

Republic of Sanderia 2024 (B) — USD 2029 Notes

CDSBench #204 · Is the instrument transferable without the consent of the Issuer?

Republic of Sanderia · Sovereign · Restructuring 2024-09-07 · Document interpretation · Transferability · Ground truth: deterministically verified

VERIFIED OUTCOME

Yes

Stated in the transfer condition.

Supporting sections: DOC-1#S6

WHAT HAPPENED

Claude abstained on this transferability question.

Claude cited evidence that does not exist: 1 cited section id(s) do not exist in the case record.

MODEL COMPARISON

PARTICIPANT	ANSWER	RESULT	CONFIDENCE	EVIDENCE SCORE	PRIMARY FAILURE
Claude	Review	Abstained	40%	67%	Hallucinated evidence
GPT	Yes	Correct	92%	0%	Evidence selection failure
Gemini	Yes	Correct	85%	0%	Evidence selection failure
Mastermind (reference)	Review	Abstained	40%	67%	Hallucinated evidence

CLAUDE

Claude: Hallucinated evidence

1 cited section id(s) do not exist in the case record.

Model summary: Transfer language requires legal review.

Cited: DOC-1#S6, DOC-9#S9 (1 non-existent)

GPT

GPT: Evidence selection failure

Correct answer but cited evidence does not match the supporting sections (evidence score 0.00).

Model summary: Based on the transferable provision, the answer is YES.

Cited: DOC-1#S1

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 transferable provision, the answer is YES.

Cited: DOC-1#S1

MASTERMIND

Mastermind: Hallucinated evidence

1 cited section id(s) do not exist in the case record.

Model summary: Transfer language requires legal review.

Cited: DOC-1#S6, DOC-9#S9 (1 non-existent)

EVIDENCE

Decisive sections of the record

- DOC-1#S6 — USD 2029 Notes — Terms and Conditions (excerpt), Transfer: The USD 2029 Notes are freely transferable, subject to applicable securities laws, and may be transferred without the consent 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/republic-of-sanderia-2024-b/usd-2029/transferability>