



# DataLab WestSax

*Zukunfts- und Wertschöpfungslabor Westsachsen: Ein regionaler Katalysator  
Für Datenbasierte Wertschöpfung*

## OCR-TECHNOLOGIEN

*... "Optical Character Recognition" oder auf Deutsch "Optische Zeichenerkennung". OCR-Technologien ermöglichen es, gedruckten oder handgeschriebenen Text auf Papier oder in digitalen Dokumenten in maschinenlesbaren Text umzuwandeln, der dann von Computern und anderen elektronischen Geräten verarbeitet werden kann.*

---

## Inhalt

<i>OCR-Technologien</i> .....	4
<i>Indikatoren für die Nutzung</i> .....	4
<i>Vorteile</i> .....	5
<i>Nachteile</i> .....	6
<i>Use Case</i> .....	7
<i>OCR-Technologien im Vergleich</i> .....	8
<i>OCR-Technologien im Vergleich</i> .....	9
<i>Kontakt</i> .....	10

---

# OCR-TECHNOLOGIEN

*OCR-Technologien sind in der heutigen digitalen Welt allgegenwärtig und bieten Unternehmen und Einzelpersonen zahlreiche Vorteile. OCR steht für "Optical Character Recognition" und bezieht sich auf eine Technologie, die es Computern ermöglicht, Text aus Bildern oder gescannten Dokumenten zu erkennen und zu konvertieren. Diese Technologien ermöglichen es Unternehmen, Papierdokumente in digitale Formate zu konvertieren, wodurch die Verwaltung und Speicherung von Dokumenten vereinfacht und beschleunigt wird.*

*Je größer der Bedarf an Datenverarbeitung in einem Unternehmen ist, desto klarer zeigt sich die Notwendigkeit solcher Technologien, um in der Zeit der zunehmenden Digitalisierung und Automatisierung anzukommen.*



## Indikatoren für die Nutzung

*... umfassen die Notwendigkeit des Zugriffs von mehreren Benutzern auf eine große Menge an Dokumenten und Daten, die täglich manuell erfasst werden müssen. Dies führt zu einem erhöhten Personalaufwand und längeren Bearbeitungszeiten. Zudem resultiert aus der manuellen Eingabe eine hohe Fehlerquote. Besonders relevant ist auch die Verwendung sensibler Daten, die den Datenschutzbestimmungen unterliegen und entsprechend geschützt werden müssen.*

---

# VORTEILE

Die Vorteile von OCR-Technologien sind zahlreich und reichen von der Steigerung der Produktivität und Effizienz bis hin zur Verbesserung der Datensicherheit und -integrität. Einer der größten Vorteile von OCR-Technologien besteht darin, dass sie die manuelle Dateneingabe reduzieren oder sogar komplett eliminieren können. OCR-Technologien ermöglichen eine schnelle und effektive Verarbeitung von großen Mengen an Informationen, was zu einer erheblichen Zeitersparnis führen kann. Ein weiterer Vorteil von OCR-Technologien ist die Verbesserung der Datenqualität. Durch die Verwendung von OCR-Technologien wird die Wahrscheinlichkeit von Tippfehlern und anderen menschlichen Fehlern bei der Dateneingabe minimiert. OCR-Technologien können auch helfen, die Genauigkeit und Konsistenz von Daten zu verbessern, indem sie die Formatierung und Struktur von Dokumenten standardisieren.

Darüber hinaus können OCR-Technologien die Datensicherheit und -integrität verbessern. Durch die automatische Verarbeitung von Dokumenten können OCR-Technologien sicherstellen, dass vertrauliche Informationen nicht in falsche Hände geraten. OCR-Technologien können auch dazu beitragen, die Nachverfolgbarkeit von Änderungen an Dokumenten zu verbessern und so die Integrität von Daten zu gewährleisten.



**Zeitersparnis**



**Schnelle Suche**



**Hohe Genauigkeit**

---

# NACHTEILE

Bei der Verwendung von OCR-Technologien zur Texterkennung können verschiedene Herausforderungen auftreten. Ein häufiges Problem ist die fehlerhafte Erkennung von Text, insbesondere wenn das Originaldokument schwer lesbar ist oder eine ungewöhnliche Schriftart/-größe verwendet wurde. Auch handgeschriebene Texte können Schwierigkeiten bereiten und zu einer geringeren Genauigkeit der Texterkennung führen, was sich negativ auf die Datenqualität auswirkt. Ein weiteres Problem besteht darin, dass OCR-Technologien oft Schwierigkeiten haben, die ursprüngliche Formatierung des Textes beizubehalten. Insbesondere bei der Umwandlung von PDF-Dateien mit komplexen Layouts oder Tabellen können Formatierungsfehler und unleserliche Texte entstehen.

Sprachbarrieren können ebenfalls zu Problemen führen, da OCR-Technologien Schwierigkeiten haben können, Texte in verschiedenen Sprachen zu erkennen und zu verarbeiten. Insbesondere bei der Verarbeitung von Texten mit Schriftarten, die nicht in der Standardbibliothek des OCR-Systems enthalten sind, kann die Genauigkeit der Texterkennung leiden. Ein weiteres wichtiges Thema bei der Verwendung von OCR-Technologien ist der Kostenfaktor. OCR-Systeme können teuer sein, insbesondere wenn sie in umfangreichen Projekten eingesetzt werden. Zusätzliche Kosten können auch für Wartung, Schulung und technischen Support entstehen. Schließlich ist es ebenfalls wichtig, darauf zu achten, dass OCR-Technologien ausreichend abgesichert sind, um die Vertraulichkeit und Integrität personenbezogener Daten wie Namen, Adressen und Sozialversicherungsnummern zu gewährleisten. Datenschutz sollte bei der Verwendung von OCR-Technologien immer eine zentrale Rolle spielen.

**Genauigkeits- und Kostenprobleme sowie Datenschutzbedenken können bei der Verwendung von OCR-Technologien auftreten.**

---

# USE CASE

Ein Beispiel für die Anwendung von OCR-Technologien in der Rechnungsverarbeitung ist das Unternehmen ABC. ABC erhält jeden Tag Hunderte von Papierrechnungen, die in einem manuellen Prozess bearbeitet werden müssen. Hierfür setzt das Unternehmen eine OCR-Software ein, die die Daten auf den Rechnungen erkennt und automatisch in digitale Formate umwandelt. Die OCR-Software liest die Daten auf der Rechnung, wie beispielsweise den Rechnungsbetrag, den Lieferanten- oder Kundenname, das Rechnungsdatum und die Artikelbeschreibung.

Die OCR-Software überprüft auch die Genauigkeit der Daten und korrigiert etwaige Fehler automatisch. Nachdem die Daten extrahiert wurden, werden sie in die Buchhaltungssoftware des Unternehmens importiert, wodurch der manuelle Aufwand für die Eingabe der Daten entfällt. Durch den Einsatz von OCR-Technologien konnte ABC den Prozess der Rechnungsverarbeitung beschleunigen und die Fehlerquote reduzieren, was zu einer erheblichen Kosteneinsparung und Effizienzsteigerung führte.



## Use Cases

- *Digitale Archivierung*
- *Automatisierte Dateneingabe*
- *Texterkennung in Bildern*
- *Automatisierte Formularverarbeitung*
- *Übersetzung von Texten*
- *Automatisierte Texterkennung bei medizinischen Untersuchungen*

# OCR-TECHNOLOGIEN IM VERGLEICH

	TESTVERSION	ABONNEMENT	VIDEO TUTORIALS	VERGLEICH VON DOKUMENTEN	MAC OS KOMPATIBEL
ABBYY Finereader 16	✓	✓	✓	✓	✓
Nuance Omnipage 19 Ultimate	✓	✗	✗	✓	✗
Avanquest Expert PDF 15 Ultimate	✓	✗	✗	✓	✗
Perfect PDF 11 Premium Edition	✓	✗	✓	✓	✗

	TESTVERSION	ABONNEMENT	VIDEO TUTORIALS	VERGLEICH VON DOKUMENTEN	MAC OS KOMPATIBEL
Adobe Acrobat Pro DC	✓	✓	✓	✗	✓
Microsoft OneNote	✓	✗	✓	✗	✓
Aiseeoft PDF Converter Ultimate	✓	✓	✗	✗	✓
Soda PDF 11 Premium + OCR	✓	✗	✓	✗	✗

	TESTVERSION	ABONNEMENT	VIDEO TUTORIALS	VERGLEICH VON DOKUMENTEN	MAC OS KOMPATIBEL
PDF Director Pro	✓	✗	✓	✗	✗
Konfuzio OCR Software	✓	✓	✓	✗	✓



# OCR-TECHNOLOGIEN IM VERGLEICH

	VORTEILE	NACHTEILE	SCHNITTSTELLEN
ABBYY Finereader 16	KI-basierte OCR-Technologie -> hohe Erkennungsrate	Abo-Modell, teurer Jahrespreis	.pdf, .csv, .xls, .jpg
Nuance Omnipage 19 Ultimate	planbare Stapelverarbeitung, 120+ Sprachen werden erkannt, akustische Ausgabe	hohe Systemvoraussetzungen, begrenzte Aktivierungen bei Neuinstallation	.pdf, .csv, .xls, .jpg
Avanquest Expert PDF 15 Ultimate	gruppierte Ausgabe in Formate, wie z.B. Word, Excel, lebenslange Lizenz	Durchsuchen in eingescannten Dokumenten in nur 8 Sprachen	.pdf, .csv, .xls, .jpg, .html
Perfect PDF 11 Premium Edition	gedrucktes Handbuch	Lizenz für nur 3 Geräte	.pdf, .csv, .xls, .jpg
	VORTEILE	NACHTEILE	SCHNITTSTELLEN
Adobe Acrobat Pro DC	Mobile App verfügbar, kostenlose Updates	Abo-Modell	.pdf, .csv, .xls, .jpg
Microsoft OneNote	erkennt handschriftliche Notizen, Mobile App verfügbar, kostenlose Updates	Office 365-Abonnement erforderlich, langsame Verarbeitung	.pdf, .csv, .xls, .jpg
Aiseeoft PDF Converter Ultimate	200+ Sprachen werden erkannt, Layouteinstellungen möglich, Batch-Konvertierung	teures Lizenzmodell (5 Geräte ab 250 €)	.pdf, .csv, .xls, .jpg
Soda PDF 11 Premium + OCR	Stapelverarbeitung, Zugriff auf Dokumente auch mobil möglich (Cloud)	langsame Verarbeitung der Dateien, zusätzliche Kosten trotz Vollversion	.pdf, .csv, .xls, .jpg
	VORTEILE	NACHTEILE	SCHNITTSTELLEN
PDF Director Pro	Texterkennung und -weiterverarbeitung in neuen Dokumenten möglich, kostenlose Updates	Lizenz nur für 3 Geräte	.pdf, .csv, .xls, .jpg
Konfuzio OCR Software	planbare Stapelverarbeitung, 120+ SprKI basierte Technologie, unlimitierte Anzahl an Nutzern pro Projekt, erkennt 70+ Sprachen im Text und 7 Sprachen in Handschrift, akustische Ausgabe	Abonnement, Programmierung und ständige Kontrolle der KI notwendig, kundenspezifische Anpassung sehr teuer, erst bei großen Projekten sinnvoll	.pdf, .csv, .xls, .jpg, .json

---

**Melden Sie sich gern!  
Daten. Wissen(s)Wert.**

## **KONTAKT**

*E-Mail: [kontakt@datalab-vestsax.de](mailto:kontakt@datalab-vestsax.de)*

*Webseite: <https://datalab-vestsax.de/>*

**Fraunhofer IMW Leipzig**  
*Neumarkt 9-19  
04109 Leipzig*

**Westfälische Hochschule**  
*Zwickau  
Kornmarkt 1  
08056 Zwickau*

### **Besucheradresse**

*Westfälische Hochschule Zwickau  
Fakultät Wirtschaftswissenschaften  
Professur Wirtschaftsinformatik  
Scheffelstraße 39  
Haus 3, Zimmer 3111  
08066 Zwickau*