



jaan altosaar, phd

I design AI, culture, and tools to enhance agency at global scale, and teach others to do the same if they choose.

 jaan@onefact.org
 <https://jaan.io>

areas of specialization

Health Care • Systems Thinking • Artificial Intelligence • Open Source • Health Equity • Physics

languages

English (native) • Estonian (native) • French (fluent) • Spanish (working) • Mandarin (beginner)

immigration

Canadian and Estonian passport holder • 7 years on F1 • 3 months on CPT • 10 months on OPT • 4 months on EAD • Received NIW EB-2 green card • Now helping other immigrants get green cards and cap-exempt H-1B visas

professional experience

2022–present

Chief Executive Officer, [One Fact Foundation](#)

Developed and built a foundation whose mission is to change global health care using open source principles. In first six months: 10+ person team, acquired initial funding (\$55,000 grant from Columbia & Stanford), academic and industry partnerships, and received 501(c)(3) status determination from the Internal Revenue Service.

2013–present

Founder & Board Member, [Useful Science](#)

Built a non-profit science organization (200+ volunteers, 2M+ pageviews, 20k+ subscribers, 1M+ podcast downloads). "Won \$50,000" on the Canadian Dragon's Den.

2016

Research Internship, [Google Brain](#). Host: Eugene Brevdo

Contributed to variational inference support in TensorFlow; developed time series models.

2015

Research Internship, [DeepMind](#).

Collaborated with Andriy Mnih and Koray Kavukcuoglu in the deep learning group.

2013

UI and UX Designer, [Ottawa Hospital Research Institute](#)

Led design and testing of a federally-funded mobile app ([CANImmunize](#)) used to submit vaccination profiles to the government; now used for COVID vaccine tracking across Nova Scotia.

research experience

2020–2022

Officer of Research

Columbia University, Department of Computer Science

Columbia University Irving Medical Center, Vagelos College of Physicians and Surgeons

Developed machine learning and statistical methods for women's health, mental health, and health disparity. Advised graduate students, undergraduate students, and high school students in research, which resulted in several publications.

2018–2020

Visiting Researcher, Host: [Kyle Cranmer](#)

New York University, Center for Data Science & Department of Physics

Applied probabilistic modeling approaches to study statistical physical systems.

2014–2020

Graduate Research Fellowship, Advisors: [David Blei](#) & [Shivaji Sondhi](#)

Columbia University, Departments of Computer Science and Statistics

Princeton University, Department of Physics

Developed deep learning and variational inference methods with applications to recommender systems and physics.

education

2020 **Ph.D., Physics, Princeton University**. Advisors: David Blei and Shivaji Sondhi.

2015 **M.A., Physics, Princeton University**

2013 **B.Sc. First Class Honours in Mathematics and Physics, McGill University**

Top 10% cumulative GPA, Dean's Honour List, Dean's Multidisciplinary Undergraduate Research List.

consulting

2022-present [Morphood](#)

2020-2021 [The Browser](#)

2016-2019 IllumeMed (acquired by Skyscape Inc. in 2019)

honors, awards, & fellowships

2023 Carriage House Birth Partial Scholarship for Doula Training

2023 TEAK Fellowship National Mentoring Month Spotlight

2022 HmntyCntrd Scholarship to attend Humanity-Centered Design Masterclass

2021 Columbia scholarship to attend PI Crash Course: Skills for Future or New Lab Leaders workshop

2020 Princeton Physics Departmental Teaching Award

2014-2017 NSERC Doctoral Postgraduate Scholarship: ranked 3rd of 204 (\$63,000)

2014 Google Summer of Code grant to work at Columbia University

2013 Julie Payette NSERC Research Scholarship: awarded to the top 24 out of 1575 applicants in the Canada-wide Postgraduate Scholarships M competition (\$25,000)

2013 Commonwealth Scholarship, DPhil studies at University of Oxford (declined, £95,625)

2013 The Faculty of Science Moyses Travelling Scholarship, McGill University (\$10,000)

2013 Delta Upsilon Graduate Scholarship, McGill University (\$5,000)

2013 Travel award, KAUST WEP Conference

2012 First Prize for best poster, Canadian Undergraduate Physics Conference (Vancouver)

2012 Second Prize, McGill Faculty-wide Undergraduate Research Conference

2012 Third Prize, McGill Department of Physics Poster Conference

2010-2012 Estonian Foundation of Canada Scholarship

2009 Annette S. Hill McGill Scholarship

2008 Harry Elton Memorial Award, Embassy of the People's Republic of China in Canada

theses

2020 **Altosaar, J.** 2020. "Probabilistic Modeling of Structure in Science: Statistical Physics to Recommender Systems". Philosophiae doctor thesis. Princeton University

2012 **Altosaar, J.** 2012. "Detecting Methylation of Single Molecules of DNA". Honours research thesis. McGill University

journal papers

- 2021 **Altosaar, J.**, Tansey, W., and Ranganath, R. 2021. "RankFromSets: Scalable Set Recommendation with Optimal Recall". *Stat*
- 2015 Henelius, P., Lin, T., Enjalran, M., Hao, Z., Rau, J. G., **Altosaar, J.**, Flicker, F., Yavors'kii, T., and Gingras, M. J. P. 2015. "Refrustration and Competing Orders in the Prototypical Dy₂Ti₂O₇ Spin Ice Material". *Physical Review B*.
- Featured on *Phys. Rev. B*. [front page](#).

conference proceedings

- 2020 **Altosaar, J.**, Tansey, W., and Ranganath, R. 2020a. "RankFromSets: Scalable Set Recommendation with Optimal Recall". American Statistical Association, Symposium on Data Science and Statistics
- Huang, K., **Altosaar, J.**, and Ranganath, R. 2020b. "ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission". ACM Conference on Health, Inference, and Learning.
- Featured on [VentureBeat](#), [Towards Data Science](#), and included in [Apache MXNet](#).
- 2018 **Altosaar, J.**, Ranganath, R., and Blei, D. M. 2018a. "Proximity Variational Inference". AISTATS
- Dieng, A. B., Ranganath, R., **Altosaar, J.**, and Blei, D. M. 2018b. "Noisin: Unbiased Regularization for Recurrent Neural Networks". ICML
- 2016 Liang, D., **Altosaar, J.**, Charlin, L., and Blei, D. M. 2016a. "Factorization Meets the Item Embedding: Regularizing Matrix Factorization with Item Co-Occurrence". ACM RecSys
- Ranganath, R., **Altosaar, J.**, Tran, D., and Blei, D. M. 2016b. "Operator Variational Inference". NeurIPS
- 2015 Benjamin, E. and **Altosaar, J.** 2015a. "MusicMapper: Interactive 2D Representations of Music Samples for in-Browser Remixing and Exploration". International Conference on New Interfaces for Musical Expression.
- Featured and interviewed on [The Wire magazine](#).
- Mercer-Taylor, A. and **Altosaar, J.** 2015b. "Sonification of Fish Movement Using Pitch Mesh Pairs". International Conference on New Interfaces for Musical Expression
- Zhang, J., Gerow, A., **Altosaar, J.**, Evans, J., and So, R. J. 2015c. "Fast, Flexible Models for Discovering Topic Correlation across Weakly-Related Collections". EMNLP

refereed workshop, symposium, and short papers

- 2021 Stengel, A., **Altosaar, J.**, Dittrich, R., and Elhadad, N. 2021. "Assisted Living in the United States: An Open Dataset". Machine Learning for Public Health. NeurIPS
- 2020 Bansal, R., Olmstead, J., Bram, U., Cottrell, R., Reder, G., and **Altosaar, J.** 2020a. "Recommending Interesting Writing Using a Controllable, Explanation-Aware Visual Interface". Workshop on Interfaces and Human Decision Making for Recommender Systems, ACM Recommender Systems
- Reder, G. K., **Altosaar, J.**, Rajniak, J., Elhadad, N., and Fischbach, M. 2020b. "Supervised Topic Modeling for Predicting Chemical Substructure from Mass Spectrometry". Machine Learning for Molecules. NeurIPS
- 2019 **Altosaar, J.**, Ranganath, R., and Cranmer, K. 2019. "Hierarchical Variational Models for Statistical Physics". Machine Learning and the Physical Sciences. NeurIPS
- 2016 **Altosaar, J.**, Ranganath, R., and Blei, D. M. 2016a. "Proximity Variational Inference". Advances in Approximate Bayesian Inference. NeurIPS

- Bell, E. and **Altosaar, J.** 2016b. "Word Embedding Models Applied to Classical Music Recover the Circle of Fifths in Embedding Space." Music Discovery. ICML
- Bhatia, A., **Altosaar, J.**, and Gu, S. 2016c. "Proximity-Constrained Reinforcement Learning". Advances in Approximate Bayesian Inference. NeurIPS

preprints and technical reports

- 2022 Sun, T. Y., Bhave, S., **Altosaar, J.**, and Elhadad, N. 27, 2022. Assessing Phenotype Definitions for Algorithmic Fairness
- 2021 Ketenci, M., Adams, G., **Altosaar, J.**, Perotte, A., and Elhadad, N. 2021a. "Pre-Training Variational Inference for Cold-Start Recommendation". In preparation
- Reder, G. K., **Altosaar, J.**, Rajniak, J., Elhadad, N., and Fischbach, M. 2021b. "Predicting Molecular Structure from Tandem Mass Spectrometry". In preparation
- Reder, G. K., Young, A., **Altosaar, J.**, Rajniak, J., Elhadad, N., Fischbach, M., and Holmes, S. 19, 2021c. "Supervised Topic Modeling for Predicting Molecular Substructure from Mass Spectrometry". F1000Research
- 2020 Whitney, W. F., Song, M. J., Brandfonbrener, D., **Altosaar, J.**, and Cho, K. 2020. "Evaluating Representations by the Complexity of Learning Low-Loss Predictors"
- 2013 **Altosaar, J.** 2013. "The Resonant Recognition Model: Long-Range Protein Interaction via Transition Dipole Couplings". McGill Honours Research Project

technical writing

- 2017 J. Altosaar. [How does physics connect to machine learning?](#)
 Authored longform article that generated 30k pageviews with an average read time of 8 minutes.
- 2016 J. Altosaar. [Variational autoencoder tutorial.](#)
 Authored longform article that generated 400k pageviews with an average read time of 10 minutes. Used as a reference in courses at the University of Toronto and New York University.

teaching experience

- 2019–2020 **Assistantship in Instruction, Princeton University** PHY301: Thermal Physics.
- 2018–2020 **Assistantship in Instruction, Princeton University** PHY525: Introduction to Condensed Matter Physics.
- 2018 **Instructor, Summer Program on Applied Rationality and Cognition** (<https://sparc-camp.org/>)
 Taught machine learning and emotional intelligence to high schoolers. Rated easiest to connect with by students. Sample anonymous student feedback: "particularly easy to approach."
- Spring 2014 **Instructor, Princeton University Splash.** Taught high school students; average rating 4.38/5 teaching quality.
- Winter 2013 **Teaching Assistant, McGill University.** Applied Linear Algebra (Prof. Adam Oberman)
- Winter 2012 **Teaching Assistant, McGill University.** Honours Complex Variables (Prof. Robert Seiringer)
- Fall 2011 **Teacher, Montreal Estonian Society Kindergarten**
- Fall 2011 **Mentor, McGill University Buddy Program**

advising and mentorship

- Work with PhD, Master's, undergraduate, and high schoolers has resulted in several publications.
- 2021 Benjamin Guzovsky (Princeton University)
- 2021 Anton Stengel (Princeton University)
- 2021 Alexander Pesendorfer (Princeton University)
- 2020 [Gabe Reder](#) (Stanford University)

- 2020–2021 [Rohan Bansal](#) (Central High School '20, MO → Stanford University)
- 2017 [Abhishek Bhatia](#) (M.Sc. '18, Columbia University)
- 2016 [Eamonn Bell](#) (Ph.D. '18, Columbia University)
- 2015–2019 [Smiti Kaul](#) (Wake Forest University)
- 2014 [Ethan Benjamin](#) (M.Sc. '14, Columbia University)
- 2014 [Jingwei Zhang](#) (M.Sc. '14, Columbia)
- 2014 [Andrew James Mercer-Taylor](#) (B.Sc. '15, Columbia University)
- 2014 [Anjishnu Kumar](#) (M.Sc. '14, Columbia University)
- 2014 [Tony Paek](#) (M.Sc. '15, Columbia University)
- 2014 [Drishan Arora](#) (M.Sc. '14, Columbia University)

talks

- 2022 Google Health Bioethics Summit Panelist
- 2022 Music Tech Festival Labs Keynote
- 2022 NIH AIM-AHEAD Inaugural Conference, Invited Keynote: <https://bit.ly/onefact-keynote>
- 2022 AI For Health Equity Symposium, 6h Deep Learning Workshop
- 2022 Society for Digital Mental Health Flash Talk
- 2022 OpenMRS Technical Action Committee Invited Talk
- 2022 University of Maryland Colloquium
- 2021 Columbia University, Data Science Institute Scholars seminar series
- 2021 Andrew Marks Lab, Physiology and Cellular Biophysics Department, Columbia University
- 2021 Weight Watchers International, Inc. invited seminar to data science team
- 2021 Invitae invited talk for computational biology group
- 2021 Johnson & Johnson invited talk on ClinicalBERT for the Office of the Chief Medical Officer
- 2021 Panelist, New York University AI School
- 2020 Lena Mamykina lab seminar, Columbia University
- 2020 Probabilistic Modeling in Support of Science; invited talk. Caltech; University of California, Irvine; University of Southern California; Scripps Research Institute; University of Toronto, Vector Institute; Stanford University; University of Pennsylvania; MSKCC
- 2018 Food recommendation with deep exponential families. Keynote. North Star AI Conference, Estonia
- 2017 Bloomberg L.P. Machine Learning Group
- 2017 New York Times, Machine Learning & Cooking editorial teams
- 2017 Northeastern University, Network Science Institute seminar
- 2016 Imperial College, London, machine learning seminar.
- 2016 Machine Intelligence Research Institute Colloquium Series on Robust and Beneficial AI
- 2016 Columbia University, Comparing Domains of Improvisation seminar
- 2012 Canadian Undergraduate Physics Conference, University of British Columbia

service

- Reviewer Nature Biomedical Engineering; AMIA '22; JMLR; NeurIPS '16–'22; ICML '17, '19–'22; AAAI '18; ICLR '17–'22; AISTATS '18–'22; PLOS ONE '17–'22; Consciousness and Cognition '17; Advances in Approximate Bayesian Inference '15–'22; NeurIPS Machine Learning and the Physical Sciences Workshop '19–'20; NeurIPS Machine Learning for Health '20–'21; NeurIPS Algorithmic Fairness through the Lens of Causality and Interpretability '20; NeurIPS I Can't Believe It's Not Better '21; NeurIPS Bayesian Deep Learning '21

organizing

- 2022 Founded ai@columbia, and grew this community to 600+ faculty, postdocs, undergrads, and researchers with monthly happy hours sponsored by venture capital.
- 2021 Workshop on Motivational Interviewing with Dr. Prantik Saha, Columbia
- 2021 Workshop on Failure in Academia with Dr. Anna Womack
- 2021 ICML Workshop on Computational Biology, Organizing Committee

selected poster presentations

- 2021 Columbia University Data Science Day
- 2021 Columbia University Data Science Institute Health Analytics Center
- 2021 New York Academy of Sciences, AI for Chemical Biology
- 2017 New York Academy of Sciences, Proximity Variational Inference
- 2014 ComSciCon: Communicating Science, Harvard University: ranked top 50 of 870 applicants
- 2012 Canadian Undergraduate Physics Conference, University of British Columbia
First Prize for best poster
- 2012 Faculty of Science Undergraduate Research Conference, McGill University
Second Prize: induction to Sigma Xi Society
- 2012 Department of Physics Poster Conference, McGill University
Third Prize: travel award for Canadian Undergraduate Physics Conference
- 2011 Department of Physics Poster Conference, McGill University - Honourable Mention

professional associations

Sigma Xi Scientific Society, New York Academy of Sciences, Park Slope Food Coop

activities & interests

- 2020–present Mentor, [TEAK Fellowship](#)
- 1996–present Classical and jazz piano, electronic music production
- 2017 FIRST LEGO League regional robotics competition judge, Brooklyn, NY
- 2014–2015 Resident Graduate Student, Wilson College, Princeton University. Taught weekly meditation.
- 2014 Hopewell Elementary School science fair judge
- 2010–2014 Mentor, McGill University Mentorship Program for First-Year Students
- 2012 University of Waterloo Choir (Director: Professor Gerard Yun)
- 2011 Milton Park Recreation Association Beach Volleyball
- 2010 Montreal Estonian Society Kindergarten Teacher
- 2009–2010 Meditation (Enpuku-ji Zen Center, Abbess Zengetsu Myōkyō)
- 2009 McGill Choral Society (Director: Mary-Jane Puiu)

selected press

- 2022 [Liberty, Equity, Data](#) podcast interview
- 2021 [The Browser](#), podcast interview
- 2019 [VentureBeat](#), "AI predicts hospital readmission rates from clinical notes"
- 2016 Editorial, [The Conversation](#), "Accurate science or accessible science in the media - why not both?"
- 2016 Interview, [The Wire magazine](#)
- 2016 MusicMappr featured on [Prosthetic Knowledge blog](#)
- 2015 Featured on [Dragons' Den](#) episode, Canadian Broadcasting Corporation

- 2015 [In Training](#), "Medical Student Startup Improves Science Communication"
- 2014 [Reddit](#) front page
- 2014 [Boing Boing](#), "Useful Science, accessible by all"
- 2014 [Lifehacker](#), "Excel shortcuts, article summaries, and web notes"
- 2014 [Fitbit](#) corporate blog, "7 science-backed numbers to improve your life"
- 2014 [New Zealand Herald](#), "10 top sites to visit this weekend"
- 2014 [AweSci](#), "A chat with Jaan Altosaar from Useful Science"
- 2014 [IT World](#), "Useful Science headlines that apply to your weird little computer life"
- 2014 [McGill Tribune](#), "Useful Science bridges communication gap in research"
- 2014 [McGill News](#), Alumni Magazine, "Better living through science"
- 2014 [Betakit](#), "McGill grad launches curated list of science articles"
- 2014 National Canadian radio show, Spark [episode](#) features Useful Science