

3 September 2025

Tēnā koe

## Official Information Act Request

Thank you for your request of 8 August 2025, under the Official Information Act 1982, for the following information:

*I would love some more information to help me understand how AI is being used by NZQA in terms of marking. Erica Stanford was recently explaining in an interview that our country was world leading and the envy of other countries and I would love to see what we are doing in this space. Specifically:*

- *Who are the experts in AI for marking at NZQA and what experience do these experts have?*
- *What specific model is being used, and if its more than one model, what difference have been found between models?*
- *What specific training have you done with the models?*
- *Data about accuracy and reliability, and what methods were used to test for reliability and accuracy.*
- *What mitigation has been carried out to discourage cheating such as against prompt injections?*

Your request has been considered under the Official Information Act 1982.

To shorten the timeframe of the release of results for the NCEA co-requisite assessments, NZQA has utilised Automated Text Scoring (ATS) to mark digital responses for the Literacy – Writing assessment. Assessment against the NCEA co-requisite standards occurs during two two-week periods in the year, commonly referred to as Assessment Event 1 and Event 2. More information about the co-requisite assessments can be found on our [Literacy and Numeracy - NZQA](#) webpage.

### **Who are the experts in AI for marking at NZQA and what experience do these experts have?**

For Assessment Event 1 in May 2025, NZQA engaged an external vendor with over 20 years of experience in using AI for marking written assessment.

NZQA also has in-house expertise in AI marking, currently utilising a small team of Data Scientists with expertise in Artificial Intelligence and Machine learning (ML). Both our Data Scientists hold PhDs and bring several years of experience in applying AI and ML to practical problems, including their work at NZQA.

### **What specific model is being used, and if its more than one model, what difference have been found between models?**

NZQA is currently using an Automated Text Scoring (ATS) unimetric model, which is different to a large language model such as ChatGPT.

### **What specific training have you done with the models?**

The ATS model used by NZQA has been trained using human scoring from previous Literacy-Writing assessments (Sept 2023, May 2024, Sept 2024). The model's rule-based system learns

from previous New Zealand student responses to ensure the engine performs correctly in the New Zealand context.

**Data about accuracy and reliability, and what methods were used to test for reliability and accuracy.**

NZQA conducted a small-scale trial of Automated Text Scoring (ATS) on the Literacy-Writing assessment responses from the May 2024 assessment event and a large-scale pilot using 35,000 Writing assessment responses from the September 2024 assessment event. These allowed us to compare the results with scores provided by our regular human markers and showed that ATS would be suitably reliable and accurate.

With circa 35,000 student responses, the at-scale pilot showed approximately an 80% achievement agreement rate between ATS marking and human marking. To put this machine to human variation in context, human-to-human marking agreement rates also vary. Key human-to-human data points we have used to inform our decision making to proceed are check marking of end of year external assessments (87% agreement), and NZQA external moderation of school-based assessment (75% agreement).

To ensure confidence with moving to an automated marking solution for the Writing assessment, we implemented a human check-marking process for the May 2025 Writing assessment. Human markers re-marked responses at and within 1.5 score points of the achievement boundary. A similar assurance approach with selected human re-marking will be repeated for Assessment Event 2.

**What mitigation has been carried out to discourage cheating such as against prompt injections?**

The ATS model is not impacted by prompt injections because it has been trained using human scoring from previous assessments.

NZQA takes any possible [breach of external assessment](#) seriously and may investigate any report of a possible breach.

Our response to your request may be published on our website after five working days. Your name and contact details will be removed before publication.

If you require further assistance or believe we have misinterpreted your request, please contact Elizabeth Templeton in the Office of the Chief Executive, email [elizabeth.templeton@nzqa.govt.nz](mailto:elizabeth.templeton@nzqa.govt.nz) or telephone (04) 463 3339.

You have the right to seek an investigation or review by the Ombudsman of this decision under section 28(3) of the Official Information Act 1982. Details of how to make a complaint can be found at [www.ombudsman.parliament.nz](http://www.ombudsman.parliament.nz). You can also telephone 0800 802 502 or write to the Ombudsman at PO Box 10152, Wellington, 6143.

Nāku nā



Dr Grant Klinkum  
Pouwhakahaere/Chief Executive