# **Spend Categorisation Tool**





#### Background

General Ledger data is collected from over 150 NSW Government Agencies

\$30 billion in annual spend data

More than 2 million transactions per quarter

The rule-based **system in place** utilises **over a million mapping rules; difficult** and **time-consuming** to maintain

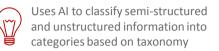
The DAC has been engaged to develop an automated and robust Spend Categorisation Tool to replace this process

#### Implementation

Successful completion of 'proof of concept' with random sample

Currently undergoing acceptance testing, understanding how the model may be operationalised

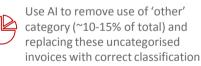
The	So	lutio



Can learn mapping rules from Constant of the second data Constant of the second data

Self-learning: accuracy will improve over time with use

Accuracy for Whole of Government spend data has exceeded proposed 90% benchmark



"The DAC has engaged really motivated scholars and used cutting-edge technology to deliver an outstanding result."

- Amol Chavan, NSW Procurement

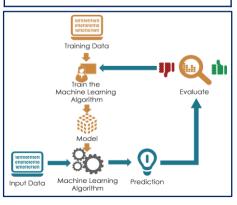
### Impact

Greater transparency of government spending

Demonstrates spending patterns through categorisation at a point in time, or over months or years

Allows elimination of silos where different departments purchase same assets. Encourages interdepartmental collaboration and economies of scale

Identification of areas where resources are over or under-utilised



"Machine learning has allowed us to automatically match category of invoice against agency, supplier and a particular kind of spend, using data science and deep learning technologies."

- Dr. Khimji Vaghjiani, NSW DAC

## **Next Steps**

Improve accuracy to greater than 95%

Use tool for classifying whole of government spend data from multiple perspectives

Potential to commercialise the model into a solution that would be offered to other agencies across government, and even the private sector