CURRICULUM VITAE

- **Personal Details**

Name	Dan Vilenchik
Date and place of birth	3/9/1977, Israel
E-mail	vilenchi@bgu.ac.il
Address	School of Electrical and Computer Engineering Ben-Gurion University of the Negev Beer-Sheva, Israel, 8410501
Work phone number	08-6428067
ORCID ID	0000-0002-3895-688X
Web page	www.bgu.ac.il/~vilenchi

- **Education**

B.Sc.	Technion – Israel Institute of Technology Computer Science (cum laude)	1998–2001
M.Sc.	The Weizmann Institute of Science Advised by Prof. Uriel Feige, Computer Science	2001–2004
Ph.D.	Tel-Aviv University Adivsed by Prof. Michael Krivelevich, Computer Science	2004–2009

- **Employment History**

Senior Lecturer 2020–now
School of Electrical and Computer Engineering
Ben-Gurion University of the Negev, Israel

Machine Learning Expert (Sabbatical) 2023-2025
AI Department, Playtika ltd.
Herzlya, Israel

Lecturer 2014–2020
Department of Communication Systems Engineering
Ben-Gurion University of the Negev, Israel

Postdoctoral Fellow 2011–2014
Department of Computer Science
The Weizmann Institute of Science, Israel

Hedrick Assistant Professor 2009–2011
Department of Mathematics
University of California, Los Angeles (UCLA), USA

Postdoctoral Fellow 2008–2009
Department of Computer Science
University of California, Berkeley , USA

- **Professional Activities**

(a) Positions in academic administration

Steering Committee Member, Data Science Research Center 2022–now
Ben-Gurion University, Israel

Head, Computer Engineering Track (Undergraduate and Graduate Programs) 2022–2023
Ben-Gurion University, Israel

Member, Undergraduate Teaching Committee 2021–2023
Ben-Gurion University, Israel

Member, School's Seminar Organizing Committee 2019–2021
Ben-Gurion University, Israel

Coordinator for Physics and Math Courses 2014–2019
Ben-Gurion University, Israel

Organizer, Departmental Seminar 2014–2015
Ben-Gurion University, Israel

(b) Professional functions outside universities/institutions (inter-university, national, international)

Member, COST Action eVoiceNet - CA24128 2026–2029
European Network to Advance the Development and Implementation of Vocal Biomarkers

Associate Editor 2021–now
Journal of Artificial Intelligence Research (JAIR) (Q2, IF 3.635)

Associate Editor 2022–now
Mathematics (Q1, IF 2.59)

Committee Member, "Aloni Postdoctoral Fellowship" 2023
Israeli Ministry of Science and Technology

Member, COST Action HiTec - CA21163 2022–2025
European Cooperation in Science and Technology

Guest Editor 2021–2023
Special Issue in Mathematics: Multidisciplinary Models and Applications of Machine Learning and Computational Statistics

Member, COST Action CRoNoS - IC1408 2017
European Cooperation in Science and Technology

(c) Significant professional consulting

Chief Data Officer 2017–2023
SOOS

AI Expert 2025–Present
Playtika

(d) Ad-hoc Reviewer for Journals and Conferences

Senior Program Committee Member 2023
European Conference on Artificial Intelligence (ECAI)

Reviewer Various Years
Journal of the ACM

Reviewer Various Years
SIAM Journal on Computing

Reviewer Various Years
SIAM Journal on Discrete Mathematics

Reviewer Various Years
IEEE Transactions on Computers

Reviewer Various Years
IEEE Transactions on Information Theory

Reviewer Various Years
Theory of Computing

Reviewer Various Years
Random Structures and Algorithms

Reviewer Various Years
Entropy

Reviewer Various Years
Proceedings of the ACM Symposium on Theory of Computing (STOC)

Reviewer Various Years
IEEE Symposium on Foundations of Computer Science (FOCS)

Reviewer Various Years
International Colloquium on Automata, Languages, and Programming (ICALP)

Reviewer Various Years
AAAI Conference on Artificial Intelligence

Reviewer International Joint Conference on Artificial Intelligence (IJCAI)	Various Years
Reviewer European Conference on Artificial Intelligence (ECAI)	Various Years
Reviewer International Conference on Web and Social Media (ICWSM)	Various Years
Reviewer Empirical Methods in Natural Language Processing (EMNLP)	Various Years
Reviewer European Chapter of the Association for Computational Linguistics (EACL)	Various Years

- **Educational activities**

(a) Courses taught

Linear Algebra	Undergraduate	UCLA	2009-2011
Calculus of Several Variables	Undergraduate	UCLA	2009-2011
Introduction to Stochastic Processes	Undergraduate	BGU	2019
Data Structures	Undergraduate	BGU	various years
Introduction to Computer Science in Python	Undergraduate	BGU	2017–now
Data Mining and Statistical Inference	Graduate	BGU	2014–2023
Causal Inference in the era of AI	Graduate	BGU	2025

(b) Research students

M.Sc. Students

Barak Yichye	Graduated 2016
Haviv Hershcovits (co-advised with Prof. Kobi Gal)	Graduated 2017
Maor Abutbul	Graduated 2018
Itamar Elmakias	Graduated 2018
Michael Sidorov	Graduated 2021
Sheli Kohan (co-advised with Prof. Carmel Sofer)	Graduated 2021
Liel Olesya Sharon (co-advised with Dr. Eran Ben-Elia)	Graduated 2022
Janna Vilkin	Graduated 2022
Idan Weiss	Graduated 2022
Khalil Wattad	Graduated 2022

Neta Bar-gil	Graduated 2022
Tal Ilan	Graduated 2022
Vladyslav Kozukhov	Graduated 2022
Hadar Cohen	Graduated 2023
Yossi Solomon	Graduated 2024
Benjamin Berend (co-advised with Prof. Aryeh Kontorovich)	Graduated 2025
Gal Shubeli (co-advised with Dr. Tal Golan)	Expected 2026
Omri Rafaeli	Expected 2026
Dolev Shaked (co-advised with Prof. Carmel Sofer)	Expected 2026
Tuvia Hausdorff	Expected 2026
Moshe Braunstein (co-advised with Prof. Ilana Niski)	Expected 2027
Josh Lorch	Expected 2027
Rafi Michaeli	Expected 2027
Shaked David Vaknin	Expected 2027
Zohar Zamir	Expected 2027

Ph.D. Students

Elad Shoham. *PhD Title: "Using the combinatorial structures of the graph-based inputs to understand the decision-making process of neural networks."*	Expected 2026
Itamar Elmakias. *PhD Title: "Characterization of high-dimensional datasets in light of the curse of dimensionality and feature selection algorithms."*	Expected 2027
Chanel Michaeli. *PhD Title: "Modeling the Affective Dimension in Language: From Generation to Subjective Language Understanding."*	Expected 2029
Ofer Feinstein. *PhD Title: "Using Semantic Dimensions and Causal Simulations to Understand Social Structures and Phenomena for Policy Design."* (co-advised with Prof. Michal Grinstein-Weiss and Prof. Eitan Bachmat)	Expected 2029
Or Maimon. *PhD Title: "Voice as a Marker of Emotional Distress in Women with Thyroid Dysfunction and Diabetes."* (co-advised with Prof. Julie Cwikel)	Expected 2030

- **Awards, Citations, Honors, Fellowships**

(a) Honors, Citation Awards

B.Sc. in Computer Science, cum laude Technion – Israel Institute of Technology	2001

Dean's Excellence Award for Ph.D. Students Tel-Aviv University	2008
Ben-Gurion University Excellence in Research Award (Top 18% Researchers) Ben-Gurion University	2026

(b) Fellowships

Fellow Social Policy Institute, Washington University in St. Louis	2022–present
Postdoctoral Grant, $18,000 Israel Ministry of Absorption, Returning Scientist Program	2011–2013
E.R. Hedrick Assistant Professorship, $48,000/year Department of Mathematics, UCLA	2009–2011

- **Scientific Publications**
  - H-index: ISI 14, GS 20.
  - Total number of citations: ISI 383 (368 without self-citations), GS 1,085.
  - The $*$ sign marks papers published after the beginning of the current review period.The superscripts on co-author names indicate career stage where S=Student, PD=Post Doc., C=Co-Researcher, CA=Corresponding Author (when available), and PI=Principal Investigator. When available, citations are taken from Google Scholar (GS) and SCOPUS (ISI).

(c) Refereed chapters in collective volumes, conference proceedings

1. M. Krivelevich[PI] and **D. Vilenchik**[S], "Semi-random models as benchmarks for coloring algorithms," in *Proceedings of the Third Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, SIAM, 2006, pp. 211–221 (citations ISI 14/GS 26)
2. M. Krivelevich[PI] and **D. Vilenchik**[S], "Solving random satisfiable 3cnf formulas in expected polynomial time," in *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2006, pp. 454–463 (citations ISI 26/GS 49; rank A*)
3. U. Feige[PI], E. Mossel[PD], and **D. Vilenchik**[S], "Complete convergence of message passing algorithms for some satisfiability problems," in *Proceedings of RANDOM 2006, Lecture Notes in Computer Science*, vol. 4410, Springer, 2006, pp. 339–350 (citations ISI 19/GS 58; rank A)
4. A. Coja-Oghlan[PD], M. Krivelevich[PI], and **D. Vilenchik**[S,CA], "Why almost all k-colorable graphs are easy to color," in *Proceedings of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS), Lecture Notes in Computer Science*, vol. 4393, Springer, 2007, pp. 121–132 (citations ISI 4/GS 11; rank A)
5. S. Ben-Shimon[S] and **D. Vilenchik**[S], "Message passing for the coloring problem: Gallager meets alon and kahale," in *Proceedings of the 13th International Conference on Analysis of Algorithms*, DMTCS, 2007, pp. 217–226 (citations ISI 0/GS 1; rank C)
6. A. Coja-Oghlan[PD], M. Krivelevich[PI], and **D. Vilenchik**[S], "Why almost all satisfiable k-cnf formulas are easy," in *Proceedings of the 13th International Conference on Analysis of Algorithms*, DMTCS, 2007, pp. 89–102 (citations ISI 0/GS 35; rank C)

7. A. Coja-Oghlan[PD], U. Feige[PI], A. Frieze[PI], M. Krivelevich[PI], and **D. Vilenchik**[S], "On smoothed k-cnf formulas and the walksat algorithm," in *Proceedings of the 20th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2009, pp. 451–460 (citations ISI 13/GS 23; rank A*)
8. T. Friedrich[PD], T. Sauerwald[PD], and **D. Vilenchik**[PD], "Smoothed analysis of balancing networks," in *ICALP 2009, Lecture Notes in Computer Science*, vol. 5556, Springer, 2009, pp. 472–483 (citations ISI 0/GS 15; rank A)
9. L. Minder[PD] and **D. Vilenchik**[PD], "Small clique detection and approximate nash equilibria," in *RANDOM 2009, Lecture Notes in Computer Science*, vol. 5687, Springer, 2009, pp. 673–685 (citations ISI 9/GS 23; rank A)
10. A. Sinclair[PI] and **D. Vilenchik**[PD], "Delaying satisfiability for random 2sat," in *RANDOM 2010, Lecture Notes in Computer Science*, vol. 6302, Springer, 2010, pp. 710–723 (citations ISI 2/GS 14; rank A)
11. V. Braverman[S], R. Ostrovsky[PI], and **D. Vilenchik**[PD], "How hard is counting triangles in the streaming model," in *ICALP 2013, Lecture Notes in Computer Science*, vol. 7965, Springer, 2013, pp. 244–254 (citations ISI 32/GS 73; rank A)
12. I. Levi[PI], **D. Vilenchik**[PD], M. Langberg[PI], and M. Effros[S], "Zero vs. epsilon error in interference channels," in *IEEE Information Theory Workshop (ITW)*, IEEE, 2013 (citations ISI 0/GS 2;)
13. A. Coja-Oghlan[PI] and **D. Vilenchik**[PD], "Chasing the k-colorability threshold," in *Proceedings of the 54th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2013, pp. 380–389 (citations ISI 33/GS 78; rank A*)
14. V. Bapst[PD], A. Coja-Oghlan[PI], S. Hetterich[S], F. Rassmann, and **D. Vilenchik**[PD], "The condensation phase transition in random graph coloring," in *Proceedings of the 18th International Conference on Random Structures and Algorithms (RANDOM)*, Springer, 2014, pp. 449–464 (citations ISI 38/GS 73; rank A)
15. **D. Vilenchik**[PI], "The million tweets fallacy: Activity and feedback are uncorrelated," in *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*, 2018, pp. 688–691 (citations ISI 0/GS 7; rank A)
16. Y. Segal[S], **D. Vilenchik**[PI], and O. Hadar[PI], "Detecting and coloring anomalies in real cellular network using principal component analysis," in *The 2nd International Symposium on Cyber Security Cryptography and Machine Learning (CSCML)*, 2018 (citations ISI 0/GS 0;)

    Contribution: Conceptualization, Methodology, Formal analysis.
17. **D. Vilenchik**[PI], B. Yichye[S], and M. Abutbul[S], "To interpret or not to interpret pca? this is the question," in *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2019, pp. 655–658 (citations ISI 0/GS 10; rank A)

∗ 18. G. Holtzman[S], A. Soffer[S], and **D. Vilenchik**[PI], "A greedy anytime algorithm for sparse pca," in *Proceedings of the 33rd Annual Conference on Learning Theory (COLT)*, 2020 (citations ISI 0/GS 22; rank A*)

∗ 19. R. Korenblum[S], V. Kozhukhov[S], **D. Vilenchik**[PI], and O. Tsur[PI], "Stem: Unsupervised structural embedding for stance detection," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022 (citations ISI 7/GS 17; rank A*)

    Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 20. R. Fairstein[S], R. Meir[PI], **D. Vilenchik**[PI], and K. Gal[PI], "Welfare vs. representation in participatory budgeting," in *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 2022 (citations ISI 0/GS 21; rank A*)

Contribution: Formal analysis, Writing.

∗ 21. K. Danilchenko[S], M. Segal[C], and **D. Vilenchik**[PI], "Opinion spam detection: A new approach using machine learning and network-based algorithms," in *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2022 (citations ISI 0/GS 16; rank A)

∗ 22. T. Ilan[S] and **D. Vilenchik**[PI], "Harald: Augmenting hate speech data sets with real data," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP) + Workshop on NLP4PositiveImpact*, 2022 (citations ISI 0/GS 6; rank A*)

∗ 23. Y. Salomon[S], E. Vax[C], A. Osowizky[C], Y. Knafo[C], N. Ben David[C], and **D. Vilenchik**[PI], "Direction of arrival estimation for radionuclides based on neural network approach," in *Proceedings of Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA)*, 2024 (citations ISI 0/GS 1)

∗ 24. G. Barel[S], O. Tsur[PI], and **D. Vilenchik**[PI], "Acquired taste: Multimodal stance detection with textual and structural embeddings," in *Proceedings of the 30th International Conference on Computational Linguistics (COLING)*, 2025 (citations ISI 0/GS 2; rank B)

Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 25. J. Barrett[S], K. Gal[PI], L. Michael[PI], and **D. Vilenchik**[PI], "Beyond the echo chamber: Modelling open-mindedness in citizens assemblies," in *Proceedings of the 2025 International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2025 (citations ISI 0/GS 0; rank A*)

Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 26. E. Shoham[S], O. Haber[S], H. Rika[C], and **D. Vilenchik**[PI], "Learning to rank: How gnns solve max-clique and sparse pca," in *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI)*, Oral Presentation, to apear., 2026 (citations ISI 0/GS 0; rank A*)

(d) Refereed articles and refereed letters in scientific journals, running numbers

1. **D. Vilenchik**[S], "It's all about the support: A new perspective on the satisfiability problem," *Journal on Satisfiability, Boolean Modeling, and Computation*, vol. 3, pp. 125–139, 2007 (citations ISI 0 / GS 5)

2. J. Böttcher[S] and **D. Vilenchik**[S], "On the tractability of coloring semirandom graphs," *Information Processing Letters*, vol. 108, no. 3, pp. 143–149, 2008 (citations ISI 1 / GS 3; IF 0.7;Q3)

3. A. Coja-Oghlan[PD], E. Mossel[PD], and **D. Vilenchik**[S], "A spectral approach to analyzing belief propagation for 3-coloring," *Combinatorics, Probability and Computing*, vol. 18, no. 6, pp. 881–912, 2009 (citations ISI 20 / GS 37; IF 0.608; Q2)

4. M. Krivelevich[PI], B. Sudakov[PI], and **D. Vilenchik**[S], "On the random satisfiable process," *Combinatorics, Probability and Computing*, vol. 18, pp. 775–801, 2009 (citations ISI 8 / GS 10; IF 0.608; Q2)

5. U. Feige[PI], A. Flaxman[PD], and **D. Vilenchik**[S], "On the diameter of the set of satisfying assignments in random satisfiable k-cnf formulas," *SIAM Journal on Discrete Mathematics*, pp. 736–749, 2011 (citations ISI 4 / GS 6; IF 0.662; Q3)

6. A. Coja-Oghlan[PD], M. Krivelevich[PI], and **D. Vilenchik**[S], "Why almost all k-colorable graphs are easy to color," *Theory of Computing*, vol. 46, no. 3, pp. 523–565, 2010 (citations ISI 4 / GS 11; Q3)

7. T. Friedrich[PD], T. Sauerwald[PD], and **D. Vilenchik**[PD], "Smoothed analysis of balancing networks," *Random Structures & Algorithms*, vol. 39, no. 1, pp. 115–138, 2011 (citations ISI 0 / GS 15; IF 1.048; Q1)

8. C. Lee[S], B. Sudakov[PI], and **D. Vilenchik**[PD], "Getting a directed hamilton cycle two times faster," *Combinatorics, Probability and Computing*, vol. 21, pp. 773–801, 2012 (citations ISI 3 / GS 6; IF 0.608; Q2)

9. U. Feige[PI], E. Mossel[PI], and **D. Vilenchik**[PD], "Complete convergence of message passing algorithms for some satisfiability problems," *Theory of Computing*, vol. 9, no. 19, pp. 617–651, 2013 (citations ISI 19 / GS 58; Q2)

10. A. Sinclair[PI] and **D. Vilenchik**[PD], "Delaying satisfiability for random 2sat," *Random Structures & Algorithms*, vol. 43, no. 2, pp. 251–263, 2013 (citations ISI 2 / GS 14; IF 1.048; Q1)

11. I. Pak[PI] and **D. Vilenchik**[PD], "Constructing uniquely realizable graphs," *Discrete & Computational Geometry*, vol. 50, no. 4, pp. 1051–1071, 2013 (citations ISI 4 / GS 10; IF 0.649; Q2)

12. R. Krauthgamer[PI], B. Nadler[PI], and **D. Vilenchik**[PI], "Do semidefinite relaxations solve sparse pca up to the information limit?" *Annals of Statistics*, vol. 4, no. 3, pp. 1300–1322, 2015 (citations ISI 40 / GS 91; IF 2.78; Q1)

    Contribution: Methodology, Formal analysis, Software, Writing.

13. A. Coja-Oghlan[PI] and **D. Vilenchik**[PI], "The chromatic number of random graphs for most average degrees," *International Mathematics Research Notices*, vol. 2016, no. 19, pp. 5801–5859, 2016 (citations ISI 16 / GS 31; IF 1.031; Q1)

    Contribution: Conceptualization, Methodology, Formal analysis, Writing.

14. V. Bapst[PD], A. Coja-Oghlan[PI,CA], S. Hetterich[S], F. Rassmann, and **D. Vilenchik**[PI], "The condensation phase transition in random graph coloring," *Communications in Mathematical Physics*, vol. 341, no. 2, pp. 543–606, 2016 (citations ISI 38 / GS 73; IF 2.375; Q1)

    Contribution: Conceptualization, Methodology, Formal analysis.

15. S. Ilic[PD], T. Akabayov[S], R. Froimovici[S], S. Meiry[S], **D. Vilenchik**[PI], A. Hernandez[PD], H. Arnathani[PI], and B. Akabayov[PI,CA], "Modulation of the activity in the mn-substituted t7 dna primase has a structural origin," *Scientific Reports*, vol. 7, p. 5797, 2017 (citations ISI 2 / GS 4; IF 5.228; Q1)

    Contribution: Formal analysis, Software, Writing.

16. M. Langberg[PI,CA] and **D. Vilenchik**[PI], "Constructing cospectral graphs via a new form of graph product," *Linear and Multilinear Algebra*, pp. 1–15, 2017 (citations ISI 5 / GS 7; IF 1.0; Q1)

    Contribution: Conceptualization, Methodology, Formal analysis, Writing.

17. C. Sofer[PD], **D. Vilenchik**[PI], R. Dotsch[PI], and G. Avidan[PI], "Emotion algebra reveals the richness of meanings of facial expressions," *Journal of Vision*, vol. 18, no. 10, p. 193, 2018 (citations ISI 0 / GS 0; Q2)

    Contribution: Methodology, Formal analysis, Software, Writing.

18. Y. Neuman[PI,CA], N. Israeli, **D. Vilenchik**[PI], and Y. Cohen, "The adaptive behavior of a soccer team: An entropy-based analysis," *Entropy*, vol. 20, no. 10, p. 758, 2018 (citations ISI 30 / GS 38; IF 2.3; Q2)

    Contribution: Formal analysis, Software, Writing.

19. **D. Vilenchik**[PI], "Simple statistics are sometimes too simple: A case study in social media data," *IEEE Transactions on Knowledge and Data Engineering*, 2019. DOI: 10.1109/TKDE.2019.2899355 (citations ISI 4 / GS 9; IF 2.775; Q1)

20. Y. Neuman[PI,CA] and **D. Vilenchik**[PI], "Modeling small systems through the relative entropy lattice," *IEEE Access*, 2019 (citations ISI NA / GS 16; IF 3.56; Q1)

    Contribution: Methodology, Formal analysis, Software, Writing.

21. B. Tam[S], D. Sherf[S], S. Cohen[S], S. Eisdorfer[S], M. Peretz[S], A. Soffer[S], **D. Vilenchik**[PI], T. Akabayov[S], G. Wagner[PI], and B. Akabayov[PI], "Discovery of small-molecule inhibitors targeting the ribosomal peptidyl transferase center (ptc) of m. tuberculosis," *Chemical Science*, 2019 (citations ISI 12 / GS 20; IF 9.55; Q1)

    Contribution: Methodology, Formal analysis, Software, Writing.

22. H. Hershcovits[S], **D. Vilenchik**[PI], and K. Gal[PI], "Modeling engagement in self-directed learning systems using principal component analysis," *IEEE Transactions on Learning Technologies*, 2019 (citations ISI 13 / GS 24; IF 2.3; Q1)

    Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 23. Y. Neuman[PI,CA], **D. Vilenchik**[PI], and N. Israeli[C], "From physical to social interactions: The relative entropy model," *Scientific Reports*, 2020 (citations ISI 3 / GS 8; IF 4.01; Q1)

    Contribution: Methodology, Formal analysis, Writing.

∗ 24. I. Elmakiass[S] and **D. Vilenchik**[PI], "An oblivious approach to machine translation quality estimation," *Mathematics*, 2021 (citations ISI 6 / GS 11; Q1)

∗ 25. A. Soffers[S], S. Eisdorfer[S], M. Ifrach[S], S. Ilic[PD], A. Afek[PI], H. Schussheim[S], **D. Vilenchik**[PI], and B. Akabayov[PI], "Inferring primase-dna specific recognition using a data driven approach," *Nucleic Acids Research*, 2021 (citations ISI 3 / GS 4; IF 16.6; Q1)

    Contribution: Methodology, Formal analysis.

∗ 26. R. Portnikhs[S], K. Kenzis[S], N. Krohns[S], **D. Vilenchik**[PI], and I. Marienberg-Milikowsky[PI], "An experimental undogmatic modelling of (hebrew) literature," *Digital Journal for Philology, Special Issue on Digital Methods in Literary Studies*, 2022 (citations ISI 0 / GS 2)

    Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 27. C. Sofer[PI,CA], R. Dotsch[C], **D. Vilenchik**[PI], and G. Avidan[C], "The compositionality of facial expressions," *Perception*, vol. 51, no. 3, pp. 172–186, 2022. DOI: 10.1177/03010066221077573 (citations ISI 1 / GS 1; IF 1.6; Q3)

    Contribution: Methodology, Formal analysis, Writing.

∗ 28. I. Weiss[S] and **D. Vilenchik**[PI,CA], "Predicting churn in online games by quantifying diversity of engagement," *Big Data*, vol. 11, no. 4, pp. 282–295, 2023. DOI: 10.1089/big.2022.0109 (citations ISI 3 / GS 8; Q1)

∗ 29. C. Cordeiro[C], N. Gonceer[C], S. Dorus[C], J. Crill[C], V. Moshayoff[C], A. Lachman[S], A. Moran[S], **D. Vilenchik**[C], and S. Fedida-Metula[PI], "Fast, accurate, and cost-effective poultry sex genotyping using real-time polymerase chain reaction," *Frontiers in Veterinary Science*, 2023 (citations ISI 3 / GS 3; IF 2.6; Q1)

∗ 30. Y. Salomons[S], E. Vax[C], Y. Knafo[C], N. Ben-David[C], A. Osovizky[C], and **D. Vilenchik**[PI], "Direction of arrival estimation for radionuclides based on neural network approach," *IEEE Transactions on Nuclear Science*, vol. 71, no. 5, pp. 1124–1133, 2024. DOI: 10.1109/TNS.2024.3384011 (citations ISI 0 / GS 0; IF 1.9; Q1)

∗ 31. A. Soffer[S], S. Viswas[S], S. Alon[S], N. Rozenberg[S], A. Peled[S], D. Piro[S], **D. Vilenchik**[C], and B. Akabayov[PI], "Moloptimizer: A molecular optimization toolkit for fragment-based drug design," *Molecules*, 2024 (citations ISI 1 / GS 2; IF 4.2; Q2)

∗ 32. J. Barrett[S], K. Gal[PI], L. Michael[C], and **D. Vilenchik**[PI], "Beyond the echo chamber: Modelling open-mindedness in citizens' assemblies," *Autonomous Agents and Multi-Agent Systems, Special Issue on Citizen-Centric AI Systems*, 2024 (citations ISI 0 / GS 2; IF 2.0; Q3)

Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 33. **D. Vilenchik**[PI,CA], J. Cwikel[PI], Y. Ezra[C], T. Hausdorff[S], M. Lazarov[S], R. Sergienko[C], R. Abramovitz[S], I. Schmidt[C], and A. Perez[S], "Method matters: Enhancing voice-based depression detection with a new data collection framework," *Depression & Anxiety*, 2025 (citations ISI 1 / GS 2; IF 3.3; Q1)

Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 34. E. Shoham[S,CA], H. Cohen[S], K. Wattad[S], H. Rika[C], and **D. Vilenchik**[PI,CA], "Concept learning for algorithmic reasoning: Insights from SAT-solving GNNs," *Information Sciences*, vol. 726, p. 122 754, 2026, ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2025.122754 (citations ISI 0 / GS 0; IF 6.8; Q1)

∗ 35. O. Rafaeli[S], I. Marienberg-Milikowsky[PI], and **D. Vilenchik**[PI,CA], "Mind the gap: Word-embedding and multi-layered literary networks," *Digital Scholarship in the Humanities*, 2025. DOI: 10.1093/llc/fqaf112 (citations ISI 0 / GS 0; IF 1.1; Q1)

Contribution: Conceptualization, Methodology, Formal analysis, Writing.

∗ 36. T. M. Schwartz Tayri[PI,CA], N. Cohen-Inger[S], O. Seadia[C], A. Gal[S], and **D. Vilenchik**[PI], "A case-specific psychological first aid ai recommendations app for rescue and first responders," *European Journal of Psychotraumatology*, vol. 16, no. 1, p. 2 591 567, 2025. DOI: 10.1080/20008066.2025.2591567 (citations ISI 0 / GS 0; IF 4.1; Q1)

Contribution: Methodology, Writing.

∗ 37. I. El-Makias[S] and **D. Vilenchik**[PI,CA], "Choosing the right dataset: Hardness criteria for feature selection benchmarking," *Knowledge-Based Systems*, p. 115 022, 2025, ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2025.115022 (citations ISI 0 / GS 0; IF 7.6; Q1)

(e) Unrefereed professional articles and publications

∗ 1. T. Schwartz Tayri[PI], N. Cohen Inger[PI], and **D. Vilenchik**[PI], "Chapter 23: Data is power: The use of big data and artificial intelligence in welfare services in israel," in *Management of Welfare Services in Israel- Theory, research and practice*, H. Shmid and Y. Sabah, Eds., In Hebrew, Resling, 2025, pp. 607–634

∗ 2. Y. Neuman[PI], M. Danesi[PI], and **D. Vilenchik**[PI], *Using AI for Dialoguing with Texts: From Psychology to Cinema and Literature*. Routledge, 2022 (citations ISI 0 / GS 15)

∗ 3. **D. Vilenchik**[PI], "Biological approaches to mathematics," in *Handbook of Cognitive Mathematics*, M. Danesi, Ed., Springer Cham, 2021, pp. 413–537

∗ 4. **D. Vilenchik**[PI], "Characterizing users of online social and e-learning platforms," in *Mathematics (Education) in the Information Age*, S. Costa, M. Danesi, and D. Martinovic, Eds., Springer, 2020

5. U. Feige[PI] and **D. Vilenchik**[S], "A local search algorithm for 3sat," Computer Science and Applied Mathematics, The Weizmann Institute of Science, Tech. Rep. MCS 04-07, 2004 (citations ISI 0 / GS 25)

• **Lectures and Presentations at Meetings and Invited Seminars**

(a) Invited plenary lectures at conferences/meetings - (running numbers)

1. The "Rector's Podium" Lecture, Ben-Gurion University 2026
Title: Using Voice to Detect Depression: Challenges and Breakthroughs

2. Industry–Academia Relations Club, Faculty of Engineering, Ben-Gurion University 2026
Title: Engineering the Future with Artificial Intelligence
Conference Host and Chair; Moderator of Talks and Industry Q&A Panel

(b) Presentation of papers at conferences/meetings

1. The 16th ACM-SIAM Symposium on Discrete Algorithms (SODA), Miami 2006
Title: Solving random satisfiable 3CNF formulas in expected polynomial time

2. The 24th International Symposium on Theoretical Aspects of Computer Science (STACS), Aachen 2007
Title: Why almost all k-colorable graphs are easy to color

3. The 13th International Conference on Analysis of Algorithms, Antibes 2007
Title: Message passing for the coloring problem: Gallager meets Alon and Kahale

4. The 13th International Conference on Analysis of Algorithms, Antibes 2007
Title: Why almost all satisfiable k-CNF formulas are easy

5. Warwick-Weizmann Workshop, University of Warwick, UK 2011
Title: Constructing Uniquely Realizable Graphs

6. Workshop on Probability and Graphs, TU Eindhoven 2014
   Title: Chasing the k-colorabilty threshold

7. International Colloquium on Automata, Languages and Programming (ICALP), Riga 2013
   Title: How hard is counting triangles in the streaming model

8. IEEE Information Theory Workshop (ITW), Sevilla 2013
   Title: Zero vs. e Error in interference channels

9. Israeli Statistical Association Annual Conference, Ramat Efa'l 2016
   Title: User Classification in Twitter via Sparse PCA

10. The 12th International AAAI Conference on Web and Social Media (ICWSM), Stanford 2018
    Title: The Million Tweets Fallacy: Activity and Feedback are Uncorrelated

11. The 13th International AAAI Conference on Web and Social Media, Munich 2019
    Title: To interpret or not to interpret PCA? This is the question

12. AI Week, Tel Aviv, Israel 2019
    Title: Generalizing the Million Followers Fallacy

13. 33rd Annual Conference on Learning Theory (COLT), (online) 2020
    Title: A Greedy Anytime Algorithm for Sparse PCA

14. Bernoulli-IMS One World Symposium 2020
    Title: A Greedy Anytime Algorithm for Sparse PCA

15. The Second International Israel Data Science Initiative Conference (IDSI) 2023
    Title: Two new approaches for data augmentation

16. The 16th Bar-Ilan Symposium on Foundations of Artificial Intelligence (BISFAI) 2023
    Title: The challenges of feature selection in high-dimensional data

17. The 2nd Annual Data for Health Conference, Ben-Gurion University 2023
    Title: Should we be scared of high-dimensional data? A new look on the 'curse of dimensionality'

18. Cognitive Approaches to Mathematics Education Today, Fields Institute, Toronto (online) 2024
    Title: A Quantitative Approach for Measuring the Engagement of Students in e-learning Platforms

19. Book Launch Event, Israel's President's Residence (attended by the President of Israel) 2025
    Title: Data is Power: The Use of Big Data and Artificial Intelligence in Welfare Services in Israel

20. The 40th AAAI Conference on Artificial Intelligence (AAAI-26), Singapore 2026
Title: Learning to Rank: How GNNs Solve Max-Clique and Sparse PCA

(c) Seminar presentations at universities and institutions

1. Seminar on Algorithms for the SAT Problem, Humboldt-Universität zu Berlin 2006
Title: Message passing algorithms for the SAT problem

2. Dagstuhl Seminar on Probabilistic Methods in the Design and Analysis of Algorithms 2007
Title: On the tractability of the k-colorabilty problem

3. Computer Science, Georgia Tech, ARC Colloquium 2008
Title: On satisfiable k-CNF formulas above the threshold

4. Mathematics, UCLA, Math Dept. Colloquium 2009
Title: Recent results in average case analysis of the satisfiability problem

5. Statistics, Hebrew University, Departmental Seminar 2013
Title: Sparse Principal Component Analysis: Dealing with high-dimensional data

6. Phase Transitions in Discrete Structures, Math Department Colloquium, Goethe University Frankfurt 2016
Title: The denoising problem, or, Statistics and Computer Science, water and oil?!

7. Computer Science, Tel-Aviv University, Machine Learning Seminar 2019
Title: Polynomial time algorithms are not always welcome – sparse PCA as a case study

8. Statistics, Haifa University, Departmental Seminar 2019
Title: Polynomial time algorithms are not always welcome – sparse PCA as a case study

9. Computer Science, Bar-Ilan University, Departmental Seminar 2019
Title: Polynomial time algorithms are not always welcome – sparse PCA as a case study

10. Electrical Engineering, Technion, Machine Learning Seminar 2019
Title: Polynomial time algorithms are not always welcome – sparse PCA as a case study

11. Statistics, Tel-Aviv University, Departmental Seminar 2020
Title: Computational statistical tradeoffs in the problem of finding sparse Principal Components in high-dimensional data

12. KnowDive Seminar, University of Trento (online) 2020
Title: An unsupervised approach to characterizing users in online social platforms

13. Inference Problems: Algorithms and Lower Bounds, Goethe University Frankfurt (online) 2020
Title: A Greedy anytime algorithm for sparse PCA

14. Workshop on Computational Methods in the Humanities, Lausanne 2022
Title: Highlighted Gaps: Toward Undogmatic Modeling of Literary Character Networks

15. University of Edinburgh ILCC/CDT NLP Seminar Series 2022
Title: From theory to practice and back – Stance detection, a case study

16. University of Edinburgh Statistics Seminar 2022
Title: The Denoising Problem in the Lens of Statistics and Computer Science

17. Google Tel Aviv, NLP Seminar 2022
Title: Two new approaches for data augmentation

18. Computer Science, The Hebrew University, Learning Club Seminar 2022
Title: From theory to practice and back – Stance Detection as a case study

19. Computer Science, Bar-Ilan University, NLP Group Seminar 2022
Title: From theory to practice and back – Stance Detection as a case study

20. Computer Data Science Seminar, Technion 2022
Title: From theory to practice and back – Stance Detection as a case study

21. IBM Tel Aviv, Language Journal Club 2023
Title: Two new approaches for data augmentation

22. University of Edinburgh ILCC/CDT NLP Seminar Series (online) 2023
Title: Two new approaches for data augmentation

23. Statistics Departmental Seminar, Tel-Aviv University 2025
Title: Beyond the Curse – Redefining Feature Selection Benchmarks in High-Dimensional Data

- **Grants**

Competitive Research Grants listed by VATAT (2.45M NIS)

Personal Grant — ISF (Israel Science Foundation) Research Grant No. 2469/25 2025–2028
Subject: Voice Biomarkers in Mental Health: Can AI Detect When Depression and Anxiety Are Faked?
Grantees: **Dan Vilenchik**[PI],Julie Cwikel (Social Work)[PI]
Total amount: 705K NIS

MOST (Ministry of Science and Technology) 2025–2028
Subject: Investigating the Impact of the First AI Chatbot on Delivering Psychological First Aid and Mitigating Acute Stress
Grantees: Talya Schwatz-Tayri (Social Work)[PI], **Dan Vilenchik**[PI]
Total amount: 500K NIS

MOST (Ministry of Science and Technology) 2022–2025
Subject: Decision-making Support System for Early Identification of Non-accidental Burns in Children Using Machine Learning
Grantees: **Dan Vilenchik**[PI], Talya Schwatz-Tayri (Social Work)[PI]
Total amount: 500K NIS

MOST (Ministry of Science and Technology) 2022–2025
Subject: Highlighted Gaps: Toward Undogmatic Modeling of Literary Character Networks
Grantees: **Dan Vilenchik**[PI], Itay Marienberg-Milikowsky (Literature)[PI]
Total amount: 400K NIS

Personal Grant — ISF (Israel Science Foundation) Research Grant No. 1388/16 2016–2020
Subject: New Directions in Average Case Complexity of NP-Hard Problems
Grantees: **Dan Vilenchik**[PI],Uriel Feige[PI]
Total amount: 340K NIS

Competitive Research Grants

The Leona M. and Harry B. Helmsley Charitable Trust 2022–2025
Subject: COBMINDEX Goes Forward: Smartphone Application for Psychological Wellbeing in Crohn's Disease
Grantees: Schwartz D.[PI], Odes S.[C], Sarid O[C], Slonim-Nevo V.[C], Monsonogo A.[C], Greenberg D.[C], **Dan Vilenchik**[C], Friger M.[C]
Total amount: 2.77M USD

Kamin (Israel Innovation Authority) 2019–2021
Subject: Research and Development of Anti-Deepfake
Grantees: **Dan Vilenchik**[PI], Carmel Sofer[PI]
Total amount: 860K NIS

MAGNET (Israel Innovation Authority) 2016–2018
Subject: Big-data Multimedia Metadata Enrichment
Grantees: Ofer Hadar[PI], **Dan Vilenchik**[C]
Total amount: 530K NIS

Other Grants

BGU Interdisciplinary Program (Faculty of Humanities and Social Sciences) 2025
Subject: Beyond the Text – Citizen Science, Computational Modeling and Literary Experience
Grantees: **Dan Vilenchik**[PI], Itay Marienberg-Milikowsky[PI]
Total amount: 25K NIS

BGU–Clalit Joint Research Fund 2022–2025
Subject: Pilot Study of Early Mental Health Detection of Depression and Anxiety in Primary Care Clinics
Grantees: **Dan Vilenchik**[PI], Julie Cwikel[PI], Aya Biderman[C], Temira Feinsilver[C]
Total amount: 100K NIS

BGU FOR 2023–2024
Subject: EnvironMENTAL – Unveiling the Matrix of Multilevel Risk and Resilience through AI and 30-Year Linked Data
Grantees: Schwartz-Tayri[PI], **Dan Vilenchik**[PI], Avner Ronen[C], Tarin Paz-Kagan[C], Yaron Ziv[C]
Total amount: 150K NIS

Data Science Research Center, BGU 2023–2024
Subject: The BGU Multidisciplinary Human Resilience Project – Advancing Wellness and Resilience Research Using Data Science
Grantees: Schwartz-Tayri[PI], **Dan Vilenchik**[PI], Nir Grinberg[PI]
Total amount: 88K NIS

BGU Interdisciplinary Program (Faculty of Humanities and Social Sciences) 2022–2023
Subject: Preventing Recurrence of Abuse: Identifying Abusive Burns Using Artificial Intelligence
Grantees: **Dan Vilenchik**[PI], Talya Schwatz-Tayri[PI], Michal Maimon [C], Yuval Krieger[C], Yaron Shoham[C], Ricardo Nachman[C]
Total amount: 30K NIS

Duet Center Grants (BGU) 2022–2023
Subject: Identifying Abusive Injuries Using Artificial Intelligence
Grantees: **Dan Vilenchik**[PI], Talya Schwatz-Tayri[PI]
Total amount: 10K NIS

BGU President Special Task Force on COVID-19 2020–2021
Subject: Assessment of Voice Markers of Depression and Other Clinical States Among Medical Staff and Patients During the 2020 Crisis
Grantees: **Dan Vilenchik**[PI], Julie Cwikel[PI]
Total amount: 15K NIS

BGU VP R&D Office 2020–2021
Subject: AmosOz2Vec: Hebrew Literature Meets Word2Vec
Grantees: **Dan Vilenchik**[PI], Itay Marienberg-Milikowsky[PI]
Total amount: 25K NIS

BGU Smart Transportation Center 2018–2019
Subject: Geospatial Big Data Fusion of Transit Performance and its Travelers' Satisfaction
Grantees: **Dan Vilenchik**[PI], Eran Ben-Elia[PI]
Total amount: 40K NIS

NR Soos Technology (Private Sector) 2018–2019
Subject: The Feasibility of Sex Reassignment Treatment in Chicken Embryos
Grantees: **Dan Vilenchik**[PI]
Total amount: 65K NIS

BGU Interdisciplinary Program (Faculty of Humanities and Social Sciences) 2016–2017
Subject: Natural vs. Constructed Languages: A Linguistic and Computational Investigation of the Uniqueness of Natural Languages
Grantees: **Dan Vilenchik**[PI], Ariel Cohen[PI]
Total amount: 60K NIS

- **Synopsis of Research**

Many natural and social phenomena are not random but structured—they follow patterns and constraints. In social networks, for example, people connect with similar others; while individual links are stochastic, the overall structure reveals clear communities. The same holds for language, facial expressions, genetic sequences, and images—all shaped by internal regularities that give rise to low-dimensional, interpretable, and computationally useful representations.

My research explores the data properties that generate such low-dimensional structures and designs algorithms to uncover and interpret them efficiently. Depending on the question and domain, my work ranges from classical algorithmic analysis to machine-learning models, bridging theory and empirics across social systems, vision, and language.

I begin with the multidisciplinary dimension of this research—a reflection not only of academic curiosity but of a personal journey that has shaped my scientific outlook.

Interdisciplinary Applications. My interdisciplinary work has personal roots. When my mother was diagnosed with schizophrenia (when I was a young kid), I learned to detect subtle changes in her voice—signs that her medication or mood had shifted—often before anyone else noticed. This early sensitivity to vocal cues, which helped prevent crises, later became the foundation of my scientific interest in understanding voice as a window into mental health.

In collaboration with Prof. Julie Cwikel from the Social Work department at Ben-Gurion University, we began collecting voice data from clinical populations and building machine learning pipelines capable of detecting mental health conditions from acoustic features of the human voice, focusing on depression to start. We recently published a paper in a Q1 psychology journal, [58], underscoring the potential of our unique approach to data collection and data analysis.

We are deeply committed to translating this work into practice: to develop a non-invasive, objective biomarker for mental health based on voice, a tool that could transform screening and early detection. Yet, before such technology can be reliably applied in real-world settings, it is essential to address a core foundational question that has never been systematically studied: to what extent is voice truly a biomarker of mental health, rather than a surface-level correlate?

This scientific challenge is at the heart of our ongoing work and was recently recognized with a competitive Israeli Science Foundation (ISF) grant. In this project, we will rigorously test whether

the proposed voice-based biomarker is robust to manipulation—that is, whether individuals can intentionally alter their vocal patterns to fake depression.

While the ISF project anchors the core scientific inquiry, we recognized that real-world impact requires tools beyond academia. This insight led to the launch of **Voxwell**, our initiative to translate voice-based biomarkers into practice. Together with CEO Efrat Kaul-Granot, we joined the state-supported Resilience Accelerator in Sderot, working directly with affected communities and acquiring key entrepreneurial skills. Voxwell has since earned early recognition, winning a $10,000 prize in the Merage Institute Israel Startup Competition and ranking among the top three ventures. Building on this momentum, we are now preparing a pilot with the Civil Aviation Authority of Israel (CAA) to test how voice technology can help reduce human-factor errors in aviation.

A core principle of my research is that scientific advances should, when possible, translate into real-world impact. This motivation has led to several multidisciplinary collaborations. Early in my career, I applied machine learning to drug design [40], [46], [50], when such methods were still emerging. More recently, I have worked on a project to detect burns in non-speaking children, developing tools to help clinicians distinguish accidental from suspicious injuries—funded by the Ministry of Science and Technology (MOST) grant *"Decision-making Support System for Early Identification of Non-accidental Burns in Children Using Machine Learning"*.

Modeling Human Interaction and Behavior. Understanding structured behavior in social systems—both large-scale (e.g., social media) and small-scale (e.g., soccer teams)—has been a recurring theme in my research. In [15], [44], we have laid the conceptual and empirical foundation for a new paradigm we term "semantic dimension analysis". Traditional dimensionality-reduction techniques such as PCA operate purely in algebraic space, decorrelating directions of maximal variance. We extend this view to the *semantic domain*, showing that principal components can correspond to distinct and interpretable semantic axes (in terms of feature loading significance). For example, in social-network data, one principal component may capture patterns of activity (posting frequency, engagement), while another captures feedback (likes, comments).

When these components are not only algebraically orthogonal but also semantically uncorrelated, we describe this phenomenon as *semantic shattering*. Semantic shattering indicates that the system's latent factors (e.g., activity vs. feedback) vary independently, revealing domain-level structure. Empirically, we found that such shattering is prominent in generic social networks (e.g., Instagram), but absent in specialized or interest-focused communities (e.g., photography or gaming platforms). Our framework thus provides a scientific language to describe a gut-feeling or widely recognized phenomenon: in generic networks, users often give feedback out of courtesy or platform norms rather than genuine engagement with content, whereas in specialized communities, feedback tends to reflect true appreciation of effort or quality, leading to correlated activity and feedback dimensions.

We are now extending this paradigm, in collaboration with my PhD student Ofer Feinstein, to large-scale social-mobility datasets provided by the Israeli Central Bureau of Statistics. There, semantic shattering across socio-economic indicators appears to signify a healthy separation of structural factors, offering actionable insights for policymakers in planning targeted interventions.

We now turn from the macro to the micro—shifting focus from large-scale societal patterns to individual and group behavior in smaller, structured systems. In [43], [45] we developed an information-theoretic framework to model team behavior in soccer, allowing us to quantify adaptability and predict season outcomes. In [48], we proposed a method to distinguish meaningful from trivial interactions using relative entropy. Our work on product-review graphs [21] leveraged structural anomalies to identify opinion spam efficiently. We also modeled deliberation in Citizens' Assemblies (CAs) as a multi-agent system [24], introducing formal criteria for success and demonstrating—using real data from the Scottish CA—that a latent open-mindedness variable is key to capturing opinion change.

This opens new directions for computational CA design.

In the realm of user engagement with online platforms such as learning or gaming, [47] introduced a theory-grounded framework for modeling engagement in math e-learning platforms, achieving performance on par with state-of-the-art ML models but with stronger interpretability. This was extended in [53], where we applied similar techniques to quantify diversity in online gaming behavior, grounded in the theory of "cognitive flow."

Building on this trajectory, I spent an 18-month sabbatical in Playtika's AI department—a fascinating experience that exposed me to the real-world scale, messiness, and practical challenges of behavioral data. There, I gained hands-on experience in coding and large-scale data analysis, enriching my experimental skill set and sharpening my ability to guide students through applied research. The sabbatical also opened a new direction for my work: causal inference, which I studied and practiced extensively at Playtika and am now ready to develop as a new line of research.

From Theory to Practice and Back. A unifying theme in my research is the interplay between theory and practice—using mathematical insights to understand real-world data and letting empirical findings inspire new theoretical questions.

The first example comes from Natural Language Processing (NLP). In [19], we modeled multi-speaker interactions as graphs and applied an SDP relaxation of Max-Cut to embed speakers in a stance space, achieving state-of-the-art unsupervised stance detection without using textual content. A follow-up study [25] integrated linguistic cues into this geometric framework, showing that textual valence forms a smooth layer over structural embeddings. Together, these works illustrate how concepts from rigidity and random graph theory can yield principled, interpretable alternatives to large-scale NLP models.

A similar theory–practice dialogue drives my work in high-dimensional and combinatorial learning. In our COLT paper [18], we studied the spiked covariance model of Johnstone and Lu and asked whether meaningful structure can be recovered under vanishing signal-to-noise ratio (SNR) without exhaustive search. We provided a rigorous affirmative answer via a new sparse PCA algorithm tailored to low-SNR regimes. Motivated by this, I turned to practice: together with my PhD student, Itamar Elmakais, we conducted a large-scale analysis of 102 feature-selection datasets [67]. Our study revealed that many widely used benchmarks are surprisingly easy, exposing a disconnect between the theoretical literature on high-dimensional data—dominated by the "curse of dimensionality"—and the practical datasets used to evaluate feature-selection methods.

Finally, with my PhD student Elad Shoham, I study how Graph Neural Networks (GNNs) learn to solve combinatorial optimization problems. In our work on NeuroSAT [65], a GNN trained to solve satisfiability, we demonstrated that the network implicitly learns classical algorithmic concepts, such as *support*, revealing how deep models internalize handcrafted combinatorial reasoning. We then extended this framework to *Max-Clique*, showing in [66] how the learned concept can be leveraged to modify the GNN decoder and improve performance. Building on these insights, we further developed a non-spectral, hyperparameter-free GNN-based algorithm for *Sparse PCA*, achieving performance competitive with the state-of-the-art covariance-thresholding methods of Bickel–Levina and Montanari–Deshpande.

- **Teaching Statement**

Teaching Philosophy

Teaching has always struck me as a deceptively complex task—how to transfer knowledge that feels intuitive and obvious to the teacher to someone encountering it for the first time. I draw inspiration from Bertrand Russell's *Unpopular Essays*, which advocate curiosity, engagement, and independent

thought as pillars of effective teaching. These ideals guide my classroom approach: never take things at face value, ask questions, think critically, and stay excited. Parenthood has also shaped my philosophy. I've seen how self-image influences a child's behavior, and I believe the same applies to students: when they perceive themselves as curious and capable, they begin to act accordingly.

My overarching goals go beyond mastering course content. I want students to build resilience in the face of difficulty, stay confident, and learn how to seek help effectively. In theoretical courses, such as algebra, probability, or algorithms, I emphasize real-world examples to make abstract concepts relevant. In hands-on courses like programming, I stress good practices and show how theoretical foundations support practical skills. I view teaching as performance—engaging, thoughtful, and paced for learning. Whether using the blackboard or slides, I design each lecture to keep students focused and motivated. Graduate courses offer an opportunity to treat students as junior researchers, exposing them to open problems and encouraging deeper engagement with the field. Across all levels, I aim to foster an appreciation of how diverse topics interconnect into a coherent whole.

Mindfulness and yoga are part of my daily routine, helping me remain attentive to the classroom environment and responsive to students' needs. I'll close with a favorite quote from Bertrand Russell's "Ten Commandments for Teachers": *Find more pleasure in intelligent dissent than in passive agreement, for, if you value intelligence as you should, the former implies a deeper agreement than the latter.* This spirit of thoughtful engagement is at the heart of my teaching.

Teaching Experience

My teaching journey began in middle school tutoring and continued through high school, military service, and throughout my academic career. Over the years, I've taught a wide range of courses, both theoretical and applied, across multiple institutions and student levels.

Even before my appointment as a PI at Ben-Gurion University, I maintained a strong and continuous involvement in teaching. During my PhD, I served as a teaching assistant at Tel-Aviv University and concurrently taught independent courses at the Open University of Israel, including Data Structures and Introduction to Algorithms. While on my postdoctoral fellowship in the U.S., I taught undergraduate courses in Linear Algebra and Calculus of Several Variables at UCLA. Upon returning to Israel and continuing my postdoc at the Weizmann Institute, I taught both undergraduate and graduate students at the Technion, giving courses such as Data Mining and Quantitative Methods in Finance for the Executive MBA program. These varied experiences reflect a longstanding dedication to teaching, well before my formal academic appointment.

At Ben-Gurion University, my teaching spans from foundational undergraduate courses like Data Structures and Introduction to Computer Science in Python, to more advanced undergraduate and graduate courses, including Introduction to Stochastic Processes, Data Mining and Statistical Inference, and Causal Inference. These experiences have deepened my ability to adapt material to varied levels of abstraction and professional relevance.

I put a premium on preparation—rehearsing each lecture thoroughly, even when I've taught the material before—so I can focus fully on student engagement and classroom dynamics. In theoretical courses, I prefer blackboard teaching to align the pace of delivery with students' comprehension. In applied or large-format settings, I use slide decks enhanced with quizzes, paradoxes, and illustrative examples to keep attention sharp in an age of distraction. I actively seek ways to improve my teaching: I review student evaluations, exchange ideas with colleagues, and explore new platforms such as Piazza to foster interaction. I also replaced the final exam in my graduate Machine Learning course with a Kaggle competition, which challenges students to apply course concepts in real-world, team-based settings—often giving them strong portfolio material for job applications. Whether in lecture halls or online platforms, I remain committed to creating intellectually rich, motivating, and respectful learning environments.