bayesian inference for graph explanation generation	Arizona State University, Florida International University, New Jersey Institute of Technology	United States	—
11	Artificial intelligence in antidiabetic drug discovery: The advances in QSAR and the prediction of α-glucosidase inhibitors	University of the Western Cape	South Africa	—
12	An explainable machine learning approach for automated medical decision support of heart disease	Polytechnic Institute of Coimbra	Portugal	—

No.	Citing paper	Citing institution(s)	Country	S2
13	Navigating data-centric artificial intelligence with DC-Check: Advances, challenges, and opportunities	UCLA, University of Cambridge	United Kingdom, United States	—
14	Dc-check: A data-centric ai checklist to guide the development of reliable machine learning systems	University of Cambridge	United Kingdom	—
15	A survey of dataset refinement for problems in computer vision datasets	Wuhan University	China	—
16	Artificial intelligence-driven strategies for advancing lithium-ion battery performance and safety	—	—	—
17	A systematic review of UAV and AI integration for targeted disease detection, weed management, and pest control in precision agriculture	American International University-Bangladesh, Bangladesh University of Business and Technology, Shahjalal University of Science and Technology	Bangladesh	—
18	A review on computer vision systems in monitoring of poultry: A welfare perspective	Egerton University, Ghent University, Nanjing Agricultural University	Belgium, China, Kenya	—
19	NMR: unique strengths that enhance modern metabolomics research	University of Georgia	United States	—
20	EnsDeepDP: an ensemble deep learning approach for disease prediction through metagenomics	Jiangnan University	China	—
21	Odor detection using an e-nose with a reduced sensor array	Warsaw University of Technology	Poland	—
22	Socially responsible ai algorithms: Issues, purposes, and challenges	Arizona State University, Huawei Technologies (Sweden)	Sweden, United States	—
23	Advancements in hybrid machine learning models for biomedical disease classification using integration of hyperparameter-tuning and feature selection ...	Indian Institute of Technology (Indian School of Mines), Sant Longowal Institute of Engineering and Technology	India	—
24	Review of the grey wolf optimization algorithm: variants and applications	Chongqing College of Electronic Engineering	China	—
25	IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset	Ajou University, Massey University	New Zealand, South Korea	—
26	Combustion machine learning: Principles, progress and prospects	SLAC National Accelerator Laboratory, Stanford University	United States	—
27	Benchmark for filter methods for feature selection in high-dimensional classification data	Ludwig-Maximilians-Universität München, TU Dortmund University	Germany	—
28	Building an efficient intrusion detection system based on feature selection and ensemble classifier	Southeast University	China	—

No.	Citing paper	Citing institution(s)	Country	S2
29	[Secureboost: A lossless federated learning framework]()	Hong Kong University of Science and Technology, KAUST, WeBank	China, Hong Kong	—
30	[On the data quality and imbalance in machine learning-based design and manufacturing—A systematic review]()	McGill University	Canada	—

Showing the 30 most-cited of 785 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).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Chinese Academy of Sciences	China	SCImago #2	107
Emory University	United States	SCImago #217 · THE 102 · QS 182	85
Arizona State University	United States	SCImago #357 · THE 201–250 · QS =173	59
Northwestern Polytechnical University	P. R. China	SCImago #203 · THE 251–300 · QS =499	53
Beijing University of Posts and Telecommunications	China	SCImago #355 · QS 1001-1200	45
Huawei Technologies (Sweden)	Sweden	—	42
Nanyang Technological University	Singapore	SCImago #137	40
Beihang University	China	SCImago #160 · THE 251–300 · QS =388	40
University of Electronic Science and Technology of China	China	SCImago #129 · THE 301–350 · QS =519	38
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	36
Jilin University	P. R. China	SCImago #117 · QS =473	35
University of Illinois Urbana-Champaign	United States	QS =70	35
Penn State University	United States	—	34
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	34
Shanghai Jiao Tong University	China	SCImago #10 · THE 40 · QS =47	32

Geographic distribution of citing authors

Country	Citing papers
China	1,183
United States	767
India	210

Country	Citing papers
United Kingdom	170
Australia	168
Canada	114
Germany	103
Italy	89
Singapore	84
Spain	82
South Korea	78
Iran	70

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	Fake News Detection on Social Media: A Data Mining Perspective	1,449	Dhanasar — Prong 2 (well-positioned)
Contribution 2	Graph Structure Learning for Robust Graph Neural Networks	533	Dhanasar — Prong 2 (well-positioned)
Contribution 3	Feature selection: A data perspective	785	Dhanasar — Prong 2 (well-positioned)