Science and Technology Facilities Council

Architecting Intelligence : *AI-Driven Decision Support for High-Energy Physics Project Management*

Sandeep Rao-Gopalam

STFC PPD – Project Office
Ai4PPD (PPD Ai Group)

Contains speculative AI imagery to illustrate evolution of the project's data journey

Overview

Business Problem

Business Problem (Challenge)

Methodology

Methodology / The AI Engine

IT Infrastructure & Optimization

IT Architecture and Optimisation

Validation & Ethics

Validating and Ethics

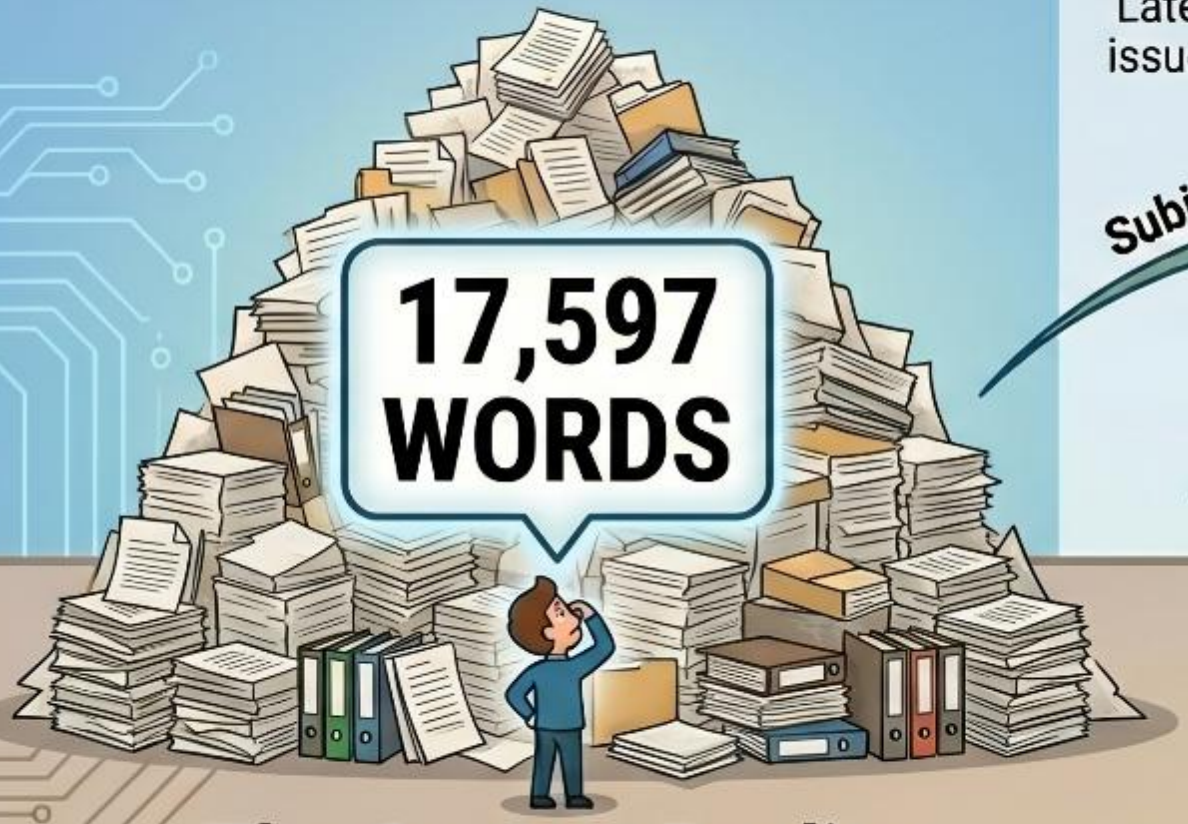
Science and Technology Facilities Council

Development Context – implementation developed as part of a Level 7 Data AI Specialist apprenticeship (Sep 2024 – Sep 2026)

Challenge

PROJECT OFFICE

Issues Management Issues Log - since 2020



The Current Reality

The Manual Mountain

THE PRIORITY GAP

Latency between logging an issue and making a decision.

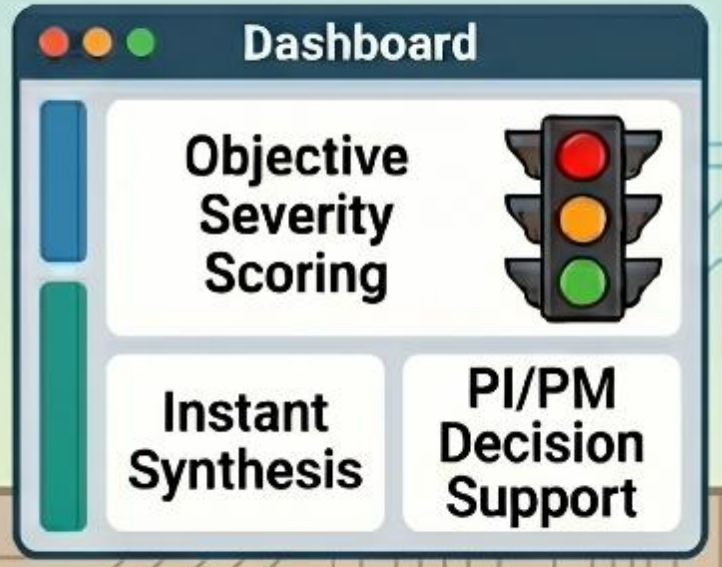
Subjectivity & Slow Triage

The Barrier

The Gap

THE GOAL

Actionable Intelligence

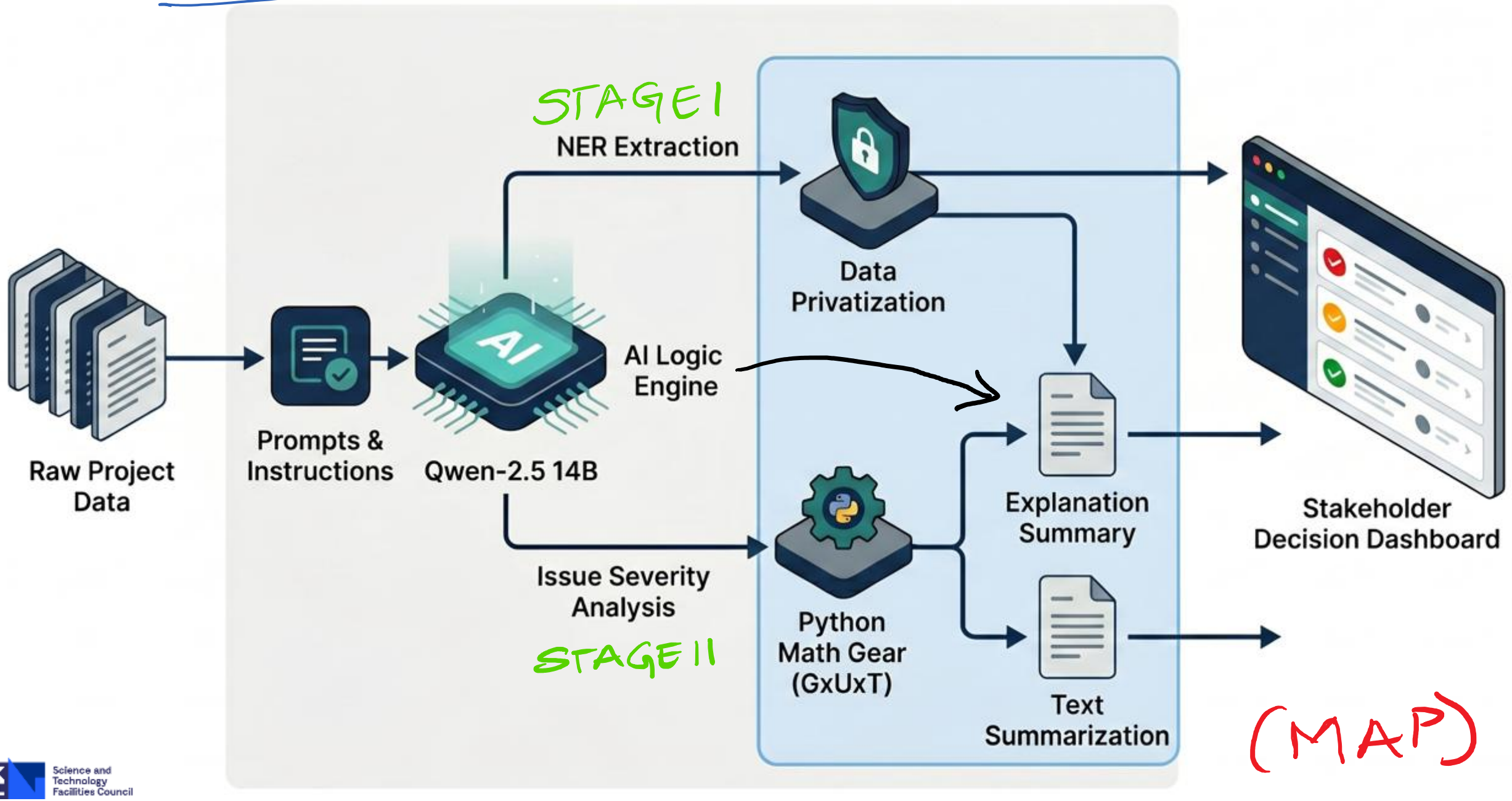


The Goal

Actionable Intelligence

Science and Technology Facilities Council

METHODOLOGY / THE AI ENGINE

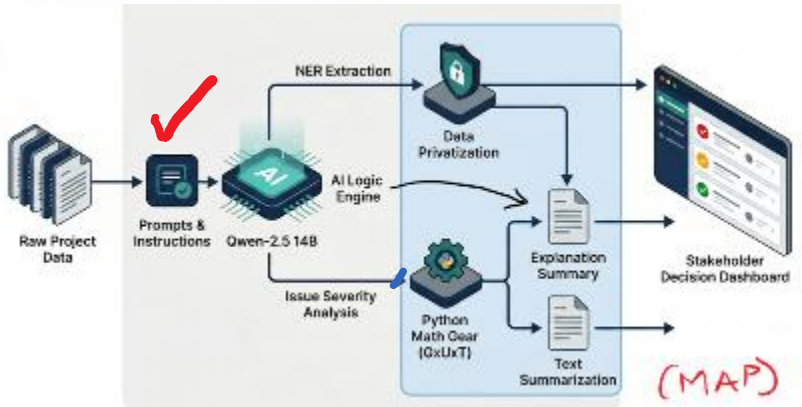


The Blueprint: PROMPT / INSTRUCTION MODULE

NER
↳ FINE TUNING

GUI
↳ INSTRUCTIONS

Science and Technology Facilities Council



(MAP)

The AI ENGINE

System Prompt
(The Job Description)

User Prompt
(The Task)

Fine Tuning

{regex

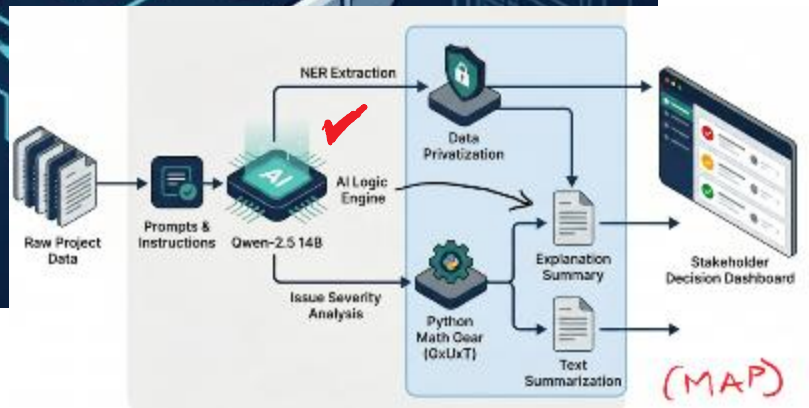
TOKENIZATION

Qwen-2.5
14B

Data Formatting Functions (Pre-processing)

Tokenization:
650k Tokens Processed

4-bit NF4 Optimization



(MAP)

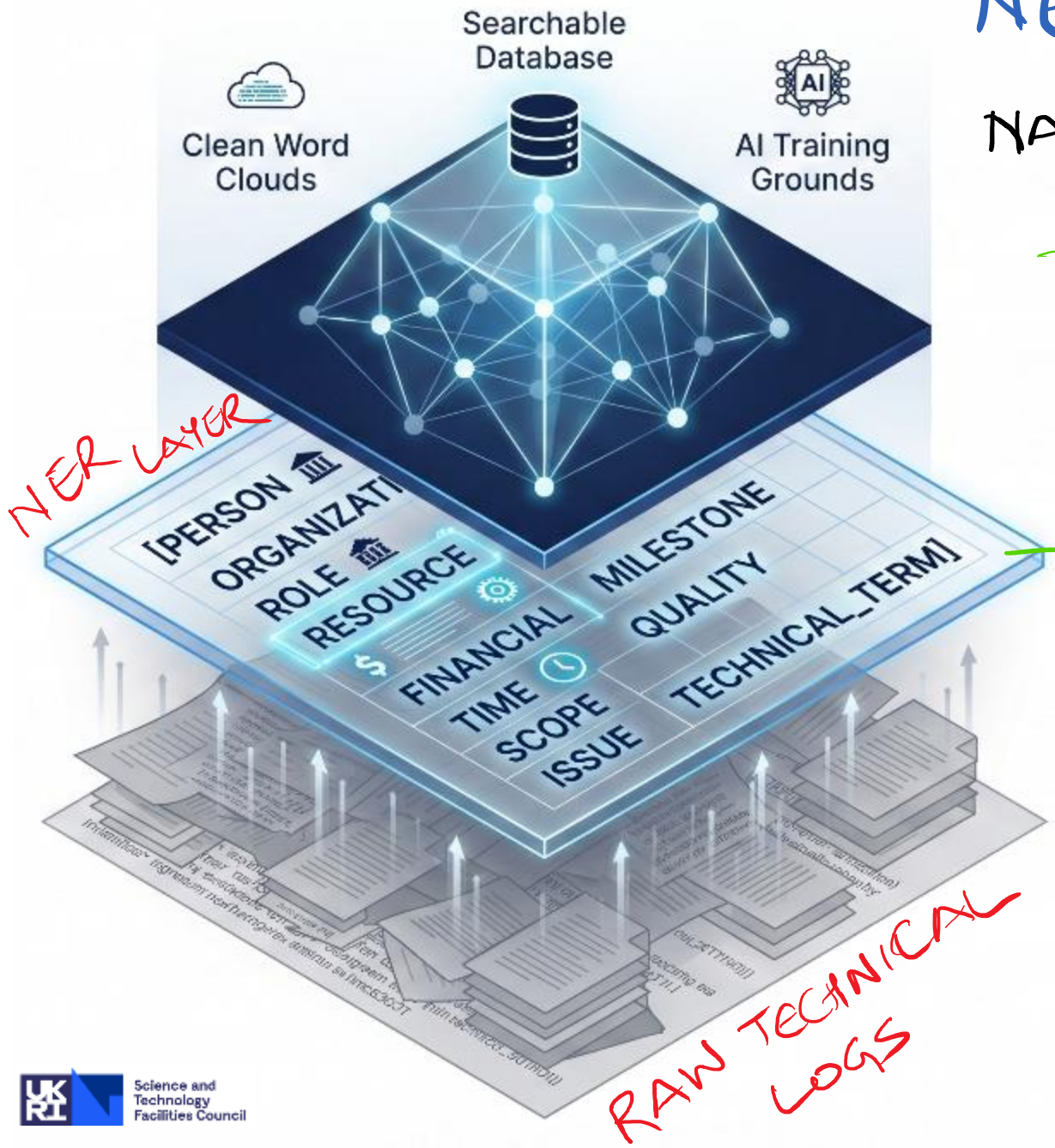
Science and Technology Facilities Council

NER EXTRACTION

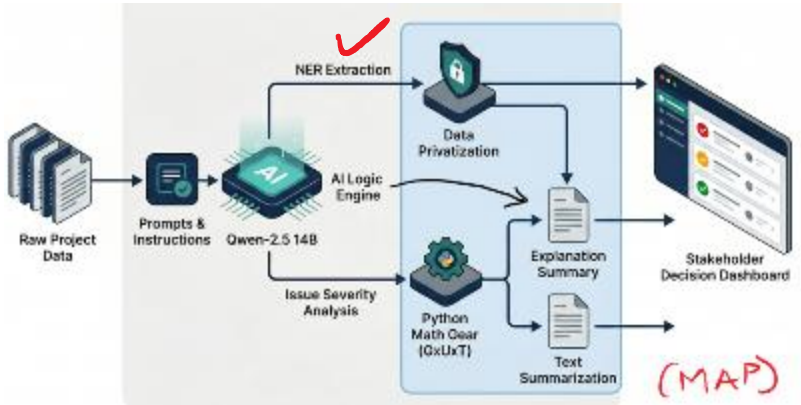
STAGE 1

NAMED ENTITY RECOGNITION

- KEY INFORMATION / NAMED ENTITIES
  - ↳ INTO DEFINED CATEGORIES



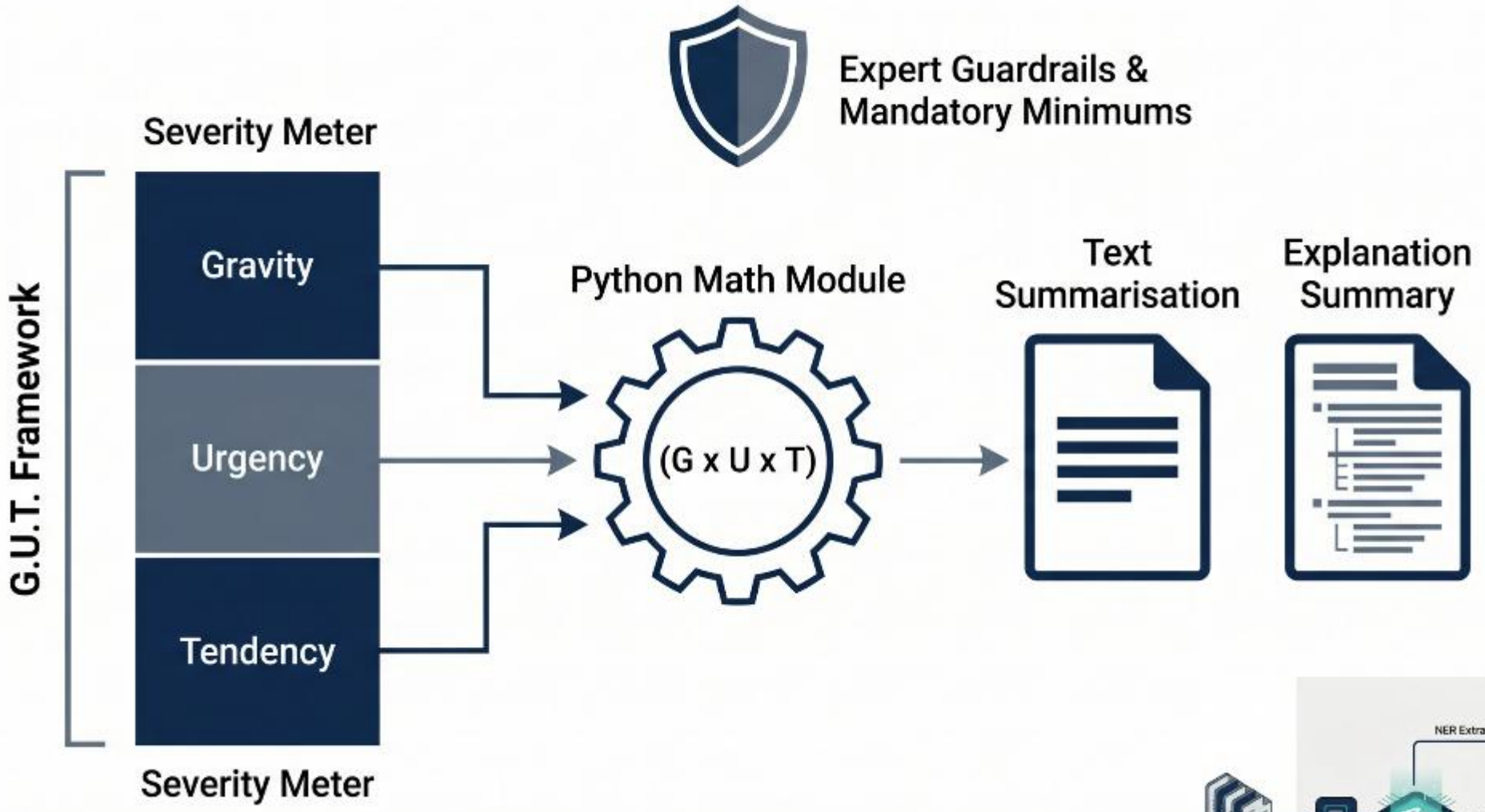
→ ELEVEN CATEGORIES (OR) ENTITES

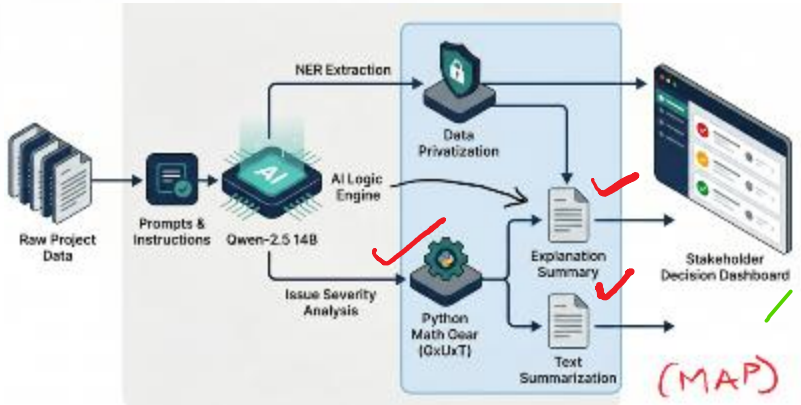


(MAP)

Issue Severity Analysis (GUT)

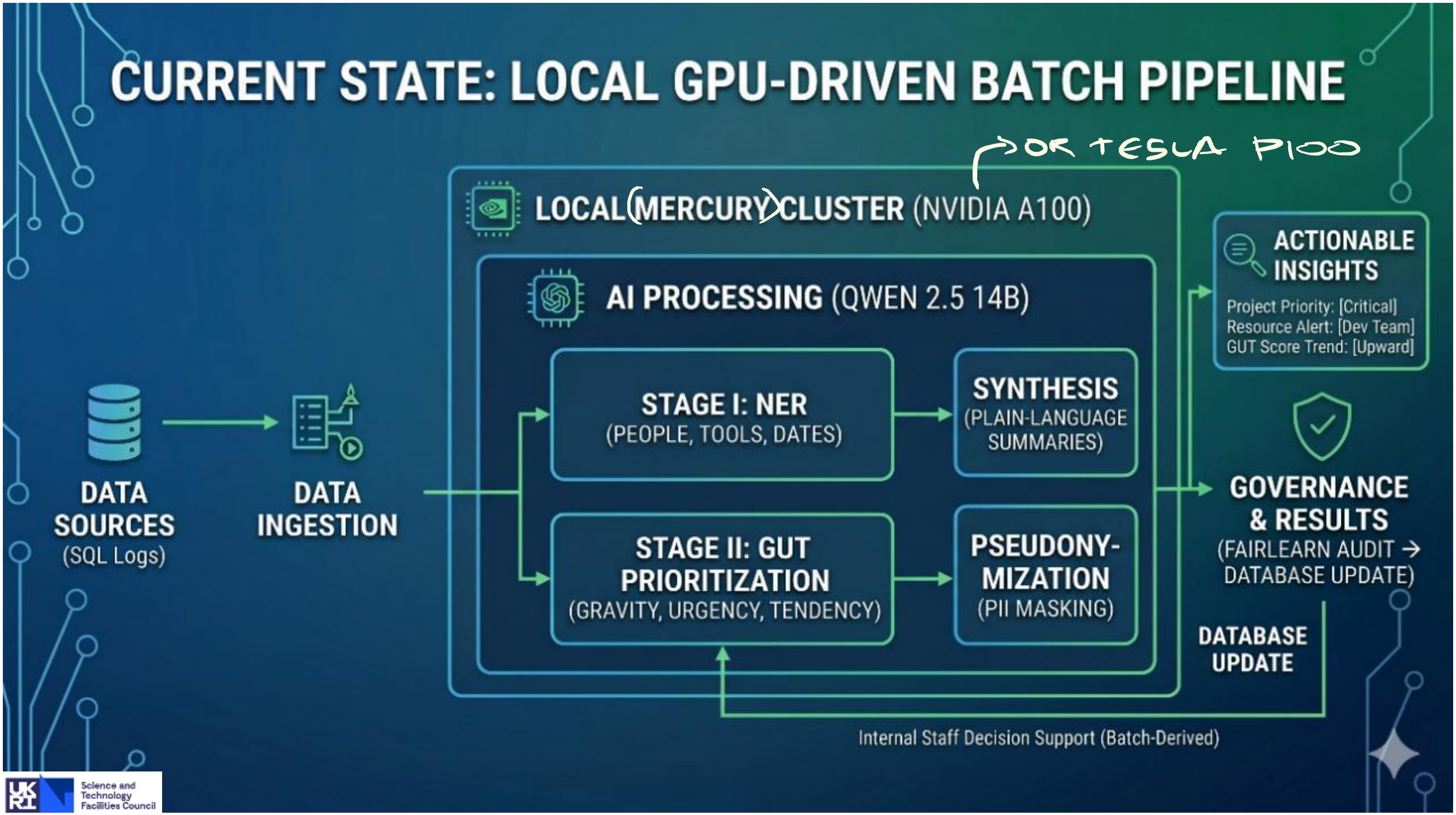
STAGE II



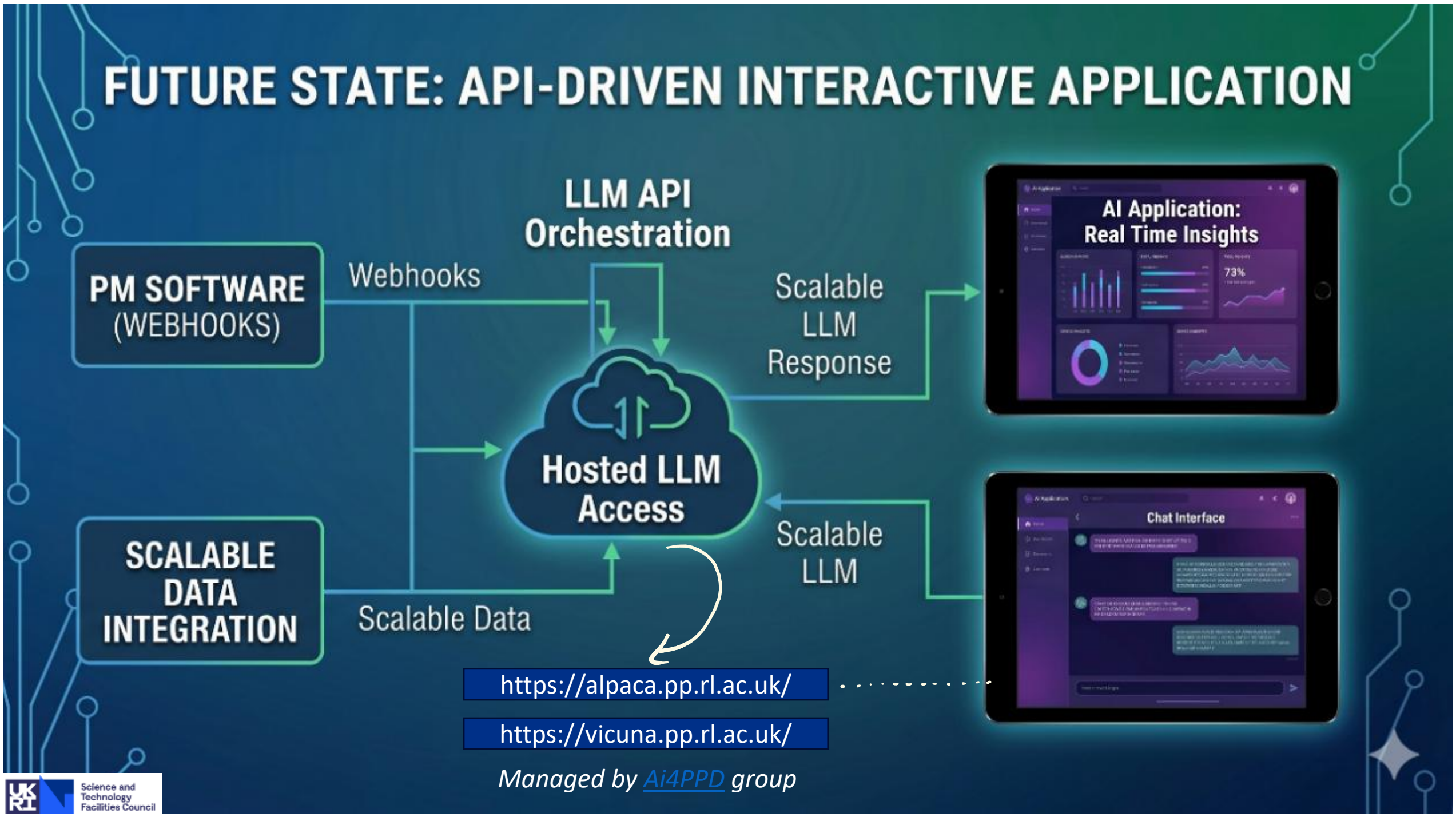


Science and Technology Facilities Council

CURRENT STATE: LOCAL GPU-DRIVEN BATCH PIPELINE



FUTURE STATE: API-DRIVEN INTERACTIVE APPLICATION



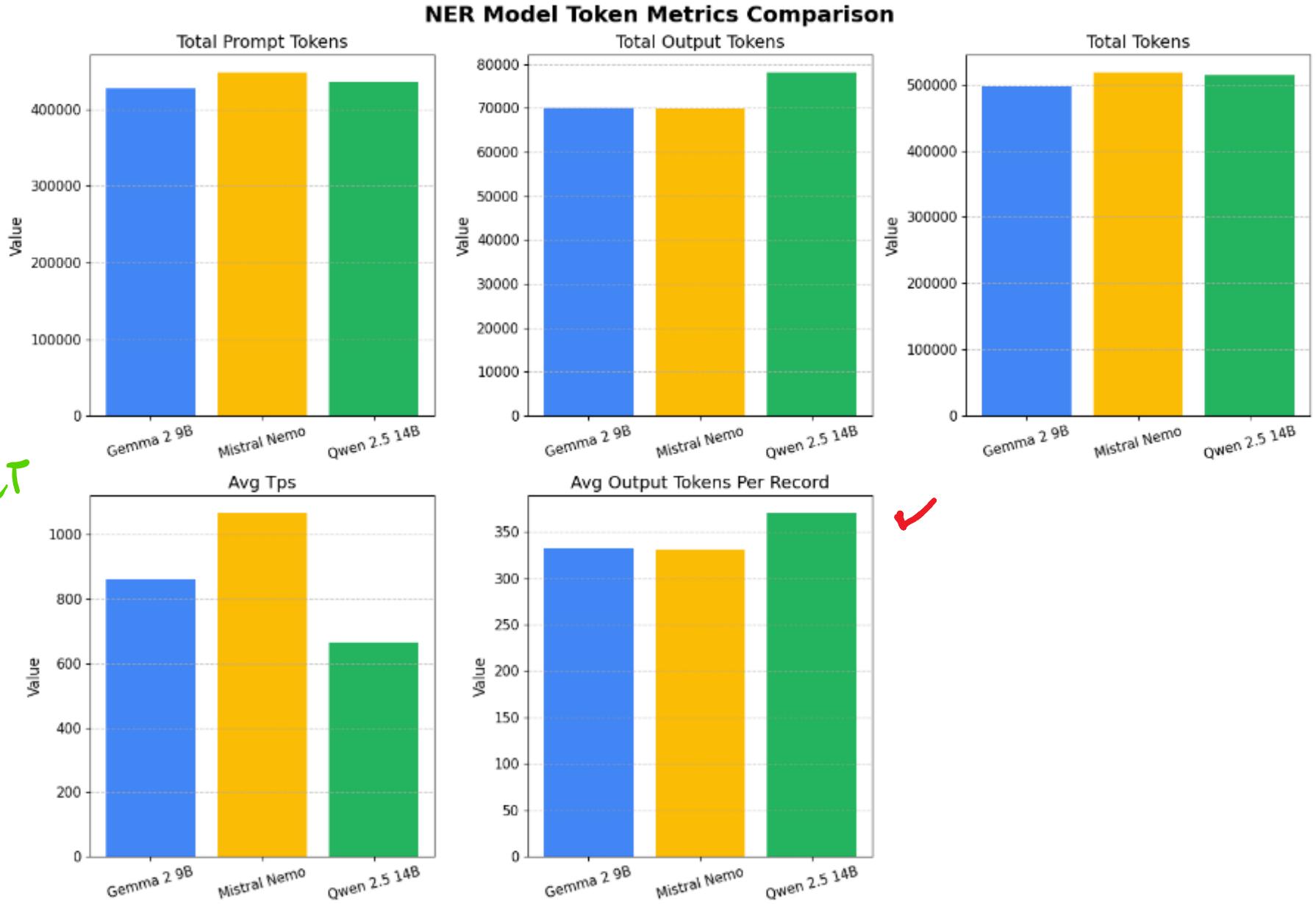
Managed by Ai4PPD group

Science and Technology Facilities Council

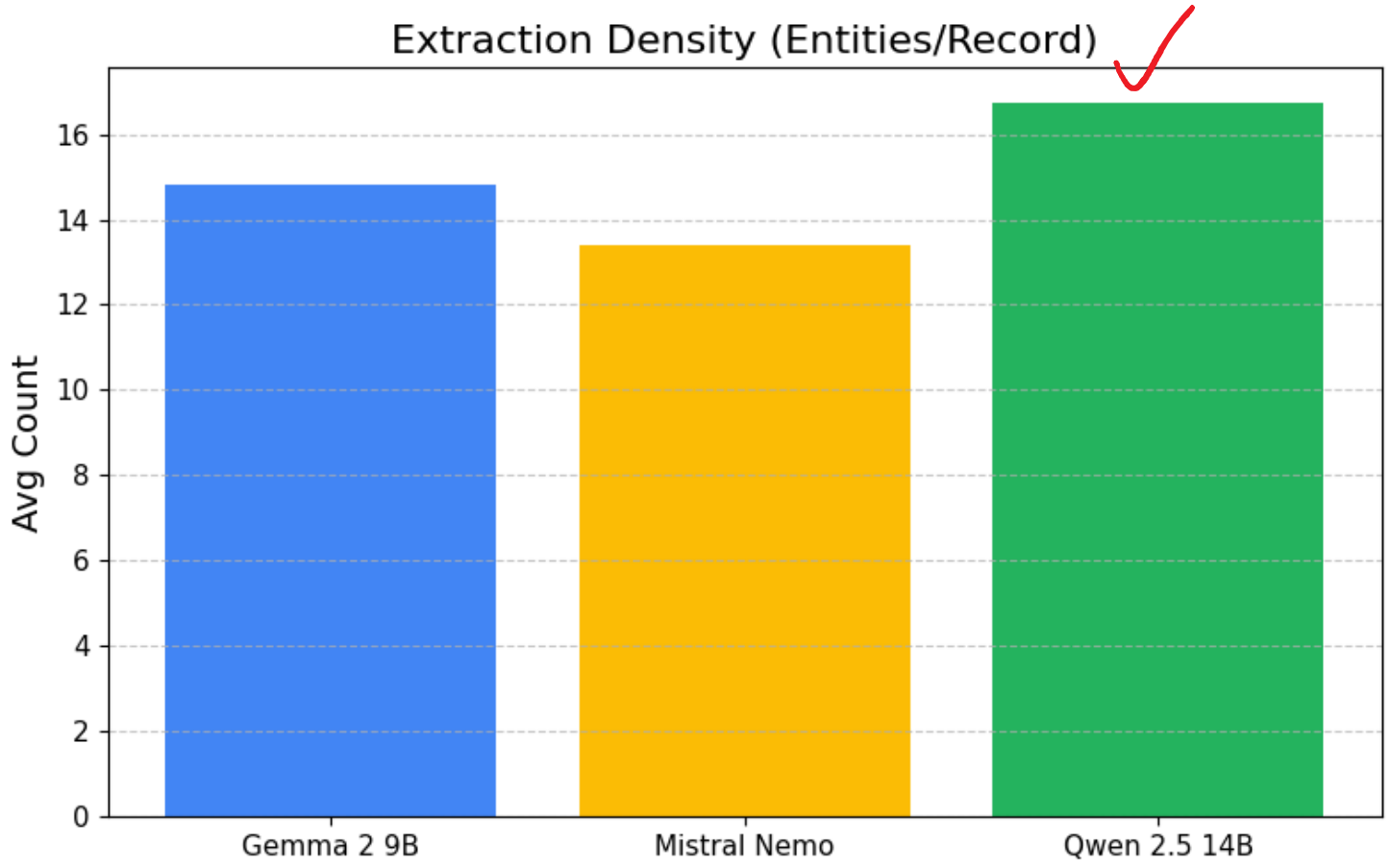
MODEL VALIDATION

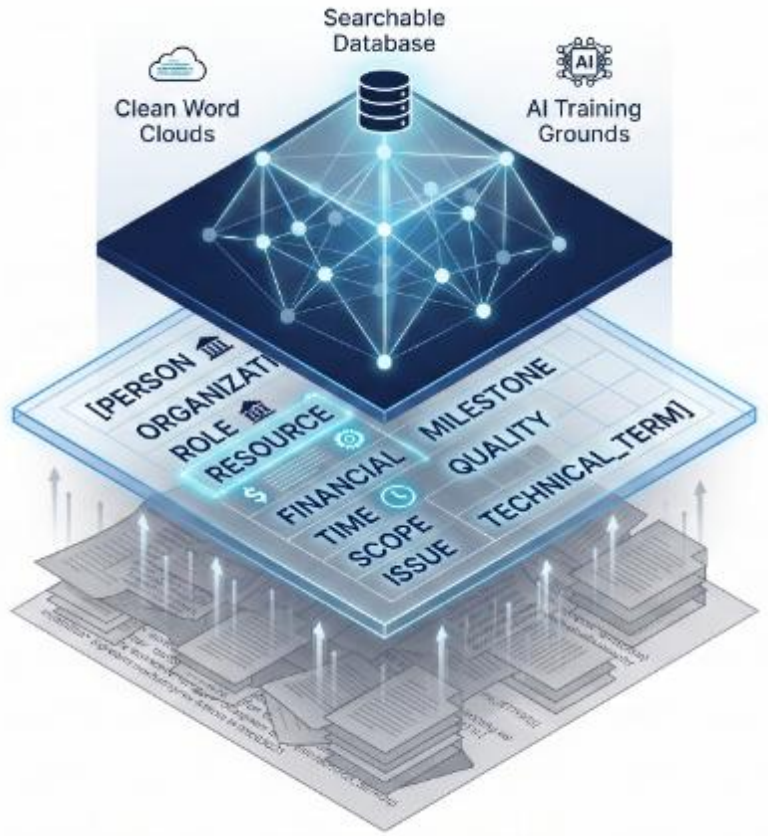
Model Names

- GEMMA 2.0 (9B)
- MISTRAL NEMO (12B)
- QWEN 2.5 (14B) INSTRUCT



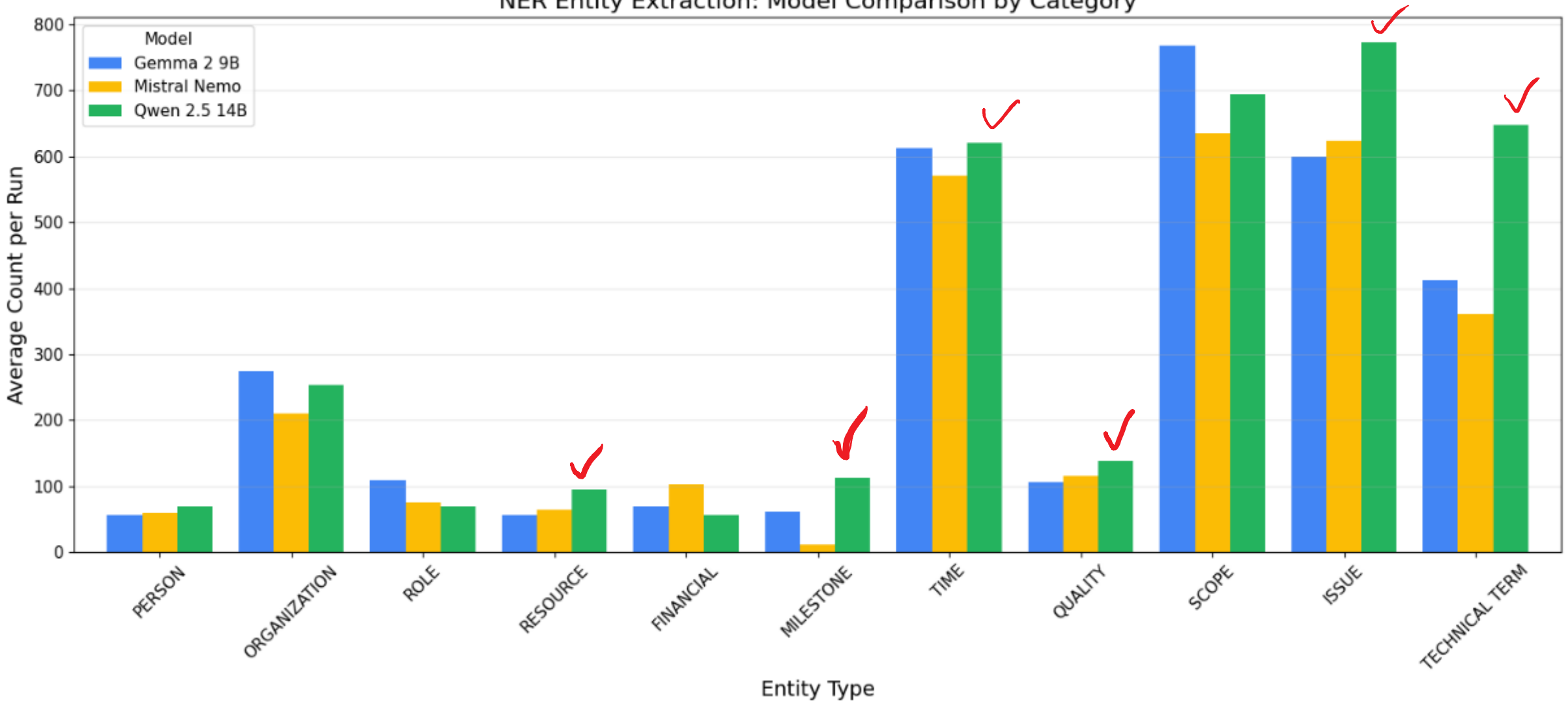
Science and Technology Facilities Council





Science and Technology Facilities Council

NER EXTRACTION



Science and Technology Facilities Council

NER EXTRACTION

PERFORMANCE METRICS QWEN 2.5

1. Run Summary

Metric	Value
Start Time	2026-04-26 10:26:59
End Time	2026-04-26 10:39:00
Duration	12.01 min
Records Processed (this run)	211
Records Skipped (incremental)	0
After AI Model Loaded - Free GPU Memory (GB)	69.33
Total Records in DB	211

2. Model Output Quality

Metric	Value	Status
JSON Parse Errors	0	OK
Parse Error Rate	0.00%	OK
Parse Success Rate	100.00%	✅

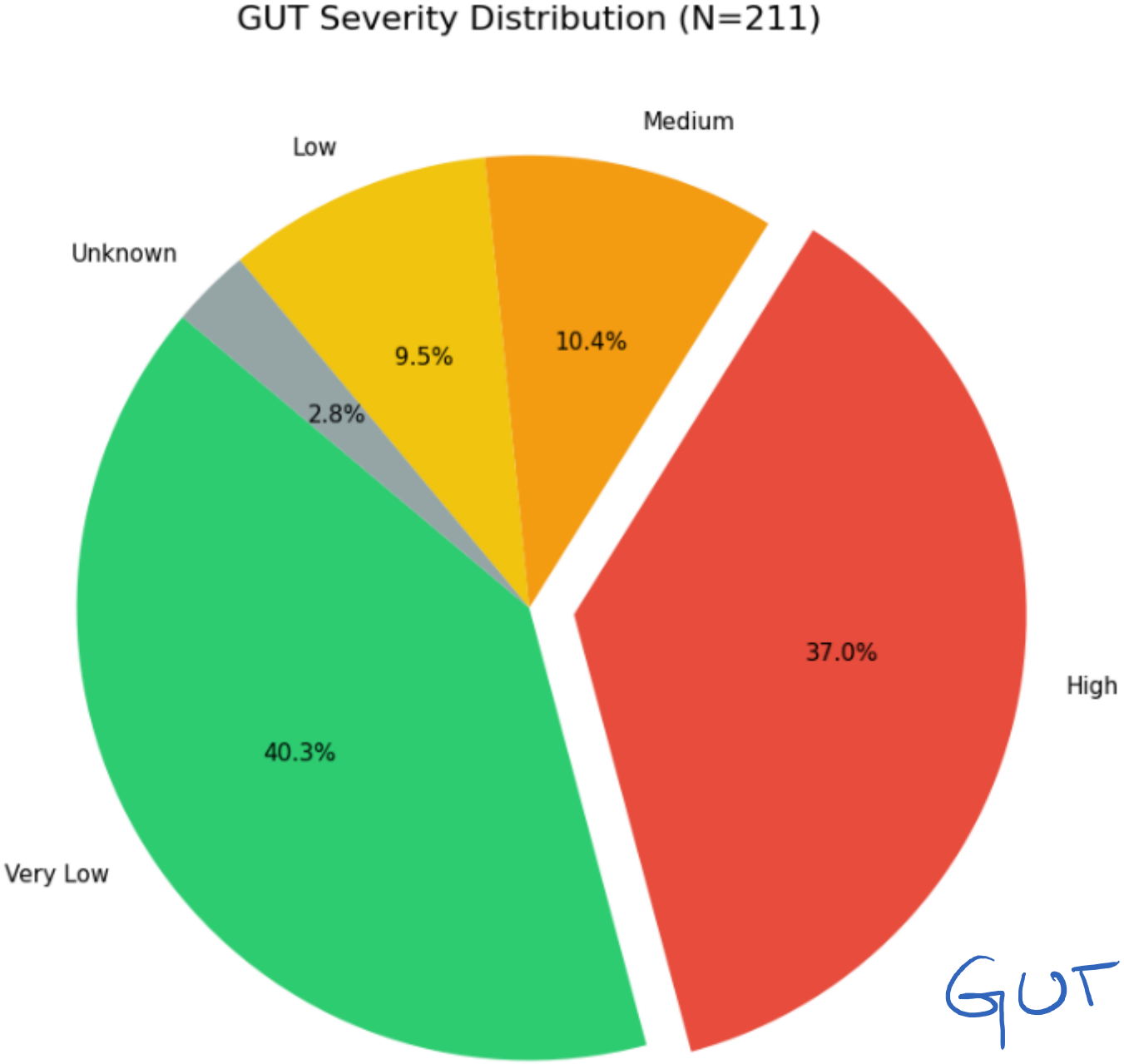
3. Throughput & Token Efficiency

Metric	Value
Total Prompt Tokens	447,760
Total Output Tokens	77,330
Total Tokens	525,090
Average TPS (tokens/sec)	728.91
Records per Minute	17.57
Avg Output Tokens per Record	366.5

7. GPU Infrastructure

Metric	Min	Avg	Max
GPU Utilisation %	0.0	90.8	100.0
Memory Utilisation %	0.0	49.6	57.0
Power Draw (W)	40.6	275.2	304.6

UKRI Science and Technology Facilities Council



GUT ANALYSIS

Science and Technology Facilities Council

PERFORMANCE METRICS QWEN-2.5 (GUT)

1. Run Summary

Metric	Value
Start Time	2026-04-26 10:40:20
End Time	2026-04-26 10:48:31
Duration	8.18 min
Records Processed	211

2. Model Output Quality

Metric	Value
Parse Success Rate	97.16%
Parse Error Rate	2.84%
Invalid JSON Rate	2.84%
Empty Output Rate	0.00%

3. Throughput & Token Efficiency

Metric	Value
Total Prompt Tokens	546,260
Total Output Tokens	104,082
Total Tokens	650,342
Average TPS (tokens/sec)	1325.51
Records per Minute	25.80

4. Severity and GUT Score Profile

Metric	Value
Mean GUT Total	115.85
Mean Gravity	4.60
Mean Urgency	3.73
Mean Tendency	3.40

Severity	Count	Percentage
Critical	0	0.00%
Very High	0	0.00%
High	78	36.97%
Medium	22	10.43%
Low	20	9.48%
Very Low	85	40.28%
Unknown	6	2.84%

UKRI Science and Technology Facilities Council

TOKEN ANALYSIS - GUT

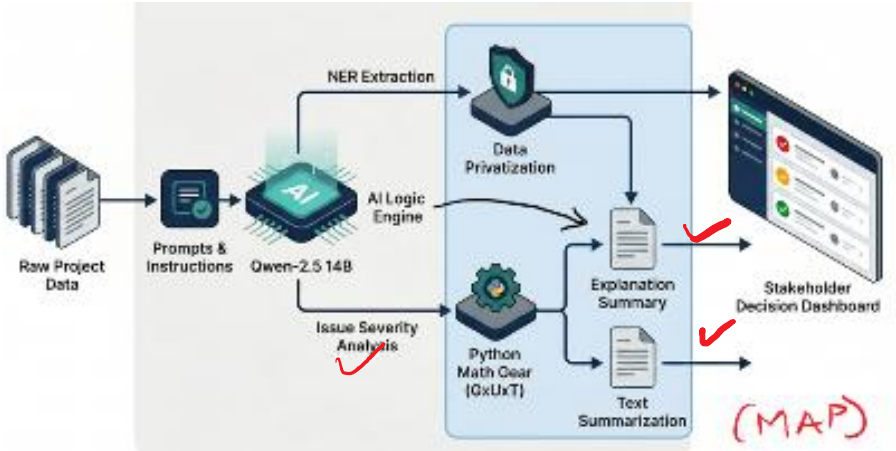
Three-Step Timing Split

Step	Total Seconds	Avg Seconds / Record
Score step (G/U/T)	58.43	0.2769
Explanation step	351.51	1.6659
Text summarization step	79.44	0.3765

~6 mins

Step-wise Token Throughput

Step	Prompt Tokens	Output Tokens	Total Tokens	Tokens / Sec
Score step (G/U/T)	80,693	44,900	125,593	2149.61
Explanation step	465,567	59,182	524,749	1492.86
Text summarization step	177,691	7,349	185,040	2329.19



Science and Technology Facilities Council

AI Ethics

- Technical Excellence is baseless if the entire system is biased
- How do we know if AI is favouring on one project than the other, during the analysis?
- How do we know if AI is hallucinating?
- Is the AI design system

  - Fair
  - Transparent
  - Safe
  - Impact on Environment
  - Accountable
  - Have data protection
  - Impact on Society

Science and Technology Facilities Council

AI Ethics – Measures (Test)

- Ensuring Human in the Loop

- Fairlearn (Open-source Python Library) – assessing the fairness of AI and ML systems.

- Data and AI Ethics Framework (Gov.uk)
  - https://www.gov.uk/government/publications/data-ethics-framework/data-and-ai-ethics-framework
  - Data and AI Ethics Self-Assessment Tool

- Fair	- Accountable
- Transparent	- Have data protection
- Safe	- Impact on Society
- Impact on Environment	

Science and Technology Facilities Council

AI Ethics – Measures Output (Test)

1. Algorithmic Fairness and Fairlearn Framework

Fairness lens: operational parity of parse success across `projectCode` groups. Groups with fewer than 20 records are excluded when Fairlearn is available.

Group	Records	Parse Success Rate
	21	100.00%
	40	100.00%
	35	100.00%
	30	93.33%

2. Fairlearn Summary Metrics

Metric	Value	Assessment
Fairlearn Enabled	Yes	computed
Sensitive Feature / Group Column	projectCode	—
Parse Success Disparity (max − min)	0.0667	moderate disparity
Demographic Parity Difference	0.066667	moderate disparity
Demographic Parity Ratio	0.933333	low disparity

AI Ethics – Measures (Test)

3. Hallucination Risk and Semantic Fidelity (Proxy Metrics)

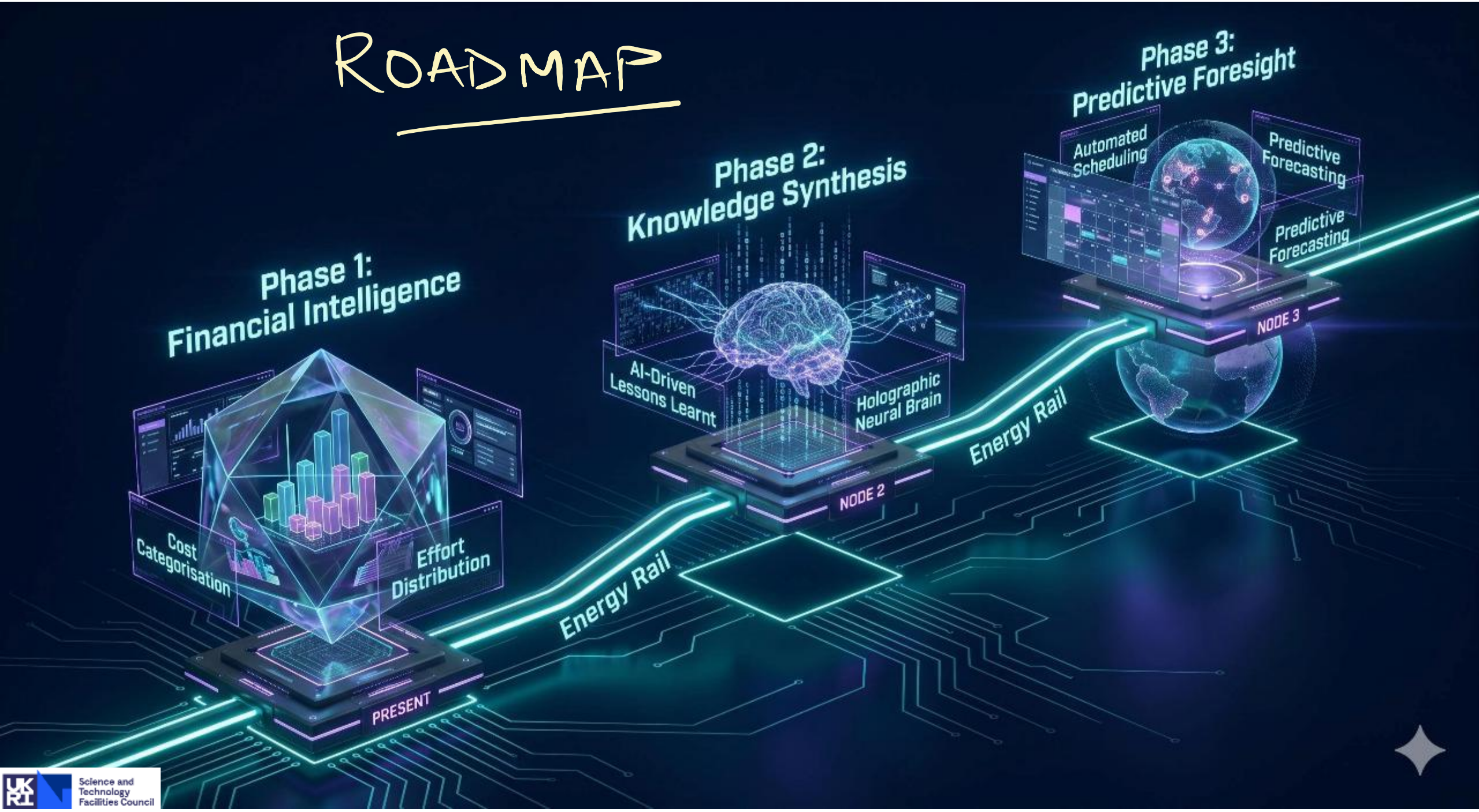
Metric	Value	Interpretation
Parse Success Rate	97.16%	Structural reliability of model output
Parse Error Rate	2.84%	JSON/schema failure risk
Invalid JSON Rate	2.84%	Hallucinated/malformed structure proxy
Empty Output Rate	0.00%	Non-response risk
Median GUT Total	60.00	Typical risk density
P95 GUT Total	336.00	High-risk tail

Current run status: **Warning**

4. Data Governance and Accountability

Governance Control	Status	
Data lineage fields captured	Yes	activityId, projectCode, a explanation_summary, s
Telemetry audit columns (parse_status/fallback/tokens/time)	Yes	Supports traceability and
Name pseudonymization mapping	Yes	
Governance posture	OK	Based on parse reliabilit

Science and Technology Facilities Council


ROADMAP
Phase 1:
Financial Intelligence
Cost Categorisation
Effort Distribution
PRESENT
Energy Rail
Phase 2:
Knowledge Synthesis
AI-Driven Lessons Learnt
Holographic Neural Brain
NODE 2
Energy Rail
Phase 3:
Predictive Foresight
Automated Scheduling
Predictive Forecasting
Predictive Forecasting
NODE 3
UKRI
Science and Technology Facilities Council

Science and Technology Facilities Council

Thank you for your time

AI Image credits – Google Gemini Nano Banana

Facebook: Science and Technology Facilities Council

Twitter:@STFC_matters

YouTube: Science and Technology Facilities Council