APS5 Data Scientist - Expression of Interest

Job ID: 4152, **APS Code:** 270308

What Inspires My Passion for Data Science

Minimising the "Tax Gap" will ensure that the Federal Government has the financial resources to implement its stated programs. This is a worthwhile and rewarding way for me to spend my final decade in the work place. Using my well developed analytical skills to ensure that the Australian Taxation System is applied fairly is what motivates me to work at the Australian Taxation Office.

Relevant Experience - STAR Examples

Analytical Solution Development & Stakeholder Liaison (Duties 2, 4)

Situation: ATO needed feasibility assessment of vision-language AI models for extracting structured data from work-related expense documents. As sole developer, I coordinated across data engineering and senior data scientists within strict timeframes.

Task: Evaluate two models against 195 business documents, extracting 17 expense fields, while liaising with data engineering for ATO infrastructure deployment and reporting progress to senior data scientists.

Action: Completed technical documentation and security assessments for model deployment. Designed comparative evaluation framework with batch processing, implementing multiple matching strategies (exact for monetary values, fuzzy for text, position-aware for transactions) and created ground truth datasets. Met all milestone deadlines through regular progress demonstrations.

Result: Evaluated 195 documents achieving 52% overall accuracy across 17 fields. Identified which field types (monetary amounts, dates) performed well versus those requiring development (complex transaction tables). Provided data-driven recommendations on model selection and highlighted technical constraints (120 documents/hour processing speed), enabling evidence-based decisions on next-phase investment.

Technical Reporting & Data Visualization (Duty 5)

Situation: Proof-of-concept evaluation results needed clear communication to senior data scientists and business stakeholders for AI investment decisions.

Task: Translate technical findings (accuracy breakdowns, performance differences, processing constraints) into accessible insights while maintaining scientific rigor.

Action: Developed dashboards and automated reporting visualising performance by document type and field category. Created executive summaries in plain language and comparative visualisations showing which documents suited automated extraction versus manual review.

Result: Dashboards enabled quick assessment showing certain fields (GST amounts, total amounts) achieved viable accuracy while others (transaction tables) required development. Visualisations communicated that models showed promise for pre-screening but constraints meant extraction should augment, not replace, human verification. This honest assessment enabled realistic planning for incremental improvements.