

OCR

Overview

JioCloud's OCR Service is built to extract text from documents, forms, IDs, receipts, and handwritten inputs—across Indian languages and English. Whether it's a scanned petition, a bank form, or a government application in Hindi or Tamil, our OCR engine reads and converts it into structured, machine-readable text. You can test your inputs in a live Playground, integrate easily through REST APIs, and manage everything through a single dashboard—including service provisioning, credit management, and usage.



Key Features

- Image-to-text extraction**
Reads text from scanned documents, forms, IDs, and receipts—in both printed and handwritten.
- Indian language support**
Recognises scripts including Devanagari (Hindi), Tamil, Telugu, Kannada, Bengali, Marathi, and more—including code-mixed content.
- Printed and handwritten OCR**
Works reliably on both typed and freehand input, even with irregular formatting.
- Live web playground**
Test your own documents or images and view results before integration.
- REST API integration**
Embed OCR into onboarding, KYC, archiving, or case intake systems through secure APIs.
- Self-service provisioning**
Manage service lifecycle, credits, and quotas from a unified interface.
- Transparent billing**
Pay only for the pages or images processed, with full usage visibility.

Benefits

- Digitise documents written in Indian languages with high accuracy
- Save time by automating manual data entry from forms, petitions, or receipts
- Speed up customer, citizen, and student onboarding with faster data extraction
- Validate documents faster—across banking, legal, and public workflows
- Preview results before going live through a self-serve Playground
- Scale flexibly with usage-based pricing and full operational control

Supported Language

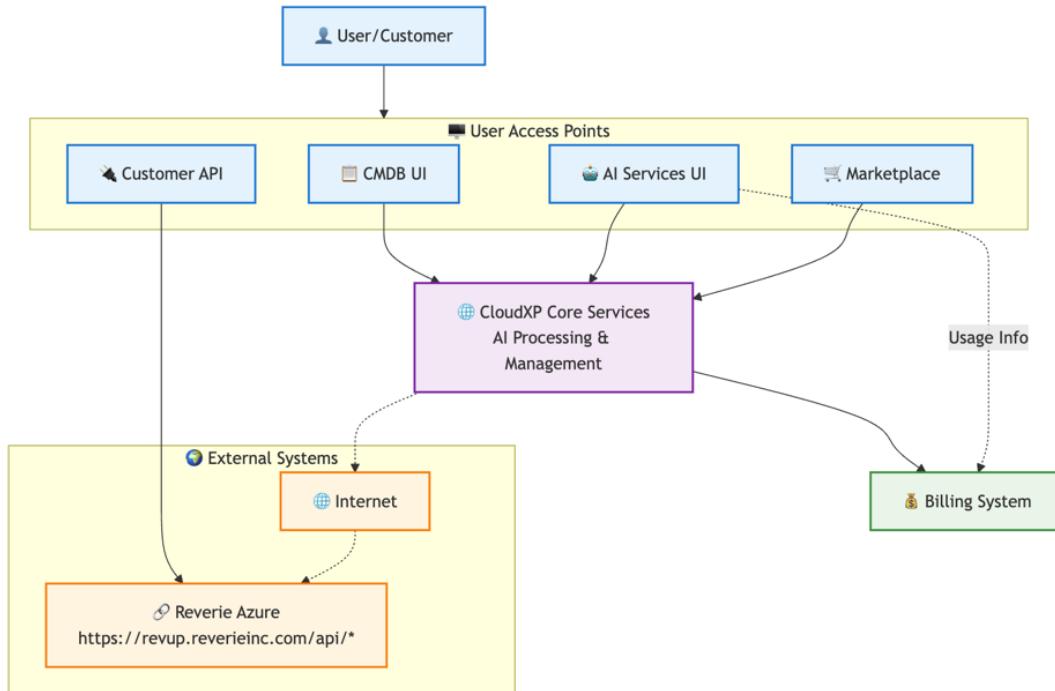
Language	Support
Hindi	Full
Bengali	Full
Tamil	Full
Telugu	Full
Kannada	Full
Malayalam	Full

Language	Support
Gujarati	Full
Punjabi	Full
Odia	Full
Assamese	Full
Urdu	Full
Sanskrit	Full
English	Full

Technical Specifications

Category	Details
Input Format	Image (JPEG, PNG, PDF – scanned)
Output Format	Text data
Image Requirements	~
Access Protocol	REST API over HTTPS
Authentication	Bearer Token
Average Latency	~
Web Playground	Upload image and view extracted text
Deployment Model	Fully managed SaaS
Billing Model	Per-image or per-character pricing
Security	HTTPS; no image storage or logging

Architecture Diagram



Use Cases

- **eGovernance scheme application processing**
Extract applicant details from physical forms in Hindi, Tamil, or Bengali—speeding up verification, reducing data entry, and improving transparency.
- **Court petition filing automation**
Scan handwritten petitions and auto-extract case data like petitioner name, category, and hearing type—reducing clerical load and improving intake speed.
- **Bank KYC and loan form digitisation**
Convert filled-in application forms into digital records instantly, enabling faster validation, fewer manual errors, and quicker decision-making.