



AZURE MARKETPLACE · TRANSACTABLE SAAS · METERED PER PAGE

Turn any document into a decision.

DocuFii is an API-first document-intelligence platform for lenders, fintechs and finance teams. Send a bank statement, payslip, ID or invoice and get back **structured, decision-ready data** — affordability, income, categorisation, fraud forensics and identity fields — in seconds. One API, one key, billed per page against your Azure commitment.

```
POST /api/ocr/v2/process/bankstatementandclassify
```

```
"income"      : { "net_monthly": 28 450.00, "payslip_match": 0.97 },
"affordability": { "disposable": 9 120.50, "dti": 0.31 },
"fraud"       : { "tamper": "none", "score": 0.04 },
"decision"    : "PASS"
```

28 endpoints

across two API families,
behind one subscription

Structured JSON

decision inputs, not raw
OCR text

Free monthly tier

integrate and evaluate in an
afternoon

Microsoft Azure

hosted, transactable, on
your Azure bill

One offer. Two ways to call it.

The Document Intelligence offer exposes two complementary API families behind a single subscription. Reach for **typed processors** when you know the document and want the deepest structured output; reach for **generic read & extract** for anything else, custom field schemas, and high-volume batches.

OCRV2 — TYPED PROCESSORS

Deep, decision-ready extraction

Purpose-built processors per document type — each returns a rich, validated schema plus derived signals, not just fields.

- > **Financial.** Bank-statement analysis & classification, payslip, lending analytics, affordability.
- > **Identity.** SA ID, smart-ID, passport, driver's licence and civil certificates.
- > **Commerce.** Invoice and receipt extraction.
- > **Matching.** ID-to-value, face/identity and address matching with confidence.
- > **Auto-route.** One endpoint that detects the type and dispatches.

OCRREAD — GENERIC READ & EXTRACT

Any document, any schema, at scale

Flexible OCR for everything the typed set doesn't cover — including fields shaped to a schema you supply, and bulk pipelines.

- > **Read.** Plain text plus optional layout from any single document.
- > **Extract.** OCR to a caller-supplied field schema across 22+ document families.
- > **Batch.** Up to 10 documents synchronously in one call.
- > **Async.** Submit up to 1,000 documents; poll a job for results.

28

endpoints across the two families

22+

schema-driven extract families

1,000

documents per asynchronous job

<10 min

to a first extraction after subscribing

What the platform reads — and what it decides.

Every capability returns structured data; the financial and matching capabilities add derived signals your systems can act on directly.

Financial document intelligence

01

Every transaction, balance and statement period extracted — even across page breaks — with **per-transaction categorisation**, income & affordability metrics, and an optional embedded PDF report. Payslips analysed and reconciled gross-to-net.

`/process/bankstatement``/bankstatementandclassify``/bankstatement/report``/process/payslip`

Lending analytics & affordability

02

Cross-validate a payslip against the bank statement: income-match confidence, disposable income, debt-service ratio and a **refer / decline signal** — the raw material of an automated credit decision.

`/lending/analytics`

Identity & civil documents

03

SA national ID, smart-ID card, passport and driver's licence, plus birth, death, marriage, asylum, voter and residence-permit certificates — each to a **typed schema**, with international format handling.

`/process/id``/id-smartid``/passport``/driverslicense``+7 certificates`

Invoices, receipts & finance docs

04

Line-item invoice and receipt extraction for AP automation — plus **22+ further families** (tax certificates, financial statements, proof of address, lease, insurance and more) via schema-driven extract.

`/process/invoice``/process/receipt``/ocr/extract`

Matching & verification

05

Extract an ID and compare it to supplied identity values; match a person's name and address against an address document — each with a **verdict and a confidence score** for straight-through verification.

`/process/id-match``/address-match``/address-name`

Generic OCR, custom schemas & bulk

06

Plain-text read with optional layout; field extraction shaped to **your own schema**; synchronous batches of 10 and asynchronous jobs of up to 1,000 documents for portfolio-scale processing.

`/ocr/read``/ocr/extract``/read/batch``/read/batch/async`

From subscription to decision, in the browser.

Every subscription includes the **no-code DocuFii portal** alongside the REST API — sign in with Microsoft and process documents without writing a line of code. Same subscription, same Azure bill.

DocuFii turns financial documents into decisions.

One API - and a no-code portal - that extracts, structures, and *interprets* financial documents - bank statements and payslips from virtually any bank, in any country. Same subscription, same Azure bill.

Off by default - we don't store your data

✓ Your subscription is active

Starter · 750 pages / month included

YOUR API SUBSCRIPTION KEY

b7f2 · 9c41 [Copy](#)

[Open the DocuFii portal →](#)

[View API reference](#)

Bank statement analysis

Every transaction, balance & period - even where tables span page breaks.

Affordability & fraud

Income, disposable & DTI, plus fraud signals - tamper forensics, UIF/PAYE reconciliation and cross-document name & income matching.

Identity & payslips

ID, passport, licence & payslip extraction - any country's formats - with declared-vs-actual checks.

API or no-code portal

Build against REST, or sign in with Microsoft and process in the browser.

Your subscription, live. Included monthly pages, your API subscription key, and one-click routes to the portal and the API reference — off by default, DocuFii doesn't store your data.

Identity & payslip extraction

ID documents, passports and driver's licences, plus payslip parsing across international formats - with declared-vs-actual income cross-checks and face-match against the ID Trust module. Two ways to consume: a REST API for developers, or the no-code **portal** - same subscription, same bill.

workspace / document-intelligence / bank-statement

Bank Statement Intelligence

Structured account data, a **Suspicious** tamper flag and confidence score, and an affordability **recommendation** - income, obligations and DTI in one response.

workspace / document-intelligence / lending-analytics

Lending Analytics

Cross-validates a payslip against the bank statement - income match confidence, disposable income, debt-service ratio and a **Refer / manual-review** decision.

Decisions on screen. Bank Statement Intelligence returns structured account data, a tamper flag with confidence score and an affordability recommendation; Lending Analytics cross-validates payslip against statement into a refer / manual-review decision.

Every endpoint, at a glance.

All operations accept a document (file or URL) and return structured JSON. Authenticate with your subscription key over HTTPS; usage is metered per page.

OCRV2 • FINANCIAL

POST /api/ocr/v2/process/bankstatement	Analyze a bank statement — transactions, balances, periods, totals.
POST /api/ocr/v2/process/bankstatementandclassify	Analyze and classify every transaction into spend categories.
POST /api/ocr/v2/process/bankstatement/report	Analyze, classify, and embed a formatted PDF report of the statement.
POST /api/ocr/v2/process/payslip	Analyze a payslip — earnings, deductions, net pay, gross-to-net checks.
POST /api/ocr/v2/lending/analytics	Payslip-vs-bank-statement matcher — income match, DTI, refer/decline signal.

OCRV2 • IDENTITY

POST /api/ocr/v2/process/id	Extract identity-document fields (auto identity type).
POST /api/ocr/v2/process/id-nationalid • id-smartid	Extract national-ID and smart-ID-card fields.
POST /api/ocr/v2/process/passport • driverslicense	Extract passport and driver's-licence fields, incl. MRZ where present.
POST /api/ocr/v2/process/id-{birth·death·marriage·...}certificate	Extract fields from the six civil-certificate types.

OCRV2 • COMMERCE, MATCHING & ROUTING

POST /api/ocr/v2/process/invoice • receipt	Extract invoice and receipt fields and line items.
POST /api/ocr/v2/process/qualification	Extract qualification-certificate fields.
POST /api/ocr/v2/process/id-match	Extract an ID and compare it to supplied identity values — verdict + score.
POST /api/ocr/v2/process/address-match • address-name	Extract an address document; match to a supplied name/address.
POST /api/ocr/v2/process/auto	Auto-detect the document type and route to the right processor.

OCRREAD • GENERIC & BULK

POST /api/ocr/read	Extract plain text (and optional layout) from a single document.
POST /api/ocr/extract	OCR a document and extract fields shaped to a caller-supplied schema.
POST /api/ocr/read/batch	Synchronously process up to 10 documents in a single call.
POST /api/ocr/read/batch/async	Submit up to 1,000 documents for asynchronous processing.
GET /api/ocr/read/jobs/{jobId}	Get the status and, when complete, the results of an async batch job.

22+ document families, one schema-driven endpoint.

Beyond the typed processors, `/api/ocr/extract` returns structured fields for a growing library of business and finance documents — you name the document family and the fields you want back.

bank-confirmation proof-of-address financial-statements individual-tax-return employer-tax-certificate
tax-compliance-status property-tax-clearance company-registration director-appointment employment-contract
employment-confirmation credit-agreement lease-agreement insurance-policy property-title trust-deed
regulatory-licence medical-certificate settlement-offer settlement-confirmation bill-of-lading
supplier-diversity

USE CASES

Where teams put it to work.

Customer & borrower onboarding

Capture an ID and proof of address, extract the fields, and match them to what the applicant told you — straight-through where it clears, flagged where it doesn't. No manual keying, no re-typing off a photo.

USES `/process/id` · `id-match` · `address-match` · `proof-of-address extract`

Lending & affordability decisioning

Turn a bank statement and payslip into an underwriting-ready view: categorised transactions, verified income, disposable income, debt-service ratio and a refer/decline signal.

USES `/bankstatementandclassify` · `/payslip` · `/lending/analytics`

Finance & AP workflow automation

Extract invoices and receipts to line-item JSON for accounts-payable, expense and reconciliation systems — batch a day's documents in one async job and pull the results when they're ready.

USES `/process/invoice` · `/process/receipt` · `/read/batch/async`

Fraud, forensics & reconciliation

Surface tamper signals and internal inconsistencies — running-balance math, gross-to-net payslip reconciliation and cross-document name/income matching — before a document reaches a human reviewer.

USES `/bankstatement/report` · `/payslip` · `/lending/analytics`

Qualification & credential capture

Read a qualification certificate to structured fields — issuer, qualification, dates, holder — as the data-capture step ahead of any formal verification, replacing manual transcription.

USES `/process/qualification` · `/ocr/extract`

Any document, your schema

Have a document type that isn't in the typed set? Post it to the generic extractor with the fields you want back, and get clean JSON — the same pipeline, shaped to your data model.

USES `/ocr/read` · `/ocr/extract (22+ families)`

Off by default — we don't store your data.

✓ Process-and-discard is the default posture

DocuFii processes each document and returns the result; it does **not** retain your files or extracted data unless you explicitly turn on history storage for your organisation.

You control storage & deletion

- History storage is an **opt-in** organisation setting (off by default).
- When on, stored documents auto-expire after a retention window (default **90 days**) and are hard-deleted.
- Delete any document on demand — individually or in bulk — a true hard delete, recorded in an audit trail.

Encrypted & isolated

- Processed in **Microsoft Azure**, encrypted in transit and at rest.
- Any stored content is encrypted at rest with a **per-organisation key** — one tenant is cryptographically isolated from another.
- Secrets held in Azure Key Vault via managed identity; API fronted by APIM with per-subscription keys & rate limits.

Compliance-aware

- Retention, deletion and consent behaviour are configurable to your regulatory posture (POPIA, GDPR, NCA).
- Storage and deletion events are captured in an audit trail.
- Deploy in your own tenant / self-hosted where required.

[roadmap](#)

WHY DOCUFII

Decisions, not OCR text.

Decision-ready outputs

Affordability, income, categorisation and fraud signals come out of the box — not a text dump you have to interpret yourself.

Two API families, one key

Typed processors for depth, generic read & extract for everything else — under a single subscription and meter.

Azure-native & transactable

Buy against your Microsoft Azure commitment, billed on your Azure invoice — no new vendor procurement cycle.

API or no-code portal

Every subscription includes the DocuFii web portal — sign in and process documents in the browser, no integration required.

Built for financial documents

Statements that span page breaks, payslips reconciled gross-to-net, cross-document matching — tuned for lending, not generic forms.

Predictable per-page pricing

Transparent metering with a free monthly tier — integrate and evaluate before you commit, then scale with usage.

Live in an afternoon.

01

Subscribe

Buy the **DocuFii Document Intelligence** plan on Azure Marketplace — billed on your Azure invoice, against your existing commitment.

02

Activate

Activate your tenant and get an API key — or sign in to the **no-code DocuFii portal** with your Microsoft account.

03

Call the API

POST a document over HTTPS with your key, or upload in the browser. **First extraction in under 10 minutes.**

04

Get structured data

Receive JSON, classifications, affordability metrics and fraud signals — **metered per page, free monthly tier.**

Find DocuFii on the Azure Marketplace

Search "DocuFii Document Intelligence" — or book a walkthrough at deistech.co.za

[Start with the free tier →](#)