

# Carlos Gemmell

[carlos.gemmell@glasgow.ac.uk](mailto:carlos.gemmell@glasgow.ac.uk) | [carlos-gemmell.github.io](https://github.com/carlos-gemmell) | Twitter: [@carlos\\_gemmell](https://twitter.com/carlos_gemmell) | [Google Scholar](#)

TL;DR

---

I am an NLP PhD at the University of Glasgow. I led the winning Amazon 2021-2022 Alexa Prize team.

## EDUCATION

---

### University of Glasgow

Scotland, UK

*PhD in Machine Learning advised by [Jeff Dalton](#), EPSRC full scholarship September 2019 – Expected 2023*

- Fundamental research in NLP exploring how language models (LMs) can use tools like external knowledge and programs. Having models generate code as part of their reasoning in order to make them more interpretable and capable of symbolic manipulation.
- Worked on semantic parsing, conversational query understanding, and extrapolation capabilities of large LMs.

### University of Glasgow

*Bachelor in Computer Science*

*September 2015 – June 2019*

*Magna Cum Laude, Honours of the First Class*

Thesis: [Deep recurrent networks for language conditioned code generation](#)

## RESEARCH & WORK EXPERIENCE

---

### Team Lead - Glasgow Alexa Prize Team (1st Place) [\[link\]](#)

June 2021 – July 2022

*Advised by Jeff Dalton*

- Put together and led a team of 5 AI researchers to build and deploy at scale GRILLBot, a multi-modal task-oriented virtual assistant for the 2021-2022 Amazon Alexa Prize Taskbot competition.
- Won 1st place and 500,000\$ prize from an initial pool of 125 global university teams. As we progressed through 3 selection stages we were also awarded 250,000\$ throughout the competition.
- I designed and built the GPU Kubernetes cluster for deployment and scaling on AWS. GRILLBot was the only system built agnostic of Amazon's Alexa Prize infrastructure and has 99.9% up-time. It is currently deployed on all US Alexa devices. Try it! Say "Alexa, assist me".
- As team lead I created a fun and productive environment centred around agile software engineering practices like kanban using Jira, peer review, sprints, stand-ups, and compassion.
- As a team member, after jointly building the core of the assistant, I designed and trained all neural models (T5, GPT-J, ...) for dialogue management and contextual question answering. To measure their effectiveness, new train/dev/test collections were created leveraging data synthesis and prompting with GPT-3.

### Research Scientist - Internship

June. 2018 – Dec. 2018

*Bird.i*

- Developed a real-time object detection pipeline for construction sites satellite in images.
- Surveyed and reproduced results for several state of the art object detection models: Faster-RCNN, YOLO

### Software Engineer - Internship

June 2017 – Sep 2017

*Bird.i*

- Co-authored a customer facing model-to-production pipeline in AWS Lambda with Docker based CI/CD.
- Developed a novel sweep & prune algorithm for real-time geographical analytics.

### Software Engineer

September 2016 – June 2017

*Bcomp Switzerland*

- Prototyped and shipped a roof-mounted LED display for live metrics of an electric race car including real-time software for monitoring.

## PUBLICATIONS (SEE [GOOGLE SCHOLAR](#))

---

- GRILLBot: A conversational agent for real-world tasks** Amazon Alexa Proceedings 2022  
*Carlos Gemmell, S. Fischer, I. Mackie, P. Owoicho, F. Rossetto, Jeff Dalton*
- VILT: Video Instructions Linking for Complex Tasks** IMuR 2022  
*Sophie Fischer, Carlos Gemmell and Iain Mackie, Jeffrey Dalton* (under submission, accepted)
- GRILLBot: An Assistant for Real-World Tasks with Neural Semantic Parsing and Graph-Based Representations** SIGDial 2022  
*Carlos Gemmell, I. Mackie, P. Owoicho, F. Rossetto, S. Fischer, Jeff Dalton* (under submission, accepted)
- CODEC: Complex Document and Entity Collection** SIGIR 2022  
*Iain Mackie, P. Owoicho, Carlos Gemmell, S. Fischer, S. MacAvaney, Jeffrey Dalton*
- Relevance Transformer: Generating Code Snippets with Relevance Feedback** SIGIR 2020  
*Carlos Gemmell, Federico Rossetto, and Jeff Dalton*
- Glasgow GRILL Lab at the Conversational Assistance Track** TREC 2020  
*Carlos Gemmell, and Jeff Dalton*

## TEACHING AND OUTREACH

---

- Workshop talk at Search-Oriented Conversational AI (SCAI)** July 2022 – SIGIR, Madrid, Spain  
*Question answering with large language models in GRILLBot*
- Invited talk at Stanford NLP** March 2022 – Stanford, Palo Alto, US  
*Conversation Search and Language Understanding*
- Multimodal NLP workshop at USC** March 2022 – Los Angeles, US  
*Multimodal Conversation Search*
- Invited talk at Universidade NOVA de Lisboa:** February 2022 – Lisbon, Portugal  
*TREC CAsT: Current Challenges and Future Directions*
- Teaching assistant for bachelors NLP course (COMPSCI4074):** University of Glasgow, Scotland  
*Lectures: Word Embeddings and Sequence to Sequence Models*
- University NLP Workshops:** October 2019 – University of Glasgow, Scotland  
*Machine Learning from Square One*
- Organiser:** November 2020 – January 2022  
*University of Glasgow Information Retrieval seminar and NLP reading group*
- Reviewer for SIGIR 2021, EMNLP 2021, ECIR 2021, ACL 2021, EMNLP 2022** .

## TECHNICAL SKILLS

---

**Languages:** Python, Java, C/C++, JavaScript, SQL  
**Libraries:** Pytorch (extensive), Tensorflow, AllenNLP, NumPy/SciP, Matplotlib, Terrier, GRPC  
**Developer Tools:** Git, Docker, Kubernetes, Google Cloud Platform, AWS

## PERSONAL

---

**Languages:** English (Native), Spanish (Native), French (Native)  
**Interests:** Rock Climbing, Longboarding, Latin Dancing, Photography, Hardware Hacker