

Dr. Rosa Filgueira Vicente

Research Fellow
EPCC
University of Edinburgh , Edinburgh, UK

April 23, 2020

Email: rosa.filgueira@ed.ac.uk
<http://www.rosafilgueira.com>
<https://github.com/rosafilgueira>

Professional Experience

- **Current position:** Research Fellow, University of Edinburgh, UK, since June 2019
- Data Architect, University of Edinburgh, UK, since March 2018 - May 2019
- Senior Data Scientist, British Geological Survey, Edinburgh, UK, since October 2016 - February 2018
- Senior Research Associate, School of Informatics, University of Edinburgh, June 2011 - October 2016
- Research and Teaching Assistant, University Carlos III Madrid, Spain, September 2005 - September 2010
- Software Developer, Soluziona, Enterprise Consultancy, Spain, June 2004 - September 2004

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 Experience

- Research Visitor, ISI, University of Southern California, US, October 2017 - December 2017
- Honorary Research Fellow, School of Informatics, University of Edinburgh, UK, since January 2017
- Research Visitor, ISI, University of Southern California, US, March 2016 - May 2016
- Consultancy Services for Accelogic company, US, August 2014 - April 2016
- Research Visitor, AIST, Informatics Technology Research Institute, Japan, July 2013 - July 2013
- Research Visitor, EPCC, University of Edinburgh, UK, February 2011 - October 2011

Research Expertise

- *City Region Deal Text and Data Mining*: funded by the Edinburgh and South East Scotland City Region Deal 2019-2020 – role: improving our distributed, scalable and easy-to-use text data mining tools for enabling digital humanities research using the extensive digital collection of the National Library of Scotland
- *Graph-Based Data Federation for Healthcare Data Science* : funded by the UK Research and Innovations Industrial Strategy Challenge Fund, 2019 – role: prototyping a cloud-base architecture for the knowledge driven phenotype computation framework developed within the project
- *A Shared Data Repository for Genomics Data In Scotland*: This was a collaboration between the NES Digital Service (for which I was seconded) and the Scottish Clinical Genetics Laboratories, 2019 – role: exploring opportunities for creating shared data repository for clinical genetics laboratories in Scotland, and scoping the architecture necessary for such repository

- *Living with Machines*: funded by UKRI's Strategic Priorities, 2018-2023, GBP 9.2 million – role: developing new scalable tools for complex text mining analysis by using distributed systems and NLP processing techniques
- *Delivering Agile Research Excellence on European e-Infrastructures (DARE)*: funded by EU H2020, 2018-2020, EUR 2.9 million – role: empowering domain experts to invent and improve their methods and models by providing a new cloud-based platform and a working environment
- *Bioexcel*: funded by EU H2020, 2018-2019, EUR 4.7 million – role: benchmarking two computing intensive Spark-based tools to measure their performance under different settings and computational infrastructure
- *GCRF-REAR*: funded by NERC Global Challenges grant, 2016 – 2017, GBP 198,904 – role: helping to understand how mobile technology can be used to collect aftershock hazard information
- *ENVRIplus*: funded by EU H2020, 2015 – 2016, EUR 7 million – role: capturing the requirements and delivering common data functionalities for 20 pan-European marine, atmospheric, biological and solid-Earth ESFRI endorsed Research Infrastructures
- *VERCE*: funded by EU FP7, 2014 – 2015, EUR 4.6 million – role: a) implementing stream-based data-intensive seismic applications with `dispel4py` and b) optimising them at run-time by applying different adaptive techniques
- *EFFORT*: funded by UK NERC, 2011 – 2014, GBP 526,969 – role: to design and implement a) a new science gateway for Rock Physicists and Volcanologists and b) a new Python toolbox (called *Varpy*) to analyse volcanic and rock physics data

Track Record of Research Funding

- Two *Postdoctoral and Early Career Researcher Exchanges (PECE)*, SICSA grants, 2017/2015, GBP 7,500 and GBP 13,580
- Co-I, *OII: Observation Imaging and Interpretation*, UoE grant, 2016 – 2017, GBP 3000
- Co-I, *SKA-Link: 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*, New Orleans, 2014, GBP 2,500
- Two *HPC-Europa2* PRACE grants, 2011, GBP 6,000

Awards and Scholarships

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

Teaching, Supervision, Hosting, Training and Examination Experience

- Hosting three HPC-Europa3 visitors at University of Edinburgh, from 2019
- Co-Supervision of three PhD students at University of Edinburgh, since 2015
- Co-Supervision of two PRACE Summer of HPC (SoHPC) students at University of Edinburgh, from 2017
- Supervision of 13 MSc students at University of Edinburgh and University Carlos III, from 2006
- Supervision of six Open Science Data Cloud (OSDC)-PIRE students, at University of Edinburgh, since 2012 to 2015
- Teaching Assistant at University Carlos III of Madrid, from 2005 to 2010
- Coordination and preparation of ten training courses, since 2014
- Examined six external PhDs, since 2012

Community Service

- Reviewer for the *MAESTRO* project funded by the European Commission under the *Future and Emerging Technologies initiatives* in H2020, since February 2020
- Expert external evaluator for *EU H2020-ICT* proposals, 2019
- Lead author of *The data-driven geoscience platform* research line for the BGS Informatics Science Directorate, February 2017
- Participation in the *SC Mentor-Protege Program* as a mentor, Super Computing, Salt Lake City, US, November 2016
- Reviewer for the *EXCAPE* project funded by the European Commission under the *Future and Emerging Technologies initiatives* in H2020, from 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

Workshops Organizer, Programme Committee Memberships and Journal Reviewing

1. Co-chair for the 15th Workshop on Workflows in Support of Large-Scale Science (WORKS20) at SC2020, 2020
2. Invited speaker's chair for Women in HPC at ISC2020, 2020
3. Programme committee member for the 12th International Workshop on Science Gateways, 2020.
4. Co-organizer of the session Data-Oriented and Self-Optimizing Architectures and Workflows for Earth, Space and Climate Science Applications, AGU, 2019
5. Co-chair of the Bridging from Concepts to Data and Computation (BC2DC) workshop collocated with the IEEE eScience, 2019
6. Programme Committee member for the SC19 edition of the Women in HPC workshop, 2019
7. Programme committee member for the IEEE eScience conference, 2019
8. Programme committee member for the 11th International Workshop on Science Gateways, 2019

9. Co-organizer of the session Data-Driven Platforms and VREs for Complex Earth, Space, and Environmental Science Applications, AGU, 2018
10. Programme committee member for the 10th International Workshop on Science Gateways, 2018
11. Reviewer for the Journal of Frontiers of Information Technology & Electronic Engineering, 2017
12. Poster Committee Member for the IEEE Conference on Cluster Computer, 2016
13. Reviewer for the IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid 2016) Conference, 2016
14. Programme committee member for the TPD 2016 on Reproducible Open Science Workshop, 2016
15. Programme committee member for the first International Workshop on Interoperable infrastructures for Interdisciplinary Big Data Sciences (IT4RIs 15), 2015
16. Programme committee member for the ACM Conference on Service-Oriented Architectures and Programming (SOAP 2015), 2015
17. Programme committee member for the IEEE Conference on International Symposium on Parallel and Distributed Processing with Applications (ISPA 2013), 2012 and 2013
18. Reviewer for the Journal of Systems and Software, 2013
19. Reviewer for the Journal of Simulation Modelling Practice and Theory, 2013
20. Reviewer for the IEEE Symposium on Computers and Communications (ISCC 2010) Conference, 2010

Selected Publications

• Journal Articles

1. Paolo Marangio, Vyron Christodoulou, **Rosa Filgueira**, *et al.* *Highlights Automatic detection of Ionospheric Alfvén Resonances in magnetic spectrograms using U-net.* (Submitted to Computers Geoscience, 2020)
2. 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
3. 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
4. Luca Trani, **Rosa Filgueira**, *et al.* *WFCatalog: a catalogue for seismological waveform data.* Computers and Geosciences, vol. 106, pp. 101-108, 2017
5. Rafael Ferreira da Silva, **Rosa Filgueira**, *et al.* *A characterization of workflow management systems for extreme-scale applications.* Future Generation Computer Systems, vol. 75, pp. 228-238, 2017
6. **Rosa Filgueira**, Amrey Krause, *et al.* *dispe14py: A Python Framework for Data-Intensive Scientific Computing.* Special issue on Data-Intensive High Performance Computing. International Journal of High Performance Computing Applications, vol. 31, no. 4, pp. 316-334, 2016
7. Jordi Vilaplana, **Rosa Filgueira**, *et al.* *The Cloud Paradigm applied to e-Health.* BMC Medical Informatics and Decision Making, vol. 35, no. 13, pp. 1-10, 2013
8. Alberto Nuñez, **Rosa Filgueira**, *et al.* *SIMCAN: A Flexible, Scalable and Expandable Simulation Platform for Modeling and Simulating HPC Architectures and Applications.* Simulation Modelling Practice and Theory, vol. 20, no. 1, pp. 12-32, 2012

9. **Rosa Filgueira**, David E. Singh, *et al.* *Adaptive CoMPI: Enhancing MPI-based Applications Performance and Scalability by Using Adaptive Compression*, The International Journal of High Performance Computing Applications, vol. 25, no. 1, pp. 93-114, 2012
10. **Rosa Filgueira**, Jesús Carretero, *et al.* *Dynamic-CoMPI: Dynamic Optimization Techniques for MPI Parallel Applications*. The Journal of Super Computing, vol. 59, no. 1, pp. 361-391, 2012

- **Conference and Workshop Papers**

1. **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, 2020.
2. **Rosa Filgueira**, Michael Jackson, *et al.* *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. IEEE eScience 2019.
3. Malcolm Atkinson, **Rosa Filgueira**, *et al.* . *Comprehensible Control for Researchers and Developers facing Data Challenges*. IEEE eScience 2019.
4. **Rosa Filgueira**. *DARE Platform: Enabling Easy Data-Intensive Workflow Composition and Deployment*. Research Objects (RO2019) during IEEE eScience 2019.
5. Vyron 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 15-17, 2019
6. **Rosa Filgueira**, Rafael Ferreira da Silva, *et al.* *IoT-Hub : New IoT Data-Platform for Virtual Research Environments*. 10th International Workshop on Science Gateways (IWSG 2018), 13-15 June 2018
7. Rafael Ferreira da Silva, **Rosa Filgueira**, *et al.* *Using Simple PID Controllers to Prevent and Mitigate Faults in Scientific Workflows*. WORKS'16 Workshop, SC16, November 14, 2016
8. **Rosa Filgueira**, Rafael Ferreira da Silva, *et al.* *Asterism: Pegasus and dispe14py Hybrid Workflows for Data-intensive Science*. DataCloud'16 Workshop, SC16 Computing, , November 13, 2016
9. Ole Weidner, **Rosa Filgueira**, *et al.* *Rethinking High Performance Computing Platforms: Challenges, Opportunities and Recommendations*. DIDC16 Workshop, Kyoto, Japan, June 1, 2016
10. Rafael Ferreira da Silva, **Rosa Filgueira**, *et al.* . *Automating Environmental Computing Applications with Scientific Workflows*. ECW Workshop, 12th IEEE eScience , Baltimore, US, October 23, 2016
11. **Rosa Filgueira**, Amrey Krause, *et al.* *dispe14py: A User-friendly Framework for Describing eScience Applications*. 11th IEEE eScience, Munich, Germany, September 1-4, 2015
12. Malcolm Atkinson, **Rosa Filgueira**, *et al.* *VERCE Delivers a Productive e-Science Environment for Seismology Research*. 11th IEEE eScience, Munich, Germany, September 1-4, 2015
13. **Rosa Filgueira**, Malcolm Atkinson, *et al.* *VarPy: A Python Library for Volcanology and Rock Physics Data Analysis*. PyHPC'14 Worksop, Super Computing, New Orleans, US, November 16-21, 2014
14. Sandra Gesing, **Rosa Filgueira**, *et al.* *Workflows in a Dashboard: A New Generation of Usability*. WORKS'14 Worksop, Super Computing, New Orleans, US, November 16-21, 2014
15. **Rosa Filgueira**, Malcolm Atkinson, *et al.* *eScience Gateway Stimulating Collaboration in Rock Physics and Volcanology*. 10th IEEE eScience, Guarujá, Brazil, October 20-24, 2014
16. **Rosa Filgueira**, Malcolm P. Atkinson, *et al.* *Applying Selectively Parallel I/O Compression to Parallel Storage Systems*. European Conference on Parallel and Distributed Computing (Euro-Par), Porto, Portugal, August 27-29, 2014
17. **Rosa Filgueira**, Iraklis Klampanos, *et al.* *FAST: Flexible Automated Synchronization transfer*. In proceedings of High-Performance Parallel and Distributed Computing. ACM International Workshop on Data Intensive Distributed Computing (DICT), Vancouver, Canada, June 23-27, 2014

18. Yusuke Tanimura, **Rosa Filgueira**, *et al.* *Two-Phase Collective I/O based on Advanced Reservations to Obtain Performance Guarantees from Shared Storage Systems*. Workshop on Interfaces and Architectures for Scientific Data Storage (IASDS), Indianapolis, IN, US, September 1-9, 2013
19. Yusuke Tanimura, **Rosa Filgueira**, *et al.* *MPI Collective I/O based on Advanced Reservations to Obtain Performance Guarantees from Shared Storage Systems*. IEEE International Conference on Cluster Computing (CLUSTER), Indianapolis, IN, US, September 23-2, 2013
20. Alberto Nunez, **Rosa Filgueira**, *et al.* *SANComSim: A Scalable, Adaptive and Non-intrusive Framework to Optimize Performance in Computational Science Applications*. International Conference on Computational Science (ICCS), Barcelona, Spain, June 5-7, 2013
21. **Rosa Filgueira**, Malcolm Atkinson, *et al.* *An Adaptive, Scalable, and Portable Technique for Speeding Up MPI-Based Applications*. International European Conference on Parallel and Distributed Computing (Euro-Par), Rhodes Island, Greece, August 27-31, 2012
22. Gabriel G. Castañe, **Rosa Filgueira**, *et al.* *Dimensioning Scientific Computing Systems to Improve Performance of Map-Reduce based Applications*. International Conference on Computational Science (ICCS), pp.: 226-235, Omaha, Nebraska, June 4-6, 2012
23. **Rosa Filgueira**, David E. Singh, *et al.* *CoMPI: Enhancing MPI Based Applications Performance and Scalability Using Run-Time Compression*. EURO PVM/MPI, Espoo, Finland, Sept. 7-10, 2009
24. **Rosa Filgueira**, David E. Singh, *et al.* *Exploiting Data Compression in Collective I/O Techniques*. IEEE International Conference on Cluster Computing (CLUSTER), Tsukuba, Japan, Sept. 29 - Oct. 1, 2008
25. **Rosa Filgueira**, David E. Singh, *et al.* *Data Locality Aware Strategy for Two-Phase Collective I/O*. In Proceedings of High Performance Computing for Computational Science (VECPAR), Toulouse, France, June 24-27, 2008
26. **Rosa Filgueira**, David E. Singh, *et al.* *Optimization and Evaluation of Parallel I/O in BIPS3D Parallel Irregular Application*. 21st IEEE International Parallel and Distributed Processing Symposium (IPDPS), Long Beach, California US, March 26-30, 2007

Books and Book chapters

1. **Rosa Filgueira**, Malcolm Atkinson, *et al.* Chapter: *dispe14py: Agility and Scalability for Data-Intensive HPC*. Springer eBook on Conquering Big Data Using High Performance Computing, 2016
2. **Rosa Filgueira**, David E. Singh, *et al.* Book: *Optimization Techniques for Parallel Applications on Clusters*. Lambert Academic Publishing, 978-3-8433-70, 2010