

Dr. Rosa Filgueira Vicente

Chancellor Fellow, EPCC, University of Edinburgh
<http://www.rosafilgueira.com>

October 10, 2024
Email: r.filgueira@ed.ac.uk

Professional Experience

- Chancellor Fellow, EPCC, University of Edinburgh, UK, since April 2024
- Lecturer, School of Computer Science, University of St Andrews, UK, since February 2022 - April 2024
- Assistant Professor, MACS, Heriot-Watt University, UK, March 2021 - February 2022
- Vice President, Applied Research, JP Morgan Chase & Co, UK, October 2020 - March 2021
- Research Fellow, EPCC, University of Edinburgh, UK, June 2019 - October 2020
- Data Architect, EPCC, University of Edinburgh, UK, March 2018 - May 2019
- Senior Data Scientist, British Geological Survey, Edinburgh, UK, October 2016 - February 2018
- Senior Research Associate, School of Informatics, University of Edinburgh, July 2016 - October 2016
- Research Associate, School of Informatics, University of Edinburgh, June 2011 - July 2016
- Research Assistant and Teaching Fellow, University Carlos III Madrid, Spain, Sept. 2005 - Sept. 2010

Education

- PhD in Computer Science, University Carlos III Madrid, Spain, September 2010
- Master of Technology in Computer Science, University Carlos III, Madrid, Spain, July 2007
- Master in e-Commerce, University Carlos III Madrid, Spain, July 2004
- Master of Computer Science, University of Deusto, Spain, June 2003
- Bachelor of Computer Science, University of Deusto, Spain, June 2001

Visiting Positions and Consultancy

- Honorary Lecturer, School of Computer Science, University of St Andrews, since May 2024
- Honorary Research Fellow, School of Informatics, University of Edinburgh, April 2021 - March 2024
- Research Visitor, ISI, University of Southern California, October 2017 - December 2017
- Honorary Research Fellow, School of Informatics, University of Edinburgh, January 2017 - March 2018
- Research Visitor, ISI, University of Southern California, Los Angeles, March 2016 - May 2016
- Consultancy Services for Accelogic, Edinburgh, August 2014 - April 2016
- Research Visitor, AIST, Informatics Technology Research Institute, Tsukuba, Japan, July 2013 - July 2013
- Research Visitor, EPCC, University of Edinburgh, February 2011 - October 2011

Research Funding

- Decoding Democracy: Analyzing Public Participation in Constitutional Reform with Deep Learning, The British Academy, 2023, GBP 2,150
- Towards Large-scale Cultural Analytics in the Arts and Humanities, UKRI/AHRC, GBP 100,000
- Research Fellowship in Digital Scholarship, National Library of Scotland, 2021, GBP 7,500
- COVID-19 Response & Recovery Small Grant Funding, Data-Driven Innovation grant, 2020, GBP 10,000
- Text and Data Mining, Data-Driven Innovation grant, 2019, GBP 133,000
- Two Postdoctoral and Early Career Researcher Exchanges grant, SICSA, 2015/17, GBP 7,500/13,580
- Observation Imaging and Interpretation, UoE grant, 2016-2017, GBP 3,000
- Combining Knowledge to Pioneer Big-Data Solutions for SKA, CSIC grant, 2016-2018, EUR 26,000
- Driving UK HPC Enabled Science and Innovation through US Collaborations to SC14, 2014, GBP 2,500
- Two HPC-Europa2 PRACE grants, 2011, GBP 6,000

Awards and Scholarships

- International HPC Summer school student travel award, Ljubljana, Slovenia, 2016
- Best poster award, SACSIS 2012, Kobe, Japan, 2012
- Scholarship for Doctoral Training Programme, University Carlos III Madrid, Spain, 2005, EUR 12,000
- Student travel award, TCPP PhD Forum at IPDPS Conference, Miami, 2008, USD 500

Teaching

- Course Leader, *Principles of Computer Communication Systems*, University of St Andrews, since 2022/23
- Course Co-Leader, *Databases*, University of St Andrews, since 2021/22
- Invited two-day intensive course for the MSc in High Performance Computing at the UC3M, Spain, 2024
- Invited three-day intensive course for the MSc in Artificial Intelligence at the UPM, Spain, 2022
- Invited lectures (twice) for the MSc in Informatics at the University of Lleida, Spain 2013-2014
- Teaching Assistant, *Operating Systems, Computer Architecture, Fault-tolerant Systems, Introduction to Programming, Computer Structure and Distributed Systems*, University Carlos III of Madrid, Spain, from 2005 to 2010

Supervision

- PhD Supervision:
 - Supervision of a PhD student at the University of St Andrews, since January 2023
 - Co-Supervision of PhD students (all successfully completed), at the University of Edinburgh, from 2015 to 2021
 - * Ole Weidner, *A Telemetry Framework for High-Performance Computing Environments*, June 2021
 - * Luca Trani, *Holistic Catalogues for Data-Intensive Science*, September 2019
 - * Alessandro Spinuso, *Active Provenance for Data Intensive Research*, June 2018
- MSc and BSc Supervision:
 - Supervision of 12 BSc and MSc students at the University of St Andrews, since 2021/22
 - Supervision of two BSc students at Heriot-Watt University, since 2021/22
 - Supervision of ten MSc students at the University of Edinburgh from 2012 to 2020
 - Supervision of seven MSc students at the University Carlos III Madrid, 2006 to 2010
- Supervision of three interns within the *St Andrews Research Internship Scheme (StARIS)*, since 2022
- Co-Supervision of two PRACE students at the University of Edinburgh, from 2017 to 2020

Examination

- John Burns, *Automatic Evaluation of Geopolitical Risk*, University of St Andrews, UK 2024
- Giannis Evagorou, *Methods for Efficient Storage, Distribution and Scalable Analysis of Spatial Data*, Imperial College London, UK, 2021
- Pablo Cerro, *Modelling and validation of cloud systems using model driven engineering, metamorphic and mutation testing*, Complutense University of Madrid, Spain, 2020
- Hector Ortega, *Parallel Approaches to Shortest-Path Problems for Multilevel Heterogeneous Computing*, University of Valladolid, Spain, 2015
- Idafen Santana Pérez, *Conservation of Computational Scientific Execution Environments for Workflow-based Experiments Using Ontologies*, Polytechnic University of Madrid, Spain, 2015
- Jordi Villaplana, *Management of Cloud systems applied to eHealth*, University of Lleida, Spain, 2015
- Miquel Orobitg, *High performance computing on biological sequence alignment*, University of Lleida, Spain, 2013
- Laura Prada, *New techniques to model energy-aware I/O architectures based on SSD and hard disk drives*, University Carlos III Madrid, Spain, 2012

Hosting and Training Engagements

- Co-hosted a SICSA Distinguished Visiting Fellow at the University of St Andrews, 2022
- Hosted six HPC-Europa3 visitors at the University St Andrews, at the University of Edinburgh and Heriot-Watt University, from 2019 to 2022
- Coordination and preparation of five *Apache Spark* training courses at the University of Edinburgh and BGS, from 2017 to 2020.
- Coordination and preparation of seven training courses at the University of Edinburgh, from 2014 to 2020: *Introduction to Machine Learning*, *Introduction to Data Science*, and *dispel4py*.

Community Service (selection)

- Member of the Steering Committee, IEEE eScience Conferences, since January 2024
- Invited participant to *Integrating HPC, AI, and Workflows for Scientific Data Analysis*, Dagstuhl, 2023
- Group Leader of the Systems Research Group at the University of St Andrews, since February 2022
- Reviewer for the H2020 *MAESTRO* project funded by the European Commission, February 2022
- Expert for the *Workflow Community Summit* track *Improving workflow systems' interface with legacy and emerging HPC software and hardware stack*, April 2021
- Participation in the *Workflow Community Summit* events, since January 2021
- Expert external evaluator for *EU H2020-ICT* proposals, 2019
- Reviewer for the H2020 *EXCAPE* project, May 2016 to September 2018
- Participation in the *European eXtreme Data and Computing Initiative* (EXDCI) as an expert in the field of *Weather, Climatology and Solid Earth Sciences*, since May 2016

Event organisation

- Chair for the Data Analytics, Visualization, & Storage track at Super Computing (SC) conferences, 2024
- Program Co-Chair for 19th IEEE International Conference on eScience, October 2023
- Co-Chair for the 15th, 16th and 17th editions of the Workshop on Workflows in Support of Large-Scale Science (WORKS) at Super Computing (SC) conferences, from 2020 to 2022
- Co-Chair of the Bridging from Concepts to Data and Computation (BC2DC) workshop co-located with IEEE eScience, 2019
- Co-Organiser of the session Data-Oriented and Self-Optimizing Architectures and Workflows for Earth, Space and Climate Science Applications, The American Geophysical Union (AGU), 2019
- Co-Organiser of the session Data-Driven Platforms and VREs for Complex Earth, Space, and Environmental Science Applications, AGU, 2018

Professional Memberships

- UK Young Academy, since January 2023
- The British Academy Early Career Researcher Network, since September 2022

Programme Committee Membership (selection)

- Programme Committee member for SC, since 2021
- Programme Committee member for WORKS23, 2023
- Programme Committee member, Euro-Par'2023, 2023
- Programme Committee member for 29th IEEE HiPC, 2022
- Programme Committee member for IEEE Cluster, 2021
- Programme Committee member for IEEE IPDPS, 2020 and 2021

Journal Reviewing (selection)

- Reviewer for the Future Generation Computer Systems, 2019 and 2020
- Reviewer for the Journal of Frontiers of Information Technology & Electronic Engineering, 2018

Editorial Duties

- Associate Editor, Royal Society Open Science Journal, since 2023

Speaking Invitations (selection)

1. *Current and Future Trends of AI.*, Demystifying trends, showcasing female talent and making data relevant to all, Bayes Center, University of Edinburgh, 2023
2. *Decoding Democracy: Analysing Public Participation in Constitutional Reform with Deep Learning*, Exploring interdisciplinary research: A hands on approach, University of Glasgow, 2023
3. *Supercomputing and Artificial Intelligence: accelerating each other's pace.*, 17th Spanish Supercomputing Network (RES) Users Conference, BSC, Spain, 2023
4. *frances: A Deep Learning NLP and Text Mining Digital Platform for Analysis of Historical Texts at Scale*, Digital Humanities Research Software Engineering Summer School, Online, 2023
5. *A Hands-On Introduction to Apache Spark and to its Programming Model*, UPM, Spain, 2022

6. *NLS Digital Scholarship Seminars*, National Library of Scotland (NLS), UK, 2022
7. *Advanced data-driven frameworks to develop reproducible data science applications at scale*, Workshop on Re-envisioning Extreme-Scale I/O for Emerging Hybrid HPC Workloads, online, 2021
8. *Improving workflow systems' interface with legacy and emerging HPC software and hardware stack*, Workflows Community Summit, online, 2021
9. *Analysing Digital Historical Textual Data at Scale with Defoe Toolbox*, RLUK, Text and Data Mining Seminar, UK, 2020
10. *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. USC, Los Angeles, US, 2019
11. *Empowering domain experts with DARE, a new cloud-based platform and working environment*. RSEConUK, Birmingham, UK, 2019
12. *Delivering easy-to-use frameworks to empower data-driven research*. UK-France Bilateral International Meeting, Royal Society, Chicheley, UK, 2019

Publications (selection)

- A full list of publications is available at <https://dblp.org/pid/37/4672.html> and <https://scholar.google.com/citations?user=1MA0NAAAAAJ&hl=es>.
- **Journal Articles: Listing 11 out 20**
 1. Pablo C. Cañizares, Alberto Núñez, **Rosa Filgueira**, *et al.* Parallel mutation testing for large scale systems. *Cluster Computing*, 2024
 2. Suzanne R. Black, **Rosa Filgueira**, Lesley McAra, *et al.* *Cultural Analytics in the UK: Current Work and Future Potential*. *Cultural Trends*, 2024
 3. Pablo C. Cañizares , Alberto Núñez, **Rosa Filgueira**, *et al.* *Parallel Mutation Testing for Large Scale Systems*. *Cluster Computing*, 2023
 4. Suzanne R. Black, **Rosa Filgueira**, Lesley McAra, *et al.* *Cultural analytics in the UK: events data potential for the creative and cultural industries*. *Cultural Trends Journal*, 2023
 5. Suzanne R. Black, **Rosa Filgueira**, Lesley McAra, *et al.* *Cultural Analytics in the UK: Current Work and Future Potential*. *Cultural Trends*, 2022 [Under Review]
 6. Liang Liang, **Rosa Filgueira**, *et al.* *Scalable Adaptive Optimizations for Stream-based Workflows in Multi-HPC-Clusters and Cloud Infrastructures*. *Future Generation Computer Systems (FGCS)*, vol. 128, pp. 102-116, 2022
 7. Johanna M. Scheidegger, ..., **Rosa Filgueira**. *Integration of 2D Lateral Groundwater Flow into the Variable Infiltration Capacity (VIC) Model and Effects on Simulated Fluxes for Different Grid Resolutions and Aquifer Diffusivities*. *Water*, vol. 13, March 2021
 8. Sandra Engwell, ... , **Rosa Filgueira**. *Near-real-time volcanic cloud monitoring: insights into global explosive volcanic eruptive activity through analysis of Volcanic Ash Advisories*. *Bulletin of Volcanology*, vol. 83, pp. 1-17, January 2021
 9. Iraklis A. Klampanos, Chrysoula Themeli, **Rosa Filgueira**, *et al.* *DARE Platform: a Developer-Friendly and Self-Optimising Workflows-as-a-Service Framework for eScience on the Cloud*. *The Journal of Open Software*, vol. 5, October 2020
 10. Paolo Marangio, **Rosa Filgueira**, *et al.* *Highlights Automatic detection of Ionospheric Alfvén Resonances in magnetic spectrograms using U-net*. *Computers & Geoscience*, vol. 145, 2020
 11. Rafael Ferreira da Silva, **Rosa Filgueira**, *et al.* *Exploring PID Controllers for Online Resilient Resource Management of Distributed Scientific Workflows*. *Future Generation Computer Systems*, vol. 95, pp. 615-628, 2019
- **Refereed Conference and Workshop Papers: Listing 14 out 46**
 1. Daniel Rotchford, ..., **Rosa Filgueira**, *Laminar 2.0: Serverless Stream Processing with Enhanced Code Search and Recommendations*. *Proceedings of the SC'24 Workshops, WORKS 2024*

2. Honglin Zhang, Leyu Zhang, ..., **Rosa Filgueira**. Multi-Level AI-Driven Analysis of Software Repository Similarities, IEEE 20th International Conference on eScience, Osaka, Japan, September 2024.
3. Lilin Yu, Ash Charlton, ..., **Rosa Filgueira**. Advancing frances: New Heritage Textual Ontology, Enhanced Knowledge Graphs, and Refined Search Capabilities, IEEE 20th International Conference on eScience, Osaka, Japan, September 2024.
4. Zaynab Zahra, Zihao Li, **Rosa Filgueira**. Laminar: A New Serverless Stream-based Framework with Semantic Code Search and Code Completion. Proceedings of the SC'23 Workshops, WORKS 2023.
5. Liang Liang, Heting Zhang, ..., **Rosa Filgueira**. Optimization towards Efficiency and Stateful of dispel4py. Proceedings of the SC'23 Workshops, WORKS 2023.
6. Michael Rovatsos and **Rosa Filgueira**. Building lightweight semantic search engines. IEEE 19th International Conference on eScience, Limassol, Cyprus, October 2023.
7. Zihao Li and **Rosa Filgueira**. Mapping the repository landscape: harnessing similarity with RepoSim and RepoSnipy. IEEE 19th International Conference on eScience, Limassol, Cyprus, October 2023.
8. Christopher Williams and **Rosa Filgueira**. RepoGraph: a novel semantic code exploration tool for python repositories based on knowledge graphs and deep learning. IEEE 19th International Conference on eScience, Limassol, Cyprus, October 2023.
9. Lilin Yu, ..., **Rosa Filgueira**. frances: cloud-based historical text mining with deep learning and parallel processing. IEEE 19th International Conference on eScience, Limassol, Cyprus, October 2023.
10. **Rosa Filgueira**, Daniel Garijo. *Inspect4py: A Knowledge Extraction Framework for Python CodeRepositories*. The 2022 Mining Software Repositories Conference, Online, May 2022.
11. **Rosa Filgueira**, Claire Grover, *et al.* *Extending defoe for the Efficient Analysis of Historical Texts at Scale*. IEEE eScience, Online, September 2021.
12. **Rosa Filgueira**, Beatrice Alex, *et al.* *Geoparsing the historical Gazetteers of Scotland: accurately computing location in mass digitised texts*. 8th Workshop on the Challenges in the Management of Large Corpora, Online, June 2020.
13. **Rosa Filgueira**, Michael Jackson, *et al.* *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. IEEE eScience, Online, September 2019.
14. Malcolm Atkinson, **Rosa Filgueira**, *et al.* *Comprehensible Control for Researchers and Developers facing Data Challenges*. IEEE eScience, San Diego, US, September 2019.