

# Peter Anderson

Austin, TX | [peteanderson80@gmail.com](mailto:peteanderson80@gmail.com) | [www.panderson.me](http://www.panderson.me)

## EDUCATION

---

### Australian National University

*PhD in Computer Science (Machine Learning)*

Thesis: *Vision & Language Learning: From Image Captioning & Visual Question Answering towards Embodied Agents*

Canberra, Australia

Feb 2015 – Apr 2018

### University of New South Wales

*Bachelor of Computer Engineering, 1st class Honors and University Medal*

Sydney, Australia

Jul 2009 – Dec 2012

### University of Sydney

*Bachelor of Commerce (Finance and Economics), 1st class Honors and University Medal*

Sydney, Australia

Mar 2000 – Dec 2004

## EXPERIENCE

---

### Senior Research Scientist

*Google Research*

Jan 2020 – Aug 2023

*Austin, TX*

- Image Generation: Collaborated with a cross-functional team to develop and evaluate [Imagen Editor](#) [1], a state-of-the-art diffusion model for language-guided image editing launched on Google Cloud. Guided development of an LLM to automatically optimize users' text prompts for Imagen, launched as dogfood to 170k Googlers, and created new datasets for RL from human feedback (RLHF).
- Embodied AI: Tech lead devising and executing a research agenda to develop grounded LLM-based agents that can generate and follow detailed natural language navigation instructions in buildings. Created [new multilingual datasets](#) [6], new evaluations, and new generative models to synthesize [3D buildings](#) [4] and near-human-quality navigation instructions [3] as synthetic training data, improving the state-of-the-art by 8% [2] in a mature task.
- Mentored and trained 4 early-career researchers to promotion or full-time conversion. Assigned 4 direct reports in recognition of my mentoring and leadership skills.

### Research Scientist

*Georgia Institute of Technology*

Jul 2018 – Jan 2020

*Atlanta, GA*

- Developed a robot to follow natural language navigation instructions in a previously unseen building [5], identifying challenges in the sim-to-real transfer of LLM agents to robots.
- Mentored and advised PhD candidates, resulting in 10 publications, multiple new datasets (e.g. [7]), and advances in the state-of-the-art for language-guided navigation, image captioning, and visual question answering.

### Research Fellow

*Macquarie University*

Apr 2018 – Jul 2018

*Sydney, Australia*

- Developed a method to train image caption generators on images labeled only with noisy keywords.

### Research Intern

*Microsoft Research*

Apr 2017 – Jun 2017

*Redmond, WA*

- Proposed a new attention mechanism to combine vision and language representations [10], leading to state-of-the-art performance in multiple tasks and widespread adoption of the technique in academia and industry.

### Robotic Software Engineer

*Sabre Autonomous Solutions*

Feb 2014 – Feb 2015

*Sydney, Australia*

- Overhauled the planning algorithms of an autonomous grit-blasting robot, reducing planning time by 75% and missed surfaces by 80%. Added functionality to precisely align CAD models with depth scans.

### Founder

*FrameFish*

Jan 2013 – Feb 2014

*Sydney, Australia*

- Created and commercialized a SaaS virtual try-on plugin for online glasses and sunglasses retailers.

### Associate

*Equities Research, Credit Suisse*

Jan 2005 – Mar 2009

*Sydney, Australia*

- Second lead in a team conducting sell-side analysis across the Australian banks sector, securing a top-3 ranking.

### Intern

*Macquarie Airports, Macquarie Group*

Dec 2002 – Feb 2003

*Sydney, Australia*

### Intern/Analyst

*Equities Research, Goldman Sachs*

Dec 2001 – Oct 2002

*Sydney, Australia*

## AWARDS & RECOGNITION

---

Outstanding Reviewer Award, EMNLP	2020
Outstanding Reviewer Award, CVPR   <i>Among the top 0.9% of reviewers</i>	2019
Outstanding Reviewer Award, NeurIPS   <i>Among the top 3.6% of reviewers</i>	2018
Facebook ParlAI Research Award   <i>1 of 7 awards</i>	2017
Winner, Visual Question Answering Challenge, CVPR   <i>1st out of 24 teams</i>	2017
ANU Finalist, Three Minute Thesis (3MT)   <i>Among the top 12 PhD thesis talks at ANU</i>	2017
Australian Postgraduate Award (APA), ANU   <i>PhD Scholarship</i>	2015 – 2018
Research Supplementary Scholarship, ANU   <i>PhD Scholarship</i>	2015 – 2018
Runner Up, CiSRA Extreme Imaging Competition   <i>For computer vision research on soccer robots</i>	2013
3rd, RoboCup Soccer SPL using the Nao humanoid robot   <i>3rd out of 25 teams</i>	2012
Best Paper Finalist, RoboCup International Symposium	2012
1st class Honors & University Medal in Computer Eng, UNSW   <i>Top student in Computer Engineering</i>	2012
1st class Honors & University Medal in Finance, USYD   <i>Top student with a major in Finance</i>	2004
Joye Prize for 1st in Finance Honors, USYD	2004
Finance Honors Scholarship, USYD   <i>Honors Year Scholarship</i>	2004

## SELECTED PUBLICATIONS

---

Full publication list on [Google Scholar](#)

- [1] Su Wang\*, Chitwan Saharia\*, Ceslee Montgomery\*, Jordi Pont-Tuset, Shai Noy, Stefano Pellegrini, Yasumasa Onoe, Sarah Laszlo, David J. Fleet, Radu Soricut, Jason Baldridge, Mohammad Norouzi<sup>†</sup>, **Peter Anderson**<sup>†</sup>, William Chan<sup>†</sup>. *Imagen Editor and EditBench: Advancing and Evaluating Text-Guided Image Inpainting*, in CVPR 2023 (**Highlight Presentation; Top 2.5%**)
- [2] Aishwarya Kamath\*, **Peter Anderson**\*, Su Wang, Jing Yu Koh, Alexander Ku, Austin Waters, Yinfei Yang, Jason Baldridge, Zarana Parekh. *A New Path: Scaling Vision-and-Language Navigation with Synthetic Instructions and Imitation Learning*, in CVPR 2023
- [3] Su Wang, Ceslee Montgomery, Jordi Orbay, Vighnesh Birodkar, Aleksandra Faust, Izzeddin Gur, Natasha Jaques, Austin Waters, Jason Baldridge, **Peter Anderson**. *Less is More: Generating Grounded Navigation Instructions from Landmarks*, in CVPR 2022
- [4] Jing Yu Koh, Honglak Lee, Yinfei Yang, Jason Baldridge, **Peter Anderson**. *Pathdreamer: A World Model for Indoor Navigation*, in ICCV 2021
- [5] **Peter Anderson**, Ayush Shrivastava, Joanne Truong, Arjun Majumdar, Devi Parikh, Dhruv Batra, Stefan Lee. *Sim-to-Real Transfer for Vision-and-Language Navigation*, in Conference on Robot Learning (CoRL), 2020
- [6] Alex Ku\*, **Peter Anderson**\*, Roma Patel, Eugene Ie, Jason Baldridge. *Room-Across-Room: Multilingual Vision-and-Language Navigation with Dense Spatiotemporal Grounding*, in EMNLP 2020.
- [7] Harsh Agrawal\*, Karan Desai\*, Yufei Wang, Xinlei Chen, Rishabh Jain, Mark Johnson, Dhruv Batra, Devi Parikh, Stefan Lee and **Peter Anderson**. *nocaps: novel object captioning at scale*, in ICCV 2019
- [8] Mark Johnson, **Peter Anderson**, Mark Dras and Mark Steedman. *Predicting accuracy on large datasets from smaller pilot data*, in ACL 2018 (**Oral Presentation**)
- [9] **Peter Anderson**, Qi Wu, Damien Teney, Jake Bruce, Mark Johnson, Niko Sünderhauf, Ian Reid, Stephen Gould and Anton van den Hengel. *Vision-and-Language Navigation: Interpreting visually-grounded navigation instructions in real environments*, in CVPR 2018 (**Spotlight Presentation; Top 8.9%**)
- [10] **Peter Anderson**, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould and Lei Zhang. *Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering*, in CVPR 2018 (**Oral Presentation; Top 2.1%**)
- [11] **Peter Anderson**, Basura Fernando, Mark Johnson and Stephen Gould. *SPICE: Semantic Propositional Image Caption Evaluation*, in ECCV 2016
- [12] **Peter Anderson**, Yongki Yusmanthia, Bernhard Hengst, and Arcot Sowmya. *Robot Localisation using Natural Landmarks*, in Proceedings of the RoboCup International Symposium, 2012 (**Oral Presentation, Best Paper Finalist: Top 4.7%**)

\*<sup>†</sup>denotes equal contribution

## SELECTED INVITED TALKS

---

<b>Training Instruction-Following Agents with Synthetic Instructions</b>	2022
<ul style="list-style-type: none"><li>• ECCV Workshop on Language for 3D Scenes</li><li>• ECCV Cross-modal Human-robot Interaction Workshop</li></ul>	
<b>Synthetic Data for Language-Guided Agents</b>	2022
<ul style="list-style-type: none"><li>• AI2 Embodied AI Lecture Series</li></ul>	
<b>Massive Datasets for Language-Guided Navigation Agents and Where to Find Them</b>	2021
<ul style="list-style-type: none"><li>• CVPR Embodied AI Workshop</li><li>• ICCV Workshop on Human Interaction for Robot Navigation</li></ul>	
<b>People don't live in images</b>	2020
<ul style="list-style-type: none"><li>• VizWiz Grand Challenge Workshop at CVPR</li></ul>	
<b>Vision-and-Language Navigation</b>	2020
<ul style="list-style-type: none"><li>• UT Austin Forum for Artificial Intelligence</li></ul>	
<b>The Increasingly Real Problem of Grounded Language Learning</b>	2019
<ul style="list-style-type: none"><li>• Facebook AI Research (FAIR), Menlo Park, CA</li><li>• Google Research, Mountain View, CA</li></ul>	
<b>Both Sides Now: Generating and Understanding Visually-Grounded Language</b>	2019
<ul style="list-style-type: none"><li>• Allen Institute for Artificial Intelligence (AI2), Seattle, WA</li><li>• Microsoft Research, Redmond, WA</li></ul>	
<b>Vision and language: attention, navigation, and making it work 'in the wild'</b>	2018
<ul style="list-style-type: none"><li>• Vision and Language Seminar, ACVT, University of Adelaide, Australia</li><li>• Visual Question Answering Challenge Workshop at CVPR</li></ul>	

## TEACHING & ADVISING

---

**Invited Guest Lectures:** Brown, UPenn, UT Austin, Georgia Tech, UIUC, UVA

**PhD Thesis Committees:** [Meera Hahn](#), [Samyak Datta](#), [Shurjo Banerjee](#), [Harsh Agarwal](#)

## PROFESSIONAL SERVICE

---

**Area Chair:** ICML 2023, NeurIPS 2019 & 2022, NAACL 2022

**Reviewer:** CoRL, EMNLP, ACL, ICCV, ICRA, ICML, ICLR, NeurIPS, CVPR, ECCV, NAACL, EACL

**Workshop Organizer:**

- RxR-Habitat Challenge at the Embodied AI Workshop | CVPR 2021-22
- Embodied Vision, Actions & Language (EVAL) Workshop | ECCV 2020
- Advances in Language and Vision Research Workshop | ACL 2020
- Visual Question Answering and Dialog Workshop | CVPR 2019
- Habitat Embodied Agents Workshop | CVPR 2019
- Visually-Grounded Interaction and Language (ViGIL) | NeurIPS 2018
- Visual Learning and Embodied Agents in Simulation Environments | ECCV 2018
- New Benchmarks, Metrics, and Competitions for Robotic Learning | RSS 2018

**Tutorial Presenter:**

- From VQA to VLN: Recent Advances in Vision-and-Language | CVPR 2021
- Connecting Language and Vision to Actions | ACL 2018

## TECHNICAL SKILLS

---

**Research:** NLP, Computer Vision, Vision and Language, Robotics, Text-to-Image Generation

**Languages:** Python, C/C++, Java, JavaScript, HTML/CSS

**Frameworks:** PyTorch, JAX, TensorFlow