#### **Research Interests and Experience**

#### Deep learning and machine learning

- · Deep learning and computer vision for wildlife monitoring
- Graph neural networks for pharmaceutical discovery
- Rank learning for drug recommendation systems
- Machine learning for prediction of disease-causing genomic variants

#### **Biomedical data science**

- · Integrative statistical analysis of large-scale molecular and clinical data
- · Bioinformatics analysis of high-dimensional next-generation sequencing data

## Education

2011–2015 WGRIERSTY NICTA	<ul> <li>PhD in Applied Machine Learning<sup>1,2</sup></li> <li><sup>1</sup>Institute for Integrated and Intelligent Systems, Griffith University, Australia</li> <li><sup>2</sup>NICTA – Australia's ICT Research Centre of Excellence, Australia</li> <li>Advisers: Prof Bela Stantic<sup>1</sup>, Prof Yaoqi Zhou<sup>1</sup>, Prof Abdul Sattar<sup>1,2</sup></li> <li>Project: Predicting stability and functional changes induced by protein mutations with a machine learning approach</li> </ul>
2007–2010 ())	Master's Degree in Informatics Faculty of Informatics, Masaryk University, Czech Republic <u>Awards</u> : Dean's awards for excellent study results and outstanding thesis
2007	<b>Visiting Student</b> School of Computer Science, University College Dublin, Ireland <u>Awards</u> : Erasmus Student Exchange Scholarship
2004–2007 W	<b>Bachelor's Degree in Applied Informatics</b> Faculty of Informatics, Masaryk University, Czech Republic <u>Awards</u> : Average grade ranked within the best 2% of 2006 and 2007 graduates

## Professional Experience

#### 2022–now Research Fellow

Coastal and Marine Research Centre, Australian Rivers Institute, Griffith University Advisers: Prof Kylie Pitt, Prof Bela Stantic

## Project: Deep learning and computer vision for marine wildlife monitoring

In this Blue Economy CRC-funded project, I am developing deep learning and computer vision algorithms for jellyfish detection and fish welfare monitoring. The system will be deployed to salmon farms operated by our three industry partners. Main objectives are achieving a real-time and high-accuracy monitoring across hundreds of cameras in a challenging low-visibility environment.

#### 2018–2022 Postdoctoral Researcher<sup>1,2</sup>

<sup>1</sup>Institute for Artificial Intelligence, Medical University of Vienna, Austria (since 2022) <sup>2</sup>CeMM Research Center for Molecular Medicine, Austrian Academy of Sciences

- Ce-M-M-
- Adviser: Prof Christoph Bock (https://www.bocklab.org)

# <u>Projects</u>: Statistical, bioinformatics, and machine-learning modelling for highdimensional molecular and clinical data

I was the lead data scientist for an international clinical trial to test non-specific effects of tuberculosis vaccine. I designed and conducted integrative statistical and bioinformatics analysis of high-dimensional ATAC-seq, single-cell RNA-seq, and immunopheno-typing data. My additional projects focused on graph neural networks for pharmaceutical discovery and explainable deep learning for single-cell RNA-seq.

## 2016–2018 Postdoctoral Researcher

**ETH***zürich* Department of Biosystems Science and Engineering, ETH Zurich, Switzerland Adviser: Prof Karsten Borgwardt (https://bsse.ethz.ch/mlcb)

Project: Machine learning for precision medicine drug recommendation systems

In this precision medicine project we developed a novel approach to recommend cancer treatments based on high-dimensional transcriptomics profiles of cancer cell lines. Our data-driven learning-to-rank approach achieved high clinical relevance compared to other regression-based methods by addressing the sparse nature of clinical records and their bias towards effective treatments. This work received two best poster awards at premiere national and international conferences.

## 2015–2016 Visiting Researcher

Adviser: Prof Yaogi Zhou (https://sparks-lab.org)

## **<u>Project</u>**: Machine learning for prediction of disease-causing genomic variants

Institute for Integrated and Intelligent Systems, Griffith University, Australia

Capitalising on research from my doctoral thesis, we developed a new machine-learning method for prediction of disease-causing synonymous variants, which are particularly challenging as the protein sequence remain unchanged despite a variant occurring at the DNA level. I addressed this challenge with mining and engineering new discriminative features from large biomedical databases. The accuracy of our prediction model superseded all other available methods.

# 2008–2010 Software Developer and Consultant

ARTIN, Czech Republic

## Funding

- 2019–2021 Marie Skłodowska-Curie Actions Individual Fellowship (186,000 EUR) Success rate of 12.4%; competitive fellowship sponsored by the European Commission <u>Project</u>: Deep learning for single-cell transcriptomics data analysis
   2011–2014 NICTA Project Award Scholarship (28,000 AUD)
- 2011–2014 Griffith University Postgraduate Research Scholarship (74,000 AUD + tuition fee waiver)
- **Dec 2010** NICTA Summer Research Scholarship (5,000 AUD)

#### Awards

- Jan 2024 "Multi-omics analysis of innate and adaptive responses to BCG vaccination reveals epigenetic cell states that predict trained immunity" was selected for *Nature Research Hinglights*
- **Sep 2017** Online Poster Award (1st place) at [BC]<sup>2</sup> Basel Computational Biology Conference 2017 for "Kernelized rank learning for personalized drug recommendation", Basel, Switzerland
- **Jul 2017** F1000Research Poster Award at the international conference ISMB/ECCB 2017 for "Kernelized rank learning for personalized drug recommendation", Prague, Czech Rep.
- Jun 2014 Winner of Three-Minutes Thesis Competition, School of ICT, Griffith University
- Apr 2014 Gold Coast Association of Postgraduates Conference Funding Award, Griffith University
- Jan 2013 Griffith University Conference Travel Grant, Griffith University
- Dec 2010 2nd best paper in the workshop Keymaker 2010, Brno, Czech Rep.
- May 2010 2nd place in Czech and Slovak university research competition, Ostrava, Czech Rep.
- Mar 2010 Dean's award for outstanding thesis, Faculty of Informatics, Masaryk University
- Mar 2010 Dean's award for excellent study results, Faculty of Informatics, Masaryk University

## **Publications**

## # These authors contributed equally and they are equal "first" authors

Moorlag<sup>#</sup> SJCFM, **Folkman<sup>#</sup> L**, ter Horst<sup>#</sup> R, Krausgruber<sup>#</sup> T, …, Netea<sup>\*</sup> MG, Bock<sup>\*</sup> C: "Multi-omics analysis of innate and adaptive responses to BCG vaccination reveals epigenetic cell states that predict trained immunity".

Immunity (2024), 57(1), 171–187, impact factor 32.4, 🕮

- de Bree<sup>#</sup> LCJ, Mourits<sup>#</sup> VP, …, **Folkman L**, …, Netea MG: "Circadian rhythm influences induction of trained immunity by BCG vaccination". *The Journal of Clinical Investigation* (2020), 130(10), 5603–5617.
- Pejaver V, Babbi G, Casadio R, **Folkman L**, ..., Bromberg Y: "Assessment of methods for predicting the effects of PTEN and TPMT protein variants". *Human Mutation* (2019), 40(9), 1495–1506.
- Savojardo<sup>#</sup> C, Petrosino<sup>#</sup> M, …, **Folkman L**, …, Capriotti E: "Evaluating the predictions of the protein stability change upon single amino acid substitutions for the FXN CAGI5 challenge". *Human Mutation* (2019), 40(9), 1392–1399.
- Clark WT, …, Folkman L, …, LeBowitz JH: "Assessment of predicted enzymatic activity of alpha-Nacetyl-glucosaminidase (NAGLU) variants of unknown significance for CAGI 2016". *Human Mutation* (2019) 40(9), 1519–1529.
- He<sup>#</sup> X, Folkman<sup>#</sup> L, Borgwardt K: "Kernelized rank learning for personalized drug recommendation". *Bioinformatics* (2018), 34(16), 2808–2816.
- Livingstone<sup>#</sup> M, **Folkman<sup>#</sup> L**, Yang Y, Zhang P, Mort M, Cooper DN, Liu Y, Stantic B, Zhou Y: "Investigating DNA-, RNA-, and protein-based features as a means to discriminate pathogenic synonymous variants". *Human Mutation* (2017), 38(10), 1336–1347.
- **Folkman L**, Stantic B, Sattar A, Zhou Y: "EASE-MM: sequence-based prediction of mutation-induced stability changes with feature-based multiple models". *Journal of Molecular Biology* (2016), 428(6), 1394–1405.
- Folkman L, Yang Y, Li Z, Stantic B, Sattar B, Mort M, Cooper DN, Liu Y, Zhou Y: "DDIG-in: detecting disease-causing genetic variations due to frameshifting indels and nonsense mutations by sequence and structural properties at nucleotide and protein levels". *Bioinformatics* (2015) 31(10), 1599–1606.
- Folkman L, Stantic B, Sattar A: "Feature-based multiple models improve classification of mutationinduced stability changes". BMC Genomics (2014) 15(Suppl 4), S6. 44
- **Folkman L**, Stantic B, Sattar A: "Towards sequence-based prediction of mutation-induced stability changes in unseen non-homologous proteins". *BMC Genomics* (2014) 15(Suppl 1), S4.
- Folkman L, Stantic B, Sattar, A: "Sequence-only evolutionary and predicted structural features for the prediction of stability changes in protein mutants". BMC Bioinformatics (2013) 14(Suppl 2), S6. 4
- Higgs T, **Folkman L**, Stantic B: "Combining protein fragment feature-based resampling and local optimisation". *in* 'IAPR International Conference on Pattern Recognition in Bioinformatics (PRIB)', Vol. 7986 of *LNCS*, pp. 114–125, Springer, 2013.
- **Folkman L**, Pullan W, Stantic B: "Generic parallel genetic algorithm framework for protein optimisation". *in* 'Algorithms and architectures for parallel processing (ICA3PP)', Vol. 7017 of *LNCS*, pp. 64–73, Springer, 2011.
- Stetsko A, **Folkman L**, Matyas V: "Neighbor-based intrusion detection in wireless sensor networks". *in* 'Wireless and Mobile Communications (ICWMC)', pp. 420–425, IEEE, 2010.

#### Teaching 2022 **Teaching Assistant** Medical University of Vienna, Austria MEDICAL UNIVERSITY OF VIENNA Biomedical Informatics and Genomic Medicine (master's degree course) Developed and delivered practical sessions on statistics and machine learning 2022 Lecturer University of Applied Sciences Technikum Vienna, Austria FH Un TECHNIKUM Special Topics of Artificial Intelligence (master's degree course) WIEN Developed and delivered two lectures on deep learning in biomedicine 2021 **Guest Lecturer** MEDICAL UNIVERSITY OF VIENNA Medical University of Vienna, Austria Biomedical Informatics and Genomic Medicine (master's degree course) Developed and delivered the lecture and practical session on machine learning 2020-2021 **Guest Lecturer** CeMM Research Center for Molecular Medicine, Austrian Academy of Sciences, Austria Ce-M-M- Machine-learning workshop for PhD students • Developed and delivered the lecture and practical session 2016 **Teaching Assistant ETH** zürich Department of Biosystems Science and Engineering, ETH Zurich, Switzerland Data Mining I (master's degree course) Co-developed and delivered the practical sessions, marked homework and exam 2012-2013 **Teaching Assistant** School of Information and Communication Technology, Griffith University, Australia Griffith • Introduction to Programming (bachelor's degree course) • Object-Oriented Programming (bachelor's degree course) Database Systems and Administration (bachelor's degree course) Received a mean score of 4.78 (out of 5) and 5 award nominations: "Lukas was also an excellent tutor. He explains things so clearly and breaks up complex subjects into something simple which helped me understand topics better." Database Technology (master's degree course) Received a mean score of 4.48 (out of 5) and 2 award nominations: "The nature of this lecturer was very engaging, and kept the students focused and interested." **Student Mentoring and Co-Supervision** 2022-now Quynh Le Khanh Vo, PhD candidate (Griffith University, Gold Coast, Australia) My role: co-supervisor Project: Advanced monitoring to maximise fish welfare in offshore aquaculture 2020-2022 Stephan Reichl, PhD candidate (CeMM and Medical University of Vienna, Austria) My role: mentor, a co-authored paper in preparation Project: Statistical and bioinformatics analysis of high-dimensional molecular data 2017 Elli Tzini, Master's degree student (ETH Zurich, Basel, Switzerland) My role: co-supervisor Project: Machine learning for personalised drug recommendation systems Elli continued her career as a machine-learning engineer. 2016-2017 Udo Gieraths, PhD candidate (ETH Zurich, Basel, Switzerland) My role: co-supervisor

<u>Project</u>: Machine learning for prediction of disease-linked genes Udo continued his career as a data scientist.

- 2016 Caroline Weis, Master's degree student (ETH Zurich, Basel, Switzerland) <u>My role</u>: co-supervisor <u>Project</u>: Machine learning for prediction of disease-causing genomic variants *Caroline continued her research as a PhD candidate at ETH Zurich.*
- 2013 Mark Livingstone, Honours thesis, (Griffith University, Gold Coast, Australia) <u>My role</u>: mentor, co-authored a paper with Mark <u>Project</u>: Impact of single amino acid change upon protein structure *Mark continued his research as a PhD candidate at Griffith University*

## **Institutional Service**

- **2020–2022** Data ambassador and manager, part of a working group for implementing and enforcing institute-wide "FAIR" data principles
- **2017** Contributed to the conceptualisation and presentation of a video game to popularise machine learning for medicine at the Swiss national scientific outreach event Scientifica
- 2012–2013 Founding chair of the IIIS/ICT Higher-Degree-Research Student Community

## Academic Service

ReviewingBioinformatics (impact factor 5.8)<br/>Journal of Machine Learning Research (A\*-ranked by CORE)<br/>European Conference on Machine Learning (A-ranked by CORE)<br/>Journal of Molecular Biology (impact factor 5.6)<br/>Scientific Reports (impact factor 4.6)

#### **Skills and Competencies**

Soft Skills\* Academic writing, public speaking, time management, project management, supervising Master and PhD projects (\*including formal training via workshops)
 Programming Python (including Scikit-Learn, PyTorch, Pandas, Statsmodels, NumPy), Unix-based

- systems, bash, R, Java, C and C++
- Languages English fluent, German beginner, Slovak native