Craig A. Glastonbury

Curriculum vitae

10 Cherwell Street
OX4 1BG, Oxford
https://github.com/GlastonburyC
📱 *+44 753 552 5258*
✉ *craig@well.ox.ac.uk*

Work

2017

Postdoctoral Researcher, *University of Oxford - Big Data Institute (BDI)*, UK.

- Developed an automated cell size estimator (Dilated U-net implementation) to quantify millions of cells from histology imaging data.
- Company formed to speed up a commonly used diagnostic test routinely used throughout hospitals and GP offices. Implemented a CNN that will form part of a mobile app for rapid diagnosis.
- Building Denoising autoencoders (DAEs) for capturing genetic variability and predicting disease burden from genotype alone.

2012
2013

Ethical Hacker, *Deloitte*, London.

Security cleared by the Ministry of Defense (MoD) and Employed by Deloitte UK to penetration test networks, web applications, mobile applications and perform social engineering for several FTSE100 companies, government agencies and departments.

Education

2013
2016

Statistical Genomics PhD, *King's College*, London.

My PhD was focused on using statistics and computational methods to understand the genetic control of gene expression in humans and how genetic variation that modulates expression influences cardio-metabolic traits and obesity. A primary focus of my thesis was to understand gene-by-environment interactions and how they contribute to complex disease. This work resulted in several papers, including first authorship in The American Society of Human Genetics, and several contributions to papers in Nature Genetics, Diabetes and Obesity.

2016

BSc Mathematics, *The Open University*.

Whilst studying for a Ph.D and now in Postdoctoral Research I am entering the second year of an undergraduate mathematics degree, part-time. Current modules cover Essential mathematics 1 & 2 - Calculus methods, number theory, Linear algebra.

2009
2012

BSc Biological Sciences, *Imperial College*, London.

Specialism in Genetics and Statistics.

Github/Deep learning implementations

- **DenseNet and DenseNet-BC implementation**. I implemented DenseNet and applied it to the CIFAR10 dataset. The test error is lower than the original DenseNet paper. Implementation here.
- **U-net/Dilated U-net implementation** - I have used this in currently unpublished work to segment 13.5 million cells from histology imaging data. Publicly, I have also applied it to the Cell Tracking Challenge Dataset (ISBI) - ranking in the top 100 academic submissions (68th). Implementation here.
- **Nature Conservancy Kaggle competition** - Classification task of 8 species of fish, imaged from cameras on fishing traulers. Used Overfeat (a bounding box regression CNN) and also

built an InceptionV3 pre-trained network and ensembled it with a ResNet50 to obtain a private leaderboard score of 30th/2293. Implementation here and blog here.

○ **Intel Cervix type** - **Kaggle Competition** - Trained both a localisation (bounding box) network (YOLO9000/darknet) and several pre-trained deep nets to classify different types of cervix's to improve cervical cancer diagnosis. Obtained a private leaderboard score of 18th/848. Implementation here and blog here.

Additional Training/Courses

○ **Kaggle competition expert (top 0.7% overall - 507th/70,500)**. Kaggle is a world leading machine learning competition website in which companies provide data and problems for the kaggle community to solve. Competitions I have participated in include fish species detection from hauler boat video footage, lung cancer nodule detection and prediction from CT scans, cervical cancer type classification and assessing duplicate questions from the Quora database. For all these competitions I have used several Deep learning methods (CNNs - Regression & classification / LSTMs).

○ **Machine Learning MRes module (2016)** - **King's College London.** Course covering machine learning, probability and statistics. Taught by Professor Giovanni Montana

○ **Lisbon Machine Learning Summer School (LxMLS, 2017)**. Selected to participate in a highly competitive machine learning school/conference co-sponsored by Google.

○ **Learning From Big Data** - **University of Oxford Doctoral Training Program.** Course taught by Andrew Zisserman covering SVMs, kernel trick, Deep learning (ANNs,CNNs, autoencoders, generative models).

○ **Blog** My blog (https://glastonburyc.github.io) where I write about Deep learning - both theory and applied to competitions and genetics research.

Teaching/Mentoring

○ **Jonas Bovijn (DPhil - 2017-2020) (Christ Church)** - Using Machine learning with genetics and genomics to improve drug target validation.

○ **Peter Strain (DPhil - 2017-2020) (St Annes)** - The genetic basis of organ structure variability. Deep learning applications to 100k whole-body MRIs

○ **Interviewing** - I am routinely involved in recruitment efforts for staff at the Big Data Institute

Languages	Python(*4+ years*), R (*8+ years*), & Bash
Frameworks	Keras, Tensorflow,Theano, Tensorboard
Cluster Computing	AWS, OpenStack cluster configuration, elasticluster, virtualisation, SLURM & SGE environments

Grants

○ **Novo Nordisk Pump Priming Grant** - *Principal Investigator* - Using deep learning to perform automated cellular phenotyping with imaging data (Awarded £25k)

○ **NVIDIA GPU Grant** - *Principal Investigator* - Titan X Pascal GPU for my work on using image segmentation methods for histology imaging data. (Awarded £1.5k)

Awards

○ **Genetics Society Junior Scientist Conference award** - **2017** - Awarded funding to attend the

Lisbon machine learning school (LxMLS)

○ **Genetics Society Junior Scientist Conference award 2016** - Awarded funding to attend the American Society of Human Genetics (ASHG16)

○ **Hackseq 2016** - Selected to attend and participate in a genomics hackathon at the University of British Columbia, Vancouver. Over a weekend our team came up with a method to optimize parameters for genome assembly https://f1000research.com/articles/6-197/v1

○ **American Society of Human Genetics** (ASHG 2015) - Reviewers Abstract award (Top 10% of abstracts submitted)

○ **American Society of Human Genetics (ASHG15)** - Selected for ASHG Poster highlight in the cardiometabolic trait section (Top 3% of abstracts)

○ **Leena Peltonen School of Human Genetics (2015)** - Selected as 1 of 20 students to attend a highly competitive school in which students are selected on merit to present to, and network with, 20 world leading professors in the field of human genetics.

Platform Presentations

○ **ASHG Platform Presentation, 2016** - Adiposity-dependent interactions on multi-tissue transcriptomes

○ **Quantitative Genomics 2014 & 2015 Presentation** - Detecting gene by environment interactions on expression in multiple tissues (Quantitative Genomics 2014/2015)

○ **King's College London RNA-Seq day** - Deconvolution of Adipose tissue cell type composition using RNA-Seq (Invited speaker - 2015)

Poster Presentations

○ **ASHG Poster Presentation, 2017** - In-silico characterization of cell-type composition in adipose tissue: implications for 'omic analyses and associations to adiposity measures

○ **ASHG Poster presentation, 2016** - Population level variability in adipose tissue cell type composition

○ **King's College London RNA-Seq day** - Deconvolution of Adipose tissue cell type composition using RNA-Seq (Invited speaker - 2015)

○ **ASHG Poster presentation, 2015** - Adiposity dependent regulatory effects on multiple tissue transcriptomes - Glastonbury et al., (ASHG Poster presentation, 2015)

Journal Publications

2017   Cohen, Joseph Paul, Genevieve Boucher, **Craig A. Glastonbury**, Henry Z. Lo, and Yoshua Bengio. "Count-ception: Counting by Fully Convolutional Redundant Counting". In: *ICCV Bioimaging* abs/1703.08710. URL: http://openaccess.thecvf.com/content_ICCV_2017_workshops/papers/w1/Cohen_Count-ception_Counting_by_ICCV_2017_paper.pdf.

2017   Pallister, T, MA Jackson, TC Martin, **Glastonbury, CA**, A Jennings, M Beaumont, RP Mohney, KS Small, A MacGregor, CJ Steves, et al. "Untangling the relationship between diet and visceral fat mass through blood metabolomics and gut microbiome profiling". In: *International Journal of Obesity*.

2017   Pulit, Sara L, Samantha Laber, Craig A Glastonbury, and Cecilia M Lindgren. "The genetic underpinnings of body fat distribution". In: *Expert Review of Endocrinology & Metabolism*.

2016   Bailey, et al.*. "Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open angle glaucoma". In: *Nature Genetics*.

2016   **Glastonbury, C**, A. Vinuela, A. Buil, R. Durbin, E. Dermitzakis, T. Spector, and K. Small. "Adiposity-Dependent Regulatory Effects on Multi-tissue Transcriptomes". In: *AJHG*.

2016   Kaul, S, H Xu, E Maruko, **Glastonbury, C**, K Small, G Dallinga-Thie, M Civelek, M Thomas, I Goldberg, and M Sorci-Thomas. "Procollagen C-endopeptidase enhancer protein 2 (PCPE2) Deficiency Profoundly Affects Adipose Distribution in Mice and Humans Linking HDL Metabolism to Adipocyte Biology". In: *Arteriosclerosis, Thrombosis, and Vascular Biology*.

2016   Menni, C, **Glastonbury, C**, K Nikolaou, K Small, K Mahney, T Spector, and A.M Valdes. "Metabolomic profiling to dissect the role of visceral fat in cardiometabolic health". In: *Obesity*.

2016   Small, K, L Quaye, A Hough, M Todorcevic, A Mahajan, M Horikoshi, A Buil, A Viñuela, **Glastonbury, C**, J Brown A Bell, R Cox, Gloyn A, Karpe F, and McCarthy M. "Characterisation of the KLF14 trans-regulatory network". In: *Submitted (Nature)*.

2016   Tsai, PC*, **Glastonbury, C***, A Vineula, R Durbin, E Dermitzakis, T Spector, and K. Small. "Tobacco smoke modulates gene expression and DNA methylation via genetic variation in multiple human tissues". In: *Submitted PloS genetics*.

References

Cecilia Lindgren
Assistant Professor & Director of Undergraduate Studies
Big Data Institute
University of Oxford
✉ celi@well.ox.ac.uk
☎ 01865 287591

Dr Kerrin Small
Senior Lecturer
Dept. Twin Research & Genetic Epidemiology
King's College London
✉ kerrin.small@kcl.ac.uk
☎ 207 188 7188 ext 51719

Professor Tim D. Spector
Director of the Dept. of Twin Research
Dept. Twin Research & Genetic
Epidemiology
King's College London
✉ tim.spector@kcl.ac.uk
☎ +44 (0) 207 188 6765