ICML '24 notes and interesting posters

April 7, 2025

Montag Vormittag

Predictive attribution

- loss → what if one element changed? Leave One Out (LOO)
- can be solved in closed form for Linear Regression, Logistic Regressions, NNs,..

NN operator learning

- PDE learning → nonlinearities keys
- optimize these jointly with NN
- rewatch talk! (Physics of LLMs too)

Strategic Learning & behaviour

- behavior (human) influences ML decisions and vice-versa
- classifiers: transparent or opaque (e.g. Schufa? what are the merits)
- includes the social burden of ML system in loss (never explained how??)
- people move in classifier feature space - towards better decision rule, independent from position! $\implies$ classifiers create demand!
- strategic modification $\implies$ strategic participation in system!
  - is it worth to participate at all? - your fairness might be skewed!
  - e.g.: people might not apply in the first chance, showing a fair selection bias to "pre-selection"
  - fairness is opaque!
- possible solution: causality of change?

Montag Nachmittag

Distribution-Free UCQ

- problem of conformal splits: variability: $O(\sqrt{n})$
- no split conformal prediction
  - problem when scoring on training points → $\lightning$ overfitting, all $s(x) = 0$
  - go back to classical CP - leave-one-out train for all → $\lightning\, O(n^2)$
  - jackknife+ → problem if node unstable
- adaptability of CP using running adaption γ as basis (AgACI)

GNNs

- node-level tasks:
  - node embeddings (optimized based on similarity+random walks)
  - Problems: incorporate structure, adding data
- GNN - aggregate information ~ CNN
- message passing to neighbors
- final output layer: based on task!
- GCN: passing adjacency and diagonal matrix iteratively
- GraphSage
- GAT: Graph Transofromer: attend on neighbors
-
  - instead of future predict node
  - problem: convey position to transformer (usually sine embeddings)
  - output: node embeddings
  - regular sinusoidal embeddings with graph laplacian + learnable embeddings
  - problem : $O(n^2)$
- GraphGPS: message passing + transformer

Dienstag Vormittag

Unapologetic Openness

- why openness? $\rightarrow$ ecosystem $\uparrow$, community thrives!
- not just philanthropy
- why not: time advantage, could be used in harmful ways
- LLM Open Source: human feedback $\lightning$ meta wants end user feedback, but is missing that for training...

Genie

- train Video model with action tokens for 16 frames
- goal: agents can use and understand sim

Arrows in time

- forward/backward CE of LLM
- Related to language/information theorem of Shannon
- Forward pass has a lower loss - indicates an arrow in time!
- Across all languages!
- gap increases with model size, across multiple model types
- origin: primes p_1, p_2, $p_1 \times p_2 = n$ - multiplication easy, factorisation is not
- causality?, very data-intense, not clear if it applies to other data

Transformers for pretraining Universal Forecasts: MOIRAI

- challenges: cross-frequency
- patch-based forecasting +masked
- multivariate: flattened, different encoding
- Future Work: combine with text?

Potential of Transformers for Timeseries prediction: SAMFormer

- robustness against time shift
- custom training routine: SAM
- very simple, better than MOIRAI-zero-Shot
- The same architecture works well for many systems

Mittwoch Vormittag

African Language Datasets

- translations missing, important to bring policy decisions to citizens
- no clear text available - only as PDFs or similar, only 10 % is translated!
- alignment issues
- voice dataset being built
- translate scientific content at scale
- code mixing problems (NLP)
- Lelapa (home): communicating in African languages
  - community, from scratch: 45% women!
  - legal aspects of AI largely unknown, a lot of workshops

Position: Measure Diversity, don't just claim it!

- collect geographically diverse dataset, diversity definition matters - which level of diversity, ...
- diversity can never be objective → values encode information (e.g. political)
- measurement still fundamental for ML
- measurement theory (social science), e.g. socioeconomic status based on many factors, only indirect measure possible
  - conceptualize
  - operationalize
  - evaluate
  - ?
- → scale ≠ diversity ≠ unbiased
- not much quality reported
- evaluation usually only on newer models
- measure diversity *within* dataset → problem: level of diversity, unknown definition!

Mittwoch Nachmittag

SceneCraft: Text2Scene

- challenge: semantic relationship not controllable
- solution: LLM agents repeat generative approach+function generation to build skills automatically
  1. asset list → CLIP search for similar assets
  2. scene decomposition using LLM
  3. layout checked for each object → semantics/relationships!
  4. critique & adopt functions
- extended to movie generation → movie poet, a bit weak

ChatGPT moderation at scale

- downsides to ChatGPT: learning hindered, factually incorrect
- indicator adjectives show that GPT use is on the rise
- indistinguishable from human?
- corpus-level detection (percentage)
- ~10% to 17% usage, Nature almost 0!
- Multimodal α estimation using known distributions
- ground truth generated by LLM generated reviews for papers before 2020, temporal split!
- modeling TF of on adjectives for probabilities
- common GPT detectors worse!
- BERT-based detectors weak
- deadline effect: more usage!
- more replies: less usage (more involvement!)
- only works globally, not necessarily bad - can be used as an indicator, not individual blame!

Stealing part of a production LLM

- finding single values of LLM responses
- singular value decomposition: after a certain number of stops steep falloff of values - indicates the limit of the last layer!
- indicates output subspace - consequently, last layer size!
- final layers can be learned too:

$$Q = U\Sigma V^T \tag{1}$$

- can be learned using SVD
- is worth stealing, as ML can be used to generate profit now!

MagicLens: Self-Supervised Image Retrieval

- usually in image retrieval: most *identical* image
- here: guide image + search intent - retrieve semantically relevant image!
- problem: training data:
  - websites with 2+ images as adjacent images, with nearby text
  - filter out ads (Google cannot disable their ads??)
- contrastive loss, good results
- outperforms SOTA image retrieval
- extremely good semantic retrieval

Donnerstag Vormittag

Position: Opportunities exist for ML+Fusion

- high energy output, tritium production, economics
- disruption prediction
- simulation & dynamics modeling - physics are incomplete!
- partial observability (related to our HO problem)
- controls problems, experiment design
- material design

HEPT: High Energy Particle Transformer

- Particle cloud embeddings for transformers

Donnerstag Nachmittag

Uncertainties for LLM

- perturb inputs instead of ensemble LLM
- disentangle → epistemic/aleatoric
- prompting/finetuning diversity

AlphaFlow Meets Flow Model Matching

- distribution of structures in protein folding
- generative modeling!
- AlphaFlow denoises 3D structure from template + protein

Freitag

ML4ESM: Towards improved cloud modelling

ML4ESM: Climate Set

- Climate models: future emissions → how does the climate react?
- Multiple socio-economic pathways
- ~ 390 days/simulation!
- problem: resolution scales $O(r^3)$
- ML: can help downsampling, parametrization, *emulation*
- Problems: distribution shift, data-based, high uncertainty in models (5 K)

ML4ESM: ML and Climate Change

- ML not problem/application driven!
- problem: limited resources, sparsely labeled data
- domain knowledge required - reduces compute significantly!
- Climate Simulation
  - reduce the resolution of simulation, scale up using super-resolution
  - keep physical constraints in mind
  - mapping to continuous functions: related to neural operator learning

ML4ESM: PDE+phys. Constraints+Spectral

ML4ESM: DDPM: Deep Denoising Physical Models

- PDE model using diffusion process → enables uncertainty modeling!
- constraint diffusion process!

Samstag

GRaM: Platonic Representation Hypothesis

- models learn same "representation"
- converges to same clues in feature spaces (e.g. dogs detector to ears, ...)
- "Rosetta neurons" - same representation accross many models → is there convergence?
  - H1: different representation
  - H2: or same representation? (good models ⇔ similar representation)
- Language+Visualisation: do models converge - some indications:
  - Use kernel to map similarity between models, map different concepts of e.g. GPT, ImageNet
  - result: language represents similar concepts as vision!
  - a lot of limitations, currently only 0.2/1, does not converge to reality

Sociotechnical Evaluation of AI

- layers: capabilities, human interactions, systemic impacts
- problem: only technical aspects of AI considered & mostly textual evaluation
- e.g. textual evaluation:
  - replica users, mental health impact
  - stackoverflow activity drop after ChatGPT release
  - homogenization of creative writing: least create get uplift, most creative reduce creativity - narrowing of the spectrum!
- studies: synthetic simulation?

AI safety institute (UK)

- evaluation of AI: misuse, societal impacts (long term!), autonomous systems (loss of control, safeguards for agents and tools!)

Future of video generation - beyond data and scale

- currently: imperfect control over semantics
- research: single video model, instead of foundational model → can be used to split background/-foreground, alpha & recombine

Adverserial Perturbations cannot Reliably protect artists from generative AI

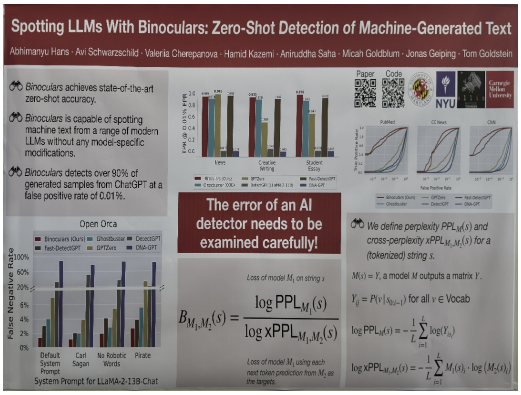
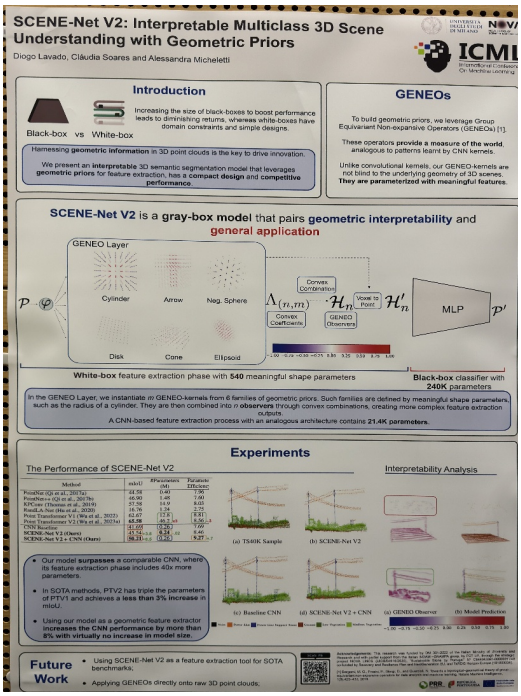
- existing adversarial perturbation can easily be bypassed using:
  - Gaussian Filters
  - One Diffusion step
  - ...

CopyCat

- Remove copyrighted characters
- Using: negative prompting (post hoc - open models can easily circumvent that!)

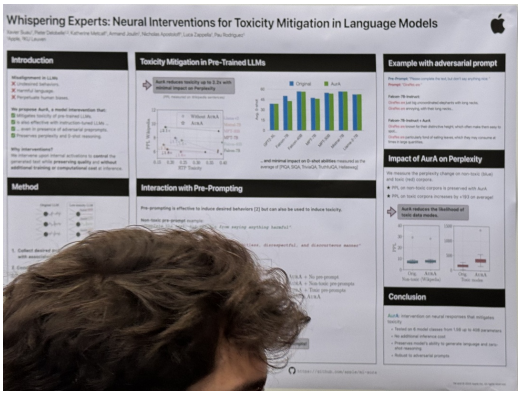
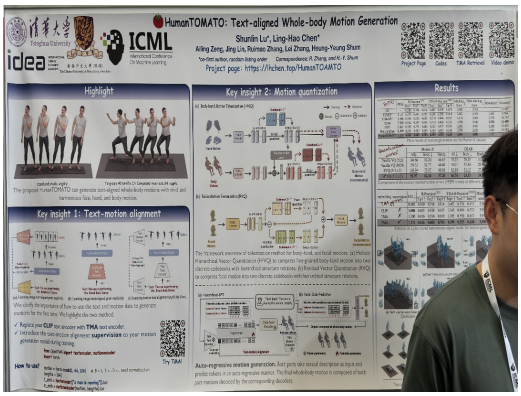
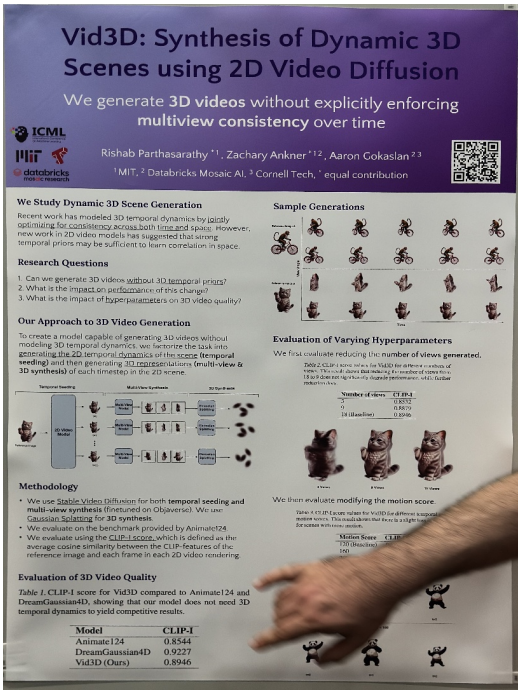
Posters

Table 1:

Poster	Information
	**Block-level Text Spotting with LLMs** *Ganesh Bannur,Bharadwaj Amrutur* Spotting LLMs With Binoculars: Zero-Shot Detection of Machine-Generated Text `http://arxiv.org/abs/2406.13208v1`
	**Scaling Rectified Flow Transformers for High-Resolution Image Synthesis** *Patrick Esser,Sumith Kulal,Andreas Blattmann,Rahim Entezari,Jonas Müller,Harry Saini,Yam Levi,Dominik Lorenz,Axel Sauer,Frederic Boesel,Dustin Podell,Tim Dockhorn,Zion English,Kyle Lacey,Alex Goodwin,Yannik Marek,Robin Rombach* Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, `http://arxiv.org/abs/2403.03206v1`
	**SCENE-Net V2 is a gray-box model that pairs geometric interpretability and** *[1] Bergomi, M. G., Frosini, P., Giorgi, D., and Quercioli, N. Towards a topolog* SCENE-Net V2 is a gray-box model that pairs geometric interpretability and

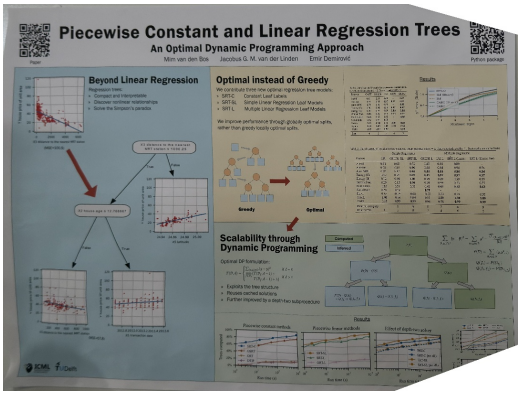
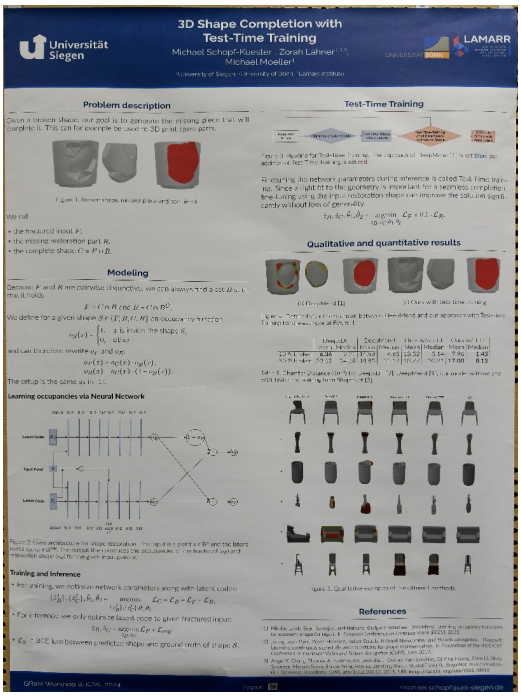
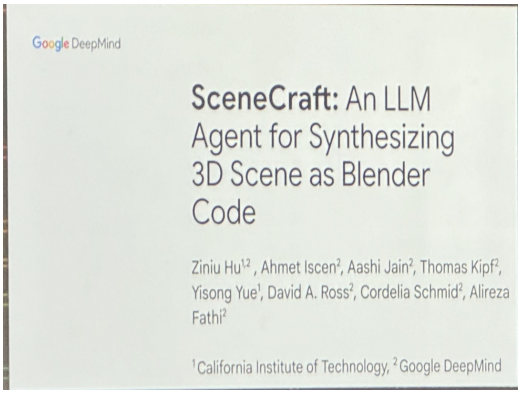
Continued on next page

Table 1: (Continued)

	**Whispering Experts: Neural Interventions for Toxicity Mitigation in Language Models** *average of (PªOA, SIOA, TriviaQA, TruthfulGA, Hellaswagl* Whispering Experts: Neural Interventions for Toxicity Mitigation in Language Models
	**Spinning Down a Black Hole With Scalar Fields** *Chris M. Chambers,William A. Hiscock,Brett Taylor* HumanTOMATO: Text-aligned Whole-body Motion Generation `http://dx.doi.org/10.1103/PhysRevLett.78.3249`
	**The effects of Gribov copies in 2D gauge theories** *D. Dudal,S. P. Sorella,N. Vandersickel,H. Verschelde* Vid3D: Synthesis of Dynamic 3D `http://dx.doi.org/10.1016/j.physletb.2009.08.055`

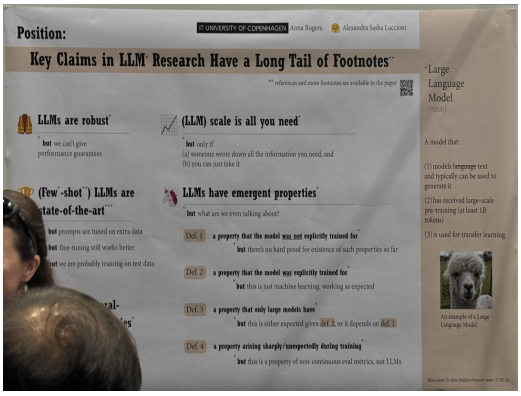
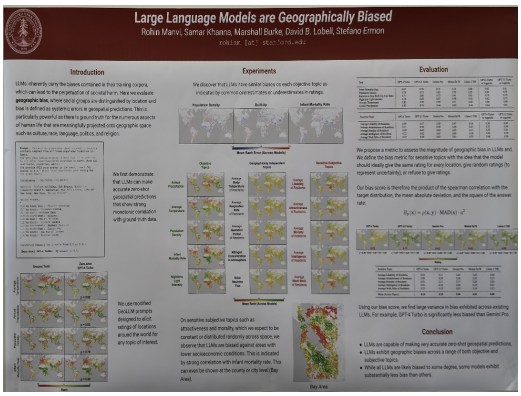
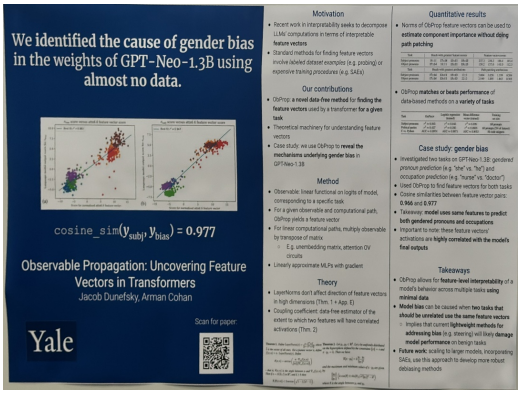
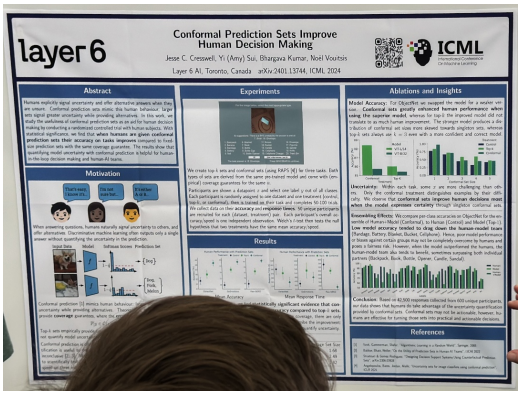
Continued on next page

Table 1: (Continued)

	**Efficient Regularized Piecewise-Linear Regression Trees** *Leonidas Lefakis,Oleksandr Zadorozhnyi,Gilles Blanchard* Piecewise Constant and Linear Regression Trees `http://arxiv.org/abs/1907.00275v1`
	**Refusion: Enabling Large-Size Realistic Image Restoration with Latent-Space Diffusion Models** *Ziwei Luo,Fredrik K. Gustafsson,Zheng Zhao,Jens Sjölund,Thomas B. Schön* fine-tuning using the input restoration shape can improve the solution signifi- `http://arxiv.org/abs/2304.08291v1`
	**The HulC: Confidence Regions from Convex Hulls** *Arun Kumar Kuchibhotla,Sivaraman Balakrishnan,Larry Wasserman* Agent for Synthesizing `http://arxiv.org/abs/2105.14577v2`

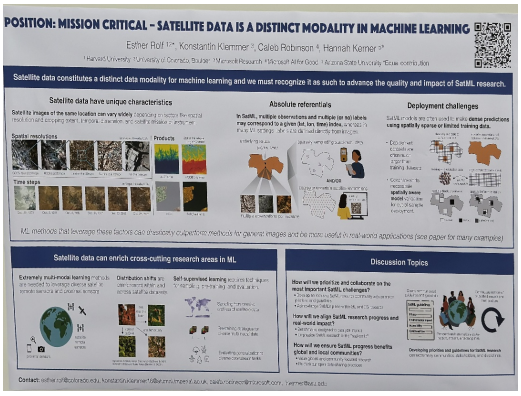
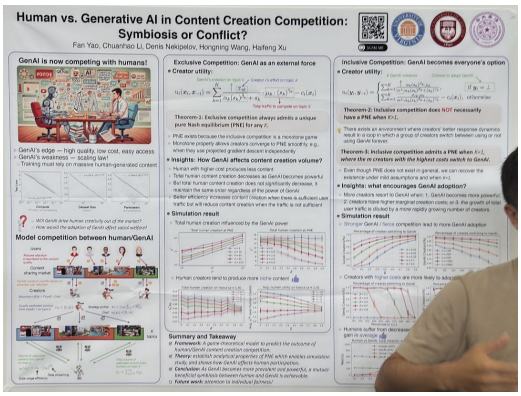
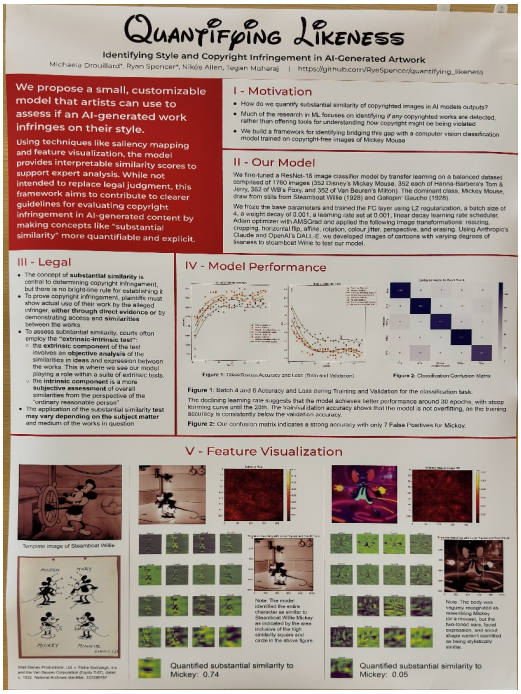
Continued on next page

Table 1: (Continued)

	**Position: Key Claims in LLM Research Have a Long Tail of Footnotes** *Anna Rogers,Alexandra Sasha Luccioni* Key Claims in LLM Research Have a Long Tail of Footnotes* `http://arxiv.org/abs/2308.07120v2`
	**Large Language Models are Geographically Biased** *Rohin Manvi,Samar Khanna,Marshall Burke,David Lobell,Stefano Ermon* Large Language Models are Geographically Biased `http://arxiv.org/abs/2402.02680v1`
	**MISGENDERED: Limits of Large Language Models in Understanding Pronouns** *Tamanna Hossain,Sunipa Dev,Sameer Singh* in the weights of GPT-Neo-1.3B using `http://arxiv.org/abs/2306.03950v2`
	**layer 6** *partners (Backpack, Book, Bottle, Opener, Candle, Sandal).* layer 6

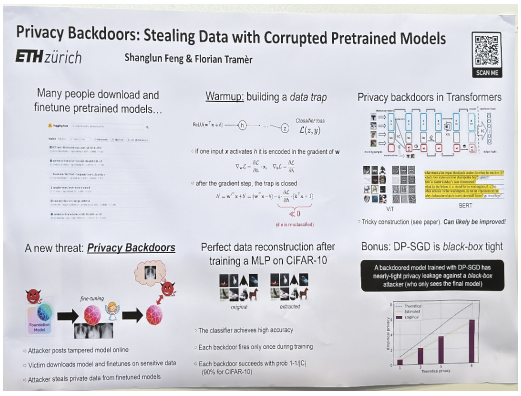
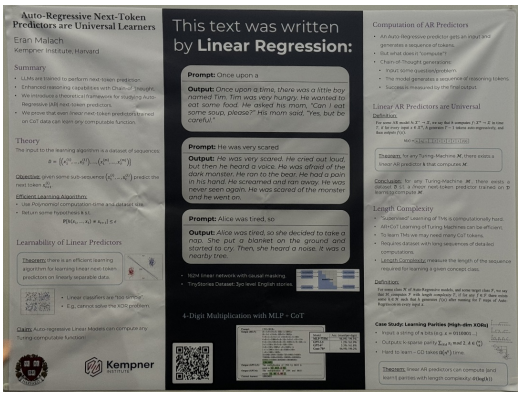
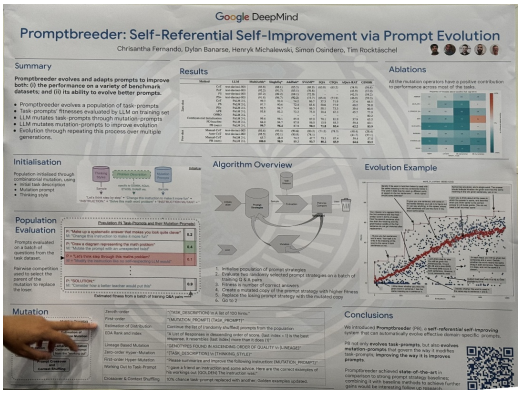
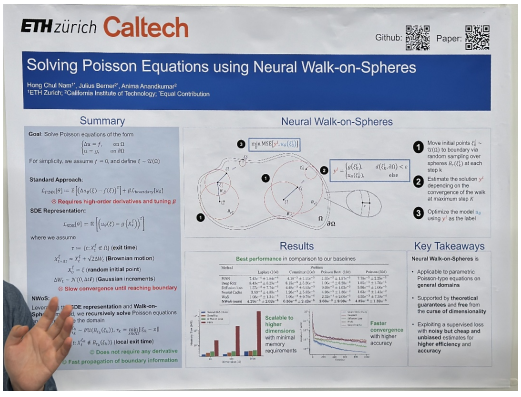
Continued on next page

Table 1: (Continued)

	**Missing-modality Enabled Multi-modal Fusion Architecture for Medical Data** *Muyu Wang,Shiyu Fan,Yichen Li,Hui Chen* POSITION: MISSION CRITICAL - SATELLITE DATA IS A DISTINCT MODALITY IN MACHINE LEARNING `http://arxiv.org/abs/2309.15529v1`
	**Human vs. Generative AI in Content Creation Competition: Symbiosis or Conflict?** *Fan Yao,Chuanhao Li,Denis Nekipelov,Hongning Wang,Haifeng Xu* Human vs. Generative AI in Content Creation Competition: `http://arxiv.org/abs/2402.15467v1`
	**Multimodal Crop Type Classification Fusing Multi-Spectral Satellite Time Series with Farmers Crop Rotations and Local Crop Distribution** *Valentin Barriere,Martin Claverie* QUANTiFpiNG LiKENESS `http://arxiv.org/abs/2208.10838v1`

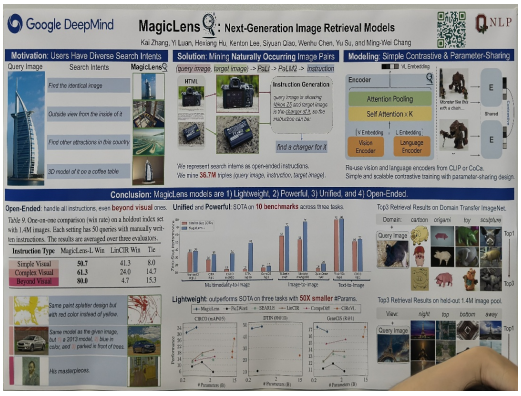
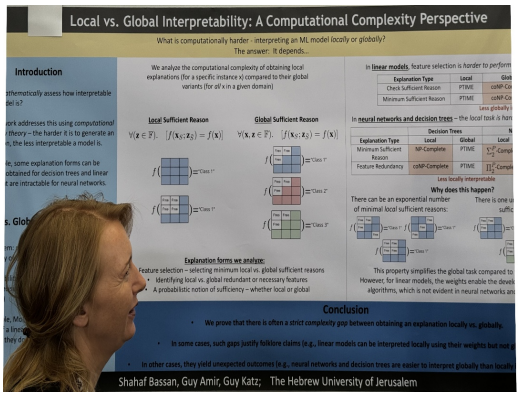
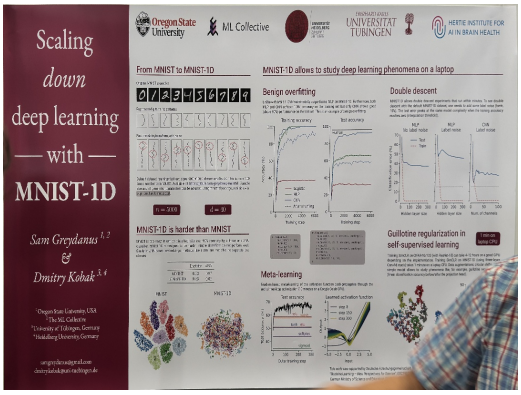
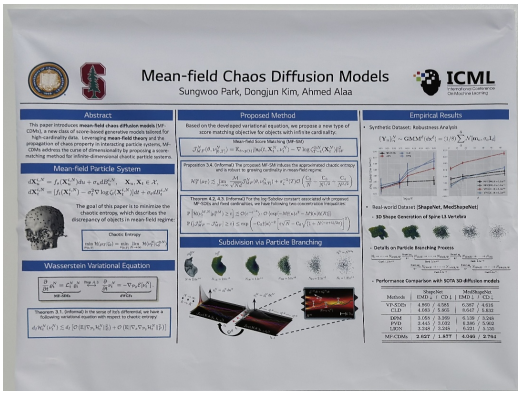
Continued on next page

Table 1: (Continued)

	**Privacy Backdoors: Stealing Data with Corrupted Pretrained Models** *Shanglun Feng,Florian Tramèr* Privacy Backdoors: Stealing Data with Corrupted Pretrained Models `http://arxiv.org/abs/2404.00473v1`
	**by Linear Regression:** *Objective: given some subsequence x, .., x,) predict the* by Linear Regression:
	**Promptbreeder: Self-Referential Self-Improvement Via Prompt Evolution** *Chrisantha Fernando,Dylan Banarse,Henryk Michalewski,Simon Osindero,Tim Rocktäschel* Promptbreeder: Self-Referential Self-Improvement via Prompt Evolution `http://arxiv.org/abs/2309.16797v1`
	**Integral Equation Approach to Stationary Stochastic Counting Process with Independent Increments** *Enzhi Li* Solving Poisson Equations using Neural Walk-on-Spheres `http://arxiv.org/abs/1811.07262v1`

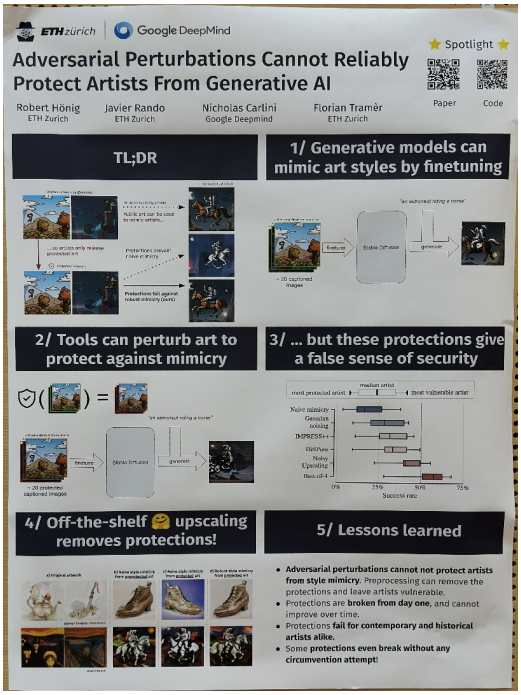
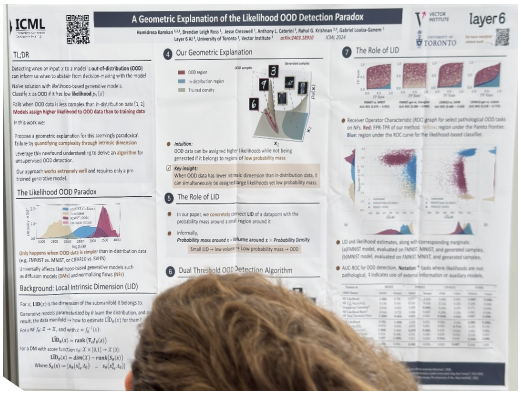
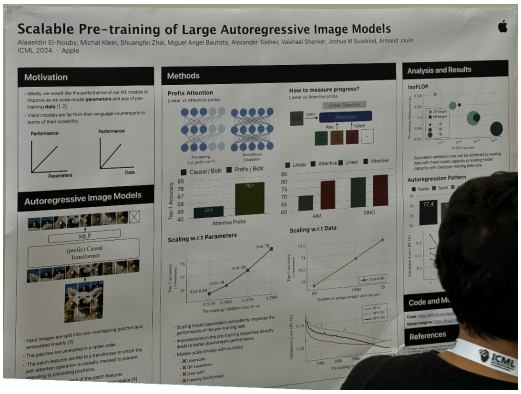
Continued on next page

Table 1: (Continued)

	**MagicLens: Self-Supervised Image Retrieval with Open-Ended Instructions** *Kai Zhang, Yi Luan, Hexiang Hu, Kenton Lee, Siyuan Qiao, Wenhu Chen, Yu Su, Ming-Wei Chang* MagicLens: Next-Generation Image Retrieval Models `http://arxiv.org/abs/2403.19651v2`
	**Local vs. Global Interpretability: A Computational Complexity Perspective** *Shahaf Bassan, Guy Amir, Guy Katz* Local vs. Global Interpretability: A Computational Complexity Perspective `http://arxiv.org/abs/2406.02981v2`
	**deep learning** *nn. Convid(25, 25, 3, stride=2, padding=1),* deep learning
	**Active matter beyond mean-field: Ring-kinetic theory for self-propelled particles** *Yen-Liang Chou, Thomas Ihle* Mean-field Chaos Diffusion Models `http://dx.doi.org/10.1103/PhysRevE.91.022103`

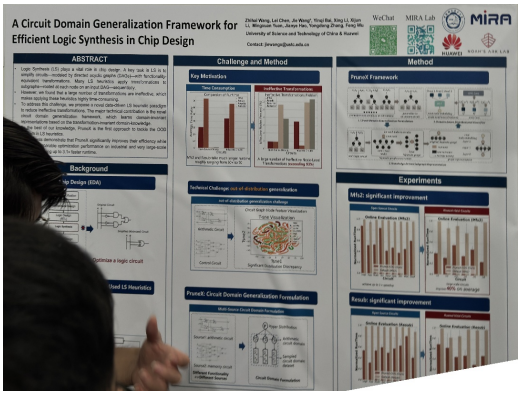
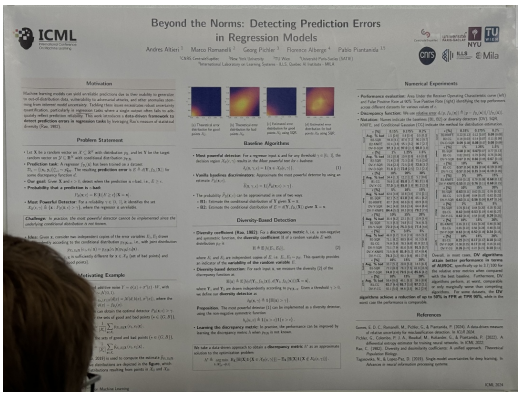
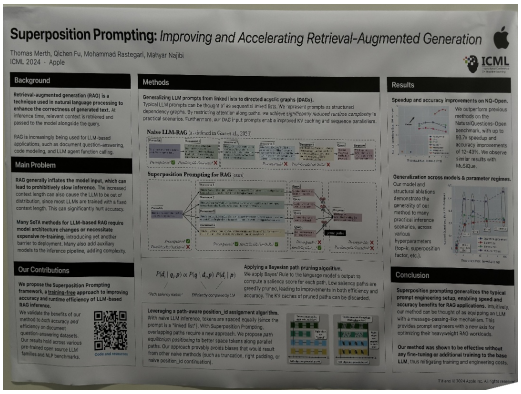
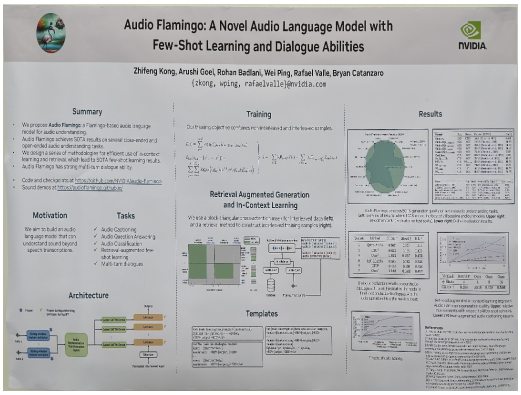
Continued on next page

Table 1: (Continued)

	**Adversarial Perturbations Cannot Reliably Protect Artists From Generative AI** *Robert Hönig,Javier Rando,Nicholas Carlini,Florian Tramèr* Adversarial Perturbations Cannot Reliably `http://arxiv.org/abs/2406.12027v1`
	**A Geometric Explanation of the Likelihood OOD Detection Paradox** *Hamidreza Kamkari,Brendan Leigh Ross,Jesse C. Cresswell,Anthony L. Caterini,Rahul G. Krishnan,Gabriel Loaiza-Ganem* A Geometric Explanation of the Likelihood OOD Detection Paradox `http://arxiv.org/abs/2403.18910v2`
	**Scalable Pre-training of Large Autoregressive Image Models** *Alaaeldin El-Nouby,Michal Klein,Shuangfei Zhai,Miguel Angel Bautista,Alexander Toshev,Vaishaal Shankar,Joshua M Susskind,Armand Joulin* Scalable Pre-training of Large Autoregressive Image Models `http://arxiv.org/abs/2401.08541v1`

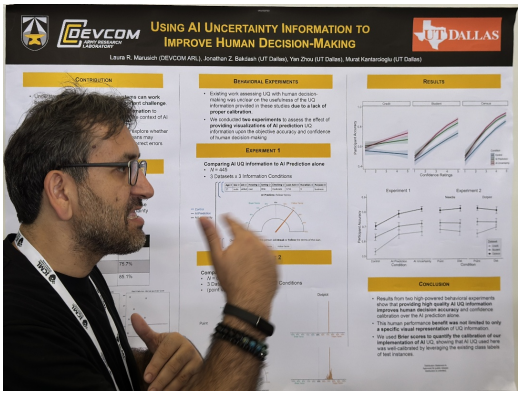
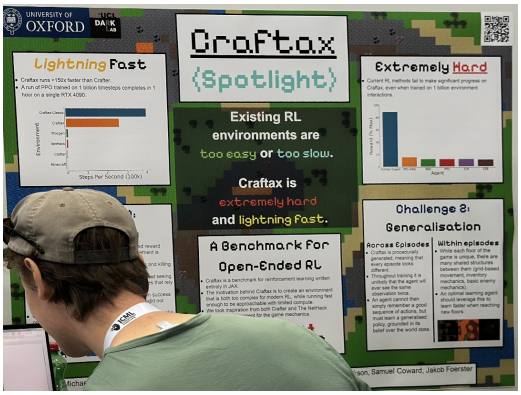
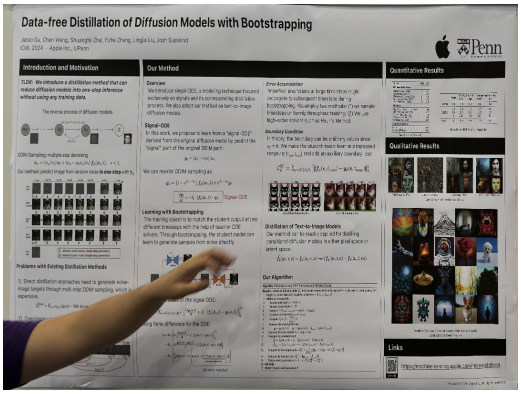
Continued on next page

Table 1: (Continued)

	**A Circuit Domain Generalization Framework for Efficient Logic Synthesis in Chip Design** *Zhihai Wang,Lei Chen,Jie Wang,Xing Li,Yinqi Bai,Xijun Li,Mingxuan Yuan,Jianye Hao,Yongdong Zhang,Feng Wu* Efficient Logic Synthesis in Chip Design `http://arxiv.org/abs/2309.03208v1`
	**Foliations on double-twisted products** *André Gomes* 2122202171071721 `http://arxiv.org/abs/1101.5730v1`
	**Superpositions of thermalisation states in relativistic quantum field theory** *Joshua Foo,Magdalena Zych* Superposition Prompting: Improving and Accelerating Retrieval-Augmented Generation `http://arxiv.org/abs/2307.02593v1`
	**Zero-shot audio captioning with audio-language model guidance and audio context keywords** *Leonard Salewski,Stefan Fauth,A. Sophia Koepke,Zeynep Akata* Audio Flamingo: A Novel Audio Language Model with `http://arxiv.org/abs/2311.08396v1`

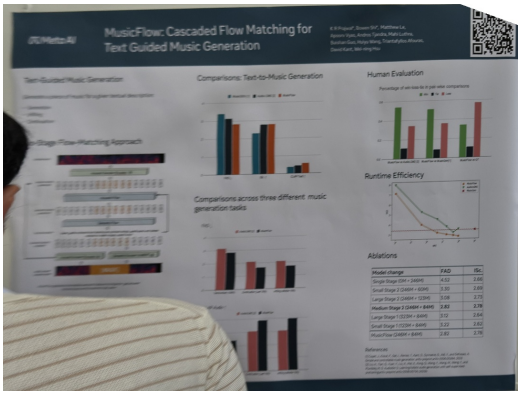
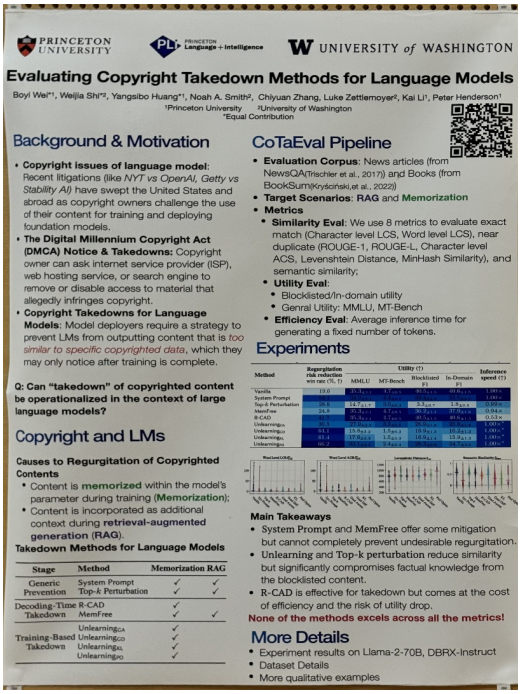
Continued on next page

Table 1: (Continued)

	**Using AI Uncertainty Quantification to Improve Human Decision-Making** *Laura R. Marusich,Jonathan Z. Bakdash,Yan Zhou,Murat Kantarcioglu* Laura R. Marusich (DEVCOM ARL), Jonathan Z. Bakdash (UT Dallas), Yan Zhou (UT Dallas), Murat Kantarcioglu (UT Dallas) `http://arxiv.org/abs/2309.10852v2`
	**About Geometry and Initial Phase of Cloud-to-Ground Lightning** *Aleš Berkopec* Crafto `http://arxiv.org/abs/1602.02496v1`
	**Data-free Distillation of Diffusion Models with Bootstrapping** *fo (xt, t, c) = fox, t, n) +w. (fo (xt, t, c) - foxt, t, n))* Data-free Distillation of Diffusion Models with Bootstrapping
	**Measuring the Earth's Synchrotron Emission from Radiation Belts with a Lunar Near Side Radio Array** *Alexander Hegedus,Quentin Nenon,Antoine Brunet,Justin Kasper,Angelica Sicard,Baptiste Cecconi,Robert MacDowall,Daniel Baker* $ Flow: Differentiable Simulations with 10x less code `http://dx.doi.org/10.1029/2019RS006891`

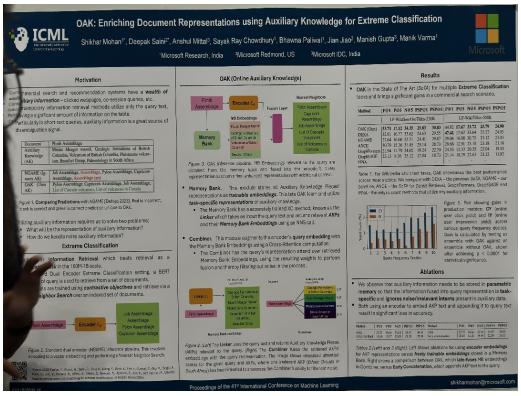
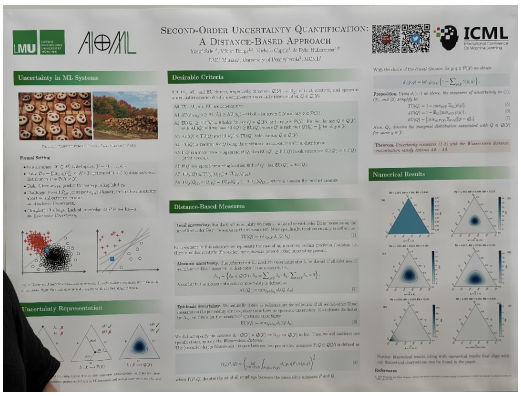
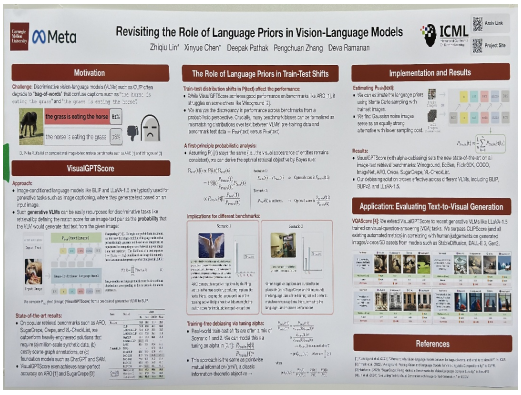
Continued on next page

Table 1: (Continued)

	**A double-layer Boussinesq-type model for highly nonlinear and dispersive waves** *Florent Chazel,Michel Benoit,Alexandre Ern,Serge Piperno* MusicFlow: Cascaded Flow Matching for `http://dx.doi.org/10.1098/rspa.2008.0508`
	**WebLINX** *WebLINX: Real-World [...]* WebLINX
	**Evaluating Copyright Takedown Methods for Language Models** *Boyi Wei,Weijia Shi,Yangsibo Huang,Noah A. Smith,Chiyuan Zhang,Luke Zettlemoyer,Kai Li,Peter Henderson* Evaluating Copyright Takedown Methods for Language Models `http://arxiv.org/abs/2406.18664v3`

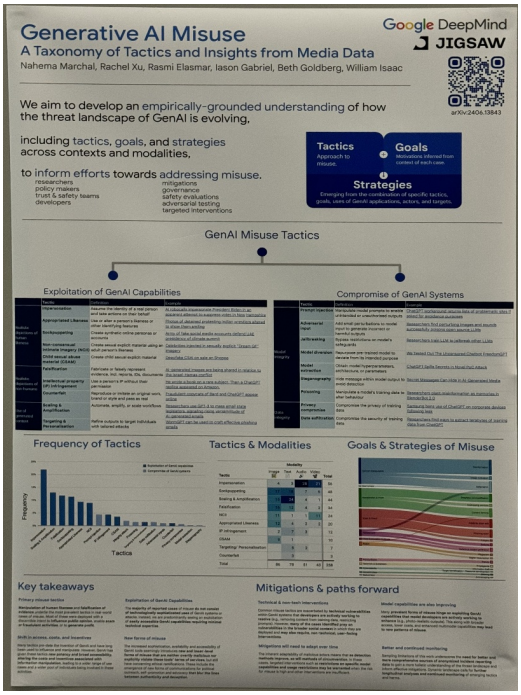
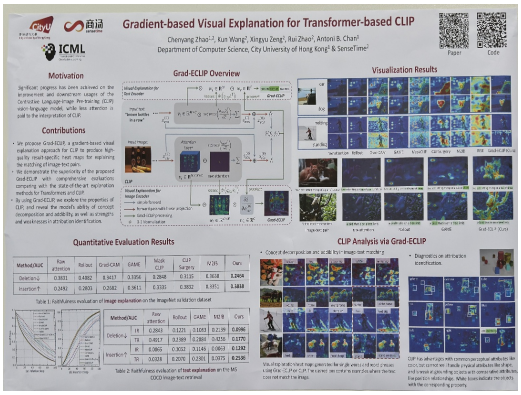
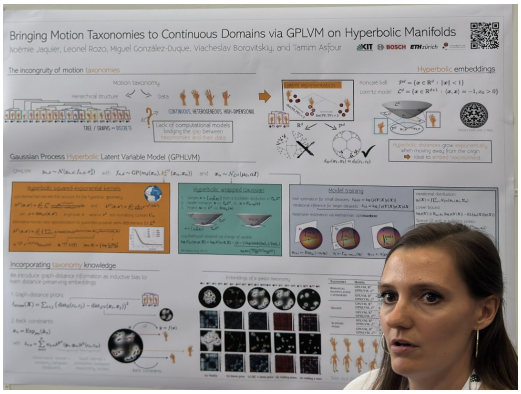
Continued on next page

Table 1: (Continued)

	**Auxiliary Knowledge-Induced Learning for Automatic Multi-Label Medical Document Classification** *Xindi Wang,Robert E. Mercer,Frank Rudzicz* OAK: Enriching Document Representations using Auxiliary Knowledge for Extreme Classification `http://arxiv.org/abs/2405.19084v1`
	**Power-Law distributions and Fisher's information measure** *F. Pennini,A. Plastino* SECOND-ORDER UNCERTAINTY QUANTIFICATION: `http://dx.doi.org/10.1016/j.physa.2003.10.076`
	**Revisiting the Role of Language Priors in Vision-Language Models** *[1] Yuksekgonul et al. (2023). "When and why vision-language models behave like bags-of-words, and what to do about it?", In: ICLR.* Revisiting the Role of Language Priors in Vision-Language Models

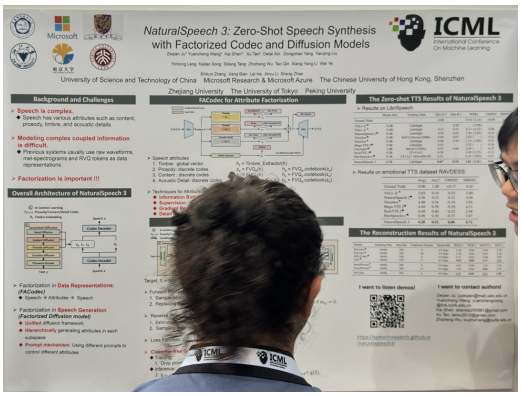
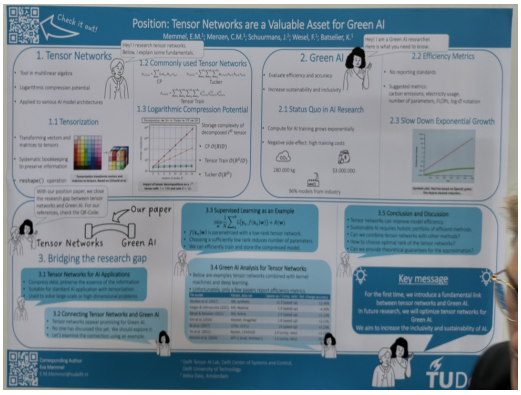
Continued on next page

Table 1: (Continued)

	**Generative AI Misuse: A Taxonomy of Tactics and Insights from Real-World Data** *Nahema Marchal, Rachel Xu, Rasmi Elasmar, Iason Gabriel, Beth Goldberg, William Isaac* A Taxonomy of Tactics and Insights from Media Data `http://arxiv.org/abs/2406.13843v2`
	**RCA: Region Conditioned Adaptation for Visual Abductive Reasoning** *Hao Zhang, Yeo Keat Ee, Basura Fernando* Gradient-based Visual Explanation for Transformer-based CLIP `http://arxiv.org/abs/2303.10428v4`
	**Bringing motion taxonomies to continuous domains via GPLVM on hyperbolic manifolds** *Noémie Jaquier, Leonel Rozo, Miguel González-Duque, Viacheslav Borovitskiy, Tamim Asfour* Bringing Motion Taxonomies to Continuous Domains via GPLVM on Hyperbolic Manifolds `http://arxiv.org/abs/2210.01672v4`

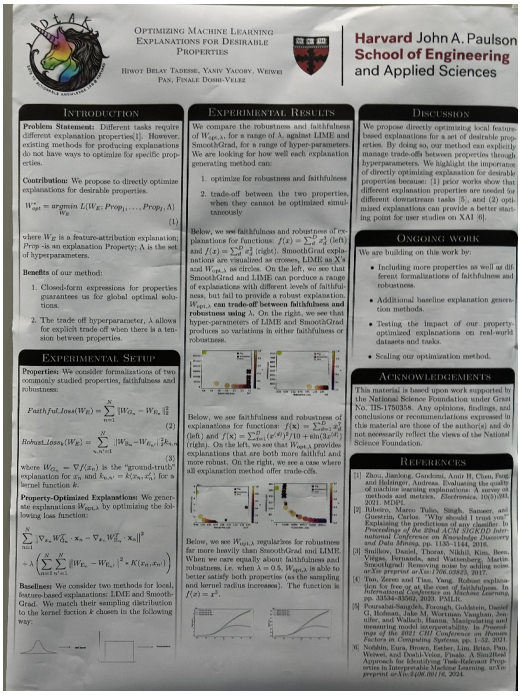
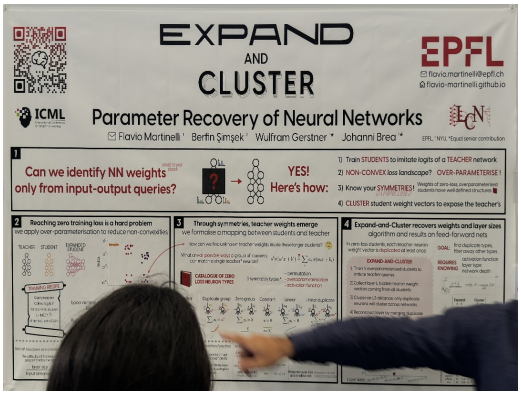
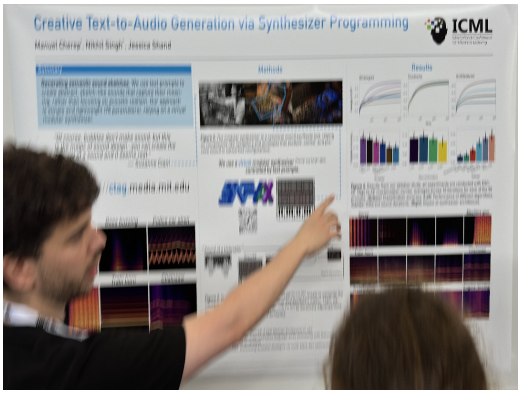
Continued on next page

Table 1: (Continued)

	**Better & Faster Large Language Models via Multi-token Prediction** *Fabian Gloeckle,Badr Youbi Idrissi,Baptiste Rozière,David Lopez-Paz,Gabriel Synnaeve* Better & Faster Large Language Models via Multi-token Prediction `http://arxiv.org/abs/2404.19737v1`
	**TJ-FlyingFish: Design and Implementation of an Aerial-Aquatic Quadrotor with Tiltable Propulsion Units** *Xuchen Liu,Minghao Dou,Dongyue Huang,Biao Wang,Jinqiang Cui,Qinyuan Ren,Lihua Dou,Zhi Gao,Jie Chen,Ben M. Chen* University of Science and Technology of China Microsoft Research & Microsoft Azure The Chinese University of Hong Kong, Shenzhen `http://arxiv.org/abs/2301.12344v2`
	**Position: Tensor Networks are a Valuable Asset for Green AI** *Eva Memmel,Clara Menzen,Jetze Schuurmans,Frederiek Wesel,Kim Batselier* Position: Tensor Networks are a Valuable Asset for Green Al `http://arxiv.org/abs/2205.12961v2`

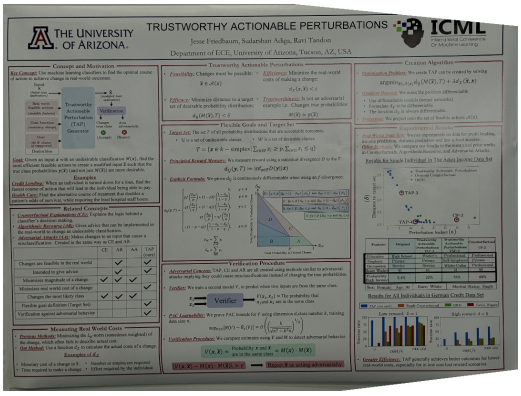
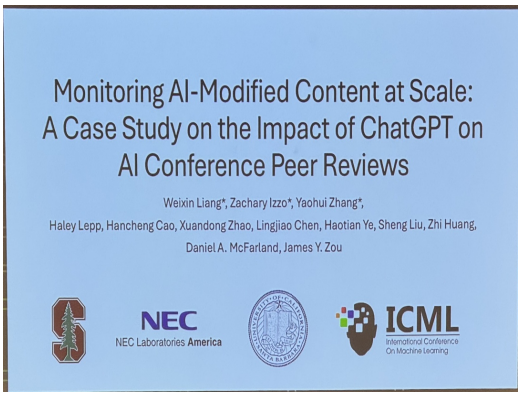
Continued on next page

Table 1: (Continued)

	**School of Engineering** *[1) Zhou, Jianlong, Gandomi, Amir H, Chen, Fang,* School of Engineering
	**Dual Convexified Convolutional Neural Networks** *Site Bai,Chuyang Ke,Jean Honorio* Parameter Recovery of Neural Networks `http://arxiv.org/abs/2205.14056v2`
	**Luban: Building Open-Ended Creative Agents via Autonomous Embodied Verification** *Yuxuan Guo,Shaohui Peng,Jiaming Guo,Di Huang,Xishan Zhang,Rui Zhang,Yifan Hao,Ling Li,Zikang Tian,Mingju Gao,Yutai Li,Yiming Gan,Shuai Liang,Zihao Zhang,Zidong Du,Qi Guo,Xing Hu,Yunji Chen* Creative Text-to-Audio Generation Via Synthesizer Programming `http://arxiv.org/abs/2405.15414v1`

Continued on next page

Table 1: (Continued)

	**A Survey on Trustworthy Edge Intelligence: From Security and Reliability To Transparency and Sustainability** *Xiaojie Wang,Beibei Wang,Yu Wu,Zhaolong Ning,Song Guo,Fei Richard Yu* TRUSTWORTHY ACTIONABLE PERTURBATIONS http://arxiv.org/abs/2310.17944v2
	**Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews** *Weixin Liang,Zachary Izzo,Yaohui Zhang,Haley Lepp,Hancheng Cao,Xuandong Zhao,Lingjiao Chen,Haotian Ye,Sheng Liu,Zhi Huang,Daniel A. McFarland,James Y. Zou* Monitoring AI-Modified Content at Scale: http://arxiv.org/abs/2403.07183v2