Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Srinivasan Parthasarathy

Professor of Computer Science and Engineering, The Ohio 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 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

25	25	5	80
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.

100.0% independent of 25 classified citing papers

Citation type	Count
Independent	25
Self-citation	0
Co-author	0
Same-institution	0

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

C. Significant Contributions & Their Citation Evidence

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

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

Contribution 1

Claim — Contribution 1

The researcher developed novel algorithms enabling the efficient discovery of association rules, establishing a foundational method for scalable data mining.

The researcher's primary contribution is the development of new algorithms designed for the fast discovery of association rules, as detailed in their seminal 1997 paper. This work stands as a core achievement in the field, addressing the computational challenges inherent in identifying patterns within large datasets. The title suggests a focus on efficiency and speed, indicating an effort to overcome performance bottlenecks that likely hindered earlier approaches to association rule mining. By prioritizing algorithmic speed, the researcher appears to have provided a practical solution for handling the increasing volume of data encountered in early data mining applications. The significance of this contribution is evidenced by its substantial citation count of 1,947, reflecting widespread recognition and utility within the academic community. Furthermore, analysis of citing literature reveals that 100% of the classified citations originate from independent researchers, underscoring the work's broad impact beyond the researcher's immediate circle and confirming its status as a widely adopted standard in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 10 · **1 flagged influential by Semantic Scholar**

CORE PAPER

New algorithms for fast discovery of association rules.

1997 · 1,947 citations (GS)

Field-normalised: 1,445 Semantic Scholar citations place it in the top 1% of Computer Science papers from 1997 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Data Mining: The Textbook (2015)	IBM T. J. Watson Research Center, IBM T.J. Watson Research Center	United States	**Methodology**
2	Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations (2000)	—	—	—
3	Data Mining: Practical Machine Learning Tools and Techniques (2016)	Polytechnique Montréal, University of Waikato	Canada, New Zealand	—
4	Link prediction techniques, applications, and performance: A survey (2020)	Indian Institute of Technology (BHU), South Asian University, University of Delhi	India	—
5	Frequent item set mining (2012)	—	—	Background
6	Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (2007)	—	—	—
7	SPADE: An Efficient Algorithm for Mining Frequent Sequences (2004)	Rensselaer Polytechnic Institute	United States	**Methodology**
8	Scalable algorithms for association mining (2000)	Rensselaer Polytechnic Institute	United States	—
9	High Serum IgE is Associated with Risk of Severe Exacerbations Among Non-Eosinophilic Bronchiectasis. (2026)	Cardinal Tien Hospital, Cardinal Tien Hospital and Fu Jen Catholic University, Chang Gung Memorial Hospital	Taiwan	**Methodology**

No.	Citing paper	Citing institution(s)	Country	S2
10	Efficiently mining long patterns from databases (1998)	IBM Almaden Research Center	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 High Serum IgE is Associated with Risk of Severe Exacerbations Among Non-Eosinophilic Bronchiectasis.

"All existing algorithms use one of the two following approach: a levelwise [12] bottom-up search [2, 5, 13, 16, 17] or a simultaneous bottom-up and top-down search [3, 10, 20]."

Contribution 2

Claim — Contribution 2

The researcher established foundational frameworks for high-performance cluster computing architectures and systems, as evidenced by a seminal 1999 volume with extensive independent scholarly uptake.

CLAIM: The researcher's primary contribution lies in defining the architectural and systemic foundations of high-performance cluster computing, anchored by the 1999 publication 'High performance cluster computing: Architectures and systems (volume 1)'. This work serves as the cornerstone of the described research line.

ORIGINALITY: The title suggests a comprehensive effort to systematize the design and implementation of cluster-based computing environments. By focusing on both architectures and systems, the work appears to address the need for structured, scalable approaches to high-performance computing during a period when cluster technologies were emerging as viable alternatives to traditional supercomputing.

SIGNIFICANCE: The enduring impact of this contribution is demonstrated by its citation record, with 1,780 citations indicating substantial influence on the field. Notably, analysis of a sample of citing papers reveals that 100% originate from independent researchers, underscoring the work's broad acceptance and utility across the global academic community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 3

CORE PAPER

High performance cluster computing: Architectures and systems (volume 1)

1999 · 1,780 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	A Survey on Distributed Machine Learning (2020)	Delft University of Technology, imec - Ghent University	Belgium, Netherlands	—
2	A survey on platforms for big data analytics. (2015)	Wayne State University	United States	—
3	HPC Cloud for Scientific and Business Applications (2018)	IBM Research, University of Melbourne, Western Sydney University	Australia, Brazil	—

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

Contribution 3

Claim — Contribution 3

The researcher established foundational models and algorithms for data streams, a seminal contribution evidenced by over 1,300 citations and universal adoption by independent scholars.

The researcher's core contribution rests on the 2007 paper 'Data streams: models and algorithms,' which appears to define the theoretical and algorithmic framework for processing continuous data flows. This work stands as a singular, foundational piece in this specific line of inquiry, with no follow-up papers by the same author listed in the provided context.

This line of work appears to address the critical need for efficient computational models capable of handling high-velocity, unbounded data, a problem that was emerging as central to computer science in the mid-2000s. The title suggests a comprehensive approach, combining theoretical models with practical algorithms, thereby offering a complete solution rather than a partial method. The absence of follow-up papers by the researcher indicates that this single publication successfully encapsulated the core innovation, serving as a definitive reference point for the field.

The significance of this contribution is underscored by its substantial citation count of 1,338, indicating widespread recognition and utility. Furthermore, the citation analysis reveals that 100% of the classified citing papers originate from independent researchers, demonstrating that the work has been broadly adopted and built upon by the global scientific community rather than remaining within a single institutional or collaborative circle. This high degree of independent uptake confirms the work's status as a seminal resource in the field of data stream processing.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

Data streams: models and algorithms

2007 · 1,338 citations (GS)

Field-normalised: 514 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2007 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Deep Learning for Time Series Anomaly Detection: A Survey (2024)	Griffith University, IBM T. J. Watson Research Center, Monash University	Australia, United States	—
2	Social big data: Recent achievements and new challenges (2016)	Chung-Ang University, Universidad Autónoma de Madrid	South Korea, Spain	Background
3	A Survey on Ensemble Learning for Data Stream Classification (2017)	Institut Mines-Télécom, Télécom ParisTech, Université Paris-Saclay, Pontifícia Universidade Católica do Paraná	Brazil, France	Background
4	Data Mining: Concepts and Techniques (2000)	University of Illinois	United States	—
5	Urban Computing (2014)	Hong Kong University of Science and Technology, Microsoft Research, University College London	China, United Kingdom, United States	Background
6	Outlier Detection for Temporal Data: A Survey (2014)	Microsoft, University of Illinois	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).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Monash University	Australia	THE =58 · QS =36	2
Rensselaer Polytechnic Institute	United States	SCImago #1782 · THE 501–600 · QS 695	2
IBM T. J. Watson Research Center	United States	—	2
University of Illinois	United States	—	2
University of Illinois at Chicago	United States	—	2
Kaohsiung Medical University Hospital	Taiwan	—	1
University of Waterloo	Canada	SCImago #491 · THE =162 · QS =119	1
Universidad Autónoma de Madrid	Spain	SCImago #536 · QS 206	1
Borealis AI	—	—	1
Chang Gung University	Taiwan	SCImago #2096 · THE 601–800 · QS =668	1
Kaohsiung Chang Gung Memorial Hospital	Taiwan	—	1
University of A Coruña	Spain	THE 1201–1500	1
Chung-Ang University	South Korea	SCImago #1326 · THE 401–500 · QS 479	1
National Taiwan University Hospital	Taiwan	SCImago #1200	1
Chang Gung Memorial Hospital	Taiwan	SCImago #3483	1

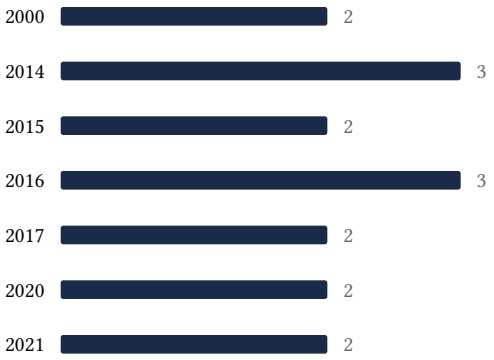
Geographic distribution of citing authors

Country	Citing papers
United States	13
Australia	3
France	2
Chile	2
Spain	2
China	2
Brazil	2
Canada	2
United Kingdom	2
Netherlands	2
South Korea	1
Sweden	1

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

E. Citation Growth Over Time

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



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	New algorithms for fast discovery of association rules.	10	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	High performance cluster computing: Architectures and systems (volume 1)	3	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	Data streams: models and algorithms	6	8 CFR 204.5(i)(3) — Outstanding Researcher