JioCloud

Entity Extraction

Overview

JioCloud Entity Extraction Service helps you identify key information from unstructured content in Indian languages, Hinglish, and regional code-mixed inputs. It detects person names, locations, phone numbers, email IDs, hashtags, and more—helping you structure messy inputs for downstream automation, search, analytics, or routing. Test results instantly in a web-based Playground or integrate directly into your backend with REST APIs. Full-service control, usage-based pricing, and no vendor lock-in included.

Key Features

- **Comprehensive entity coverage**
  Extracts names, locations, phone numbers, email IDs, organisations, URLs, amounts, dates, times, hashtags, and more.

- **Built for India's language landscape**
  Trained on Hindi, Tamil, Telugu, Bengali, Kannada, Marathi, Malayalam, Assamese, Odia, and English—including transliterated and code-mixed inputs.

- **Works across text types**
  Handles chats, support messages, emails, feedback forms, and other informal or semi-structured content.

- **Live playground**
  Upload real samples and preview detected entities instantly before deployment.

- **REST API integration**
  Seamlessly plug into CRMs, helpdesks, form parsers, or analytics pipelines.

- **Self-service operations**
  Set usage limits, top up credits, and manage services independently.

- **Pay-as-you-go billing**
  No minimums, no lock-in—pay only for what you use.

![JioCloud]

Benefits

- Structure unorganised content in regional Indian languages

- Automate tagging, classification, and routing of incoming requests

- Save manual effort and reduce response turnaround time

- Improve analytics and search on multilingual datasets

- Enable AI use cases with accurate, ready-to-use extracted metadata

- Monitor usage and spend with real-time transparency

Supported Language

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

| Language	Support Type
Marathi	Full
Gujarati	Full
Punjabi	Full
Odia	Full
Assamese	Full
English	Full

Supported Entities

-["location","amount","money","number","organisation","person_name"

Technical Specifications

Category	Details
Processing Type	Named Entity Recognition (NER)
Entities Extracted	Person, Location, Organisation, Date, Time, Quantity, etc.
Input Format	UTF-8 encoded plain text
Output Format	JSON with entity types, spans, and confidence scores
Access Protocol	REST API over HTTPS
Authentication	Bearer Token
Average Latency	~
Rate Limiting	Tier-based; adjustable quotas available

Web Playground	Browser-based UI for instant validation
Deployment Model	Fully managed SaaS
Billing Model	Per-character consumption-based
Security	HTTPS with zero data retention

Architecture Diagram
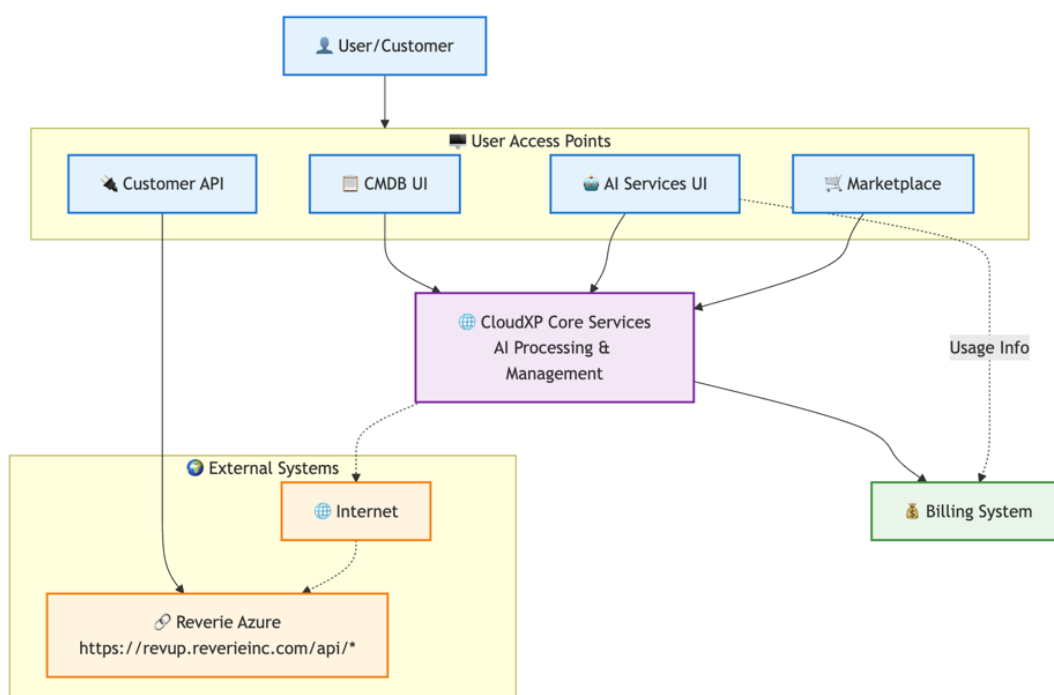
Use Cases

- **Customer support automation**
  Extract user details and issue types from multilingual emails
  and messages to auto-create support tickets.

- **Brand monitoring and social listening**
  Scan multilingual feedback for brand mentions, campaign
  hashtags, or complaints involving money or service.

- **Smarter eCommerce search**
  Parse phrases like "Nike white sneakers under 5k" to identify
  brand, color, category, and filters—even in code-mixed search queries.