Workshop Summary

Abstract

Learning Meaningful Representation Learning (LMRL) Workshop 2026 aims to identify the key bottlenecks in the development of virtual cells. Virtual cells are an *in silico* representation of a cell's behaviour and dynamics in both health and disease, with immense implications for research, diagnostics and therapeutic development. Building towards such a system begins with learning meaningful representations within individual modalities, which form the foundations for scaling the complex heterogeneous biological signals into a coherent model of a cell, and combining them into integrative models that capture biology's complexity. LMRL 2026 highlights emerging directions for overcoming these challenges by focusing on four core ingredients - causality in biological systems, generative modelling, interpretable representations, and *leveraging virtual cells for real-world impact*. This workshop aims to catalyse the advances in how we learn meaningful representations by bringing together the AIxBio community around a shared scientific roadmap.

Motivation & Themes

Following the resounding success of our revival of the *Learning Meaningful Representations of Life (LMRL)* workshop at ICLR 2025 in Singapore (https://www.lmrl.org/), which attracted over 500 attendees throughout the day, it is evident that biology-themed topics resonate deeply within the broader ICLR community. This enthusiastic engagement underscores that the **intersection of AI for biological applications** is both one of the next frontiers for AI for Science, and is central to the future of biomedical discovery.

One of the main drivers behind this surge of interest is the growing excitement for building large models to simulate cellular function and behavior *in silico* with AI models. The promise of leveraging AI to capture the complexity and structure of biological systems has captivated both computational biologists and core ML researchers. Over the past year alone, the community has expressed a multitude of visions for how these foundation models should be built, including "virtual cells" [Bunne et al., 2024; Noutahi et al., 2025] and causal cell atlases [Rood et al., 2024], as well as valid reasons to be cautious about the progress made thus far [Ahlmann-Eltze et al., 2025], and new methods published toward their development [e.g. Adduri et al. 2025., Klein et al. 2025, Wenkel et al. 2025].

As the field is still in its early yet rapidly consolidating phase, this presents a remarkable opportunity to foster open discussion, collaboration, and insight sharing about *how* to design virtual cells in ways that are rigorous, interpretable, and ultimately useful. To this end, we have identified what we see as the **core ingredients** of meaningful virtual cell design, and curated our speaker line-up to represent expertise across these areas: (i) *causality in biological systems,* (ii) *generative modelling,* (iii) *interpretable representations,* (iv) *leveraging virtual cells for real-world impact.*

Our program, refined through prior editions, will combine **four invited keynote speakers** as well as **contributed oral and spotlight talks** from the best full and tiny paper tracks to foster in-depth exchange and networking opportunities. Beyond these structured talks, *LMRL 2026* aims to provide a space where the anticipated differing perspectives on the feasibility and scope of virtual cells can naturally surface. To this end, our agenda will open the floor to an engaging **panel discussion** on the topic of *"The Core Ingredients for Meaningful Biological Representations: Causality, Interpretability & Generative Modelling"* with key opinion leaders in this space. To bridge academic and industry perspectives, we aim to spark a stimulating discussion with each panelist sharing their vision for how virtual cell models could transform drug discovery and personalized medicine in the coming decade(s). To further encourage networking opportunities, we will

extend both of our **poster sessions** to 90 minutes (vs. 60 minutes last year) to enable fruitful and unstructured debates about the relevance, contribution and meaningfulness of the presented works, inclusive of multiple layers of biological information.

Together, we aim to (i) identify the computational, methodological, and conceptual advances that will shape the path toward *learning meaningful representations of life*, (ii) channel this discussion toward constructive and actionable directions, including applications of virtual cells in drug discovery and personalized medicine, and (iii) emphasize foundational approaches that transcend modalities, engaging both biological and core ML researchers and offering an academically grounded forum for the main conference audience. We believe that ICLR, a premier AI conference with a strong emphasis on cutting-edge advances in representation learning, provides the ideal platform for uniting researchers at the forefront of AIxBio and fostering breakthroughs in how we understand and apply biological data representations to real-world challenges.

Invited Speakers

Invited Keynote Line-up

We have extended our invitations to 4 keynote speakers, each covering their specific research area, all of whom have **already accepted our invitation** as well as an **in-person attendance** of our proposed workshop**:**

- **Jiaqi Zhang** (PhD Candidate | Massachusetts Institute of Technology (MIT); Cambridge, MA, USA)
  - Research Area & Workshop Fit: *Beyond correlation - developing causal & mechanistic cell biology insights*
  - Bio: Jiaqi's research in the world-renowned Caroline Uhler causality lab focuses on developing statistical and computational methods for analyzing biological data, with a particular interest in understanding complex biological systems such as gene regulatory networks. Jiaqi's work spans various areas, including graphical models, causal inference, and algebraic statistics, and she has made significant contributions to advancing computational biology through interdisciplinary collaborations. Her theoretical background will help enhance the workshop through her unique perspective on causality for learning meaningful representations.

- **Alexander Tong** (Principal Investigator | AITHYRA Research Institute for Biomedical AI; Vienna, Austria)
  - Research Area & Workshop Fit: *Generative biology revolution via flow matching & vector fields modelling*
  - Bio: Alex leads research on generative modeling for biological systems. His work integrates ideas from optimal transport, flow matching and causal discovery, to develop generative models of both proteins and cellular responses. His recent work at the intersection of flow-based modeling, causality, and biological inference aligns closely with our workshop's mission to highlight meaningful generative modeling capabilities for understanding cellular function and behavior.

- **Virginie Uhlmann** (Institute Director | BioVision Center @ University of Zürich; Zürich, Switzerland)
  - Research Area & Workshop Fit: *New era of Interpretable intelligence for quantifiable microscopy imaging*
  - Bio: Virginie's research focuses on developing computational and statistical approaches to quantify, describe, and model biological structures observed in microscopy images. With a background in bio- and electrical engineering, her work combines computational geometry, statistical shape analysis, and machine learning to extract meaningful representations of morphology and organization from large-scale bioimage datasets. As the Director of the BioVision Centre and a Visiting Group Leader at EMBL-EBI, she is a leading advocate for open, collaborative, and theory-driven approaches to biological discovery, which are values that resonate strongly with our workshop's community to build interpretable and biologically grounded models of life.

- **Marinka Zitnik** (Associate Professor | Harvard University; Boston, MA, USA)
  - Research Area & Workshop Fit: *Multimodal models in personalized medicine & agentic scientific discovery*
  - Bio: Marinka is a leading researcher at the intersection of artificial intelligence, biology, and medicine. Her lab develops foundational AI models that integrate genetic, molecular, and clinical data to advance agentic scientific discovery and enable individualized diagnosis and treatment. By pioneering approaches grounded in geometry, structure, and biomedical knowledge, her work is transforming how AI drives therapeutic design and biomedical understanding. She also leads the Therapeutics Data Commons and the International AI4Science initiative, advancing global efforts to harness AI for discovery and health.

Invited Panelist Line-up

To foster thought-provoking discussions & insights from the field, we plan to host a panel discussion on the topic of *"The Core Ingredients for Meaningful Biological Representations: Causality, Interpretability & Generative Modelling",* featuring the following panelists*.

- **Nathan Frey** (Co-founder & CTO | Coefficient Bio; New York City, NY, USA)
  invitation accepted
  - Research Area & Workshop Fit: *Lab-in-the-loop approach via cell-based representation learning*
  - Bio: Nathan is a pioneer in effectively utilising active AI models for laboratory experiment design, which he introduced as the 'lab-in-the-loop' paradigm. Formerly as a Principal Scientist and Group Leader at Prescient Design @ Genentech, Nathan spearheaded efforts in generative modeling and biological AI and led the development of biological foundation models that capture the language of cellular and molecular processes to accelerate drug discovery and therapeutic innovation.

- **Sara-Jane Dunn** (VP of Computational Research | Relation Therapeutics; London, UK)
  invitation accepted
  - Research Area & Workshop Fit: *Conceptual and computational architect of virtual cell representations*
  - Bio: SJ is an exceptional addition to our panel discussion on meaningful representations of virtual cells given her pioneering work at the intersection of computational modeling, information theory, and cellular biology. Her research has long focused on understanding how biological systems can be formalized, modeled, and ultimately programmed, offering a deeply conceptual and rigorous lens on what it means to represent cellular processes faithfully. At Relation Therapeutics, and previously at Microsoft Research and DeepMind, she has advanced frameworks that bridge abstract computation with biological function, which is precisely the type of cross-disciplinary thinking needed to discuss the principles underlying virtual cells.

- **Yusuf Roohani** (Machine Learning Group Leader | Arc Institute; Palo Alto, CA, USA)
  invitation accepted
  - Research Area & Workshop Fit: *Challenging virtual cells via the Virtual Cell Challenge*
  - Bio: As the co-author of the *Virtual Cell Challenge* (https://virtualcellchallenge.org/), Yusuf is an ideal panelist for the discussion on *meaningful representations of virtual cells* given his pioneering work at the intersection of machine learning and biological systems. As the leader of a machine learning group at the Arc Institute, he focuses on building virtual cells by integrating large-scale AI models with high-throughput experimental data generation, directly embodying the workshop's vision of constructing biologically faithful and computationally grounded representations of life. His research on AI-driven experiment design and cellular engineering offers a unique perspective on how generative and causal modeling principles can guide meaningful scientific discovery. With his deep grounding in both computational modeling and experimental biology, Yusuf brings an invaluable, integrative viewpoint to the discussion of what makes virtual cells *trustworthy, interpretable, and actionable*.

- **Jean-Philippe Vert** (Co-founder & CEO | Bioptimus; Chief R&D Officer | Owkin; Paris, France)
  invitation accepted
  - Research Area & Workshop Fit: *Drug discovery via multi-scale AI foundation models for biology*
  - Bio: JP is a prominent leader in the field of machine learning and computational biology, focused on transforming medicine through AI-powered solutions. He directs groundbreaking efforts to use AI for optimizing treatment discovery, particularly in cancer, by analyzing complex biological data to guide personalized therapies. Through the Bioptimus pioneering initiative, JP creates universal AI foundation models across biological scales to accelerate breakthroughs in biomedicine and drug development.

- **Daniel Burkhardt** (Digital Biology | NVIDIA; Toronto, ON, Canada)
  invitation accepted
  - Research Area & Workshop Fit: *Leading NVIDIA's strategy in single-cell and multiscale digital biology*
  - Bio: Daniel leads NVIDIA's strategy and partnerships in single-cell and multiscale biology, fostering a global ecosystem for AI-driven drug discovery and digital biology innovation. Previously, he co-founded the Open Problems in Single-Cell Analysis initiative, funded in part by the Chan Zuckerberg Initiative, bringing together thousands of international researchers to advance benchmarking and reproducibility in single-cell research through NeurIPS competitions. He previously led machine learning research at Cellarity, developing models to accelerate therapeutic design.

- **Mor Nitzan** (Assistant Professor | The Hebrew University of Jerusalem, Israel)
  invitation extended
  - Research Area & Workshop Fit: *Single-cell network dynamics & interactions in collaborative ecosystems*
  - Bio: Mor's group is interested in how local changes, responses and interactions between cells lead to global coordinated effects, such as multicellular self-organization. She studies how to infer and design optimal interactions on single-cell level to reveal biological processes and contexts that cellular populations are driven by. Mor implements geometrical structures to determine the global states of (dys-)functioning tissues.

- **Christina Theodoris** (Assistant Investigator | Gladstone Institute; San Francisco, CA, USA)
  invitation extended
  - Research Area & Workshop Fit: *Learning causal representations of cellular networks driving disease*
  - Bio: Christina's research sits at the interface of gene regulation and disease modeling, using AI to model and perturb complex gene regulatory networks in human cells. At the Gladstone Institutes, her group develops interpretable ML approaches that connect predictive modeling with experimental validation. Her expertise in uncovering mechanistic insights from biological data makes her an ideal contributor to our panel on *meaningful representations of virtual cells*, bringing a clinician-scientist's perspective on how AI models can faithfully capture cellular function and disease mechanisms.

- **Maria Brbić** (Assistant Professor | EPFL; Lausanne, Switzerland)
  invitation extended
  - Research Area & Workshop Fit: *Reveal structure, diversity, and meaning in biological systems*
  - Bio: Maria leads the Machine Learning in Biology Lab at EPFL, where she designs algorithms to learn from large, heterogeneous biological datasets, ranging from microbial genomes to single-cell atlases. Her work bridges *representation learning, graph-based modeling,* and *biomedical discovery*, including pioneering methods for cross-experiment cell type discovery. With experience spanning Stanford, the Chan Zuckerberg Biohub, and now EPFL, Maria brings a deep understanding of how to translate abstract representations into biological insight, making her a perfect voice on our panel exploring *how to learn truly meaningful representations of life.*

** Notes: (i) We recognize that travel logistics can be challenging, and while we prioritize in-person attendance for all speakers, we understand that final confirmations may be affected by unforeseen constraints (such as visa*

delays or childcare responsibilities). That said, for the panel discussion in particular, we will strictly enforce in-person attendance to facilitate rich discussion and debate, as this approach proved highly successful in last year's workshop and was among the most valued aspects of participant feedback. (ii) To ensure a dynamic and balanced discussion, we will be limiting the final panel to a manageable number of speakers. We initially extended a few extra invitations to accommodate the possibility of last-minute cancellations. Once the workshop is approved, we will confirm a final subset of panelists with an extended official invitation, prioritizing in-person participation to foster meaningful exchange and debate.

Preliminary Agenda & Thematic Areas

By maintaining a streamlined program with **four keynote speakers** and subsequent **spotlight talks** highlighting the best posters per theme (chosen to match the keynote speakers' areas of research), we aim to keep the workshop focused and elevate the prestige of the spotlight sessions to showcase the most significant contributions within each area. Likewise, our two 1.5-hour long **poster sessions** are aimed to ensure a high level of impact and engagement for all attendees, which will be further enhanced via their questions to our **6 invited panelists**. Moveover, three tiny paper submissions will be included to present their **tiny paper pitches.**

Table 1 | Preliminary workshop agenda.

Time	Duration	Activity	Presenter(s)	Theme / Description
09:00 - 09:10	10 mins	Opening remarks	Lead organisers	Welcome + Sponsor Announcement
09:10 – 09:45	30 mins + 5 mins	Keynote #1 + Qs	**Alexander Tong**	*Generative biology revolution via flow matching modelling*
09:45 – 10:00	15 min total (3 mins each)	Spotlight talks		Tiny papers #1 - #5
10:00 – 11:30	90 mins	Poster Session #1	–	10:00–10:30 Coffee Break (with ICLR)
11:30 – 12:05	30 mins + 5 mins	Keynote #2 + Qs	**Virginie Uhlmann**	*New era of Interpretable intelligence for quantifiable microscopy image analysis*
12:05 – 12:17	10 mins + 2 mins	Oral #1 + Qs		Full paper #1
12:18 – 12:30	10 mins + 2 mins	Oral #2 + Qs		Full paper #2
12:30 – 13:45	75 mins	Lunch Break	–	Partially overlapping with ICLR lunch break (12:00 – 14:00)
13:45 – 14:45	60 mins	Panel Discussion	Panelists + moderator	*The Core Ingredients for Meaningful Biological Representations: Causality, Interpretability & Generative Modelling*

14:45 – 15:20	30 mins + 5 mins	Keynote #3 + Qs	**Jiaqi Zhang**	*Beyond correlation - developing causal & mechanistic cell biology insights*
15:20 – 15:50	90 mins	Poster Session #2	–	15:00–15:30 Coffee Break (with ICLR)
15:50 – 16:25	30 mins + 5 mins	Keynote #4 + Qs	**Marinka Zitnik**	*Foundation models for therapeutic design and treatment prediction*
16:25 – 16:37	10 min + 2 mins	Oral #3 + Qs		Full paper #3
16:37 – 16:50	10 min + 2 mins	Oral #4 + Qs		Full paper #4
16:50 – 17:00	10 min	Closing remarks	Sponsors' representatives	Award ceremony
17:00 – 18:00	60 min	Networking	All attendees	Drinks & nibbles

***** All time slots (+) include the time for questions (Qs) from the audience.

Submission Themes

What this workshop is: We invite topics spanning diverse areas of biology, including AI for omics analysis, cell profiling, plant biology, microbiome research, tissue and systems biology, and clinical or diagnostic tool development, and we also welcome submissions from related scientific domains. Our primary goal is to attract participants who are **not native to ICLR** and bring expertise from the broader scientific community into an AI-oriented conference. Therefore, a submission being "misaligned" with conventional biological topics will not automatically constitute grounds for rejection.

What this workshop is not: To maintain a focused discussion, we deliberately limit the scope to our core topics and will generally exclude protein design, molecular design, docking/binding, and chemistry-focused work. That said, if other workshops in this space do not cover these areas, we have both the experience and the interest in hosting them, and are open to including them to ensure coverage of complementary topics. We may also consider other modalities, but only if they complement our focus and are not addressed elsewhere.

However, to align all submissions to the topic of the workshop and to encourage deeper reflection on the relevance and impact of their work, we require each submission to include a brief "**meaningfulness statement**". This brief statement (100 words max) should address the question of *"How does this work help us learn meaningful representations of life?"* and it is aimed to ensure that the contributors thoughtfully consider the significance, impact and relevance of their work to the workshop. Importantly, this statement will not be visible to reviewers nor influence the acceptance decision, but it serves to discourage the resubmission of previously rejected papers without proper refinement.

Paper Tracks

We recognize that workshops often attract work-in-progress submissions that may not be ready for a full conference paper. To accommodate varying levels of progress, we offer **two submission track:**

Full paper track | We invite submissions of 4–8 pages, excluding references. These papers should outline the biological problem, the AI methods used, implementation details, benchmarking metrics, and comparisons to related work. Based on past LMRL experience, we expect over 150 submissions and plan to accommodate up to 80 posters. The top three ranked papers will be selected for a spotlight in the main program as "Contributed Talks", and all full papers will compete for various awards.

Tiny paper track | We also invite 1–2 page submissions to encourage participation from under-represented and under-funded communities. These "extended abstracts" should focus on the methodology or novel approach to solving a biological problem, even if results are not yet available. Lack of results will not be grounds for rejection, but the proposed methodology must be well-defined and framed within the biological context. We expect 50+ submissions and plan to accept 20 "tiny posters" (smaller format, *e.g.* A1). Additionally, three "Tiny Paper Pitches" will be selected for short spotlight talks in the main program.

Table 2 | Comparison of the ‘full’ and ‘tiny’ paper tracks.

Criteria	**Item**	**Paper Track**	
		Full	**Tiny**
Submission	Meaningfulness statement	Required	Required
	Biological areas	Restricted (see above)	Any
	Page limit	4-8 pages	1-2 pages
	Paper structure (comparable to)	Conference submission	Extended abstract
	Expected no. of submission	Over 100 papers	Over 50 papers
	Expected no. of acceptances	~70 posters	~15 posters
	AI-assistance with paper writing	Occasionally allowed	Unallowed
Reviewing	Reviewer guidance	Stricter	Relaxed
	Reviewer nomination	Required	Optional
	Double-blind review	Yes	Yes
On the day	Number of spotlights	4 talks	5 talks
	Spotlight talk duration	10 mins + 2 mins Qs	3 mins + no Qs
	Spotlight talk coverage	Unlimited slides	1 slide only
	Poster size	A0	A1

Both paper tracks will share the same submission deadline and will be reviewed by a common program committee. The main distinctions are: (i) full paper submissions must nominate at least one reviewer,

whereas this will remain optional for tiny papers; and (ii) review criteria will be slightly less stringent for shorter, work-in-progress submissions. Our structure draws inspiration from the *Women in Machine Learning (WiML) workshop* guidelines at NeurIPS. Reviewers will be informed of the track for each submission and, in exceptional cases, may recommend moving a paper between tracks. All submissions will be non-archival, and papers will be posted online unless authors request otherwise.

To ensure accessibility, authors who cannot attend in person may share **pre-recorded talks** with the workshop audience. Keynote, oral and spotlight talks will be recorded to enable remote participation, and we will **display posters on behalf of authors** unable to be physically present.

Timeline & Key Deadlines

To ensure timely submissions and prompt communication of decisions to authors, we have established a preliminary unified timeline for both the full and tiny paper tracks. This schedule **follows the recommended ICLR workshop deadlines**, with a few buffer days included to accommodate any last-minute organizational adjustments.

- Workshop Acceptance Notification: 1 December 2025
- Social Media Campaign to Announce Workshop: 3 December 2025
- Paper Track Criteria Published (on the website): 17 December 2025
- Social Media Campaign to Invite Submissions: 7 January 2026
- Paper Submission Deadline: **30 January 2026, 11.59pm AoE**
- Reviewer Assignment Deadline: 4 February 2026
- Review Submission Deadline: 25 February 2026
- Decision Notification + Spotlight Talk Invitations: **27 February 2026, 11.59pm AoE**
- Camera-Ready Deadline: 9 March 2026
- Import Workshop Program & Accepted Papers to iclr.cc: **11 March 2026, 11.59pm AoE**
- Electronic Poster Upload Deadline: 17 April 2026
- Spotlight Slides Submission Deadline: 17 April 2026

Workshop Attendees

Building on the success of the past LMRL workshops and the growing interest in AIxBio sessions at major AI conferences, we anticipate approximately 300 attendees at peak times. While tracking exact numbers at an open venue is challenging, we estimate that up to 500 individuals could benefit from our diverse program throughout the day, including poster presenters. We are committed to enabling remote participation, with plans to (i) livestream all talks and sessions for virtual attendees, and (ii) record the presentations for later access via the ICLR platform and/or our custom LMRL website (https://www.lmrl.org/). This option will be particularly valuable for participants facing exceptional circumstances, such as visa issues or inability to travel to the venue due to underfunding. Additionally, we may consider documenting the outcomes of the workshop in a correspondence letter or review article, or a less formal series of blog posts to be communicated over social media.

Related Workshops

The LMRL workshop complements other efforts at the AIxBio interface but stands out through its focus on *principles* rather than *domains*.

We recognise that the proposal of the ***"Virtual Cells across Scale, Space, and States"*** workshop concentrates on bio-specific challenges, such as noisy data, experimental reproducibility, and modeling across biological

scales. However, we are continuing the LMRL proposal as a timeless brand which broadens the scope to explore *meaningfulness* itself: causality, interpretability, and generative modelling. These concepts are central not only to biology but to the entire machine learning community.

Similarly, unlike the ***"MLGenX"*** workshop series, which is focused on drug discovery, LMRL focuses on learning representations of biological systems themselves, whether cells, tissues, or patients. By emphasizing foundational approaches that transcend modalities, LMRL aims to engage both biological and core ML researchers and offer a more academically grounded forum for the ICLR audience, setting our workshop apart. LMRL workshop (i) has historically attracted more researchers studying genomics and cellular phenotypes beyond drug discovery, and (ii) highlights the importance of learning meaningful representations that provide biological insights rather than just improving task performance.

While we believe our proposal stands out, we recognize that participants may wish to attend sessions across multiple workshops. To facilitate this, we kindly request that the ICLR workshop committee either schedule our workshop on a separate day from other biology-themed workshops, or co-locate related sessions in nearby rooms, if venue capacity permits. This arrangement will allow participants to move between rooms based on their interests and expertise, enhancing their engagement with concurrent workshops.

Organizing Committee

Kristina Ulicna* (kristina.smith.ulicna@gmail.com) is an incoming Research Scientist at Iambic Therapeutics, specializing in **interpretable single-cell representations from time-lapse image embeddings**. Previously, she was a Research Scientist at Valence Labs @ Recursion, after completing her post-doctoral fellowship at The Alan Turing Institute, working on spatial patterning in imaging data. She holds a PhD in Computational Biology from Alan Lowe's lab at UCL, a BSc in Biomedical Science from King's College London, internship experience gained from MIT, Cancer Research UK & Microsoft Research. *Previous organizing experience:* Kristina led the LMRL 2025 organising committee & serves as a Senior Program Chair at the Women in Machine Learning (WiML) workshop at NeurIPS 2025 and long-term volunteer since ICML 2023.

Rebecca Boiarsky* (rebeccaboiarsky@gmail.com) is a Machine Learning Research Scientist at Genesis Therapeutics, where she builds AI for small molecule drug design. In her PhD research, co-advised by David Sontag at MIT CSAIL and Gad Getz at the Broad Institute, she focused on **representation learning for single-cell RNA-sequencing data with an emphasis on applications in cancer and translational medicine**. Before pursuing her Ph.D., she worked in the Molecular Profiling department at Regeneron Pharmaceuticals. Rebecca holds a PhD in EECS from MIT, a Master's in Biomedical Engineering from Columbia University, and a B.A. in Physics from Yeshiva University. *Previous organizing experience:* Rebecca co-organised the previous LMRL workshops at ICLR 2025 & NeurIPS 2022.

Till Richter* (till.richter@helmholtz-munich.de) is a final year PhD candidate at Technische Universität München (TUM) and Helmholtz Munich, advised by Fabian Theis and Niki Kilbertus. His research focuses on **self-supervised and generative deep learning for single-cell genomics and dynamical modeling**. Till completed an internship in the BioML team at Microsoft Research in Cambridge, MA, holds an MSc degree in Robotics, Cognition, Intelligence from TUM & a BSc degree in Industrial Engineering from Hannover University, Germany. *Previous organizing experience:* Till co-organised the previous LMRL workshop edition at ICLR 2025 and the Explainable AI workshop at ECCB in 2024.

Soo-Jeong Kim (sjk99@cam.ac.uk) is a PhD candidate at the University of Cambridge and the European Bioinformatics Institute (EMBL-EBI). Her research focuses on **developing generative models for protein design in low-data regimes, specifically for diverse protein families** that are underrepresented in existing datasets. She is an incoming fellow at the Sovereign AI Unit in the UK government, where she will focus on AI

for Science policy and develop new datasets that accelerate AI-driven scientific discovery. *Previous organizing experience:* Soo previously organized the 2025 AI × Bio Unconference in San Francisco, which brought together researchers, founders and engineers to explore the next frontiers of biological discovery.

Lazar Atanackovic (latanack@broadinstitute.org) is currently a Postdoctoral Associate at the Eric & Wendy Schmidt Center of the Broad Institute of MIT and Harvard. He is an incoming Assistant Professor at the University of Alberta and an incoming Research Fellow at the Alberta Machine Intelligence Institute. He completed his PhD at the University of Toronto and the Vector Institute with Brendan Frey and Bo Wang. During his PhD, Lazar spent time as an intern at Mila, Valence Labs / Recursion Pharmaceuticals, and Deep Genomics. His research focuses on **developing new generative methods for modelling and deciphering complex biological systems from data.**

Jason Hartford* (jason@valencelabs.com) is Dame Kathleen Ollerenshaw Fellow at the University of Manchester and Research Unit Lead and Staff Research Scientist at Valence Labs @ Recursion. Previously, he was a postdoctoral fellow with Yoshua Bengio at Mila & Université de Montréal. Before joining Mila, he completed his PhD at the University of British Columbia with Kevin Leyton-Brown. His research focuses on **developing new techniques for causal representation learning and causal inference**. *Previous organizing experience:* Jason is a previous co-organizer of the LMRL workshop 2025, UKRI Generative Biology Workshop and multiple workshops on machine learning at the University of Manchester. He also previously ran the Causality, Abstraction, Representation, and Extrapolation (CARE) reading group virtual seminar series.

Romain Lopez* (romain.lopez@nyu.edu) is an Assistant Professor at New York University, where he leads the NYU Biological Machine Learning Group. He is interested in interdisciplinary research focused on **understanding the complex interactions and dynamics of cellular populations, particularly in the immune system**. Romain completed his joint post-doctoral appointment between Genentech and Stanford University, hosted by Aviv Regev and Jonathan Pritchard. He holds a PhD from University of California, Berkeley, where he worked on computational, inferential and modelling perspectives of single-cell states, under guidance from Michael I. Jordan & Nir Yosef. *Previous organizing experience:* Romain co-organised the previous LMRL workshops at NeurIPS 2021 & NeurIPS 2022.

Thouis Ray Jones* (thouis@gmail.com) is a Senior Principal Computational Scientist at Arboretum LifeSciences, where he works on **developing computational and machine learning approaches to understand gene regulation and single-cell biology**. With over two decades of experience spanning computational bioinformatics, image analysis, and deep learning, his research bridges quantitative modeling with biological discovery. Ray has over 2 decades of experience, including his 12-year long appointment at the Broad Institute of MIT and Harvard as both Senior Scientist and Group Leader, pioneering high-throughput imaging techniques such as cell painting. He holds a PhD in Computer Science from MIT, where he focused on image processing and statistical modeling of biological data. *Previous organizing experience:* Ray was at the inception of the early iterations of the LMRL workshop and supported the organising committee throughout.

** These organisers have co-organised the LMRL workshop in the past.*

Committee Structures

Our **Organizing Committee** (OC) is intentionally larger to ensure a diverse representation of institutions, geographies, affiliations, seniority, and gender balance. This approach comprehensively covers many relevant research topics, creating an inclusive platform for attendees to submit their papers. By embracing diversity within the OC, we aim to enhance the experience for all participants. Our team has extensive experience to create this workshop, having previously served as invited speakers, poster award winners, and organizers of similar workshops at ICLR, NeurIPS, and other major AIxBio conferences. Importantly, our entire OC is

dedicated solely to proposing and organizing this workshop for ICLR 2025, reflecting our full commitment to this event.

Our **Programme Committee** (PC) currently consists of over 50 members from more than 15 institutions worldwide, recruited through our extensive connections in the field. Given LMRL's past successes and the rising popularity of ICLR workshops, excitement about our proposal is growing within our network. As the PC member list (see **Table 3** below) continues to expand, we will actively manage volunteer involvement to ensure broad topic coverage and a range of reviewer seniority for each submission. Each full paper track submission will require the nomination of *at least* one reviewer, allowing our PC to focus on reviewing submissions for the tiny paper track, which demands less time and seniority.

With the addition of nominated reviewers, our PC structure will ensure timely, comprehensive, and thoughtful reviews, with a maximum of five submissions per reviewer and a minimum of two reviews per submission in either paper track. To manage conflicts of interest, our extensive PC will guarantee that no OC member is involved in writing or curating the paper review process. We will implement a double-blind review policy and manage conflicts using the standard email domain matching system through OpenReview.

Diversity Statement

Diverse Committees Structure | Diversity is essential to our organizing team, fostering innovation, creativity, inclusivity, and cross-talk between academia and industry. We take pride in establishing a **gender-balanced** organizing committee, complemented by a diverse line-up of speakers and panelists for our workshop. Our team includes individuals across **various career stages**, from PhD students in their early to final years, to professors actively engaged in academia and university environments, as well as early-career Research Scientists to principal-level staff in **research-driven industries**. With 8 organizers from **8 affiliated institutions** spanning 4 distinct countries, we ensure a broad range of perspectives. Moreover, the scale and richness of our programme committee's affiliations and career stages will help us bring together a diverse community and provide insightful, domain-specific reviews.

Diverse Speaker & Panelist Line-Up | For this workshop, we have curated an exceptionally diverse lineup of researchers who will share their insights on a **wide array of AI research topics**. We are proud to have created **gender-balanced** and **ethnically diverse lineups** with a mix of seniority (principal investigators to industry executives) and **institutional diversity**, encompassing industrial (budding biotech companies to established technological giants) to academic and non-profit affiliations (universities and independent institutes). Prioritizing **geographical diversity**, we have successfully nominated speakers from institutions across the US, Europe, and the Middle East. We also recognize the unique opportunity presented by the conference in Rio de Janeiro to involve **local researchers** from the Latin America region, who are at the forefront of the meaningful AI representation learning community.

Diverse Authors & Poster Presenters | We remain committed to diversity, equity, and inclusion, believing that our workshop represents a significant opportunity for PhD candidates, post-doctoral researchers, and industry professionals to submit their work with the intention to (i) receive high-quality feedback, (ii) present at an event featuring world-renowned speakers and attendees, and (iii) contribute to discussions on advancing the field collectively. To additionally encourage underrepresented groups to submit their work to either paper track, we will advertise through affinity group email lists (WiML, Black in AI, MLDS-Africa, etc.) to foster a truly inclusive environment that embraces a variety of perspectives. Additionally, we plan to reach a broad range of researchers through a coordinated social media campaign to (i) announce workshop acceptance and (ii) invite submissions.

Diverse Attendees & Representatives | We believe in fostering meaningful networking and collaboration opportunities throughout the workshop, aiming for mutual engagement rather than a one-sided experience. It

is essential for us to encourage poster and spotlight presenters to actively engage with established leaders and innovators in the field, ensuring their contributions receive ample attention and constructive feedback. To this end, we are inviting prominent industry leaders to join us for the entire day of the workshop. We will finalize these nominations as the date approaches, recognizing that travel arrangements for these attendees may not yet be fully confirmed. In the meantime, we are actively generating support and excitement for this workshop. Our outreach efforts will include a social media campaign and targeted communications aimed at attracting high-profile individuals from industrial research institutions to engage with our poster presenters. By planning outreach and advertising in this way, we aim to build a strong audience while also addressing diversity and inclusion: specifically, increasing exposure of women to AI in the context of life sciences and biological research. ML research is often male-dominated, while practical AI applications in biology are typically overrepresented by women, so these efforts help close the gender gap while providing networking opportunities that meet the diverse interests and career aspirations of our attendees.

Programme Committee

Tentative Programme Committee Members include: Xinming Tu, Kexin Huang, Yanay Rosen, Kasia Kedzierska, Michael Vinyard, Ruitong Li, Alejandro Tejada-Lapuerta, Anna Schaar, Berton Earnshaw, Michael Craig, Cristian Gabelini, Frederick Wenkel, Dominique Beaini, Shawn Whitfield, Hanene Benyedder, Semih Canturk, Konstantin Donhauser, Ihab Bendidi, Auguste Genovesio, Thouis “Ray” Jones, Elizabeth Wood, Johann Wenckstern, Stefan Stark, Frederike Lübeck, Vignesh Ram Somnath, César Miguel Valdez Córdova, Matthew Scicluna, Joseph Viviano, Velina Kozareva, Delaram Pouyabahar, Duncan Forster, Sam Cooper, Elias Williams, Conor Tilinghast, Safiye Celik, Youssef Barhomi, Stephen MacKinnon, Kyle Hansen, Ben Fogelson, John Urbanik, Peter McLean, Marissa Saunders, Nathan Lazar, Saber Saberian, Grant Watson, Rory Blucher, Jay Huang, Arin Minasian, Kian Kenyon-Dean, and other willing to volunteer their free time to make LMRL a successful event.

References

Adduri, A. K., Gautam, D., Bevilacqua, B., Imran, A., Shah, R., Naghipourfar, M., Teyssier, N., Ilango, R., Nagaraj, S., Dong, M., Ricci-Tam, C., Carpenter, C., Subramanyam, V., Winters, A., Tirukkovular, S., Sullivan, J., Plosky, B. S., Eraslan, B., Youngblut, N. D., Leskovec, J., Gilbert, L. A., Konermann, S., Hsu, P. D., Dobin, A., Burke, D. P., Goodarzi, H., & Roohani, Y. H. (2025). *Predicting cellular responses to perturbation across diverse contexts with State. bioRxiv.* https://doi.org/10.1101/2025.06.26.661135

Ahlmann-Eltze, C., Huber, W., & Anders, S. (2025). *Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. Nature Methods, 22*, 1657–1661. https://doi.org/10.1038/s41592-025-02554-0

Boiarsky, R., Singh, N. M., Buendia, A., Amini, A. P., Getz, G., & Sontag, D. (2024). *Deeper evaluation of a single-cell foundation model. Nature Machine Intelligence, 6*, 1443–1446. https://doi.org/10.1038/s42256-024-00981-9

Bunne, C., Roohani, Y., Rosen, Y., … Lundberg, E., Leskovec, J., & Quake, S. R. (2024). *How to build the virtual cell with artificial intelligence: Priorities and opportunities. Cell, 187*(25), 7045–7063. https://doi.org/10.1016/j.cell.2024.11.001

Noutahi, E., Hartford, J., Tossou, P., Whitfield, S., Denton, A. K., Wognum, C., Ulicna, K., Craig, M., Hsu, J., Cuccarese, M., Bengio, E., Beaini, D., Gibson, C., Cohen, D., & Earnshaw, B. (2025). *Virtual cells: Predict, explain, discover. arXiv preprint arXiv:2505.14613.* https://doi.org/10.48550/arXiv.2505.14613

Klein, D., Fleck, J. S., Bobrovskiy, D., Zimmermann, L., Becker, S., Palma, A., Dony, L., Tejada-Lapuerta, A., Huguet, G., Lin, H.-C., Azbukina, N., Sanchís-Calleja, F., Uscidda, T., Szalata, A., Gander, M., Regev, A., Treutlein, B., Camp, J. G., & Theis, F. J. (2025). *CellFlow enables generative single-cell phenotype modeling with flow matching. bioRxiv.* https://doi.org/10.1101/2025.04.11.648220

Rood, J. E., Hupalowska, A., & Regev, A. (2024). *Toward a foundation model of causal cell and tissue biology with a Perturbation Cell and Tissue Atlas. Cell, 187*(17), 4520–4545. https://doi.org/10.1016/j.cell.2024.07.015

Wenkel, F., Tu, W., Masschelein, C., Shirzad, H., Eastwood, C., Whitfield, S. T., Bendidi, I., Russell, C., Hodgson, L., El Mesbahi, Y., Ding, J., Fay, M. M., Earnshaw, B., Noutahi, E., & Denton, A. K. (2025). *TxPert: Leveraging biochemical relationships for out-of-distribution transcriptomic perturbation prediction. arXiv preprint arXiv:2505.14919.* https://doi.org/10.48550/arXiv.2505.14919