Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Andrew Ng

Stanford University

Google Scholar profile

Generated 2026-05-31 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

2,537	2,557	4	160
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.

95.8% independent of 1,107 classified citing papers

Citation type	Count
Independent	1,060
Self-citation	1
Co-author	14
Same-institution	32

1,430 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 introduced Latent Dirichlet Allocation, a foundational probabilistic model for topic discovery that has become a standard method in machine learning and natural language processing.

The researcher's primary contribution is the development of Latent Dirichlet Allocation, published in the Journal of Machine Learning Research in 2003. This work stands as a seminal core paper in the field, establishing a robust framework for analyzing document collections. The titles and metadata indicate that this single publication represents a distinct and complete line of inquiry, without subsequent follow-up papers by the same researcher building directly upon it.

This line of work appears to address the need for scalable, generative models capable of uncovering latent thematic structures within large text corpora. By introducing a hierarchical Bayesian approach, the researcher provided a novel method for topic modeling that likely improved upon previous techniques in flexibility and interpretability. The absence of follow-up papers suggests the core contribution was sufficiently comprehensive to stand alone as a definitive solution to this specific problem.

The significance of this work is evidenced by its extensive adoption within the scientific community. With over 62,000 citations, the paper has clearly influenced a vast array of subsequent research. Furthermore, citation analysis reveals that 96.9% of citing papers originate from independent researchers, indicating that the contribution has been widely validated and utilized by the broader field rather than just the researcher's immediate circle. This high level of independent uptake underscores the work's foundational status and broad impact on machine learning and data science.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 490 · **77 flagged influential by Semantic Scholar**

CORE PAPER

Latent Dirichlet Allocation

2003 · Journal of Machine Learning Research (JMLR) · 62,524 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Learning Transferable Visual Models From Natural Language Supervision	OpenAI	United States	Background
2	The simple macroeconomics of AI	Massachusetts Institute of Technology	United States	—
3	HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models	Renmin University of China, Université de Montréal	Canada, China	**Methodology**
4	BERTopic: Neural topic modeling with a class-based TF-IDF procedure	Comprehensive Cancer Organisation Netherlands (IKNL)	Netherlands	**Methodology**
5	Artificial intelligence for literature reviews: opportunities and challenges	The Open University	United Kingdom	—
6	A comprehensive survey of loss functions and metrics in deep learning	Autonomous University of Queretaro, Instituto Politécnico Nacional, National Polytechnic Institute	Mexico	—
7	Artificial intelligence in E-Commerce: a bibliometric study and literature review	ICN Business School, TBS Business School, University of Wollongong	Australia, France	—

No.	Citing paper	Citing institution(s)	Country	S2
8	Machine learning in medicine: a practical introduction	Harvard Medical School, University of Cambridge	United Kingdom, United States	—
9	Data Mining: The Textbook	IBM T. J. Watson Research Center, IBM T.J. Watson Research Center	United States	Methodology
10	"What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter	Mahidol University International College	Thailand	—
11	SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks	VIB Center for Brain & Disease Research	Belgium	—
12	Deep learning for urban land use category classification: A review and experimental assessment	The University of Hong Kong, Tsinghua University	China, Hong Kong	—
13	Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects	Jiujiang University, Taizhou University	China	—
14	Networks beyond pairwise interactions: Structure and dynamics	CENTAI	Italy	Background
15	Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems	Technische Universität Berlin, University of Cambridge, University of Luxembourg	Germany, Luxembourg, United Kingdom	—
16	GAN Inversion: A Survey	ETH Zürich, Tsinghua University, University College London	China, Switzerland, United Kingdom	Methodology
17	Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations	—	—	—
18	A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques	University of Georgia	United States	—
19	Data Mining: Practical Machine Learning Tools and Techniques	Polytechnique Montréal, University of Waikato	Canada, New Zealand	—
20	Lessons from a decade of adaptive pathways studies for climate adaptation	Delft University of Technology, Deltares, Victoria University of Wellington	Netherlands, New Zealand	—
21	A Model of Text for Experimentation in the Social Sciences	Harvard University, University of California, Irvine Medical Center	United States	—
22	A Survey on Data Collection for Machine Learning: A Big Data-AI Integration Perspective	KAIST, Korea Advanced Institute of Science and Technology	South Korea	Background
23	Transformative privacy calculus: Conceptualizing the personalization-privacy paradox on social media	Université Jean Moulin Lyon 3	France	—

No.	Citing paper	Citing institution(s)	Country	S2
24	A review of stochastic block models and extensions for graph clustering	Newcastle University	United Kingdom	Methodology
25	Uniting the Tribes: Using Text for Marketing Insight	Columbia University, Emory University, Northwestern University	United States	Methodology
26	Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda	Aarhus University	Denmark	Methodology
27	The Role of Big Data and Predictive Analytics in Retailing	Dartmouth College, Indian School of Business, University of Pennsylvania	India	Background
28	Climate Change Concerns and the Performance of Green vs. Brown Stocks	Ghent University, HEC Montréal	Belgium, Canada	—
29	Recent Trends in Deep Learning Based Natural Language Processing	Beijing Institute of Technology, Nanyang Technological University, National University of Singapore	China, Singapore	Background
30	Using natural language processing to analyse text data in behavioural science	Columbia University, JLU Giessen, LMU Munich	Canada, Germany, United States	—

Showing the 30 most-cited of 490 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 HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models

"To further understand the failures of ChatGPT, we visualize the topics of those failed samples via Latent Dirichlet Allocation (LDA) (Blei et al., 2003)."

METHODOLOGY BERTopic: Neural topic modeling with a class-based TF-IDF procedure

"For Trump, it outperforms LDA on all measures whereas it only achieves the top score on topic coherence for the UN dataset."

METHODOLOGY GAN Inversion: A Survey

"For example, we can formulate GAN inversion as decomposing signals into components (matrix factorization problems) and use non-linear Factor Analysis (FA) [208], Independent Component Analysis (ICA) [209], Latent Dirichlet Allocation (LDA) [210], [211] to decompose the network weight and to find interpretable directions of latent space."

METHODOLOGY A review of stochastic block models and extensions for graph clustering

"Blei et al. (2003), along with Wang et al. (2011) used variational inference, the general formulation of which is described in "Variational methods" section."

METHODOLOGY Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda

"We performed Latent Dirichlet Allocation (LDA) based topic modeling [16] along with co-occurrence"

Contribution 2

Claim — Contribution 2

The researcher established a foundational theoretical framework and algorithm for spectral clustering, providing rigorous analysis that has become a seminal reference in the field.

The researcher's contribution centers on the seminal 2001 paper 'On spectral clustering: Analysis and an algorithm,' which serves as the core of this line of work. This publication appears to provide both a theoretical analysis and a practical algorithmic approach to spectral clustering, establishing a unified foundation for the method.

This work appears to address the need for rigorous theoretical grounding in spectral clustering techniques. By combining analysis with an algorithm, the researcher likely bridged the gap between heuristic applications and formal mathematical understanding, offering a robust framework that subsequent studies could rely upon.

The significance of this contribution is evidenced by its extensive citation record, with over 14,000 citations indicating widespread adoption. Furthermore, the high proportion of independent citations suggests that the work has had a broad impact across the global research community, influencing scholars outside the researcher's immediate network and institution.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 214 · **17 flagged influential by Semantic Scholar**

CORE PAPER

On spectral clustering: Analysis and an algorithm

2001 · Advances in neural information processing systems 14, 2001 · 14,207 citations (GS)

Field-normalised: 10,664 Semantic Scholar citations place it in the top 1% of Mathematics papers from 2001 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Deep Clustering: A Comprehensive Survey	Lehigh University, University of Electronic Science and Technology of China, University of Illinois at Chicago	China, United States	—
2	A Comprehensive Survey on Transfer Learning	Baidu Inc., University of Chinese Academy of Sciences	China	Methodology
3	Artificial intelligence for literature reviews: opportunities and challenges	The Open University	United Kingdom	—
4	Data Mining: The Textbook	IBM T. J. Watson Research Center, IBM T.J. Watson Research Center	United States	Background
5	Unmasking Clever Hans predictors and assessing what machines really learn	Fraunhofer Heinrich Hertz Institute, Singapore University of Technology and Design, Technische Universität Berlin	Germany, Singapore	Methodology
6	Multi-view evidential c-means clustering with view-weight and feature-weight learning	Hainan University, National Defence University, Universiti Sains Malaysia	China, Malaysia, Turkey	—
7	Feature Selection: A Data Perspective	Arizona State University, Michigan State University	United States	Methodology
8	Community detection in graphs	ISI Foundation	Italy	Methodology
9	Diffusion Improves Graph Learning	Technical University of Munich	Germany	Methodology
10	Community Detection and Stochastic Block Models: Recent Developments	Princeton University	United States	—
11	Consistency of spectral clustering in stochastic block models	Carnegie Mellon University	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
12	Social big data: Recent achievements and new challenges	Chung-Ang University, Universidad Autónoma de Madrid	South Korea, Spain	Methodology
13	Graph based anomaly detection and description: a survey	Stony Brook University, University of Illinois at Urbana-Champaign, University of Michigan	United States	Background
14	Quantum computing for finance	Argonne National Laboratory, Fujitsu Research of America, Inc., JPMorgan Chase	United States	—
15	Semi-Supervised Learning	LIFE Biosystems GmbH, Max Planck Institute for Intelligent Systems, Yahoo	Germany	—
16	Manifold Learning: What, How, and Why	ByteDance, University of Washington	United States	Background
17	Spectral Hashing	Massachusetts Institute of Technology, New York University, The Hebrew University of Jerusalem	Israel, United States	Methodology
18	Line Graph Neural Networks for Link Prediction	Texas A&M University, University of Science and Technology of China, University of Virginia	China, United States	—
19	Dimensionality Reduction: A Comparative Review	Maastricht University	Netherlands	Background
20	Classic GNNs are Strong Baselines: Reassessing GNNs for Node Classification	Beihang University, The Hong Kong Polytechnic University	China	Background
21	Machine Learning for Text: An Introduction	—	—	—
22	Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data	—	—	—
23	Data Clustering: Theory, Algorithms, and Applications	Hunan University, York University	Canada, China	—
24	Structural Deep Network Embedding	Tsinghua University	China	—
25	A fast, scalable and versatile tool for analysis of single-cell omics data	University of California, San Diego School of Medicine	United States	—
26	Structural Deep Clustering Network	Beijing University of Posts and Telecommunications, Tencent Ltd, Tsinghua University	China	Methodology
27	A Comprehensive Survey on Deep Clustering: Taxonomy, Challenges, and Future Directions	Macquarie University, Simon Fraser University, Tsinghua University	Australia, Canada, China	—
28	Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering	Iowa State University, University of Minnesota	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
29	Unsupervised machine learning in urban studies: A systematic review of applications	National University of Singapore	Singapore	Methodology
30	Generalized Latent Multi-View Subspace Clustering	Chinese Academy of Sciences, Inception Institute of Artificial Intelligence, Institute of Automation, Chinese Academy of Sciences	Australia, China, Singapore	Background

Showing the 30 most-cited of 214 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 A Comprehensive Survey on Transfer Learning

"3) SFA: In this step, a spectral clustering algorithm is adapted and performed to align domain-specific features [108], [109]."

METHODOLOGY Unmasking Clever Hans predictors and assessing what machines really learn

"The SC [53,147,148] used in step (3) is a clustering algorithm with interesting analytic properties, which constitutes the foundation of the analysis pipeline."

METHODOLOGY Feature Selection: A Data Perspective

"In the first step, spectral clustering (Chan et al., 1994; Ng et al., 2002) is applied on the dataset to detect the cluster structure of the data."

METHODOLOGY Community detection in graphs

"In the algorithm by Ng et al. (Ng et al., 2001) one adopts the normalized Laplacian Lsym (Section A.2)."

METHODOLOGY Diffusion Improves Graph Learning

"The degreecorrected stochastic block model (DCSBM) [30], spectral clustering (using Lsym) [55], DeepWalk [61], and Deep Graph Infomax (DGI) [73] are unsupervised models."

Contribution 3

Claim — Contribution 3

The researcher developed ROS, an open-source Robot Operating System, establishing a foundational software framework that has become a standard infrastructure for robotics research and development.

CLAIM: The researcher's primary contribution is the creation of ROS, an open-source Robot Operating System, as detailed in the seminal 2009 paper. This work stands as a singular, foundational achievement in the field, with no subsequent follow-up papers by the researcher listed in this specific line of work.

ORIGINALITY: The titles indicate that this work addressed the need for a standardized, open-source software infrastructure for robotics. By providing a unified operating system, the researcher appears to have solved critical interoperability and development challenges, enabling a more cohesive and accessible environment for building complex robotic systems.

SIGNIFICANCE: The impact of this contribution is evidenced by 14,759 citations, marking it as a highly influential work. Furthermore, analysis of 1,107 citing papers reveals that 96.9% originate from independent researchers, demonstrating that the robotics community widely adopted ROS as a standard tool beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 367 · **65 flagged influential by Semantic Scholar**

CORE PAPER

ROS: an open-source Robot Operating System

2009 · ICRA workshop on open source software 3 (3.2), 5, 2009 · 14,759 citations (GS)

Field-normalised: 10,833 Semantic Scholar citations place it in the top 1% of Engineering papers from 2009 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Neuromorphic computing at scale	Google DeepMind, Indian Institute of Science, Intel Labs	China, Germany, India	—
2	A survey on deep learning-based monocular spacecraft pose estimation: Current state, limitations and prospects	University of Luxembourg	Luxembourg	—
3	Symbiotic human-robot collaborative assembly	Case Western Reserve University, Hungarian Academy of Sciences, KTH Royal Institute of Technology	Germany, Hungary, Sweden	—
4	Secure Robotics: Navigating Challenges at the Nexus of Safety, Trust, and Cybersecurity in Cyber-Physical Systems	Bluerydge, University of Canberra	Australia	—
5	Survey of Human–Robot Collaboration in Industrial Settings: Awareness, Intelligence, and Compliance	Rochester Institute of Technology	United States	Background
6	A guide for construction practitioners to integrate robotic systems in their construction applications	New York University Abu Dhabi	United Arab Emirates	—
7	An all-round AI-Chemist with a scientific mind	University of Science and Technology of China	China	**Methodology**
8	Visual dexterity: In-hand reorientation of novel and complex object shapes.	Massachusetts Institute of Technology, Meta AI, Tsinghua University	China, United States	—
9	Application of Automated Guided Vehicles in Smart Automated Warehouse Systems: A Survey	Beijing University of Chemical Technology	China	—
10	Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping	Massachusetts Institute of Technology, Stevens Institute of Technology	United States	**Methodology**
11	Orbit: A unified simulation framework for interactive robot learning environments	Carnegie Mellon University, ETH Zurich, Georgia Institute of Technology	Canada, Switzerland, United States	**Methodology**
12	Dexterous helical magnetic robot for improved endovascular access	Cantonal Hospital Aarau, ETH Zurich, University Hospital Zurich	Switzerland	**Methodology**
13	Tumtraf v2x cooperative perception dataset	Technical University of Munich	Germany	—
14	RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation	Université de Sherbrooke	Canada	—
15	Teaser: Fast and certifiable point cloud registration	Massachusetts Institute of Technology	United States	**Methodology**

No.	Citing paper	Citing institution(s)	Country	S2
16	DS-SLAM: A semantic visual SLAM towards dynamic environments	Beihang University, Carnegie Mellon University, Tsinghua University	China, United States	Methodology
17	Analogcoder: Analog circuit design via training-free code generation	Chinese University of Hong Kong, The University of Hong Kong, The University of Texas at Austin	China, Hong Kong, United States	Methodology
18	Scientific exploration of challenging planetary analog environments with a team of legged robots	ETH Zurich, ETH Zurich; Harbin Institute of Technology, ETH Zurich; University of Bern	Switzerland, Switzerland; China	Methodology
19	Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours	Carnegie Mellon University	United States	—
20	Mm-fi: Multi-modal non-intrusive 4d human dataset for versatile wireless sensing	Nanyang Technological University, University of California, Irvine Medical Center, University of Edinburgh	Singapore, United Kingdom, United States	Methodology
21	Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for mapless navigation	City University of Hong Kong; the HKUST, ETH Zurich, HKUST	Hong Kong, Switzerland	Background
22	Core challenges of social robot navigation: A survey	Carnegie Mellon University, Honda Research Institute, Honda Research Institute and California Institute of Technology	United States	—
23	Homerobot: Open-vocabulary mobile manipulation	Carnegie Mellon, Carnegie Mellon University, Georgia Tech	Canada, United States	Methodology
24	Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017	Centre de Développement des Technologies Avancées	Algeria	Background
25	Milestones in autonomous driving and intelligent vehicles—Part I: Control, computing system design, communication, HD map, testing, and human behaviors	Beihang University, Chinese Academy of Sciences, Cranfield University	China, Hong Kong, Singapore	Background
26	AutoTrack: Towards high-performance visual tracking for UAV with automatic spatio-temporal regularization	National University of Singapore, New York University, Tongji University	China, Singapore, United States	Methodology
27	Can an Embodied Agent Find Your "Cat-shaped Mug"? LLM-Based Zero-Shot Object Navigation	University of Maryland, University of Maryland at College Park	United States	—
28	Robots that can adapt like animals	Inria, Sorbonne Universités, University of Wyoming	France, United States	—
29	Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems	Massachusetts Institute of Technology, U.S. Army Combat Capabilities Development Command	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
30	Agilicious: Open-source and open-hardware agile quadrotor for vision-based flight	UZH	Switzerland	—

Showing the 30 most-cited of 367 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 An all-round AI-Chemist with a scientific mind

"Control system software based on a Robot Operating System (ROS) [16] is developed to control the mobile robot with a six-degree-of-freedom robotic arm to move around and operate all apparatuses."

METHODOLOGY Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping

"All the methods are implemented in C++ and executed on a laptop equipped with an Intel i7-10710U CPU using the robot operating system (ROS) [24] in Ubuntu Linux."

METHODOLOGY Orbit: A unified simulation framework for interactive robot learning environments

"The modular nature of the agent makes it easy to switch between different control architectures for each task while using the same interface for the real robot. b) Using ROS: A variety of existing robots come with their ROS software stack."

METHODOLOGY Teaser: Fast and certifiable point cloud registration

"Furthermore, we provide a ROS [117] wrapper to enable easy integration and deployment in real-time robotics applications."

METHODOLOGY DS-SLAM: A semantic visual SLAM towards dynamic environments

"The system is also integrated with Robot Operating System (ROS), and its performance is verified by testing DS-SLAM on a physical robot in the real environment."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	50
University of California, Irvine Medical Center	United States	—	41
Stanford University	United States	SCImago #18 · THE =5 · QS 3	35
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	29
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	28
Nanyang Technological University	Singapore	SCImago #137	27
Princeton University	United States	SCImago #386 · THE =3 · QS =25	21
University of Washington	United States	SCImago #45 · THE 25 · QS 81	20
ETH Zurich	Switzerland	THE 11 · QS 7	20
Georgia Institute of Technology	United States	SCImago #270 · THE =41 · QS =123	19
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	17
Columbia University	United States	SCImago #65 · THE 20 · QS =38	17
UC Berkeley	United States	—	17
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	17
Harvard University	United States	SCImago #4 · THE =5 · QS 5	16

Geographic distribution of citing authors

Country	Citing papers
United States	549
China	201
United Kingdom	97
Germany	89
Australia	62
Canada	62
Singapore	51
Switzerland	46
Italy	40
Netherlands	35
France	29
Hong Kong	27

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	Latent Dirichlet Allocation	490	Dhanasar — Prong 2 (well-positioned)
Contribution 2	On spectral clustering: Analysis and an algorithm	214	Dhanasar — Prong 2 (well-positioned)
Contribution 3	ROS: an open-source Robot Operating System	367	Dhanasar — Prong 2 (well-positioned)