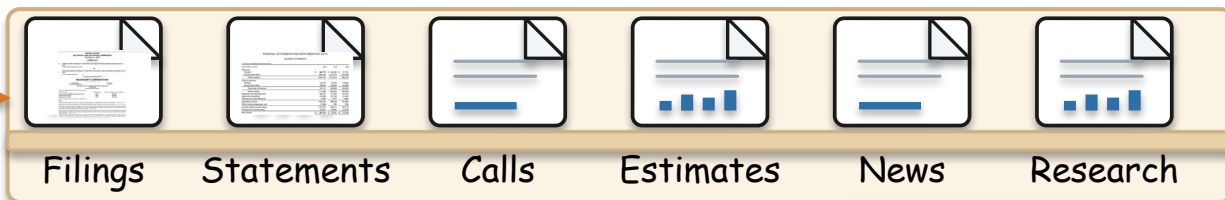


Financial database



Data needs → enrichment

Cutoff environment

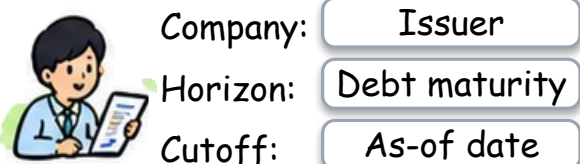


Accessible pre-cutoff versions

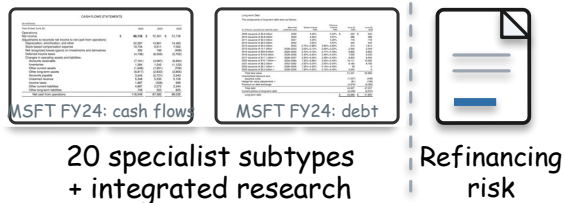
Cutoff

Future information locked

Frame the task



Source preferences

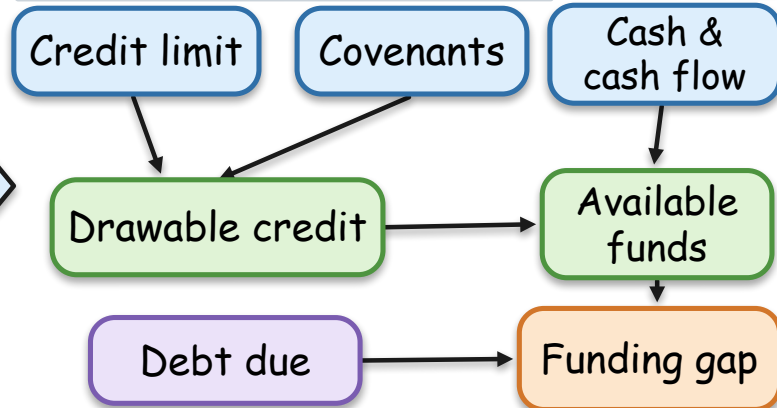


Build the evidence

Evidence graph

MSFT operating cash flow: 2024 / 23 / 22 (USDm)

Net cash from operations	118,548	87,582	89,035
--------------------------	---------	--------	--------



Source documents



Write & refine

Question

Example: MSFT FY2026 Q4

Forecast revenue + a central 80% interval.
USD millions
3 segments sum to total.
State drivers.

Cutoff: 29 Apr 2026, 16:00 ET

Rubric

- ✓ Forecast accuracy
- ✓ Segments = total
- ✓ Interval quality
- ✓ Reported base
- ✓ Recompute within 2%

Selected MSFT criteria

Benchmark task

Expert-authored task

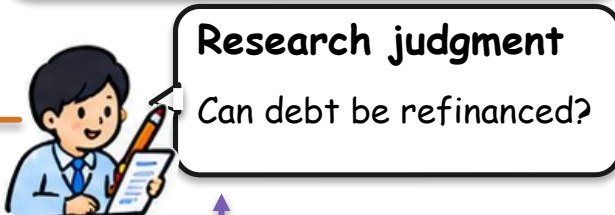
Public question

Can the company refinance its debt?

Private evidence & rubric



Expert-led refinement



SOTA models

Search & challenge

- Find missing evidence
- Challenge assumptions
- Test counterexamples

Expand expert insight

Review

Unused credit ≠ drawable credit



Reassess & refine

- Broaden the evidence
- Revise question, graph and rubric