Catch, Adapt, and Operate: Monitoring ML Models Under Drift

Abstract

Machine learning systems are increasingly deployed in high-stakes domains such as healthcare, finance, robotics, and autonomous systems, where data distributions evolve continuously. Without robust monitoring and timely adaptation, even high-performing models can degrade silently, compromising reliability, safety, and fairness. Continuous monitoring is therefore an absolute necessity.

While there has been rapid progress in drift detection, test-time and continual adaptation, and the deployment of ML systems at scale, these topics are often studied separately. The **Catch, Adapt, and Operate** workshop brings them together around three themes: **sensing** drift through statistical and representation-based monitoring, **responding** through adaptive and self-supervised updates, and **operating** at scale in production pipelines. By connecting theory, systems, and real-world practice, the workshop aims to build a shared foundation for reliable, fair, and continuously adaptive machine learning under real-world drift.

Workshop Description

Scope. This workshop addresses methods for ensuring the reliability of machine learning systems under *distributional drift*. Such drift may occur as *temporal*, *label/concept*, *domain*, or *representation* shift, each presenting unique challenges for monitoring and adaptation. We organize the program into three complementary themes:

- **Sensing the Drift:** Developing tools to detect when models encounter distributional change, using statistical tests, kernel methods, uncertainty estimation, and representation-based monitoring. This theme focuses on identifying drift early and minimizing false alarms in real-world pipelines.
- **Responding to Drift:** Designing strategies that adapt models once drift is detected, including test-time adaptation, continual and online learning, regularization, and selective prediction. The focus is on maintaining accuracy and stability while avoiding catastrophic forgetting.
- **Operating at Scale:** Extending monitoring and adaptation to large-scale production environments, where heterogeneous data streams, governance requirements, and real-time costs amplify the challenge. This theme emphasizes system design, infrastructure, benchmarks, and protocols for reliable deployment at scale.

Together, these themes establish a unified research agenda for developing robust and trustworthy ML systems in dynamic, non-stationary environments.

Community Impact. Monitoring is a foundational requirement for deploying reliable ML systems. Most real-world failures arise after deployment, when models face shifting distributions, evolving user behaviors, or adversarial conditions. Its importance is amplified in domains such as healthcare, finance, cybersecurity, and autonomous driving. By consolidating research on these challenges, this workshop aims to establish shared practices for responsible and reliable ML deployment, bridging academic advances with industrial needs.

Research Goals and Key Questions. Our primary goal is to consolidate and advance research on reliable ML under drift, bridging theory and practice. We seek to create a forum where statisticians, ML researchers, and systems practitioners engage with domain experts from high-stakes applications. Key questions include:

- **Detection:** How can we reliably detect temporal, covariate, label/concept, and representation drift in real-world pipelines? What are the statistical limits of detection, and how should false alarms or missed detections be managed?
- **Adaptation:** Which strategies such as test-time adaptation, online learning, continual learning, are most effective for adapting under drift, and how can we avoid catastrophic forgetting?
- **Scalability:** How can monitoring and adaptation extend to foundation and multimodal models, where both data and model complexity are vastly larger?
- **Evaluation:** How should we benchmark monitoring and adaptation? Can standardized metrics, datasets, and protocols be defined to capture diverse drift scenarios and ensure fair comparison?
- **Safety and Trustworthiness:** How can monitoring and adaptation improve reliability in safety-critical settings without introducing new risks? What governance, auditing, and accountability mechanisms are needed?
- **Lessons Learned:** What can we learn from failed deployments or unexpected drift phenomena, and how should these lessons guide future research?

By addressing these questions, the workshop will promote a cross-disciplinary agenda that strengthens methodological foundations while producing actionable insights for real-world deployment.

Topics and Call for Papers. We invite contributions spanning monitoring, adaptation, fairness, and governance of ML systems under distributional drift, including (but not limited to):

- **Drift detection and characterization:** Statistical tests, kernel methods, density-ratio estimation, causal or graph-based change detection, sequential monitoring, and early-warning systems for drift identification.
- **Representation and uncertainty:** Calibration, uncertainty estimation, explainability, and representation-based monitoring for identifying concept, covariate, or representation drift.
- **Adaptation and recovery:** Test-time adaptation, online learning, continual learning, domain generalization, and meta-learning approaches for handling evolving data distributions.
- **Fairness and bias under drift:** Methods for maintaining fairness, subgroup robustness, and demographic parity as data or population shifts occur; detection and mitigation of emerging biases during deployment.
- **Governance and auditing:** Frameworks for model monitoring under regulatory and compliance constraints; accountability, audit trails, and explainable intervention mechanisms in high-stakes domains.
- **Reliability at scale:** Systems and infrastructure for large-scale monitoring, data logging, alerting, and automated response in production pipelines.

- **Evaluation and reproducibility:** Benchmarks, datasets, drift taxonomies, standardized metrics, and protocols for evaluating monitoring and adaptation methods.
- **Drift in large and multimodal models:** Challenges of monitoring and adapting foundation and LLM-based systems with heterogeneous input modalities and dynamic contexts.
- **Applications:** Monitoring and adaptation in finance, healthcare, robotics, law, cybersecurity, autonomous driving, climate science, and sensor-based systems.

Beyond the main track, we will host a dedicated track on **Lessons from Failures: Understanding What Did Not Work and Why**. This track highlights methods that fell short, failed experiments, or unexpected challenges, encouraging openness and reducing duplicated effort across the community.

We welcome both **full papers** (ICLR main conference format) and **tiny papers** (up to 4 pages, excluding references). Accepted contributions will be presented as contributed talk, lightning, or poster sessions, and will be made available through both the workshop website and the official conference platform.

This workshop prioritizes **work-in-progress and novel research that has not been previously published at major machine learning venues**.

Previous Related Workshops.

- **DistShift (NeurIPS 2023)** and **Foundation Models in the Wild (ICLR 2024)**. These recurring workshops focus on distribution shifts and have recently pivoted toward foundation models. While they discuss adaptation and evaluation under shift, their emphasis lies in finetuning foundation models on labeled target data rather than general-purpose monitoring and drift response.
- **Model, Adapt Thyself (MAT) (CVPR 2024)** and **Putting Updates to the Test (ICML 2025)**. Both focus on adaptation methods, aligning with our *responding to drift* theme, but overlook the equally critical challenge of *detecting* when adaptation should occur, leaving a gap between detection and response.
- **Challenges in Deploying and Monitoring Machine Learning Systems (ICML 2021, NeurIPS 2022)**. These workshops established monitoring as a research priority, yet no direct follow-up has occurred in the four years since. With the rapid deployment of AI in high-stakes domains, the need for new approaches to monitoring and adaptation has become more urgent. Our workshop revives this agenda for the era of foundation and large-scale models.
- **CLVision (CVPR 2024)**. This workshop addresses continual learning where data streams evolve over time and models adapt to new labeled distributions. In contrast, our workshop targets broader scenarios where drift can be abrupt, unlabeled, or unstructured.
- **Shift Happens (ICML 2022)**. This workshop emphasized dataset contributions for robustness evaluation. While aligned with our *sensing the drift* theme, it focused on benchmarks rather than actionable methods for detection or adaptation.

Taken together, these efforts show growing interest in robustness and adaptation but reveal a persistent gap: no existing venue jointly explores *catching drift* (monitoring and detection) and *responding to drift* (adaptation and intervention). Bridging these perspectives is crucial, without detection, adaptation lacks reliable triggers; without adaptation, detection provides no actionable response. Our workshop addresses this gap by creating a dedicated venue that unifies

these viewpoints and advances the study of reliable and adaptive AI under distributional shift. This workshop directly extends these efforts by unifying detection and adaptation perspectives and by introducing a dedicated **'Lessons from Failures'** track to promote openness and reproducibility

Expected Audience and Participation. With the rapid expansion of AI across sectors such as finance, healthcare, robotics, and other high-stakes domains, the need for reliable monitoring and adaptation has become increasingly critical. Combined with the growing momentum of related workshops in recent years, we anticipate attracting a large and diverse audience. We estimate approximately 300-400 participants, spanning academia, industry, and policy communities worldwide.

We project 120-160 submissions, with accepted contributions presented through a carefully structured program: contributed talks for high-impact work, lightning talks to maximize exposure for emerging ideas, and poster sessions to foster collaborative discussion and knowledge exchange across the community.

Tentative Schedule and Format

Timeline.

- Call for papers released: December 14, 2025
- Paper submission opens: January 2, 2026
- Paper submission deadline: January 30, 2026
- Reviewer bidding and assignment: February 3, 2026
- Review deadline for reviewers: February 23, 2026
- Author notification: March 1, 2026
- Camera-ready deadline for accepted papers: March 9, 2026
- Import workshop program and accepted papers to `iclr.cc`: March 11, 2026
- Workshop schedule and invited speaker information released: April 1, 2026
- Talks videos and poster files uploaded: April 18, 2026
- Workshop day: April 26 or 27, 2026

Format and Program. The workshop will run as a one-day event from 9:00 to 18:00. It will feature seven invited talks, six contributed talks, two lightning sessions, two poster sessions, and one panel discussion.

- **Invited talks:** 20-minute talk followed by 10 minutes of discussion.
- **Contributed talks:** 10-minute talk with 5 minutes of discussion.
- **Lightning talks:** 1-minute pitch per paper to provide broader visibility.
- **Poster session:** 45-minute interactive session for in-depth discussion of accepted papers.
- **Panel discussion:** 45-minute interactive dialogue between experts and participants.

The program is organized around three themes: *sensing drift*, *responding to drift*, and *operating at scale*. The day will conclude with closing remarks and the presentation of the **Best Paper**

Award. The schedule is designed to balance depth and breadth: invited talks offer in-depth perspectives from leading experts; Contributed talk sessions highlight high-impact contributions; lightning talks broaden visibility; and poster sessions maximize interaction. This structure ensures meaningful participation across senior and junior researchers, as well as academia and industry practitioners.

09:00 – 09:15	Opening Remarks: Workshop introduction and goals
Section 1: Sensing Drift	
09:15 – 09:45	Invited Talk 1: Arthur Gretton (20 + 10 Q&A)
09:45 – 10:15	Invited Talk 2: Rahaf Aljundi (20 + 10 Q&A)
10:15 – 10:30	Contributed Talk 1 (10 + 5 Q&A)
10:30 – 10:45	Contributed Talk 2 (10 + 5 Q&A)
10:45 – 11:00	Lightning Session 1 (each paper 1 min flash talk)
11:00 – 11:45	Poster Session 1 + Coffee Break
Section 2: Responding to Drift	
11:45 – 12:15	Invited Talk 3: Chelsea Finn (20 + 10 Q&A)
12:15 – 12:30	Contributed Talk 3 (10 + 5 Q&A)
12:30 – 13:00	Lunch Break
13:00 – 13:30	Invited Talk 4: Taesup Moon (20 + 10 Q&A)
13:30 – 13:45	Contributed Talk 4 (10 + 5 Q&A)
13:45 – 14:00	Lightning Session 2 (each paper 1 min flash talk)
14:00 – 14:45	Poster Session 2 + Coffee Break
Panel Discussion	
14:45 – 15:30	Panel Discussion (45 min)
Section 3: Operating at Scale	
15:30 – 16:00	Invited Talk 5: Anima Anandkumar (20 + 10 Q&A)
16:00 – 16:30	Invited Talk 6: Elahe Arani (20 + 10 Q&A)
16:30 – 16:45	Coffee Break
16:45 – 17:00	Contributed Talk 5 (10 + 5 Q&A)
17:00 – 17:15	Contributed Talk 6 (10 + 5 Q&A)
17:15 – 17:45	Invited Talk 7: Masashi Sugiyama (20 + 10 Q&A)
17:45 – 18:00	Closing Remarks: Future directions, and Best Paper Award

Diversity, Accessibility, and Outreach

Diversity and Inclusion. Our workshop demonstrates strong diversity across **geography**, **sector**, **career stage**, and **gender**. The organizers, speakers, and panelists collectively represent **11 countries and 5 continents**, Australia, Belgium, Brazil, Canada, China, Japan, the Netherlands, Singapore, South Korea, the United Kingdom, and the United States, ensuring that perspectives from both established and emerging research communities are well represented.

We maintained a deliberate balance between **academia and industry**, including participants from leading universities (e.g., Caltech, Stanford, UCL, University of Tokyo, Seoul National University, York University, Shanghai Jiao Tong University) and major research labs or companies (e.g., Amazon, Wayve, Google DeepMind, Qualcomm, RBC Borealis), alongside national research institutions such as the Singapore-MIT Alliance for Research and Technology and

National Institute of Advanced Industrial Science and Technology (AIST, Japan), and public sector institutions such as the Office of the Comptroller General of Brazil.

The team (Speakers, panelists, organizers) spans all career stages, from early-career researchers and postdoctoral fellows to mid-career faculty and senior scientists. We further highlight **gender diversity**, with both male and female researchers among invited speakers, panelists, and organizers (52% women, 48% men).

Beyond the organizing team, a program committee of 80+ members further strengthens inclusivity, bringing together voices from **20+ countries and 6 continents** across academia, industry, and research labs. This breadth supports fair, high-quality reviewing and equitable representation across regions, sectors, and career stages.

Finally, the workshop format promotes inclusion by creating opportunities for **junior researchers** through contributed talks, lightnings, and poster sessions, encouraging active dialogue across regions, organizations, and experience levels.

Grants and Financial Assistance. To recognize excellence, we will present a **Best Paper Award** to one outstanding contribution. We are also finalizing resources to offer a number of **travel and participation grants** for students and early-career researchers, with priority given to those with financial need, individuals from underrepresented groups, and first-time conference attendees. The support may cover registration, accommodation, or partial travel expenses.

Accessibility and Modality. We strongly encourage in-person participation to maximize interaction and community building. However, accepted papers and posters will also be made available on the official workshop website and open review, and all talks will be recorded and hosted on the conference platform for long-term access. We will accommodate Contributed Talks to be given remotely if authors cannot attend in person and will use the official conference interaction platform (e.g., Whova) to connect remote and in-person audiences.

To ensure inclusivity across **different time zones and participation formats**, we will combine live sessions with flexible access options. In addition to streaming and recorded talks, we will invite authors (5-min short video) and invited speakers to share pre-recorded summaries and presentation material in advance, ensuring that participants unable to join live sessions can still engage with the content. During live sessions, one organizer will moderate and selects questions submitted online, maintaining active discussion across both in-person and remote participants. With authors' consent, all presentation materials and recordings will remain publicly available on the workshop website after the event, extending the workshop's reach and accessibility beyond the conference day.

Outreach and Promotion. We will actively promote the workshop through a coordinated outreach strategy across social media platforms (e.g., X/Twitter, LinkedIn), academic and university mailing lists, and institutional and partner research networks spanning academia and industry. Calls for papers, invited talks, and accepted contributions will be shared through targeted announcements to maximize visibility. We will also collaborate with invited speakers, panelists and organizers to circulate information within their departments and communities, ensuring broad participation from both established and emerging research groups worldwide.

Review Process and Conflicts of Interest. All submissions will undergo **double-blind** review on OpenReview. Organizers will not handle or discuss submissions from their own institutions or recent collaborators; reviewer assignments will be both automatically and manually screened to avoid institutional or recent collaboration overlaps (within the past three years).

Each paper will receive **at least three reviews**, including a senior reviewer, to ensure balanced, high-quality feedback while providing mentorship opportunities for junior reviewers. This process promotes fairness, diverse perspectives, and constructive author feedback. Additionally, workshop chairs will not serve as organizers or deliver invited talks. However, they are welcome to submit papers and present contributed talks.

Use of Large Language Models. Following ICLR LLM usage policy, any use of LLMs in authoring, analysis, coding, or editing will be explicitly disclosed in the paper and submission form, and human authors remain fully responsible for all content. The review process maintains human oversight, with all reviewing and moderation conducted by human committee members, reviewers, and organizers who must likewise disclose any LLM assistance and must not compromise submission confidentiality (e.g., by pasting non-public text into third-party tools).

Workshop Committee, and Confirmed Speakers and Panelists

This workshop centers on three themes, *sensing drift*, *responding to drift*, and *operating at scale*. We balance theory and practice across academia and industry, while also ensuring diversity across career stage, gender, and geography. **All invited speakers, panelists, and organizers listed below have confirmed their participation**. Talk titles and abstracts will be posted before the event, and all names are listed alphabetically by surname.

Speakers.

- **Rahaf Aljundi**; Research Scientist, Toyota Motor Europe, *Expertise: continual/lifelong learning, class-incremental learning, OOD detection*
- **Anima Anandkumar**; Bren Professor, Caltech, *Expertise: foundation models and AI for science (neural operators), large-scale deep learning systems*
- **Elahe Arani**; Head of AI Research, Wayve; Assistant Professor, Eindhoven University of Technology, *Expertise: foundation models for autonomous/embodied intelligence, continual learning, self-supervised learning, learning under noisy labels*
- **Chelsea Finn**; Assistant Professor, Stanford University, *Expertise: meta-learning, adaptation, robot learning*
- **Arthur Gretton**; Professor, UCL; Research Scientist, Google DeepMind, *Expertise: kernel methods, nonparametric hypothesis testing, distribution-shift testing/detection*
- **Taesup Moon**; Professor, Seoul National University, *Expertise: test-time adaptation, continual learning, domain generalization*
- **Masashi Sugiyama**; Director, RIKEN Center for Advanced Intelligence Project; Professor at the University of Tokyo, *Expertise: covariate shift, density-ratio estimation, weakly supervised learning*

Panelists.

- **Ticiana L. Coelho da Silva**; Researcher, Brazilian Office of the Comptroller General (CGU), *Expertise: government oversight, auditing, transparency, responsible AI in the public sector*
- **Flora D. Salim**; Professor, University of New South Wales; Deputy Director, UNSW AI Institute, *Expertise: spatio-temporal and time-series machine learning, human-centered AI, urban computing, continual learning*

- **Evan Shelhamer**; Assistant Professor, University of British Columbia; Canada CIFAR AI Chair; Vector Institute faculty member, *Expertise: computer vision, representation learning, adaptation, self-supervised learning*
- **Frederick Tung**; Research Director, RBC Borealis, *Expertise: foundation models, trustworthy machine learning, large-scale deployment in finance*

Organizers.

- **Motasem Alfarra**; Research Scientist, Qualcomm AI Research in Amsterdam, *Research Expertise: Adaptation and Safety.*
- **Chung-Chi Chen**; Researcher, AIST, *Expertise: Concept drift, NLP for Finance*
- **Elham Dolatabadi**; Assistant Professor, York University; PI, i4Health Research Lab, *Expertise: Fairness, interpretability, healthcare ML*
- **Sepid Hosseini**; Machine Learning Researcher, RBC Borealis, *Expertise: Fraud detection, time Series forecasting, domain adaptation, trustworthy ML in finance*
- **Bo Li**; Associate Professor, University of Illinois Urbana-Champaign, *Expertise: Robustness, trustworthy ML, adversarial learning*
- **Murat Sensoy**; Senior Applied Scientist, Amazon AGI (London), *Expertise: Uncertainty quantification (Evidential Deep Learning), continual learning, LLM-based agents, trustworthy ML*
- **Dequan Wang**; Assistant Professor, Shanghai Jiao Tong University, *Expertise: Agentic AI for science, test-time scaling*
- **Teresa Yeo**; Postdoctoral Researcher, Singapore-MIT Alliance for Research and Technology, *Expertise: Robustness & adaptation, computer vision, embodied AI*

Program Committee. We have successfully secured 80+ confirmed Program Committee members from a wide range of institutions and organizations worldwide. This broad membership ensures comprehensive coverage of the workshop's topics and provides a balance of senior and early-career researchers. Such a large and diverse committee demonstrates the community's strong interest and support for the workshop theme, highlighting the timeliness and relevance of reliable and adaptive ML under drift. With this strong committee size, we can guarantee high-quality, timely feedback while keeping the workload light for individual reviewers. In addition, we are continuously expanding our program Committee by reaching out through professional networks, academic mailing lists, and community channels to ensure wide representation and fresh perspectives.

Confirmed Program Committee Members

- Adiel Teixeira de Almeida Filho, Federal University of Pernambuco
- Afsaneh Hasanebrahimi, University of Melbourne
- Ali Gholami, Augmodo
- Alvaro Corriea, Qualcomm AI Research
- Amira Fathy, eHealth, Cairo University
- Amirhossein Zarezadeh, Sharif University
- Arian Khorasani, Tech3Lab; HEC Montréal
- Ariana M. Villegas Suarez, University of Engineering and Technology in Peru
- Aurelien Pelissier, Yale University
- Azmine Toushik Wasi, Shahjalal University of Science and Technology
- Betty Shear, University of British Columbia
- Cheng-Long Wang, King Abdullah University of Science and Technology

- Christos Ziakas, Imperial College London
- Debangshu Banerjee, University of Illinois
- Dharmesh Tailor, Qualcomm AI Research
- Edward Smith, RBC Borealis
- Ehsan Hoseinzade, Simon Fraser University
- Eric Nuertey Coleman, University of Pisa
- Eunjin Roh, Oregon State University
- Faeze Ghorbanpour, TU Munich
- Fatemeh Amerehi, University of Limerick
- Golnoosh Samei, RBC Borealis
- Hala Sheta, University of Waterloo
- Hamed Shirzad, University of British Columbia
- Heitor Rapela Medeiros, École de Technologie Supérieure
- Hossein Firooz, Finnish Center for AI
- Hossein Hajimirsadeghi, RBC Borealis
- Ishan Jindal, Samsung
- Issam Laradji, ServiceNow
- Jack Lu, New York University
- Juexiao Zhou, King Abdullah University of Science and Technology
- Julio Hurtado, University of Warwick
- Junhyug Noh, Ewha Woman University
- Korbinian Pöppel, Johannes Kepler University Linz
- Leo Feng, RBC Borealis
- Mahsa Keramati, DarkVision Technologies
- Mahsa Maleki, Prenuvo
- Mandana Ghanavati, University of Melbourne
- Maternus Herold, BMW Group
- Melissa Mozifian, Amazon
- Mengyao Zhai, RBC Borealis
- Merey Ramazanova, KAUST
- Mingyu Kim, University of British Columbia
- Mohamed Mahmoud, RAMYRO Inc.
- Mohamed Osama Ahmed, RBC Borealis
- Mohammad Amin Shabani, RBC Borealis
- Mohammad Fares, Istanbul University-Cerrahpasa
- Mohammadreza Mohseni, Google
- Mojgan Kouhounestani, University of Melbourne
- Myungjoon Kim, KAIST ; Cornel University
- Nina da Hora, Universidade Estadual de Campinas
- Omid Reza Heidari, Concordia University
- Paria Mehrbod. Concordia University/Mila Quebec
- Qiaoyue Tang, University of British Columbia
- Qincheng Lu, McGill University
- Radha Poovendran, University of Washington
- Rozhin Nobahari, University of Montreal
- Ruizhi Deng, RBC Borealis
- Saber Saberian, Recursion
- Saeed Ilchi Ghazaan, Second Foundation
- Sagar Srinivas Sakhinana, Tata Research Development and Design Center
- Sai Srujana Buddi, Apple
- Sarinasadat Hosseini, Panasonic
- Shyma Alhuwaider, KAUST
- Sumanth Varambally, University of California San Diego
- Sungkyu Park, Korea Development Institute
- Tarun Prasad, Harvard University
- Thapelo Sindane, University of Pretoria
- Thibaut Durand, RBC Borealis
- Thomas Lee, Edinburgh University
- Vishnu Sarukkai, Stanford University
- Weillian Song, Integrated Projects
- Wen-Ding Li, Cornell University
- William Toner, Huawei
- Yasaman Etesam, Simon Fraser University
- Yasir Ghunaim, KAUST
- Yasutaka Furukawa, Wayve; Simon Fraser University
- Yuefan Deng Stonybrook University
- Yunbei Zhang, Tulane University
- Yuzhen Mao, Stanford University
- Zhaorun Chen, University of Chicago
- Zhengqing (Eric) Wang, Wayve; Simon Fraser University
- Zhijie Deng, Shanghai Jiao Tong University

Biographies

Speakers.

Rahaf Aljundi: is a Senior Research Scientist at Toyota Motor Europe. Her work focuses on continual and lifelong learning, novelty and out of distribution detection, active learning, and domain adaptation, building models that keep learning from streams, detect unknown inputs, and escalate for annotation when needed. She develops methods that balance stability and plasticity under changing conditions to maintain performance over time. She received her Ph.D. in computer science from KU Leuven (Belgium), where she worked extensively on continual learning.

Anima Anandkumar: is the Bren Professor of Computing at Caltech and a former senior director of AI research at NVIDIA. Her research spans large scale learning, robustness, and foundation models, with contributions to tensor methods, distributed training, and out of distribution generalization. She is a recipient of the Alfred P. Sloan Fellowship, NSF CAREER Award, and IEEE Fellowship.

Elahe Arani: is Director of AI at Wayve and an Assistant Professor at Eindhoven University of Technology (TU/e). Her research focuses on foundation models for embodied intelligence and real world perception–action systems, with contributions to generative world models, vision language action policies, reinforcement learning, continual learning, and learning under noisy labels. At Wayve, she leads teams building real world driving AI using world models (e.g., GAIA-2) and vision–language–action driving (e.g., LINGO-2).

Chelsea Finn: is an Assistant Professor of Computer Science and Electrical engineering at Stanford University and a Co-founder of Physical Intelligence. Her research develops algorithms that enable robots and agents to rapidly adapt to new environments, tasks, and conditions, with a focus on meta-learning and large scale robot learning. She has pioneered influential methods for few-shot learning, such as Model Agnostic Meta-Learning, and vision based robotic manipulation, advancing the foundations of adaptive and robust AI.

Arthur Gretton: is a Professor at the Gatsby Computational Neuroscience Unit, University College London (UCL), Director of the Centre for Computational Statistics and Machine Learning at UCL, and a Research Scientist at Google DeepMind. His research develops kernel based nonparametric methods for machine learning, including the Maximum Mean Discrepancy for two sample testing and the Hilbert–Schmidt Independence Criterion for dependence measurement. He applies these tools to causal inference, representation learning, and the design and training of generative models. These statistical tests are widely used for distribution shift detection, model criticism, and monitoring data and model alignment in real-world ML systems.

Taesup Moon: is a Professor of Electrical and Computer Engineering at Seoul National University, where he leads the MindLab research group. His work focuses on continual learning, explainable AI, and fairness, aiming to overcome the practical limitations of machine learning systems. He has developed adaptive learning methods that maintain performance across changing environments, with applications in medical imaging, neuroscience, semiconductors, and environmental data. His group also develops auditing methods for bias and performance drift in deployed systems, supporting trustworthy AI in realworld use.

Masashi Sugiyama: is a Professor at the University of Tokyo and Director of the RIKEN Center for Advanced Intelligence Project. His research develops the theoretical foundations of machine learning under distribution shift, including covariate shift adaptation, density ratio estimation, and weakly supervised learning. He has co-authored influential books such as Machine Learning in Non-Stationary Environments and Machine Learning from Weak Supervision, which have shaped approaches to model monitoring and domain adaptation.

Panelists.

Ticiana L. Coelho da Silva: is a Researcher at the Brazilian Office of the Comptroller General (CGU). Her work focuses on data driven public sector oversight and responsible AI, including audit analytics, risk scoring and sampling, anomaly and out of distribution detection, and model explainability and documentation. She collaborates with policy and engineering teams to design monitoring and governance frameworks for ML systems in high stakes administrative workflows, ensuring accountability, fairness, and transparency.

Flora Salim: is a Professor in the University of New South Wales (UNSW) School of Computer Science and Engineering, Deputy Director of the UNSW AI Institute, and the Cisco Chair of Digital Transport & AI. Her work advances multimodal and time series ML foundation models for spatio-temporal data, sensor and wearable modeling, and trustworthy ML with applications in mobility, transport, energy, and urban systems. She also contributes to human centered sensing for health and well-being. She serves as vice chair of the IEEE Task Force on AI for time series & spatio-temporal data and is a member of the ARC College of Experts.

Evan Shelhamer: is an Assistant Professor of Computer Science at the University of British Columbia, a Faculty Member at the Vector Institute, and a Canada CIFAR AI Chair. He is known for foundational work on Fully Convolutional Networks (FCN) for semantic segmentation and for test-time adaptation methods such as Tent for adapting models under distribution shift. His research spans visual recognition, self-supervised learning, and robust adaptation for deployment time data shift. Prior to UBC, he was a research scientist at Adobe and Google DeepMind, and received his Ph.D. in computer science from UC Berkeley.

Frederick Tung: is a Research Director at RBC Borealis, leading the development and deployment of ATOM, RBC's proprietary foundation model for financial services. Since 2023, ATOM has been integrated into several products and services. Frederic's work involves training and aligning large models for banking use cases, covering data governance and privacy constraints, domain adaptation, and evaluation under distribution shift. He drives the translation of research into production, ensuring that advanced AI can be safely deployed in high stakes financial environments. He has co-authored over thirty publications in machine learning and computer vision, including in ICLR, ICML, CVPR, ICCV, ECCV, and TPAMI.

Organizers Biographies and Previous Workshop Experiences.

Motasem Alfarra: is a Research Scientist at Qualcomm AI Research in Amsterdam, the Netherlands. His research focuses on robustness and alignment under domain shifts. His current research develops test-time adaptation and continual learning methods that keep models reliable as distributions evolve. He also studies safety protocols for large language models, connecting robustness objectives with pragmatic evaluation. He co-organized the first and second Workshops on test-time adaptation, held at CVPR 2024 and ICML 2025, respectively. His publications on adaptation appeared at ICML, ICLR, ICCV, AAAI, and TMLR. Further, his publications on robustness appeared at CVPR, BMVC, and TPAMI.

Chung-Chi Chen: is a Researcher at the Artificial Intelligence Research Center (AIRC), National Institute of Advanced Industrial Science and Technology (AIST), Japan. His research focuses on financial NLP for noisy, dynamic, and domain specific text, with an emphasis on robustness and adaptation. He has established key community resources, including ACL SIG-FinTech, the FinNum and FinArg benchmark series at NTCIR, and co-organized the FinNLP/FinWeb workshops at venues such as IJCAI, WWW, EMNLP, and AACL-IJCNLP. Chung-Chi Chen has been a speaker at multiple finance focused workshops, including the KDF Workshop at SIGIR 2023, the Financial AI Workshop at ICLR 2025, the FinNLP Workshop series, and tutorials at AACL 2020 and EMNLP 2021. His work bridges academic research

and real world impact, advancing FinTech/LegalTech applications and promoting responsible, governance oriented evaluation.

Elham Dolatabadi: is an Assistant Professor at York University and Principal Investigator of the i4Health Research Lab. Her research develops trustworthy machine learning for healthcare, combining fairness, interpretability, causality, and uncertainty modeling to support clinical decision making from electronic health records and medical imaging. She designs shift aware evaluation and monitoring protocols, and methods robust to missingness, label noise, and distribution change, with the goal of safe, reliable deployment in real clinical workflows. She has served on the organizing committee for the ML4H 2021 Workshop at NeurIPS and as an area chair for NeurIPS 2021.

Sepidehsadat (Sepid) Hosseini: is a Machine Learning Researcher at RBC Borealis. She works on fraud and anomaly detection for high volume financial time series, building detectors that stay calibrated and robust under non-stationarity (distributional and temporal shift). Her broader interests include domain adaptation and foundation model use in finance, alongside techniques for bias mitigation and machine unlearning to satisfy risk and compliance constraints. She has published in leading venues including ICLR, NeurIPS, CVPR, and ICML.

Bo Li: is an Associate Professor of Computer Science at the University of Illinois at Urbana Champaign. Her research develops the theory and practice of trustworthy machine learning at the intersection of ML, security, privacy, and game theory, covering adversarial robustness, data poisoning/backdoor defenses, and privacy preserving learning and data publishing. She has designed scalable frameworks for robust training and evaluation in safety and risk sensitive settings. Her contributions have been recognized with the IJCAI Computers and Thought Award, an Alfred P. Sloan Research Fellowship, and an NSF CAREER Award. She has co-organized the ICML 2019 Workshop on Security and Privacy of Machine Learning and the CVPR 2019 Workshop on Adversarial Machine Learning in Real World Computer Vision Systems, and served as tutorial chair for ICML 2024.

Murat Sensoy: is a Senior Applied Scientist at Amazon's Artificial General Intelligence (AGI) Organization in London, UK. He is known for Evidential Deep Learning (EDL), a probabilistic approach to uncertainty quantification in neural networks. His current work develops LLM based agents that adapt via experience and memory, integrating continual learning with reliability constraints. Previously, he was an associate professor at Ozyegin University, a visiting scholar at University College London, and a senior research scientist at Blue Prism AI Lab.

Dequan Wang: is an Assistant Professor at Shanghai Jiao Tong University. His research focuses on computer vision and representation learning, with interests in transfer and domain adaptation, generalization under distribution shift, and efficient training/inference for large scale models. He develops methods that improve robustness and data/compute efficiency for recognition and related tasks in real world settings. He has organized multiple workshops at top computer vision venues such as ICCV and CVPR.

Teresa Yeo: is a Postdoctoral Researcher at the Singapore-MIT Alliance for Research and Technology, where she develops adaptive systems using neuro symbolic methods across domains ranging from mathematics to embodied AI. Her work focuses on making models robust and adaptive under distributional and task shift. She was a speaker at the CVPR 2024 Workshop on Test-Time Adaptation and co-organizer of the ICML 2025 edition.