# Position Details

## Technical Services- CSOF5

|  |  |
| --- | --- |
| The following information is for applicants | |
| Advertised Job Title | Data Engineer/ Analyst |
| Job Reference | 78526 |
| Tenure | Indefinite or Specified Term of 24 months  Full-time or Part Time |
| Salary Range | CSOF5 Level  AU$102,724 to AU$111,165 pa (pro-rata for part-time) + up to 15.4% superannuation |
| Location(s) | Canberra, Melbourne preferred. |
| Relocation Assistance | Will be provided to the successful candidate if required |
| Applications are open to | * Australian Citizens Only * Australian/New Zealand Citizens and Australian Permanent Residents Only * All Candidates |
| Position reports to the | Peggy Newman |
| Client Focus – Internal | 80% |
| Client Focus – External | 20% |
| Number of Direct Reports | 0 |
| Enquire about this job | Contact Peggy Newman via email at peggy.newman@csiro.au |
| How to apply | Apply online at <https://jobs.csiro.au/>  Internal applicants please apply via **Jobs Central**  If you experience difficulties when applying, please email [careers.online@csiro.au](mailto:careers.online@csiro.au) or call 1300 984 220. |

### Role Overview

* Do you want to apply your data skills to be part of science and research at CSIRO?
* Are you passionate about open-source software and open data?
* Would you like to work on international collaborations?

We are offering an exciting data role working at CSIRO in biodiversity science, using modern open- source technology and collaborating with stakeholders all around the country and the world.

The Atlas of Living Australia (ALA) is Australia’s national biodiversity data aggregator funded under the National Collaborative Research Infrastructure Strategy (NCRIS) and hosted by CSIRO. The ALA is the Australian node of the Global Biodiversity Information Facility (GBIF). Our digital infrastructure is developed in-house to support research activities, government decision-making and community events.

The ALA uses and produces open-source software and tools to aggregate Australian biodiversity data from a variety of providers and make it discoverable and reusable online. Our technology stack is reused by over 25 countries.

The ALA Data Management team is seeking a data engineer for a **two-year contract opportunity** to work on data acquisition, transformation, loading and quality assurance**.** Our team is technically oriented and uses multiple technologies and platforms to explore and manipulate large datasets into a standardised format, which we then ingest into our processing pipeline.

### Duties and Key Result Areas:

As a successful candidate you will develop new and support existing automated jobs to harvest data from a series of data providers, ensuring data currency and quality is consistent with expectations. You will need to be effective both as a team member and as a reliable point of contact for data providers. We’re looking for strong collaboration and communication skills, and the ability to develop great rapport with stakeholders.

Suitable candidates located in either Canberra or Melbourne are encouraged to apply. Our main office is in Canberra, but we have a Melbourne cohort who work completely remotely. Our workplace culture facilitates remote and flexible work.

Duties and Key Result Areas:

* Work to the Data Manager in the Data Management team to build and manage both automated and manual data loading processes
* Map datasets to the Darwin Core standard
* Implement, deploy, schedule and maintain data load processes
* Implement quality assurance and verification on datasets to ensure loaded records meet expectation
* Engage professionally with external stakeholders offering technical guidance on data management issues such as data mapping, automation, and loading
* Contribute to team meetings and planning and review activities

## **Required Competencies:**

* **Teamwork and Collaboration:** Cooperates with others to achieve organisational objectives and may share team resources in order to do this. Collaborates with other teams as well as industry colleagues.
* **Influence and Communication:** Uses knowledge of other party's priorities and adapts presentations or discussions to appeal to the interests and level of the audience. Anticipates and prepares for others reactions.
* **Resource Management/Leadership:** Sets up and maintains effective and efficient work teams and manages performance and resources, to achieve objectives. Chooses appropriate management strategies and communication styles to maintain high levels of motivation and productivity. Gives feedback for development purposes and provides support and direction for improvement.
* **Judgement and Problem Solving:** Investigates underlying issues of complex and ill-defined problems and develops appropriate response by adapting/creating and testing alternative solutions.
* **Independence:** Plans, sets and works to meet challenging standards and goals for self and/or others. Recognises where endeavours will make the most impact or difference, decides on desired outcome and sets realistic goals to reach this target.
* **Adaptability:**Copes with ambiguity or situations that lack clarity. Adapts readily to changing circumstances and new responsibilities (which may include activities outside own preferences) in the interests of achieving team objectives. Recognises the need for and undertakes personal development as a result of changes.

## **Selection Criteria**

Our preference is to hire a talented programmer with tertiary qualifications in Computer Science/Software Engineering, Quantitative Ecology, Science or another relevant field.

#### Essential

We understand that women and other marginalised groups don’t tend to apply for these roles unless they meet all of the criteria, and we recognise that there can be other things that make a candidate a great fit. If you’re enthusiastic about working in biodiversity science, have strengths in just some of these areas and a willingness to learn fast, please get in touch.

* 2+ years demonstrated operations experience in a data driven production system
* Strong knowledge of scripting languages – Python, Spark, Scala, SQL, bash, R, JavaScript
* Strong ETL skills with large datasets with a focus on efficiency and scale
* Experience with Linux OS
* Experience with a variety of open source relational and non-relational databases
* Experience in both delivering and consuming REST services
* Source code management using git, svn or Bitbucket
* Knowledge of SOLR and/or Elasticsearch administration and queries
* Effective stakeholder engagement and technical liaison skills

**Desirable Criteria:**

* Background or strong interest in biodiversity/ecology/taxonomy
* Enthusiasm and knowledge of open data standards, procedures and policy
* Experience with Apache Beam/Spark/AVRO, Jenkins, ELK, Zabbix, Ansible
* Experience with geospatial data systems and development
* Experience with Darwin Core standard

## *CSIRO thrives on innovation; innovation that is harnessed by the diversity of the minds and lived experiences brought to our greatest challenges. If you meet some, but not all requirements, we encourage you to submit your application. Our priority is to improve how we recruit diverse, talented people and empower them to flourish.*

Special Requirements

Appointment to this role may be subject to conditions including provision of a national police check as well as other security/medical/character clearance requirements.

## **About CSIRO:**

We solve the greatest challenges through innovative science and technology. To find out more visit us [online](http://www.csiro.au/)!