TRISTAN THRUSH

tristant33@gmail.com
http://www.tristanthrush.com
https://github.com/TristanThrush
https://scholar.google.com/citations?user=qDDmq54AAAAJ

EDUCATION

Massachusetts Institute of Technology *2017 - 2019*
Master of Engineering in Computer Science with concentration in Artificial Intelligence.
Thesis: SAL: a Self-Aware Learning system.
Supervisor: Patrick Winston, Ford Professor of Artificial Intelligence and Computer Science.
GPA: 5.0 (5.0 scale)

Massachusetts Institute of Technology *2015 - 2019*
Bachelor of Science, Computer Science and Engineering. Minor in Mathematics. Minor in Linguistics.
GPA: 4.7 (5.0 scale)

Graduated high school a year early as valedictorian after taking 7 engineering and upper division math classes at UCSD; attended MIT a year early as a result.

RESEARCH EXPERIENCE

Facebook AI Research

- Research Associate *2020 - Present*
  - Supervisor: Douwe Kiela
  - Engineering lead for the Dynabench AI benchmarking project (published)
    - several governments (UK, EU) and companies (Facebook, Microsoft, Huawei, Tencent, Wikipedia, Amazon, ML Commons, etc.) impacted, 3000+ users, 50+ model evaluation datasets hosted, 5+ large datasets generated, 10+ research papers enabled, 500+ models uploaded, 1 large scale machine translation competition hosted
  - Joint lead for the Dynaboard holistic model evaluation system (published)
  - Lead for the Dynatask holistic training/evaluation data collection system (working paper)
  - Lead for the WinoGround vision+language model evaluation dataset (in review at CVPR)
  - Research on hate speech detection (published)
  - Research on question answering (published)
  - Research on natural language inference (published)
  - Research on visual question answering (published)
  - Engineering for the WMT 2021 machine translation competition (published)
  - Engineering for DataPerf core-set selection research

MIT Brain and Cognitive Sciences

- Research Associate, Computational Psycholinguistics Lab *2019 - 2020*
  - Supervisor: Roger Levy
  - Lead research on few-shot learning in foundation language models (published)
  - Lead research on cognitively inspired neural machine translation models (published)

MIT Computer Science and Artificial Intelligence Lab

- Graduate Researcher, Genesis Group *2018 - 2019*
  - Supervisor: Patrick Winston, Randall Davis
  - Cognitive AI and reinforcement learning research (thesis).
- Undergraduate Researcher, Genesis Group *2016 - 2018*
  - Supervisor: Patrick Winston
  - Cognitive AI and reinforcement learning research (published).
- Undergraduate Researcher, Robot Locomotion Group *Summer 2017*

- Supervisor: Russ Tedrake
  - Integrated Kaelbling's BHPN planner with the Drake robot simulation and control toolbox.
  - Integrated BHPN with a language production system to explain itself.
  - Contributed to the Drake software engineering effort

NASA/Caltech Jet Propulsion Lab

- Research Intern, Perception Systems Group *Summer 2018*
  - Supervisor: Renaud Detry
  - Developed a stereo visual odometry algorithm for sample tube localization (published)
  - Developed a dataset and data collection tools for sample tube localization on mars (published)
  - Research on neural dense segmentation and feature extraction modules (published)
  - Impacted the group's approach to autonomously recovering the first sample tubes from Mars
- Research Extern, Perception Systems Group *Winter 2017*
  - Supervisor: Renaud Detry
  - Combined inverse kinematics, motion planning, and vision systems for a robot arm.
  - System was used as part of a project that won the best paper award at IROS.

PAPERS

Preprints

1. Winoground: Probing vision and language models for visio-linguistic compositionality.
   Tristan Thrush*, Candace Ross*, Ryan Jiang, Max Bartolo, Amanpreet Singh, Adina Williams, and Douwe Kiela.
   2021.

2. Dynatask: a generic framework to specify AI tasks for dynamic data collection and model evaluation.
   Tristan Thrush, Anmol Gupta, Kushal Tirumala, and Douwe Kiela.
   2021.

3. Hatemoji: A Test Suite and Adversarially-Generated Dataset for Benchmarking and Detecting Emoji-based Hate.
   Hannah Rose Kirk, Bertram Vidgen, Paul Röttger, **Tristan Thrush**, and Scott A Hale.
   arXiv, 2021.

4. ANLIzing the Adversarial Natural Language Inference Dataset.
   Adina Williams, **Tristan Thrush**, and Douwe Kiela.
   arXiv, 2020

Peer-Reviewed Publications

1. Human-Adversarial Visual Question Answering.
   Sasha Sheng, Amanpreet Singh, Vedanuj Goswami, Jose Alberto Lopez Magana, **Tristan Thrush**, Wojciech Galuba, Devi Parikh, and Douwe Kiela.
   NeurIPS, 2021.

2. Dynaboard: An Evaluation-As-A-Service Platform for Holistic Next-Generation Benchmarking.
   Zhiyi Ma*, Kawin Ethayarajh*, **Tristan Thrush***, Somya Jain, Ledell Wu, Robin Jia, Christopher Potts, Adina Williams, and Douwe Kiela.
   NeurIPS, 2021.

3. Improving Question Answering Model Robustness with Synthetic Adversarial Data Generation.
   Max Bartolo, **Tristan Thrush**, Robin Jia, Sebastian Riedel, Pontus Stenetorp, and Douwe Kiela.
   EMNLP, 2021.

4. Dynabench: Rethinking benchmarking in NLP.
   Douwe Kiela, Max Bartolo, Yixin Nie, Divyansh Kaushik, Atticus Geiger, Zhengxuan Wu, Bertie Vidgen, Grusha Prasad, Amanpreet Singh, Pratik Ringshia, Zhiyi Ma, **Tristan Thrush**, Sebastian Riedel, Zeerak Waseem, Pontus Stenetorp, Robin Jia, Mohit Bansal, Christopher Potts, and Adina Williams.
   NAACL, 2021.

5. Rover Relocalization for Mars Sample Return by Virtual Template Synthesis and Matching.
   Tu-Hoa Pham, William Seto, Shreyansh Daftry, Barry Ridge, Johanna Hansen, **Tristan Thrush**, Mark Van der Merwe, Gerard Maggiolino, Alexander Brinkman, John Mayo, Yang Cheng, Curtis Padgett, Eric Kulczycki, and Renaud Detry.
   IEEE Robotics and Automation Letters, 2021.

6. Learning from the worst: Dynamically generated datasets to improve online hate detection.
   Bertie Vidgen, **Tristan Thrush**, Zeerak Waseem, Douwe Kiela.
   ACL, 2021.

7. Investigating Novel Verb Learning in BERT: Selectional Preference Classes and Alternation-Based Syntactic Generalization.
   Tristan Thrush, Ethan Wilcox, and Roger Levy.
   BlackboxNLP at EMNLP, 2020.

8. Compositional neural machine translation by removing the lexicon from syntax (Abstract).
   Tristan Thrush.
   CogSci, 2020.

9. The partial mental state inducer: Learning intuition with few training examples and k-line theory.
   Tristan Thrush, and Patrick Winston.
   Advances in Cognitive Systems, 2018

10. A neural model for learning a humanlike vowel feature space.
    Tristan Thrush.
    Northeastern Computational Phonology, 2018.

Workshop Committee Papers

1. Findings of the WMT 2021 Shared Task on Large-Scale Multilingual Machine Translation.
   Guillaume Wenzek, Vishrav Chaudhary, Angela Fan, Sahir Gomez, Naman Goyal, Somya Jain, Douwe Kiela, **Tristan Thrush**, and Francisco Guzmán.
   WMT at EMNLP, 2021.

2. The First Workshop on Dynamic Adversarial Data Collection (DADC)
   Max Bartolo, Hannah Rose Kirk, **Tristan Thrush**, Katerina Margatina, Pedro Rodriguez, Mimansa Jaiswal, Pontus Stenetorp, Robin Jia, and Douwe Kiela.
   DADC at ACL, 2022.

Theses

1. SAL: a Self-Aware Learning system (Master's thesis).
   Tristan Thrush.
   MIT Computer Science and Artificial Intelligence Laboratory, 2019

Reports

1. Convolutions inspired by the human retina enable learning of more robust features.
   Tristan Thrush.
   MIT Department of Brain and Cognitive Sciences, 2018.

2. Machine learning approaches to capture the reliability of news articles
   Rares Buhai*, and **Tristan Thrush***.
   MIT Department of Electrical Engineering and Computer Science, 2017

3. A self-aware and hypothetical question-answering BHPN with drake control.
   Tristan Thrush.
   MIT Computer Science and Artificial Intelligence Laboratory, 2017.

4. Probabilistic lattice learning and backward chaining.
   Tristan Thrush.
   MIT Computer Science and Artificial Intelligence Laboratory, 2016.

HONORS

Funding from grants, MIT Computational Psycholinguistics Lab	*2019 - 2020*
MEng thesis proposal was selected for RA funding	*2018 - 2019*
Author of three research proposals accepted for MIT lab sponsored funding	*2017 - 2018*
Certificate in Advanced Undergrad Research in AI and Machine Learning, MIT	*2017 - 2018*
MIT EECS Undergraduate Research and Innovation Scholar	*2017 - 2018*
Author of research proposal accepted for MIT institute funding	*2017*

TALKS

NeurIPS, Virtual	*2021*
CogSci, Virtual	*2020*
Northeastern Computational Phonology, MIT	*2019*
Advances in Cognitive Systems, Stanford	*2018*

SERVICE

Workshops Organized

Dynamic Adversarial Data Collection (DADC) at ACL	*2022*
Conference on Machine Translation (WMT) at EMNLP	*2021*

Reviewing

Data-Centric AI at NeurIPS	*2021*
Conference on Machine Translation (WMT) at EMNLP	*2021*
Workshop on Online Abuse and Harms (WOAH) at ACL	*2021*

TEACHING

MIT Brain and Cognitive Sciences

- Lead NLP tutorial for undergrads, Computational Psycholinguistics Group *Spring 2020*
- Teaching Assistant, Computational Cognitive Science Class *Fall 2019*
  - Supervisor: Josh Tenenbaum
  - Supervised grad students' NLP research projects, focusing on grammar induction.
  - Received highest possible ratings from student evaluations at the end of the course.
  - Other normal TA responsibilities: helped students at office hours, recorded lectures, etc.

MENTORING

Major League Hacking fellows working on Facebook AI projects

- Ishita Dasgupta *2021*
- Anand Rajaram *2021*

- Wong Kok Rui *2021*
- Fatima Zahra Chriha *2021*

Facebook AI SWE interns

- Anmol Gupta (partial supervision) *2021*

SELECTED PRESS AND MEDIA

Dynatask [2]: Analytics Drift, Facebook AI Blog, Talk at Toronto Machine Learning Summit

Dynaboard [2]: VentureBeat, Tech Times, Facebook AI Blog

Dynabench [4]: MIT Tech Review, Wired, Market Tech Post, NLP Highlights, Facebook AI Blog

LANGUAGES AND FRAMEWORKS

Huggingface Transformers. Amazon Mechanical Turk experiment design. Mephisto. Python. Cython. C++. C. Java. JavaScript. Lua. R. MATLAB. React js. WebPPL. Assembly. SQL. YAML. Batch Files. Shell Scripts. XML-ish (HTML, URDF, SDF). Android App development. AWS. Production Model Deployment. Cluster Computing and Machine Learning. ROS. Drake. LCM. PyTorch. Torch. Keras. TensorFlow. OpenCV. OpenAI Gym. Pandas. scikit-learn. NumPy. SciPy. GNU/UNIX. Android. LIBVISO2. Linguistic Research Databases (e.g. CHILDES, Treebanks, VerbNet).