



DATALEHTI · LUO · AI-OCR

Skannattu, mutta ei haettavissa? Nyt jokainen skannaus muuttuu tekstiksi.

AI-OCR lukee skannaukset, valokuvat ja PDF-tiedostot ja palauttaa jokaiselle riville tekstin, sijainnin ja luotettavuusarvon. Lisäksi QR- ja viivakoodit, taulukot sekä haettavan PDF-tiedoston. Paperia ja kuvia käsitteleville tiimeille: sovelluksessa, REST-API:n ja MCP:n kautta.

YHDELLÄ SILMÄYKSELLÄ

Kolme tasoa: Basic lukee tekstin sijaintitietoineen, Premium lisää QR- ja viivakoodit sekä haettavan PDF-tiedoston, Ultra tuo lisäksi taulukot, asettelun ja käsinkirjoituksen.

01 · MISTÄ ON KYSE

Skannaus on hakua varten pelkkä kuva. AI-OCR tekee siitä luettavan.

Jokaisen haun kannalta skannattu sivu on vain kuva. Jos sisältö tarvitaan, se kirjoitetaan näppäimistöllä tai selataan käsin. AI-OCR tunnistaa jokaisen sivun tekstin, tallentaa kunkin rivin sijainnin ja antaa kaiken tekstinä, Markdownina, JSON-muodossa tai haettavana PDF-tiedostona.

SYÖTTEET

PNG, JPG, TIFF, BMP, WEBP ja PDF, myös monisivuiset; sovelluksessa enintään 20 MB tiedostoa kohti

TULOS

Teksti, Markdown ja JSON rivikohtaisilla rajauseroilla ja luotettavuusarvolla sekä haettava PDF

KÄYTTÖ

Sovelluksessa, REST-API:n kautta ja MCP-työkaluna `po_aiocr_analyze`

KIELI

Jokaisen sivun kieli tunnistetaan automaattisesti ja sisällytetään tulokseen

**Teksti sijaintitietoineen**

Jokainen tunnistettu rivi tulee rajauseroilla ja luotettavuusarvon kanssa. Näin jokaisen tuloksen voi tarkistaa sivulla ja merkitä.

QR- ja viivakoodit

Premium-tasosta alkaen AI-OCR lukee sivun QR- ja viivakoodit ja palauttaa niiden sisällön sijaintineen tulokseen.

Haettava PDF

Premium-tasosta alkaen syntyy Sandwich-PDF: alkuperäinen skannaus näkymättömällä tekstikerroksella, josta teksti on haettavissa ja kopioitavissa.

Taulukot datana

Ultra tunnistaa rivit ja sarakkeet sisältävät taulukot ja palauttaa ne rakenteisena datana JSON-muodossa sivun asetteluun kanssa.

Käsinkirjoitus

Ultra tunnistaa myös käsinkirjoituksen, kuten lomakkeiden ja lähetekirjojen merkinnät, ja palauttaa sen painetun tekstin tapaan sijaintitietoineen.

Osa AI-DMS:ää

Workspacessa AI-OCR toimii AI-DMS-käsittelyn osana. Käsiteltyjen sivujen teksti on sen jälkeen haussa.

02 · MITEN SE TOIMII

Ladatkaa tiedosto, valitkaa taso. Tarkistakaa tulos ja käyttekää sitä.

Sovelluksessa lataatte tiedoston ja näette tuloksen suoraan sivulla. REST-API:n ja MCP:n kautta sama työ kulkee samoilla kolmella tasolla.

1

Tiedoston lataaminen

Vetäkää kuva tai PDF kenttään tai valitkaa se napsauttamalla: PNG, JPG, TIFF, BMP, WEBP tai PDF enintään 20 MB.

2

Tason valinta ja käynnistys

Valitkaa Valitse OCR-tila : Basic, Premium tai Ultra ja napsauttakaa Aloita AI-OCR . Monisivuiset PDF-tiedostot luetaan sivu kerrallaan.

3

Tuloksen tarkistaminen

Avatkaa välilehti Tekstisijainnit ja napsauttakaa aluetta: teksti, sijainti ja luotettavuus näkyvät vieressä. Taulukot näyttää tunnistetut taulukot.

4

Työn jatkaminen

Käyttäkää toimintoa Lataa hakukelpoinen PDF tai Vie JSON . API:n kautta: POST /job/add/paperoffice_aiocr___generate ja ocr_mode.

TEKNISET TIEDOT

OMINAISUUS	TIEDOT
Syötemuodot	PNG, JPG, TIFF (myös monisivuinen), BMP, WEBP, PDF (monisivuinen)
Tuloste	Teksti, Markdown, JSON rajauslaatikoilla ja luotettavuusarvolla; taulukot ja asettelu JSON-muodossa; Sandwich-PDF
Tasot (ocr_mode)	basic = teksti, premium = ruudukko, ultra = täydellinen
REST-API	POST /job/add/paperoffice_aiocr___generate, tulos vastauksessa tai työtunnuksella
MCP-työkalu	po_aiocr_analyze, tiedosto upload_id-, file_url- tai pofid-tunnuksella Workspacesta
Workspacessa	AI-DMS-käsittelyn yhteydessä, käynnistys: Toiminnot › Agentit ja työpolut › AI-DMS
Kieli	Automaattinen tunnistus sivukohtaisesti

03 · KÄYTÄNNÖSSÄ

Paperista tulee dataa. Ilman näppäilyä ja ilman mallipohjaa.

Vanha arkisto haettavaksi

Skannatut asiakirjat käsitellään Premium-tasolla ja palautetaan Sandwich-PDF-tiedostoina. Sen jälkeen tekstiä voi hakea, merkitä ja kopioida jokaisesta PDF-tiedostosta.

Lähetekirjan rivit

Ultra tunnistaa lähetekirjan rivitaulukon. Rivit tulevat JSON-muodossa rakenteisena taulukkona, joten niitä voi käsitellä edelleen ilman näppäilyä.

Tekstintunnistus omassa sovelluksessa

Kehittäjätiimi lähettää omasta sovelluksestaan kuvia REST-API:lle ja saa takaisin tekstin sijaintitietoineen, esimerkiksi kenttien merkitsemiseksi suoraan kuvaan.

USEIN KYSYTYT KYSYMYKSET

Miten Basic, Premium ja Ultra eroavat?

Basic palauttaa tekstin, Markdownin ja rajauslaatikot luotettavuusarvoineen. Premium lisää QR- ja viivakoodit sekä haettavan PDF-tiedoston. Ultra tuo lisäksi taulukot, asettelun lukujärjestyksineen ja käsinkirjoituksen.

Voiko AI-OCR lukea Workspacen asiakirjoja?

Kyllä. MCP-työkalu po_aiocr_analyze ottaa tallennetun asiakirjan pofid-tunnuksen, esimerkiksi Claudessa tai Cursorissa. AI-DMS-käsittelyn yhteydessä Workspacessa AI-OCR toimii joka tapauksessa.

Pitääkö asiakirjan kieli ilmoittaa?

Ei. AI-OCR tunnistaa jokaisen sivun kielen automaattisesti ja sisällyttää sen tulokseen.

Miten näen, kuinka varma tunnistus on?

Jokaisella rivillä on luotettavuusarvo, ja sovellus näyttää sen sijainnin sivulla. Arvo perustuu PaperOffice AI:n vertailuasteikkoon.

ALOITA

Kokeilkaa vaikeinta skannaustanne.

Kirjautukaa PaperOffice-tilillenne, avatkaa AI-OCR, ladatkaa tiedosto ja napsauttakaa **Aloita AI-OCR**. Veloitus perustuu käyttöön ja riippuu tasosta.

paperoffice.ai/fi/ai-ocr-tekstintunnistus-llm-dokumenttien-digitointi



LISÄTIETOA MODUULISTA

paperoffice.ai/fi/ai-ocr-tekstintunnistus-llm-dokumenttien-digitointi

HINNAT

paperoffice.ai/en/pricing

YHTEYSTIEDOT

paperoffice.ai/fi/yhteystiedot