**1. Overview**

Australian general practitioners (GPs) bill Medicare using Medical Benefits Schedule (MBS) item numbers. Choosing the correct item number depends on service type, consult duration, time of day (after‑hours vs in‑hours), and co‑billing restrictions.  
This project delivers an AI‑powered service that ingests a transcript of a doctor–patient consultation plus metadata (date, start & end times, duration). It returns **a deterministic, repeatable list of valid MBS item numbers with short, human‑readable rationales**.

**2. Goals & Objectives**

* **Accuracy ≥ 93 %** on a curated gold‑standard test set of GP encounters.
* **Determinism** – identical inputs always yield identical item numbers.

|  |  |
| --- | --- |
|  | **Requirement** |
| F‑1 | Accept JSON payload with: transcript, consultDateTime, durationSeconds. |
| F‑2 | Return array of objects {itemNumber, rationaleText} for **GP‑only** MBS items. |
| F‑3 | Never suggest pathology, radiology, or other non‑GP specialties. |
| F‑4 | Apply co‑billing rules – exclude invalid combinations; suggest alternates when conflicts detected. |
| F‑5 | Account for after‑hours rules based on consultDateTime. |

### High‑Level Architecture

1. **Retrieval & Knowledge Layer**
   * Azure AI Search index loaded nightly with:
     + data from https://www.mbsonline.gov.au/ (provided as XML) => filtered to GP items.
     + Co‑billing matrix curated from Medicare guidelines.
   * **Knowledge‑Graph Agent** (see KGAREVION paper) for complex queries.
2. **Reasoning Service**
   * Orchestrates an **OpenAI GPT‑4o (or latest)** LLM chain.
   * RAG prompt supplies: top‑k MBS entries, co‑billing rules, and consult metadata.
   * Deterministic decoding (temperature = 0, top\_p = 1) guarantees repeatability.
3. **Rule‑Engine Micro‑service** – Post‑LLM deterministic validator (Drools/JsonLogic) enforces:
   * Co‑billing bans
   * After‑hours eligibility

### Minimum/maximum duration thresholds Acceptance Criteria (Traceability)

|  |  |
| --- | --- |
| Req | Test |
| F‑1 | Submit valid/invalid payloads → expect 200 / 400 codes. |
| F‑2 | Known scenario → expected item numbers exactly. |
| F‑4 | Scenario with illegal pair → engine returns single valid code or error. |
| F‑5 | After‑hours transcript 19:00 Sat → returns 5020 not 23. |

Check references for context.

[https://arxiv.org/pdf/2410.04660](https://www.youtube.com/redirect?event=video_description&redir_token=QUFFLUhqblpkeXVrSl9OR09SUUFmWFZMVFdBTl93VGd3Z3xBQ3Jtc0trU0xMejBHMWZmZTUzOVFQcEhHcXlBbElHTkVWQ1RqeGs5TjRzZDh2YmxhV3ppWkxUNGhqbU9EN2VMWGdjN01Uajd2akNFbUJQY0I3R1FqbWtnVjNtdENvX0hsQ0o4WHk4RDVwc2RZTEdYejkzZ0luYw&q=https%3A%2F%2Farxiv.org%2Fpdf%2F2410.04660&v=Fm68I-phaiY)   
<https://www.youtube.com/watch?v=Fm68I-phaiY>

### References

1. Australian Department of Health. MBS Online https://www.mbsonline.gov.au/
2. G. Shin et al. “KGAREVION: An AI Agent for Knowledge‑Intensive Biomedical QA.” arXiv:2410.04660 (2024).
3. Microsoft Azure. Azure AI Search documentation.
4. YouTube. Architecting production RAG systems, Oct 2024. https://www.youtube.com/watch?v=Fm68I-phaiY