DR. MARGARET MITCHELL

Computer Scientist & Researcher

margarmitchell@gmail.com | Seattle, WA, U.S.A. | m-mitchell.com
https://scholar.google.com/citations?user=5na92fcAAAAJ | margaret-mitchell-9b13429/

About Me

Focused on foundational machine learning research and AI ethics operationalization, bringing together machine learning, social science, law and policy. My work is interdisciplinary, and tends to include computer vision, natural language processing, social media, statistical methods, and insights from cognitive science. I also like to work on applied assistive and augmentative technology. I wear many hats when I work, adapting my role to the needs of each project.

Research Areas

Artificial/Machine Intelligence, Ethical AI, Responsible AI, ML Fairness, Personal identity in ML, Vision & Language, Assistive/Clinical Technology, Natural Language Generation, Natural Language Processing, Referring Expression Generation/Object Reference, Cognitive Modeling, Information Extraction, Syntax

Some Notable Things

First woman (??) to win a deep learning competition, image captioning at CVPR 2015.

Research Lead for Microsoft's Seeing AI product, which won the prestigious Helen Keller Achievement Award from the American Foundation for the Blind, 2018.

Research Lead for Google Cloud Model Cards product, which won US Secretary of Defense Ash Carter's Tech Spotlight award, 2020.

Work

Staff Research Scientist, Google Brain, Ethical AI — 2018-2021
Manager: Samy Bengio

Grew team to 12+ people with my new co-lead Dr. Timnit Gebru. Worked on defining and operationalizing AI practices aligned to human values under the broad umbrella of "ethics" across Google. Work included developing methods for accountability, transparency, design processes, dataset construction and usage, and building tools to measure bias issues. Worked closely with Google Cloud on auditing practices and ethical launch protocols. Published more theoretical and experimental aspects of our work externally.

Senior Research Scientist, Google Cerebra — 2016-2017
Manager: Blaise Agüera y Arcas

Worked on defining and operationalizing AI practices aligned to human values, focusing on fairness. Developed "ML Fairness" organization and created an "Ethical AI" team.

Researcher, Microsoft Research, Cognition Group — 2014-2016
Manager: Pushmeet Kohli

Focused on advancing AI to maximally benefit people. Worked on image description, storytelling, and visual descriptions for the visually impaired. Won 1st and 3rd place in the first deep learning image captioning competition at CVPR. Developed Seeing AI, leveraging our work in image captioning and Microsoft's cloud platform.

Researcher, Microsoft Research, NLP Group 2013-2014
Manager: Bill Dolan

Worked on natural language processing for Bing & Cortana, focusing on microsummarization. Advanced conversation generation, image description. Led clinical NLP research.

Postdoctoral Researcher, Johns Hopkins University 2012-2013
Supervisor: Benjamin Van Durme

Worked on graphical models for semantic role labeling, named entity recognition, and sentiment detection. With Dr. Matt Gormley, developed the Pacaya graphical modelling toolkit.

Johns Hopkins Workshop Graduate, Center for Language and Speech Processing 2011
Supervisors: Tamara Berg, Alex Berg

Built data-driven language generation system, "Midge", that takes vision output as input, producing syntactically/semantically well-formed descriptions of images. Worked on automatic classification of visually descriptive text, and characterizing description.

Visiting Scholar, Center for Spoken Language Understanding 2009-2012
Supervisors: Brian Roark, Richard Sproat

Developed system to generate personal language for teenagers with cerebral palsy. Completed work on system that helps predict Mild Cognitive Impairment.

Research Assistant/Associate, Center for Spoken Language Understanding 2005-2007
Supervisors: Brian Roark, John-Paul Hosom, Jan van Santen

In charge of all transcription work for projects on Alzheimer's, Aphasia, Apraxia, Autism, Dysarthria, Parkinson's. Developed system to help automatically diagnose Mild Cognitive Impairment by extracting phonemic and syntactic features from narratives.

Recent Public Launches

Know Your Data 2020-2021

Interactive tool to explore datasets with the goal of improving data quality and identifying bias issues. Conceived the idea of & led the math behind the tool's "Relations" module, which show how closely associated different characteristics are.

Diversity and Inclusion Explorable 2020-2021

Building from my work on formal methods for measuring diversity and inclusion, this interactive tool focused on explaining diversity metrics methods and visually demonstrating tradeoffs between different diversity goals.

Model Card Toolkit 2019-2020

Tensorflow library that streamlines and automates generation of Model Cards.

Data Cards Playbook 2019-2021

Playbook to help interdisciplinary teams adopt a people-centered approach to transparency in dataset documentation for responsible AI systems.

Fairness Indicators 2017-2019

A suite of tools to enable simple computation and visualization of fairness metrics for binary and multi-class classification.

Google Cloud Model Cards 2017-2019

Transparent documentation on how well machine learned models work. Focused in particular on the performance of face detection models, with evaluation results disaggregated by face size, facial orientation, perceived gender presentation, age, and skin tone.

Machine learning "Crash Course" on Fairness. 2019

Teaching module looking at different types of human biases that can manifest in training data, and providing strategies to identify them and evaluate their effects.

Responsible AI Practices 2017-2021

Created and shared recommended practices for responsible AI. Continued to take latest findings into account, working to incorporate them as appropriate, and adapting as we learned more over time.

Operationalization of fairness and ethics across Google - Basic Steps
Operationalization of fairness and ethics across Google - My story 2016-2021

Ensured that a robust framework for ethical AI principles are in place, and that Google's products do not amplify or propagate unfair bias, stereotyping, or prejudice.

Education

University of Aberdeen, PhD: Computing Science 2009-2012
Thesis: Generating Reference to Visible Objects
Advisors: Kees van Deemter, Ehud Reiter

Examined the factors affecting reference in visual settings, including the production of color and size words, visual perception, and stored object representations. Created an algorithm that generates human-like reference to visible real world objects.

University of Washington, MA/MS: Computational Linguistics 2007-2008
Thesis: Towards the Generation of Natural Reference
Advisors: Scott Farrar, Emily Bender

Introduced an algorithm to generate referring expressions comparable to human-produced expressions, and developed a class-based system for the prenominal ordering of modifiers.

Reed College, BA: Linguistics, Allied Field: Psychology 2001-2005
Senior Thesis: On the Generation of Referring Expressions
Advisors: John Haviland, Matt Pearson

Critiqued the problem of generating human-like reference from a knowledge base, examining linguistic, psychological, and computational factors.

Favorite Publications

For full list of publications, please see my Google Scholar profile:
https://scholar.google.com/citations?hl=en&user=5na92fcAAAAJ&view_op=list_works&sortby=pubdate
h-index: 33; i10-index: 52

2021

Aka, O. and Burke, K. and Bäuerle, A. and Greer, Christina, and Mitchell, M. Measuring Model Biases in the Absence of Ground Truth. *Arxiv.*

Gebru, T. and Bender, E. and McMillan-Major, A. and Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜 *Proceedings of FAccT 2021.*

Ben Hutchinson, Andrew Smart, Alex Hanna, Emily Denton, Christina Greer, Oddur Kjartansson, Parker Barnes, Margaret Mitchell. Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure. *Proceedings of FAccT 2021*.

2020 Margaret Mitchell, Dylan Baker, Nyalleng Moorosi, Emily Denton, Ben Hutchinson, Alex Hanna, Timnit Gebru, Jamie Morgenstern. Diversity and inclusion metrics in subset selection. *Proceedings of AIES 2020*.

Inioluwa Deborah Raji, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, Emily Denton. Saving face: Investigating the ethical concerns of facial recognition auditing. *Proceedings of FAccT 2020*.

Inioluwa Deborah Raji, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, Parker Barnes. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *Proceedings of FAccT 2020*.

2019 Vinodkumar Prabhakaran, Ben Hutchinson, Margaret Mitchell. Perturbation sensitivity analysis to detect unintended model biases. *Proceedings of EMNLP 2019*.

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru. Model Cards for Model Reporting. *Proceedings of FAT* 2019*.

Ben Hutchinson, Margaret Mitchell. 50 years of test (un) fairness: Lessons for machine learning. *Proceedings of FAT* 2019*.

2018 Brian Hu Zhang, Blake Lemoine, Margaret Mitchell. Mitigating unwanted biases with adversarial learning. *Proceedings of AIES 2018*.

2017 Adrian Benton, Margaret Mitchell, Dirk Hovy. Multi-task learning for mental health using social media text. *Proceedings of EACL 2017*.

Margaret Mitchell and Alexander Todorov Blaise Agüera y Arcas. Physiognomy's New Clothes. *Medium*.

Hee Jung Ryu, Margaret Mitchell, Hartwig Adam. Improving smiling detection with race and gender diversity. *Proceedings of FAT/ML 2017*.

2016 Mostafazadeh, N. and Misra, I. and Devlin, J. and Zitnick, C. L. and Mitchell, M. and He, X. and Vanderwende, L. (2016). Generating Natural Questions About an Image. *Proceedings of ACL 2016*.

Huang, T.-H., and Ferraro, F., and Mostafazadeh, N. and Misra, I. and Agrawal, A. and Devlin, J. and Girshick, R. and He, X. and Kohli, P. and Batra, D. and Zitnick, C. L. and Parikh, Devi and Vanderwende, L. and Galley, M. and Mitchell, M. (2016). Visual Storytelling. *Proceedings of NAACL 2016*.

Misra, I. and Zitnick, C. L. and Mitchell, M. and Girshick, R. (2016). Seeing through the Human Reporting Bias: Visual Classifiers from Noisy Human-Centric Labels. *Proceedings of CVPR 2016*.

Mason, R. and Gaska, B. and Van Durme, B. and Choudhury, P. and Hart, T. and Dolan, B. and Toutanova, K. and Mitchell, M. (2016). Microsummarization of Online Reviews: An Experimental Study. *Proceedings of AAAI 2016*.

2015 Antol, S. and Agrawal, A. and Lu, J. and Mitchell, M. and Batra, D. and Zitnick, C. L. and Parikh, D. (2015). VQA: Visual Question Answering. *Proceedings of ICCV 2015*.

Devlin, J. and Cheng, H. and Fang, H. and Gupta, S. and Deng, L. and He, X. and Zweig, G. and Mitchell, M. (2015). Language Models for Image Captioning: The Quirks and What Works. *Proceedings of ACL 2015*.

Mitchell, M. and Hollingshead, K. and Coppersmith G. (2015). Quantifying the Language of Schizophrenia in Social Media. *Proceedings of the 2nd CLPsych Workshop, NAACL 2015*.

Fang, H. and Gupta, S. and Iandola, F. and Srivastava, Rupesh K. and Deng, L. and Dollar, P. and Gao, J. and He, X. and Mitchell, M. and Platt, J. C. and Zitnick, L. and Zweig, G. (2015). From Captions to Visual Concepts and Back. *Proceedings of CVPR 2015*.

2014 Gormley, M., and Mitchell, M., and Van Durme, B., and Dredze, M. (2014). Low Resource Semantic Role Labeling. *Proceedings of ACL 2014*.

Beller, C., and Knowles, R., and Harman, C., and Bergsma, S., and Mitchell, M., and Van Durme, B. (2014). I'm a Belieber: Social Roles via Self-identification and Conceptual Attributes. *Proceedings of ACL 2014*.

2013 Mitchell, M., and Aguilar, J., and Wilson, T., and Van Durme, B. (2013). Open Domain Targeted Sentiment. *Proceedings of EMNLP 2013*.

Mitchell, M., and van Deemter, K., and Reiter, E. (2013). Attributes in Visual Reference. *Proceedings of PRE-CogSci 2013*.

Mitchell, M. and Reiter, E., and van Deemter, K. (2013). Typicality and Object Reference. *Proceedings of CogSci 2013*.

Mitchell, M. and van Deemter, K., and Reiter, E. (2013). Generating Expressions that Refer to Visible Objects. *Proceedings of NAACL 2013*.

2012 Mitchell, M., Dodge, J., Goyal, A., Yamaguchi, K., Stratos, K., Han, X., Mensch, A., Berg, A., and Berg, T. L., Daumé III, H. (2012). Midge: Generating Image Descriptions From Computer Vision Detections. *Proceedings of EACL 2012*.

2011 Mitchell, M., Dunlop, A., and Roark, B. (2011). Semi-Supervised Modeling for Prenominal Modifier Ordering. *Proceedings of ACL 2011*.

Mitchell, M., van Deemter, K., and Reiter, E. (2011). On the Use of Size Modifiers When Referring to Visible Objects. *Proceedings of CogSci 2011*.

Roark, B., Mitchell, M., Hosom, J., Hollingshead, K., and Kaye, J. (2011). Spoken Language Derived Measures for Detecting Mild Cognitive Impairment. *IEEE Transactions on Audio, Speech, and Language Processing*.

2010

Mitchell, M., van Deemter, K., and Reiter, E. (2010). Natural Reference to Objects in a Visual Domain. *Proceedings of INLG 2010*.

Some Favorite (Public) Projects

Founder & Lead, Model Cards 2017-2021

Worked cross-Alphabet to operationalize "Model Cards", transparent documentation that accompanies machine learning models. Multiple launches across Google Cloud, Google Perception, and Jigsaw, as well as launches from OpenAI, Nvidia, Allen AI, Salesforce, Hugging Face.

Google Representative, Partnership on AI 2018-2020

Worked within the Fairness, Transparency, and Accountability group to establish norms on these issues across AI-focused organizations.

Co-founder & Research Lead, ABOUT ML in the Partnership on AI 2019-2021

Building from my work on Model Cards, created collaboration with Microsoft and IBM on transparency documentation for machine learning models.

Technical & Research Lead, Seeing AI 2015-2016

Leveraged my image captioning work to produce visual descriptions for the visually impaired. Streamed technology through cloud services to land on user's mobile device or smartglasses. Connected groups throughout Microsoft to move forward as one: MSR Labs, Cognitive Services, Windows, Narrator, Bing, Garage, Outreach, NExT.
Won the prestigious Helen Keller Achievement Award from the American Foundation for the Blind in 2018.

Microsoft Strategic Intern Project Lead, Visual Storytelling 2015-2016

Led group of 8 researchers, 2 visiting professors, and 5 students on summer project to drive forward research on selecting key frames in a photo album and generating a coherent story for them. Project resulted in full working system, with all 4 full-time interns first-authoring a paper at a top-tier conference.
Has been the basis of multiple Storytelling workshops and competitions.

Senior Researcher, JSALT Mental Health Summer Workshop 2016-2017

Exploring how we can utilize shared patient data alongside shared patient social media feeds to monitor PTSD and depression.
Some of the first work on fairness in health settings.

Top Dishes Researcher, Cortana and Bing Local Recommendations. 2014-2015

Developed end-to-end system to mine social media for positive/negative sentiment expressed towards most common items.
Incorporated into Microsoft products.

Academic Service (Subset)

FAccT Program Chair 2020-2021 and FAccT/FAT* Steering Committee 2018-2021

Co-Founder, Women and Underrepresented Minorities in NLP (WiNLP) group and annual workshop

Co-Founder, Annual Ethics in NLP workshop in ACL

Co-Founder, Program & Publication Chair, Annual Storytelling workshop in ACL

Co-Founder, Computational Linguistics and Clinical Psychology Workshop (CLPsych) in ACL. Program & Publication Chair for ACL 2014 and NAACL 2015.

First-of-its-kind gathering of medical practitioners and computational linguists. Became an annual event, still ongoing.

Area Chair, Generation ACL 2016

Publication Chair & Area Chair, Generation, NAACL 2016

Publication chair contributions still used in official style guidelines.

Area Chair, Generation & Summarization, EMNLP 2015

General, Program & Publication Chair, the International Natural Language Generation Conference (INLG) 2014

Knowledge Base Population - Sentiment Slot Filling Track Owner, Text Analysis Conference 2014

For the first time, defined and organized task for the National Institute of Standards and Technology (NIST) Text Analysis Conference (TAC) to predict sentiment between knowledge base entities.

Regular Reviewer for:

- Various workshops on fairness, vision-to-language, clinical NLP work
- NeurIPS
- Association for the Advancement of Artificial Intelligence (AAAI)
- Empirical Methods in Natural Language Processing (EMNLP)
- The Annual Meeting of the Association for Computational Linguistics (ACL)
- The North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)
- The International Natural Language Generation Conference (INLG)

Systems & Languages

Program in Python, Bash, C#, Java. Experience with Tcl, C++, HTML, JavaScript, PHP, sed, awk, etc.

Comfortable working with Windows, MacOS 9 and X, Linux (Ubuntu, Fedora), all MS Office software, most Google Enterprise software, all OpenOffice software, LaTeX, Praat, Wavesurfer, version control systems, Tensorflow

Personal Details

Citizenship: U.S.
Residence: Seattle, Washington
Date of birth: 18th November, 1983
Hobbies: Gardening, exercising, political analysis. Occasionally all at once.
Nerdy obsession with collecting vinyl records: Yes