APS5 Data Scientist - Expression of Interest

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

What Inspires My Passion for Data Science

The ATO's AI Strategy vision to "industrialise responsible AI with scale and speed to enable lawful, transparent, and efficient tax collection" represents a transformative opportunity to modernise how we serve government and taxpayers. I am particularly motivated by the Document Understanding initiative—the ability to automatically read, understand, and analyse taxpayer documents, saving significant human hours while improving accuracy. My recent work evaluating vision-language models for document extraction demonstrates my commitment to advancing this vision. Contributing my analytical skills to ensure the Australian Taxation System is applied fairly through scalable AI solutions is what drives me to work at the ATO during this pivotal period of technological transformation.

Relevant Experience - STAR Examples

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

Situation: ATO required an urgent feasibility assessment of vision-language AI models for extracting information from taxpayer documents (invoices, receipts and bank statements) to facilitate the substantiation of work related expense claims. As this technology had never been deployed at ATO, the Document Understanding Group needed an assessment of extraction accuracy and processing speed to inform investment decisions.

Task: Design and implement a comprehensive evaluation system to measure the accuracy and throughput of two vision-language models over 195 documents.

Action: I pioneered ATO's first vision-language model deployment: liaising with data engineers to establish the environment, configuring AAP within resource constraints, and designing the evaluation framework with senior data scientists (17 fields, matching strategies, batch processing). As sole developer, I implemented the system, completed technical documentation and security assessments, and met all milestones through regular demonstrations.

Result: Findings enabled the team to prioritize development efforts: monetary amounts and dates showed production-ready accuracy while transaction tables required further work. By providing data-driven model comparisons and throughput constraints (120 documents/hour), I equipped stakeholders with evidence to guide ATO's Document Understanding investment strategy.

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: Design and build the dashboard to translate technical findings (accuracy breakdowns, performance differences, processing constraints) into accessible insights while maintaining scientific rigor.

Action: I designed the reporting interface to visualise 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. Implementing the visualisation solution through a dashboard. Built the process and corresponding data so it can be displayed in an automated manner.

Result: Dashboards enabled quick assessment of the model evaluation including 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 data-driven assessment enabled realistic planning for incremental improvements.