

Liam Dugan

ldugan@seas.upenn.edu | [linkedin.com/in/liam-dugan](https://www.linkedin.com/in/liam-dugan) | liamdugan.com

RESEARCH FOCUS

I specialize in AI safety research, focusing on human and automated detection of AI-generated content. My work explores the limitations of detection and strategies for deploying detectors with minimal harm. Additionally, I contribute to evaluating and applying Large Language Models (LLMs) to educational question generation, code prompting, machine translation, and multi-document question answering.

EDUCATION

UNIVERSITY OF PENNSYLVANIA	Philadelphia, PA
Ph.D, Computer Science (Advisor: Chris Callison-Burch)	<i>Aug. 2021 – Present</i>
M.S.E, Robotics	<i>Aug. 2017 – Dec. 2020</i>
B.S.E, Computer Engineering & East Asian Studies	<i>Aug. 2015 – Aug. 2020</i>

WORK EXPERIENCE

Roblox - PhD Research Intern	San Mateo, CA (Summer 2022)
John's Hopkins University - Visiting Research Scholar	Baltimore, MD (Summer 2021)
NVIDIA - Autonomous Driving Software Intern	Santa Clara, CA (Summer 2019)
Robotic Research LLC - Software Engineering Intern	Clarksburg, MD (Summer 2018)

PRESENTATIONS, POSTERS AND TALKS

Conference Talks

- “Exploring The Curious Case of Code Prompts”: NLRSE Workshop @ ACL 2023, Toronto ([video](#)) ([slides](#)) ([poster](#))
- “Real or Fake Text: Investigating Human Ability to Detect Boundaries between Human-Written and Machine Generated Text”: AAI 2023, Washington D.C. ([video](#)) ([slides](#)) ([poster](#)) ([code](#))

Invited Talks

- “Should we still use Text for Speech-to-Speech Translation? Promise meets Practice”: John's Hopkins University HLTCOE Seminar, Baltimore MD, May 2023 ([slides](#))
- “Real or Fake Text: Investigating Human Ability to Detect Boundaries between Human-Written and Machine Generated Text”: Brown University, Providence RI, March 2023 ([slides](#))
- “Detecting Generated Text from ChatGPT and other LLMs”: Penn Critical Writing Seminar, Philadelphia PA (Virtual), Feb. 2023 ([slides](#))
- “Intro to Machine Learning and AI Research”: St. Joe's Prep High School, Philadelphia PA, Feb. 2022 ([slides](#))

Poster Presentations

- “Kani: A Lightweight and Highly Hackable Framework for LM Applications” - EMNLP 2023 ([poster](#)) ([demo](#))
- “Learning When to Speak: Latency and Quality Trade-offs for Simultaneous Speech-to-Speech Translation with Offline Models” - Interspeech 2023 ([poster](#)) ([demo](#))
- “A Feasibility Study of Answer-Agnostic Question Generation for Education” - ACL 2022 ([video](#)) ([slides](#)) ([poster](#))
- “RoFT: A Tool for Evaluating Human Detection of Machine Generated Text” - EMNLP 2020 ([poster](#))

Other Presentations

- “Learning Formality from Japanese-English Parallel Corpora” Master's Thesis Defense ([video](#)) ([slides](#))
- “Scene++ VR” Penn Engineering School-Wide Senior Design Top 15 Finalist Demo ([video](#)) ([poster](#)) ([slides](#))
- “Cloud Chaser” PennApps XVII Hackathon Grand Prize Presentation ([video](#))
- “Todd: The Inter-Dimensional Robot” PennApps XVI Hackathon 3rd Place Presentation ([video](#))

TECHNICAL SKILLS

Natural Languages: English (native), Japanese (advanced - 7+ years [JLPT N2])

Programming Languages: Python, C, C++, bash, CUDA, Java, JavaScript, Go, HTML/CSS, Verilog, MATLAB

Frameworks: PyTorch, HuggingFace, Pandas, Numpy, Tensorflow, OpenCV, DXR, Vulkan

Developer Tools: Git, Slurm, QSub, Docker, Google Cloud Platform, VS Code, emacs, Atom, tmux

- 2024 Runsheng Huang, Yue Yang, **Liam Dugan**, and Chris Callison-Burch. MiRAGeNews: Multimodal Realistic AI-Generated News Detection. IC2S2 2024 (In Prep)
- Salvatore Giorgi, David M. Markowitz, Nikita Soni, Vasudha Varadarajan, Siddharth Mangalik, **Liam Dugan**, João Sedoc, Lyle H. Ungar, and H. Andrew Schwartz. Estimating Human Traits from AI-written Text. *Frontiers in Artificial Intelligence*, (In Prep)
- Alyssa Hwang, **Liam Dugan**, Andrew Head, and Chris Callison-Burch. Developing Grounded Intuition of Large Language Models. COLM 2024 (In Prep)
- Andrew Zhu, Alyssa Hwang, **Liam Dugan**, and Chris Callison-Burch. FanOutQA: Multi-Hop, Multi-Document Question Answering for Large Language Models. ACL 2024 (In Submission)
- Liam Dugan**, Alyssa Hwang, Filip Trhlik, Josh Magnus Ludan, Andrew Zhu, Hainiu Xu, Daphne Ippolito, and Chris Callison-Burch. RAID: A Shared Benchmark for Robust Evaluation of Machine-Generated Text Detectors. ACL 2024 (In Submission)
- 2023 Andrew Zhu*, **Liam Dugan***, Alyssa Hwang, and Chris Callison-Burch. Kani: A Lightweight and Highly Hackable Framework for Building Language Model Applications. In *Proceedings of the 3rd Workshop for Natural Language Processing Open Source Software (NLP-OSS 2023)*, pages 65–77, Singapore, Singapore, December 2023. Empirical Methods in Natural Language Processing
- Josh Ludan, Qing Lyu, Yue Yang, **Liam Dugan**, Mark Yatskar, and Chris Callison-Burch. Interpretable-by-Design Text Classification with Iteratively Generated Concept Bottleneck. ArXiv, October 2023. COLM 2024 (In Prep)
- Liam Dugan**, Anshul Wadhawan, Kyle Spence, Chris Callison-Burch, Morgan McGuire, and Victor Zordan. Learning When to Speak: Latency and Quality Trade-offs for Simultaneous Speech-to-Speech Translation with Offline Models. In *Proc. INTERSPEECH 2023*, pages 5265–5266, August 2023
- Hannah Gonzalez, **Liam Dugan**, Eleni Miltsakaki, Zhiqi Cui, Jiaxuan Ren, Bryan Li, Shriyash Upadhyay, Etan Ginsberg, and Chris Callison-Burch. Enhancing Human Summaries for Question-Answer Generation in Education. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*, pages 108–118, Toronto, Canada, July 2023. Association for Computational Linguistics
- Li Zhang*, **Liam Dugan***, Hainiu Xu*, and Chris Callison-Burch. Exploring the Curious Case of Code Prompts. In *Proceedings of the 1st Workshop on Natural Language Reasoning and Structured Explanations (NLRSE)*, pages 9–17, Toronto, Canada, June 2023. Association for Computational Linguistics **Selected for Oral Presentation**
- Liam Dugan***, Daphne Ippolito*, Arun Kirubarajan, Sherry Shi, Chris Callison-Burch. Real or Fake Text?: Investigating Human Ability to Detect Boundaries between Human-Written and Machine-Generated Text. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37(11), pages 12763–12771, Washington, D.C., Jun. 2023. **Selected for Oral Presentation**
- Aarohi Srivastava, Abhinav Rastogi, and (**440 others**). Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models. *Transactions on Machine Learning Research*, May 2023
- 2022 S. Upadhyay, E. Ginsberg, **Liam Dugan**, E. Miltsakaki, H. Gonzalez, D. Choi, C. Yuan, and C. Callison-Burch. Question Generation for Textbook Flashcards. In *EDULEARN22 Proceedings*, 14th International Conference on Education and New Learning Technologies, page 3412. IATED, July 2022
- Daphne Ippolito, **Liam Dugan**, Emily Reif, Ann Yuan, Andy Coenen, and Chris Callison-Burch. The Case for a Single Model that can Both Generate Continuations and Fill-in-the-Blank. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2421–2432, Seattle, United States, July 2022
- Liam Dugan**, Eleni Miltsakaki, Shriyash Upadhyay, Etan Ginsberg, Hannah Gonzalez, DaHyeon Choi, Chuning Yuan, and Chris Callison-Burch. A Feasibility Study of Answer-Agnostic Question Generation for Education. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 1919–1926, Dublin, Ireland, May 2022
- 2020 **Liam Dugan**. Learning Formality from Japanese-English Parallel Corpora. Master’s thesis, University of Pennsylvania, December 2020
- Liam Dugan***, Daphne Ippolito*, Arun Kirubarajan*, and Chris Callison-Burch. RoFT: A Tool for Evaluating Human Detection of Machine-Generated Text. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 189–196, Online, October 2020. Association for Computational Linguistics
- 2019 Zhengyi Luo, Austin Small, **Liam Dugan** and Stephen Lane. Cloud Chaser: Real Time Deep Learning Computer Vision on Low Computing Power Devices. *Eleventh International Conference on Machine Vision (ICMV 2018)*

NEWS ARTICLES

- (1/2/24) Can Humans Learn To Spot Fake Text? - Penn Engineering Magazine ([link](#))
- (9/20/23) Unlocking AI Potential: Unveiling Kani, the Groundbreaking Open-Source Framework Revolutionizing Large Language Model Applications - CJ&CO ([link](#))
- (9/19/23) Researchers from the University of Pennsylvania Introduce Kani: A Lightweight, Flexible, and Model-Agnostic Open-Source AI Framework for Building Language Model Applications - MarkTechPost ([link](#))
- (9/19/23) Kani: A Lightweight and Customizable Framework for Language Model Applications - TS2 ([link](#))
- (8/7/23) AI ‘Watermarking’ Tools Emerging to Tag Machine-Made Content - Bloomberg Law ([link](#))
- (7/27/23) CNN Features Penn Engineering AI Research - Penn Engineering Today ([link](#))
- (7/19/23) Academic Integrity and AI: Is Detection the Answer? - Temple Center for Teaching ([link](#))
- (7/11/23) Bot or not? How to tell when you’re reading something written by AI - CNN ([link](#))
- (5/18/23) NewsChannel12 Investigates: Artificial Intelligence Part III - ABC News North Carolina ([video](#))
- (4/26/23) Alien Minds, Immaculate Bullshit, Outstanding Questions - The Pennsylvania Gazette ([link](#))
- (4/18/23) How can people navigate AI-generated misinformation? - Canvas 8
- (4/11/23) Reddit Moderators Brace for a ChatGPT Spam Apocalypse - Vice ([link](#))
- (3/10/23) Real or fake text? We can learn to spot the difference - Penn Today ([link](#))
- (3/8/23) A Bot Isn’t Going to Take Your Place, But AI Will Make Your Job Harder - CCI ([link](#))
- (3/8/23) New Study Shows People Can Learn to Spot Machine-Generated Text - UniteAI ([link](#))
- (3/6/23) How can humans detect AI writing? These Penn researchers have some tips - Technically Philly ([link](#))
- (3/3/23) Can Humans Detect Text by AI Chatbot GPT? - Psychology Today ([link](#))
- (3/2/23) People can learn to detect AI writing - Cosmos Magazine ([link](#))
- (2/27/23) Real or Fake Text? We Can Learn to Spot the Difference - Penn Engineering Today ([link](#))
- (1/23/18) Object-Seeking Robot Wins PennApps XVII - Penn Engineering Today ([link](#))
- (9/10/17) At PennApps XVI, students made inter-dimensional robots and hung out with the founder of Quora - The Daily Pennsylvanian ([link](#))

TEACHING EXPERIENCE

Teaching Assistant

- CIS530 (Computational Linguistics) – Spring 2020, Fall 2020, Summer 2023
- CIS700 (Interactive Fiction & Text Generation) – Spring 2022
- CIS565 (GPU Programming & Architecture) – Fall 2021
- CIS380 (Operating Systems) – Fall 2018, Spring 2019, Fall 2019 (Head TA)
- CIS240 (Introduction to Computer Systems) – Spring 2017, Fall 2018, Spring 2018

Guest Lectures

- CIS565 (GPU Programming): “Optimizing Machine Learning with CUDA” – Fall 2021, Fall 2022, Fall 2023
- CIS565 (GPU Programming): “Introduction to Machine Learning” – Fall 2021, Fall 2022, Fall 2023
- CIS380 (Operating Systems): “Linux Page Replacement Algorithms and Belady’s Anomaly” – Fall 2019

Authored Homework Assignments

- CIS530 (Computational Linguistics): “HW7: Fine-Tuning Pre-Trained Language Models” – June 2023
- CIS530 (Computational Linguistics): “HW7: Transformers and State-of-the-Art Language Models” – Nov. 2020
- CIS530 (Computational Linguistics): “HW10: Neural Machine Translation” (with Li “Harry” Zhang) – Apr. 2020

FELLOWSHIPS, AWARDS AND HONORS

Fellowships and Grants

- (August 2022) Roblox Research Grant – \$100,000
- (October 2021) Google Cloud Platform Research Grant – \$5,500
- (October 2018) FLAS: Foreign Language and Area Studies Undergraduate Fellowship (East Asia) – \$15,000

Academic Honors & Awards

- (May 2020) Penn Engineering Exceptional Service Award
- (May 2020) Moore School Council Cwikla Award (Most Improved Student) [Nominated]
- (May 2019) Third Place: Penn Engineering Computer Science Senior Design

Hackathon Awards

- (February 2018) Most Innovative Use of Technology: Wharton Undergrad FinTech (WUFT) Hacks
- (January 2018) Grand Prize & Best use of Cloud Hosting: PennApps XVII (**1st out of 156 teams**)
- (September 2017) Third Prize: PennApps XVI (**3rd out of 158 teams**)

PROJECTS

- SCENE++ VR | *Oculus Rift, ZED Mini, Unity 3D, Python, YOLOv3, Paperspace* May 2019
- **Won 3rd Prize in Penn Computer Science Senior Design Competition**
 - We developed a Unity API that allows VR & AR Developers to query real-world objects around the user
 - Hardware: Oculus Rift headset with head-mounted ZED mini depth camera for pass-through Augmented Reality
 - We use Spatial Feature Mapping of environment to allow localization and stabilization of queried objects in depth
 - Offloading object detection to cloud server allows Scene++ to run on any platform with *virtually no drop in FPS*
- RTX EXPLORE | *C++, DirectX Raytracing, NVIDIA TitanV* December 2018
- **(50+ stars on GitHub)** Built the first open-source path tracer in the DirectX Raytracing GPU framework
 - Features include: Dynamic model loading from .gltf and .obj, support for texture and normal maps, live editing of scene transformations through GUI interface, specular/refractive/dispersive/transmissive materials, subsurface scattering, anti-aliasing, depth of field
- BANKING WITH A VISION | *Python, Javascript, TCP, Bootstrap* February 2018
- **Won Most Innovative use of Technology at WUFT Hacks**
 - Use facial key-point mapping algorithm to perform face recognition through front facing webcam
 - Faces are used to access database of customer information to save bank tellers having to pull up user information
- CLOUD CHASER | *Python, C, TCP, AWS, YOLOv3, Alexa, Raspberry Pi, 3D Printing* January 2018
- **Won Grand Prize and Best use of Cloud Hosting at PennApps XVII (1st out of 150 teams)**
 - Presented a platform that allows low resource IoT devices to do high level image processing on the cloud
 - Hardware: Raspberry Pi + camera, 3D printed robot chassis & camera mount, 4 servo motors, Amazon Echo Dot
 - Built robot “Chase” to demonstrate our platform. Commands are given to Chase through Echo Dot
 - Paper outlining our techniques to reduce latency of streaming for IoT image detection accepted to ICMV 2018
- TODD: THE INTER-DIMENSIONAL ROBOT | *C, HC-05 Bluetooth, Arduino, Unity 3D* September 2017
- **Won Third Prize at PennApps XVI (3rd out of 158 Teams)**
 - Made multiplayer game where player controlling Todd has to dodge objects only visible in virtual world
 - Hardware: Arduino, breadboard, 2 servo motors, Bluetooth HC-05 controller
 - USB connected Bluetooth controller allows Unity to communicate with “Todd” the robot

REFERENCES

Chris Callison-Burch, Associate Professor
University of Pennsylvania
ccb@cis.upenn.edu

Lyle Ungar, Professor
University of Pennsylvania
ungar@cis.upenn.edu

Daphne Ippolito, Assistant Professor
Carnegie Mellon University
daphnei@cmu.edu

Morgan McGuire, Chief Scientist
Roblox Research
morgan@roblox.com