Investigating Science KG Validation Report

HSC Past Papers (2019-2024)

NSW Stage 6 Investigating Science Examiner Bot

2025-08-27

Table of contents

1. Executive Summary

This report details the validation and enrichment of the Investigating Science Knowledge Graph (KG). The analysis was performed by parsing the official NESA Marking Guidelines for the Higher School Certificate exams from 2019, 2020, 2021, 2022, 2023, and 2024.

The primary goal was to create explicit, evidence-based links between exam questions and the syllabus modules and outcomes they assess. The process successfully mapped 288 questions and their parts, significantly enriching the KG.

Key Findings: - **Total Questions Processed:** 288 questions/parts across 6 years. - **Coverage:** 100% of processed questions were successfully mapped to at least one Module and one Syllabus Outcome based on the explicit mapping grids in the Marking Guidelines. - **Recommendations:** The generated NDJSON patch should be ingested to enrich the KG.

2. Coverage Analysis

Coverage is 100% for all processed years. The following tables summarise the number of questions and total marks allocated per module and outcome.

Coverage by Module

Module	Title	Question Count	Total Marks
M5	Scientific Investigations	91	201
M8	Science and Society	77	151
M7	Fact or Fallacy?	64	113
M6	Technologies	56	104

Coverage by Outcome

Outcome	Question Count	Total Marks
INS12-12	63	131
INS12-15	34	82
INS12-14	40	74
INS12-13	39	74
INS12-5	38	68
INS12-7	27	68
INS12-2	27	63
INS12-6	26	58
INS12-4	28	57
INS12-3	8	27
INS12-1	8	19

3. Proposed KG Improvements

A patch file (`kg-patch.ndjson`) containing 864 new relations has been generated. This patch is non-destructive and only adds new entities (Questions/Parts) and relations.

Sanity Checks (for `kg-patch.ndjson`)

- **Total Relations:** `jq -s 'length' out/proposals/kg-patch.ndjson` (Expected: 864)
- `ASSESSES_OUTCOME` **Relations:** `jq -s 'map(select(.type=="relation" and .predicate=="ASSESSES_OUTCOME")) | length' out/proposals/kg-patch.ndjson` (Expected: 326)
- `ASSESSES_MODULE` **Relations:** `jq -s 'map(select(.type=="relation" and .predicate=="ASSESSES_MODULE")) | length' out/proposals/kg-patch.ndjson` (Expected: 288)

4. Next Steps

1. **Review and Apply Patch:** A human operator should review the generated `kg-patch.ndjson` and `surreal-import.surql` files before applying them to the production knowledge graph.
2. **Update Quarto Page:** Integrate the `exam-mapping-quarto-update.md` snippet into the main `exam-mapping.qmd` file to reflect the new data.