EDITORIAL

AI's 10 to Watch, 2022

Jürgen Dix, *Clausthal University of Technology, 38678, Clausthal, Germany*

Zhongfei Zhang, *Binghamton University, State University of New York, Binghamton, NY, 13902-6000, USA*

IEEE Intelligent Systems is promoting young and aspiring artificial intelligence (AI) scientists and recognizing the rising stars as "AI's 10 Watch." This biennial 2022 edition is slightly different from the previous editions: We solicited submissions from individuals who had obtained their Ph.D. up to 10 years prior (as opposed to 5 years in all of the previous editions). This led to more applications of the highest quality. The selection committee finally had to select 10 outstanding contributors from a pool of 30+ highly competitive and strong nominations, which made the selection decisions rather difficult. After a careful and detailed selection process through many rounds of discussions via e-mails and live meetings, the committee voted unanimously on a short list of 10 top candidates who have all demonstrated outstanding achievements in different areas of AI. The selection was based solely on scientific quality, reputation, impact, and expert endorsements accumulated since their Ph.D. It is our honor and privilege to announce the following 2022 class of "AI's 10 to Watch."

- **Bo Li.** She is working on trustworthy machine learning (ML) at the intersection of ML, security and privacy, and game theory. She was able to integrate domain knowledge and logical reasoning abilities into data-driven statistical ML models to improve learning robustness with guarantees, and she has designed scalable privacy-preserving data-publishing frameworks for high-dimensional data. Her work has provided rigorous guarantees for the trustworthiness of learning systems and been deployed in industrial applications. She is an assistant professor with the University of Illinois at Urbana-Champaign.
- **Tongliang Liu.** He is working in the fields of trustworthy ML. His work in theories and algorithms of ML with noisy labels has led to significant contributions and influence in the fields of ML, computer vision, natural language processing (NLP), and data mining, as large-scale datasets in those fields are prone to suffering severe label errors. He is a senior lecturer at the School of Computer Science, University of Sydney, and a visiting associate professor at the Department of Machine Learning, Mohamed bin Zayed University of Artificial Intelligence.
- **Liqiang Nie.** He is the dean of and a professor with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen). He works on multimedia content analysis and search, with a particular emphasis on data-driven multimodal learning and knowledge-guided multimodal reasoning. He pioneered the explicit modeling of consistent, complementary, and partial alignment relationships among modalities.
- **Soujanya Poria.** He is an assistant professor at Singapore University of Technology and Design (SUTD). His seminal research on fusing information from textual, audio, and visual modalities for diverse behavioral and affective tasks significantly improved systems reliant on multimodal data, paving the way to various novel research avenues. His latest works are on information extraction, vision–language reasoning, and understanding human conversations in terms of common sense-based, context-grounded causal explanations.
- **Deqing Sun.** He is a staff research scientist at Google. He has made significant contributions to computer vision, in particular in motion estimation. His work on optical flow ("Classic+NL" and "PWC-Net") has been very influential and has been powering commercial applications such as Super SloMo in NVIDIA's RTX platform, Face Unblur, and Fusion Zoom on Google's Pixel phone.
- **Yizhou Sun.** She is a pioneer in heterogeneous information network (HIN) mining, with a recent focus on deep graph learning, neural symbolic reasoning, and providing neural solutions to multiagent dynamical systems. Her work has a wide spectrum of applications, ranging from e-commerce, health care, and material science to hardware design. She is currently an

1541-1672 © 2023 IEEE
Digital Object Identifier 10.1109/MIS.2023.3252919
Date of current version 4 April 2023.

associate professor at the University of California, Los Angeles (UCLA).

- **Jiliang Tang.** He is a University Foundation Professor at Michigan State University. He works on graph ML and trustworthy AI and their applications in education and biology. His contributions to these fields include highly cited algorithms, well-received systems, and popular books.
- **Zhangyang "Atlas" Wang.** He works on efficient and reliable ML. Recently, his core research theme is to leverage, understand, and expand the role of sparsity, from classical optimization to modern neural networks (NNs), whose impacts span the efficient training/inference of large-foundation models, robustness and trustworthiness, generative AI, graph learning, and more.
- **Hongzhi Yin.** He has worked on trustworthy data intelligence to turn data into privacy-preserving, robust, explainable, and fair intelligent services in various industries and scenarios. He is also a leading expert researching and developing next-generation intelligent systems and algorithms for lightweight on-device predictive analytics as well as recommendation and decentralized ML on massive and heterogeneous data. He is an associate professor and ARC Future Fellow at the University of Queensland.
- **Liang Zheng.** He is a senior lecturer at the Australian National University and works on data-centric computer vision, where he seeks to improve the quality of training and validation data, predict test data difficulty without labels, and more. These efforts provide a complementary perspective to model-centric developments. He has also made significant contributions to object re-identification and the broader smart city initiative through the introduction of widely used benchmarks and baseline methods.

IEEE Intelligent Systems thanks the selection committee members, Thomas Eiter (Technical University of Vienna, Austria), Niloy Ganguly (Indian Institute of Technology, Kharagpur, India), Sarit Kraus (Bar-Ilan University Ramat Gan, Israel), Eugene Santos Jr. (Dartmouth College, USA), and Fei Wu (Zhejiang University, China), who devoted many hours studying the nomination materials and discussing and selecting the best young members of our AI community.

We would like to warmly congratulate these awardees—they are truly outstanding as bright rising stars in their research areas and truly deserve this honor. We would also like to mention that the final selection result very much reflects the pool of submissions: there were an overwhelming number of nominations from the ML, data mining, and computer vision areas from Australia and the United States, and only very few from Europe and other parts of the world. While the public typically and often thinks of AI as deep learning, we all know that this is not true. We would, therefore, like to encourage all young scientists working in all the areas of AI—in particular, those underrepresented, less popular, classic areas of AI—in all parts of the world to consider applying in the next edition. Specifically, for those candidates who were nominated this year and were fewer than 5 years after receiving their Ph.D., please do reapply. The selection was based on overall cumulative achievements, not on the ratio of "achievements per time."

—Jürgen Dix and Zhongfei (Mark) Zhang, Selection Committee cochairs and Editorial Board members

2022 AI's 10 TO WATCH

Bo Li, University of Illinois at Urbana–Champaign

Dr. Bo Li is an assistant professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign. She is the recipient of the IJCAI Computers and Thought Award; Alfred P. Sloan Research Fellowship; National Science Foundation (NSF) CAREER Award; Massachusetts Institute of Technology (MIT) Technology Review TR-35 Award; Dean's Award for Excellence in Research; C.W. Gear Outstanding Junior Faculty Award; Intel Rising Star Award; Symantec Research Labs Fellowship; Rising Star Award; research awards from tech companies such as Amazon, Facebook, Intel, Google, and IBM; and best paper awards at several top ML and security conferences. Her research focuses on both theoretical and practical aspects of trustworthy ML, security, ML, and game theory. She has designed several scalable frameworks for robust learning and privacy-preserving data publishing. Her work has been featured by major publications and media outlets, such as *Nature*, *Wired*, *Fortune*, and *The New York Times*. Her website is http://boli.cs.illinois.edu/.

Trustworthy ML

As intelligent systems become pervasive, safeguarding their security and trustworthiness is critical. For instance, our prior work has shown that adversarially manipulating the perceptual systems of autonomous vehicles (AVs) may lead to misreading road signs, with possibly catastrophic consequences; commercial face recognition application programming interfaces (APIs) can be attacked to misrecognize faces with potential economic losses; and learned personal information could be leaked, among other trustworthiness issues. The goal of our research is to design trustworthy ML algorithms and systems for diverse real-world applications. Specifically, our research focuses on three interconnected perspectives—robustness, privacy, and generalization—and their underlying connections as well as bringing trustworthy ML to real-world applications. Our work along this direction has provided fundamental understanding, novel theories, effective algorithms, and a series of open source toolkits and benchmarks. For instance, we have designed certification for robustness, generalization, and fairness; uncovered the bidirectional indications between robustness and generalization; and characterized the relationship between robustness and privacy.

In terms of robustness, many ML methods, in practice, assume that the training and testing data distributions are similar without considering active adversaries seeking to manipulate either distribution, leading to various train-time (poisoning) and test-time (evasion) attacks. Our work aims to tackle the ML robustness problem from three perspectives: threat model exploration, game-theoretic and knowledge-enabled robustness, and certified robustness for general ML paradigms. In particular, we have 1) performed a series of black-box attacks against real-world APIs, 2) generated unrestricted adversarial attacks by manipulating semantic spaces for both image and text data, 3) demonstrated physical attacks against AVs and software testing systems (our work on generating the first physical adversarial perturbations on road signs has become a permanent collection in the Science Museum of London), and 4) bridged data-driven learning and logical reasoning to improve the certified robustness of ML against adversarial attacks.

Data privacy is an important aspect of trustworthy ML. As existing and our prior works show that credit card numbers can be pulled out of language models (LMs) and faces reconstructed from trained models, our research aims to protect data privacy, focusing on three aspects: unified privacy attack framework exploration, privacy-preserving data generation, and privacy-preserving learning.

As robustness and privacy mainly focus on data distribution shifts under an adversarial setting, ML generalization—a never-ending pursuit of the ML community—tackles such distribution mismatches appearing in a natural setting. Toward analyzing and improving ML generalization, our work uniquely focuses on 1) uncovering the underlying connections between generalization and robustness/privacy and 2) certifying ML generalization. Specifically, one of our works has proven that adversarial transferability (a robustness property) and domain transferability (a generalization property) are bidirectional indicators for each other, which has great implications for a range of applications, such as pretrained model selection. We have also provided the necessary conditions of model generalization to characterize its connection with robustness and bridge different learning paradigms. In addition, we have generalized our analysis on certified robustness to certify ML generalization given a bounded distribution shift.

Our research not only focuses on the main aspects of trustworthy ML but also pays attention to their underlying connections. For instance, we have proven that models with tighter privacy guarantees are more certifiably robust and that more generalizable models will achieve higher robustness and privacy guarantees.

Complementing the outlined research, we have also made contributions to trustworthy quantum ML, uncertainty estimation, fairness, and data valuation, among other topics, each of which grapples with the central trustworthy ML problem. Our approaches for building trustworthy ML systems serve as a back end to many applications, such as safe autonomous driving, trustworthy LMs, resilient federated learning, and knowledge-enabled certifiably robust learning systems for cybersecurity tasks.

Tongliang Liu, University of Sydney

Tongliang Liu is the director of the Sydney AI Centre and is a senior lecturer in ML with the School of Computer Science at the University of Sydney. He is also an associate professor in ML with the Mohamed bin Zayed University of Artificial Intelligence. He received

his Ph.D. from the University of Technology Sydney. His research interests lie in understanding and designing ML algorithms for problems in the field of trustworthy ML. He received the ARC Future Fellowship Award in 2022, was named in the Early Achievers Leaderboard by *The Australian* in 2020, and received the ARC Discovery Early Career Researcher Award (DECRA) Fellowship Award in 2018. He regularly serves as area chair for ICML, NeurIPS, and ICLR. He is an action editor of *Transactions on Machine Learning Research* and serves on the editorial board of *Journal of Machine Learning Research* and *Machine Learning*. Contact him at tongliang.liu@sydney.edu.au.

Learning With Noisy Labels

Modern ML is migrating to the era of complex models (e.g., deep NNs), which usually require a plethora of well-annotated data. However, it is expensive and sometimes even infeasible to accurately label large-scale datasets. Small businesses and nonprofit organizations may only have access to cheap datasets, which contain heavy label errors. Label errors not only extensively exist in industry but also widely exist in the datasets frequently used in academia. For example, in the well-known ImageNet, 6% of labels are incorrect in the validation set. As the complex models will easily overfit label errors and have significant performance degeneration, label noise now has become a serious issue impeding the development, deployment, and trustworthy of AI.

The learning algorithms dealing with noisy labels developed in my research group can be divided into two categories: algorithms that result in statistically inconsistent or consistent classifiers. Methods in the first category usually employ heuristics to reduce the side effect of label noise—for example, extracting confident/incorrect examples (whose labels are more likely to be correct/incorrect) by exploiting empirical properties of the model and data. Methods in the second category aim to design statistically consistent algorithms, where classifiers learned by exploiting noisy data will converge to the optimal classifiers defined on the clean domain.

Our long-term mission is twofold: 1) to detect and correct label errors to contribute to data health and AI sustainability and 2) to enable industry partners to directly rely on data with noisy labels. The research on learning with noisy labels is at its early stage. There are many unsolved challenging problems. For example, most of the existing work focus on preventing models from overfitting label noise during the training procedure. However, very little research has been done for model selection when the validation datasets contain label errors. The method of modeling label noise will make the classification algorithms robust to the data generation procedure and label noise type. Although there are many research works on modeling label noise—e.g., we have introduced the anchor point assumption and the sufficiently scattered assumption for identifying and estimating class-dependent label noise—how to identify and estimate the instance-dependent, annotator-dependent, or incentive-dependent label noise remain elusive when only noisy data are available because the truth labels are unavailable. We believe the work in theories and algorithms of ML with noisy labels will lead to significant contributions and influence in many fields, including computer vision, NLP, and data mining, as large-scale datasets in those fields are prone to suffering severe label errors.

Liqiang Nie, Harbin Institute of Technology (Shenzhen)

Dr. Liqiang Nie, fellow of IAPR, is the dean of the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen). He received his B.Eng. and Ph.D. from Xi'an Jiaotong University and National University of Singapore, respectively. His research interests lie primarily in multimedia content analysis and information retrieval. Dr. Nie has published more than 100 papers and five books at the first-tier venues, with 17,000+ Google citations. He is an area chair or senior program committee member of *IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology*, and the Association for Computing Machinery (ACM) *Transactions on Multimedia Computing, Communications, and Applications*. Meanwhile, he is the regular AC or SPC of ACM MM, NeurIPS, ICML, IJCAI and AAAI. He is a member of the IEEE ICME steering committee. He has received many awards over the past years, like SIGMM rising star in 2020, MIT TR35 China 2020, DAMO Academy Young Fellow in 2020, SIGIR best student paper in 2021, and ACM MM best paper award in 2022.

Multimodal Correlation Learning and Reasoning

The widespread availability of cheap hardware media technologies, like digital cameras and mobile devices, has exponentially increased the volume of multimedia data, such as videos and animations. Meanwhile, the emergence of social platforms, like Twitter and TikTok, has sped up the multimedia data to be disseminated and shared to a wide audience of users. Multimedia data consist of data in multiple modalities, including text, images, videos, audio, and time sequences, which allow fast and effective expression and communication in our daily lives. As it is true for human perception that we gather information from different sources in natural and multimodality forms, learning from multimodalities has become a long-standing research problem.

Rather than independence, these multiple modalities are usually correlated in a complex and sophisticated way. Therefore, modeling the interactions among them is the key for multimedia content representation, categorization, captioning, and search. In view of this, my research interests primarily lie in multimodal correlation learning and reasoning, with emphasis on data-driven multimodal learning and knowledge-guided multimodal reasoning.

I have to mention that several efforts have been dedicated to multimodal learning via either early fusion, late fusion, or common subspace learning. They, however, overlook the specific and explicit correlations among multiple modalities. Instead, my team and I explicitly clarify the consistent, complementary, and partial alignment relationships among modalities. We propose a series of data-driven multimodal learning models, such as multimodal subspace learning, adaptive multimodal cooperative learning, and multimodal explainable attribute-level learning. These models fill the theoretical gap in this field and have been justified over the task of microvideo recommendation and search, yielding 32.5% performance improvement regarding Macro-F1.

Data-driven multimodal learning models toward explicit multimodal correlations have achieved great success over the past decade and are associated with ML algorithms, like deep NNs, to extract nonexplainable features and patterns through the training process. They, however, heavily rely on well-labeled data. As a supplement to the data-driven models, the knowledge-guided paradigm has been gaining momentum in many research communities. Inspired by this, my team and I are pioneers of leveraging prior knowledge or rules to enhance multimodal reasoning. To be more specific, we design and justify a series of original multimodal reasoning approaches, including but not limited to hierarchy-regularized multimodal hashing, graph-constrained multimodal multitask learning, and multimodal learning with probabilistic knowledge distillation. These models greatly alleviate the data dependence and strengthen the model generalization and reasoning abilities.

Soujanya Poria, SUTD

Soujanya Poria is an assistant professor at SUTD. He earned his Ph.D. in computer science from the University of Stirling, U.K., and was honored with the prestigious Nayang Technological University Presidential Postdoctoral Fellowship in 2018. Prior to his fellowship, he served as a scientist at A*STAR and Temasek Laboratory, NTU. With more than 100 published papers and articles in leading conferences and journals such as ACL, EMNLP, AAAI, NAACL, ECCV, *Neurocomputing*, *IEEE Transactions on Affective Computing, IEEE Computational Intelligence Magazine*, *IEEE CIM*, and *Information Fusion*, his cutting-edge research has been highly cited and received substantial funding from both the government and industry. His research has been recognized internationally, including with the IEEE CIM Outstanding Paper Award and ACM ICMI Best Paper Award Honorable Mention. He has held prominent roles at numerous conferences and workshops, including serving as area cochair at ACL, NAACL, and EMNLP and as workshop chair at AACL 2022. He had given invited talks at events such as CICLing 2018, SocialNLP 2019, MICAI 2020, and ICON 2020. Currently, he is serving as an associate editor for *Cognitive Computation*, *Information Fusion*, and *Neurocomputing*.

Multimodal AI and NLP

A large quantity of online user-generated data is multimodal—data that include visual, audio, and text channels. This content is highly valued by enterprises for its rich information that can be applied to various purposes, such as enhanced user engagement for better recommender systems and AdSense. However, effectively analyzing such data requires the integration of information from multiple modalities, which presents significant technical challenges for the development of AI agents. Our research group is focused on advancing

the field of AI by developing cutting-edge techniques aimed at solving complex multimodal tasks such as emotion recognition. We focus on addressing the issue of multimodal information fusion through representation learning and mutual information maximization. We have demonstrated that fusing information from multiple modalities leads to improved performance compared to unimodal systems. The open source codes we developed have been widely adopted by both academia and industry.

Multimodal data can contain corrupted channels due to intermediate noise. Additionally, not all modalities carry equal amounts of information, making it important to develop robust systems that can effectively handle such challenges. To address these issues, we are actively researching robust multimodal ML algorithms. Our research leverages state-of-the-art techniques, including modality imputation using optimal transports, denoising of inputs, and making the backbone network more robust through the use of modality dropouts. Our ultimate goal is to develop robust multimodal approaches that enhance the performance of nonrobust systems.

We have been actively working to address the challenge of commonsense reasoning in NLP. The literature demonstrates that LMs exhibit subpar performance in this area, motivating us to focus on developing AI models and tasks specifically tailored for contextual commonsense reasoning. We incorporated commonsense knowledge into deep learning models to enhance their performance on various downstream tasks, such as emotion recognition, dialogue understanding, and sentence ordering. We also introduced a novel and challenging commonsense reasoning task that requires AI models to answer causal questions by leveraging in-context speculation and creative thinking.

Over the past few years, our team has made significant contributions to the field of dialogue understanding. We focused on key problems, such as extracting implicit knowledge triplets from dialogues and emotion recognition in conversations. These tasks are important for enterprises that rely on chatbots for customer interactions, as they aid in better understanding of conversations and, ultimately, lead to improved customer engagement. To tackle these challenges, we developed several open source dialogue context modeling algorithms using advanced techniques, such as transformers and graph neural networks (GNNs). Additionally, we created multiple large-scale datasets, which have helped to establish this research direction as a key subfield of dialogue system research.

Currently, we are tackling the following pressing issues in the field of AI: 1) the time-consuming process of annotating data for supervised learning and 2) the environmental impact of running large AI models, which have high carbon emissions from GPU utilization. To combat these challenges, we have been actively exploring resource- and parameter-efficient techniques. Our efforts have resulted in the development of innovative approaches, including the use of contextual prompts for language understanding, LM prompting for dataset augmentation for zero-shot NLP, and the deployment of adapters to improve the performance of language models in domain adaptation.

Deqing Sun, Google

Deqing Sun is a staff research scientist at Google. He received a Ph.D. in computer science from Brown University. He has made significant contributions to computer vision, particularly in motion estimation. His work on optical flow ("Classic+NL" and "PWC-Net") has been very influential in the research community and powers commercial applications such as Super SloMo on NVIDIA's RTX platform, Face Unblur, and Fusion Zoom on Google's Pixel phone. He served as an area chair for CVPR/ICCV/ECCV and co-organized several workshops/tutorials at CVPR/ECCV/SIGGRAPH. He is a recipient of the Best Paper Honorable Mention award at CVPR 2018, the Best Paper Finalist at CVPR 2022, the PAMI Young Researcher award in 2020, and the Longuet–Higgins prize at CVPR 2020. Contact him at deqingsun@google.com.

Learning to Perceive the Dynamic 3-D World

With seemingly little effort, we perceive the world around us from light projected from surfaces onto photosensors in the retina. Indeed, our ability to understand and interact with the 3-D world and, in doing so, infer the shapes, sizes, and motions of objects from images, is truly a remarkable property of human vision. It seems so easy to reach out and catch a baseball or predict whether a stack of books on a table is about to topple or whether a car is going to veer left or right. While deep learning methods have surpassed humans on somewhat artificial tasks, such as the "ImageNet classification benchmark" and the generation of stunning images

from text prompts, enabling computers to infer the 3-D structure of the world and the dynamics of objects from 2-D images remains a major challenge and the topic of extensive research.

One of my main interests is motion estimation from image sequences, also known as optical flow, which tells us how objects move and interact and enables us to discover new objects. My collaborators and I introduced the first fully learnable optical flow model in 2008. However, it underperformed hand-designed methods due to inadequate data, shallow architectures, and non-end-to-end training. A decade later, we developed a compact and effective architecture using classical optical flow principles as an inductive bias. It was the first deep learning method to surpass hand-designed methods and won first place in the optical flow challenge by performing robustly on four benchmarks. Recently, we compared several deep learning architectures using the same training protocol and found that older architectures were still competitive when trained with recent datasets and techniques. This analysis helped identify new research avenues, such as learning to render synthetic datasets, which can make synthetic data more valuable for training AI models.

The advances in optical flow have enabled progress on other tasks, such as reconstructing the shapes of articulated objects, e.g., people and animals, from monocular videos. Prior approaches typically relied on a category-specific template, e.g., people with tight clothing, and could fail when inputs violated these assumptions, e.g., people with loose clothing. Instead, we took a template-free, analysis-by-synthesis approach that enforced renderings from the 3-D shape, articulation, and camera pose to be consistent with observed texture, segmentation, and optical flow. Our approach reconstructed articulated shapes faithful to the input for categories in the wild and may eventually make casual videos into editable 3-D assets for every object in the world.

Our work in optical flow powers commercial products like Super SloMo on NVIDIA's RTX platform, Face Unblur, and Fusion Zoom on Google's Pixel phone, which have helped popularize optical flow as a core video-processing tool. With increasing amounts of videos captured and created daily, new motion representations could unlock the potential of videos without requiring massive amounts of human labeling, similar to how infants might learn the visual world from motion.

We have also made progress on other fundamental vision tasks, such as the efficient processing of 2-D images and 3-D point clouds using a sparse lattice network; robust visual recognition via pyramid adversarial training; and the joint estimation of depth, camera, optical flow, and segmentation through geometric constraints. I believe that a unified approach to visual perception will provide a foundation to enable intelligent systems to adapt to new tasks flexibly.

Yizhou Sun, University of California, Los Angeles

Yizhou Sun is an associate professor at the Department of Computer Science of UCLA. She received her Ph.D. in computer science from the University of Illinois at Urbana-Champaign in 2012. Her principal research interest is mining graphs/networks and, more generally, data mining, ML, and network science, with a focus on modeling novel problems and proposing scalable algorithms for large-scale, real-world applications. She is a pioneer researcher in mining HINs, with a recent focus on deep learning on graphs/networks. Contact her at yzsun@cs.ucla.edu.

Graphs as a "Middleware" for Modern AI

Graphs are such a powerful data structure and are widely used to represent 1) data, such as molecules and e-commerce systems; 2) symbolic knowledge, such as knowledge graphs; 3) multiagent dynamical systems, such as *n*-body systems; and 4) computational engines, such as NNs. My research focuses on mining and learning from these graphs.

In earlier days, my research concentrated on data-related graphs. Specifically, I created a new research field called HINs, which are graphs with different types of nodes and different types of edges. These graphs are ubiquitous in the real world, ranging from e-commerce to health care. My seminal paper, "PathSim" (VLDB'11), defined HINs for the first time and proposed metapaths to systematically capture different types of semantic relationships between objects in the graphs. This research has inspired considerable follow-up work in both academia and industry and has had a profound impact on a wide range of applications, such as drug discovery, recommender systems, and cybersecurity. Many companies have embraced the concept of HINs and/or metapaths in their products and systems,

including Microsoft, Amazon, Meta, Alibaba, and Twitter. Recently, the "PathSim" paper won VLDB's 2022 Test of Time Award, awarded to the most influential papers published 10–12 years ago.

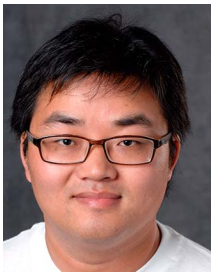
Recently, I have been working on graph representation learning and GNNs and have made significant contributions to a wide spectrum of aspects of GNNs, including 1) heterogeneous graph representation learning [HGT (WWW'20)] to handle heterogeneity, 2) pretraining GNNs to handle label scarcity [GPT-GNN (KDD'20)], 3) novel sampling techniques to accelerate training runtime and reduce memory cost [LADIES (NeurIPS'19)], 4) knowledge distillation techniques to significantly reduce the inference time of GNNs [GLNN, (ICLR'22)], and 5) addressing the explanability of GNN models [GStarX (NeurIPS'22) and PaGE-Link (WWW'23)].

More recently, I am obsessed with graphs that are beyond data representation. By studying knowledge graphs, I am working on integrating traditional symbolic reasoning and modern representation learning. I would like to combine the two worlds from a graph perspective. To do so, the first step is to build connections between logic expressions and graph terminologies. For example, a "triple" in knowledge graph (KG) is a "predicate" in a logic formula, and a "conjunction of predicates" is a "path" in KG. Based on this understanding, we developed algorithms that combine logical reasoning and representation learning for KG completion [UniKER (EMNLP'21)], first-order logic KG queries [FuzzQE (AAAI'22)], and logical rule mining [RLogic (KDD'22) and NCRL (ICLR'23)]. We have also successfully integrated KG reasoning with large LMs for open-domain question answering and answers [OREO-LM (EMNLP'22)], which requires many fewer parameters to achieve a comparable performance and is able to quickly absorb new knowledge without retaining LMs.

I am also fortunate to work with many collaborators for very novel applications where graphs are related to dynamical systems or computational engines. For example, I am now working with innovative applications that aim to use GNNs to boost material design and hardware high-level synthesis. For the former application, we propose using GNNs to model ordinary differential equations (ODEs) for dynamical systems, called GraphODE, which can be learned with irregularly observed data without the need to fully understand the underlying mechanism [LG-ODE (NeurIPS'20) and CG-ODE (KDD'21)]. For the latter application, preliminary results have already shown that GNNs can predict the performance of a hardware design in milliseconds without running expensive simulators. The technique can find designs that are up to 64× faster with much shorter search time [GNN-DSE (DAC'22) and GNN-DSE-MAML (DAC'22)].

I envision that future AI will be more intelligent with graphs as a middleware, which will be better in understanding the world, simulating complex systems, creating logic-consistent content, and accelerating innovations in providing reliable design solutions.

Jiliang Tang, Michigan State University



Jiliang Tang is a University Foundation Professor in the Computer Science and Engineering Department at Michigan State University. He was an associate professor (2021–2022) and an assistant professor (2016–2021) in the same department. Before that, he was a research scientist at Yahoo Research and got his Ph.D. from Arizona State University in 2015. His research interests include ML on graphs and their applications in social media and biology. He authored the first comprehensive book, *Deep Learning on Graphs*, with Cambridge University Press and developed various well-received open source tools, including scikit-feature for feature selection, DeepRobust for trustworthy AI, and DANCE for single-cell analysis. He was the recipient of various awards, including the 2022 IAPR J. K. Aggarwal Award, 2022 SIAM/IBM Early Career Research Award, 2021 IEEE ICDM Tao Li Award, 2021 IEEE Big Data Security Junior Research Award, 2020 ACM SIGKDD Rising Star Award, 2019 NSF CAREER Award, and eight best paper awards (or runners-up). He has published his research in highly ranked journals and top conference proceedings; these have received tens of thousands of citations, with an h-index of 77 and extensive media coverage. Contact him at tangjili@msu.edu.

Data Mining and ML on Graphs

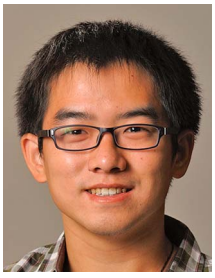
Graphs, which abstract complex systems of relations and interactions, provide a universal representation for a wide range of real-world data, including social networks, transportation networks, transaction networks, integrated circuits, and chemical molecules. In the past decade, my research emphasis has primarily been on harnessing this natural structure of data through mining, learning, and optimization. More specifically, my research strives to 1) develop fundamental principles to identify actionable patterns and insights and

extract knowledge from graphs, 2) build computational and trustworthy graph methods to advance various real-world applications, and 3) design systems and tools to enable researchers with limited or no expertise to perform necessary data analysis via graph methods.

Feature selection has been proven to be an effective and efficient way to prepare data for mining and learning. Before my work, existing feature selection methods were designed for independent identically distributed data. However, data are often associated with graphs. I proposed original problems and developed innovative algorithms, established a new research direction of feature selection on graphs, and opened doors to new research opportunities. I led a team to build a repository that is the most famous feature selection system. To further facilitate graph mining and learning, I was one of the early researchers on graph embedding. I was the first to provide ingenious solutions—employing social theories to explain phenomena and adapt deep learning to capture complicated patterns. These solutions offer a principled and flexible way to develop algorithms for numerous types of graphs. For example, I was the first to introduce deep architectures for heterogeneous graph embedding, and I adopted the social balance theory to build start-of-the-art signed network embedding. Moreover, its flexibility allows me to introduce graph embedding to advance a boarder range of real-world applications. For example, I introduced bipartite graph embedding to web search and utilized signed network embedding for recommendations.

Due to their great promise to intelligently advance innumerable real-world applications, recently, GNNs have emerged as a new frontier of data mining and ML in the era of deep learning. I has significantly contributed to developing GNNs from new perspectives. I proposed a unified framework from the perspective of graph signal denoising, which covers many existing GNNs as special cases, bridging the gap between multiple proposed GNNs. The significance of this work is twofold. First, it provides a unified understanding of existing graph filters. Second, it paves the way for us to design graph filters via optimization. Though GNNs are powerful in learning representations of graph-structured data and have permeated numerous areas of science and technology, there are also limitations. For example, GNNs are often treated as black boxes and lack human-intelligible explanations; they are easily fooled by adversarial attacks, and they have algorithmic biases toward certain groups when making these decisions. Without mitigating these limitations, GNNs cannot be fully trusted. It will prevent their use in critical applications pertaining to fairness and safety, such as autonomous driving and health care. Thus, I have done important research on trustworthy GNNs. I published the first comprehensive book in this area and released two well-received systems to build trustworthy GNNs (i.e., DeepRobust) and advance single-cell analysis via GNNs (i.e., DANCE).

Zhangyang "Atlas" Wang, University of Texas–Austin



Prof. Zhangyang "Atlas" Wang is the Jack Kilby/Texas Instruments Endowed Assistant Professor in the Chandra Family Department of Electrical and Computer Engineering at the University of Texas (UT) at Austin. He is also affiliated with UT Computer Science and the Oden Institute. He received his Ph.D. in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2016 and his B.E. in electronic engineering and information science from the University of Science and Technology of China in 2012. He has broad research interests spanning from the theory to the application aspects of ML. At present, his core research mission is to leverage, understand, and expand the role of sparsity, from classical optimization to modern NNs, whose impacts span many important topics, such as efficient training/inference/transfer of large-foundation models, robustness and trustworthiness, learning to optimize, generative AI, and graph learning. He has received many awards and honors, including an NSF CAREER Award, an ARO Young Investigator Award, an IEEE "AI's 10 to Watch" Award, an INNS Aharon Katzir Young Investigator Award, a Best Paper Award from the inaugural LoG Conference 2022, and several more industry faculty research awards as well as research competition prizes. He is an ACM Distinguished Speaker and a Senior Member of IEEE.

Sparse NN: Simplicity Is the Final Achievement

Recent breakthroughs in deep NNs have fueled a growing demand for intelligent edge devices. However, many real-world applications require real-time inference and in situ learning. The limited computing and energy resources available at the edge stand at odds

with the massive and growing learning costs for state-of-the-art deep NNs. My group has engaged extensive research efforts on efficient training and inference algorithms for deep networks. At the core of our methodology is leveraging understanding, and expanding the role of sparsity in NNs: one of the longest-standing concepts in ML.

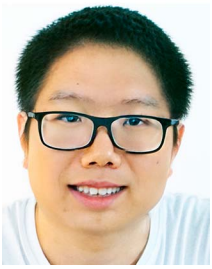
Substantial efforts have been devoted to scaling deep NNs to enormous sizes, and parameter counts are frequently measured in billions. Sparse NNs, whose large portions of parameters are zero, have been studied to address those gaps. Early approaches first train dense NNs and then prune the trained NNs to high levels of sparsity. Those methods significantly reduce the inference complexity yet cost even greater computational resources and memory footprints at training. An emerging subfield has explored the prospect of directly training smaller, sparse subnetworks in place of the full models without sacrificing task performance. Over the past few years, my group has contributed many works that lay both empirical and theoretical foundations for the efficiency, optimization, and generalization of sparse NNs.

On the empirical side, recent findings demonstrated that standard dense NNs contain sparse subnetworks (called "winning tickets") capable of training in isolation to full accuracy. While finding such winning tickets was initially expensive and impractical, this key hurdle was addressed by a series of my works. First, my group revealed the existence of "universal winning tickets": one can identify the winning ticket from a pretrained model and then directly transfer this sparse submodel to various downstream tasks. In this way, the extraordinary cost of winning ticket finding is amortized by its reusability. This finding can make gigantic pretrained models accessible to TinyML applications. We later found that those winning ticket subnetworks can be structured in a hardware-friendly way, further closing the practical gap.

Next, my group discovered that the winning ticket structure can be found very early in dense NN training, sometimes even without any training. The sparse NN performance can improve more if its sparse mask is dynamically changed during training, which is also biologically plausible. Those efforts together make it much more efficient to find good sparse NNs in general, yielding remarkable computation and energy savings by directly training sparse NNs in place of dense ones. Besides those significant efficiency gains, my group has also empirically observed the competence of sparse NNs in data-efficient learning, communication-efficient federated learning; robust learning under data noise or attacks, and diverse applications beyond computer vision and natural language processing (even multitask and multimodal learning). This strong evidence suggests sparse NNs to be a "winning-all" key knob for future deep learning and AI.

My group has, meanwhile, strived to establish the theoretical foundations underlying sparse NNs, and we have achieved some of the first results in this new frontier of deep learning theory. In one recent paper, my student and I study the behavior of ultrawide NNs when their weights are randomly pruned at the initialization through the lens of neural tangent kernels. We show that, like the layer width, the optimization and generalization of the pruned sparse NNs will approach and converge to those of the original dense NN. Other recent works have demonstrated a favorable generalization bound and transferability result of NNs that involve sparsity-dependent parameters.

Hongzhi Yin, University of Queensland



Hongzhi Yin is an associate professor at the University of Queensland's School of Information Technology and Electrical Engineering. He received his Ph.D. in computer science from Peking University. He has made notable contributions to the fields of predictive analytics, recommendation systems, and decentralized machine intelligence and has received numerous awards and recognition for his research achievements. He was named Field Leader of Data Mining and Analysis in the *Australian Research* 2020 magazine and received the 2022 AI 2000 Most Influential Scholar Honorable Mention in Data Mining. In addition, he has received the Research.com Rising Star of Science Award 2022, Australian Research Council Future Fellowship 2022, DECRA 2016, and the University of Queensland Foundation Research Excellence Award 2019. He was featured among the 2022 and 2021 Stanford's World's Top 2% of Scientists (career-long and single-year impacts). His research work has won the Best Paper Award at the 35th IEEE International Conference on Data Engineering (ICDE 2019), Best Student Paper Award at the 25th International Conference on Database Systems for Advanced Applications (DASFAA 2020), Best Paper Award nomination at the 20th International Conference

on Data Mining (ICDM 2018), and Peking University Distinguished Ph.D. Dissertation Award 2014. Contact him at db.hongzhi@gmail.com.

Decentralized Collaborative Learning for Trustworthy Predictive Analytics

With the fast development of wireless communication and mobile chip techniques, connected user devices, such as smartphones, smart watches, wearable devices, drones, and other personal edge devices, are now widespread and unprecedentedly powerful. Moreover, these devices constantly collect data about where we go; what we eat, purchase, and see; how active we are; our temperature; etc. It, therefore, makes sense that we should harness all of this computing power and the data we are collecting to solve the world's biggest and most wicked problems, such as health, privacy, economic depression, etc.

Currently, the most common predictive analytics approach is building a massive data center that collects, stores, and processes data, i.e., cloud computing. This centralized paradigm is convenient for training ML models, but its drawbacks are becoming increasingly evident, especially as the Internet of Things (IoT) becomes more ubiquitous. For service providers, the issues include the high infrastructure costs associated with gathering, storing, and analyzing data on an enormous central server. For users, the disadvantages include the privacy risks that come with limited control over personal data. Another inherent limitation of the centralized paradigm is the high latency caused by communication between IoT devices and central servers, especially when the communication is unstable. Further, a subtler risk raising public concern is that the global or central models can be biased. This is because most models achieve high accuracy by favoring the majority classes to the exclusion of outliers. However, when outlying classes reflect marginalized populations (like, say, patients over 65 or Aboriginals), in a blood sugar-monitoring app, that bias can become dangerous or even discriminatory.

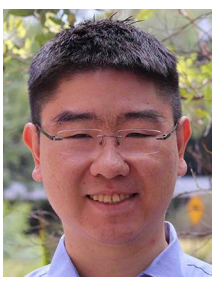
In light of these issues, my research over the past few years has mainly focused on designing and developing a new learning paradigm, decentralized collaborative learning of personalized models, where a large number of connected personal devices, each acting as an agent, collaborate to learn their own personal predictive analytics model in an environment that offers protection against adversaries. Compared to existing learning schemes, this new learning paradigm has the following advantages: 1) fairness by design in the realm of personal matters, as nothing is fairer than a model built solely for you; 2) the ability to self-organize a dynamic collaboration network; 3) lightweight model representation and optimization mechanisms small enough to operate on resource-constrained devices; 4) no single point of vulnerability, making the scheme robust to failure and attacks; 5) scalability by design up to a massively large number of devices; and 6) full security against privacy leaks and adversary attacks.

To implement such a new learning paradigm for trustworthy predictive analytics, we have addressed numerous technical challenges in the following aspects.

- Efficient model optimization and representation on resource-constrained devices.
- Self-organizing dynamic collaboration networks.
- Decentralized collaborative learning across heterogeneous devices and models.
- Secure decentralized collaborative learning against adversary attacks (privacy inference attacks and model poisoning attacks).
- Decentralized machine unlearning for the right to be forgotten.

Our research outcomes translate to Web 3.0 applications in smart health care, e-tourism, digital economy, online services, and manufacturing, representing significant new market advantages.

Liang Zheng, Australian National University



Dr. Liang Zheng is a senior lecturer at the Australian National University. He is best known for his contributions in object re-identification. He and his collaborators designed widely used datasets and algorithms such as Market-1501 (ICCV 2015), part-based convolutional baseline (ECCV 2018), random erasing (AAAI 2020), and joint detection and embedding (ECCV 2020). His recent research interest is datacentric computer vision, where improving leveraging and analyzing and improving data instead of algorithms are of primary concern. He is a co-organizer of the AI City workshop series at CVPR and the Vision Datasets Understanding workshop series at CVPR, and he serves as area chair for important conferences such as CVPR, ICCV, and

ECCV. He is an associate editor for *IEEE Transactions on Circuits and Systems for Video Technology*. He received his B.S. in 2010 and Ph.D. in 2015 from Tsinghua University, China. Contact him at liang.zheng@anu.edu.au.

Datacentric Computer Vision

Models and data are two indispensable pillars in the computer vision area. While people generally focus on model developments, the role of data has not been well studied. Without high-quality datasets, stable and robust system performance cannot be achieved, even with advanced models.

To resolve this significant knowledge gap, I focus on datacentric computer vision, a fundamental problem aiming to analyze and improve various properties of data. My long-term goal is to improve system robustness in real-world computer vision applications by the coupling of data and models that are both the state of the art. On the one hand, I aim to design methods allowing for effective analysis of the quality of training data. It is generally assumed that the training set is fixed, but such fixation will inevitably bias the system toward certain scenarios. If the test distribution is very different, there might be severe performance degradation. Therefore, it would be very interesting to efficiently customize the training data to fit the test distribution, which facilitates the use of computer vision models in extreme environments.

On the other hand, I aim to analyze the difficulty of the test environment, which is a critical problem in the deployment of vision systems because hard environments would lead to system failure. To this end, I study how to find effective dataset representations that, without requiring test ground truths, reflect the distribution difference between the test and training domains. Such representations might come from feature statistics, self-supervised signals, or data simulators. For example, if a vision system is poor at predicting the rotation angle of images, this system might have a poor accuracy in semantic recognition.

Regarding optimizing training data, my collaborators and I developed the random erasing data augmentation technique, an automatic and inexpensive protocol for large-scale data collection. Published at AAAI 2020, this technique has been officially included in Pytorch and has now become a de facto tool for a wide range of AI problems associated with AVs, robotic navigation, speech recognition, and automatic medical diagnosis. I have also been working with students, designing innovative ways of using 3-D simulators to create training data and improve the real-world performance of AI models, including applications for intelligent transport. For 2020–2023, the resulting training datasets have been shared globally through collaborations with NVIDIA for international competitions for AI City applications that have attracted hundreds of the world's leading researchers in this field.

In studying test data difficulty, my students and I proposed a new research problem: evaluating model performance without access to ground truth labels. This problem is important and useful in practice: predicting potential model failure before real harm is caused. Using simulators to create a diverse set of environments, they quantitatively measured how an AI model performs in different environments and demonstrated that the results can be effectively applied to address AI model failure cases. This research has the potential to benefit many safety-sensitive applications, such as AVs.

Previously, in my Ph.D. and postdoctoral research, I made significant contributions to the object re-identification field. I made early attempts at defining the pipeline, dataset, and evaluation metric of this task, which brings significant social benefits, such as improving driving safety, animal husbandry, and smart retail.

JÜRGEN DIX is with the Clausthal University of Technology, 38678, Clausthal, Germany.

ZHONGFEI ZHANG is with Binghamton University, State University of New York, Binghamton, NY, 13902-6000, USA.