Job Description – JD250 v1.1

Role Title	Senior Research Software Developer – Data Science Specialist
Department/section	Research Software Development Group Research IT Services
Base location	UCL Bloomsbury Campus, London
Grade	8
JDO Reference	Technical Specialist / Technical Operations NJ11-70 (grading ref NJ14-233)
Reporting to	Head of Research Software Development
Direct reports	0-3 Data Scientists and Research Software Developers
Works closely with	Team and group colleagues Head of Research Software Development Scientific and Executive Directors, UCL Centre for Data Science Director of Research IT UCL academic leaders in computational and data intensive research RITS Research Computing Group RITS Research Data Services Group
Date updated	30 May 2017

Working context

The Information Services Division supports and enhances learning, teaching, research and a range of administrative processes across UCL.
ISD has a recurrent budget of £30m pa and a capital budget of £13m-£20m pa and provides technology related services to over 38,000 students and 11,000 staff at UCL and associated institutions.

Facilities and services provided by ISD include all core data centres, email, desktop, printing, networking and software/hardware purchasing. The Division also supports research computing platforms, e-learning and IT security across UCL. The key administrative systems including HR, Finance and Student Information are provided by ISD, as are website development, support, and creative design.

With some 380 staff, the Division is structured into the **Director's Office, Service Delivery, IT Change and Project Delivery Services**, **Service Strategy and Improvement, Learning Technology and Media Services, Research IT Services, IT for SLMS, IT for IOE and Humanities**, **IT for Professional Services.**

Research IT Services and teams in **Learning and Technology and Media Services** focus on the two key missions of the university - research and teaching & learning. They aim to ensure services in these areas are world leading.

Director's Office provides architecture services, strategic management, finance services, staffing support and communications.

Service Delivery provides infrastructure and other technology services such as data centres, networks, server and virtualisation support plus the central IT service desk. In addition, **Service Delivery** also provides application support and enhancement services including database administration.

IT Change and Project Delivery Services (ITCPD) owns the project delivery methodology and is responsible for delivering new and enhanced technology enabled change and services which support UCL's strategic goals.

The **Service Strategy & Improvement Department** is responsible for ensuring that ISD continuously improves and promotes is services and processes, including ITIL. The department coordinates the development of the ISD service strategy and ensures it is aligned across the service portfolio.

IT for SLMS provides local support for users in the School of Life and Medical Sciences (SLMS).

IT for IOE and Humanities provides local support to users in the Institute of Education (IOE) and to departments in the Faculties of Arts & Humanities and Social & Historical Sciences.

IT for Professional Services oversees the provision of technology enabled services to the central divisions such as Finance and HR.

Job purpose

- The UCL research software development group collaborates with UCL researchers using computational methods, to help deliver reliable, reproducible and efficient cutting edge compute- and data-intensive research.
- This post calls for a senior specialist research software engineer to focus on data science and machine learning. The appointee will take the lead in the delivery of expertise in coding for data analysis, computational statistics, machine learning, and big data engineering to the UCL research community.

- They will collaborate with research colleagues from across UCL to maximize insight from the wealth of data generated across all disciplines, applying expertise in a wide range of statistical, mathematical and computational methods.
- They will rapidly assimilate research context through publications and conversation with research groups, and understand the computational algorithms, requirements and interfaces needed within the research effort.
- They will express UCL research analyses as reproducible computer software, resulting in robust software tools providing a sustained impact on research.
- Working with colleagues in the research computing group, and deploying expertise in cutting edge software platforms and tools, they will ensure that big-data analyses can be efficiently and reliably deployed to UCL's High Performance Computing (HPC) platforms, and other external facilities to which UCL researchers have access, including national supercomputing facilities such as JADE and ARCHER and cloud services. They will maintain and support the state-of-the-art infrastructure and services needed for effective data science, including ensuring that UCL's shared tools in this space are up-to-date.
- They will review and revise the work of junior colleagues, amending as necessary to ensure consistency and quality.
- They will monitor the team's processes ensuring that internal and external processes are followed, including coordination of code review and issue tracking.
- They will contribute to the design of the team's ways of working, ensuring continuous improvement and adoption of appropriate techniques, technologies and tools.
- They will manage external funding opportunities from opportunity through to successful funding, liaising with funding bodies, UCL Professional Services staff and Principal Investigators, preparing proposal costings and contributing to bid documents.
- They will coordinate the delivery of collaborative data science projects, allocating team resources efficiently to ensure timely delivery of research outcomes.
- They will ensure Data Science services are effectively embedded in Divisional processes and may act as Service Operational Manager, preparing documentation for Divisional projects, and liaising with Divisional colleagues.
- They will represent RITS at University, national and international events, enhancing the recognition and reputation of the Department and UCL, and creating opportunities for collaboration.
- Working with the Head of Group, and academic leaders in the area, they will contribute to strategic leadership, and may deputise for these colleagues as necessary on strategic and governance bodies.
- They will train UCL research staff and students in the effective use of code for data science, through designing and leading hands-on training in software development for computational statistics, machine learning tools, and data science best practice.
- They will continually study new and existing technologies, tools and ideas in software for data science, maintaining expertise through engagement with UCL's world-class research community in computational statistics and machine learning.

- They will build and maintain relationships within the research and e-Infrastructure communities in UCL and beyond, seeking opportunities to contribute to research, and to generate and prepare opportunities for new research projects and funding.
- They will contribute to the wider ecosystem of support for compute- and data-intensive research in UCL, working with departmental and group colleagues, departmental IT staff and other ISD colleagues to help build integrated systems and services which meet the needs of researchers.
- Working with colleagues in research data services and the library, they will support release and dissemination of UCL research data, through open source, scholarly, and commercial channels, ensuring data sets and results are easy to re-use.
- They will explain and document the software and data they help create, contributing to research publications and documentation.

Main accountabilities and tasks

1) The table below summarises the main tasks undertaken for this role.

Accountabilities	Key tasks	% time spent
Software Engineering for Data Science	• Collaborate with research colleagues from across UCL to construct, improve, and maintain analyses of research data. • Rapidly assimilate research context and understand the computational algorithms, requirements and interfaces involved in a research project. • Design and construct high-quality software for research colleagues which will result in a sustained impact on their research programme(s). • Ensure analyses perform efficiently, by making effective use of UCL-owned and other High Performance Computing platforms. • Assist colleagues in analysis and problem-solving tasks, sharing knowledge and expertise with team members. • Manage and mentor other team members ensuring the code they produce meets rigorous quality standards • Review and revise code and documentation produced by team members, amending as necessary to ensure consistency and quality	25
Leadership and Management	• Source and manage external funding opportunities from opportunity generation to successful funding, liaising with	10

	funding bodies, UCL professional services staff and principal investigators. • Prepare proposal costings and contribute to bid documents. • Coordinate and lead the delivery of collaborative research programming projects, organising meetings, reviews and preparing interim reports. • Provide supervision and leadership and/or line management to up three Software Engineers and Data Scientists, which may include: performance management and coaching, objective setting, identification and follow-up of training and development needs, and appraisal. • Monitor processes, metrics, and dashboards, ensuring that internal and external processes are followed. • Coordination of code review and issue tracking. • Contribute to the design of the team's ways of working, ensuring continuous improvement and adoption of appropriate techniques, technologies and tools. • Contribute to strategic leadership, working with senior colleagues to develop new services. • Deputise as necessary on strategic and governance bodies.	
Provide training relating to data science	• Train UCL research staff and students in coding for data analysis for research. • Develop and design teaching and training in software for data science, suitable for a range of audiences with a very variable degree of computational and statistical experience. • Advise researchers on data science software practices, techniques, design, and architecture.	15
Maintain and enhance expertise	• Maintain expertise in many areas of data science and software through independent study, training courses, and engagement with UCL's statistical and computer science research communities. • Practice and enhance skills by contributing to relevant open source projects. • Maintain and develop expertise in technical leadership and best practices in data science.	15

Publish, document and support use of research outputs	• Support release and dissemination of UCL research, through open source, scholarly, and commercial channels. • Explain and document analyses created with the team's involvement, contributing to research papers published in the academic literature, project reports and case studies, and code documentation and manuals. • Provide online and face-to-face user support for software and data they create or maintain to both UCL and external users.	10
Engage with and contribute to wider research community	• Build and maintain relationships within the research and e-Infrastructure communities in UCL and beyond, actively seeking opportunities for collaboration with researchers including programming and grant preparation. • Contribute to community activities such as seminars and networking events. • Seek out and develop opportunities for new research projects and funding. • Attend conferences and community events in a variety of software engineering/research computing fields in the UK and abroad. • Represent RITS at University, national and international events, enhancing the recognition and reputation of the Department and the University, and creating opportunities for collaboration. • Contribute ideas, experience and thinking to technical working groups in and beyond UCL.	15
Maintain and support infrastructure and services	• Maintain and support the state-of-the-art infrastructure and services needed for effective big data analyses. • Work with the research computing platforms team to help create systems and servers used to deliver data science infrastructure. • Author and maintain documentation relating to data science infrastructure services. • May act as Service Operational Manager for one or more data science services. • Contribute to the wider ecosystem of support for computational research in UCL, working with departmental IT staff and ISD colleagues to help build and maintain integrated systems and services that meet the needs of researchers.	10

	• Contribute to wider departmental and divisional activities through discussions and meetings. • Prepare and maintain project management documentation for divisional projects. • Liaise effectively with divisional colleagues on security, platforms, project management and design matters.	

Person Specification

Essential

1. PhD degree in a field making significant use of quantitative data analysis, or equivalent professional experience.
2. Experience analysing research data, with evidence such as authorship of research publications or other analyses.
3. Significant experience of using and developing computer software for reproducible data analysis.
4. Ability to rapidly acquire fluent knowledge of new statistical and computational methods, programming languages, libraries and platforms.
5. Experience using Unix-based operating systems and Unix system tools and utilities.
6. Experience with the advanced use of high-level dynamic languages for numerically-intensive research, preferably Python, Julia or R.
7. Experience of the application of computational statistics, Bayesian inference and machine learning to the analysis of data.
8. Excellent written and verbal communication skills including the ability to effectively present complex or technical information to a range of audiences.
9. Desire to keep up-to-date and learn about new developments in computational research.
10. Knowledge of and commitment to the use of software development best practice to make data analyses reproducible, including issue tracking, testing, documentation, version control, and continuous integration.
11. Experience with programming libraries for computational statistics and machine learning, such as Python's statsmodels and scikit-learn.
12. Proven ability to forge effective professional relationships at all levels.
13. Proven ability to manage multiple concurrent tasks and activities, both alone and in a team, working to deadlines and prioritising as appropriate.
14. Experience mentoring and leading other researchers (formal line management experience is not essential, but such candidates should be able to show they have effectively guided the work of more junior colleagues).

Commented [CJ1]: These could perhaps be merged with 9 & 10.

Desirable

1. Expertise with platforms and file systems for efficient data science, such as Spark and HDFS.
2. Experience with advanced high-performance data science libraries such as TensorFlow.

Commented [CJ2]: Is it worth adding a sentence at the start of this section along the lines of "We are not expecting applicants to have experience in all areas, but try to ensure a diverse range of backgrounds among the wider team." To make the long list less scary.

3. Experience with use of public cloud platforms for data science, including higher level as-a-service layers, such as Amazon or Azure.
4. Experience of using literate programming tools for the presentation of research analyses, such as the Jupyter notebook.
5. Experience of the deployment of analysis tools and research outputs to the web, including experience programming with web frameworks.
6. Experience using parallel programming, GPUs and accelerators for efficient data analysis.
7. Advanced software process and project management expertise, with knowledge of multiple process frameworks and theories, from a variety of cultures (SCRUM, XP, Lean, PRINCE II, ITIL), with an understanding of commonalities and differences.
8. Knowledge of and interest in leadership, demonstrable through awareness of appropriate books, websites or blogs.
9. Understanding of commercial aspects of the data science industry, with experience of organisations of multiple maturity levels and sizes.
10. Understanding of research and academic funding sources, with authorship on one or more research grants.
11. Demonstrable leadership in one or more software or data science communities, such as through membership of standards bodies, organisation of community events, or a lead maintainer role for open source tools or packages.
12. Experience designing and/or delivering teaching or training courses.
13. Experience of open source software practices, with at least one accepted contribution to an open source project.

Commented [CJ3]: Partially covered by 13.

Special working conditions

None

Conditions of Service

The normal hours of work are 36.5 hours per week. However, this is a senior post and flexibility will be expected in response to varying workload. Reasonable notice will be given and where properly authorised such work will be recompensed as stated in the terms and conditions. The annual leave entitlement is 27 days per year, plus 6 College closure days, plus public holidays.

Additional Information

- The job description reflects the present requirements of the post, and as duties and responsibilities change/develop, the job description will be reviewed and be subject to amendment in consultation with the post holder. The post holder will carry out any other duties as are within the scope, spirit and purpose of the job as requested by the line manager or Head of Department/Division.

- The post holder will actively follow UCL policies including Equal Opportunities policies and be expected to give consideration within their role as to how they can actively advance equality of opportunity and good relations between people who share a relevant protected characteristic and people who do not share it.
- The post holder will maintain an awareness and observation of Fire and Health and Safety Regulations.
- The post holder must ensure organisational compliance, and conformance with the Data Protection Principles. All data, whether stored electronically or by other means must be processed in accordance with the Data Protection Act 1998.
- The post holder will attend staff meetings and training as required.
- The post holder will maintain their knowledge and skills through professional activities with agreement from their line manager and readership of relevant publications.