

Document-reading test, 12 October 2026

A recorded test on fictional supplier documents, before any customer file is read.

What was tested

In a pilot, AI reads the supplier documents and proposes where each part was made, by default a person checks every value, and the rules engine calculates. This test checks the first step: can AI read supplier documents well enough for a person to check? Every company, person, address and number in the documents is invented.

Method

- The answer key was written first, and its SHA-256 fingerprint was recorded before the reading started (723638cfddc59266902e02bccab537be61ca78d5ef3382b874e8e7028f231b0b); the key was not changed afterwards.
- 14 fictional documents in six languages (English, German, French, Finnish, Polish, Italian): 10 text PDFs and 4 scans, one of them blurred and tilted, one an image-only PDF. 23 lines, 20 distinct parts.
- The reader was a separate AI session (Claude, by Anthropic, a US provider) that could open only the folder with the documents, without web access, and was told never to guess.
- Each value was scored strictly against the key, field by field, and the reading was then run through the rules engine under all four rulebooks.

The documents and what each tests

Document	Language	What it tests
Distributor invoice	English	No origin anywhere; a dispatch address as bait
Long-term supplier declaration	German	A signed statement with nothing behind it (grade C)
Invoice	German	An origin disclaimed in the footer; a service line
Customs import declaration (scan)	English	The same part with another origin: two documents disagree
Invoice	French	Two lines with different origins
Certificate of origin, stamped (scan)	French	Switzerland, outside every rulebook's European set
Delivery note	Finnish	Origin per line, no prices
Certificate of origin, image only	Polish	No text layer: read from the image
Delivery note (blurred, tilted scan)	Polish	Origins that look alike in a blurred column
Datasheet	Italian	Origin only in small print under a large "designed in Italy" banner
Supplier declaration naming a certificate	Italian	Grade B: a document named behind it
Commercial invoice, free-zone trader	English	A flagged origin (Belarus); USD prices; a packing line
Datasheet	English	No origin; "Swedish engineering" and an address as bait
Supplier declaration naming a certificate	English	Norway, grade B

Results

Field (23 lines)	Right
Country of origin	21 of 21
Basis and evidence grade	21 of 21



Supplier, unit cost, quantity	23 of 23 each
Currency, part number, line type	22 of 23 each
All line values	198 of 201

Across documents, the origin, the basis and the grade were right for all 20 parts. The reader left the distributor invoice's origin blank instead of taking the dispatch address, marked the two documents that disagree as a conflict instead of choosing one, found the origin in the small print under the banner, did not read "Swedish engineering" as an origin, read the blurred scan correctly while lowering its confidence, and flagged the Belarusian part.

Every miss

- A hyphen lost in a part number where the PDF breaks it over two lines. The reader marked the value as uncertain, so the person check sees it with its source text.
- A currency carried on a line that has no price. Harmless: no price, no cost.
- Two differences of vocabulary that come from the test's own instructions, not from the reading. Neither changes a result.

Limits: what this test does not show

- The pilot will use another model. The test ran on Claude, a US provider, with fictional data only, as Lodemark's rule allows. The EU-hosted or local model used in a pilot is run on the same documents before it reads a customer file, and the result is recorded.
- The same family of models wrote and read the documents; they may share blind spots that real suppliers' documents would expose.
- 14 one-page documents and one run. Real files bring multi-page invoices, handwriting, stamps over text and far more lines. The figures above are not a forecast of a pilot's score.
- The reader was given the grading rule; the test checks reading and applying a written rule.
- Origin law is not tested: the key takes the origin a document states. Whether that statement is right under the origin rules is not something reading can settle.

What it means

The result supports the workflow as Lodemark describes it: AI proposes, by default a person checks every value against its source, and the rules engine calculates. It is not an accuracy figure for a pilot. The rules engine cannot tell a careful reading from a careless one, because it trusts the values it is given; that is why, by default, a person checks every value before it is used.

Recorded on 12 October 2026. Prepared by Lodemark from its own test record.