



Document Entity Extraction

Extract Key Entities from Indian Language Documents – with Speed and Accuracy

JioCloud Document Entity Extraction automates data capture from unstructured regional-language documents — for faster workflows and compliance.

The Challenge

- Businesses and government teams regularly deal with forms, contracts, applications, and emails in Indian languages.
- Manually picking out key details like names, dates, or contact numbers is slow and often inconsistent.
- Most extraction tools aren't built for Indian scripts or the way people mix languages while writing.
- It's also hard to find tools that offer simple APIs (Application Programming Interface), quick testing, or clear pricing.
- Teams often struggle to pull useful, structured information from these documents at scale



The JioCloud Solution

JioCloud Document Entity Extraction is built specifically for Indian documents and languages. It captures key information like names, phone numbers, dates, and organisations from scanned, printed, or typed text. The service works across multiple Indian scripts, supports API-based integration, and includes a live testing playground. With consumption-based billing and a self-service dashboard, teams can automate document processing at scale — without depending on manual workflows or external support.

Key Features

- **Support for Indian languages**
Works with documents written in Hindi, Tamil, Bengali, Telugu, Marathi, Kannada, and other widely used Indian scripts.
- **Multiple entity types**
Detects and extracts names, organisations, phone numbers, email addresses, locations, dates, numbers, hashtags, and URLs.
- **Handles unstructured content**
Processes data from paragraphs, scanned forms, email text, and handwritten inputs.
- **Live testing via playground**
Upload samples and instantly check how entities are extracted before going live.
- **Easy API integration**
Designed to connect smoothly with CRMs, document workflows, or e-governance systems.
- **Self-service control panel**
Creates, monitors, and manages your service and credits independently.
- **Usage-based pricing**
Pay per character or document, with clear visibility into how much you consume.

What You Gain

- Speed up backend processes by automating document parsing
- Reduce errors and manual effort in form-based workflows
- Capture data from Indian scripts and mixed-language text
- Maintain consistency for audits, KYC, and compliance
- Feed structured data into analytics or downstream systems
- Scale entity extraction without infrastructure costs



Use Cases in Action

Resume screening

Extract candidate details from regional-language resumes and auto-fill recruitment systems.

KYC form processing

Capture names, addresses, and contact info from scanned forms in Indian scripts.

Government scheme applications

Identify Aadhaar, income, and location data from regional application forms.

Legal document analysis

Extract parties, dates, and addresses from contracts and petitions across languages.

Hospital intake forms

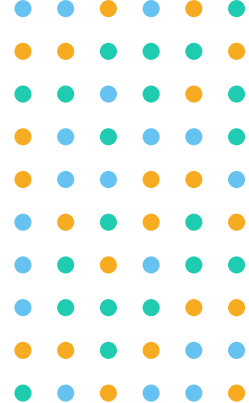
Pull emergency contacts and insurance data from handwritten or typed patient forms.

Ideal for

- Banks and insurance firms.
- Government and e-governance platforms.
- Law firms and judicial systems.
- Hospitals and healthcare providers.
- Recruitment platforms and enterprises.
- EdTech, Fintech, and document-focused startups.

Why JioCloud

- **Tailored for Indian contexts**
Trained on regional formats, informal inputs, and domain-specific documents.
- **Playground for instant testing**
Test real documents before integration.
- **API-first design**
Easy to embed into your systems with minimal setup.
- **Full operational control**
Self-serve provisioning, usage tracking, and credit management.
- **Transparent, affordable billing**
Pay only for what you process — no hidden costs or lock-ins.



Accelerate Document Workflows with JioCloud

Reach us at jpl.cloudsales@ril.com or visit our (website) to get started.

