

Khoa D. Doan

🏠 khoadoan.me

✉ khoadoan106@gmail.com

ABOUT

I lead **MAIL-Research**, where we develop computational frameworks that enable the *reliability, trustworthiness, and practicality* of existing complex/deep models in *constrained* and especially *low-resource* settings. I focus on improving the following aspects of existing models: (i) training, (ii) inference, (iii) realistic assumptions, and (iv) trustworthiness. Most of my ML/AI solutions center around generative-based approaches that have low computational complexity and require less human effort. My research often translates into high-impact projects that benefit low-and-middle-income countries; these works have been featured in several national and international news outlets, such as **khoahocphattrien** (Vietnam's top scientific newspaper), **Thanh Nien** (Vietnam's top newspaper), **VnExpress** (Vietnam's top online newspaper), **DanTri**, **Yahoo Finance**, **Benzinga**, **Macau Business**, **Taiwan News**, and **TNGlobal**. Read more about my research interests **HERE**.

EDUCATION

Ph.D in Computer Science

Virginia Tech (VT)

Virginia, USA

August 2021

MS in Computer Science

University of Maryland, College Park (UMCP)

College Park, Maryland, USA

August 2010-May 2015

BS in Computer Science with a Minor in Mathematics

Webster University

Saint Louis, Missouri, USA

August 2003-May 2006

PROFESSIONAL EXPERIENCE

Associate Director – **VinUni-Illinois Smart Health Center**, Vietnams

2024-

Assistant Professor – Computer Science, **VinUniversity**, Vietnam

2022-

Research Scientist – **Baidu Research USA**, Seattle, WA

2020-2022

Senior Data Scientist – Verve Mobile/Carlsbad, CA

2015-2019

Faculty Research Associate – University of Maryland, College Park, MD

2012-2015

Research Scientist – NASA Goddard Space Flight Center, Greenbelt, MD

2012-2015

Senior Developer – Aquilent/Laurel, MD (acq. by Booz Allen Hamilton)

2008-2012

Software Developer – Insight Distribution Software/Portland, OR

2007-2008

Software Engineer – Conde Nast Publication/New York, NY

2006-2007

FUNDINGS AND SPONSORED PROGRAMS

- **Google Academic Research Award** 2024 (South/Southeast Asia), "Accelerate Poisoning and Backdoor Research with Large-scale Generative Models". **PI**. Amount: \$30K. *International*.
- **AWS Health Equity Initiative** Promotional Credit. **PI**, Amount: \$340K, Duration: 11/2024-06/2026. *International*.
- Green Industrial AI Challenge (Computing Credit). **Co-PI**, Amount: \$170K, Duration: 12/2024-12/2026. *National*.
- **Trinity Challenge/MIT Solve**, "Farm2Vet: Combatting AMR on the Farm Frontier", **PI**, Amount: \$1.3M, Duration: 11/2024-11/2027. *International*.

- VinUni-Illinois Smart Health Center, “Autonomous and Knowledge-guided Health Misinformation Defender Platform”, **PI**, Amount: over \$1M to support 2 PhD and 5 Master students at UIUC (~\$70K/Student-Year), and 5 PhD students VinUni (~\$10K/Student-Year). Duration: 02/2024-02/2029. *Cross-Institution*.
- VinUni-Illinois Smart Health Center, “From Causal Understanding of Healthy Longevity to Smart Health”, **PI**, Amount: over \$1M to support 2 PhD and 5 Master students at UIUC (~\$70K/Student-Year), and 5 PhD students VinUni (~\$10K/Student-Year), Duration: 02/2024-02/2029. *Cross-Institution*
- **Bill and Melinda Gates Foundation/Grand Challenges**, “Generative AI Technologies for Gynecological Healthcare in Vietnam”, **PI**, Amount: \$150,000, Duration: 07/2023-04/2024. *International*.
- Center for Environmental Intelligence at VinUniversity, “Improving NLP Applications in Low-resource Languages: One Country and One Use Case At A Time”, **Co-PI**, Amount: \$170,000, Duration: 11/2024-11/2026. *Internal*.
- Center for Environmental Intelligence at VinUniversity, “Carbon Stock Estimation and Biodiversity Assessment in Vietnam Forests using Remotely Sensed Data and Deep Learning Neural Networks”, **Co-PI**, Amount: \$210,000, Duration: 07/2023-07/2025. *Internal*.
- VinUniversity Seed Funding, “Evaluating and Defending against Advanced Backdoor Attacks on Deep Learning”, **PI**, Amount: \$75,000, Duration: 12/2023-12/2026. *Internal*.
- Smarthealth Center at VinUniversity, “Privacy-Preserving, Robust, and Explainable Federated Learning Framework for Healthcare System”, **Co-PI**, Amount: funding support 2 PhD and 5 Master students at UIUC (~\$70K/Student-Year), and 5 PhD students VinUni (~\$10K/Student-Year). Duration: 02/2022-02/2027. *Cross-Institution*.

TEACHING

- **Instructor** for “COMP4040 (Undergraduate)/COMP5110 (Graduate): Data Mining and Big Data Analysis”, VinUniversity, Fall (2024).
- **Instructor** for “COMP2050: Artificial Intelligence”, VinUniversity, Spring (2023, 2024, 2025).
- **Instructor** for “COMP3010: Algorithm Design”, VinUniversity, Spring 2023.
- **Instructor** for “COMP3020: Machine Learning”, VinUniversity, Fall (2022, 2023, 2024, 2025).

ADVISING AND MENTORING ROLES

Advising or Co-advising the following PhD students:

- *Thu Le*, Computer Science, VinUniversity, Fall 2025-present
- *Chau H Tran*, Computer Science, UIUC, co-advised with **Heng Ji**, Fall 2025-present
- *Duy Nguyen*, Computer Science, UIUC, co-advised with **Minh N. Do**, Fall 2024-present
- *Nghia Dai Nguyen*, Computer Science, UIUC, co-advised with **Lav R. Varshney** and **Pablo D Robles Granda**, Fall 2024-present
- *Thinh Tran*, Computer Science, VinUniversity, co-advised with **Kok-seng Wong**, Fall 2024-present
- *Tuan Minh Nguyen*, Computer Science, VinUniversity, co-advised with **Kok-seng Wong**, Spring 2023-present

Advising or Co-advising the following Master students:

- *Nguyen K Hoang*, Computer Science, UIUC, co-advised with **Dilek Hakkani-Tür**, Fall 2025-present

- *Oanh Ngoc Tran*, Computer Science, VinUniversity, Spring 2026-present

Advising or Co-advising the following Research Residents/Assistants (RAs):

- *Quang MT Nguyen* (from VNU-HCMC), VinUni RA, Fall 2025-present
- *Cat Le* (from VNU-HCMC), VinUni RA, Fall 2025-present
- *Nam Nguyen* (from HCMUT), VinUni RA, Fall 2025-present
- *Vinh Pham* (from HCMUT), VinUni RA, Fall 2025-present
- *Phan Hoang* (from VinUniversity), VinUni RA, Fall 2025-present
- *Quoc-Hung Le Thanh* (from VinUniversity), VinUni RA, Fall 2025-present
- *Oanh Ngoc Tran* (from HCMUT), VinUni RA, co-advised with **Heng Ji**, Fall 2024-Fall 2025 → **Next:** MS Student at VinUni
- *Chau Hoai Tran* (from VNU-HCMC), VinUni RA, co-advised with **Heng Ji**, Summer 2024-Fall 2025 → **Next:** PhD Student at UIUC
- *Hieu Ngoc Nguyen* (from VNU-Hanoi), VinUni RA, Spring 2024-Fall 2025 → **Next:** PhD Student at Penn State University
- *Phuc Minh Nguyen* (from VNU-HCMC), VinUni RA, Spring 2024-present
- *Quan Huu Do* (from VNU-HCMC), VinUni RA (Spring 2024-Fall 2025) → **Next:** PhD Student at Institut Polytechnique de Paris
- *Hung-Quang Nguyen* (from VNU-Hanoi), VinUni RA, Spring 2023-present
- *Cao-Duy Hoang* (from Viet-German University), VinUni RA, Fall 2022-present
- *Chinh Duc La* (from Australian National University), VinUni RA, Summer 2023-present
- *Jason Yang* (from University of Malaya), VinUni RA, Summer 2023-Fall 2023 → **Next:** AMD
- *Hieu Nguyen*, VinAI Resident (Fall 2022-present) → **Next:** PhD Student at Chinese University of Hong Kong.
- *Hoang Pham* (HUST-Hanoi), VinUni RA (Summer 2023-Fall 2023) → **Next:** PhD Student at University of Warwick

Advising or Co-advising the following Research Undergraduate students:

- *Chi H Nguyen* (Electrical Engineering, VinUni), Summer 2025-present
- *Quoc-Hung Le Thanh* (Computer Science, VinUni), Summer 2024-Spring 2025 → **Next:** RA at VinUni
- *Hoang Phan* (Electrical Engineering, VinUni), Summer 2024-Spring 2025 → **Next:** RA at VinUni
- *Duy-Minh Nguyen* (Computer Science, VinUni), Fall 2024-Spring 2025 → **Next:** RA at VinUni
- *Huong-Lan Tran* (Computer Science, VinUni), Fall 2023-Summer 2024 → **Next:** Master Student at EPFL
- *Nghia Dai Nguyen* (Computer Science, VinUni), Fall 2023-Summer 2024 → **Next:** PhD Student at UIUC

ACADEMIC SERVICE

Organizers

- *Workshop Program Chair/Asian Conference on Machine Learning (ACML) 2024*
- *Main Organizer/CAI 2024 Workshop:* AI for Environmental Intelligence
- *Main Organizer/NeurIPS 2023 Workshop:* Backdoors in Deep Learning: The Good, the Bad, and the Ugly (BUGS). **ICML 2025 Workshop:** Data in Generative Models: The Good, the Bad, and the Greats (DIG-BUGS).

Area Chair

- International Conference on Machine Learning (**ICML**): 2025-2026
- International Conference on Artificial Intelligence and Statistics (**AISTATS**): 2025-2026
- Conference on Neural Information Processing Systems (**NeurIPS**): 2024-2025
- International Conference on Learning Representations (**ICLR**): 2026
- Asian Conference on Machine Learning (**ACML**): 2024-2025

Program Committee

- International Conference on Learning Representations (**ICLR**): 2021-2025
- Conference on Neural Information Processing Systems (**NeurIPS**): 2020-2023
- International Conference on Machine Learning (**ICML**): 2020-2024
- Conference on Computer Vision and Pattern Recognition (**CVPR**): 2020-2025
- International Conference on Computer Vision (**ICCV**): 2021-2023
- European Conference on Computer Vision (**ECCV**): 2020-2024
- AAAI Conference on Artificial Intelligence (**AAAI**): 2021-2025
- IEEE International Conference on Big Data (**BigData**): 2020, 2022
- 1st International Workshop on Industrial Recommendation Systems (**IRS**): 2020-2021

Journal Editorial Board

- Discover Data (Springer Nature): 2023-
- ACM AI Letters (ACM Journals): 2025-

Journal Reviewer

- Transactions on Machine Learning Research (**TMLR**): 2025
- IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**): 2025
- Journal of Machine Learning Research (**JMLR**): 2024
- IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**): 2024
- ACM Transactions on Internet Technology (**TOIT**): 2020
- ACMD Transactions on Knowledge Discovery from Data (**TKDD**): 2018, 2019, 2020, 2021

Conference Reviewer

- ACM SIGKDD International Conference on Knowledge discovery and data mining (**KDD**): 2017-2019
- ACM International Conference on Information and Knowledge Management (**CIKM**): 2017-2019
- ACM International Conference on Web Search and Data Mining (**WSDM**): 2017-2019
- The Web Conference (WWW): 2017-2019
- International Joint Conference on Artificial Intelligence (**IJCAI**): 2017-2019

Service Award

- 2024: ECCV (**Outstanding Reviewer**)
- 2022: NeurIPS (**Top Reviewer**)
- 2021: CVPR (**Outstanding Reviewer**)

HONORS, AWARDS, AND DISTINCTIONS

VinUniversity's Outstanding Contribution Award – VinUniversity	2024
VinUniversity's Rising Star Award – VinUniversity	2024
Google Research Award (South/Southeast Asia) – VinUniversity	2024

VinUniversity's Outstanding Research Award – VinUniversity	2024
Green Industrial AI Challenge - Winner – VinUniversity	2024
Criteo Research Award – Virginia Tech	2018
NSF Urban Computing Fellowship – Virginia Tech	2016-2017
Graduation Honor, Summer Cum Laude – Webster University	2006
Regional ACM Collegiate Computing Contest Honor – Webster University	2004
University Scholarship – Webster University	2003

PATENTS

B. E. Crook, **Khoa D Doan**, G. K. Ng, C. G. Nicotra, M. J. Wrona. Systems, methods, and apparatus for reverse geocoding. US Patent US20160330592A1.

PUBLICATIONS

Note: Students advised or co-advised directly

1. Duy MH Nguyen, **Trung Quoc Nguyen**, et al. FACET: A Fragment-Aware Conformer Ensemble Transformer. *Under Submission*.
2. **Phuc M Nguyen**, **Chinh D. La**, Duy HM Nguyen, Nitesh V. Chawla, Binh T Nguyen, & **Khoa D Doan**. The Reasoning Boundary Paradox: How Reinforcement Learning Constrains Language Models. *Under Submission*. <https://arxiv.org/abs/2510.02230>
3. **Duy Cao Hoang**, **Thanh Quoc Hung Le**, Rui Chu, Ping Li, Weijie Zhao, Yingjie Lao, **Khoa D Doan**. SpARK: An Embarrassingly Simple Sparse Watermarking in LLMs with Enhanced Text Quality. *Under Submission*.
4. **Duy A Nguyen**, **Quan H Do**, **Khoa D Doan** & Minh N Do. Are you SURE? Enhancing Multi-modal Pretraining with Missing Modalities through Uncertainty Estimation. *Under Submission*.
5. **Sze Jue Yang**, **Chinh Duc La**, **Hung-Quang Nguyen**, Eugene Bagdasaryan, Kok-Seng Wong, Anh Tuan Tran, Chee Seng Chan, & **Khoa D Doan**. Synthesizing Physical Backdoor Datasets: An Automated Framework Leveraging Deep Generative Models. *Under Submission*.
6. **Sze Jue Yang**, **Hung-Quang Nguyen**, Chee Seng Chan, & **Khoa D Doan**. Everyone Can Attack: Exploiting Image Compression Artifacts as a Natural Backdoor Attack. *Under Submission*.
7. **Thinh Dao**, **Khoa D Doan**, Kok-Seng Wong. Clean-Label Physical Backdoor Attacks with Data Distillation. *AAAI Conference on Artificial Intelligence 2026 (AAAI)*.
8. **Phuc M Nguyen**, **Ngoc-Hieu Nguyen**, Binh T Nguyen, & **Khoa D Doan**. Mitigating Reward Over-optimization in Direct Alignment Algorithms with Adaptive Importance Sampling. *The Thirty-ninth Conference on Neural Information Processing Systems 2025 (NeurIPS)*.
9. **Hung-Quang Nguyen**, **Hoang Phan**, & **Khoa D Doan**. Unveiling Concept Attribution in Diffusion Models. *The Thirty-ninth Conference on Neural Information Processing Systems 2025 (NeurIPS)*. <https://arxiv.org/abs/2412.02542>.
10. Tuan Anh Tran, Duy Minh Ho Nguyen, **Hoai-Chau Tran**, Michael Barz, **Khoa D Doan**, Roger Wattenhofer, Vien Anh Ngo, Mathias Niepert, Daniel Sonntag, Paul Swoboda. How Many Tokens Do 3D Point Cloud Transformer Architectures Really Need? *The Thirty-ninth Conference on Neural Information Processing Systems 2025 (NeurIPS)*.

11. Duy Minh Ho Nguyen, [Trung Quoc Nguyen](#), Ha Thi Hong Le, Mai Thanh Nhat Truong, Trung Tin Nguyen, Nhat Ho, **Khoa D Doan**, Duy Duong-Tran, Li Shen, Daniel Sonntag, James Zou, Mathias Niepert, Hyojin Kim, Jonathan E Allen. From Fragments to Geometry: A Unified Graph Transformer for Molecular Representation from Conformer Ensembles. *ICML 2025 Workshop - Generative AI and Biology (GenBio)*.
12. Parshin Shojaee*, [Ngoc-Hieu Nguyen*](#), Kazem Meidani, Amir Barati Farimani, **Khoa D Doan**, Chandan K Reddy. Llm-srbench: A new benchmark for scientific equation discovery with large language models. *International Conference on Machine Learning 2025 (ICML)*. ORAL. (*Equal Contribution)
13. [Phuc M Nguyen](#), [Chinh Duc La](#), Heng Ji, **Khoa D Doan**. Weak-to-strong Generalization via Formative Learning from Student Demonstrations & Teacher Evaluation. *ICML 2025 Workshop - Data in Generative Models-The Bad, the Ugly, and the Greats (DIG-BUGS)*.
14. [Hung-Quang Nguyen](#), [Ngoc-Hieu Nguyen](#), The-Anh Ta, Thanh Nguyen-Tang, Kok-Seng Wong, Hoang Thanh-Tung, & **Khoa D Doan**. Wicked Oddities: Selectively Poisoning for Effective Clean-Label Backdoor Attacks. *International Conference on Learning Representations 2025 (ICLR)*. <https://openreview.net/forum?id=1Z3C49JQVf>
15. [Hoang Pham](#), The-Anh Ta, Anh Tran, & **Khoa D Doan**. Flatness-aware Sequential Learning Generates Resilient Backdoors. *European Conference on Computer Vision 2024 (ECCV)*. ORAL.
16. [Tran Ngoc Huynh](#), Anh Tuan Tran, **Khoa D Doan**, & Tung Pham. Data Poisoning Quantization Backdoor Attack. *European Conference on Computer Vision 2024 (ECCV)*.
17. Banibrata Ghosh, HARIPRIYA HARIKUMAR, **Khoa D Doan**, Svetha Venkatesh, & Santu Rana. Composite Concept Extraction through Backdooring. *International Conference on Pattern Recognition 2024 (ICIP)*. ORAL.
18. Grigorii Khvatskii, Nuno Moniz, **Khoa D Doan**, & Nitesh V Chawla. Class-Aware Contrastive Optimization for Imbalanced Text Classification. *The Ninth International Conference on Data Science and Management of Data 2024 (CODS-COMAD)*. ORAL.
19. [Hieu Nguyen](#), Duy Nguyen, **Khoa D Doan**, & Viet-Anh Nguyen. Cold-start Recommendation by Personalized Embedding Region Elicitation. *Conference on Uncertainty in Artificial Intelligence 2024 (UAI)*.
20. [Duy Cao Hoang](#), [Hung-Quang Nguyen](#), Minlong Peng, Kok-seng Wong, & **Khoa D Doan**. "Fooling the Textual Fooler via Randomizing Latent Representations". *Findings of the Association for Computational Linguistics 2024 (ACL)*.
21. [Hung-Quang Nguyen](#), Yingjie Lao, Tung Pham, Kok-seng Wong, & **Khoa D Doan**. "Understanding the Robustness of Randomized Feature Defense Against Query-Based Adversarial Attacks". *International Conference on Learning Representations 2024 (ICLR)*.
22. [Thuy-Dung Nguyen](#), [Tuan M Nguyen](#), Anh T Tran, **Khoa D Doan**, & Kok-seng Wong. "IBA: Towards Irreversible Backdoor Attacks in Federated Learning". *The Thirty-seventh Conference on Neural Information Processing Systems 2023 (NeurIPS)*.
23. [Thuy-Dung Nguyen](#), [Tuan Nguyen](#), Phi Le Nguyen, Hieu H Pham, **Khoa D Doan**, & Kok-Seng Wong. Backdoor attacks and defenses in federated learning: Survey, challenges and future research directions. *Engineering Applications of Artificial Intelligence 2024 (EAAI)*.

24. **Thai-Hung Nguyen**, Hong-Phuc Vu, Dung Thuy Nguyen, Tuan Minh Nguyen, **Khoa D Doan**, & Kok-Seng Wong. An Empirical Study of Federated Unlearning: Efficiency and Effectiveness. *Asian Conference on Machine Learning 2023 (ACML)*.
25. **Hung-Quang Nguyen**, **Ngoc-Hieu Nguyen**, The-Anh Ta, Thanh Nguyen-Tang, Thanh-Tung Hoang, & **Khoa D Doan**. Clean-label Backdoor Attacks by Selectively Poisoning with Limited Information from Target Class. *NeurIPS 2023 Workshop on Backdoors in Deep Learning: The Good, the Bad, and the Ugly*.
26. **Ngoc-Hieu Nguyen**, **Hung-Quang Nguyen**, The-Anh Ta, Nguyen-Tang Thanh, **Khoa D Doan**, & Thanh-Tung Hoang. "A Cosine Similarity-based Method for Out-of-Distribution Detection". *ICML 2023 Workshop on Spurious Correlations, Invariance and Stability*.
27. **Khoa D Doan**, Shulong Tan, Weijie Zhao, & Ping Li. "Asymmetric Hashing for Fast Ranking via Neural Network Measures". *International ACM SIGIR conference on research and development in Information Retrieval 2023 (SIGIR)*.
28. **Khoa D Doan**, Yingjie Lao, Yang Peng, & Ping Li. "Defending backdoor attacks on vision transformer via patch processing". *AAAI Conference on Artificial Intelligence 2023 (AAAI)*.
29. **Khoa D Doan**, Sarkhan Badirli, & Chandan K Reddy. "Unified Learning of Multipurpose Energy Based Generative Hashing Network". *Asian Conference on Computer Vision 2022 (ACCV)*.
30. **Khoa D Doan**, Yingjie Lao, & Ping Li. "Marksman Backdoor: Backdoor Attacks with Arbitrary Target Class ". *The Thirty-sixth Conference on Neural Information Processing Systems 2022 (NeurIPS)*.
31. **Khoa D Doan**, Yang Peng, & Ping Li. "One Loss for Quantization: Deep Hashing with Discrete Wasserstein Distributional Matching". *Conference on Computer Vision and Pattern Recognition 2022 (CVPR)*.
32. **Khoa D Doan**, Yingjie Lao, & Ping Li. "Backdoor Attack with Imperceptible Input and Latent Modification". *The Thirty-fifth Conference on Neural Information Processing Systems 2021 (NeurIPS)*.
33. **Khoa D Doan**, Yingjie Lao, Weijie Zhao, & Ping Li. "LIRA: Learnable, Imperceptible and Robust Backdoor Attacks". *IEEE International Conference on Computer Vision 2021 (ICCV)*.
34. **Khoa D Doan**, Saurav Manchanda, Suchismit Mahapatra, & Chandan K Reddy. "Interpretable Graph Similarity Computation via Differentiable Optimal Alignment of Node Embeddings". *International ACM SIGIR conference on research and development in Information Retrieval 2021 (SIGIR)*.
35. **Khoa D Doan**, Fengjiao Wang, Saurav Manchanda, Sathiya S Keerthi, Avaradeep Bhowmik, & Chandan K Reddy. "Image Generation via Minimizing Fréchet Distance in Discriminator Feature Space. <https://arxiv.org/abs/2003.11774>.
36. Sarkhan Badirli, Xuanqing Liu, Zhengming Xing, Avradeep Bhowmik, **Khoa D Doan**, & Sathiya S Keerthi. Gradient Boosting Neural Networks: GrowNet. <https://arxiv.org/abs/2002.07971>.
37. Saurav Manchanda, **Khoa D Doan**, & Sathiya S Keerthi. Regression via Implicit Models and Optimal Transport Cost Minimization. <https://arxiv.org/abs/2003.01296>.

38. **Khoa D Doan** & Chandan K Reddy. Efficient Implicit Unsupervised Text Hashing using Adversarial Autoencoder. *The Web Conference 2020* ([WWW](#)).
39. Saurav Manchanda, Pranjul Yadav, **Khoa D Doan**, & Sathiya S Keerthi. Targeted display advertising: the case of preferential attachment. *IEEE International Conference on Big Data 2019* ([BigData](#)).
40. **Khoa D Doan**, Pranjul Yadav, & Chandan K Reddy. Adversarial Factorization Autoencoder for Look-alike Modeling. *ACM International Conference on Information and Knowledge Management 2019* ([CIKM](#)).
41. **Khoa D Doan**, Guolei Yang, & Chandan K Reddy. An Attentive Spatio-Temporal Neural Model for Successive Point Of Interest Recommendation. *Pacific-Asia Conference on Knowledge Discovery and Data Mining 2019* ([PAKDD](#)).
42. **Khoa D Doan**, Amidu Oloso, Kwo-Sen Kuo, Thomas L Clune, Hongfeng Yu, Brian Nelson, & Jian Zhang. Evaluating the impact of data placement to Spark and SciDB with an earth science use case. *IEEE International Conference on Big Data 2016* ([BigData](#)).
43. **Khoa D Doan**, Amidu Oloso, Kwo-Sen Kuo, & Thomas L Clune (2014, December). Performance comparison of big-data technologies in locating intersections in satellite ground tracks. *In Proceedings of ASE BigData Conference, 2014*.

TALKS/PRESENTATIONS

- **Khoa D Doan**. Creating Impact as an ML Scientist in Product-driven Research: The LLM Story. (Talk). [University of South Florida](#). 2025.
- **Khoa D Doan**. Eliciting novel problem-solving ability in LLMs: where are we? (Talk). [National Taiwan University](#). 2025.
- **Khoa D Doan**. Low-resource Machine Learning and Opportunities in LMICs (Talk). [ICLR AI for Science](#). 2025.
- **Khoa D Doan**. Product-driven Research: An example from building women gynecological screening chatbo? (Talk). [AI Global South Summit](#). 2024.
- **Khoa D Doan**. Machine Learning on Machine Learning (Talk). [University of Science, Vietnam National University & Ho Chi Minh City University of Technology](#). 2024.
- **Khoa D Doan**. Retrieval Augmented Generation and LLM on the Edge (Talk). Vietnam Petroleum Institute (Hanoi), Vietnam. 2024.
- **Khoa D Doan** & Heng Ji. Toward Practical Machine Learning Applications: Adversarial Robustness and Drug Discovery (Talk). [University of Science, Vietnam National University, Ho Chi Minh City University of Technology, Hanoi University of Science and Technology, & VinAI Research](#). 2024.
- **Khoa D Doan**. Robust Machine Learning and Low-resource NLP (Talk). [University of Malaya](#), Malaysia. 2023.
- **Khoa D Doan**. Toward Practical Machine Learning Applications and Low-resource NLP (Talk). [University of Notre Dame & Lucy Family Institute](#), USA. 2023.
- **Khoa D Doan**. Toward Practical Machine Learning Applications with Generative Models: Data Generation and Beyond (Seminar). [Singapore Management University \(SMU\)](#), Singapore. 2022.
- **Khoa D Doan**. [Information Retrieval with Deep Hashing \(Seminar\)](#). [VinAI, Vietnam](#), 2022.
- **Khoa D Doan**. Generative models meet similarity search (Seminar). [Baidu Cognitive Computing Lab](#), USA, 2020.
- **Khoa D Doan**, B. Cave & Chandan K Reddy (Poster). CrimeLab: A data-driven approach. Virginia Tech Urban Computing Day, 2017.

- H. Avik, G. Takahara & **D. Khoa** (*Talk*). Social Media Analytics using Bayesian Multistate Intensity Models. 43rd Annual Meeting of the Statistical Society of Canada. <https://bit.ly/3extrE>

REFEREES

Referee 1:

Dr. **Chandan Reddy**
Professor, Computer Science, **Virginia Tech**
Email: reddy@cs.vt.edu

Referee 2:

Dr. Ping Li
CEO and Founder, **VecML Inc.**
Email: pingli98@gmail.com

Referee 3:

Dr. **Nitesh Chawla**
Frank M. Freimann Professor of Computer Science and Engineering, University of Notre Dame
Founding Director, the Lucy Family Institute for Data and Society
Email: nitesh.chawla@gmail.com

Referee 4:

Dr. **Heng Ji**
Professor, Siebel School of Computing and Data Science, University of Illinois Urbana-Champaign
Founding Director Amazon-Illinois Center on AI for Interactive Conversational Experiences (AICE)
Email: hengji@illinois.edu

Referee 5:

Dr. **Thanh Huong (Helen) Nguyen**
Ivan Racheff Professor of Environmental Engineering in CEE, University of Illinois Urbana-Champaign
Email: thn@illinois.edu