


Java Overview [HOME](#)  [GUIDES](#)  [JAVA](#)  [OVERVIEW](#)

Java OCR SDK for extracting text from PDFs



COPY PAGE



Nutrient Java SDK includes advanced optical character recognition (OCR) capabilities, enabling seamless text extraction and interaction within PDFs. By leveraging OCR, you can transform scanned documents and image-based PDFs into machine-readable, actionable data with unparalleled accuracy and ease.

TRY FOR FREE

LAUNCH DEMO

Enhancing text accessibility in PDFs

Many PDF documents, especially scanned files or photographed pages, contain inaccessible text that's stored as images rather than selectable and searchable. The OCR component transforms these documents by recognizing and converting raster- and vector-based text into an interactive format, unlocking powerful features such as:

- ✧ Text selection and extraction
- ✧ Text search functionality
- ✧ Text markup annotations



ASK AI

Key benefits of Nutrient Java SDK's OCR feature

- ✧ **Comprehensive functionality** — OCR is part of a larger suite of document processing tools, allowing seamless integration with other PDF-handling features.
- ✧ **Flexible processing** — Customize OCR workflows by limiting processing to specific page ranges or selecting recognition languages to suit your document needs.
- ✧ **Seamless integration** — Designed for effortless incorporation into existing Java applications, ensuring a smooth adoption process.
- ✧ **Automated data entry** — Extract structured data from documents and populate fields automatically, streamlining document processing.

Licensing and customization

OCR is an add-on for Nutrient Java SDK. If you're interested in enabling OCR, learning more about its development roadmap, or sharing feedback and feature requests, contact our [Sales team](#).

Language support

OCR supports text recognition in multiple languages. For a complete list of supported languages, refer to the [supported languages for Java OCR guide](#).

FAQs

What is an OCR SDK?



An OCR SDK equips developers with libraries and tools to implement optical character recognition, enabling applications to recognize and process text from diverse document types.

What is OCR used for?



Can I process handwritten text?



How do I get started?



Was this helpful?

✓ YES

✗ NO



Copyright 2025 Nutrient. All rights reserved.