

RM Assessment  
Summit 2025

# RM Focus

## Marking Engine in Action

Dr. Gráinne Watson, RM Chief Operations Officer





# Marking Engine background

RM built a Marking Engine using Generative AI and proved:

**1** We could match and complement human marking

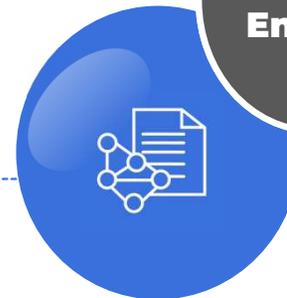
**Logic:**  
Rubrics & Framework



**Consensus Experience:**  
Real and Created

**2** We could provide appropriate feedback to explain mark provided

**Data:**  
Real and Synthetic



**Feedback:**  
Correct & Instantaneous

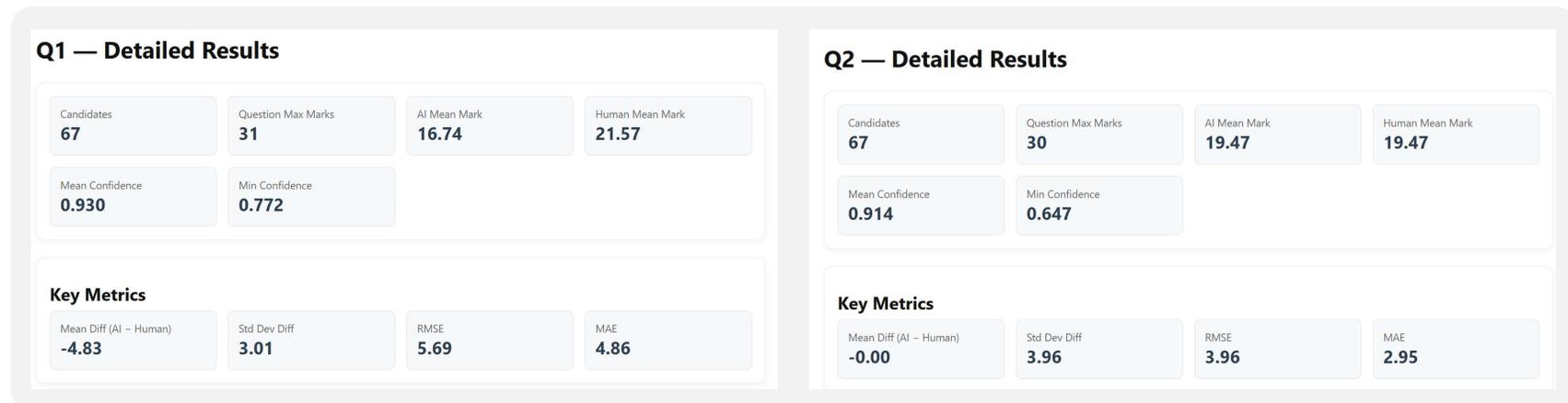
**3** We could answer nuanced and long form essays

**Marking Engine**



# What next? More subjects

We are currently working on several complex PoCs where we are marking a variety of subjects and we are seeing some commonalities from our other PoCs



## First look commonalities

- For some answers, the AI marking engine can mark as well as humans from the outset
- But overall, AI consistently awards lower marks than a human
- Distribution of scores can differ significantly, but widest gap still remains human marker vs human marker



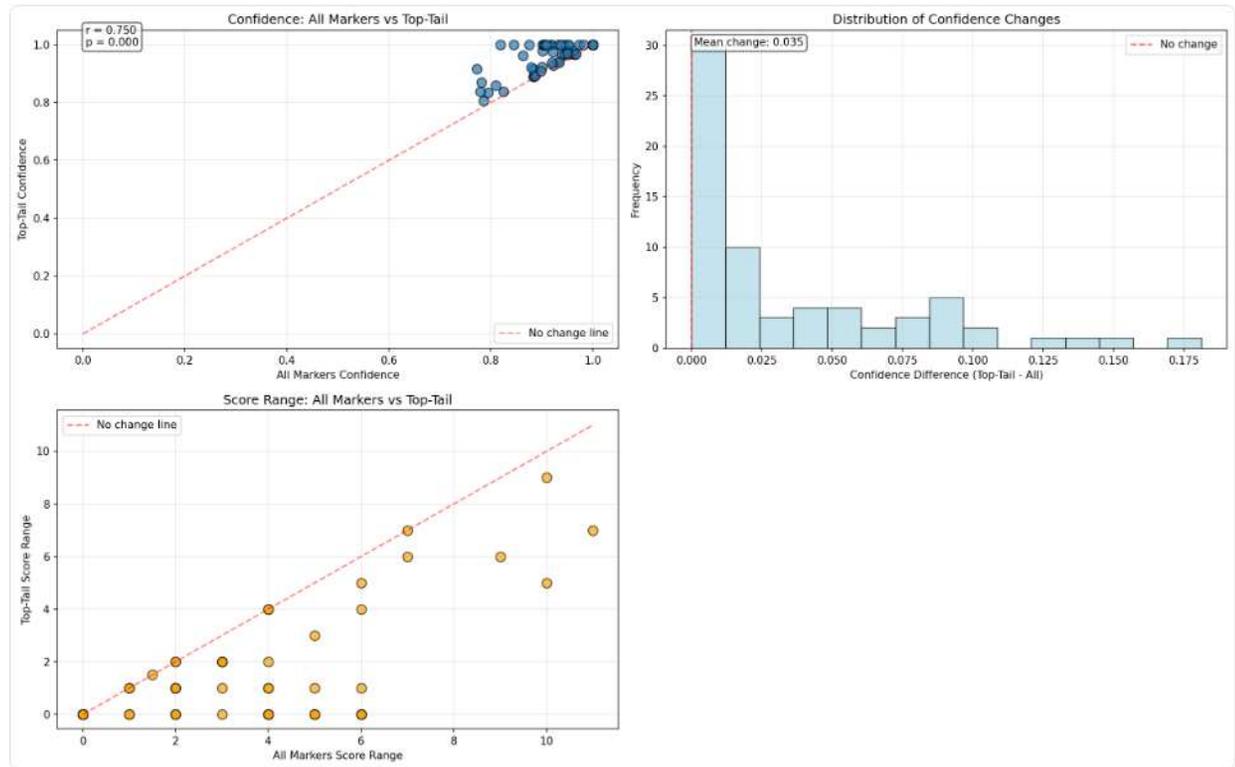
# How do we work on consistency?

## What do we assess?

- AI Marker Confidence Analysis

- Confidence vs Marker Disagreement

- Top-Tail Analysis





# What are the takeaways?

There is consistently a gap between what a human marker knows and a rubric, making three points vital:

● Feedback

● Marking Engine

● Ability to choose methods of marking

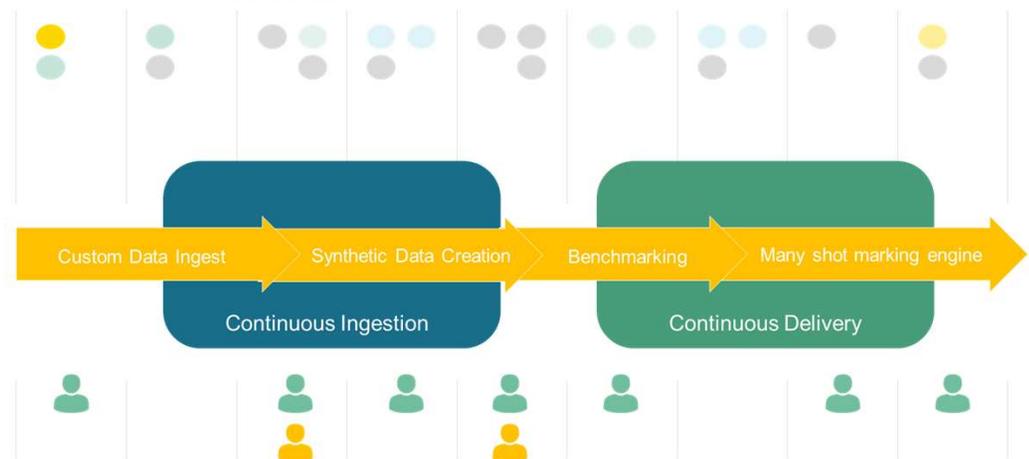
Marking Results for Question Q5

**Question:** 'At that moment she realised...' Use this as the opening words of a story.

**Marks Awarded:** 12 / 40

**Feedback**

- **A05:** The response attempts to establish a clear viewpoint and offers some coherence in structure. The setting and character emotions are described, but the vocabulary and stylistic choices are limited, affecting the overall engagement with the narrative. There is some attempt to support coherence with paragraphs and grammar, but lack of effect in language choice reflects a basic understanding of purpose and audience.  
*Marks Awarded: 8 / 24*
- **A06:** The vocabulary used is straightforward and appropriate for the narrative, but it reflects limited ambition. There are numerous spelling errors (e.g., 'beutiful', 'subcousiously', 'desparate', 'messied', 'visulise', 'immediatety'), which detract from the clarity of the piece. Simple and straightforward sentences dominate, with limited use of complex structures, leading to a basic effectiveness in conveying the narrative.  
*Marks Awarded: 4 / 16*





# Integrating a Marking Engine

