HR191	**POSITION DESCRIPTION**	UNIVERSITY OF CAPE TOWN IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

NOTES

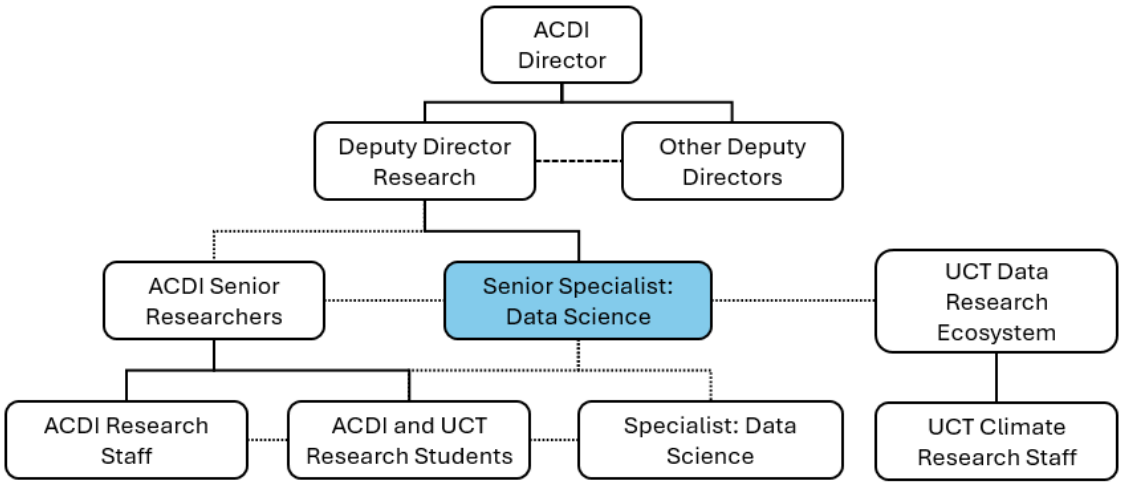
- Forms must be downloaded from the UCT website: https://forms.uct.ac.za/forms.htm
- This form serves as a template for the writing of position descriptions.
- A copy of this form is kept by the line manager and the position holder.

POSITION DETAILS

Position title	Senior Data Scientist		
Job title (HR Business Partner to provide)	Senior Research Support Specialist (PC11)		
Position grade (if known)	PC 11	Date last graded (if known)	March 2025
Academic faculty / PASS department	Science		
Academic department / PASS unit			
Division / section	African Climate and Development Initiative		
Date of compilation	20 June 2025		

ORGANOGRAM

(Adjust as necessary. Include line manager, line manager's manager, all subordinates and colleagues. Include position grades)



PURPOSE

The main purpose of this position is to provide data science leadership and support/advice to a range of research staff, projects and other activities within ACDI and its research groups and centres, such as ASCEND, the PiNC lab and Climate Risk lab, working largely with climate, environmental/ecological, and socioeconomic data.

Data science leadership responsibilities will include coordination and management of data science activities at ACDI, coordination of data science training, mentoring of junior data scientists, data infrastructure design, development, and governance, being the main point of contact with UCT's e-research capabilities. Data science technical and support responsibilities will include advanced data analysis, application of existing software and codes, development of new software and tools, including machine learning, and provision of data science consultations, support and advice to specialist researchers and research teams. The position will also require leading the data science inputs for grant proposals developed within ACDI. Initial areas of focus of the position will be data science at the intersection of climate science and climate impacts modelling (health, hydrology, ecology).

CONTENT

Key performance areas		% of time spent	Inputs (Responsibilities / activities / processes/ methods used)	Outputs (Expected results)
1	Leadership and Management	20	• Participate in ACDI management activities, with a specific focus as lead on data science, data management and data governance. • Lead or co-lead the ACDI data science management/working group. • Act as lead for research data governance within specific ACDI projects. • Ensure effective use of data science computational and human resources at ACDI. • Mentor, and where necessary, line-manage, junior data scientists. • Lead on the provision of training and capacity building in data science when necessary.	• Data science strategy for ACDI is up to date. • Data governance and management in relevant ADCI projects is fully legally compliant and meets all ethical standards.
2	Climate, environmental, socio-economic data management	20	• Set up, manage and apply standard procedures for data processing and management. • Day-to-day research project data and data portal management activities. • Source and update datasets, converting these and merging with other datasets. • Data quality control and liaising with researchers on quality and errors. • Process and manage data for analyses from local to global scales, in line with complex analytic questions and hypotheses. • Provide data management support to researchers and students, including developing data management plans. • Ensure all data is collected, managed and stored according to international data standards, including open science and reproducible research protocols.	• Up to date climate and related data is available to researchers and partners. • Up to date data-related web portals, web pages and applications. • Data quality and control procedures are followed. • All technology is up to date, supporting the work of the researchers.

3	Climate and/or Environmental Data Science Research Support	50	• Contribute to conceptualization of research questions / projects, especially experimental and data analysis designs. • Conduct processing, transformation, integration, analysis, visualisation and interpretation of data using appropriate programming languages and tools, including delivery via web applications where needed. • Work with existing code and develop new code to improve efficiency of processing and analysis of spatial and temporal big data, including using parallel computing, using best practices in code development, including collaborative code development and reproducible science. • Work closely with ACDI researchers and teams, including those in PiNC and CR labs, and visiting teams at ASCEND, to develop and improve workflows for analysis of heterogeneous climate, social and environmental spatial and temporal datasets. • Create static and interactive visualisations and use state-of-the-art tools to support data dissemination and communications for different target audiences including leading or supporting development of dissemination platforms or user interfaces, such as web applications to showcase data and research findings. • Advise researchers of data science tools and approaches available and provide input on the best fit for the research task. • Provide training and capacity building in data science, when necessary, especially for early career researchers and research teams. • Contribute to project reports and papers as required. • Assist with data management sections of proposals.	• Sound advice and assistance on all data-related issues is provided. • Data science components of all proposals and budgets are completed to high quality • Early career research staff are supported and enabled to undertake advanced data analyses • Effective visualisations are developed for various research and communication purposes • Efficient and well documented climate data analysis code is developed and maintained • Reproducible research workflows • Effective collaborative coding projects are developed

4	Stakeholder relationships	10%	• Ensure ACDI is represented by data science staff on key UCT data science working groups and committees. • Maintain strong working relationships with researchers at ACDI and affiliated groups. • Maintain good working relationships with UCT's AI Initiative, ICTS, Library staff, Research Support Hub (including UCT eResearch Centre), and keep ACDI informed in developments in the wider UCT data science ecosystem.	• Collaboration between partners continues to function at a high level • Staff at ACDI are well informed of eResearch and related developments at UCT.

MINIMUM REQUIREMENTS

Minimum qualifications	• NQF 9 (Masters) degree or equivalent postgraduate diploma qualification in an appropriate field: data science, information systems, computer science, climate science, quantitative ecology/environmental science, statistics, bioinformatics or a closely related field.
Minimum experience (type and years)	**Required** • At least five years of post-NQF9 qualification experience in data science. • Data management, data processing, coding for advanced analysis, modelling and/or machine learning (3+ years) • Coding in both R and Python (5+ years total, 2+ years in each) • Working with large spatial and/or temporal datasets (3+ years) • Working with climate and/or environmental data, as well as (optionally) socio-economic data (3+ years). • Use of online code collaboration platforms such as GitHub (4+ years) and management of collaborative projects on these platforms (2+ years) • Data quality management and use of data quality management tools (3+years). • Provision of advanced data science service to support the research enterprise (such as developing and improving data processing and analysis workflows, working with specialist scientists to identify and implement data science analyses for specific research questions) (3+ years). • Working in cloud computing and/or high-performance computing environments/contexts (3+ years) • Teaching or training of data science and/or other quantitative skills (1+ years) **Advantageous** • Tertiary level training (e.g. BSc) and/or several years work experience in one or more relevant subjects, such as climate or atmospheric science, oceanography, ecology, hydrology, health sciences, environmental economics, conservation science. • Experience in installing and running climate, hydrological and other process-based models. • Leading or support on institutional research data governance and management
Skills	• Capability to manage or coordinate a data science team and lead on data science governance and management in a research setting. • Ability to provide inputs on data science into institutional strategic planning and research proposals • Teaching/training students and/or researchers in data science • Excellent data science skills including programming in languages and software environments for advanced analysis, modelling and/or machine learning. • Highly proficient in both R and Python to support diverse analysis needs of researchers at ACDI; skills in SQL would be an advantage. • Ability to provide advanced data science service to support the research enterprise (such as developing and improving data processing and analysis workflows and/or working with specialist scientists to identify and implement data science analyses for specific research questions). • Proficient in the management and use of online collaborative code platforms such as GitHub, and ability to train/support researchers in using such a platform. • Able to maintain and install data science software in multiple operating system environments (Linux/Unix, MacOS, Windows). Ability to install and maintain process-based models on Linux/Unix would be an advantage. • Proficiency in working in a high-performance computing (HPC) environment. Capability to procure, mission and maintain small HPC machines would be advantageous. • Proficient in sourcing, processing and analysing large geospatial datasets such as climate model data, land-surface datasets, and/or earth-observation datasets. • Proficient in data quality management and use of data quality management tools in a research environment. • Works well in a team, and able to understand key scientific concepts in the natural sciences. • Strong written and verbal communication • High numeracy, demonstrating high levels of attention to detail • Enjoys working with and validating data (including spatial and temporal data) • Excellent written and verbal communication skills. • Ability to work effectively in interdisciplinary teams. • Meticulous attention to detail

Knowledge	• Expert technical specialist knowledge of data science, statistics, and/or bioinformatics • Good knowledge (e.g. through experience or tertiary level courses) in one or more fields relevant to ACDI, such as climate science, hydrology, ecology, public health, economics, geography.
Professional registration or license requirements	None
Other requirements (If the position requires the handling of cash or finances, other requirements must include 'Ability to handle cash or finances'.)	None

Competencies (Refer to UCT Competency Framework)

Competence	Level	Competence	Level
Analytical and Problem Solving	4	Building Interpersonal Relationships	3
Conceptual Thinking	4	University Awareness	2
Client/Student Service and Support	3	Work Management	3
Research Support Skills	4	Professional Knowledge and Skill	4

SCOPE OF RESPONSIBILITY

Functions responsible for	• Act as lead or co-lead for research data governance and management within ACDI. • Lead on data science in ACDI's Executive Committee • Provide input on data science component of budgets regarding resources (e.g., computing), equipment, time and staffing requirements. • Mentor and advise junior data scientists • Maintaining links to data science functions in the wider university, including the AI initiative and eResearch. • Lead on the provision of training and capacity building in data science when necessary.
Amount and kind of supervision received	• Responsible for delivery of agreed monthly activities, without direct supervision. • Regular interaction with data centre researchers. • Regular meetings to discuss projects and progress. • Monthly team meetings with Project PI(s).
Amount and kind of supervision exercised	• Mentoring of junior data scientists. • Data science support/advice/training to Post-docs, PhDs and Masters students.
Decisions which can be made	• Day-to-day and week-to-week decisions about data science, scheduling of own work, representation in global working groups.
Decisions which must be referred	• Changes in technology and approach in data architecture, curation, and governance. • Large procurements. • Changes to models and algorithms used for data linkage. • Approval of requests for data access from external parties. • Data governance changes proposed by working groups

CONTACTS AND RELATIONSHIPS

Internal to UCT	Research staff, ICTS, Library staff, AI Initiative, E-research
External to UCT	Data support staff of funders; collaborators on research projects.