

# William Callaghan

hello@willcallaghan.ca

<http://willcallaghan.ca/>

## EDUCATION

**Master of Mathematics**, Computer Science 2015 - 2018  
University of Waterloo, Waterloo, Ontario, Canada  
Advisor: Edith Law

**Honors Bachelor of Science**, Computer Science, Microbiology & Immunology 2015  
University of Western Ontario, London, Ontario, Canada  
Thesis Supervisor: Mark Daley

## WORK EXPERIENCE

**Data Engineer II** Nov 2019 – Present  
Borealis AI, Montreal, Quebec, Canada

- Driving the development of tools for unified data access and governance across the machine learning lifecycle.

**Data Engineer I** May 2018 – Nov 2019  
Borealis AI, Montreal, Quebec, Canada

- Developed Spark applications to support efficient data transformations on large datasets through all stages of the machine learning lifecycle.
- Developed and optimized feature engineering and hyperparameter tuning pipelines. Improved runtime by orders of magnitude.
- Developed internal tools to simplify querying of large, distributed datasets composing of multiple tables.
- Contributed to writing documentation and performing analysis for model validation and governance.
- Worked with various stakeholders, including business representatives, researchers and engineers to bring model from concept into production.

**Software Developer (Part-Time)** January 2017 – August 2017  
eSentire Inc., Cambridge, Ontario, Canada

- Provided development support to ongoing analytics projects.
- Network Scan Analytics Lead (May 2017 – August 2017)

**Software Developer Intern**, Data Analytics May 2016 – December 2016  
eSentire Inc., Cambridge, Ontario, Canada

- Built an engine to perform complex queries on large data sets with mission critical response times.
- Created RESTful HTTP APIs to interact with various components of the analytics pipeline.
- Worked with Analytics Lead and other stakeholders to define ongoing projects and infrastructure.
- Technologies: Spark, Cassandra, Kafka, Alluxio, HDFS
- Languages: Python, Scala

**Graduate Research Assistant** January 2017 - April 2018  
University of Waterloo, Waterloo, Ontario, Canada September 2015 - May 2016

**Teaching Assistant** January 2017 - April 2018  
University of Waterloo, Waterloo, Ontario, Canada September 2015 - May 2016

**Software Developer (Contract)**

May 2015 - September 2015

eSentire Inc., Cambridge, Ontario, Canada

- Worked on tools involved in real-time capture and analysis of raw network traffic.
- Core developer on Targeted Retrospective Analysis Platform, bringing prototype to product.
- Extended eSentire's Deep Packet Inspection tools.

**QA Developer Intern**

May 2014 - August 2014

Pelmorex Media Inc., The Weather Network, Oakville, Ontario, Canada

- Developed web automation framework and tests for web products in Java, using Selenium.

**Technology Manager**

May 2013 - August 2013

Engineering Outreach, University of Waterloo, Waterloo, Ontario, Canada

- Managed a group of 20+ people in establishing a technology camp curriculum for multiple age groups.
- Monitored daily technology camp operations.

**TECHNICAL SKILLS****Languages:** Python, Java, Bash**Knowledge/Experience:**

- Apache Spark, Python Pandas and surrounding ecosystem.
- Relational & NoSQL Databases (PostgreSQL, Cassandra, Teradata)
- Containerization & Cloud Orchestration (Docker, Openshift)
- Scheduling (Apache Airflow)
- RESTful API Development
- Continuous Integration and Deployment (Git, Jenkins)

**RESEARCH****SELECTED PROJECTS****HeartBeat**

Framework to combine machine and human intelligence for the scalable and accurate analysis of human clinical phonocardiograms. This was the topic of my Master's thesis at the University of Waterloo.

**Instance Selection Methods in Machine Learning for EEG Analysis**

Researched, implemented and evaluated active learning and instance reduction strategies for EEG analysis (spindle detection and sleep stage classification) as a part of final projects for 'Human-In-The-Loop Systems' and 'On The Synergy Between Computer Science and Biology' courses at the University of Waterloo.

**Modelling Influence in Social Networks**

Researched and proposed a new model for learning the most influential agent in a social network. This was done as a final project for 'Advanced Topics in Artificial Intelligence: Trust and Online Social Networks' course at the University of Waterloo.

**Using Kalman Filtering and Lasso Regularization to Generate Brain Networks**

Researched and proposed a new method of constructing functional brain networks as diagnostic markers for neurological disease from fMRI data. This was the topic of my Bachelor's thesis at the University of Western Ontario.

## CONFERENCE PAPERS

**Callaghan, W.**, Goh, J., Mohareb, M., Lim, A., & Law, E. (2018). MechanicalHeart: A Human-Machine Framework for the Classification of Phonocardiograms. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 28.

## WORKSHOP PAPERS

Schaekermann, M., Law, E., Williams, A. C., & **Callaghan, W.** (2016). Resolvable vs. irresolvable ambiguity: A new hybrid framework for dealing with uncertain ground truth. *In Proceedings of the 1st Workshop on Human-Centered Machine Learning at SIGCHI*.

Williams, A. C., Bradshaw, J., Schaekermann, M., Tse, T., **Callaghan, W.**, & Law, E. (2016). The Big Picture: Preserving Context in the Decomposition of Complex Expert Tasks. *In 1st Workshop on Micro-productivity at SIGCHI*.

## INVITED TALKS

**Callaghan, W.** (2017). Fighting Cybercrime: A Joint Task Force of Real Time Data and Human Analytics. Presented at the Databricks Spark Summit East Conference, Boston, MA.

## THESES

**Callaghan, W.** (2018). *A Human-Machine Framework for the Classification of Phonocardiograms* (Master's thesis, University of Waterloo).

**Callaghan, W.** (2015). *Using Kalman Filtering and Lasso Regularization to Generate Brain Networks* (Bachelor's thesis, University of Western Ontario).