I. Personal Information

Name: Latifur Khan
School: Engineering and Computer Science
Department: Computer Science
Date: 02/15/2025

II. Professional Narrative

Dr. Latifur Khan is currently a *Professor* in the *Computer Science* department at *the University of Texas, Dallas* where he has been teaching and conducting research since September 2000. He received his *Ph.D. and M.S.* degrees in Computer Science from the *University of Southern California (USC),* in August of 2000, and December of 1996 respectively. He is currently supported by grants from *NSF*, ARMY, NIST, ONR, *NSA, IBM, HPE and the Air Force Office of Scientific Research*. Dr. Khan's research emphasizes several broad areas including *big data analytics and management*, generative AI including large language models (LLM), *data mining, complex data management including semantic web, and geospatial data and cyber security*. Recently, he has become a Fellow of IEE, BCS, and IET. At UT Dallas he has introduced and taught a new regular lecture-oriented course, "Big Data Management and Analytics."

Dr. Khan applies his "Big data analytics" research in data mining/machine learning (online learning, metric learning, adversarial learning, domain adaptation), and cybersecurity areas. With regard to *data mining/machine learning*, Dr. Khan is researching semi-*supervised learning algorithms* to classify evolving data streams including brand new class or emerging class.

Dr. Khan's group develops ConfliBERT, a domain-specific pretrained language model for conflict and political violence. ConfliBERT consistently outperforms traditional transformer based large language model (LLM) BERT (trained on generic corpora) in the conflict and political violence domain. Further, it outperforms other models even with limited training data, which addresses an endemic problem for researchers due to high annotation costs.

In his political science applications, Dr. Khan's group has observed that the collected data follow a skewed distribution (data from a few classes appear much more frequently than those of other classes. His group has proposed a novel set convolution (SetConv) operation and episodic training strategy to extract a single representative for each class so that classifiers can later be trained by a balanced class distribution. Dr. Khan has proved that his proposed algorithm is permutation-invariant despite the order of inputs, and experiments on multiple large-scale benchmark text datasets show the superiority of our proposed framework when compared to other state-of-the-art (SOTA) methods.

Dr. Latifur Khan has published over 300 papers in premier journals such as *VLDB, Journal of Web Semantics, IEEE TDKE, IEEE TDSC, IEEE TSMC, and AI Research* and in prestigious conferences such as *AAAI*, *IJCAI, CIKM, ICDE*, *ACM GIS, IEEE ICDM, IEEE BigData, ECML/PKDD, PAKDD, ACM Multimedia, ACM WWW, ICWC, ACM SACMAT, IEEE ICSC, IEEE Cloud and INFOCOM*. He has been invited to give keynotes and invited talks at a number of conferences hosted by *IEEE and ACM*. In addition, he has conducted tutorial sessions in prominent conferences such *as SIGKDD 2016, 2017, IJCAI 2017, AAAI 2017, SDM 2017, IEEE CollaborateCom 2014*, *CloudCom 2013*, *PAKDD 2011& 2012, DASFAA 2012, ACM WWW 2005, MIS2005, DASFAA 2007, and WI 2008*. As of today, fifteen Ph.D. students have graduated under Dr. Khan's supervision and one of them (Dr. Feng Luo) is currently working as an *associate professor* in *Computer Science* department at *Clemson University*, USA. Dr. Khan also collaborates actively with top researchers from *MIT, UIUC, UMN, Purdue and IBM TJ Watson Research.* (For further information please check: http://www.utdallas.edu/~lkhan)

III. Scholarly Activities

Refereed Journal Publications:

1. "*An Automated Vulnerability Detection Framework for Smart Contracts.*" Feng Mi, Chen Zhao, Zhuoyi Wang, Sadaf Halim, Xiaodi Li, Zhouxiang Wu, **Latifur Khan**, Bhavani Thuraisingham. Distributed Ledger Technologies: Research & Practice DLT-2023-0109.

2. "*Heterogeneous Domain Adaptation for Multistream Classification on Cyber Threat Data.*" Yi-Fan Li, Yang Gao, Gbadebo Ayoade, **Latifur Khan**, Anoop Singhal, Bhavani Thuraisingham. IEEE Trans. Dependable Secur. Comput. 21(1): 1-11 (2024).

3. "*Dynamic Environment Responsive Online Meta-Learning with Fairness Awareness.*" Chen Zhao, Feng Mi, Xintao Wu, Kai Jiang, **Latifur Khan**, Feng Chen. ACM Trans. Knowl. Discov. Data 18(6): 153:1-153:23 (2024).

4. "*FedFusion: Adaptive Model Fusion for Addressing Feature Discrepancies in Federated Credit Card Fraud Detection.*" Nahid Ferdous Aurna, Md Delwar Hossain, **Latifur Khan**, Yuzo Taenaka, Youki Kadobayashi. IEEE Access 12: 136962-136978 (2024).

Refereed Conference Publications:

Published:

1. "*Leveraging Codebook Knowledge with NLI and ChatGPT for Zero-Shot Political Relation Classification.*" Yibo Hu, Erick Skorupa Parolin, **Latifur Khan**, Patrick T. Brandt, Javier Osorio, Vito J. D'Orazio. Annual Meeting of the Association for Computational Linguistics (ACL) (1) 2024: 583-603.

2. "*Algorithmic Fairness Generalization under Covariate and Dependence Shifts Simultaneously.*" Chen Zhao, Kai Jiang, Xintao Wu, Haoliang Wang, **Latifur Khan**, Christan Grant, Feng Chen. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2024: 4419-4430.

3. "Handling Non-IID Data in Federated Learning Using Metaheuristic Optimization Techniques." Amin Birashk, Sadaf MD Halim, **Latifur Khan**: Proceedings of the IEEE International Conference on Data Mining, ICDM 2024, 681-686, December 9-12, 2024

4. "*RFPG: Question-Answering from Low-Resource Language (Arabic) Texts using Factually Aware RAG.*" Mitha Alshammary, Md Nahiyan Uddin, **Latifur Khan**. IEEE International Conference on Collaboration and Internet Computing (CIC) 2024: 107-116.

5. "*Ensuring End-to-End IoT Data Security and Privacy Through Cloud-Enhanced Confidential Computing.*" Md Shihabul Islam, Mahmoud Zamani, Kevin W. Hamlen, **Latifur Khan**, Murat Kantarcioglu. Data and Applications Security and Privacy (DBSec) 2024: 71-91.

6. "*ALERT: A Framework for Efficient Extraction of Attack Techniques from Cyber Threat Intelligence Reports Using Active Learning.*" Fariha Ishrat Rahman, Sadaf Md. Halim, Anoop Singhal, **Latifur Khan**. Data and Applications Security and Privacy (DBSec) 2024: 203-220.

7. "*Blockchain Smart Contract Vulnerability Detection and Segmentation Using ML.*" Sabbir Ahmed, Luay Abdeljaber, Sharif Noor Zisad, Mohammad Shahadat Hossain, **Latifur Khan**. IEEE International Conference on Blockchain and Cryptocurrency (ICBC) 2024: 290-292.

8. "*Banking Malware Detection: Leveraging Federated Learning with Conditional Model Updates and Client Data Heterogeneity.*" Nahid Ferdous Aurna, Md Delwar Hossain, Hideya Ochiai, Yuzo Taenaka, **Latifur Khan**, Youki Kadobayashi. International Conference on Information Systems Security and Privacy (ICISSP) 2024: 309-319.

9. "*Extractive Question Answering for Spanish and Arabic Political Text.*" Sultan Alsarra, Parker Whitehead, Naif Alatrush, Luay Abdeljaber, **Latifur Khan**, Javier Osorio, Patrick T. Brandt, Vito D'Orazio. International Conference on Social, Cultural, and Behavioral Modeling (SBP-BRiMS) 2024: 144-153.

10. "*ConfliLPC: Logits and Parameter Calibration for Political Conflict Analysis in Continual Learning.*" Xiaodi Li, Niamat Zawad, Patrick T. Brandt, Javier Osorioc, Vito D'Orazio, **Latifur Khan**. International Conference on Big Data (BigData), pp. 6320-6329. IEEE, 2024.

11. "*LLM-Sentry: A Model-Agnostic Human-in-the-Loop Framework for Securing Large Language Models.*" Saquib Irtiza, Khandakar Ashrafi Akbar, Arowa Yasmeen, **Latifur Khan**, Ovidiu Daescu, Bhavani Thuraisingham. 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA 2024), pp 245-254. Washington D.C., USA, IEEE 2024.

12. "*VulPrompt: Prompt-Based Vulnerability Detection Using Few-Shot Graph Learning.*" Saquib Irtiza, Xiaodi Li, Mahmoud Zamani, **Latifur Khan**, Kevin W. Hamlen. IFIP Annual Conference on Data and Applications Security and Privacy (DBSEC), pp. 221-240. San Jose, CA, USA, Cham: Springer Nature Switzerland 2024.

13. "*Leveraging Multimodal Retrieval-Augmented Generation for Cyber Attack Detection in Transit Systems.*" Muhaimin Bin Munir, Yuchen Cai, **Latifur Khan**, Bhavani Thuraisingham. IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA) (pp. 341-350). IEEE, 2024.

14. "*Keep it local: Comparing domain-specific LLMs in native and machine translated text using parallel corpora on political conflict.*" Javier Osorio, Sultan Alsarra, Amber Converse, Afraa Alshammari, Dagmar Heintze, **Latifur Khan**, Naif Alatrush, Patrick T. Brandt, Vito D'Orazio, Niamat Zawad, Mahrusa Billah. 2nd International Conference on Foundation and Large Language Models (FLLM). IEEE, 2024.

15. "*LSEBMCL: A Latent Space Energy-Based Model for Continual Learning.*" Xiaodi Li, Dingcheng Li, Rujun Gao, Mahmoud Zamani, **Latifur Khan**. In 2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC).

Invited Performances, Lectures, Colloquia Presentations or Exhibitions

Invited Talk

- International Seminar(Online), Industry 4.0, JMI, India, February 2024
- North South university, "Big Data Analytics for Autonomous Vehicle," May 2024.
- BUET, "Fairness Aware ML", May 2024

Patent

- **USPTO Patent: Utility Patents.** The nominee is the primary inventor for all the patents.
- (a) Systems and methods for determining user attribute values by mining user network data and information, Patent number (PN): **US8965974B2; Citations: 67+ including by Facebook, Microsoft, LinkedIn, Yahoo, Snapchat et al.**
- (b) Systems and methods for detecting a novel data class, PN: **US9165051B2; Citations: 45+ including by IBM, Amazon, Microsoft, Qualcomm, Symantec, Oracle, Veritas, Bank of America, Intel, GE, HP**
- (c) Systems and methods for automated detection of application vulnerabilities **PN: US9977904B2; Citations: 23+ including by Twitter, IBM**
- (d)Real-time, stream data information integration and analytics system, **Serial Number: 10902468 (issued on January 26, 2021)**

IV. Proposal and Grant Activity (Align)

Grants Received:

- "FARML: Flexible, Adaptive and Robust Machine Learning for Resource Constrained Environments", US ARMY ACC-APG-RTP, June 2024-May 2027, $3,489,175
- "Improving RAG Models with Knowledge Graph Integration," Booz Allen Hamilton Inc (NSA), ,July 2024-July 2025, $125,568
- "UT Dallas Center for Secure and Trustworthy Artificial Intelligence," NIST, 10/2024-9/2026, $962,930
- "Frameworks: Infrastructure For Political And Social Event Data using Machine Learning, $1.59M, September 2023- August 2026, National Science Foundation (NSF)
- "Addressing Vaccine Hesitancy through Social-Media Data Collection, Machine Learning Algorithm Development and Targeted Use" ("Study"), Merck Sharp & Dohme Corp, March 2024-Deceber 2024, $36K.
- "SaTC EDU Directions workshop," National Science Foundation, October 2023-June 2024, $167,864

- *"*National Center for Transportation Cybersecurity and Resiliency (TraCR), U.S. Department of Transportation (USDOT), National Center on Transportation Cybersecurity, Total amount $4 million/year, UT Dallas share $350K/year over 5 years

- "Mapping of Cyber Threat Intelligence to Vulnerabilities and System Level Traces Using Ontology,"
  PI: Latifur Khan
  Funding Organization: National Institute of Standards and Technology (NIST)
  Date: March 2023-February, 2024
  Amount: $69,514.

- "FAI: A novel paradigm for fairness-aware deep learning models on data streams,"
  PI: Feng Chen, Latifur Khan
  Funding Organization: *National Science Foundation*,
  Date: September 2023-August 2026,
  Amount: $999,444

- *"Information and Software Assurance: Dynamic Multi-Layer Multi-Stage Cognition-Inspired Active Cyber Deception,"*
  PI: Kevin Hamlen, Latifur Khan
  Funding Organization: ARMY
  Date: Dec 2020- Nov 2024
  Amount: $600K + $150K (supplement funding)

- "SATC: EDU: Curriculum Development for Secure Blockchain"
  PI: Latifur Khan, Murat Kantarcioglu, Bhavani Thuraisingham
  Funding Organization: NSF
  Date: January 1, 2020 – December 31, 2025
  Amount: $499,595

- **<u>Names of Doctoral Students and Stipend Amounts (in 2024)</u>**

  - Akbar,Khandakar Ashrafi –RA for full year
  - Bin Munir,Muhaimin –RA/TA
  - Birashk,Amin--TA
  - Cai,Yuchen –RA for Fall 2024
  - Deniz,Omer Faruk—RA for Fall 2024
  - Du,Wenbo –RA/TA
  - Halim,Sadaf Md—RA for full year
  - Hosseini,Seyyed Mohammadsaleh—RA for Fall 2024
  - Irtiza,Saquib—RA for full year
  - Sakib,Md Nazmus--TA
  - Uddin,Md Nahiyan—RA for full year
  - Zaoad,Md Shahir—RA in Fall 2024
  - Zawad,Niamat—RA for full year

Hosted a Fullbright Scholar:

- Dr. Ricardo Cerri, Instituto de Ciências Matemáticas e de Computação, University of São Paulo, São Carlos, SP, Brazil from January 2025-May 2025
- Dr, Mohammad Hossain, Chittagong University from January 1, 2023-Septemeber 30, 2023.

V. Teaching Activities

Enrollment	Semester	Prefix	Number	Course Name	Format (0..4)	Syllabus Posted
13	Spring 2024	CS	7301	Recent Advances in Computing	2	Y
127	Fall 2024	CS	4371	Introduction Big Data Management and Analytics	2	Y

<u>Doctoral Students Supervised in 2024</u>

Spring 2024:

-
- Abdeen,Basel
- Akbar,Khandakar Ashrafi
- Birashk,Amin
- Du,Wenbo
- Halim,Sadaf Md
- Hosseini,Seyyed Mohammadsaleh
- Irtiza,Saquib
- Islam,Md Shihabul
- Rahman,Fariha Ishrat
- Zawad,Niamat

Summer 2024:

- Abdeen,Basel
- Rahman,Fariha Ishrat

Fall 2024:

- Akbar,Khandakar Ashrafi
- Bin Munir,Muhaimin

- Birashk,Amin
- Cai,Yuchen
- Deniz,Omer Faruk
- Du,Wenbo
- Halim,Sadaf Md
- Hosseini,Seyyed Mohammadsaleh
- Irtiza,Saquib
- Sakib,Md Nazmus
- Uddin,Md Nahiyan
- Zaoad,Md Shahir
- Zawad,Niamat
- Abdeljaber,Luay

Graduated Students in 2024

PhD

- Md Shihabul Islam, (PhD Spring 2024, Dissertation: Confidential Computing with Trusted Execution Environments, DataSecTech, Software Engineer)

- Xiaodi Li (PhD Summer 2024, Dissertation: Advanced Approaches in NLP and Security: Addressing Catastrophic Forgetting Through Continual Learning and Resolving Data Imbalance in Semi-supervised Setting, Post-doc, Mayo Clinic, Minnesota).

- Yigong Wang (PhD Spring 2024, Dissertation Title: *Continual Learning with Applications in Different Scenarios*, **Meta**, San Francisco, 2024, USA).

Employer and Position of Graduated Doctoral/MS Thesis Students in 2024

- Wooseong Yang (MS Spring 2023)

Number of Publications of PhD Students in 2024:

(J for Journal, C for Conference, A for Accepted/Published, P for Pending)

- Shihab Islam 1 C (A); 1 J (A)
- Zhuoyi Wang 2 C (A)
- Xiaodei Li 2 C (A)
- Mi Feng 1 C (A)
- Halim, Sadaf 1 C
- Saquib, Irtiza 1 C; I J (A)
- Luay, Abdeljaber 1 C

VII Professional Service Activities

UTD Administrative Duties

School Committees

- Member of *Colloquium Committee* since Fall 2006.
- Member of Course Curriculum Committee since Fall 2015
- Johsson School ADR-RAC Committee since 2023

Service Contributions External to UTD

Editorial Activities

- **Associate Editor** of *ACM Transaction on Knowledge Discovery and Data (TKDD),* (2021 -).

- **Associate Editor** of *ACM Transaction on Internet Technology (TOIT),* (2019 - 2023).

- **Associate Editor** of *IEEE Transactions on Artificial Intelligence (TAI),* (2021-2023)

- **Associate Editor** of *IEEE Transactions on Knowledge & Data Engineering (TKDE),* IEEE Computer Society, 2011-2015.

- **Associate Editor** of *International Journal of Data Mining, Modelling and Management (IJDMMM),* Inderscience Publishers, Switzerland since 2008.

- **Associate Editor** of *International Journal of Security Informatics,* Springer ISSN: 2190-8532, Journal no. 13388 (from Spring 2011).

- **Associate Editor** of *Journal of Advanced Computing and Applications,* USM, Malaysia (since July 2011)

Keynote Speech

- IEEE 2025 19th International Conference on Ubiquitous Information Management and Communication, Trustworthy and Generative AI for Social Good, Including Cybersecurity, January 2025 http://www.imcom.org/gnu/kspeakers.htm

- IEEE 13th International Conference On Electrical and Computer Engineering (ICECE), Trustworthy and Generative AI for Social Good, Including Cybersecurity, Dec 2024

- International Conference on Electronics and Informatics 2024 (ICEI2024), Generative AI for Social Good, Including Cybersecurity, Dec 2024

- 2024 2nd International Conference on Information and Communication Technology (ICICT 2024) 21-22 October 202, Generative AI including Large Language Models for Social Good & Cyber-security

Award

- **Fellow of IEEE since 2022**

- **Fellow of the British Computer Society (Fellow of BCS), IET** since 2021.

- **Fellow of SIRI (Society of Information Reuse and Integration)** award in Aug, 2018.

- **IBM Faculty Award (Research), 2016:** Fundamental Research on Data Analytics (Stream Mining).

- **ACM Distinguished Scientists,** since Dec 2012.

Chair/Co-Chair:

CHAIR/CO-CHAIR OF CONFERENCES:

- Co-chair, IEEE 23rd IEEE International Conference on Data Mining (ICDM 2023), Shanghai, China, December 2023
- Big Data Track chair, IEEE 43rd IEEE International Conference on Distributed Computing Systems (ICDCS 2023).

PROGRAM COMMITTEE MEMBER OF:

- 31th *ACM SIGKDD Conference on Knowledge Discovery and Data Mining,* August 2025 (area chair)
- IEEE 24th IEEE International Conference on Data Mining (ICDM 2023), December 2025
- 30th *ACM SIGKDD Conference on Knowledge Discovery and Data Mining,* August 2024 (area chair)

Session Chair

- *ACM SIGKDD Conference on Knowledge Discovery and Data Mining,* Spain, August 2024.

Annual Program Review Meeting:

- NSF Secure and Trustworthy CyberSpace Principal Investigators' Meeting (SaTC PI Meeting)

Reviewer:

- Invited Reviewer of Book, Database Systems by Elmasri and Navathe, Addison, Wesley.

- Invited Reviewer of Encyclopedia of Computer Science and Engineering, John Wiley & Sons.

- Invited Reviewer of *IEEE Transactions on Neural Network, TNN*.

- Invited Reviewer of *Journal of Machine Learning Research.*

- Invited Reviewer of *IEEE Intelligent Systems.*

- Invited Reviewer of *ACM/Springer Multimedia Systems Journal.*

- Invited Reviewer of *IEEE Transactions on Knowledge and Data Engineering, TKDE*.

- Invited Reviewer of *VLDB Journal, ACM/Springer Publisher.*

- Invited Reviewer of *IEEE Transactions on Dependable and Secure Computing.*

- External Reviewer of *International Journal of Knowledge and Information Systems (KAIS)*, Springer-Verlag.

VIII. Special Professional Recognition

- **Fellow of IEEE since 2022**
- **Fellow of the British Computer Society (Fellow of BCS), IET** since 2021.
- **ACM Distinguished Scientists,** since Dec 2012.
- **Fellow of SIRI (Society of Information Reuse and Integration)** award in Aug, 2018.
- **IBM Faculty Award (Research), 2016:** Fundamental Research on Data Analytics (Stream Mining).