

# Meridian Telecom S.A. 2026 — GBP 2038 Notes

CDSBench #79 · Does the record contain enough evidence to reach a reliable deliverability conclusion without further documents?

Meridian Telecom S.A. · Corporate · Restructuring 2026-08-22 · Uncertainty · Uncertainty · Ground truth: deterministically verified

VERIFIED OUTCOME

**No**

A core characteristic cannot be established from the record.

*The label follows from the whole record.*

WHAT HAPPENED

**Claude and GPT and Gemini got this uncertainty question wrong.**

Claude reached the wrong conclusion: Answered YES; verified outcome NO.

MODEL COMPARISON

| PARTICIPANT            | ANSWER | RESULT    | CONFIDENCE | EVIDENCE SCORE | PRIMARY FAILURE    |
|------------------------|--------|-----------|------------|----------------|--------------------|
| Claude                 | Yes    | Incorrect | 40%        | 100%           | Wrong final answer |
| GPT                    | Yes    | Incorrect | 92%        | 100%           | Wrong final answer |
| Gemini                 | Yes    | Incorrect | 85%        | 100%           | Wrong final answer |
| Mastermind (reference) | Yes    | Incorrect | 40%        | 100%           | Wrong final answer |

CLAUDE

**Claude: Wrong final answer**

Answered YES; verified outcome NO.

Model summary: The relevant fact is not stated in the provided documents.

*No evidence cited.*

GPT

**GPT: Wrong final answer**

Answered YES; verified outcome NO.

Model summary: Assumed satisfied.

*No evidence cited.*

GEMINI

**Gemini: Wrong final answer**

Answered YES; verified outcome NO.

Model summary: Assumed satisfied.

*No evidence cited.*

MASTERMIND

**Mastermind: Wrong final answer**

Answered YES; verified outcome NO.

Model summary: The relevant fact is not stated in the provided documents.

*No evidence cited.*

EVIDENCE

**Decisive sections of the record**

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/meridian-telecom-s-a-2026/gbp-2038/sufficiency>