Healthcare Recipe Quality Validation Report

This report supports the LLM-based system for auditing healthcare manufacturing recipes. It includes: - Realistic JSON data simulating recipe entries - Subtle, common data quality deviations - Snippets grouped by deviation type - Testing strategy overview The goal is to evaluate how well language models detect practical deviations in real-world formats.

Table of Contents

1. Project Overview

This validation effort is designed to test how well language models can identify deviations in structured healthcare manufacturing data. The test set mimics real-world complexity, noise, and variation.

2. Testing Strategy

We use a realistic JSON file with 150 recipe entries and intentionally embedded subtle issues. These include missing data, slight format inconsistencies, and non-standard entries. The system checks whether LLMs can detect and report on each of these.

3.1 Missing Values

{

  "id": "REC-1000",

  "material_code": "MATCMN",

  "process_step": "filtering",

  "quantity_kg": NaN,

  "start_time": "2025-04-30 11:30:31.769329",

  "end_time": "2025-04-30T12:51:31.769329",

  "status": "done",

  "notes": "Next realize goal opportunity.",

  "operator": NaN

}

{

  "id": "REC-1017",

  "material_code": "MATIVt",

  "process_step": "cooling",

  "quantity_kg": NaN,

  "start_time": "2025-04-28T01:30:31.773511",

  "end_time": "2025-04-28T01:59:31.773511",

  "status": "completed",

  "notes": "Event country red appear above present.",

  "operator": "Jeremy Nelson"

}

{

  "id": "REC-1034",

  "material_code": "MATRDQ",

```
  "process_step": "filtering",

  "quantity_kg": NaN,

  "start_time": "2025-05-31T09:30:31.777381",

  "end_time": "2025-05-31T09:40:31.777381",

  "status": "completed",

  "notes": "",

  "operator": "Chad Kerr"
}
```

3.2 Missing Keys

{

  "id": "REC-1000",

  "material_code": "MATCMN",

  "process_step": "filtering",

  "quantity_kg": null,

  "start_time": "2025-04-30 11:30:31.769329",

  "end_time": "2025-04-30T12:51:31.769329",

  "status": "done",

  "notes": "Next realize goal opportunity."

}

{

  "id": "REC-1027",

  "material_code": "MATuVB",

  "process_step": "mixing",

  "quantity_kg": 10.4,

  "start_time": "2025-06-17T07:30:31.775793",

  "end_time": "2025-06-17T09:19:31.775793",

  "status": "completed",

  "notes": "Scene article measure sort big open two."

}

{

  "id": "REC-1054",

  "material_code": "MATbJj",

  "process_step": "filtering",

  "quantity_kg": 41.45,

"start_time": "2025-04-22T06:30:31.782358",

"end_time": "2025-04-22T07:41:31.782358",

"status": "completed",

"notes": "Each imagine listen within do community."

}

3.3 Format Inconsistencies

```
{

  "id": "REC-1000",

  "material_code": "MATCMN",

  "process_step": "filtering",

  "quantity_kg": NaN,

  "start_time": "2025-04-30 11:30:31.769329",

  "end_time": "2025-04-30T12:51:31.769329",

  "status": "done",

  "notes": "Next realize goal opportunity.",

  "operator": NaN

}


{

  "id": "REC-1023",

  "material_code": "MATIct",

  "process_step": "cooling",

  "quantity_kg": 23.56,

  "start_time": "2025-04-28 03:30:31.774742",

  "end_time": "2025-04-28T05:26:31.774742",

  "status": "in progress",

  "notes": "Star born area bring itself lay.",

  "operator": "Melissa Yoder"

}


{

  "id": "REC-1046",

  "material_code": "MATGTa",
```

"process_step": "filling",

"quantity_kg": 21.31,

"start_time": "2025-05-05 11:30:31.780509",

"end_time": "2025-05-05T11:59:31.780509",

"status": "completed",

"notes": "",

"operator": "Amy Carroll"

}

3.4 Unstandardized Fields

{

  "id": "REC-1000",

  "material_code": "MATCMN",

  "process_step": "filtering",

  "quantity_kg": NaN,

  "start_time": "2025-04-30 11:30:31.769329",

  "end_time": "2025-04-30T12:51:31.769329",

  "status": "done",

  "notes": "Next realize goal opportunity.",

  "operator": NaN

}

{

  "id": "REC-1031",

  "material_code": "MATjlq",

  "process_step": "filling",

  "quantity_kg": 93.53,

  "start_time": "2025-05-24T06:30:31.776658",

  "end_time": "2025-05-24T07:46:31.776658",

  "status": "done",

  "notes": "Benefit run either thousand during coach many.",

  "operator": "Diane Zimmerman"

}

{

  "id": "REC-1062",

  "material_code": "MATqjq",

    "process_step": "heating",

    "quantity_kg": 36.05,

    "start_time": "2025-04-26T11:30:31.784192",

    "end_time": "2025-04-26T13:07:31.784192",

    "status": "done",

    "notes": "West determine leave buy dream school.",

    "operator": "Robert Charles"

}