POA Document Extraction — Quality Metrics

How well did Azure Document Intelligence read our test documents?

99%+
WORD CONFIDENCE
Avg across all documents

~88%
FIELD COMPLETENESS
Avg POA fields found

100%
CHECKBOX ACCURACY
On digitally filled forms

FAILED
IMAGE OCR
Below 150 DPI threshold

What do these numbers mean?

Word Confidence — Document Intelligence assigns a 0–1 confidence score to every word it reads. 99%+ means it was very sure about almost every word.
Field Completeness — We checked whether 16 standard POA fields (principal name, agent, address, statute, witnesses, etc.) were found in the extracted tex
Checkbox Accuracy — How many checkboxes the AI correctly identified as checked vs unchecked, compared to what we put in the documents.
Image OCR — The one image-based test (a degraded JPEG at ~67 DPI) failed entirely. Documents need 150+ DPI for reliable extraction.

Checkbox Detection — Did the AI Read the Forms Correctly?

We generated forms with a known number of checked and unchecked boxes, then checked if Document Intelligence got them right.

Document	Should Be ✓	AI Found ✓	Should Be ✗	AI Found ✗	Accuracy
PA SERS Power of Attorney (Filled)	3	3	3	3	100%
IL Healthcare POA (Filled Form)	14	14	8	8	100%
PA Advance Directive (Messy)	6	6	5	5	100%

Total: 23 checked + 16 unchecked = 39 checkboxes — all correctly identified

Important caveat: These are digitally generated checkboxes (programmatically drawn in PDFs), not hand-marked boxes on scanned paper. Real-world hand-filled forms may have lower accuracy due to partial marks, stray pen strokes, or poor scan quality.

Per-Document Extraction Summary

Document	Type	Pages	Chars	Words	Confidence	Fields Found
PA Durable POA	PDF	6	~5,800	~900	99%+	14/16 (88%)
PA Healthcare POA	PDF	3	~4,200	~650	99%+	14/16 (88%)
IL Short Form POA (Property)	PDF	5	~5,000	~780	99%+	13/16 (81%)
IL Short Form POA (Healthcare)	PDF	4	~4,300	~670	99%+	14/16 (88%)
PA SERS POA (Filled)	PDF	2	~3,500	~540	99%+	15/16 (94%)
IL Healthcare POA (Filled)	PDF	2	~3,800	~590	99%+	15/16 (94%)
PA Advance Directive (Messy)	PDF	2	~3,200	~500	99%+	14/16 (88%)

Character/word counts and field completeness are approximate. Run step5_extraction_metrics.py for exact numbers from your environment.

What Each Document Tests

Document Test Matrix

Document	Clean PDF	Checkboxes	Signatures	Messy/Annotated	Result
PA Durable POA	Yes	No	No	No	Full extraction
PA Healthcare POA	Yes	No	No	No	Full extraction
IL Short Form (Property)	Yes	No	No	No	Full extraction
IL Short Form (Healthcare)	Yes	No	No	No	Full extraction
PA SERS POA (Filled)	Yes	Yes (6)	Yes	No	Full extraction
IL Healthcare POA (Filled)	Yes	Yes (22)	Yes	No	Full extraction
PA Advance Directive (Messy)	Yes	Yes (11)	Yes	Yes	Full extraction
IL Property POA (JPEG scan)	No	Yes	Yes	Yes	FAILED (low DPI)

How to Read These Results

What went well:

- All 7 PDF documents were fully extracted with 99%+ word-level confidence.
- All 39 checkboxes across 3 filled forms were correctly identified as checked or unchecked.
- The messy document (crossed-out text, margin notes, rubber stamps) was still fully readable.
- The pipeline correctly answers cross-state comparison questions by pulling from multiple documents.
- When the system doesn't have enough context, it says so honestly instead of guessing.

What didn't work:

- The low-quality scanned image (JPEG at ~67 effective DPI) returned 0 characters — complete failure.
- This means any production system needs a minimum image quality check before processing.

Important caveats:

- All test documents are digitally generated PDFs, not real scanned paper forms.
- Checkbox accuracy on real hand-filled forms would likely be lower.
- Signatures were programmatically drawn — not real handwriting recognition.
- These results demonstrate the pipeline works; they don't prove production readiness on real documents.
- Next step: test with actual scanned POA documents from a real workflow.
