

Mariya I. Vasileva

☎ (626) 236-8343
✉ maria.i.vasileva@gmail.com

🌐 Website
LinkedIn

🎓 Google Scholar
GitHub

RESEARCH
INTERESTS Machine Learning, Computer Vision, Deep Learning, Generative Models
Vision and Language, Multimodal Understanding & Generation, Video Intelligence
Model Posttraining, Synthetic Data, Multimodal Trust & Safety, Scaled Evaluations

EDUCATION **University of Illinois** —Urbana-Champaign, IL

Ph.D., Computer Science, August 2020
Advisor: David A. Forsyth
Thesis Committee: Alexander Schwing, Derek Hoiem, Tamara L. Berg

California Institute of Technology —Pasadena, CA

B.S., [Mechanical Engineering](#), May 2013
Advisors: Joel W. Burdick, John O. Dabiri

B.S., Business Economics and Management, May 2013
Advisor: Colin Camerer

Control and Dynamical Systems (Minor), May 2013
Thesis Advisor: Richard M. Murray

SELECT
PUBLICATIONS

A. Xu, **M. I. Vasileva**, A. Seshadri. *HandsOff: Labeled Dataset Generation with No Additional Human Annotations*. [Highlight at CVPR 2023](#).

M. I. Vasileva, B. A. Plummer, V. Petsiuk, K. Saenko, D. Forsyth. *Why do These Match? Explaining the Behavior of Image Similarity Models*. ECCV 2020.

R. Tan, **M. I. Vasileva**, K. Saenko, B. A. Plummer. *Learning Similarity Conditions Without Explicit Supervision*. ICCV 2019.

M. I. Vasileva, B. A. Plummer, K. Dusad, S. Rajpal, R. Kumar, D. Forsyth. *Learning Type-Aware Embeddings for Fashion Compatibility*. ECCV 2018.

PATENTS

M. I. Vasileva, K. Li, K. Dusad, B. A. Plummer, Y. Shen, S. Rajpal, R. Kumar, D. Forsyth. *Search Engine Use of Neural Network Regressor for Multi-Modal Fashion Recommendations Based on Visual Semantic Embeddings*.

U.S. Patent No. US-12131365-B2. Issued Oct 29, 2024.
University of Illinois at Urbana-Champaign.

V. Shankar, A. Dave, **M. I. Vasileva**, D. Pal, Z. Corneli, J. Rehg, A. Lu, G. Medioni. *Selecting Articles of Clothing for Customers Based on Visual Relevance*. U.S. Patent Application No. 17/706,122. Mar 2022. Amazon.

M. I. Vasileva, N. Bodla, R. Sarkar, A. Beniwal, A. Lu, G. Medioni. *Voice-Driven Outfit Completion in Physical Stores*.
U.S. Patent Application No. 17/540,865. Dec 2021. Amazon.

A. Malik, A. Lu, A. Beniwal, B. Hardenstein, G. Medioni, J. McAuley, K. Nar, **M. I. Vasileva**, R. Wang, S. Sun, V. Shankar, V. Chaturvedi, W. Zhang. *A Multi-Model Approach to Address Relevance, Diversity, Physical Fit, Outfit Complementarity, Discovery, and Trend in Fashion Recommendation.*
U.S. Patent Application No. 17/218,081. Mar 2021. Amazon.

PROFESSIONAL
AND RESEARCH
EXPERIENCE

Senior Research Scientist, Meta Superintelligence Labs (MSL) Mar 2025 –
New York, NY

- Designed and implemented rigorous, large-scale evaluation frameworks for AI trust, safety, and alignment with a focus on systematic benchmarking, quantitative risk assessment, and scalable oversight of multimodal foundation and frontier models across high-impact domains (eg, the [Llama-4 herd](#))
- Built and operationalized comprehensive benchmark suites for multimodal safety, covering image and video analysis, captioning and summarization, grounding, retrieval, and generation tasks; benchmarks addressed youth protections, complex bias evaluation, cross-lingual visual understanding consistency, and multimodal content integrity analysis – enabling rigorous performance characterization of model safety and alignment under realistic deployment conditions
- Developed scalable synthetic data generation pipelines to discover and mitigate context-dependent bias in multimodal models in pre- and post-training stages, advancing research on bias generalization, safety adaptation, and eval robustness
- Architected and implemented [Youth Safety](#) evaluation benchmarks as part of a company-wide initiative to deliver a PG-13 experience for under-18 users; operationalized evolving safety policy into measurable evaluation criteria and drove data collection, benchmark design, and scaled evaluation in cross-functional partnership with product, infra, legal, and policy teams (more [here](#) and [here](#))
- Designed and implemented end-to-end scalable evaluation ecosystems spanning policy operationalization, large-scale vision data sourcing and annotation workflows, LLM/VLM-as-a-judge framework development and calibration against human baselines, and risk metric design and instrumentation, enabling continuous oversight of model behavior in close collaboration with cross-functional partners.
These evaluation frameworks now underpin all MSL pre-release model and product testing.

Senior Applied Scientist, AWS Responsible AI
Amazon, San Francisco, CA

Jun 2022 – May 2023

- Researched and developed Responsible AI (RAI) methodologies such as bias detection and evaluation, ethical data sourcing and annotation, application-specific fairness metrics, defenses against membership inference attacks, and tooling for model interpretability and explainability
- Applied RAI frameworks to designing and auditing a variety of machine learning and computer vision models in sensitive use cases, such as face recognition, face landmark detection, face attribute prediction, content moderation, gaze prediction, face liveness detection, multimodal reasoning, open-vocabulary object detection and classification, and text-to-image generation
- Prepared model service cards outlining model design, intended use cases and limitations, fairness and bias assessment, and robustness analysis for a number of services offered by Amazon Rekognition: [Compare and Search Faces](#), [Detect and Analyze Faces](#), [Detect Face Liveness](#), [Detect Objects and Concepts](#), and [Moderate Content](#)

- Consulted on model guardrails for two of Amazon Bedrock's **Titan** foundation models (FMs): **Titan Multimodal Embeddings** (used for large-scale visual search and retrieval, recommendation and personalization), and **Titan Image** (text-to-image generation with inpainting and outpainting capabilities)
- Translated complex research concepts into digestible insights for stakeholders across the organization: from scientists and engineers working on customer-facing product, to legal experts, to executives; participated in setting OKRs and designing project roadmaps
- Integrated RAI practices across AWS's AI services lifecycle to promote safe interactions between users and AI applications, and heavily influenced the focus of internal AI development ideologies with results-driven findings on model safety and fairness performance
- Developed a **framework based on Generative Adversarial Networks (GANs) inversion** for producing an unlimited number of synthetic images and their corresponding pixel-wise labels after being trained on less than 50 preexisting labeled images, and demonstrated SOTA performance on several downstream tasks like semantic segmentation, keypoint detection, and depth estimation compared to prior dataset generation approaches and transfer learning baselines (*highlight at CVPR 2023*)

Head of Research, Kaiber.ai

May 2023 – Dec 2023

Los Angeles, CA

- Headed a research team developing text-to-image, text-to-video, and image-to-video diffusion models
- Productionized an image-to-video diffusion model consisting of a text-to-image base generator finetuned on a combined dataset of images and video clips, and a motion modelling module trained to represent plausible and temporally-consistent motion (open-source code [here](#))
- Developed an efficient memory-augmented diffusion-based video editing approach competitive with then-SOTA on metrics assessing perceptual quality, structural and temporal consistency, with a 5x speedup in VRAM use and 2x speedup in inference time
- Developed a large-scale video recommendation and ranking system integrating dense frame extraction, multimodal visual-semantic embeddings, candidate feed generation, and efficient similarity search over a vector database for ranking and retrieval optimization
- Developed a research roadmap of open problems at multiple levels of complexity, risk, ambiguity, and resource demand with the goal of maximizing product impact and creating a competitive advantage

Applied Scientist, Amazon Style

Nov 2020 – Jun 2022

Amazon, San Francisco, CA

- Served as one of the first applied scientists on a new initiative for building a physical store with large inventory and a machine-learning-and-computer-vision powered intelligent item recommendation system, which surfaces items to the user in real time based on historic interaction data, collaborative filtering techniques, and visual relevance
- Built, evaluated and productionized models for computer vision applications in the fashion domain, such as visual search and retrieval of relevant clothing items

- for a user, personalized style modelling, body shape and pose estimation from a single 2D image, virtual try-on, and fashion item compatibility modelling
- Built a pipeline for iterative human-in-the-loop data collection, annotation and validation, as well as frameworks for label disambiguation and explainability, to gather a high-quality dataset of fashion images, corresponding attribute annotations, and expert-provided labels on personal style and fit
 - Developed a method for solving two critical tasks for a fashion recommendation system: compatibility prediction (ie, the task of determining whether a set of fashion items in an outfit go well together), and large-scale complementary item retrieval (ie, the task of completing a partial outfit by finding a compatible item from a large database) via a visual transformer model (*WACV 2023*)

**Graduate Research Assistant
Computer Vision Group, UIUC**

Advisor: David A. Forsyth

May 2016 – Jul 2020

- Worked on a range of research projects at the intersection of computer vision, machine learning and deep learning: specifically, building multimodal embedding models for visual search and retrieval, representation learning, reasoning about complex relationships between images, generative models for controllable semantic image editing in an attribute-guided manner, and model explainability
- Developed a contrastive learning approach for producing an image embedding for fashion items that respects clothing item type, and jointly learns notions of item similarity and compatibility in an end-to-end trained model
- Introduced the most-cited dataset of fashion images for compatibility learning: [Polyvore Outfits](#) (*ECCV 2018*)
- Developed an approach to learning type-dependent similarity conditions as a latent variable for scalable learning of item similarity and compatibility that results in semantically-meaningful embedding subspaces (*ICCV 2019*)
- Developed an interpretable explanation system for embedding models that score image similarity between a pair of input images – unlike standard single-image classifiers – and output reference-dependent saliency maps paired with natural language attribute explanations per query-reference pair (*ECCV 2020*)
- Developed a method for diverse automatic capsule wardrobes generation via an end-to-end trained multi-task model for jointly learning visual similarity between fashion items, and predicting style attributes
- Developed a method for style summarization, probing and exploiting the space of fashion outfits (modelled as "bags" of fashion items) that allows for complex search queries and produces visual outfit analogies, by learning outfit-level embeddings with the help of a contrastive model and an embedding topic model

**Machine Learning Engineering Intern
KaleidoGlobe, Boston, MA**

Feb 2019 – Aug 2019

- Developed tools for various natural language processing and information retrieval tasks, such as data collection, training of text embedding models, document clustering and summarization, entity extraction and coreference resolution, and knowledge graph mining

Machine Learning Research Intern Butterfly iQ, New York, NY	May 2018 – Sep 2018
– Developed an attribute-conditioned Generative Adversarial Network (GAN) model for the purpose of data augmentation and distributional gap filling in unbalanced medical image datasets	
 Graduate Research Assistant Virtual Reality Group, UIUC Advisor: Steven M. LaValle	Feb 2015 – May 2015
– Explored the feasibility of shared environments and virtual social spaces for applications in education such as medical training and foreign language learning	
– Prototyped virtual reality social spaces for joint problem solving in language learning; worked on avatar creation, synchronization, and shared interaction with virtual objects	
 Summer Undergraduate Research Fellow Robotic Manipulation and Sampling Group NASA Jet Propulsion Laboratory Advisor: Joel Burdick (Caltech)	Jun 2012 – Sep 2012
– Developed signal processing algorithms for inferring local geometry and contact surface mapping for rigid objects in robot manipulation challenges "in-the-wild" using tactile sensor feedback	

TECHNICAL SKILLS	<p>Post-Training & Fine-Tuning: Supervised fine-tuning (LoRA, ControlNet, prompt tuning), RLHF-adjacent human feedback evaluation and calibration, model steering and behavioral alignment via systematic prompting, evaluation of post-trained multimodal models, benchmark development for capability and safety tracking</p> <p>Model Development & Evaluation: Large-scale benchmark development and operationalization, VLM/LLM-as-a-judge frameworks, human baseline calibration, capability tracking across training iterations, model capability and safety evaluation</p> <p>ML/CV: PyTorch, Python, Torchvision, Diffusers, Transformers, ControlNet, LoRA, ffmpeg, OpenCV, scikit-learn, numpy, pandas</p> <p>Infra & MLOps: AWS, Hydra, Hugging Face, FB Learner Flow, ModelScope, Docker, Cog, Weights and Biases</p> <p>Data & Pipelines: Large-scale synthetic data generation, ETL workflows, human-in-the-loop annotation pipelines; S3, DynamoDB, Hive, Scuba, Presto, Dataswarm, Pinecone, MongoDB, MySQL</p> <p>Other: Git, GitHub, GitLab, CircleCI, Linux, bash, vim</p>
CONFERENCE WORKSHOPS	<p>16th Women in Machine Learning Workshop (WiML) cohosted with NeurIPS 2021: General Chair. I managed the end-to-end organization of the workshop, and oversaw the program preparation, the paper submission and review process, and all communications with the main conference organizers, mentors, advisors, invited speakers, and student authors.</p> <p>7th Workshop on Computer Vision for Fashion, Art, and Design (CVFAD 2024) cohosted with CVPR 2024: Organizer</p> <p>6th Workshop on Computer Vision for Fashion, Art, and Design (CVFAD 2023) cohosted with CVPR 2023: Organizer</p>

5th Workshop on Computer Vision for Fashion, Art, and Design (CVFAD 2022) cohosted with CVPR 2022: Organizer

ResistanceAI Workshop cohosted with NeurIPS 2020: Organizer

Broadening Participation in Data Mining Workshop (BPDM) cohosted with KDD 2017: Organizer

PROJECTS IN
MEDIA

Instagram Teen Accounts Will Be Guided by PG-13 Ratings

Empowering Parents, Protecting Teens: Meta's Approach to AI Safety

Helping Teens See Age-Appropriate Content

Llama Protections: An Open Approach to Protections in the Era of Generative AI

Introducing AWS AI Service Cards: A New Resource to Enhance Transparency and Advance Responsible AI

Announcing New Tools and Capabilities to Enable Responsible AI Innovation

AWS re:Invent 2023 - Responsible AI in the Generative Era: Science and Practice

Responsible AI in the Wild: Lessons Learned at AWS

Amazon Reimagines In-Store Shopping with Amazon Style

Amazon Style Seeks to Integrate Tech and Fashion

What the Amazon Style Store is REALLY Like?!

PROFESSIONAL
MEMBERSHIPS

Women in Machine Learning

- WiML Mentor at ICML, NeurIPS (2023 –)
- WiML Workshop General Chair (2021)
- WiML Workshop Social Host and Volunteer (2020)
- WiML Workshop Volunteer and Reviewer (2016 – 2019)
- WiML Workshop Author (2016 – 2017)

Women in Computer Vision

Machine Learning Collective

Queer in AI

Caltech Alumni Association

ACADEMIC
SERVICE

Area chair: CVPR, ICCV, ECCV

Conference reviewer: CVPR, ICCV, ECCV, NeurIPS, ICML, WACV, FAccT

Student volunteer: NeurIPS, ICML, ICLR, SIGGRAPH (2016 – 2019)

Organizer, Deep Learning Reading Group (UIUC, 2016-17)

TEACHING
EXPERIENCE

**CS 598LAZ: Cutting-Edge Trends in
Deep Learning and Recognition**

May 2017

CS 498DV: Data Visualization

May 2020 – Aug 2020

CS 498DF: Applied Machine Learning

Jan 2016 – May 2016

CS 445: Computational Photography

Aug 2015 – Dec 2015

FELLOWSHIPS

Mechanical Science and Engineering Excellence Fellowship (UIUC, 2014)

John W. and Herberta M. Miles Endowed Scholarship (Caltech, 2011-12, 2012-13)

Summer Undergraduate Research Fellowship (Caltech, 2012)

CS@Illinois Grace Hopper Celebration scholarship (2016, 2017)

Qualcomm Women's EmpowHERment Summit scholarship (San Diego, CA, 2015)

Bulgaria National Mathematical Linguistics Olympiad top-10 2005, 2007, 2008