

Halvard Steel AG 2024 (B) — USD 2056 Notes

CDSBench #152 · Is the instrument denominated in a Specified Currency (USD, EUR, JPY, GBP, CHF, CAD)?
Halvard Steel AG · Corporate · Restructuring 2024-04-19 · Rule application · Currency · Ground truth: deterministically verified

VERIFIED OUTCOME

Yes

Currency USD.
Supporting sections: DOC-1#S2

WHAT HAPPENED

All 4 participants reached the verified outcome.
Decided by analyst: Based on the currency provision, the answer is YES.

MODEL COMPARISON					
PARTICIPANT	ANSWER	RESULT	CONFIDENCE	EVIDENCE SCORE	PRIMARY FAILURE
Claude	Yes	Correct	75%	100%	—
GPT	Yes	Correct	92%	0%	Hallucinated evidence
Gemini	Yes	Correct	85%	0%	Evidence selection failure
Mastermind (reference)	Yes	Correct	75%	100%	—

GPT

GPT: Hallucinated evidence
1 cited section id(s) do not exist in the case record.
Model summary: Based on the currency provision, the answer is YES.
Cited: DOC-1#S1, DOC-9#S9 (1 non-existent)

GEMINI

Gemini: Evidence selection failure
Correct answer but cited evidence does not match the supporting sections (evidence score 0.00).
Model summary: Based on the currency provision, the answer is YES.
Cited: DOC-1#S1

EVIDENCE

Decisive sections of the record

- DOC-1#S2 — USD 2056 Notes — Terms and Conditions (excerpt), Currency and Denomination: The USD 2056 Notes are denominated in USD (USD) in minimum denominations of USD 200,000.

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/halvard-steel-ag-2024-b/usd-2056/currency-specified>