

NEWMIND RESEARCH · BENCHMARK REPORT 04

DOCUMENT PROCESSING & STRAIGHT-THROUGH BENCHMARK 2026

What a document really costs to handle, how much could pass without a person, and the playbooks and tools to get there.

Edition 1

Desk-research edition

October 2026

Published

Licence

Single company use

CONTENTS

What's inside

PART A • THE BENCHMARK

01	About this edition	4
02	Executive summary	6
03	Market context: intelligent document processing	8
04	Adoption: how many firms use AI on documents	11
05	E-invoicing: how much arrives machine-readable	13
06	Definitions and method	15
07	Benchmark: cost per document, manual and automated	17
08	Benchmark: straight-through rates by document type	20
09	AI and LLMs in document processing	22
10	The cost-of-gap model	24

PART B • SECTOR DEEP DIVES

11	Sector deep dives: overview	26
12	Financial services and insurance	28
13	Logistics, freight and trade	30
14	Construction and real estate	32
15	Healthcare	34
16	Professional services	37

PART C • PLAYBOOKS

17	Playbooks: overview	39
18	Playbook: supplier invoices	40
19	Playbook: insurance claims	42
20	Playbook: KYC and onboarding documents	44
21	Playbook: bills of lading and customs documents	46
22	Playbook: medical claims and prior authorization	48
23	Playbook: contracts	50
24	Designing the field schema	52
25	Human in the loop	54
26	Build or buy	55
27	Governance and risk	57

PART D • VENDORS, REGULATION AND ROADMAP

28	Vendor landscape	58
29	Regulation: EU and UK	60
30	Regulation: United States	62

31	The first 90 days	64
32	The 12-month roadmap	65
PART E • SELF-ASSESSMENT TOOLKIT		
33	Self-assessment: KPI sheet	66
34	Self-assessment: scoring worksheet	67
35	Scenario tables by role	68
36	Scenario tables: the first step	72
37	Reference grid: yearly handling cost by role	76
38	Reference grid: automated cost per document	79
39	Worksheet: baseline time study	81
40	Worksheet: field schema	84
41	Worksheet: exception and missing-information log	86
42	Worksheet: tool bake-off scorecard	88
43	Worksheet: questions for vendors	89
44	Worksheet: business case	91
45	Worksheet: regulatory checklist	92
46	Worksheet: roles and ownership	93
47	Worksheet: weekly accuracy sample	94
48	Worksheet: twelve-month KPI tracker	95
49	Worksheet: action plan	96
50	Method and limits	97
REFERENCE		
51	Glossary	98
52	Sources	99

Figures from third-party research carry a numbered source (full list in the Sources chapter). Everything else is labelled as NEWMIND analysis or model.

About this edition

This is Edition 1 of the NEWMIND Document Processing & Straight-Through Benchmark. It is a desk-research edition: it contains no survey data. It is built on official labour and price data, public filings, official legal texts, academic benchmarks and vendors' own published claims, each fetched, saved and checked for this edition.

What this edition contains

- **Cost per document:** the labour cost of handling a document for 12 document-heavy US occupations, from May 2025 wage data¹, against public list prices for machine reading^{2,3}.
- **Straight-through processing:** definitions, what is and is not published by document type, and how to measure your own rate.
- **AI and language models:** what independent benchmarks show, next to what vendors claim (tagged 'vendor claim').
- **A cost-of-gap model:** the yearly value and full-time equivalents of a better straight-through rate, labelled 'Market research model · based on published benchmarks'.
- **5 sector deep dives:** Financial services and insurance, Logistics, freight and trade, Construction and real estate, Healthcare, Professional services.
- **Playbooks:** supplier invoices, insurance claims, KYC and onboarding documents, bills of lading and customs documents, medical claims and prior authorization, contracts; plus field schema design, human in the loop, build or buy, and governance.
- **Regulation, vendors and roadmap:** the EU AI Act, eIDAS, VAT in the Digital Age, the UK Electronic Trade Documents Act 2023 and the US prior-authorization rule, a vendor landscape, and a 90-day and 12-month plan.
- **Self-assessment toolkit:** a KPI sheet, a scoring worksheet, scenario tables by role, reference cost grids and blank worksheets for a time study, field schema, exception log, tool bake-off, vendor questions, business case, regulatory checklist, ownership, weekly accuracy sampling, a twelve-month KPI tracker and an action plan.

What it does not contain yet

No original survey data. Edition 1 is desk research. Edition 2 is planned to add a survey of document teams, with straight-through, exception and rework rates by document type that companies do not publish.

Who it is for

COO and operations leaders; CIO, CTO and IT leaders. Document-heavy teams of 3+ people.

You will recognise these signals

- Staff read PDFs and re-key fields into a core system
- Missing information causes back-and-forth with senders
- Intake headcount grows with volume

What to measure

- Documents per month by type and channel
- Cost per document and minutes of handling
- Straight-through rate and field-level accuracy
- Turnaround time and rework rate
- Missing-information requests and their delay

How to read the numbers

- **Market figures** carry a numbered source. Where sources disagree, each is shown and named, never averaged.
- **Vendor figures** are tagged 'vendor claim'. They are not independently verified and are never used as model inputs.
- **Models** are transparent: every formula and input is printed so you can replace it with your own data.
- **Currency:** US dollars throughout. Figures published in pounds sterling are labelled GBP.

From benchmark to your numbers

A NEWMIND feasibility audit measures your own documents in 2–4 weeks and delivers:

- Your intake cost vs the reference ranges
- Straight-through potential per document type
- Field schema and validation rules needed
- Build-or-buy comparison of IDP options with costs

Licence: single company use. Please cite as: NEWMIND Research, Document Processing & Straight-Through Benchmark 2026, Edition 1. NEWMIND PARTNERS has no affiliation with any vendor or firm named in this report.

Executive summary

Most business documents are still read by a person and re-keyed into a system. Machine reading now costs a fraction of a cent per page, so the cost of a document is the person, and the lever is the share of documents that pass straight through without one.

\$4.96

Labour cost of 10 minutes of keying, US data entry keyer (NEWMIND calculation)

\$0.00150

List price to read one page of text by machine²

6.9%

EU enterprises using text mining, 2024⁴

Ten findings

01 A document costs labour, not reading

Ten minutes of keying by a US data entry keyer costs about \$4.96 in loaded labour (NEWMIND calculation from May 2025 wages^{1,5}). Reading a 3-page document by machine costs \$0.00450 at list price².

02 The market is sized two ways; neither is your business case

Precedence Research puts intelligent document processing at \$3.22 billion in 2025⁶; MarketsandMarkets puts a wider Document AI market at \$17.5 billion in 2026⁷. The definitions differ, so the figures are shown side by side and never averaged.

03 The straight-through rate, not the price per page, sets the cost

At a 30% straight-through rate, human review is 92% of the automated cost per document (NEWMIND model). Every ten points of straight-through rate is worth more than any difference in list price.

04 Closing a 40-point gap frees whole teams

A claims team handling 10,000 documents a month at 8 minutes each that lifts its straight-through rate from 30% to 70% frees about 3.1 FTE, worth \$233,000 a year at \$36.3 an hour (NEWMIND model).

05 AI use is spreading; on documents it is still a minority

Use of any AI technology by EU enterprises rose from 13.5% in 2024 to 20.0% in 2025; text mining reached 11.8%⁸.

06 More firms still send paper invoices than structured ones

In 2023, 70.2% of EU enterprises sent paper invoices and 38.7% sent e-invoices suitable for automated processing⁹. Capture stays part of accounts payable until mandates arrive.

07 Vendor accuracy claims are not benchmark results

Most large multimodal models score below 50 of 100 on OCRBench v2 ¹⁰, while one vendor says customers regularly achieve 99.5% accuracy **VENDOR CLAIM** ¹¹. Test any tool on your own documents before you sign.

08 Healthcare runs on document exchanges

Medicare Advantage insurers received nearly 53 million prior authorization requests ¹², and 80.7% of appeals were partly or fully overturned ¹². CMS API requirements generally apply from 1 January 2027 ¹³.

09 Regulation is moving documents to data

The EU AI Act has applied since 2 August 2026 ¹⁴. VAT in the Digital Age takes effect in stages from 1 July 2028 ¹⁵. In the UK, the Electronic Trade Documents Act 2023 has been in force since 20 September 2023 ¹⁶.

10 No one publishes straight-through rates by document type

We found no independent, verifiable straight-through benchmark by document type. This report gives definitions and a measurement method instead, and leaves those cells empty rather than filling them with vendor figures.

What to do on Monday

1. Count last month's documents by type and channel (chapter 33).
2. Price your cost per document with the reference grid (chapter 7).
3. Measure your straight-through rate to the definition in this report (chapter 6).
4. Price the gap to your target rate (chapter 10).
5. Write the field schema for your largest document type before you test any tool (chapter 24).

Market context: intelligent document processing

Most business documents still arrive as PDFs, scans, photos and email attachments. Someone reads them and re-keys the fields into a core system. Intelligent document processing (IDP) is the software category that tries to remove that step. This chapter sizes the market from published sources and sets out who is selling document work today.

\$3.22B

IDP market, 2025, Precedence Research ⁶

\$17.5B

Document AI market, 2026, MarketsandMarkets ⁷

45%

North America share of the IDP market, 2024 ⁶

Two market sizes, two definitions

Market research firms define this market differently, so their numbers cannot be compared or averaged. One measures IDP software only. The other measures a wider Document AI market that includes models and services built on them. Both are shown as published.

PUBLISHER	MARKET DEFINED AS	BASE VALUE	FORECAST	GROWTH A YEAR
Precedence Research	Intelligent document processing (IDP)	\$3.22 billion (2025)	\$43.9 billion (2034)	33.68% ⁶
MarketsandMarkets	Document AI	\$17.5 billion (2026)	\$35.3 billion (2031)	15.1% ⁷

The Precedence Research page also prints \$2.41 billion as 2025 revenue in its own summary table ⁶, next to the \$3.22 billion headline. We show both because the publisher does. Growth rates are the publishers' compound annual growth rates; we checked that each matches its own start and end values.

Why the range is so wide

The gap between the two sizes is not an error. It comes from scope. A narrow definition counts licence and subscription revenue for capture and extraction software. A wide definition adds document-focused AI models, platforms and the services sold around them. Forecast growth rates differ for the same reason: a small, narrowly defined market can grow faster in percentage terms than a broad one.

For a buyer, the market size matters less than two facts behind it. First, the published per-page prices of the large cloud providers are low and public, so basic extraction can be priced before any sales call. Second, the cost of the underlying hardware keeps falling. Stanford HAI reports that AI hardware costs have declined by 30% a year ¹⁷. The cost of reading a page by machine is now measured in fractions of a cent (chapter on cost per document).

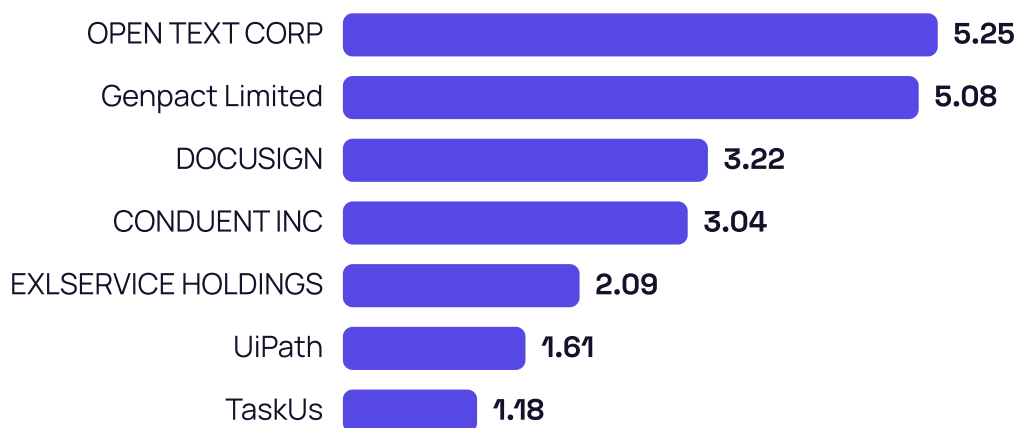
Who sells document work today

No filing used here reports intelligent document processing as a separate revenue line. The firms below sell document work as part of wider software or services businesses. Their total revenue, taken from annual filings

with the US Securities and Exchange Commission, shows the scale of the companies a buyer will meet. It is not a measure of the IDP market.

COMPANY	WHAT IT SELLS	FISCAL YEAR ENDED	REVENUE	PRIOR YEAR	CHANGE
OPEN TEXT CORP	Information management software	June 2026	\$5.25 billion	\$5.17 billion	rose 1.5% ¹⁸
Genpact Limited	Business process services	December 2025	\$5.08 billion	\$4.77 billion	rose 6.6% ¹⁹
DOCUSIGN, INC.	Agreement and e-signature software	January 2026	\$3.22 billion	\$2.98 billion	rose 8.2% ²⁰
CONDUENT INC	Business process services, incl. transaction processing	December 2025	\$3.04 billion	\$3.36 billion	fell 9.4% ²¹
EXLSERVICE HOLDINGS, INC.	Data, analytics and operations services	December 2025	\$2.09 billion	\$1.84 billion	rose 13.6% ²²
UiPath, Inc.	Automation and document understanding software	January 2026	\$1.61 billion	\$1.43 billion	rose 12.7% ²³
TaskUs, Inc.	Outsourced operations services	December 2025	\$1.18 billion	\$995 million	rose 19.0% ²⁴

Latest fiscal-year revenue, USD billion (company totals, not document revenue)



Of the 7 companies, 6 grew revenue in their latest fiscal year and 1 did not. Software firms selling automation, information management and agreements sit next to large services firms that run document operations for clients with people and tools. Buyers can choose between buying software, buying an outcome from a services firm, or building on cloud APIs. The build-or-buy playbook compares these options.

Where the document volume comes from

- **Healthcare:** US national health spending reached \$5.30 trillion in 2024 ²⁵, and Medicare Advantage insurers received nearly 53 million prior authorization requests ¹². Each one is a document exchange.
- **Trade and logistics:** the Law Commission estimates that global container shipping generates billions of paper documents a year ²⁶.

- **Financial services:** onboarding, lending and claims depend on identity documents, statements and forms, under customer due diligence rules (chapter on US regulation).
- **Construction and real estate:** US construction spending ran at a seasonally adjusted annual rate of \$2.20 trillion in the latest Census Bureau estimate ²⁷, and every dollar moves through invoices, lien waivers, change orders and contracts.

What this means for you

Ignore the headline market size when you build a business case. What matters is your own volume, your cost per document and the share that could pass straight through. The rest of this report gives you reference ranges for each and a model to price the gap.

Definitions and method

Document processing numbers are hard to compare because vendors, analysts and buyers use the same words for different things. This chapter fixes the definitions used in the rest of the report and explains how every figure was produced.

The metrics

METRIC	MEANING	HOW TO COMPUTE IT
Straight-through processing (STP) rate	Share of documents that go from receipt to posting in the system of record with no human touch.	Documents with zero touches ÷ all documents received
Touchless rate	Same idea measured per field or per step; some vendors use it for documents with no manual correction.	Fields (or steps) with no manual edit ÷ all fields (or steps)
Field-level extraction accuracy	Share of extracted fields whose value matches the ground truth exactly after normalisation.	Correct fields ÷ fields checked, on a labelled sample
Document-level accuracy	Share of documents where every required field is correct.	Documents with all required fields correct ÷ documents checked
Exception rate	Share of documents routed to a person because a rule, a confidence threshold or a match failed.	Documents sent to review ÷ all documents
Missing-information rate	Share of documents that cannot be completed without asking the sender for more.	Documents with an outbound request ÷ all documents
Cycle time	Elapsed time from receipt to posting, including waiting time.	Posting timestamp minus receipt timestamp, median and 90th percentile
Handling time	Minutes of human work spent on one document, excluding waiting.	Logged or sampled minutes per document
Cost per document	Fully loaded labour cost of handling plus software and processing cost.	Loaded hourly cost × handling minutes ÷ 60 + software cost per document
Rework rate	Share of posted documents later corrected or reversed.	Documents corrected after posting ÷ documents posted
Confidence threshold	Score below which the system sends a field or document to review.	Set per field from a labelled sample

Why STP and accuracy are not the same number

A system can read 98% of fields correctly and still pass far fewer documents straight through. If an invoice has 20 required fields and each is right 98% of the time, independently, the chance that all 20 are right is 67%. Every document with one wrong or low-confidence field goes to a person. That is why a vendor's field accuracy and a buyer's STP rate are different numbers, and why this report never converts one into the other.

STP also depends on things outside extraction: whether the purchase order exists, whether the supplier is in the master file, whether the policy is active. A perfect reader cannot pass a document whose data does not match the system of record.

How this report labels numbers

LABEL	WHAT IT MEANS
A number with a superscript	A published figure. The superscript points to the numbered source list.
VENDOR CLAIM	An outcome a vendor publishes about its own product or a named customer. Not independently verified.
NEWMIND model	A calculation by NEWMIND from published inputs and stated assumptions, labelled 'Market research model - based on published benchmarks'.
Not published in a verified source	We looked and found no figure we could fetch, read and verify this edition. We leave the cell empty rather than fill it.

Method

- **Desk research only.** This edition contains no survey data. Every figure was fetched from its source for this edition, saved, and checked against the saved copy by an automated script before the report was built.
- **Labour cost.** US occupational wages come from the Bureau of Labor Statistics Occupational Employment and Wage Statistics, May 2025 ¹. We load them with the employer cost of benefits from the Employer Costs for Employee Compensation release, 2026 2nd Quarter ⁵: for private industry workers, wages and salaries averaged \$32.8 of \$46.9 total compensation per hour worked, 70.0%. The model divides each OEWS mean hourly wage by 0.70, which is the same as multiplying it by 1.43 ($1 \div 0.70$), to add benefits and other employer costs. The same multiplier applies to every role.
- **Processing cost.** Machine reading prices come from the public price pages of two cloud providers ^{2,3}. List prices only; negotiated prices may differ.
- **Vendors.** Outcome figures come from vendors' own websites and are tagged as vendor claims. We claim no affiliation with any vendor named.
- **Regulation.** Dates come only from official legal sources: EUR-Lex, legislation.gov.uk, the Federal Register and CMS. The tense of each date is computed against the edition date.
- **Currency.** US dollars throughout. Figures published in pounds sterling are labelled GBP and never converted.
- **Rounding.** Money is shown to three significant figures. Percentages are shown as published or to one decimal place.

What this edition does not do

It does not publish an independent STP benchmark by document type where none could be verified. It does not convert vendor accuracy claims into expected results. It does not rank vendors.

Benchmark: cost per document, manual and automated

The cost of a document is mostly the cost of the person who reads it. This chapter prices that person's time from official US wage data, prices machine reading from public cloud price lists, and shows where the two meet.

\$29.7–\$47.9

Loaded hourly cost across 12 document-heavy US occupations, May 2025 wages (NEWMIND calculation)

\$4.96

Cost of 10 minutes of keying by a data entry keyer (NEWMIND calculation)

\$0.00150

List price to read one page of text by machine ²

Step 1: what an hour of document work costs

We use the mean hourly wage for each occupation from the May 2025 Occupational Employment and Wage Statistics ¹, then divide by the wage share of total compensation, 70.0% for private industry in the 2026 2nd Quarter release ⁵. That adds benefits and employer costs. It does not add office space, management or IT, so it is a floor, not a full cost.

OCCUPATION	US EMPLOYMENT	MIDDLE HALF OF WAGES	MEAN WAGE	LOADED HOURLY COST
Data Entry Keyers	127,080	\$17.2–\$23.3	\$20.8	\$29.7
Office Clerks, General	2,464,940	\$17.5–\$26.1	\$22.3	\$31.9
Medical Secretaries and Administrative Assistants	961,610	\$18.5–\$24.3	\$22.5	\$32.1
Order Clerks	75,200	\$19.1–\$26.1	\$23.2	\$33.1
New Accounts Clerks	36,860	\$21.1–\$26.5	\$23.7	\$33.9
Billing and Posting Clerks	404,060	\$20.6–\$28.0	\$24.6	\$35.1
Loan Interviewers and Clerks	164,790	\$21.7–\$29.1	\$25.3	\$36.1
Insurance Claims and Policy Processing Clerks	214,260	\$21.3–\$29.1	\$25.4	\$36.3
Bookkeeping, Accounting, and Auditing Clerks	1,373,680	\$20.9–\$29.6	\$25.8	\$36.8
Medical Records Specialists	194,720	\$20.9–\$31.2	\$27.3	\$39.0
Cargo and Freight Agents	97,670	\$22.0–\$30.5	\$27.5	\$39.3
Paralegals and Legal Assistants	392,880	\$24.2–\$38.5	\$33.5	\$47.9

Middle half of wages = 25th to 75th percentile hourly wage. Loaded hourly cost is a NEWMIND calculation: mean wage + wage share of compensation.

Step 2: cost per document by handling time

Handling time is the number you know best and the one we cannot publish for you: it depends on the document, the system and the team. The grid shows the labour cost per document for typical handling times. Find your role and your minutes.

OCCUPATION	2 MIN	5 MIN	10 MIN	20 MIN	30 MIN
Data Entry Keyers	\$0.991	\$2.48	\$4.96	\$9.91	\$14.9
Office Clerks, General	\$1.06	\$2.66	\$5.31	\$10.6	\$15.9
Medical Secretaries and Administrative Assistants	\$1.07	\$2.68	\$5.36	\$10.7	\$16.1
Order Clerks	\$1.10	\$2.76	\$5.52	\$11.0	\$16.6
New Accounts Clerks	\$1.13	\$2.83	\$5.65	\$11.3	\$17.0
Billing and Posting Clerks	\$1.17	\$2.92	\$5.85	\$11.7	\$17.5
Loan Interviewers and Clerks	\$1.20	\$3.01	\$6.01	\$12.0	\$18.0
Insurance Claims and Policy Processing Clerks	\$1.21	\$3.03	\$6.06	\$12.1	\$18.2
Bookkeeping, Accounting, and Auditing Clerks	\$1.23	\$3.07	\$6.13	\$12.3	\$18.4
Medical Records Specialists	\$1.30	\$3.25	\$6.50	\$13.0	\$19.5
Cargo and Freight Agents	\$1.31	\$3.28	\$6.55	\$13.1	\$19.7
Paralegals and Legal Assistants	\$1.60	\$3.99	\$7.98	\$16.0	\$23.9

MARKET RESEARCH MODEL · BASED ON PUBLISHED BENCHMARKS

Step 3: what machine reading costs

Cloud providers publish per-page list prices for reading documents. Basic text detection costs a fraction of a cent per page. Extracting tables, form fields and answers to specific questions costs more, but still cents per page.

SERVICE (LIST PRICE)	PER PAGE	PER 1,000 PAGES	SOURCE
Amazon Textract, text detection (first 1M pages)	\$0.00150	\$1.50	2
Amazon Textract, text detection (after 1M pages)	\$0.000600	\$0.600	2
Amazon Textract, tables	\$0.0150	\$15.0	2
Amazon Textract, forms	\$0.0500	\$50.0	2
Amazon Textract, tables + forms + queries	\$0.0700	\$70.0	2
Google Document AI, Form Parser (first tier)	\$0.0300	\$30.0	3

SERVICE (LIST PRICE)	PER PAGE	PER 1,000 PAGES	SOURCE
Google Document AI, Form Parser (above first tier)	\$0.0200	\$20.0	3
Google Document AI, custom extractor (to 1M pages)	\$0.0300	\$30.0	3

Prices are public list prices captured this edition; volume tiers and negotiated discounts apply. Platform and integration costs are not included.

Step 4: automated cost per document

Automated cost = pages × price per page + share of documents sent to review × review minutes × loaded hourly cost. Model inputs: 3 pages per document, 5 minutes per reviewed document, reviewed by data entry keyers at \$29.7 an hour.

STRAIGHT-THROUGH RATE	CUSTOM EXTRACTOR (GOOGLE LIST PRICE)	FORMS ANALYSIS (AMAZON LIST PRICE)
30%	\$1.83	\$1.89
50%	\$1.33	\$1.39
70%	\$0.834	\$0.894
90%	\$0.338	\$0.398

MARKET RESEARCH MODEL · BASED ON PUBLISHED BENCHMARKS

At a 30% straight-through rate, human review is 92% of the automated cost per document. At 90%, the automated cost falls to \$0.338, against \$4.96 for 10 minutes of manual keying. The machine reading price barely moves the result. The straight-through rate does.

The lesson for your business case

Do not choose a tool on its price per page. Choose it on the share of your documents it can pass without a person, measured on your own documents. Every ten points of straight-through rate is worth more than any difference in list price.

Sources

Each entry gives the date that applies to it: when a law was adopted or came into force, when a document was published, or, for live web pages and datasets, when it was accessed (6 October 2026). Figures quoted are short, attributed and linked; no third-party tables or charts are reproduced.

1. U.S. Bureau of Labor Statistics, Occupational Employment and Wage Statistics, national estimates (May 2025), accessed 6 October 2026 <https://www.bls.gov/oes/>
2. Amazon Web Services, Amazon Textract pricing, accessed 6 October 2026 <https://aws.amazon.com/textract/pricing/>
3. Google Cloud, Document AI pricing, accessed 6 October 2026 <https://cloud.google.com/products/document-ai/pricing>
4. Eurostat, Usage of AI technologies increasing in EU enterprises, published 23 January 2025 <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20250123-3>
5. U.S. Bureau of Labor Statistics, Employer Costs for Employee Compensation, private industry, accessed 6 October 2026 <https://www.bls.gov/news.release/ecec.nr0.htm>
6. Precedence Research, Intelligent Document Processing Market Size to Hit USD 43.92 Billion by 2034, published 7 July 2026 <https://www.precedenceresearch.com/intelligent-document-processing-market>
7. MarketsandMarkets, Document AI Market worth \$35.34 billion by 2031, press release, published 18 September 2026 <https://www.marketsandmarkets.com/PressReleases/document-ai.asp>
8. Eurostat, Artificial intelligence by size class of enterprise (dataset isoc_eb_ai), published 15 June 2026 https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/isoc_eb_ai?geo=EU27_2020&size_emp=GE10&unit=PC_ENT&nace_r2=C10-S951_X_K
9. Eurostat, Integration with customers/suppliers, supply chain management by size class of enterprise (dataset isoc_eb_ics), published 17 July 2025 https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/isoc_eb_ics?geo=EU27_2020&size_emp=GE10&unit=PC_ENT&nace_r2=C10-S951_X_K
10. Fu et al., OCRBench v2: An Improved Benchmark for Evaluating Large Multimodal Models on Visual Text Localization and Reasoning, arXiv:2501.00321, published 31 December 2024 <https://arxiv.org/abs/2501.00321>
11. Hyperscience, Hypercell: Intelligent Document Processing Platform, accessed 6 October 2026 <https://www.hyperscience.ai/platform/hypercell/>
12. KFF, Medicare Advantage Insurers Made Nearly 53 Million Prior Authorization Determinations in 2024, published 28 January 2026 <https://www.kff.org/medicare/medicare-advantage-insurers-made-nearly-53-million-prior-authorization-determinations-in-2024/>
13. CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F), fact sheet, published 17 January 2024 <https://www.cms.gov/newsroom/fact-sheets/cms-interoperability-and-prior-authorization-final-rule-cms-0057-f>
14. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal of the European Union, adopted 13 June 2024 https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689
15. Council Directive (EU) 2025/516 (VAT in the Digital Age), adopted 11 March 2025 https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202500516
16. Electronic Trade Documents Act 2023, section 8, in force 20 September 2023 <https://www.legislation.gov.uk/ukpga/2023/38/section/8>
17. Stanford HAI, The 2025 AI Index Report, accessed 6 October 2026 <https://hai.stanford.edu/ai-index/2025-ai-index-report>
18. SEC EDGAR, OPEN TEXT CORP annual report on Form 10-K, XBRL company facts, published 6 August 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001002638.json>
19. SEC EDGAR, Genpact Limited annual report on Form 10-K, XBRL company facts, published 26 February 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001398659.json>
20. SEC EDGAR, Docusign, Inc. annual report on Form 10-K, XBRL company facts, published 18 March 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001261333.json>
21. SEC EDGAR, Conduent Inc annual report on Form 10-K, XBRL company facts, published 19 February 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001677703.json>
22. SEC EDGAR, Exlservice Holdings, Inc. annual report on Form 10-K, XBRL company facts, published 24 February 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001297989.json>
23. SEC EDGAR, UiPath, Inc. annual report on Form 10-K, XBRL company facts, published 25 March 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001734722.json>
24. SEC EDGAR, TaskUs, Inc. annual report on Form 10-K, XBRL company facts, published 5 March 2026 <https://data.sec.gov/api/xbrl/companyfacts/CIK0001829864.json>

25. Centers for Medicare & Medicaid Services, NHE Fact Sheet, accessed 6 October 2026 <https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/nhe-fact-sheet>
26. Law Commission, Electronic trade documents, accessed 6 October 2026 <https://lawcom.gov.uk/project/electronic-trade-documents/>
27. U.S. Census Bureau, Monthly Construction Spending, August 2026, published 1 October 2026 <https://www.census.gov/construction/c30/current/index.html>

© 2026 NEWMIND PARTNERS. Licensed for use within one company. NEWMIND PARTNERS AG, Luzern, Switzerland · NEWMIND PARTNERS Inc., Sheridan, Wyoming, USA · hi@newmindpartners.com

Get the full report

This sample includes the cover, contents, about page, executive summary, the market context, definitions and method, the cost-per-document benchmark and the sources cited in them. The full report continues with adoption and e-invoicing, straight-through rates by document type, AI and language models, the cost-of-gap model, five sector deep dives, playbooks by document type, field schema, human in the loop, build or buy, governance, vendors, regulation, a 90-day and 12-month roadmap and the self-assessment toolkit.

Full report

newmindpartners.com/benchmark-document-processing.html#report

NEXT STEP

TURN THIS BENCHMARK INTO YOUR CUSTOM DOCUMENT WORKFLOW.

You now know what a document costs and what a better straight-through rate is worth. NEWMIND PARTNERS designs, builds and runs the workflow that closes your gap: intake, extraction, validation, exception queues and requests to senders, built around the systems you already use.

01

Feasibility audit

In 2–4 weeks we measure your document intake on your data (read-only): volume by type and channel, cost per document, straight-through potential and missing-information delays. Then we compare every option on the market, off-the-shelf or custom.

02

Design

We write the field schema and validation rules for each document type and design the workflow around your core systems, with a human checkpoint on every decision that matters.

03

Build and run

Our senior engineers build it, connect it to the systems you already use, and run it in production with an audit trail.

04

Prove it

Baseline on day one, then the change in straight-through rate, hours and cost per document, reported every month.

Start with a feasibility audit

Find out what your intake really costs, which option fits best, and the business case to back it. No commitment.

Start a feasibility audit →