NAOMI SAPHRA

Gradient Descent Spectator

@ nsaphra@nsaphra.net • www.nsaphra.github.io/ • @nsaphra • github.com/nsaphra

WORK EXPERIENCE

NLP Consultant
MosaicML

📅 2022 - present • 📍 New York, New York

Postdoctoral Researcher
NYU

📅 2021 - present • 📍 New York, New York

- Supervisor: Kyunghyun Cho.

Research Intern
Google

📅 Winter 2020 • 📍 New York, New York

- Supervisors: Dipanjan Das (Language, Google AI), Ian Tenney, Jasmijn Bastings, Thibault Sellam.
- Better understanding of how BERT learns.

Contractor
Koko

📅 Summer 2017 • 📍 New York, New York

- Developed neural classifiers for informal text at anti-abuse startup.

Software Engineering Intern
Google

📅 Summer 2013 & 2015 • 📍 Mountain View, California

- Supervisor: Marius Pasca (Common Sense, Machine Intelligence).
- Unsupervised discovery of paraphrases of compositional noun phrases.
- Supervisor: Eric Altendorf (Common Sense, Machine Intelligence).
- Integrated models of real-world knowledge into entity linking.

Undergraduate Research Fellow
CLSP Summer Workshop at Johns Hopkins University

📅 Summer 2012 • 📍 Baltimore, Maryland

- Supervisor: Matthew Blaschko.
- Unsupervised methods to verify image annotations.

Engineering Intern
Facebook

📅 Summer 2011 • 📍 Palo Alto, California

- Supervisor: Jen Burge (Engagement Team).
- Engineered new features for the People You May Know model.

EDUCATION

Ph.D. in Informatics
Edinburgh University

📅 2021

Supervisor: Adam Lopez.
Thesis: Training Dynamics of Neural Language Models.

M.Sc. in Computer Science
Johns Hopkins University

📅 2015

B.Sc. in Computer Science
Carnegie Mellon University

📅 2013

Minor in Language Technologies

Sabbatical
The Recurse Center

📅 2017

Unstructured programming retreat.

AWARDS

- ICLR outstanding reviewer 2021, 2022.
- New York Academy of Sciences - Natural Language, Dialogue and Speech Symposium: Best Poster Runner Up.
- Daniella Sciama Award for Achievement through Adversity 2019.
- Google Europe Scholarship for Students with Disabilities 2017.
- Carnegie Mellon School of Computer Science Dragon Award 2013.
- National Merit Semifinalist 2009.
- Nannina Rasulo Memorial Scholarship Award for Technology 2009.

SKILLS & HOBBIES

Python   C++   spacy   pytorch

NLP   pretraining   interpretability

Roller Derby   Standup   Clown

SELECTED PUBLICATIONS

👥 Conference Proceedings

- Sellam, Thibault et al. (2022). "The MultiBERTs: BERT Reproductions for Robustness Analysis". In: *ICLR*.
- White, Jennifer C. et al. (2021). "A Non-Linear Structural Probe". In: *NAACL*.
- Pimentel, Tiago et al. (2020). "Pareto Probing: Trading Off Accuracy for Complexity". In: *EMNLP*.
- Saphra, Naomi and Adam Lopez (2020). "LSTMs Compose—and Learn—Bottom-Up". In: *Findings of EMNLP*.
- Saphra, Naomi and Lopez, Adam (2019). "Understanding Learning Dynamics Of Language Models with SVCCA". in: *NAACL*.
- Saphra, Naomi and Adam Lopez (2015). "AMRICA: an AMR Inspector for Cross-language Alignments". In: *NAACL*.
- Vedaldi, Andrea et al. (2014). "Understanding Objects in Detail with Fine-grained Attributes." In: *CVPR*.

👥 Workshop Proceedings

- Saphra, Naomi and Adam Lopez (2019). "Sparsity Emerges Naturally in Neural Language Models". In: *ICML Workshop on Identifying and Understanding Deep Learning Phenomena*.
- Saphra, Naomi and Lopez, Adam (2016). "Evaluating Informal-Domain Word Representations With UrbanDictionary". In: *ACL Workshop on Evaluating Vector Space Representations for NLP*.
- Cotterell, Ryan et al. (2014). "An Algerian Arabic-French Code-Switched Corpus". In: *LREC Workshop on Free/Open-Source Arabic Corpora*.
- Schneider, Nathan et al. (2014). "A framework for (under) specifying dependency syntax without overloading annotators". In: *ACL Linguistic Annotation Workshop*.

SERVICE

Organizing Committee
5th Workshop on analyzing and interpreting neural networks for NLP
📅 2022

Area Chair
EMNLP
📅 2021,2021

Organizing Committee
6th Workshop on Representation Learning for NLP
📅 2021

ACL Accessibility Subcommittee
Association for Computational Linguistics
📅 2020　　　　　　　📍 Seattle, WA

Communications committee
Auld Reekie Roller Derby
📅 2017 – 2019　　　📍 Edinburgh, UK

Puzzle tester and local organizer
North American Computational Linguistics Olympiad
📅 2012 – 2014

TEACHING

Project Supervisor
NYU
📅 2021

Machine Learning MS Capstone

Tutor
University of Edinburgh
📅 2016 – 2019

Informatics Research Review
Probabilistic Modeling & Reasoning
Machine Learning & Pattern Recognition

Course Assistant
Johns Hopkins University
📅 2013

Natural Language Processing

Teacher's Assistant
Irvington High School
📅 2008 – 2009

Ancient Greek

INVITED TALKS

Sources of Variance in Pretraining and Finetuning (Keynote)
3rd Neural Scaling Laws Workshop
📅 June 2022　　　　📍 Quebec

Learning Dynamics of LMs
UC Berkeley
📅 November 2020　　📍 Berkeley, CA

Transparent, Hackable, Accessible
QueerInAI EMNLP Social
📅 November 2020　　📍 Remote (EMNLP)

Learning Dynamics of LSTMs
NLP with Friends
📅 October 21, 2020　📍 Remote

Accessible Means Hackable (Keynote)
Pydatafest Amsterdam
📅 June 15, 2020　　📍 Amsterdam, Netherlands

Blackbox NLP Panel Discussion
ACL Blackbox NLP Workshop
📅 Aug 1, 2019　　　📍 Florence, Italy

Paying the Panopticon (standup)
The Stand Edinburgh Comedy Club
📅 Feb 26, 2019　　　📍 Edinburgh, UK