

Dr. Rosa Filgueira Vicente

Lecturer, Computer Science, University of St Andrews
Group Leader, Systems Research Group

January 15, 2023

Email: rf208@st-andrews.ac.uk
<http://www.rosafilgueira.com>

Professional Experience

- Lecturer, School of Computer Science, University of St Andrews, UK, since February 2022
- 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 Research Fellow, School of Informatics, University of Edinburgh, since April 2021
- 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

- Embed Digital Skills in Arts and Humanities Research (UKRI/AHRC), 2022, GBP 236,000 [Under Review]
- 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 3000
- 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, 2022/23
- Course Co-Leader, *Databases*, University of St Andrews, 2021/22 and 2022/23
- 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 Lilin Yu, at the University of St Andrews, from 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 three BSc students at the University of St Andrews, since 2022/23
 - Supervision of four 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 four MSc students at the University of Edinburgh from 2018 to 2020
 - Supervision of six MSc students at the University of Edinburgh from 2012 to 2016
 - Supervision of seven MSc students at the University Carlos III Madrid, 2006 to 2010
- Supervision of an intern within the *St Andrews Research Internship Scheme (StARIS)*, summer 2022.
- Co-Supervision of two PRACE students at the University of Edinburgh, from 2017 to 2020
- Supervision of six Open Science Data Cloud-PIRE students at the University of Edinburgh, 2012 –2015

Examination

- 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)

- 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

- 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, AGU, 2019
- Co-Organiser of the session Data-Driven Platforms and VREs for Complex Earth, Space, and Environmental Science Applications, AGU, 2018

Memberships

- UK Young Academy, September 2023 <https://ukyoungacademy.org/members/rosa-filgueira/>
- The British Academy Early Career Researcher Network, since September 2022

Programme Committee Membership (selection)

- Programme Committee member, Euro-Par'2023, 2023
- Programme Committee member for 29th IEEE HiPC, 2022
- Programme Committee member for SC22, 2021 and 2022
- Programme Committee member for IEEE Cluster, 2021
- Programme Committee member for the IEEE IPDPS, 2020 and 2021
- Programme Committee member for the IEEE eScience conference, 2019
- Programme Committee member for the CCGrid Conference, 2016 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

Speaking Invitations (selection)

1. *frances: An AI-toolbox to discover automatically insights from digital collections*, British Library, 2022
2. *A Hands-On Introduction to Apache Spark and to its Programming Model*, UPM, 2022
3. *NLS Digital Scholarship Seminars*, National Library of Scotland (NLS), 2022
4. *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, 2021
5. *Improving workflow systems' interface with legacy and emerging HPC software and hardware stack*, Workflows Community Summit, online, 2021
6. *Analysing Digital Historical Textual Data at Scale with Defoe Toolbox*, RLUK, Text and Data Mining Seminar, 2020
7. *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. USC, Los Angeles, US, 2019
8. *Empowering domain experts with DARE, a new cloud-based platform and working environment*. RSEConUK, Birmingham, UK, 2019
9. *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=1MAnONAAAAAJ&hl=es>.

- **Journal Articles: Listing 10 out 17**

1. Suzanne R. Black, **Rosa Filgueira**, Lesley McAra, *et al.* *Evaluating Events Data for Cultural Analytics: A case study on the economic and social effects of Covid-19 on the Edinburgh Festivals*. Creative Industries Journal, 2022 [Under Review]
2. Suzanne R. Black, **Rosa Filgueira**, Lesley McAra, *et al.* *Cultural Analytics in the UK: Current Work and Future Potential*. Cultural Trends, 2022 [Under Review]
3. Pablo C. Cañizares , Alberto Núñez, **Rosa Filgueira**, *et al.* *Parallel Mutation Testing for Large Scale Systems* . Future Generation Computer Systems (FGCS), 2022 [Under Review]
4. 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

5. 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
6. 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
7. 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
8. Paolo Marangio, **Rosa Filgueira**, et al. *Highlights Automatic detection of Ionospheric Alfvén Resonances in magnetic spectrograms using U-net*. Computers & Geoscience, vol. 145, 2020
9. 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
10. Luca Trani, **Rosa Filgueira**, Rosa Filgueira et al. *Establishing Core Concepts for Information-Powered Collaborations*. Future Generation Computer Systems, vol. 89, pp. 421-437, 2018

• **Refereed Conference and Workshop Papers: Listing 15 out 37**

1. Emily K. Adams, ..., **Rosa Filgueira**, et al. *Novel proposals for FAIR, automated, recommendable, and robust workflows*. 17th Workshop on Workflows in Support of Large Scale (WORKS'22), Dallas, US, November 2022.
2. **Rosa Filgueira**. *frances: a Deep Learning NLP and Text Mining Web-Tool to Unlock Historical Digital Collections*. IEEE 18th International Conference on eScience, Salt Lake City, US, October 2022.
3. **Rosa Filgueira**, Feras M. Awaysheh, et al. *SparkFlow: Towards high-performance data analytics for spark-based genome analysis*. 22nd IEEE International Symposium on Cluster, Cloud and Internet (CCGrid), Taormina, Sicily, Italy, May 2022.
4. Pablo C. Cañizares, **Rosa Filgueira**, et al. *Validating communication network configurations in cloud and HPC systems using Metamorphic Testing*. 2022 4th International Electronics Communication Conference, June 2022.
5. **Rosa Filgueira**, Daniel Garijo. *Inspect4py: A Knowledge Extraction Framework for Python CodeRepositories*. The 2022 Mining Software Repositories Conference, Online, May 2022.
6. Henry Casanova, ..., **Rosa Filgueira**, et al. *Emerging Frameworks for Advancing Scientific Workflows Research, Development, and Education*. 16th Workshop on Workflows in Support of Large Scale (WORKS'21), Online, November 2021.
7. Rafael Ferreira da Silva, ..., **Rosa Filgueira**, et al. *A community roadmap for scientific workflows research and development*. 16th Workshop on Workflows in Support of Large Scale (WORKS'21), Online, November 2021.
8. **Rosa Filgueira**, Claire Grover, et al. *Extending defoe for the Efficient Analysis of Historical Texts at Scale*. IEEE eScience, Online, September 2021.
9. Liang Liang, **Rosa Filgueira**, et al. *Adaptive Optimizations for Stream-based workflows*. 15th Workshop on Workflows in Support of Large Scale (WORKS'20), Online, November 2020.
10. **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.
11. **Rosa Filgueira**, Michael Jackson, et al. *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. IEEE eScience, Online, September 2019.
12. Malcolm Atkinson, **Rosa Filgueira**, et al. *Comprehensible Control for Researchers and Developers facing Data Challenges*. IEEE eScience, San Diego, US, September 2019.
13. **Rosa Filgueira**. *DARE Platform: Enabling Easy Data-Intensive Workflow Composition and Deployment*. Research Objects Workshop, San Diego, US September 2019.
14. Vyrion Christodoulou, **Rosa Filgueira**, et al. *Automatic classification of Aurora-related tweets using machine learning methods*, 2nd International Conference on Geoinformatics and Data Analysis (ICGDA), Prague, Czechia, March 2019
15. **Rosa Filgueira**, Rafael Ferreira da Silva, et al. *IoT-Hub : New IoT Data-Platform for Virtual Research Environments*. 10th International Workshop on Science Gateways, Edinburgh, UK, June 2018