

EYLON CAPLAN

Email ◊ Website ◊ GitHub ◊ Google Scholar ◊ LinkedIn

4th-year PhD student in NLP | LLM Evaluation, Retrieval & Context Selection, and Data at Scale

Available for internships Spring–Summer–Fall 2027

EDUCATION

Purdue University , West Lafayette, Indiana, USA Ph.D. in Natural Language Processing, Department of Computer Science	Expected December 2027 GPA: 3.8/4.0
University of Nebraska–Lincoln , Lincoln, Nebraska, USA B.S. in Computer Science and Mathematics (Minors: Physics, Spanish)	May 2023 GPA: 3.99/4.0

RESEARCH EXPERIENCE

AWS AI Labs <i>Applied Scientist Intern</i>	Summer 2026 <i>Manager: Rashmi Gangadharaiah</i>
• Built BOUNDARYMORPH, a Gaussian Process selector that spends a limited reranking budget choosing which documents enter an LLM’s context for deep-research queries • State-of-the-art set retrieval (+5.4 nCG@100); beats a multi-hop LLM search agent at 34× lower latency; first-author, under review at ICLR 2027	
Purdue University <i>Graduate Research Assistant</i>	Aug 2023 – Present <i>Advisor: Dan Goldwasser</i>
• RL for open-ended generation (current): countering mode collapse in GRPO with real human text (500+ argument strategies mined from 338K argument corpus); early runs win 56.5% on held-out conversation generation • Data curation at scale (SPLITS!): 9.7M-post Reddit corpus split by demographic × topic; triages 23K LLM-generated hypotheses with 15–18× less human review (ACL 2026 Oral, SAC Highlight Award) • Evaluation (VIBE): co-designed a VLM benchmark isolating a new failure mode, misreading correctly seen social cues; humans beat the best VLM by 17 points (EMNLP 2025 Findings) • LLM-in-the-loop retrieval (CONCEPTCARVE): an LLM iteratively steers a retriever over 165M posts on a fixed budget; +26% over LLM query expansion (ACL 2025)	

SKILLS

Research: LLM evaluation, VLM evaluation, benchmarks, retrieval, RAG, reranking, agentic search, context selection, large-scale data curation, synthetic data, post-training, SFT, DPO, GRPO, Gaussian Processes, Bayesian optimization
Tools: Python, PyTorch, HuggingFace (Transformers, TRL, PEFT), vLLM, FAISS, ColBERT, pyserini, SLURM, Docker, Kubernetes, AWS, Dask, Weights & Biases, Hydra, Git

PUBLICATIONS

<i>BoundaryMORPH: Budgeted Reranking via Active Set Selection for Diffuse Retrieval</i> Eylon Caplan , Shamik Roy, Shib Sankar Dasgupta, Yingfan Wang, and Rashmi Gangadharaiah	Under review, ICLR
<i>GRID: Modeling Health Influencer Discourse as Argument Networks for Credibility Assessment</i> Nishanth Sridhar Nakshatri, Eylon Caplan , and Dan Goldwasser	Under review, NAACL
<i>Datasets and Methods for Improving the Cultural Capabilities of NLP Systems: A Survey</i> Tania Chakraborty*, Eylon Caplan *, Zhaoqing Wu*, Kevin Cushing, Bruce Qin, Shreya Havaladar, and Dan Goldwasser *Equal contribution	NLP+CSS 2026
<i>TAIGR: Towards Modeling Influencer Content on Social Media via Structured, Pragmatic Inference</i> Nishanth Sridhar Nakshatri, Eylon Caplan , Rajkumar Pujari, and Dan Goldwasser	ACL 2026, Oral
<i>SPLITS! Flexible Sociocultural Linguistic Investigation at Scale</i> Eylon Caplan , Tania Chakraborty, and Dan Goldwasser, SAC Highlight Award (top 0.5%)	ACL 2026, Oral
<i>VIBE: Can a VLM Read the Room?</i> Tania Chakraborty, Eylon Caplan , and Dan Goldwasser	EMNLP 2025 Findings
<i>CONCEPTCARVE: Dynamic Realization of Evidence</i> Eylon Caplan and Dan Goldwasser	ACL 2025