

# Orinoco Retail Holdings plc 2027 — JPY 2035 Notes

CDSBench #39 · Which section of the record establishes the seniority (ranking) of the instrument? Answer with the section identifier (e.g. DOC-1#S4).

Orinoco Retail Holdings plc · Corporate · Failure to pay 2027-04-10 · Fact extraction · Evidence · Ground truth: deterministically verified

VERIFIED OUTCOME

**DOC-1#S4**

The Status condition states the ranking.

Supporting sections: DOC-1#S4

WHAT HAPPENED

**GPT and Gemini got this evidence question wrong.**

GPT did not retrieve the relevant section: None of the cited sections is among those that establish the answer; the relevant evidence was not retrieved.

MODEL COMPARISON

| PARTICIPANT            | ANSWER   | RESULT    | CONFIDENCE | EVIDENCE SCORE | PRIMARY FAILURE   |
|------------------------|----------|-----------|------------|----------------|-------------------|
| Claude                 | DOC-1#S4 | Correct   | 75%        | 100%           | —                 |
| GPT                    | DOC-1#S1 | Incorrect | 92%        | 0%             | Retrieval failure |
| Gemini                 | DOC-1#S1 | Incorrect | 85%        | 0%             | Retrieval failure |
| Mastermind (reference) | DOC-1#S4 | Correct   | 75%        | 100%           | —                 |

GPT

**GPT: Retrieval failure**

None of the cited sections is among those that establish the answer; the relevant evidence was not retrieved.

Model summary: Section DOC-1#S1 establishes the seniority.

Cited: DOC-1#S1

GEMINI

**Gemini: Retrieval failure**

None of the cited sections is among those that establish the answer; the relevant evidence was not retrieved.

Model summary: Section DOC-1#S1 establishes the seniority.

Cited: DOC-1#S1

EVIDENCE

**Decisive sections of the record**

- DOC-1#S4 — JPY 2035 Notes — Terms and Conditions (excerpt), Status: The JPY 2035 Notes constitute direct, unsecured and subordinated obligations of the Issuer and rank junior to all senior indebtedness of the Issuer.

METHODOLOGY

A response is correct when its normalised answer equals the verified label. Evidence score is the F1 between cited section IDs and the supporting sections; cited IDs that do not exist count as hallucinated. Abstentions are correct only when the label calls for them. Ground truth is verified by two independent extractors plus an evidence check; no model grades itself.

Full Case Replay with documents and every model's reasoning: <https://cdsbench.com/cases/orinoco-retail-holdings-plc-2027/jpy-2035/seniority-evidence>