

Short version:

Dr. Pin-Yu Chen is a principal research scientist and manager at IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA. He is also the chief scientist of [RPI-IBM AI Research Collaboration](#) and PI of ongoing [MIT-IBM Watson AI Lab projects](#). Dr. Chen received his Ph.D. in electrical engineering and computer science from the University of Michigan, Ann Arbor, USA, in 2016. Dr. Chen's recent research focuses on AI safety and robustness. His long-term research vision is to build trustworthy machine learning systems. He received the [IJCAI Computers and Thought Award](#) in 2023. He is a co-author of the book "[Introduction to Foundation Models](#)" and the book "[Adversarial Robustness for Machine Learning](#)". At IBM Research, he received several research accomplishment awards, including IBM Master Inventor, IBM Corporate Technical Award, and IBM Pat Goldberg Memorial Best Paper. His research contributes to IBM open-source libraries including Adversarial Robustness Toolbox (ART 360), AI Explainability 360 (AIX 360), and In-Context Explainability 360 (ICX-360). He has published more than 50 papers related to trustworthy machine learning at major AI and machine learning conferences, given tutorials at NeurIPS'22, AAAI('22,'23,'24,'26), IJCAI'21, CVPR('20,'21,'23), ECCV'20, ICASSP('20,'22,'23,'24), KDD'19, and Big Data'18, and organized several workshops for adversarial machine learning. He has been an IEEE Fellow since 2025. He is currently on the editorial board of Transactions on Machine Learning Research, IEEE Transactions on Signal Processing, IEEE Transactions on Pattern Recognition and Machine Intelligence, and IEEE Access. He is also an Area Chair or Senior Program Committee member for NeurIPS, ICLR, ICML, AAAI, IJCAI, and PAKDD, and a Distinguished Lecturer of ACM. He received the IEEE GLOBECOM 2010 GOLD Best Paper Award and UAI 2022 Best Paper Runner-Up Award. In 2025, he received the IEEE SPS Industry Young Professional Leadership Award.

Full version:

Dr. Pin-Yu Chen is a principal research scientist and manager at IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA. He is also the chief scientist of [RPI-IBM AI Research Collaboration](#) and PI of ongoing [MIT-IBM Watson AI Lab projects](#). Dr. Chen received his Ph.D. in electrical engineering and computer science and M.A. in Statistics from the University of Michigan, Ann Arbor, USA, in 2016. He received his M.S. degree in communication engineering from National Taiwan University, Taiwan, in 2011 and B.S. degree in electrical engineering and computer science (undergraduate honors program) from National Chiao Tung University, Taiwan, in 2009.

Dr. Chen's recent research focuses on AI safety and robustness. His long-term research vision is to build trustworthy machine learning systems. He received the [IJCAI Computers and Thought Award](#) in 2023 for his contributions to consolidating properties of trust, robustness and safety

into rigorous algorithmic procedures and computable metrics for improving AI systems. He has published more than 50 papers related to trustworthy machine learning at major AI and machine learning conferences, given tutorials at NeurIPS'22, AAAI('22,'23,'24,'26), IJCAI'21, CVPR('20,'21,'23), ECCV'20, ICASSP('20,'22,'23,'24), KDD'19, and Big Data'18, and organized several workshops for adversarial machine learning. He is a co-author of the book "[Introduction to Foundation Models](#)" and the book "[Adversarial Robustness for Machine Learning](#)". His research interest also includes graph and network data analytics and their applications to data mining, machine learning, signal processing, and cyber security. He was a recipient of the Chia-Lun Lo Fellowship from the University of Michigan Ann Arbor. He also received the IEEE GLOBECOM 2010 GOLD Best Paper Award and UAI 2022 Best Paper Runner-Up Award. Dr. Chen is currently on the editorial board of Transactions on Machine Learning Research, IEEE Transactions on Signal Processing, IEEE Transactions on Pattern Recognition and Machine Intelligence, and IEEE Access. He is also an area chair at top AI/ML conferences, including NeurIPS, ICLR, ICML, AAAI, IJCAI, PAKDD, etc. He is an IEEE Fellow, an ACM senior member, and a Distinguished Lecturer of ACM. He is also a current member of the IEEE SPS MLSP and ASPS Technical Committees. In 2025, he received the IEEE SPS Industry Young Professional Leadership Award.

At IBM Research, Dr. Chen's research contributes to IBM open-source libraries including Adversarial Robustness Toolbox (ART 360), AI Explainability 360 (AIX 360), and In-Context Explainability 360 (ICX-360). His inventions are incorporated in IBM business products, including Watson Studio and Watsonx.governance. Dr. Chen has co-invented more than 50 U.S. patents and received the honor of IBM Master Inventor. In 2022 and 2023, he received the IBM Pat Goldberg Memorial Best Paper Awards. In 2021, he received an IBM Corporate Technical Award. In 2020, he received an IBM Research special division award for research related to COVID-19. He also received several IBM Outstanding Research Accomplishment Awards. He is an IBM representative of the U.S. Artificial Intelligence Safety Institute. His AI-generated text detector model has been downloaded more than 2.5M times on Hugging Face.