

Curious Case of CAITY

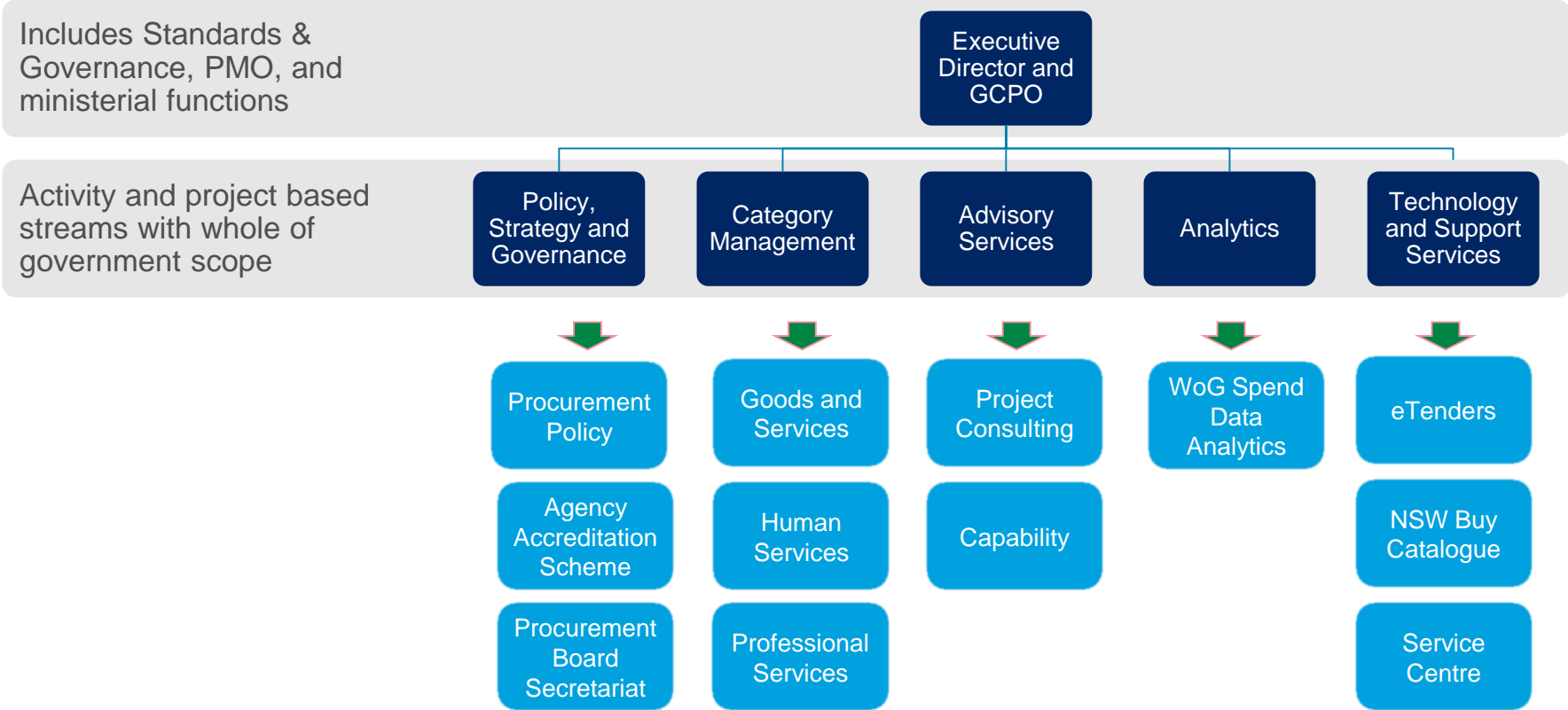
Artificial Intelligence in NSW Procurement

NSWP Analytics Team



NSW Procurement (NSWP) Overview

NSW Procurement sits under the Commercial, Commissioning & Procurement Branch of the NSW Treasury Cluster



3 Pillars of NSWAP Analytics – What we do

The Analytics Team

Spend Consolidation

Consolidate sector wide spend, supplier reporting and related ancillary data.

Analytics Platform

Provide analytics platforms (e.g. category specific dashboards) to support procurement teams derive insight and deliver their objectives.

Analytical Insight and Solutions

Support clusters derive insights and solutions, by understanding their requirements and leveraging data.

Analytics Team Evolution

NSW Procurement Spend Analytics has evolved over last 4-5 years and is now set to use advanced analytics tools like AI / Machine Learning for deriving insights from the spend data

Analytics Team

- Data enrichment
- Online Spend Dashboard (outsourced provider)

2015

Advanced Analytics

- AI / Machine Learning
- BAS Category Dashboard
 - Taxonomy Project

2018

2013

- Spend data collection
 - Desktop tools for analytics

Strategy Unit

2017

- In-house Dashboard development
 - \$1M+ savings
- Advanced Analytics
 - Open Data

Spend Cube Online

2019+

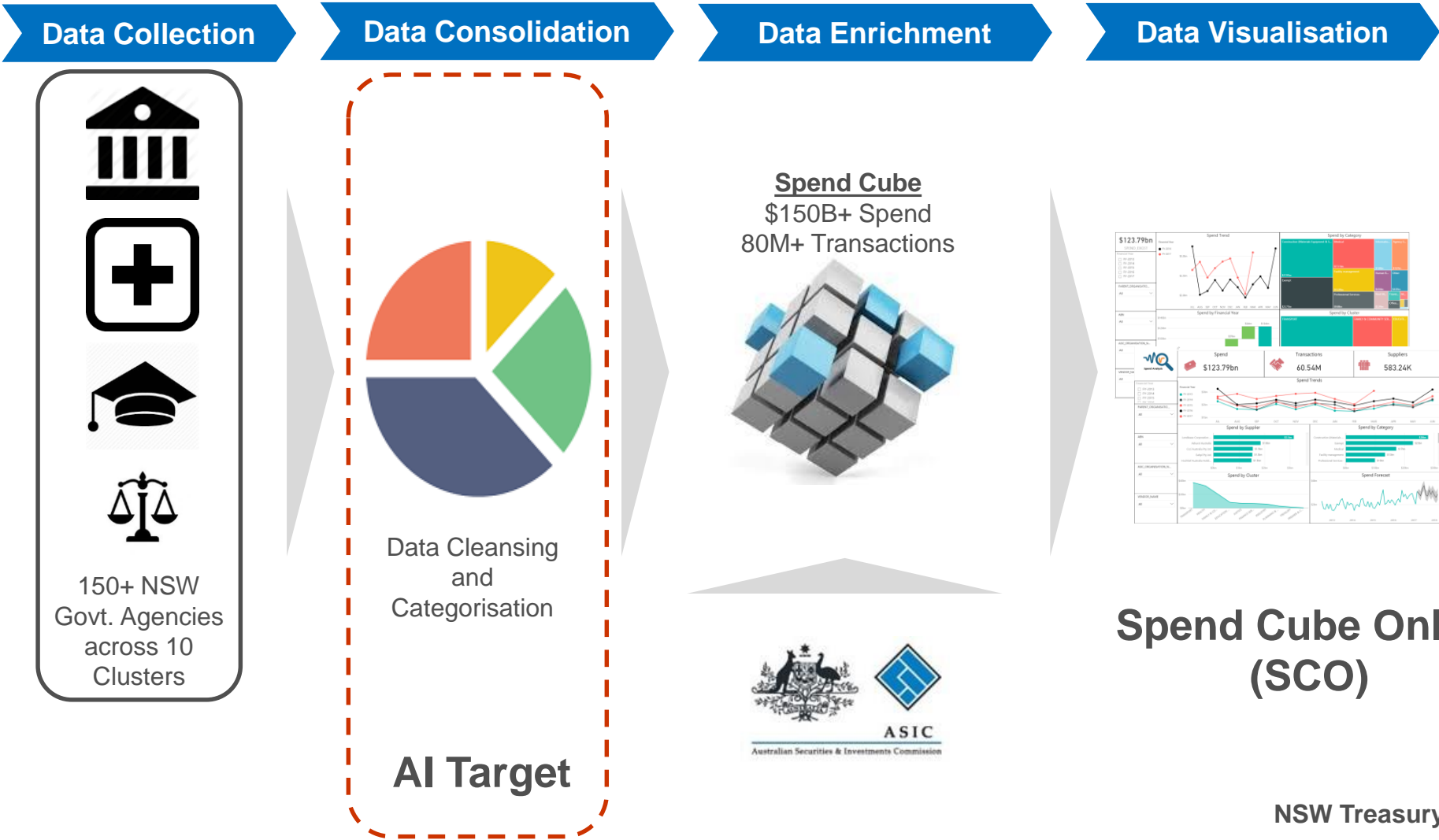
- Enhance machine learning (taxonomy project)
 - Expand category dashboards
 - Develop agency dashboards

Enhance WoG Capability (Vision 2021)



Spend Analytics – Process Overview

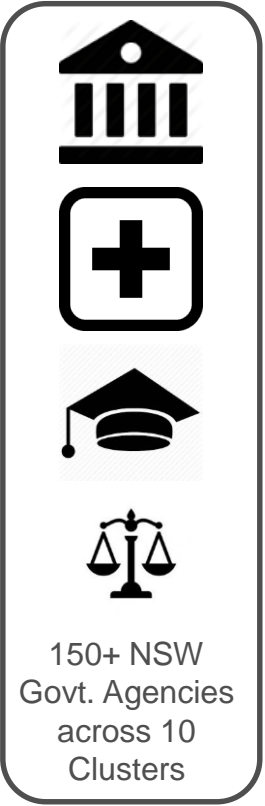
The Analytics Team collects General Ledger data from 150+ NSW Govt. Agencies, processes the data to create a Spend Cube and visualises it for further analysis and end user reporting



Spend Categorisation – Prior Process

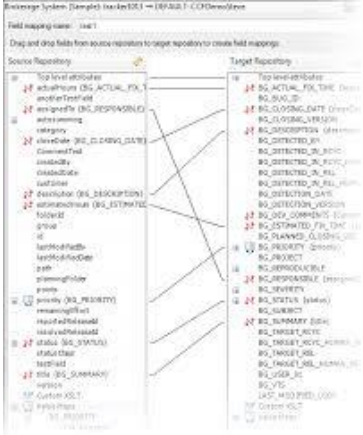
Spend data categorisation is a key step in Spend Analytics; Previous categorisation process is effort and time intensive

Input Data



150+ NSW
Govt. Agencies
across 10
Clusters

Mapping Rules



Categorised Spend Data



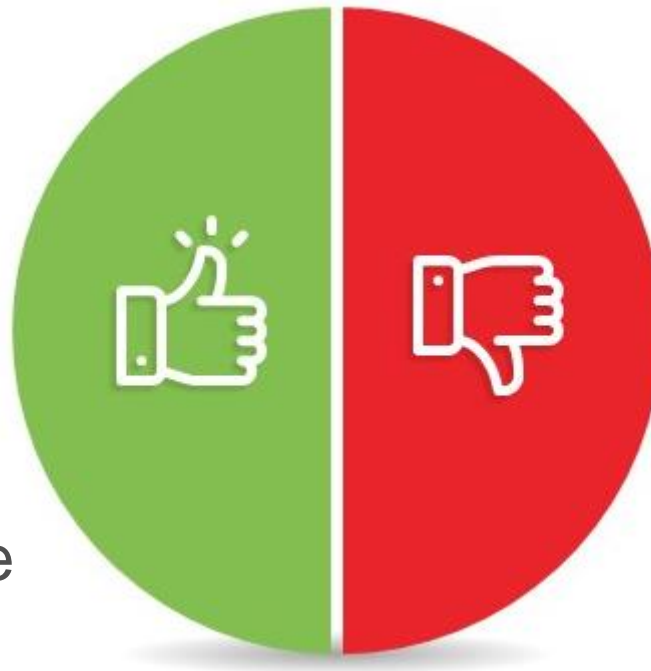
Categorisation Issues

- 1M+ Mapping Rules
- Mapping rules are cumbersome to maintain
- Difficult to check and utilise all key data fields to accurately categorise spend

Overview: Manual Categorisation Method

Pros

- ✓ Easy concept to learn and perform
- ✓ Prior human experience (orthodox method)



Cons

- ✗ Time and Resource Expensive
- ✗ **Human Inconsistency/Bias**
- ✗ Tedious and Onerous

Why we need a new solution?



2 mil+ transactions of new spend data received quarterly. A never ending process/task for analysts.

	A	B	C	D	E	F	G	H
1	Account Number	Account Name	sku	category	quantity	unit price	ext price	date
2	803666	Fritsch-Glover	HX-24728	Belt	1	98.98	98.98	9/28/2014 11:56
3	64898	O'Conner Inc	LK-02338	Shirt	9	34.8	313.2	4/24/2014 16:51
4	423621	Beatty and Sons	ZC-07383	Shirt	12	60.24	722.88	9/17/2014 17:26
5	137865	Gleason, Bogisich and Franecki	QS-76400	Shirt	5	15.25	76.25	1/30/2014 7:34
6	435433	Morissette-Heathcote	RU-25060	Shirt	19	51.83	984.77	8/24/2014 6:18
7	198887	Shanahan-Bartoletti	FT-50146	Shirt	4	18.51	74.04	9/5/2014 7:24
8	969663	Gusikowski, Reichert and Gerlach	AE-95093	Shoes	4	49.95	199.8	4/28/2014 21:51
9	1288	Wilderman, Herman and Breitenberg	FT-50146	Shirt	14	68.2	954.8	12/4/2013 13:53
10	979589	Brown Inc	HX-24728	Belt	16	52.99	847.84	2/7/2014 14:53
11	839884	Turcotte, Turner and Anderson	FT-50146	Shirt	8	21.35	170.8	9/3/2014 16:06
12	91059	Armstrong, Champlin and Ratke	FT-50146	Shirt	17	34.61	588.37	4/6/2014 20:52
13	151300	Upton, Runolfsson and O'Reilly	HX-24728	Belt	3	75.1	225.3	8/3/2014 7:55
14	185412	Goyette, Kessler and Goodwin	AE-95093	Shoes	10	10.93	109.3	7/11/2014 17:20
15	357988	Dietrich-Krajcik	QS-76400	Shirt	15	11.15	167.25	2/21/2014 8:57
16	269278	Luetzgen-Ritchie	QS-76400	Shirt	11	91.95	1011.45	9/3/2014 2:15
17	823639	Reilly-Leannon	LK-02338	Shirt	11	72.65	799.15	5/27/2014 12:36



Manual categorisation is long and cumbersome. Time consuming for analysts, where they could be focusing on other more technical tasks.



Minimise mapping errors due to human bias.

AI Background – What is AI?

Artificial Intelligence is the science of making smart machines that can perform human tasks.

Some examples that come to mind:

Self Driving Cars



SIRI



R2D2

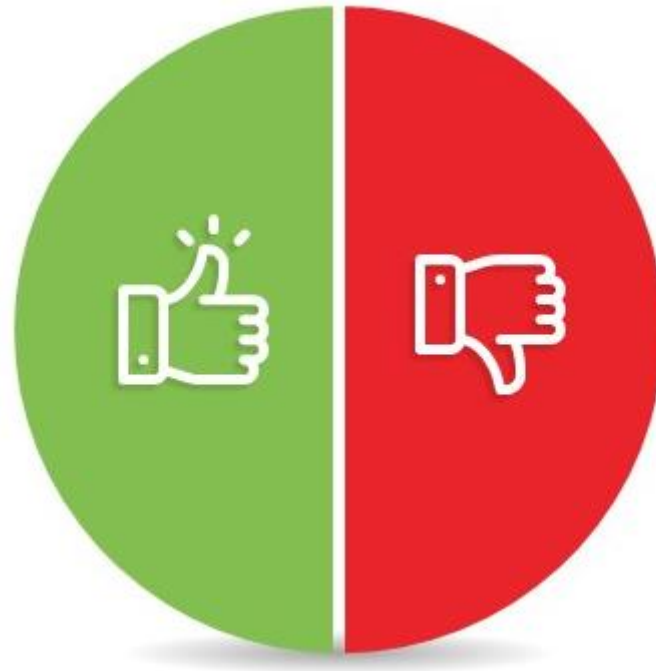


BUT
Where does AI fit into
NSW Procurement?

Overview: Spend Categorisation By AI

Pros

- ✓ Consistent & Accurate
- ✓ Fast & Efficient
- ✓ Save resources & expenses



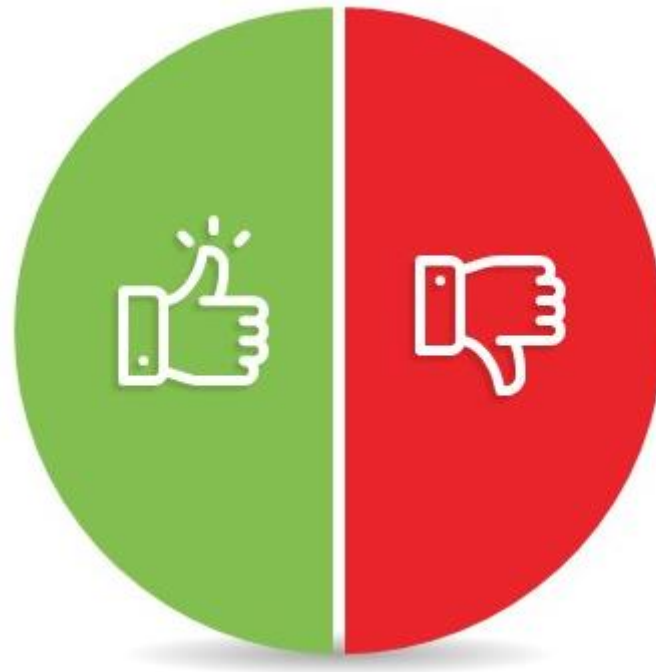
More Pros

- ✓ Minimise human intervention
- ✓ Replace 1 million+ rules with automation
- ✓ Allow analysts task prioritisation

Overview: Spend Categorisation By AI

Pros

- ✓ Consistent & Accurate
- ✓ Fast & Efficient
- ✓ Save resources & expenses



Cons

- ✓ Accuracy dependant on training data quality
- ✓ Heavy initial set-up phase

Meet CAITY

Role: Procurement Categorisation Master

Interesting facts about CAITY

Birthdate: 23 Feb 2018

Age: 1 year and 3 months old

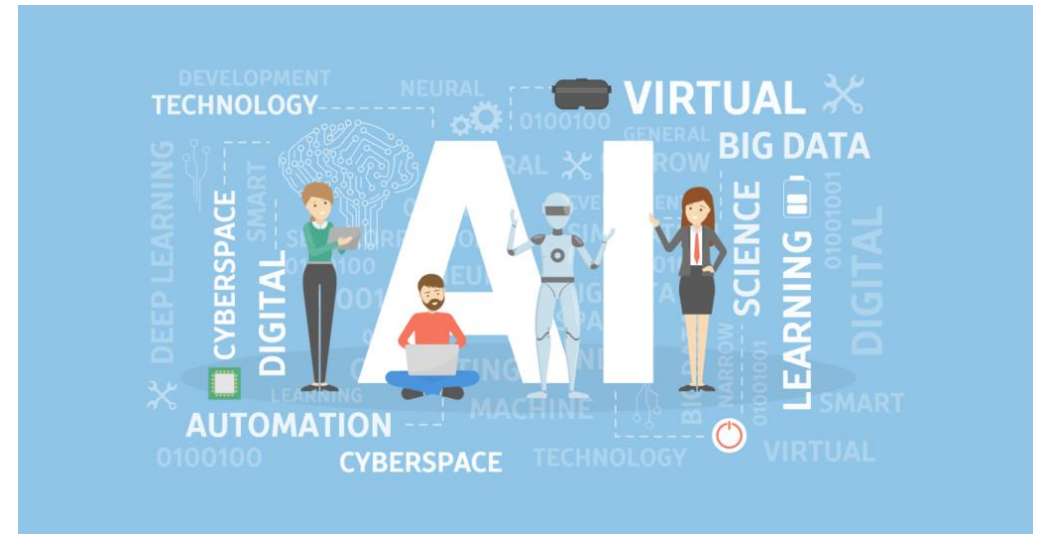
Lives: Level 11 of the Mckell Bld for now

Can work 24/7, do not need a lunch break

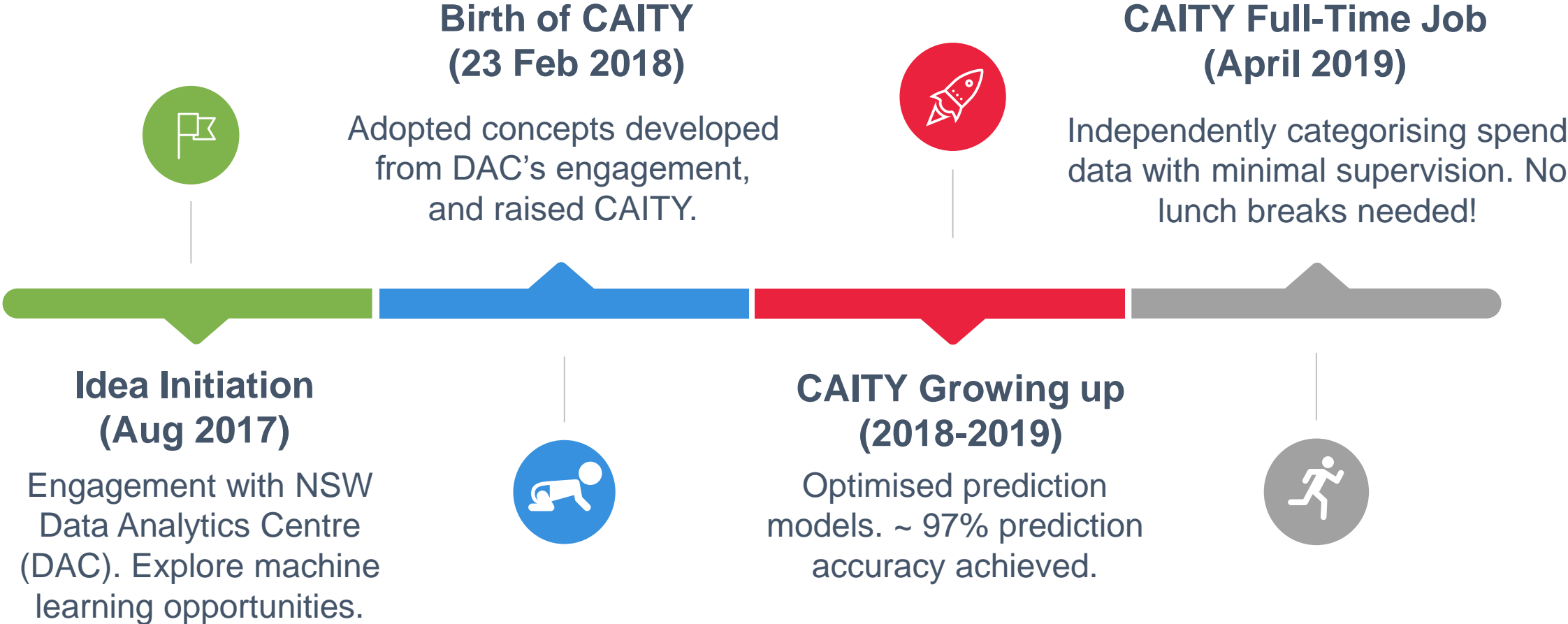
Dear child of NSW Analytics

Name stands for Categorisation AI TechnologY

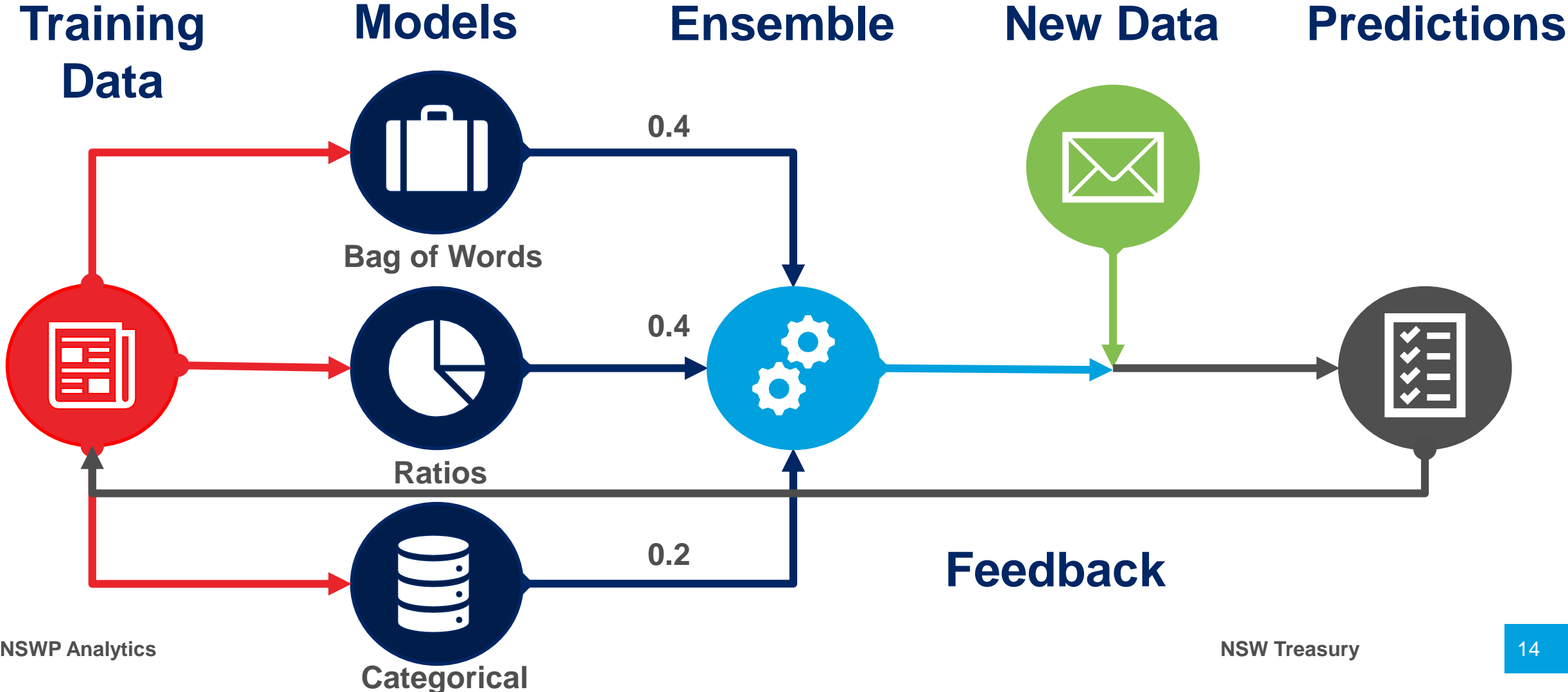
CAITY, future generation of Procurement Analytics!



CAITY Developmental Timeline



Inside CAITY's Brain - Implementation Framework



What we feed CAITY: Historical Spend Data (Training data)

Excel Spreadsheet Input File Structure

Cluster Name	Transac. ID	Vendor Name	ABN	GL Account Name	Description	ANZSIC Agency Data	ANZSIC Desc.	Category Labels
Industry	4729002	Telstra	1918 2257 835	Purchase Assets	Industry Hardware	5004581 8	Internet Service Providers	ICT Hardware
Health	5482185	Hewlett Packard	11854 5926 211	Maintenance	HP Server	N/A	N/A	ICT Maintenance
Education	15616135	Deloitte	5891 4891	Grants Paid	e-Learning Courses	6151381 2	Scientific Research Services	Media Production

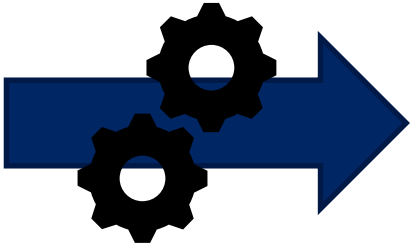
N.B Data is made-up for the purpose of demonstration

CAITY on the Job: New Data & Prediction Results

New Data

Cluster Name	Transac. ID	Vendor Name	...
Industry	4729002	Telstra	...
Health	5482185	Hewlett Packard	...
Education	1561613	Deloitte	...

CAITY Calculations



Predicted Results

...	ANZSIC Desc.	Top Predicted Category & Prediction Confidence (%)	
...	Internet Service Providers	Equipment	93%
...	N/A	Food Catering & Consumables	71%
...	Scientific Research Services	Advertising Related Services	95%

N.B Data is made-up for the purpose of demonstration

Challenges with “toddler” CAITY

Lower prediction rates on new vs. existing vendors

Handling uncompliant data inputs

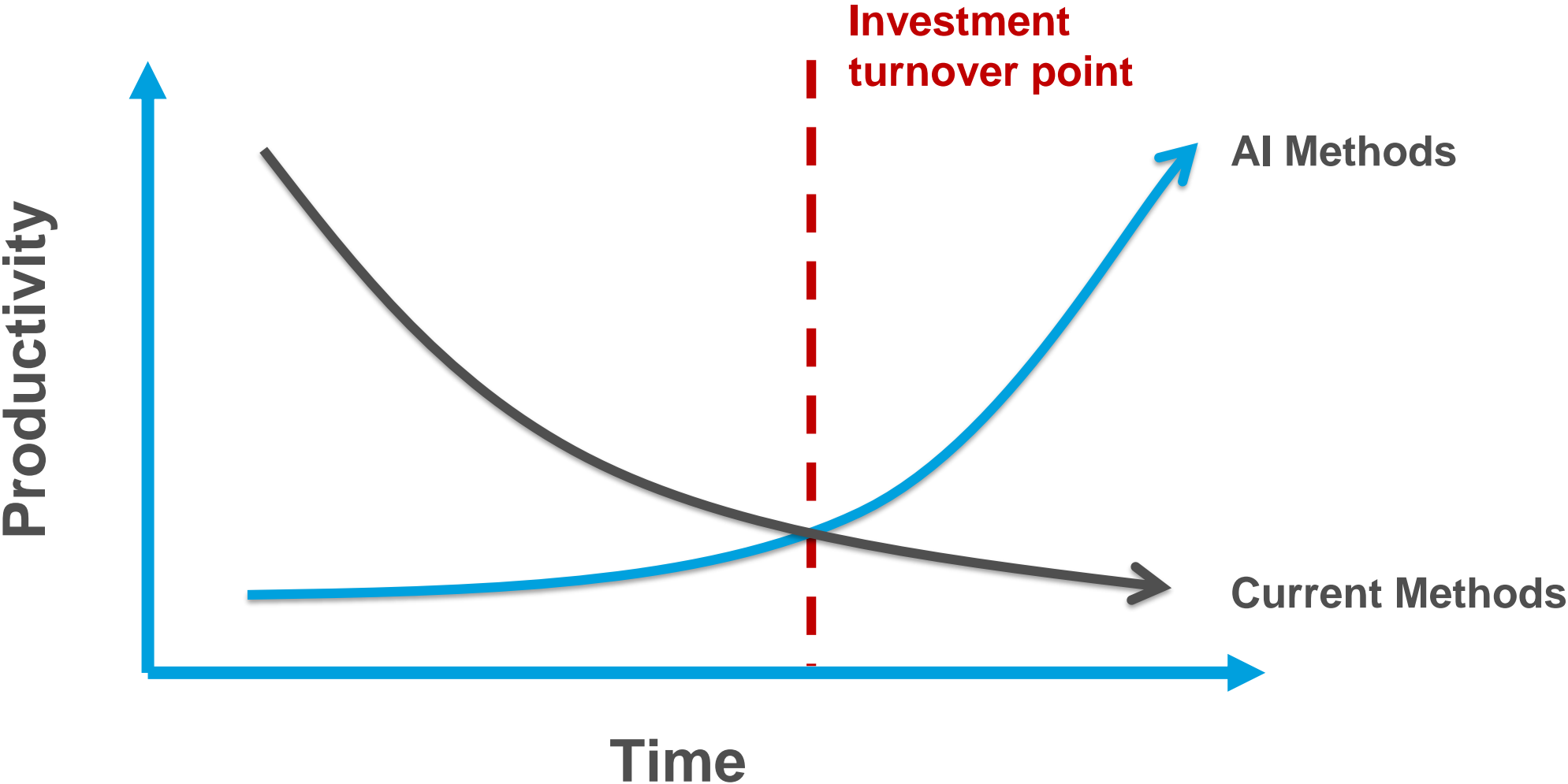
Technical troubleshooting

Lack of validated training data



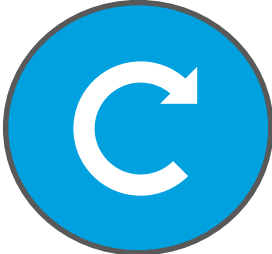
Changing the nappies is inevitable

CAITY Future Potential



CAITY Summary

Increased Spend Data Turnover



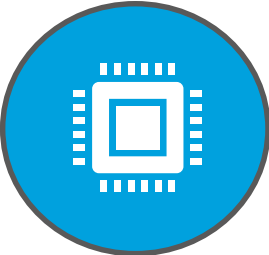
Faster data consolidation for data visualisation on the Spend Cube

Efficient Staff Utilisation



Allow analysts to focus on the analysis of transactions

Evolving with Technology



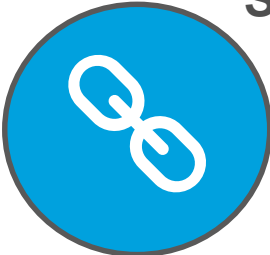
Using the latest technology instead of outdated methods

Minimise Error



Spend categorisation consistency. Eliminate human bias

Spend Cube Integration



Streamline integration into current Spend Cube framework

Time-Cost Effective



A cost effective long term solution to spend categorisation

Future of CAITY

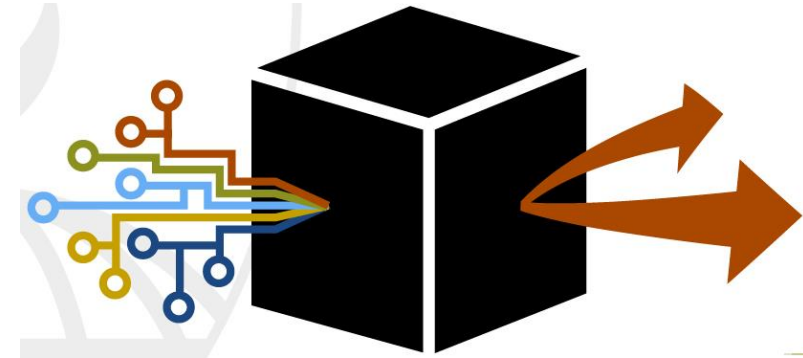
Apply to solve problems in different frameworks:

- A transferable model.
- Adaptable to similar government frameworks.
- Applicable to other tasks involving categorical data.
 - **As long as there is quality training data.**

Create a user-friendly interface:

- To categorise new data, send feedback & retrain models.
- Streamlining the process.

Putting a face to CAITY



Questions for CAITY?



Appendix - Spend Cube – Outcomes

Procurement Teams

- ✓ Understand spend using a visual tool
- ✓ Drive procurement strategy
- ✓ Derive actionable insights
- ✓ Track supplier diversity e.g. SMEs, aboriginal, disability

Small Clusters/Agencies

- ✓ Provide spend analytics support
- ✓ Help leverage buying power of WoG spend
- ✓ Help understand priorities to target savings
- ✓ Help respond to GIPA requests

Ministerial Services

- ✓ Help respond to Question on Notice
- ✓ Provide inputs to House File Notes
- ✓ Prepare for meetings with suppliers

Open Data

- ✓ Publish data on Data NSW website
- ✓ Demonstrate transparency

Appendix - Spend Cube Online - Framework

