Stats NZ
Tatauranga Aotearoa

AI and the Future of NSOs

Gary Dunnet
Interim Chief Methodologist, Stats NZ

22 September 2025

AI beyond LLMs – world leading ML solutions

Machine Learning for migration estimation

- ML model for probabilistic classification of border-crossings
- Allows NZ to produce some of the most timely migration statistics in the world
- Is able to pick up disruptions (like COVID) without manual adjustments
- First use of ML in migration estimation in the world

Address selection from multi-source admin data using Machine Learning

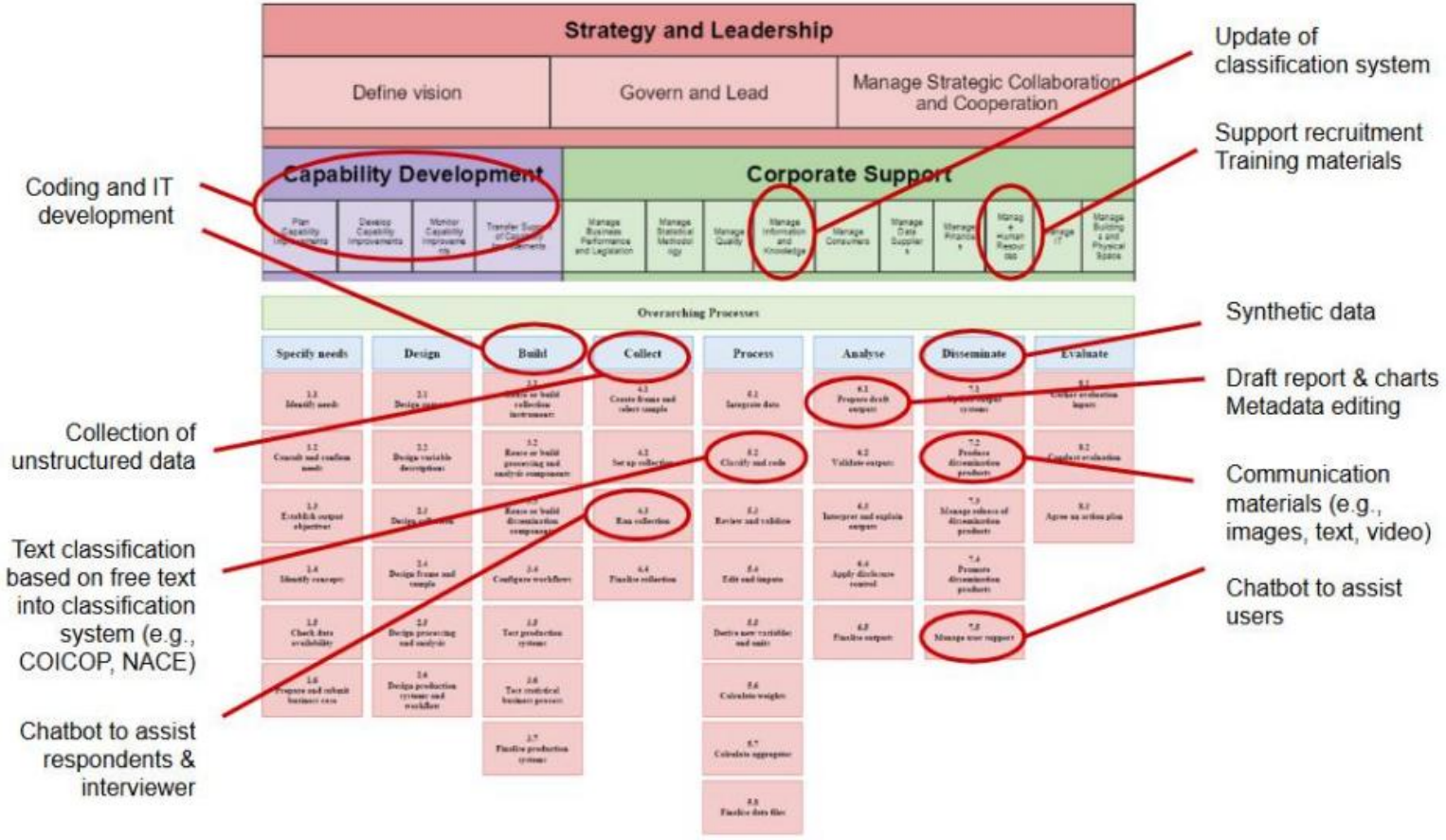
- Estimates the “most correct” address from multi-source data with many addresses for a single individual
- Used in address determination for admin data enumeration in Census 2023
- Allows for the construction of an “admin population” census (combined with other R&D)

Other AI/ML uses being considered / trialled

Land use classification using earth observation data	Useful for understanding various cropping production and could reduce the footprint of the Agriculture Production Survey (APS) Neural networks and time series data of seasonal plant changes being considered Would extend the classifications available in datasets such as Sentinel-2 (but for NZ)
LLMs for data classification, response processing, and categorisation	Classification and categorisation of survey responses has broad usefulness across our collection suite LLMs with some tuning have the ability to process responses into standard classifications and categories With human-in-the-loop guardrails, this may amplify the efficiency of survey processing
Microsoft Copilot pilot	We're running a Stats NZ-wide pilot of Microsoft Copilot (concluding end of September), exploring both chat-based productivity and the new "agents" features. Early feedback shows Copilot is helping staff summarise complex threads, search through accessible documents, draft communications, and automate repetitive tasks. We're closely monitoring during the pilot the adoption, user experience, and assurance controls to inform a safe and scalable rollout afterwards
GitHub Copilot coding assistant (Stats-wide rollout)	Our rollout of GitHub Copilot as a coding assistant is underway across Stats NZ. Early results show analysts and developers are spending around **20% less time** on repetitive programming tasks, freeing up capacity for higher-value work. This is paired also with assurance: all users complete training, code is tagged when AI-assisted, and we have peer review and configuration baselines in place to manage IP and licensing risks.
SAS code conversion (new initiative)	We've launched an initiative to use AI for converting legacy SAS code to modern languages. This is a big step for Stats NZ, as it accelerates our migration off legacy platforms, reduces technical debt, and supports our move to cloud-native analytics. The work is already surfacing reusable patterns and lessons.
AI for Census policy-submission analysis	In partnership with the Ministry for Pacific Peoples, MBIE, and Ministry of Justice, we're building a repeatable approach to process and analyse survey submissions, starting out with this years Census policy surveys. AI is helping reduce turnaround time while preserving nuance and equity considerations. While we are trying to automate repetitive tasks, the human in the loop component is a key component in the process.

Where do the Experts think there is application?

Figure 1. Generative AI Application Areas in Statistical Organizations



UNCLA

Quality Management / Metadata Management

Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Build collection instrument	4.1 Create frame & select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Build or enhance dissemination components	4.3 Run collection	5.3 Review & validate	6.3 Interpret & explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame & sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit & impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing & analysis	3.5 Test production system		5.5 Derive new variables & units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare business case	2.6 Design production systems & workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production system		5.7 Calculate aggregates			
				5.8 Finalise data files			

Guardrails

- **Our posture in one line:** *Experiment then scale, always making our process open and keeping people involved at every step, using approved tools and having someone check the results before they're used.*
- We have clear accountability with the CTO being the Senior Responsible Officer for all AI at Stats NZ.
- AI Working Group (AIWG) with a formal Terms of Reference, reporting to the Systems & Standards Committee. An AI Triage Task Force pre-screens opportunities and risks, and a RASCI clarifies who decides, who does, and who assures..
- Our AI Policy (in force since 31 Jul 2025) and AI Coding Policy are live with companion procedures. They enforce registration in the Transparency Register, human-in-the-loop, prompt hygiene & data-classification controls, and TRM-approved tooling (or explicit exceptions with assurance). This also includes some pre-training requisites (our AI Fluency) modules, with level of training dependent on the sensitivity of the data used with AI.
- Every use case, whether statistical or corporate in nature is logged in a Transparency Register and it's life cycle is logged from prototyping to production. The AI Transparency Register is up and being actively monitored.
- We've delivered an enterprise comms plan (Te Matapihi hub, leader talking points, targeted campaigns) and are preparing external transparency content to sustain public trust.
- We are actively collaborating with other agencies to make sure that best practices, learnings, and safeguards are shared widely.

AI Training: (a snip of my training page)

Stats NZ
Tatauranga Aotearoa

- Course administration
  - Grades

Understanding Artificial Intelligence at Stats NZ

- Participants
- Badges
- Grades
- Introduction
- NEW MODULES: AI 101 Basics & Safe and Responsi...
- Value capture
- Trustworthy AI, ethics, privacy, and governance
- Prompt training
- Data and information foundations
- Risks of AI and GenAI
- Threats of AI
- More information, pātai
- Information Security Classification at Stats NZ
- Animated navigation off
- Data preference



Artificial Intelligence (AI) is changing the way we work — helping with writing text, write coding, data analysis, and more. At Stats NZ, we want all kaimahi to feel confident and informed when using AI tools.

We've created this AI training to guide you through what AI is, how to use it responsibly, and what's required depending on how you plan to use AI tools in your role.

AI use at Stats NZ
To use AI tools safely and responsibly at Stats NZ, there are minimum training requirements depending on your level of use:

If you...	You must complete...
Might use or encounter AI at work	✅ AI 101: From Basics to GenAI **+** ✅ Safe and Responsible AI at Stats NZ
Actively interact with tools such as ChatGPT, Microsoft Copilot or other embedded GenAI features, **or** use GenAI-generated content (e.g. include generated text, summaries, analysis, or code in deliverables.	✅ AI 101: From Basics to GenAI **+** ✅ Safe and Responsible AI at Stats NZ **+** ✅ Information Security Classification at Stats NZ
Use AI tools classified for coding, broader use, or higher-risk use.	✅ All the above **+** ✅ AI Fluency PDF modules: Value capture, Trustworthy AI, Prompt training, Data and Information foundations, Risks of AI and GenAI, and Threats of AI.



Why data ethics matter in an AI world

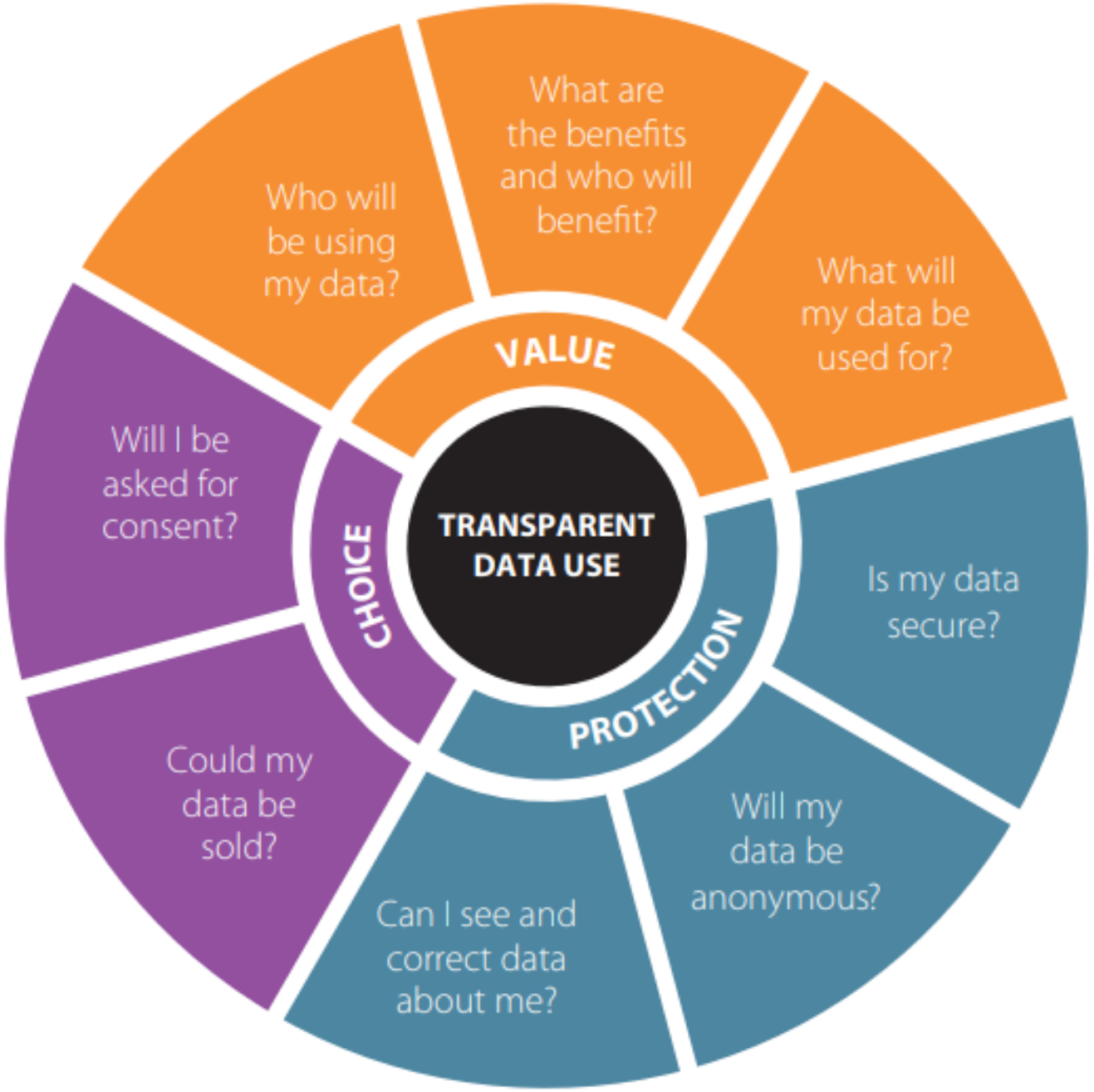
- **Scale and reach**
- **Visibility of harm**
- **Collective impact**
- **Power dynamics**
- **Consent challenges**
- **Technical complexity**

The Data Ethics Advisory Group

Building Trust and Ensuring Responsible AI

- Trust is earned through transparency, consistency, and moral responsibility
- Responsible AI Governance is essential
- Accuracy, reliability, and data integrity matter
- Actively protect privacy and prevent misuse
- Equity and cultural respect must guide all AI use.

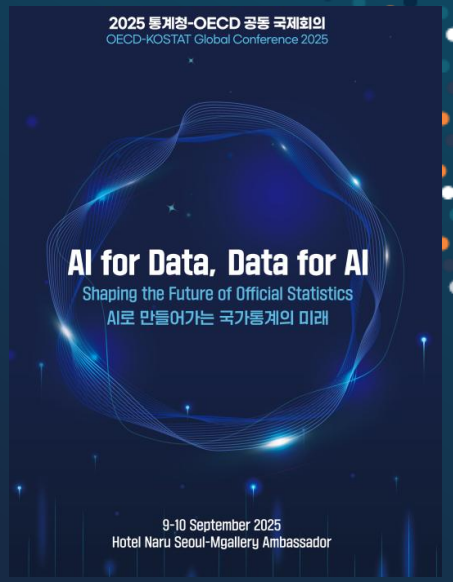




The 8 questions for trust

(2017 the NZ Data Futures Partnership)

Stats NZ
Tatauranga Aotearoa



OECD-KOSTAT Global Conference 2025

Questions?

Gary Dunnet
gary.dunnet@stats.govt.nz



Generative AI for Official Statistics (HLG-MOS Report) | UNECE