

ABBYY® FlexiCapture Engine 10

Das umfangreiche SDK für Data Capture



Was ist FlexiCapture Engine 10?

FlexiCapture Engine 10 ist das erste umfassende Data Capture SDK, welches Technologien und Tools zur intelligenten Dokumentenklassifizierung und Formularverarbeitung kombiniert. Mit dem SDK können Entwickler eigene Applikationen mit leistungsstarken Funktionalitäten zur automatisierten Extraktion von Daten aus Dokumenten ausstatten. FlexiCapture Engine erleichtert die nahtlose Integration von komplexen Data Capture Funktionalitäten zur Verarbeitung von semi-strukturierten und unstrukturierten Dokumenten wie Rechnungen, Lieferscheinen oder Verträgen. Mit FlexiCapture Engine verschaffen Sie Ihren Applikationen, egal ob Desktop oder Serveranwendung, einen zusätzlichen Mehrwert und können gleichzeitig Entwicklungskosten reduzieren.

Erweiterung Ihrer bestehenden Applikation oder Entwicklung automatisierter Data Capture Lösungen

FlexiCapture Engine 10 liefert eine große Bandbreite neuester Technologien für Datenextrahierung, Datenklassifizierung und Indizierung. Bereits entwickelte Lösungen können schnell um den Datenexport zu Backend-Systemen oder Dokumentenkonvertierung erweitert werden. Mit FlexiCapture Engine erhalten Entwickler alles was sie brauchen aus einer Hand: Tools zur Entwicklung, Training aber auch Support und Wartung.

Intelligente Dokumenten-Verarbeitung

Automatisierte Datenextrahierung aus Formularen, Verarbeitung von semi- und unstrukturierten Dokumenten und Dokumentenklassifikation in einem einzigen SDK.

Verbesserte Verarbeitung von mehrseitigen Dokumenten

Einzigartige Technologie zur vereinfachten Verarbeitung von mehrseitigen Dokumenten sowie Tabellen, die sich über mehr als eine Seite erstrecken.

Zuverlässige Verarbeitungsplattform

Das leistungsstarke SDK entspricht Data Capture Industriestandards zur Datenverarbeitung in Unternehmensprojekten.

Automatisierte Datenextraktion in Backend-Applikationen

Einfache Integration von Daten-Export-Funktionalitäten in Backend-Applikationen anhand von Standardschnittstellen.

Einsatzmöglichkeiten in der internationalen Dokumentenverarbeitung

Unterstützung verschiedener Sprachen zur Entwicklung von lokalisierten Anwendungen im internationalen Bereich.

Neue Auto-Learning Funktionen

Die neue Auto-Learning API und ein Beispiel im Quelltext erlauben es, einfache Dokumente schnell über eine eigene Point-&-Click Oberfläche zu trainieren, um dann aus ähnlichen Dokumenten die Daten auszulesen.

Kosteneinsparungen

Reduktion der Kosten bei der Datenextraktion durch Integration einer automatisierten Verarbeitung.

Effiziente Entwicklung

Standard-Komponenten ermöglichen eine schnelle Integration des SDK in existierende Applikationen und verkürzen dadurch die Entwicklungszeit.

Qualitativ hochwertige Datenextraktion

Erstklassige OCR-Erkennungstechnologien in Kombination mit regelbasierten und automatisierten Verifizierungsmethoden garantieren eine optimierte und valide Datenausgabe.

Nahtlose Integration in Applikationen

FlexiCapture Engine garantiert Entwicklern eine nahtlose Integration von Data Capture Technologien in ihre eigenen Applikationen.

Funktionen

- Integration intelligenter Data Capture Funktionalitäten zur Verarbeitung von Dokumenten
- Dokumentenklassifizierung und automatische Dokumententrennung
- Leistungsstarke Datenextrahierung für semi- und unstrukturierte Dokumente
- Verifizierungs- und Überwachungsoptionen für eine optimale Datenausgabe
- Qualitativ hochwertige Erkennungstechnologien für OCR, ICR, OMR, 1D und 2D Barcodes
- Sprachunterstützung für bis zu 198 OCR und 112 ICR Sprachen
- Automatischer Datenexport in Backend-Applikationen
- PDF, MRC PDF & PDF/A Dokumentenexport zur Archivierung

Vorteile für Entwickler

- Erweitern Sie Ihre Applikationen um Data Capture Komponenten
- Einzigartige und leistungsstarke Features sowie API für Datenextraktion
- Gebrauchsfertige SDK-Komponenten zur Beschleunigung des Entwicklungsprozesses
- Vollständige Kontrolle über Dokumentenverarbeitung, extrahierte Daten und das Routing von Dokumenten
- Kompatibel mit FineReader Engine 10 API, um weitere Funktionen zur Dokumentenkonvertierung zu integrieren

Übersicht über den Entwicklungsprozess



Integration, Entwicklung von Dokumentbeschreibungen und Verarbeitung

Die Integration von Data Capture Funktionalitäten mit FlexiCapture Engine 10 in Unternehmensprojekte ist ein zweistufiger Prozess. Zur Integration der Dokumenten- und Datenverarbeitungsfunktionalitäten können Entwickler die bereitgestellten FlexiCapture SDK Komponenten verwenden (Schritt 1). Für die Verarbeitung unterschiedlicher Dokumententypen in einem Verarbeitungsprozess müssen intelligente Dokumentenbeschreibungen oder Dokumentenvorlagen entwickelt und getestet werden (Schritt 2). Das FlexiCapture Engine Entwicklungskit enthält alle nötigen Design-Tools, einen Document Template Editor für strukturierte Formulare, FlexiLayout Studio für unstrukturierte Dokumente und eine API die es erlaubt einfache, einseitige FlexiLayouts über eine eigene Benutzeroberfläche zu erzeugen.

Applikationsintegration und Dokumentenverarbeitung

FlexiCapture Engine hat eine einfach strukturierte und leicht zu verwendende API. Die Entwicklung kann in C/C++, Visual Basic, .Net, Delphi und anderen COM-kompatiblen Tools erfolgen. Die folgenden Schritte zeigen, wie Sie Ihre Applikationen um Data Capture Funktionalitäten erweitern können:

Schritt 1 – Integration:

Integrieren Sie die FlexiCapture Engine DLLs in Ihre Applikationen, Workflows oder Geschäftsanwendungen.

Schritt 2 – Projektauswahl:

Laden Sie ein vordefiniertes Set von Dokumentvorlagen, die mit der neuen Auto-Learning API erstellt oder mit FlexiLayout Studio entwickelt und getestet wurden.

Schritt 3 – Dokumentverarbeitung:

Laden Sie die Dokumente und PDFs, die verarbeitet werden sollen. Dokumententrennung, Klassifizierung und Erkennung werden automatisch ausgeführt, basierend auf der Logik der FlexiCapture 10 Dokumentendefinitionen.

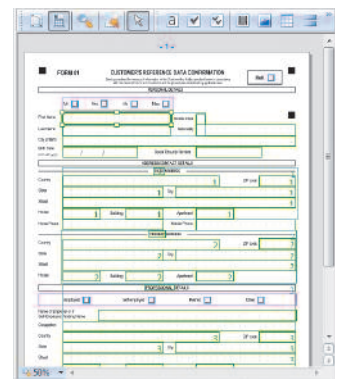
Schritt 4 – Datenverifizierung:

Die API für die Datenverifizierung von FlexiCapture Engine erlaubt Ihnen den Zugriff auf die internen Erkennungsergebnisse. Die Visual Components, erlauben eine schnelle Implementierung der manuellen Verifizierung.

Schritt 5 – Verarbeitung von Dokumenten und Daten:

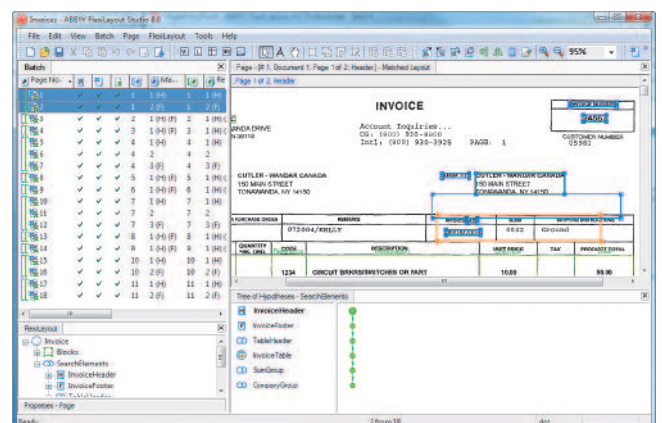
Alle extrahierten Daten können nun in Ihrer Applikation verarbeitet, an andere Applikationen geschickt oder in Form von XML-, XLS- und CSV-Dateien exportiert werden. Die verschiedenen Dokumente können sowohl als Bild-dateien als auch als durchsuchbare PDFs oder PDF/A-Dateien exportiert werden.

- 3) Anschließend werden die Erkennungsbereiche mit Hilfe eines graphischen Editors festgelegt, z. B. Textblöcke, Tabellen, Kontrollkästchen, Barcodes oder Bildbereiche.
- 4) Zuletzt werden die Erkennungsoptionen festgelegt, also die Definitionen ob der Inhalt mit OCR oder ICR gelesen wird und welche Datentypen, Wörterbücher und Verifizierungsregeln angewendet werden sollen.



Entwicklung von mehrseitigen FlexiLayout Templates mit FlexiLayout Studio

- 1) Zunächst wird eine Auswahl repräsentativer Dokumente mit unterschiedlichem Layout geladen.
- 2) Spezielle Elemente zur Identifikation und Orientierung innerhalb eines Dokumentes werden bestimmt, z. B. Texte, Barcodes, Linien oder auch Abstände zwischen Elementen.
- 3) Im letzten Schritt erfolgt die Definition der Elemente, die die gesuchten Daten aufnehmen sollen. Dies können Texte unterschiedlicher Länge, Zahlen, ein Datum oder ganze Tabellen sein. Zusätzlich werden diese Elemente in geometrischen Bezug zu den unter 2.) definierten Anker gesetzt, z. B. „Lies den Text rechts neben „Rechnungsnummer“, der zwischen 8-16 Zeichen lang ist und Zahlen enthält“.



Entwicklung von Dokumenten-Templates für Formulare

Mit dem Document Template Editor können schnell und intuitiv Dokumenten-Templates zur Verarbeitung von Formularen erstellt werden.

- 1) Verschiedene Segmente eines mehrseitigen, strukturierten Formulars werden in den Editor geladen.
- 2) Dann erfolgt die Definition der Elemente, die zum „Template Matching“ verwendet werden, z. B. Anker (Quadrat, Kreuze), statischer Text oder Linien.

Funktionen im Überblick

ABBYY FlexiCapture Technologie für Data Capture

Einzigartige Technologien für OCR, ICR, OMR und Barcode-Erkennung

Die Kerntechnologien von FlexiCapture Engine verarbeiten bis zu 198 Sprachen bei OCR und 112 Sprachen bei ICR, und bieten zudem die Erkennung von Ankreuzfeldern und Barcodes.

Verschiedene Dokumentarten auf einmal verarbeiten

FlexiCapture Engine 10 bietet fertige Komponenten, die es Entwicklern ermöglichen, die automatische Extrahierung von Daten aus einem Stapel von Formularen und Dokumenten in eigene Anwendungen zu integrieren. Die automatische Erkennung der unterschiedlichen Dokumententypen ermöglicht die Verarbeitung von verschiedenen Dokumententypen (strukturiert, semi-strukturiert, unstrukturiert) mit Hilfe von Formularvorlagen und FlexiLayouts.

Dokumententrennung

FlexiCapture Engine bietet einfache und erweiterte Dokumententrennung. Die einfache Trennung basiert auf Leerseiten und Barcodes, die erweiterte auf intelligenten Dokumentendefinitionen. Die Anwendung der Templates auf einen Dokumentenstapel kann ohne manuelles Eingreifen ausgeführt werden.

Dokumentenklassifizierung

Sobald ein Dokument zu einem Formulartemplate oder einem FlexiLayout passt, wird automatisch ein Dokumenttyp klassifiziert, z. B. Kundenumfrage, Rechnung, Vertrag.

Verarbeitung von „fixen“ Formularen

FlexiCapture Engine ist perfekt dafür geeignet, jede Art von Formularen zu verarbeiten, beispielsweise Fragebögen, Studien oder Coupons.

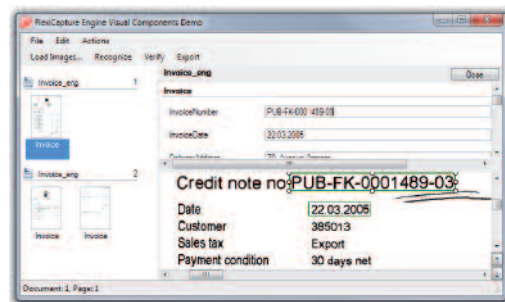
Verarbeitung von mehrseitigen Dokumenten mit verschiedenen Layouts

Geschäftsdokumente bestehen oft aus mehreren Seiten. Mit FlexiCapture Engine bekommen Entwickler eine einzigartige Technologie an die Hand, die es ihnen ermöglicht, auch mehrseitige Dokumente mit Tabellen einfach zu verarbeiten.

Datenkontrolle und Verifizierung

FlexiCapture Engine bietet effektive Möglichkeiten, um die erfassten Daten zu überprüfen und Fehler aufzudecken:

- Daten können bei Bedarf automatisch überprüft werden, das umfasst eine einfache Bereichsprüfung und den Abgleich bekannter Werte aus einer Datenbank. Zudem können durch Skripte eigene Regeln festgelegt werden.
- Die API von FlexiCapture Engine erlaubt die Entwicklung eigener Benutzeroberflächen. Mit den Visual Components (Source Code Samples) können Entwickler den Anwendern eine effektive, ergonomische Benutzeroberfläche bereitstellen.



Daten- und Dokumentenexport

ABBYY FlexiCapture Engine 10 ist in der Lage, gewünschte Daten aus einem Dokument auszulesen und es dann anhand dieser zu indexieren. Für Archivierungszwecke können zudem Dokumente in hochkomprimierte (MRC) PDFs oder PDF/A-Dokumente konvertiert werden.

Java Wrapper

Das FlexiCapture Engine 10 enthält auch Java-Klassen, so kann die Funktionalität des SDKs auch ohne Mehraufwand aus Java-Anwendungen genutzt werden.

Auto-Klassifizierung

FlexiLayout Studio 10 unterstützt neue Klassifizierungstechniken, die Raster und Inhaltsklassifizierung intelligent miteinander verbinden. Die Technologie ist einfach einzurichten und garantiert eine schnelle Verarbeitung. Dokumenten-Training ist auch über die API möglich.

Dokumententemplates

Fixed Form Templates:
Fragebögen, Coupons, Formulare, ...

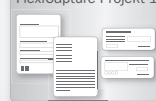


FlexiLayouts:
Rechnungen, Verträge, Briefe, ...



Sammlung von Templates

FlexiCapture Projekt 1



Dynamisches Laden von Templates



Applikation



Spezifikationen und Lizenzierung

Spezifikationen

Systemanforderungen

- PC mit x86-kompatiblen Prozessor (1 GHz oder höher).
- Betriebssystem:
Microsoft® Windows Server® 2008, Windows 7®, Windows Vista®, Windows Server 2003, Windows® XP, Windows 2000 und 64-Bit Versionen von Windows Server 2008, Windows 7®, Windows Vista, Windows Server 2003, Windows XP
- Speicher: mind. 1 GB RAM

Cloud Umgebungen

- Windows Azure
- Amazon EC2

Eingabeformate

BMP, PCX, DCX, JPEG, JPEG 2000, JBIG2, PNG, GIF, TIFF, DjVu, PDFs (bis zu PDF 1.7)

Ausgabeformate

- Dokument-Export: XLS, DBF, CSV, TXT, XML
- Dokumentexport: BMP, PCX, DCX, JPEG, JPEG 2000, PNG, TIFF, JBIG2, durchsuchbare, hochkomprimierte (MRC) PDF und PDF/A

Entwicklung

FlexiCapture Engine API unterstützt den COM-Standard und kann einfach mit C/C++, Visual Basic, .Net, Delphi und anderen Tools implementiert werden. Das SDK enthält Codebeispiele in C/C++, VB 6, VB.Net, C#, Java und Delphi.

Weitere Informationen online:

http://www.abbyy.com/flexicapture_engine/technical_specifications/

ABBYYs Lizenzpolitik

ABBYY FlexiCapture wird über eine flexible, modulare Lizenzpolitik vertrieben, die es Entwicklern ermöglicht, die beste Kombination von Tools und Preisoptionen für ihr Projekt auszuwählen. Die Lizenzierung wird angeboten als:

Entwicklerlizenzen

Diese erlauben es, FlexiCapture Engine in eigene Applikationen zu integrieren und zu testen. Die Lizenz beinhaltet FlexiCapture 10 Standalone und FlexiLayout Studio zur Entwicklung von Templates für Formulare sowie von FlexiLayouts.

Runtime Lizenzen

Erlauben den Vertrieb von Applikationen, die FlexiCapture Engine 10 Funktionalitäten beinhalten. Runtime Lizenzen (RTLs) unterscheiden sich in ihrer Funktionalität, in der Anzahl an Seiten die pro Jahr (ohne CPU Limit) verarbeitet werden können, oder der Anzahl an CPU Cores (ohne Seitenzähler). ABBYY Europe liefert standardmäßig Netzwerk-lizenzen. Diese werden zentral über einen im LAN installierten Lizenzmanager bereitgestellt, laufende Instanzen, egal ob virtuell oder physikalisch, können dann die Lizenz nutzen.

Zusätzliche Module für Runtime Lizenzen

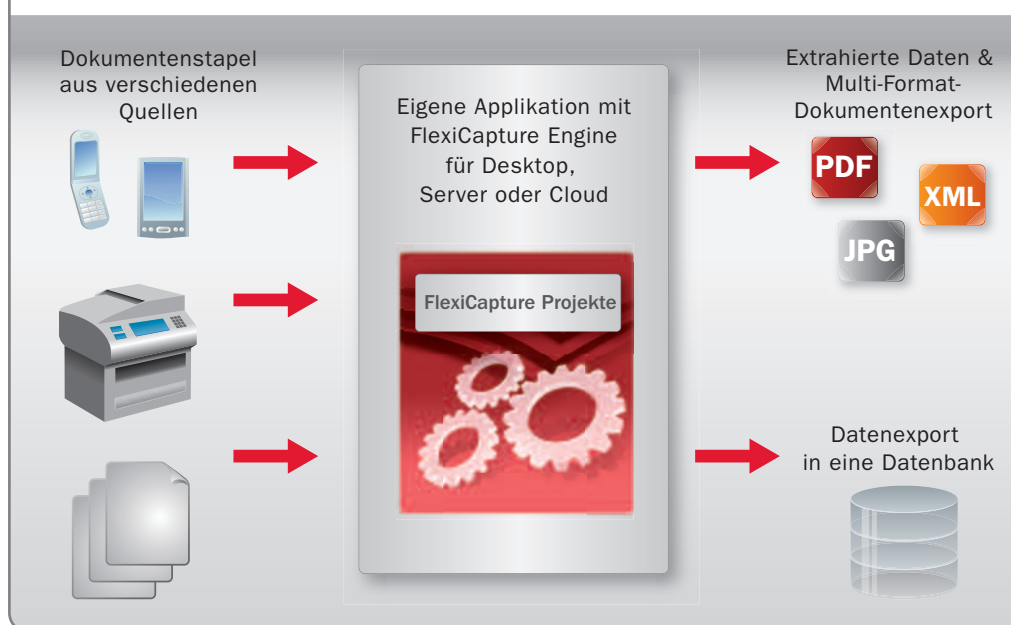
RTLs können durch das Hinzufügen von Funktionalitäten, wie z. B. weiterer OCR Sprachen oder PDF Export, die als Add-On Modul angeboten werden, erweitert werden. Auch FineReader Engine 10 API* für Volltext-OCR und Dokumentumwandlung können mit FlexiCapture Engine lizenziert werden.

Support, Wartung und Upgrade-Versicherung (SMUA), Zertifizierungstrainings und Professional Services

Um den Erfolg Ihrer Produkte zu sichern, bietet ABBYY zusätzliche Unterstützung, Trainings und Zertifizierungsprogramme für alle Produkte. Wenn Sie Ihr Projekt beschleunigen möchten, kontaktieren Sie bitte ABBYY um Details zu Professional Services zu erhalten.

* FineReader Engine 10 API ist im neuen Lizenzmodell von ABBYY Europe enthalten, dies gilt nicht für andere Vertriebsgebiete.

ABBYY FlexiCapture Engine 10 Dokumenten-Verarbeitung



ABBYY®

ABBYY Europe GmbH
Elsenheimerstr. 49, D - 80687 München
Tel: +49 89 511 159 0
sales_eu@abbyy.com
www.ABBYY.com

ABBYY UK
sales_UK@abbyy.com

ABBYY France
sales_france@abbyy.com

ABBYY Spain
sales_es@abbyy.com

ABBYY Italy
sales_it@abbyy.com

ABBYY Benelux
sales_benelux@abbyy.com

ABBYY Scandinavia
sales_nordic@abbyy.com