

Tony Bruguier

US & EU Citizen

tony.bruguier@gmail.com

Professional experience

Google DeepMind (lead individual contributor unless noted otherwise)

- Audio keyword retrieval for biasing for LLM contextualization. Training with Jax/Numpy in Python, on-device also implementation in C++ and added plumbing to Java app (IC & tech lead)
- Gemini [Astra](#) (part time IC for data preparation)
- Training on-device speech recognition models with LLMs (Jax/Numpy in Python, serialization in TensorFlow lite)
- Designed RNN-T based algorithm for pronunciation learning ([paper](#)), implemented backend in C++, used in foreign language learning ([press release](#))
- Improved automatic speech recognition of uncommon words for RNN-T end-to-end ASR models ([paper](#) and [patent](#)). Training with TensorFlow in Python, inference implementation in C++
- Large-scale learning of pronunciation from correction logs and YouTube human captions. Implementation in C++ equivalent of Apache Beam. Currently used by text-to-speech models throughout the stack.
- Specification of DNN and LSTM ASR models to reduce inference cost. Research in Python, implementation in C++ ([paper](#) and [patent](#))
- Developed Pub/Sub messaging backend system for improvement of long-tail words using pronunciation learning ([paper](#) and [patent](#)).
- All above using Mercurial and Perforce, using GPU-accelerated hardware
- Volunteer, non-lead IC, help with Quantum Computing open source library ([Github](#))

Brion Technologies (computational lithography startup, later acquired by ASML)

- Creation of a gauge selection algorithm to lower cost of metrology ([two patents](#)) C++ implementation on customers' premises in fabs
- Created and implemented a new model for resist process ([patent](#)). Research using Matlab and implementation in C++
- Improved OPC accuracy using gauge selection algorithm ([paper](#)) before SEM measurement.
- Wrote both FFTW and FPGA accelerated code

Philips Medical Systems (Intern) 3 summers

- Improved on automated EKG interpretation algorithms for cardiographs

Education

California Institute of Technology (Caltech): PhD, Electrical Engineering, 4.0 GPA

Encoding of Financial Signals in the Human Brain ([thesis](#))

Themes: Analysis of fMRI, inferential statistics, neurobiology, and behavioral economics.

ESIEE, Paris: *Diplôme d'Ingénieur*, major in signal processing

Graduated valedictorian

Complete list of publications and patents

See <https://www.bruguier.com/research.html>