

OCI Dokument AI Scanner

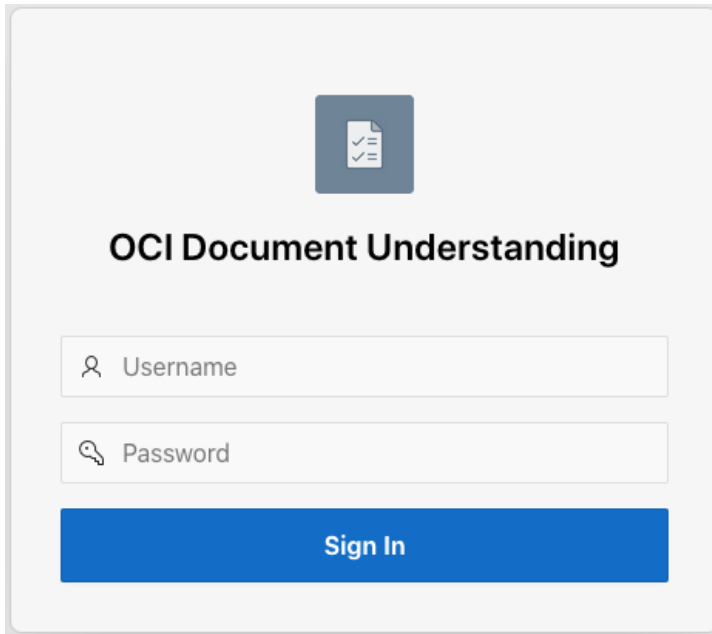
Die Anwendung erschliesst die Funktionen des Oracle Cloud – Document Understanding AI Service in einer modernen APEX Benutzeroberfläche. Das Ziel der Anwendung ist es mehrere Pdf - oder Bild-Dateien in einen Job zu verarbeiten und die Ergebnisse der Analyse (Textextraktion, Tabellenextraktion, Dokumentklassifizierung, Schlüsselwertextraktion) als ZIP-Datei zum Download oder Export in die Buchhaltungsdatenbank bereitzustellen. Auf der Seiten Details zum Prozessorauftrag und Dokumenten Analyse können die Details angezeigt und verbessert werden.

Inhaltsverzeichnis

Anmelden auf der Login Seite.....	3
Sprache der Benutzeroberfläche.....	3
Erscheinungsbild der Benutzeroberfläche.....	4
Processor Aufträge erstellen.....	5
Dokumente Verarbeiten.....	6
Ergebnisse herunterladen.....	6
Ergebnisse auswerten.....	6
Auftrag Löschen.....	7
Details zum Prozessorauftrag.....	8
Rechnungen Exportieren.....	9
Dokumenten Analyse.....	10
Dokumentattribute.....	11
Feldbeschriftungen.....	12
Einzelposten.....	14
Rechnungsposten.....	15
Tabellen.....	16
Textzeilen.....	17
Text Vorschau.....	17
Dokumente suchen.....	18
Konfiguration.....	19
AI Prozessor Konfiguration.....	19
Konfiguration – Feldtypen.....	23
Konfiguration – Einzelpostentypen.....	24
Konfiguration – Sprachen.....	25
Konfiguration - Text Index Status.....	26
Installation.....	27
Voraussetzungen.....	27
OCI-Konfiguration.....	27
APEX Anwendung installieren.....	29
Verarbeitung von Prozessor Aufträgen.....	29
1. Dateien im Objekt Store ablegen.....	29
2. Prozessorauftrag zusammenstellen und starten.....	30
3. Abwarten der Verarbeitung.....	30
4. Aufbereitung der Ergebnisse.....	30
5. Algorithmus zur Analyse von Textfragmenten.....	30

Anmelden auf der Login Seite

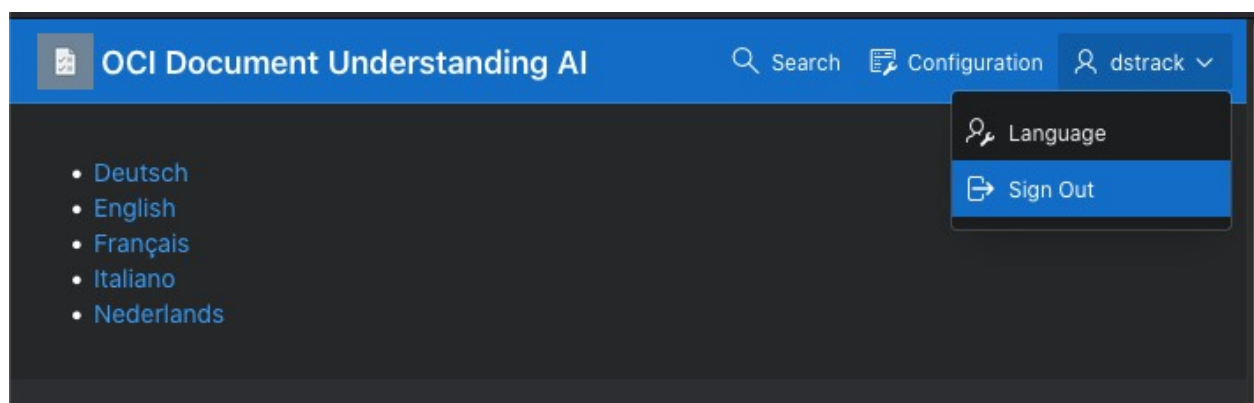
Die Anwendung benötigt den Benutzer Name und das Passwort des APEX Account der von Administrator des APEX Workspaces eingerichtet wurde.



The login form for OCI Document Understanding is displayed within a light gray rounded rectangle. At the top center is a blue square icon containing a white document with a checkmark. Below the icon, the text "OCI Document Understanding" is centered in bold black font. Underneath, there are two input fields: the first is labeled "Username" with a small person icon, and the second is labeled "Password" with a small key icon. At the bottom of the form is a prominent blue button with the white text "Sign In".

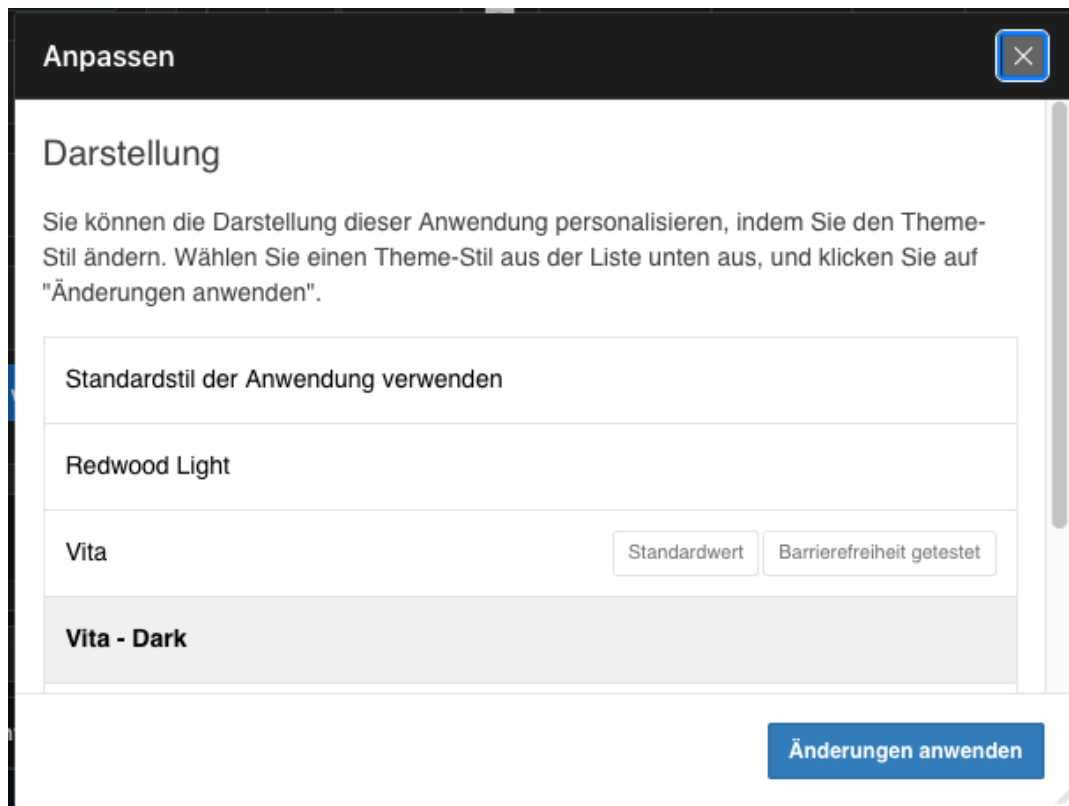
Sprache der Benutzerüberfläche

Nach der Anmeldung kann die Sprache der Benutzeroberfläche geändert werden. Klicke dazu oben rechts auf das Benutzer Icon mit ihrem Namen und wähle Language. Eine Liste von Sprachen erscheint oben links. Wähle die passende Sprache mit einem Klick auf einen Eintrag aus. Zur Zeit werden die Sprachen Deutsch, Englisch, Französisch, Italienisch und Holländisch unterstützt.



Erscheinungsbild der Benutzeroberfläche

Klicke unten links auf Anpassen und ein Dialogfenster mit dem Titel Darstellung erscheint. Sie können die Darstellung dieser Anwendung personalisieren, indem Sie den Thema-Stil ändern. Wählen Sie einen Thema-Stil aus der Liste unten aus, und klicken Sie auf "Änderungen anwenden".



Processor Aufträge erstellen

Auf der Home Page können neue Aufträge erstellt und erledigte Aufträge gesichtet werden. Für einen neuen Auftrag lade mehrere Dateien per drag und drop in die dafür vorgesehene Zone oben links auf der Seite. Stelle den Dokumenttyp, die Sprache und den Kontext der Dokumente ein.

Für den Import von Rechnungen tellen Sie den Dokumenttyp auf **Rechnung**, Sprache auf **Englisch** und den Kontext auf den passenden Projektnamen ein.

Mit der Spracheinstellung Englisch wird sichergestellt, dass auch amerikanisch formatierte Zahlen und Datumswerte konvertiert werden können, wenn diese in deutschen Dokumenten vorkommen.

Mit der Dokumenttypeinstellung Rechnung wird sichergestellt, dass die passenden Feldbezeichnungen gesucht werden, auch wenn von der Dokumentanalyse ein anderer Dokumenttyp ermittelt wurde.

Nur einzelne Rechnungen je PDF Datei

Die Scans der Rechnungen sollten nur eine einzelne Rechnung enthalten. Das Programm kann mehrere Rechnungen in einem PDF-Dokument NICHT erkennen und zerlegen! Auf dem Mac kann mit dem Programm Vorschau per Drag and Drop einer Seite auf den Desktop ein neues Dokument je Seite erstellt werden, wenn das notwendig wird.

OCI Document Understanding AI

Suchen Konfiguration dstrack

Startseite

Dokumenten-KI-Scanner

Neue Dokumente verarbeiten

Daten per Drag-and-Drop verschieben

Dateien auswählen oder löschen in dieser Region.

Dateien auswählen

Dokumenttyp

Rechnung

Sprache

Englisch

Kontext

KAP SixSenses

Tabellenextraktion

Durchsuchbares PDF generieren

Dokumente verarbeiten

KI-Prozessor-Jobs

Aktionen Bearbeiten Speichern

Zurücksetzen

	Auftragsid	Herun	Lebenszyklus	Dokumenttyp	Sprache	Kontext	Dateien	Datum der letzten Akt	Zuletzt akti	Dokum	Schlüsselw	Einzelposten	Wurde exporti
<input type="checkbox"/>	000176		SUCCEEDED	Rechnung	English	KAP SixSenses	WF_Maklerprovision_Re...	03-NOV-2023 13:04	DSTRACK	1	13	3	N
<input type="checkbox"/>	000175		SUCCEEDED	Rechnung	German	KAP SixSenses	Lindner_AR 400166.pdf	13-OKT-2023 13:14	DSTRACK	1	61	425	Y
<input type="checkbox"/>	000174		SUCCEEDED	Rechnung	German	KAP SixSenses	Webmedia_Huber_Re220...	12-OKT-2023 15:09	DSTRACK	5	92	25	N
<input checked="" type="checkbox"/>	000172		SUCCEEDED	Rechnung	English	KAP SixSenses	Oracle_Re_300002332.p...	12-OKT-2023 11:49	DSTRACK	2	42	11	N
<input type="checkbox"/>	000171		SUCCEEDED	Rechnung	English	KAP SixSenses	Stadtgem_Mittersill_Grun...	06-OKT-2023 13:00	DSTRACK	12	197	96	N
<input type="checkbox"/>	000170		SUCCEEDED	Rechnung	German	KAP SixSenses	Notar_Zimmerer_Re_1334...	06-OKT-2023 12:57	DSTRACK	12	276	310	N
<input type="checkbox"/>	000169		SUCCEEDED	Rechnung	English	KAP SixSenses	230309_Octopus.pdf, 23...	06-OKT-2023 12:53	DSTRACK	13	301	292	N

1 Zeilen ausgewählt

Gesamt 7

Dokumente

Suche: Volltextsuche

Los

Aktionen

Zurücksetzen

	Dokument	Dateiname	Sprache	Sprachbewertung	Dokumenttyp	Dokumenttypbewertung	Seiten	Einzelposten	Schlüsselwerte	Tabellen
<input type="checkbox"/>		Oracle_Re_300002332.pdf	German	75.4%	Andere	99.0%	1	5	21	0
<input type="checkbox"/>		Oracle_Re_300006785.pdf	German	75.2%	Andere	51.3%	1	6	21	0

1 Zeilen ausgewählt

Gesamt 2

Dokumenttyp

Wählen Sie den Typ der hochgeladenen Dokumente aus. Der Dokumenttyp bestimmt welche Art für Feldinhalten und Einzelposten in der Auswertung erzeugt wird.

Sprache

Wählen Sie die Sprache der hochgeladenen Dokumente aus. Die Sprache bestimmt welche länderspezifischen Datenformate bei der Konvertierung zu Zahlen und Datumsangaben verwendet werden.

Kontext

Wählen Sie den Kontext für die hochgeladenen Dokumente aus. In dem Fall das die Ergebnisse nach der Analyse z.B. einem Projekt zugeordnet werden sollen, können Sie hier das Projekt einstellen.

Tabellenextraktion

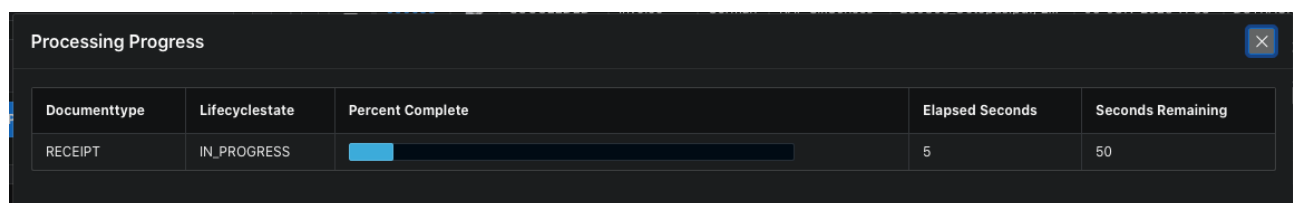
Aktivieren Sie diese Option wenn Sie die in den Dokumenten enthaltenen Tabellen auswerten möchten. Tabellenextraktion kann nur mit einigen Dokumenttypen verwendet werden.

Durchsuchbares PDF erstellen

Aktivieren Sie diese Option wenn Sie die hochgeladenen Dokumente in durchsuchbares PDF umwandeln wollen. Damit werden diese Dokumente mit einem Oracle Text Index durchsuchbar gemacht.

Dokumente Verarbeiten

Starte den Auftrag mit einem Klick auf ‚Dokumente Verarbeiten‘. Ein Dialogfenster informiert während der Ausführung über den Fortschritt und die erledigten Prozente des Auftrags.



The image shows a 'Processing Progress' dialog window with a close button in the top right corner. It contains a table with the following data:

Documenttype	Lifecyclestate	Percent Complete	Elapsed Seconds	Seconds Remaining
RECEIPT	IN_PROGRESS	<div><div></div></div>	5	50

Wenn der Auftrag erledigt ist, verschwindet das Dialogfenster und die **Details zum Prozessorauftrag** werden angezeigt.

Ergebnisse herunterladen

Klicke auf das ZIP-Datei Icon in der Spalte Herunterladen der KI-Prozessor-Jobs Liste um den Download zu starten. In der Zip-Datei ist eine Verzeichnisstruktur mit den diversen Ergebnislisten enthalten. Die Ordner hat den Namen Document_AI_Files_X und endet mit der Jobnummer. Enthalten sind die Verzeichnisse csv_files und searchable_pdf.

In csv_files finden die Dateien die sie in Numbers oder Excel auswerten können. Je nach Art des Auftrags befinden sich hier Dateien mit den Namen der Dokumente und den Endungen X_Line_Items.csv und X_Tables.csv. Außerdem die Dateien Invoices_List_X.csv, Receipts_List_X.csv, Driver_License_List_X.csv, Passport_List_X.csv mit Feldwerten für jedes entsprechende Dokument.

Das Verzeichnis searchable_pdf enthält die durchsuchbaren PDF Dateien für die hochgeladenen Dokumente.

Ergebnisse auswerten

Wähle einen Eintrag in der Liste KI-Prozessor-Jobs mit einen Klick auf die gewünschte Zeile. Für den gewählten Eintrag werden die einzelnen Dokumente in der Liste Dokumente aufgelistet. Klicke auf die Job-Nummer in der Spalte Auftragsdetails um Details zum Prozessorauftrag aufzurufen, oder klicke auf das Lupen Icon in der Spalte Dokument um die Dokumenten Analyse-Ergebnisse aufzurufen.

Auftrag Löschen

Nachdem Sie die Ergebnisse heruntergeladen oder ausgewertet haben kann der Auftrag gelöscht werden. Wählen Sie dazu ‚Zeile löschen‘ aus dem Zeilenmenü und klicken Sie danach auf Sichern. Mit dem Löschvorgang wird der verwendete Speicher in der Datenbank und auch im Object Store wieder freigegeben.

Details zum Prozessorauftrag

Auf dieser Seite können die Kenndaten des Auftrags und eine schnelle Sichtung der erkannten Felddbeschriftungen mit Inhalten für Rechnungen mit Rechnungsposten und Quittungen mit Quittungseinzelposten angezeigt werden. Klicke auf das Lupen Icon in der Spalte Dokument um die Dokumenten Analyse aufzurufen.

OCI Document Understanding AI

Suchen Konfiguration dstrack

Startseite

Details zum Prozessorauftrag

Vorherige Schlüsselklasse zuweisen In die Buchhaltung exportieren ZIP-Datei herunterladen Einen anderen verarbeiten Weiter

Job-ID
000172

Erstellungsdatum
12-Okt-2023 11:49

Erstellt von
DSTRACK

Lebenszyklusstatus
SUCCEEDED

Lebenszyklusdetails

Processorjob Message

Kontext
KAP SixSenses

Sprachcode
English

Dokumenttyp
Rechnung

Dokumente
2

Schlüsselwerte
42

Tabellen
0

Einzelposten
11

Tabelleneextraktion
Nein

Wurde exportiert
Nein

Wurde heruntergeladen
Nein

Wurde bereinigt
Ja

Löschen Adressen neu verknüpfen Speichern

Rechnungen

Suchen: Alle Textspalten Los Rechnungs-Kenndaten Aktionen Zurücksetzen

Doku	Dateiname	Anbietername	Anbieteradresse	Vendor Logo	Adresse des verknüpften Anl	Rechnungs-ID	Zwischensum	Gesamtsum	Rechnungssumme	Fälliger Betrag	Rechnungsda	Gesamtsteu	Fälligkeit
	Oracle_Re_300002332.pdf	Oracle Austria GmbH	Donau-City-Strasse 7, WI...	ORACLE	Oracle Austria GmbH, Don...	300002332	1.050,00	210,00	1.260,00	1.260,00	22.11.2021		22.12.20
	Oracle_Re_3000006785.pdf	0 Oracle Austria GmbH	Donau-City-Strasse 7, WI...	ORACLE	Oracle Austria GmbH, Don...	300006785	1.050,00	210,00	1.260,00	1.260,00	23.05.2022		22.06.20

1 Zeilen ausgewählt Gesamt 2

Rechnungsposten

Suchen: Alle Textspalten Los Aktionen

Seite

Linie	Name	Menge	Einheit	Einzelpreis	Betrag	Steuer	Datum	Produktcode
Seite: 1								
1	connectPM Plattformsteuerung von 01/11/21 bis 31/0...				1.050,00			
2				1.050,00	1.050,00			

Gesamt 2

Rechnungen Exportieren

Mit einem Klick auf ‚In die Buchhaltung Exportieren‘ wird eine Vorschau der exportieren Daten angezeigt. Mit einem Klick auf ‚Export ausführen‘ werden die angezeigten Daten als Eingangsrechnungen in die Buchhaltungsdatenbank übertragen. Die durchsuchbaren Pdf Dateien werden dabei als Anlagen zu den Rechnungen in die Buchhaltungsdatenbank übertragen. Nach dem Export wird der Button ausgeblendet um einen doppelten Export zu verhindern.

Rechnungen in die Buchhaltungsdatenbank exportieren												
Job-ID: 000093											Export ausführen	
Document_Id ↑	Job_Id	File_Name	Doc_Ustidnr	Workspace\$_Id	Eingangsdatum	Faelligkeitsdatum	Belegdatum	Belegjahr	Abrechnungsjahr	Posten_Nr	Forderung	Geaenderte_Forderung
242	93	230309_Octopus.pdf		124	05.06.2023	23.03.2023	05.04.2023	2023	2023	20230258	525,02	525,02
243	93	230329_Inng_Moser_0431_2023.pdf		124	05.06.2023		12.04.2023	2023	2023		8951,26	8951,26
244	93	230323_Archt_Werkstatt_RE_234_ZL_10.pdf	0043 5355 65898	124	05.06.2023		23.03.2023	2023	2023	234	3120	3120
245	93	230331_MBLD_RE_5522.pdf	583 0464 38	124	05.06.2023	31.03.2023	31.03.2023	2023	2023	5522	10990	10990
246	93	230401_Praschberger_RE_1010676.pdf		124	05.06.2023		01.04.2023	2023	2023		165	165
Herunterladen Drucken												

Dokumenten Analyse

Auf der Seite Dokumentenergebnisse wird links die durchsuchbare PDF Ansicht des hochgeladenen Dokuments angezeigt. Im unterem Bereich der Ansicht kann das Dokument als PDF Datei heruntergeladen werden.

Oben rechts kann aus den möglichen Auswertungen gewählt werden. Abhängig vom erkannten Dokumenttyp werden hier Dokumentattribute, Feldbeschriftungen, Einzelposten, Rechnungsposten, Quittungseinzelposten, Tabellen und Textzeilen.

The screenshot displays the OCI Document Understanding AI web interface. The main window shows a scanned invoice from Oracle with the title 'RECHNUNG'. The invoice details include a total amount of 1,260.00 EUR, a net amount of 1,050.00 EUR, and a tax amount of 210.00 EUR. The invoice is dated 22-DEC-2021 and has a reference number 300002332. The right sidebar contains a navigation menu with options like 'Dokumentattribute', 'Feldbeschriftungen', 'Textbeschriftungen', 'Einzelposten', 'Rechnungsposten', and 'Textzeilen'. The 'Dokumentattribute' section is currently active, showing a list of attributes such as 'Job-ID', 'Dateiname', 'Sprachcode', 'Dokumenttyp', 'Jobsprachencode', 'Seitenanzahl', 'Einzelposten', 'Gesamtbetrag', 'Übereinstimmende Kundenadresse', and 'Übereinstimmende Anbieteradresse'. The 'Feldbeschriftungen' section is also visible, showing a list of fields with their respective labels and values.

Dokumentattribute

Attribut	Wert
Job-ID	000172
Dokument-ID	369
Dateiname	Oracle_Re_300002332.pdf
Mime-Typ	application/pdf
Sprachcode	German
Sprachbewertung	75.4%
Dokumenttyp	Andere
Dokumenttypbewertung	99.0%
Jobsprachencode	English
Jobdokumenttyp	Rechnung
Seitenanzahl	1
Schlüsselwerte	21
Einzelposten	5
Tabellen	0
Gesamtbetrag	1,260,00
Übereinstimmende Kundenadresse	Kitzbühler Alps Projekt GmbH, Obere Gänsbachgasse 7 6370 Kitzbühel (72500)
Übereinstimmende Anbieteradresse	Oracle Austria GmbH, Donau-City-Strasse 7 1220 Wien (73502)

Feldbeschriftungen

Aktionen	Bearbeiten	Speichern	Zurücksetzen
Bestätigt	Feldbezeichnung T:2	Label	Feldalias
Feldtext	Werttyp	Zahlenw	

Dokumentattribute

Hier werden neben den Dokumentkenndaten (Job-ID, Dateiname, Mime-Typ) auch die erkannte Sprache und der Dokumenttyp angezeigt. Die Sprache und der Dokumenttyp können ggf. geändert werden damit die passenden Zahlen -und Datumsformate bei der Auswertung angewendet werden können. Für Rechnungen und Quittungen werden die Anzahl der Seiten, Feldwerte, Einzelposten und Tabellen, der Gesamtbetrag und die zugeordneten Adressen angezeigt.

Dokumentattribute	
Job-ID 000126	
Dateiname Quittung nachos.pdf	Mime-Typ application/pdf
Sprachcode English	Sprachbewertung 56.5%
Dokumenttyp Quittung	Dokumenttypbewertung 94.1%
Jobsprachencode German	Jobdokumenttyp Quittung
Seitenanzahl 1	Schlüsselwerte 7
Einzelposten 3	Tabellen 0
Processorjob Message	
Gesamtbetrag 12,00	
Übereinstimmende Händleradresse Roswitha Huber KG (60568)	
Speichern	

Feldbeschriftungen

Die von Prozessor gefundenen Feldbeschriftungen werden für die Ausgabe der Rechnung -und Quittungsdaten aufgelistet. Der Anwender kann die vom Prozessor gefundenen Daten Zuordnungen prüfen und verbessern.

1. Mehrere Möglichkeiten

Wenn Felder mehrfach vorkommen kann der Anwender mit einen Klick in der Spalte ‚Bestätigt‘ den gewünschten Eintrag auswählen.

2. Datenkonvertierung

In der Spalten Zahlenwert und Datumswert wird der Feldtext oder Feldwert als konvertierte Zahl oder Datum ausgegeben. Wenn die Daten nicht wie erwartet konvertiert werden, kann die Sprache für ein Dokument so eingestellt werden, das die passenden Formatmasken und Währungssymbole angewendet werden. Dabei können je Sprache mehrere Varianten zum Einsatz kommen. Wenn Beispielsweise für den Prozessor-Job die Sprache **German** gewählt wurde und der Prozessor findet die Sprache **English** für ein Dokument, dann werden 4 Varianten für die Erkennung und Konvertierung angewendet.

Sprache	Territory	Datumsformat	Zahlenformat	Währung Symbol	Währung Code
English	AMERICA	MM/DD/RRRR	999,999,999.99	\$	USD
English	UNITED KINGDOM	DD-MON-RRRR	999,999,999.99	£	GBP
German	GERMAN	DD.MM.RRRR	999.999.999,99	€	EUR
German	AUSTRIA	DD.MM.RRRR	999 999 999.99	€	EUR

3. Felddeschriftungen Zuordnung

In der Spalte Benutzerbezeichnung kann die Feldbezeichnung neu zugeordnet werden.

Wenn es für ein Feld mehrere Möglichkeiten gibt, kann mit einem Klick in der Spalte Bestätigt der gewünschte Eintrag ausgewählt werden.

Felddescriptions

Hauptbericht

Aktionen

Bearbeiten

Speichern

Zurücksetzen

Seite

Multiple choices

	Bestätigt	Feldbezeichnung ↑ ±2	Feldtext	Feldwert	Label-Score	Werttyp	Zahlenwert	Datumswert	Benutzerb
Seite: 1									
	<input type="checkbox"/>	CustomerAddress	5730 Mittersill, Austria AUS...	5730 Mittersill, Austria AUS...	98	STRING			
	<input type="checkbox"/>	InvoiceId	AR0000400166	AR0000400166	100	STRING			
	<input type="checkbox"/>	ServiceEndDate	31.03.2023	1680220800000	84	DATE		31.03.2023 00:00:00	
	<input type="checkbox"/>	ServiceStartDate	28.02.2023	1677542400000	77	DATE		28.02.2023 00:00:00	
	<input type="checkbox"/>	VendorNameLogo	X Lindner	X Lindner	100	STRING			
Seite: 2									
	<input type="checkbox"/>	InvoiceDate	07.02.2023	1682560000000	89	DATE		02.07.2023 00:00:00	
	<input type="checkbox"/>	VendorNameLogo	X Lindner	X Lindner	100	STRING			
	<input type="checkbox"/>	VendorTaxId	ATU74785958	ATU74785958	90	STRING			
Seite: 5									
	<input type="checkbox"/>	InvoiceTotal	12.225.418,94 EUR		85	NUMBER	12.225.418,94		
Seite: 7									
	<input type="checkbox"/>	DueDate	23.Apr.23	1682208000000	73	DATE		23.04.2023 00:00:00	
	<input type="checkbox"/>	VendorAddress	Obere Gnsbachgasse 6370...	Obere Gnsbachgasse 6370...	100	STRING			
	<input type="checkbox"/>	VendorAddressRecipient	Kitzbheler Alps Projekt Gm...	Kitzbheler Alps Projekt Gm...	99	STRING			
	<input type="checkbox"/>	VendorName	Kitzbheler Alps Projekt Gm...	Kitzbheler Alps Projekt Gm...	99	STRING			
	<input type="checkbox"/>	VendorNameLogo	Lindner	Lindner	100	STRING			
Seite: 8									
	<input type="checkbox"/>	SubTotal	11.843.578,47		79	NUMBER	11.843.578,47		
	<input type="checkbox"/>	VendorNameLogo	Lindner	Lindner	100	STRING			
Seite: 9									
	<input type="checkbox"/>	VendorNameLogo	Lindner	Lindner	100	STRING			

1 Zeilen ausgewählt

Gesamt 17

Einzelposten

Für die erkannten Dokumentarten Rechnung und Quittung werden vom Prozessor gefundene Einzelposten aufgelistet. Wenn in der Auflistung Rechnungsfelder vom Benutzer erkannt werden, kann mit einem Klick in der Spalte Benutzerfeldbezeichnung ein passender Eintrag ausgewählt werden. Diese Felder werden dann als Einträge in der Liste Felderbeschriftungen aufgeführt. So kann dann z.B. der Rechnungsbetrag markiert werden, wenn der Prozessor das nicht konnte.

Einzelposten						
<div>Suchen: Alle Textspalten Los Aktionen Bearbeiten Speichern Zurücksetzen</div>						
Seite						
<input type="checkbox"/>	Linie ↑±2	Feldbezeichnung	Feldwert	Werttyp	Benutzerbezeichnung	Feldtext
Seite: 1						
<input checked="" type="checkbox"/>	1	Description	AR0000400052	STRING		AR0000400052
<input type="checkbox"/>	1	Name	AR0000400052	STRING		AR0000400052
<input type="checkbox"/>	1	Date	1638230400000	DATE		30.11.2021
<input type="checkbox"/>	1	Amount	646.21985	NUMBER		646.219,82
<input type="checkbox"/>	2	Description	AR0000400054	STRING		AR0000400054
<input type="checkbox"/>	2	Name	AR0000400054	STRING		AR0000400054
<input type="checkbox"/>	2	Date	1640908800000	DATE		31.12.2021
<input type="checkbox"/>	2	Amount	564.4588	NUMBER		564.458,80 EUR
<input type="checkbox"/>	3	Description	AR0000400068	STRING		AR0000400068
<input type="checkbox"/>	3	Name	AR0000400068	STRING		AR0000400068
<input type="checkbox"/>	3	Date	1643155200000	DATE		26.01.2022
<input type="checkbox"/>	3	Amount	353.07324	NUMBER		353.073,25 EUR
<input type="checkbox"/>	4	Description	AR0000400078	STRING		AR0000400078
<input type="checkbox"/>	4	Name	AR0000400078	STRING		AR0000400078
<input type="checkbox"/>	4	Date	1667433600000	DATE		11.03.2022
<input type="checkbox"/>	4	Amount	560.5008	NUMBER		560.500,78
1 Zeilen ausgewählt		Gesamt 376				

Rechnungsposten

Für den Dokumenttyp Rechnung wird aus den Einzelposten die Liste der Rechnungsposten gebildet.

Erkannt und konvertierte Zahlen und Datumsfelder werden einheitlich formatiert ausgegeben.

Rechnungsposten								
Aktionen ▾								
▶ ⌵ Seite								
Linie...	Name	Menge	Einheit	Einzelpreis	Betrag	Steuer	Produktcode	Datum
▼ Seite: 1								
1	AR0000400052				646.219,82			30.11.2021
2	AR0000400054				564.458,80			31.12.2021
3	AR0000400068				353.073,25			26.01.2022
4	AR0000400078			560.599,79	67.271.975,00			03.11.2022
5	AR0000400083			119.415,33	143.298,40			04.12.2022
6	AR0000400091				522.071,87			05.10.2022
7	AR0000400096				1.060.241,33			06.08.2022
8	AR0000400106				2.055.174,14			13.07.2022
9	AR0000400112				1.273.887,42			16.08.2022
10	AR0000400117			871.518,62	1.045.822,34			13.09.2022
11	AR0000400123 AR0000400137				970.122,71			11.07.2022
▼ Seite: 2								
1				328.799,35	394.559,22			
2					553.680,18			
3					458.208,56			
▼ Seite: 3								
1	Invoice No. AR-20266 Oberrater Bau GesmbH				240.000,00			
2	10% Overhead on 240.000. € Pos. 1 (Cost plus fee)				24.000,00			
3	5% Profit on 264.000 € Pos. 1 + 2 (Cost plus fee)				13.200,00			
4	Prelims 8.623.445 € / 33 months duration (According to site meet...				261.316,52			
5	Invoice No. AR-210328 Oberrater Bau GesmbH				181.009,36			
6	10% Overhead on Pos. 2 - 181.009,36 € (Cost plus fee)				18.100,94			
7	5% Profit on Pos. 5 + 6 - 199.110,29 € (Cost plus fee)				9.955,51			
8	Prelims 8.623.445,00 € / 33 months duration (According to site m...				261.316,52			
9					27.280,54			
Gesamt 120								

Tabellen

Hier werden die vom Prozessor in den Dokumenten erkannten Tabellen aufgelistet.

In der Report-Ansicht wird die erste Zeile einer Tabelle markiert angezeigt. In dem Popup Feld ‚Tabelle‘ kann eine Tabelle aus einer Liste aller Tabellen des Dokuments ausgewählt werden.

Table								
Q								
Los Aktionen								
2 Page Number, Table Number Markieren: Header								
Page Number: 1, Table Number: 1								
Col001	Col002	Col003	Col004	Col005	Col006	Col007	Col008	Col009
Open installment invoices:	Net	VAT (%)	Gross					
AR0000400052 30.11.2021	538.516,52 EUR	107.703,30 EUR (20,00)	646.219,82 EUR					
AR0000400054 31.12.2021	470.382,33 EUR	94.076,47 EUR (20,00)	564.458,80 EUR					
AR0000400068 26.01.2022	294.227,71 EUR	58.845,54 EUR (20,00)	353.073,25 EUR					
AR0000400078 11.03.2022	560.599,79 EUR	112.119,96 EUR (20,00)	672.719,75 EUR					
AR0000400083 12.04.2022	119.415,33 EUR	23.883,07 EUR (20,00)	143.298,40 EUR					
AR0000400091 10.05.2022	435.059,89 EUR	87.011,98 EUR (20,00)	522.071,87 EUR					
AR0000400096 08.06.2022	883.534,44 EUR	176.706,89 EUR (20,00)	1.060.241,33 EUR					
AR0000400106 13.07.2022	1.712.645,12 EUR	342.529,02 EUR (20,00)	2.055.174,14 EUR					
AR0000400112 16.08.2022	1.061.572,85 EUR	212.314,57 EUR (20,00)	1.273.887,42 EUR					
AR0000400117 13.09.2022	871.518,62 EUR	174.303,72 EUR (20,00)	1.045.822,34 EUR					
AR0000400123 07.11.2022	808.435,59 EUR	161.687,12 EUR (20,00)	970.122,71 EUR					
AR0000400137 14.12.2022	1.014.136,32 EUR	202.827,26 EUR (20,00)	1.216.963,58 EUR					
AR0000400142 10.01.2023	1.244.838,09 EUR	248.967,62 EUR (20,00)	1.493.805,71 EUR					
AR0000400149 07.02.2023	946.216,34 EUR	189.243,27 EUR (20,00)	1.135.459,61 EUR					
Page Number: 1, Table Number: 2								
Col001	Col002	Col003	Col004	Col005	Col006	Col007	Col008	Col009
Headquarters: Bahnhofstrasse 29	Chairman of the Board: Hans Lindner Managing Directors:	Bank: HypoVereinsbank, IBAN DE39						
94424 Arnstorf	Heinrich Bichner, Helmut Lang, Patrick Ossiander, Martin Weber	Landshut 7432 0073 0027 9846 73 980/1173						
Germany	Registered Office: Arnstorf	HYVEDEMM433						
Phone +49 8723 20-0	HRB 12135	BIC Tax-No.						
	Commercial Register Local Court: Landshut							

Textzeilen

Hier werden die vom Prozessor in den Dokumenten erkannten Textzeilen aufgelistet.

Wird eine Textzeile entsprechend zu den Formaten der Dokumentsprache als Datum oder Zahl erkannt, dann werden in den Spalten Value Type, Number Value und Date Value entsprechende Einträge angezeigt.

Textzeilen

Suchen: Alle Textspalten

Los

Aktionen

Zurücksetzen

Seitenzahl

Zeilenr. ↑±2	Textzeile	Value Type	Number Value	Date Value
Seitenzahl: 1				
1	Octopus	STRING		
2	Werbegestaltung & Montagebau GmbH	STRING		
3	Kitzbühler Alps Projekt GmbH	STRING		
4	ATU74785958	STRING		
5	Kontakt:	STRING		
6	Obere Gnsbachgasse 7	STRING		
7	Wasserfeld 1a	STRING		
8	6370 Kitzbhel	STRING		
9	6361 Hopfgarten im Brixental	STRING		
10	Mobil: +43 /676/9616336	STRING		
11	Mobil: +43 /676 / 5339782	STRING		
12	Datum:	STRING		
13	09.03.2023	DATE		09.03.2023
14	Rechnungs-Nr.:	STRING		
15	20230258	NUMBER	20.230.258,00	
16	Kunden-Nr.:	STRING		
17	2265	NUMBER	2.265,00	
18	Sachbearbeiter/-in:	STRING		
19	Mathias Gandler	STRING		
Gesamt 169				

Text Vorschau

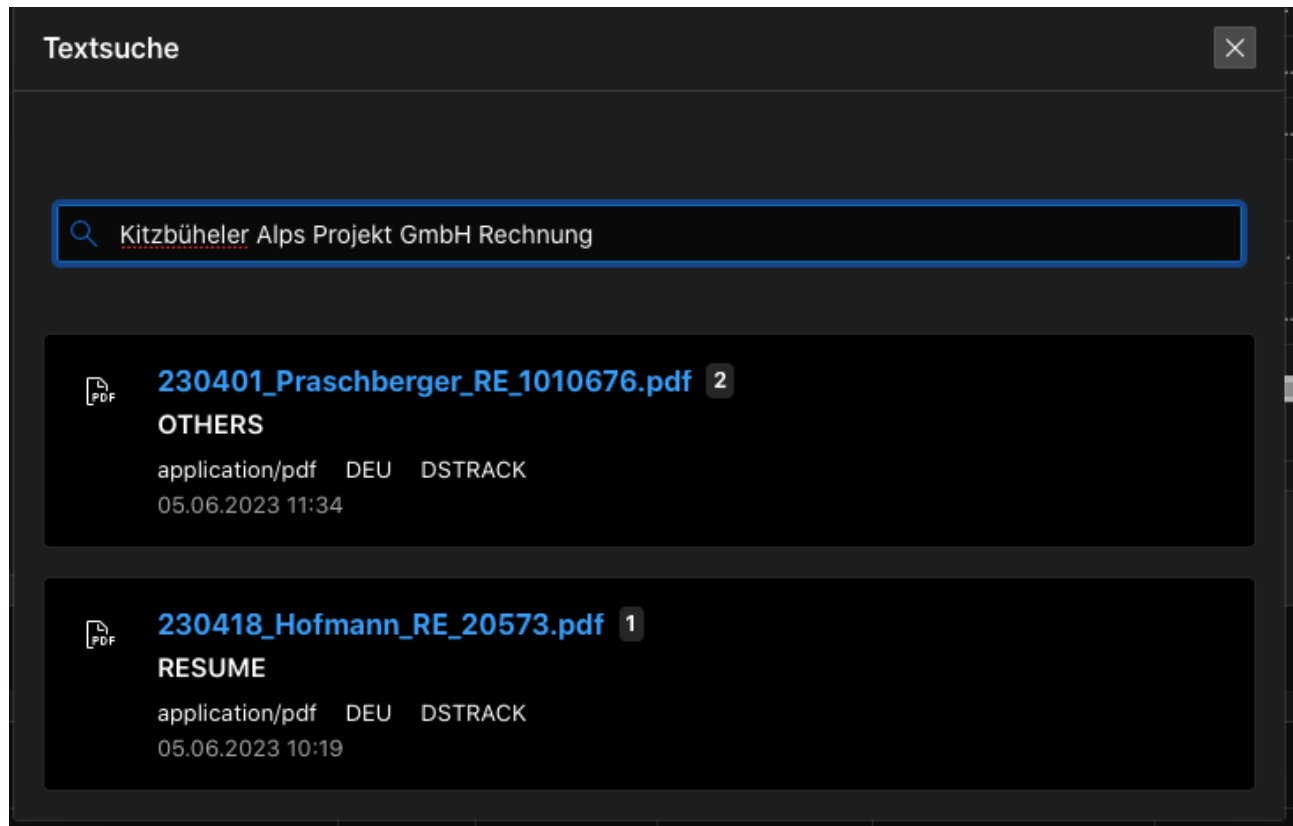
Hier wird der gesamte Text wie er im Text Index verwendete wird ausgegeben.

Text Preview

Octopus Werbegestaltung & Montagebau GmbH Kitzbühler Alps Projekt GmbH ATU74785958 Kontakt: Obere Gnsbachgasse 7 Wasserfeld 1a 6370 Kitzbhel 6361 Hopfgarten im Brixental Mobil: +43 /676/9616336 Mobil: +43 /676 / 5339782 Datum: 09.03.2023 Rechnungs-Nr.: 20230258 Kunden-Nr.: 2265 Sachbearbeiter/-in: Mathias Gandler Rechnung Sehr geehrte Damen und Herren, wir erlauben uns, wie folgt in Rechnung zu stellen: Projekt: Sujetwechsel LED Rahmenit Anfrage vom 24.02.2023 Auftraggeber:Herr Walter Fischer Pos. Anzahl Einheit Artikelnr. Bezeichnung Rabatt Einzelpreis Gesamtpreis 1 1 Std. 099 Datenbernahme, Datenhandling inkl. RIP 50% 15,00 € 15,00 € 2 2,43 m2 134 hochwertiger Sublimations_Digitaldruck 25% 42,30 € 102,79 € auf Frametex B1 220 g Format : It Naturmanahme V2 Konfektion : rundum Gummilippe vernht 1 STK 3 6,3 lfm 212 Silikonflachkeder 10% 4,68 € 29,48 € 4 4,86 m2 134 hochwertiger Sublimations_Digitaldruck 25% 42,30 € 205,58 € auf Frametex B1 220 g Format : It Naturmanahme V2 Konfektion : rundum Gummilippe vernht 1 STK 5 99 lfm 212 Silikonflachkeder 10% 4,68 € 46,33 € 6 1 Stk. 900 Montage Sujetwechsels 30,00 € 30,00 € Montagepauschale incl. An- und Abfahrt zum Montageort Montagedatum: 03.02.2023 Montageort: Sujetwechsel Montageort : Kitzbhel betrag 429,18 € octopus Werbegestaltung & Montagebau GmbH Wasserfeld 1a E-Mail: s.weinert@octopus-wm.at ; m.gandler@octopus-wm.at A-6361 Hopfgarten im Brixental UID-Nr. ATU 63536935 - FN 296280g Rechnungs-Nr.: 20230258 Seite 2 von 2 Pos. Anzahl Einheit Artikelnr. Bezeichnung Rabatt Einzelpreis Gesamtpreis bertrag 429,18 € Summe 429,18 € Mehrwertsteuer 20% auf 429,18 € netto 85,84 € Gesamtbetrag 515,02 € Zahlungsbedingung: Zahlbar nach Erhalt der Rechnung, ohne Abzug Es gilt die Zuständigkeit des Bezirksgerichts Kufstein und die AGB's octopus Werbegestaltung & Montagebau GmbH Die Ware bleibt bis zur vollständigen Bezahlung unser Eigentum. Bankverbindung: Sparkasse Wrgl-Kufstein, BLZ.: 20506, Konto Nr. 00100031731 IBAN: AT022050600100031731, BIC: SPKUAT22XXXX Unsere UID:ATU 63536935 ;FN 296280g EINGEGANGEN 0 6 April 2023 Octopus Werbegestaltung & Montagebau GmbH Bro Franz Wieser Kitzbhel Kitzbühler Alps Projekt GmbH ATU74785958 Kontakt: Obere Gnsbachgasse 7 Wasserfeld 1a 6370 Kitzbhel 6361 Hopfgarten im Brixental Mobil: +43 / 676 / 9616336 Mobil: +43 / 676 / 5339782 Datum: 05.04.2023 Sachbearbeiter/-in: Mathias Gandler Zahlungserinnerung Sehr geehrte Damen und Herren, leider mussten wir feststellen, dass folgende Rechnung Ihrerseits noch nicht beglichen wurde: 20230258 Rechnungs-Nr.: Rechnungsdatum: 09.03.2023 23.03.2023 Flig am: 515,02 € Rechnungsbetrag: Bisher gezahlt: 0,00 € 10,00 € Mahngebühren: 525,02 € Zu zahlender Betrag: Sollten Sie Ihrerseits die Rechnung schon beglichen haben,dann sehen sie dieses Schreiben als gegenstandslos an. Wir mchten Sie bitten, den noch ausstehenden Betrag umgehend zu berweisen. Es gilt die Zuständigkeit des Bezirksgerichts Kufstein und die AGB's octopus Werbegestaltung & Montagebau GmbH. Die Ware bleibt bis zur vollständigen Bezahlung unser Eigentum. Bankverbindung: Sparkasse Wrgl-Kufstein, BLZ.: 20506, Konto Nr. 00100031731 IBAN: AT022050600100031731, BIC: SPKUAT21 octopus Werbegestaltung & Montagebau GmbH Wasserfeld 1a E-Mail: s.weinert@octopus-wm.at ; m.gandler@octopus-wm.at A-6361 Hopfgarten im Brixental UID-Nr. ATU 63536935 - FN 296280g

Dokumente suchen

Mit einem Klick auf das Lupen-Symbol oben rechts in der Menüleiste wird ein Dialogfenster für die Volltextsuche geöffnet. Die im Object-Store abgelegten durchsuchbaren Pdf Dateien sind mit einem Oracle Text Index schnell durchsuchbar. Mit den Eingabe von Stichworten werden entsprechende Dokumente aufgelistet. Mit einen Klick in der Liste der Suchergebnisse kann das Dokument in der Seite Dokumenten Analyse geöffnet werden.



Konfiguration

Mit einem Klick auf den Eintrag Konfiguration oben rechts in der Menüleiste wird das Dialogfenster für die AI Prozessor Konfiguration geöffnet. Ausserdem werden hier die Feldtypen, Einzelpostentypen und Sprachen eingerichtet wie sie für die Prozessor Version in den Herstellerangabe und Ihrer Oracle Cloud Konsole definiert wurden.

AI Prozessor Konfiguration

Hier werden alle notwendigen Parameter für den Zugriff auf den Oracle Object Store und den AI Document Understanding Webservice eingerichtet.

AI-Prozessorkonfiguration bearbeiten

Aktuelle Konfiguration
20221109/processorJobs

Alle anzeigen KI-Prozessorkonfiguration Feldtypen Einzelpostentypen Sprachen Textindexstatus

Konfigurationsname
20221109/processorJobs

Eingabestandortpräfix
scan_ai_documents

Ausgabestandortpräfix
scan_ai_results

PDF-Standardpräfix
searchablePdf

Prozessorangabe löschen

Abteilungs-ID
ocid1.compartment.oc1..aaaaaackpx37s32zo44j4ymjdfpiq4uk6fnvkfl6qj6azezaefqoefda

Objekt-Bucket-Name
APEX_OCI_AI_REST_SERVICE

Objekt-Namespace-Name
frsve7ogr9o

Objektspeicher-Basis-URL
https://objectstorage.eu-frankfurt-1.oraclecloud.com

Doc Doc AI URL
https://document.aiservice.eu-frankfurt-1.oraclecloud.com/20221109/processorJobs

Wc-Anmeldeinformations-ID
APEX_OCI_CREDENTIAL

Währungszeichen
€

Zahlensymbol
., Dezimalpunkt & Gruppenkomma

Prozessorsprachencode
English

Bearbeiterdokumenttyp
Rechnung

Schlüsselwerte-Extraktion

Tabellextraktion

Durchsuchbares PDF generieren

Originaldaten behalten

Asynchron ausführen

Kontext-ID-Abfrage
SELECT NAME, ID FROM WECO_TOWER1.PROJEKTE_BT WHERE WORKSPACE_ID = 124 ORDER BY 1

Adresse-ID-Abfrage
SELECT FIRMATID||' ('||MANDANTEND||')' NAME, ID FROM WECO_TOWER1.ADRESSEN_BT WHERE WORKSPACE_ID = 124 ORDER BY 1

Funktion „Adresse suchen“
Document_Scan_AI_Find_Address

Datum der letzten Aktualisierung
22-Mai-2023 14:08:21

Zuletzt aktualisiert von
DSTRACK

Rechnungsexportname
Weco_Tower1V_DOCUMENT_SCAN_AI_INVOICE_EXPORT

Rechnungsexportverfahren
Weco_Tower1.EXPORT_INVOICE_DOCUMENTS

Abbrechen Kopie speichern Änderungen übernehmen

Feldname	Column-Name	Beispielinhalt	Bedeutung
Konfigurationsname		20221109/processorJobs	Name der Konfiguration
Eingabestandortpräfix	input_location_prefix	scan_ai_documents	Sammelordnername für die hochgeladenen Dateien im Objektspeicher
Ausgabestandortpräfix	output_location_prefix	scan_ai_results	Sammelordnername für die AI Prozessor Ergebnisse im Objektspeicher

Feldname	Column-Name	Beispielinhalt	Bedeutung
PDF-Standortpräfix	pdf_location_prefix	searchable_pdf	Sammelordnername für durchsuchbare Pdf Dokumente im Objektspeicher
Abteilungs-ID	compartment_id	ocid1.compartment.oc1..aaaaaaa	Compartment_Id im Objektspeicher
Objekt-Bucket-Name	object_bucket_name	APEX_OCI_AI_REST_SERVI	Bucketname im Objektspeicher
Objekt-Namespace-Name	object_namespace_name	CE frsvel7xgy9o	Objekt Namespacename im Objektspeicher
Objektspeicher-Basis-URL	object_store_base_url	https://objectstorage.eu-frankfurt-1.oraclecloud.com	Base_Url des Objektspeicher
OCI Doc AI Url	oci_doc_ai_url	https://document.aiservice.eu-frankfurt-1.oci.oraclecloud.com/2021109/processorJobs	Base_Url des Document Understanding AI Prozessor
Wc-Anmeldeinformationen-ID	wc_credential_id	APEX_OCI_CREDENTIAL	Static Name der Web-Credentials im APEX Workspace
Kontext-ID-Abfrage	context_id_query	select name d, id r from projects	Werteliste - Select-Statement für die Zuweisung einen Kontext für die spätere Verarbeitung
Adressensuchfunktion	find_address_function	Document_Scan_AI_Find_Addr	Funktionsname für die Adressensuche
Adressenliste Abfrage	address_id_query	Select name, id from addresses	Werteliste - Select-Statement für die Anzeige von Adressen
Rechnungsexportansicht	invoice_export_view	Weco_Tower1.V_DOCUMENT_SCAN_AI_INVOICE_EXPOR	Name der SQL View für die Exportansicht
Rechnungsexport verfahren	invoice_export_procedure	Weco_Tower1.EXPORT_INVO	Name der SQL Prozedur für den Rechnungsexport zur Buchhaltung
Währungszeichen	currency_character	€	Währungszeichen Standardvorgabe
Zahlenzeichen	number_character	„	Zahlenzeichen Standardvorgabe
Prozessorsprachencode	processor_language_code	English	Dezimal & Gruppen Zeichen Prozessor Sprachen Standardvorgabe. Muss English sein wenn Schlüsselwerte Extraktion eingeschaltet ist.
Bearbeiterdokumententyp	processor_document_type	Rechnung	Dokumenttyp Standardvorgabe
Schlüsselwerte Extraktion	Key_Values_Extraction	Ein	OCI Prozessor Option Key Value Extraction – Erzeugt Feldbeschriftungen und Einzelposten für Rechnungen, Quittungen, und Feldbeschriftungen für Reisepass, Führerschein
Tabellenextraktion	Table_Extraction	Ein	OCI Prozessor Option Table Extraction - Erzeugt Tabellen mit Angaben über die Seite, Anzahl Spalten und Zeilen aus den Dokumentinhalten.
durchsuchbares PDF generieren	generate_Searchable_Pdf	Ein	OCI Prozessor Option Generate Searchable PDF - durchsuchbares PDF aus den Dokumentinhalten generieren. Die Dateien werden im Objekt Store mit dem PDF-Standortpräfix erzeugt. Ein Oracle Text Index macht die Dokumente mit Stichworten durchsuchbar.
Prozessor Ausgabe Löschen	delete_processor_output	Ein	AI Prozessor Ergebnisse im Objektspeicher löschen, nachdem

Feldname	Column-Name	Beispielinhalt	Bedeutung
Originaldateien Behalten	keep_original_files	Aus	die Daten in der SQL Tabelle gespeichert wurden. die hochgeladenen Dateien im Objektspeicher nicht löschen. Wenn durchsuchbares PDF generieren eingeschaltet ist können die Dateien auch dann durchsucht werden wenn die Originaldateien gelöscht wurden.
Asynchron ausführen	execute_asynchronous	Ein	Wenn eingeschaltet wird während der Ausführung von OCI Prozessor Aufträgen wird ein Dialogfenster mit dem Verarbeitungsfortschritt und einer Prognose über die Verarbeitungszeit in Sekunden angezeigt. Außerdem wird dadurch ein Timeout wegen zu langer Verarbeitungszeit vermieden.

Oracle Text Index für die Adressen-Tabelle einrichten

Um die Tabelle mit den Feldinhalten aus dem Analyse-Ergebnis durchsuchbar zu machen, muss ein Oracle Text Index für die Adressen-Tabelle eingerichtet werden. In der Datei Accounting_Address_Text_Index.sql finden sie ein Beispiel dafür.

Funktion zur Adressensuche registrieren

Eine Funktion mit dem Profil

```
FUNCTION Find_Address (p_Search IN VARCHAR2,p_Language IN VARCHAR2) RETURN NUMBER
```

wird in dem Schema der Anwendung installiert und in der Konfiguration (find_address_function) eingetragen. In der Datei Accounting_Find_Address.sql finden sie ein Beispiel dafür.

Abfragen für Kontext und Adressen einrichten

Es werden Select Abfragen für Adressen-Liste (address_id_query) und Projekt-Liste (context_id_query) eingetragen.

Beispiel für address_id_query: SELECT FIRMA1||' ('||MANDANTENID||')' NAME, ID FROM WECO_TOWER1.ADRESSEN_BT WHERE WORKSPACE\$_ID = 124 ORDER BY 1

Beispiel für context_id_query: SELECT NAME, ID FROM WECO_TOWER1.PROJEKTE_BT WHERE WORKSPACE\$_ID = 124 ORDER BY 1

Export zur Buchhaltung einrichten

Im Feld Rechnungsexportansicht wird der Name der SQL View für die Exportansicht und in Rechnungsexportverfahren der Name der SQL Prozedur für den Rechnungsexport zur Buchhaltung eingerichtet. Die Exportansicht zeigt die Datenzusammenstellung die sich aus dem Kontext (Projekt), den Rechnungsdaten und den gefundenen Adressen ergibt. Die Prozedur überträgt für eine Job-Id die Daten aus der Exportansicht als Eingangsrechnungen mit Anlagen in die Buchhaltungsdatenbank. In der Datei Accounting_Invoice_Export.sql finden sie ein Beispiel dafür.

Mehrere Konfigurationen verwalten

Die aktuelle Konfiguration kann mit einem Klick auf ‚Kopie speichern‘ als Duplikat erzeugt werden.

Führen Sie eine Umbenennung durch um indem sie den Inhalt des Feld Bezeichnung anpassen.

In dem Duplikat können eigene Einstellungen für Prozessor Konfiguration, die Feldtypen, Einzelpostentypen und Sprachen vorgenommen werden.

Wenn Sie mehrere Konfiguration haben, können die eine aus der Liste am Anfang der Seite auswählen und als aktuelle Konfiguration markieren.

Konfiguration – Feldtypen

Hier werden alle Feldbezeichnungen der Schlüsselwerte Extraktion für die einzelnen Dokumenttypen definiert. Es werden ihre Bedeutung, der Wertetyp und Formatmasken wie sie für die Prozessor Version in den Herstellerangabe definiert wurden hier eingerichtet. Für die Zuordnung der Feldtypen zu einem Dokument werden der Dokumenttyp des Auftrags und der ermittelte Dokumenttyp je Dokument in Betracht gezogen und gültige Feldtypen die Ausgabe Ergebnisse übernommen.

AI-Prozessorkonfiguration bearbeiten

Alle anzeigen AI-Prozessorkonfiguration bearbeiten **Feldtypen** Werbuchungstypen Sprachen

Feldtypen

Suchen: Alle Textspalten

Los

Aktionen ▾

Bearbeiten

Speichern

Zeile hinzufügen

Zurücksetzen

Dokumenttyp

×

<input type="checkbox"/>	<div>Feldbezeichnung ↑↓</div>	Beschreibung	Api-Antwortwert	Werttyp	Formatmaske
Dokumenttyp: DRIVER_LICENSE					
<input checked="" type="checkbox"/>	Address	Address listed on the document.	Extracted Text	STRING	
<input type="checkbox"/>	BirthDate	Date of birth.	Date in YYYY/MM/DD format	DATE	YYYY/MM/DD
<input type="checkbox"/>	Country	Country listed on the document.	Extracted ISO 3166-1 country code	STRING	
<input type="checkbox"/>	DocumentNumber	Document identification number.	Extracted Text	STRING	
<input type="checkbox"/>	ExpiryDate	Date of expiration listed on the docum...	Date in YYYY/MM/DD format	DATE	YYYY/MM/DD
<input type="checkbox"/>	FirstName	First name (given name) listed on the d...	Extracted Text	STRING	
<input type="checkbox"/>	Gender	Gender listed on the document.	Extracted Text	STRING	
<input type="checkbox"/>	IssueDate	Date of issue listed on the document.	Date in YYYY/MM/DD format	DATE	YYYY/MM/DD
<input type="checkbox"/>	LastName	Last name (family name) listed on the ...	Extracted Text	STRING	
<input type="checkbox"/>	Region	Region listed on the document. For exa...	Extracted Text	STRING	
Dokumenttyp: INVOICE					
<input type="checkbox"/>	AmountDue	Total amount due to the vendor.		NUMBER	
<input type="checkbox"/>	BillingAddress	Explicit billing address for the customer.		STRING	
<input type="checkbox"/>	BillingAddressRecipient	Name referenced with the BillingAddre...		STRING	
<input type="checkbox"/>	CustomerAddress	Mailing address for the Customer.		STRING	
<input type="checkbox"/>	CustomerAddressRecipient	Name referenced with the CustomerAd...		STRING	
<input type="checkbox"/>	Customerid	Customer reference identifier.		STRING	
<input type="checkbox"/>	CustomerName	Name of invoiced customer.		STRING	

Abbrechen **Änderungen übernehmen**

Konfiguration – Einzelpostentypen

Hier werden alle Feldbezeichnungen der Einzelposten Extraktion für die die Dokumenttypen Rechnung und Quittung definiert. Es werden ihre Bedeutung und der Werttyp wie sie für die Prozessor Version in den Herstellerangabe definiert wurden hier eingerichtet.

AI-Prozessorkonfiguration bearbeiten

Alle anzeigen AI-Prozessorkonfiguration bearbeiten Feldtypen Einzelpostentypen Sprachen

Einzelpostentypen

☒

<input type="checkbox"/>	<input type="button" value="Werbebuchung ↑ ±2"/>	Beschreibung	Werttyp
▼ Dokumenttyp: INVOICE			
<input checked="" type="checkbox"/>	Amount	The amount of the line item.	NUMBER
<input type="checkbox"/>	Date	Date corresponding to each line item.	DATE
<input type="checkbox"/>	Description	The text description for the invoice line item, for example, men's ray...	STRING
<input type="checkbox"/>	Items	Concatenation of all other line item values (that is, the entire line of t...	STRING
<input type="checkbox"/>	Name	The name listed for a product or service, for example, t-shirt.	STRING
<input type="checkbox"/>	ProductCode	Product code, product number, or SKU referenced in the line item. F...	STRING
<input type="checkbox"/>	Quantity	The quantity for this invoice line item.	NUMBER
<input type="checkbox"/>	Tax	Tax explicitly referenced on the line item.	NUMBER
<input type="checkbox"/>	Unit	The unit of the line item, for example, kg, lb.	STRING
<input type="checkbox"/>	UnitPrice	The price per item identified on the invoice.	NUMBER
▼ Dokumenttyp: RECEIPT			
<input type="checkbox"/>	ItemName	Name of the item.	STRING
<input type="checkbox"/>	ItemPrice	Unit price of the item.	NUMBER
<input type="checkbox"/>	ItemQuantity	The number of each item purchased.	NUMBER
<input type="checkbox"/>	ItemTotalPrice	The total price of the line item.	NUMBER

1 Zeilen ausgewählt Gesamt 14

Konfiguration – Sprachen

Die Liste der Sprachen enthält NLS Parameter mit den die Texte in der Prozessorausgabe in Oracle Number und Date Datentypen konvertiert werden, wenn die Daten zu den Parametern passen. Je ISO-Code können mehrere Zeilen definiert werden. Um aus den Feldinhalten vom Typ Zahl und Datum die richtigen Daten zu konvertieren wird die Sprache des Auftrags und die ermittelte Sprache je Dokument verwendet um die Konvertierungsparameter zu ermitteln. Wenn z.B. die Sprache des Auftrags Deutsch ist und die ermittelte Sprache eines Dokuments Englisch ist, dann werden die Zeilen mit den Gebieten AMERICA, UNITED KINGDOM und GERMANY für die Konvertierung in Betracht gezogen und gültige Konvertierungen in die Ausgabe Ergebnisse übernommen.

AI-Prozessorkonfiguration bearbeiten

[Alle anzeigen](#)[AI-Prozessorkonfiguration bearbeiten](#)[Feldtypen](#)[Werbuchungstypen](#)[Sprachen](#)

Sprachen

Suchen: Alle Textspalten

Los

Aktionen

Bearbeiten

Speichern

Zeile hinzufügen

Zurücksetzen

<input type="checkbox"/>		Iso-Code	Sprachname	Nls Date Language	Gemeinsames Datum	Gemeinsames Zahler	Gebiet	Nls Numerische Zeicl	Nls-Währung	Nls Iso-Währung
<input checked="" type="checkbox"/>		ENG	English	AMERICAN	MM/DD/YYYY	999,999,999,99	AMERICA	„	\$	USD
<input type="checkbox"/>		ENG	English	ENGLISH	DD-MON-YYYY	999,999,999,99	UNITED KINGDOM	„	£	GBP
<input type="checkbox"/>		FRA	French	FRENCH	DD/MM/YYYY	999 999 999,99	FRANCE	„	€	EUR
<input type="checkbox"/>		DEU	German	GERMAN	DD.MM.YYYY	999.999.999,99	GERMANY	„	€	EUR
<input type="checkbox"/>		ITA	Italian	ITALIAN	DD/MM/YYYY	999.999.999,99	ITALY	„	€	EUR
<input type="checkbox"/>		ESP	Spanish	SPANISH	DD/MM/YYYY	999.999.999,99	SPAIN	„	€	EUR
<input type="checkbox"/>		PRT	Portuguese	PORTUGUESE	DD/MM/YYYY	999.999.999,99	PORTUGAL	„	€	EUR
<input type="checkbox"/>		RUS	Russian	RUSSIAN	DD.MM.YYYY	999 999 999,99	RUSSIA	„	RUB	RUB
<input type="checkbox"/>		POL	Polish	POLISH	YYYY-MM-DD	999 999 999,99	POLAND	„	zł	PLN
<input type="checkbox"/>		NLD	Dutch	DUTCH	DD-MM-YYYY	999.999.999,99	THE NETHERLANDS	„	€	EUR
<input type="checkbox"/>		SWE	Swedish	SWEDISH	YYYY-MM-DD	999 999 999,99	SWEDEN	„	kr	SEK
<input type="checkbox"/>		NOR	Norwegian	NORWEGIAN	DD.MM.YYYY	999 999 999,99	NORWAY	„	kr	NOK
<input type="checkbox"/>		DNK	Danish	DANISH	DD-MM-YYYY	999.999.999,99	DENMARK	„	kr	DKK
<input type="checkbox"/>		FIN	Finnish	FINNISH	DD.MM.YYYY	999 999 999,99	FINLAND	„	€	EUR
<input type="checkbox"/>		GRC	Greek	GREEK	DD/MM/YYYY	999.999.999,99	GREECE	„	€	EUR
<input type="checkbox"/>		TUR	Turkish	TURKISH	DD.MM.YYYY	999.999.999,99	TURKEY	„	TL	TRY
<input type="checkbox"/>		CZE	Czech	CZECH	DD.MM.YYYY	999 999 999,99	CZECH REPUBLIC	„	Kč	CZK
<input type="checkbox"/>		HUN	Hungarian	HUNGARIAN	YYYY.MM.DD	999 999 999,99	HUNGARY	„	Ft	HUF
<input type="checkbox"/>		RON	Romanian	ROMANIAN	DD/MM/YYYY	999.999.999,99	ROMANIA	„	lei	RON
<input type="checkbox"/>		BGR	Bulgarian	BULGARIAN	DD.MM.YYYY	999 999 999,99	BULGARIA	„	лв	BGN

Abbrechen

Änderungen übernehmen

Konfiguration - Text Index Status

Hier wird für alle Dokumente der Status des Text Index aufgelistet. Mit einem Klick auf ‘Rebuild Text Index’ kann der Index neu erstellt werden wenn das nötig ist.

Edit AI Processor Configuration

Current Configuration
20221109/processorJobs

Show All AI Processor ConfigurationField TypesLine Item TypesLanguagesText Index Status

Text Index Status

Q

Go

Actions

Rebuild Text Index

Document Id	Job Id	File Name	Has Text Index	Text Index Err Text
270	120	USCIS_US-Passport-Inside_180x241@2x.jpg	Y	
236	92	230413_Magenta_ReNr 901891390423.pdf	Y	
181	77	2023_05rechnung_5611287364.pdf	Y	
286	127	Booking .pdf	N	
182	77	Deine Rechnung von Apple für iCloud.pdf	Y	
238	92	230416_Salzburg_AG_RE_733715583.pdf	Y	
180	76	2023_04rechnung_5611287364.pdf	Y	
183	77	Deine Rechnung von Apple für Angry Birds.pdf	Y	
277	123	HASPA Kontoauszüge .pdf	N	
239	92	230418_Hofmann_RE_20573.pdf	Y	
195	81	Starbucks.jpeg	Y	
196	81	Walgreens.jpg	Y	
237	92	230414_Oesterreichische_Bundesforste_500228003.pdf	Y	
200	85	Chipotle.jpeg	Y	
284	127	Reifen Discounter .pdf	N	
260	115	Chipotle.jpeg	Y	
245	93	230331_MBLD_RE_5522.pdf	Y	
283	126	Quittung nachos.pdf	N	
266	119	Requirements-to-get-driving-license-for-H4-Visa-holders-No-SSN.jpg	Y	
240	92	230420_WF_RE_CD218B8A_0003.pdf	Y	
285	127	ATU.pdf	N	
213	89	Rechnung_2023_04_30254912000791.pdf	Y	

Cancel

Save a Copy

Apply Changes

Installation

Voraussetzungen

Voraussetzungen sind eine bezahlte Oracle Cloud Umgebung mit einem Admin Account für die Einrichtung der benötigten Services und Berechtigungen.

Für die Installation der Supporting Objects sind folgende Schema Berechtigungen erforderlich:

Als SYS oder ADMIN User:

GRANT EXECUTE ON CTXSYS.CTX_DDL TO <APP_SCHEMA>; – für Oracle Text Index

GRANT EXECUTE ON SYS.DBMS_LOCK TO <APP_SCHEMA>; – Für Wait in der asynchronen Ausführung

GRANT EXECUTE ON DBMS_CLOUD TO <APP_SCHEMA>; – Für Object Store Zugriffe

OCI-Konfiguration

In der OCI Service Konsole müssen Security_Group, IAM User, API keys, Compartments, Policies und Object Storage Bucket konfiguriert werden. In APEX müssen Web Credentials konfiguriert werden.

Beispiel OCI-Konfiguration

Identity / Domains / Default domain /

Groups

Name: APEX_OCI_AI_REST_Security_Group

Description: Security Group for access to AI Services

Users

User Type: IAM User

Capabilities:

API keys: Yes

Auth tokens: Yes

Name: APEX_OCI_AI_REST_Service_Account

Description: Application User for AI REST Services

Email: your.name@example.de

Groups: APEX_OCI_AI_REST_Security_Group

API keys:

Download Private Key, Download Public Key

Configuration file - Copy

[DEFAULT]

user=ocid1.user.oc1..aaaa...
fingerprint=66:40:...
tenancy=ocid1.tenancy.oc1..aaaa...
region=eu-frankfurt-1
key_file=<path to your private keyfile> # TODO

Auth tokens:

Description : Access Buckets - Authentifizierungstoken for
DBMS_CLOUD.CREATE_CREDENTIAL

Generate Token - Copy:

;...-b<FvGe

username => 'APEX_OCI_AI_REST_Service_Account'

password => ';...-b<FvGe'

Identity / Compartments

Name : APEX_OCI_SERVICES

Description: Compartment for Secure APEX Access to OCI Object Storage and AI REST Service

OCID <Copy> : ocid1.tenancy.oc1..aaa...

Identity / Policies

Name: APEX_OCI_AI_REST_Bucket_Policy

Description: Policy to assign APEX Access to OCI Services

Compartment: strackdev (root)/APEX_OCI_SERVICES

Policy Builder - Show manual editor:

Statements:

allow group APEX_OCI_AI_REST_Security_Group to read buckets in compartment APEX_OCI_SERVICES

allow group APEX_OCI_AI_REST_Security_Group to manage objects in compartment
APEX_OCI_SERVICES where all {target.bucket.name='APEX_OCI_AI_REST_SERVICE', any
{request.permission='OBJECT_CREATE', request.permission='OBJECT_INSPECT',
request.permission='OBJECT_READ',request.permission='OBJECT_OVERWRITE'}}}

allow group APEX_OCI_AI_REST_Security_Group to manage object-family in compartment
APEX_OCI_SERVICES

allow group APEX_OCI_AI_REST_Security_Group to manage ai-service-document-family in compartment
APEX_OCI_SERVICES

Object Storage

Compartment: APEX_OCI_SERVICES

Buckets: APEX_OCI_AI_REST_SERVICE

Default Storage Tier: Standard

Visibility: Private

Namespace <Copy>: frmnolrlgf8w

APEX Builder / Workspace Utilities / Web Credentials

Name: APEX OCI Credential

Credential Static ID: APEX_OCI_CREDENTIAL

Authentication Type: Oracle Cloud Infrastructure (OCI)

OCI User ID <API Keys Configuration file user>: ocid1.user.oc1..aaa...

OCI Private Key <API Keys private key file CR removed>: MIIEvQIBAD...

OCI Tenancy ID <API Keys Configuration file tenancy>: ocid1.tenancy.oc1..aaaa...

OCI Public Key Fingerprint <API Keys Configuration file fingerprint>: 5e:7b:...

Document Understanding API

Lookup your URL: <https://docs.oracle.com/en-us/iaas/api/#/en/document-understanding/20221109/>

API Endpoint: <https://document.aiservice.eu-frankfurt-1.oci.oraclecloud.com/20221109>

APEX Anwendung installieren

Importiere die Oracle APEX Installations Datei OCI_Document_Understanding_App.sql

Installiere die Application und die Supporting Objects.

Nach der Installation als APP_SCHEMA User:

```
GRANT SELECT ON V_DOCUMENT_SCAN_AI_INVOICES TO <ACCOUNTING_SCHEMA> WITH GRANT OPTION;
```

```
GRANT SELECT ON V_DOCUMENT_SCAN_AI_LINE_ITEM_INVOICE TO <ACCOUNTING_SCHEMA> WITH GRANT OPTION;
```

```
GRANT SELECT ON V_DOCUMENT_SCAN_AI_DOC_FILES TO <ACCOUNTING_SCHEMA> WITH GRANT OPTION;
```

Nach der Installation als ACCOUNTING_SCHEMA User:

```
GRANT SELECT ON V_DOCUMENT_SCAN_AI_INVOICE_EXPORT TO <APP_SCHEMA>;
```

```
GRANT EXECUTE ON Export_Invoice_Documents TO <APP_SCHEMA>;
```

Verarbeitung von Prozessor Aufträgen

1. Dateien im Objekt Store ablegen

Die im Auftrag hochgeladenen Dateien werden gemäß der Konfiguration Bucket-Name und Eingabestandortpräfix (input_location_prefix) in einem Unterverzeichnis mit Job-Nummer abgelegt. Dabei werden die originalen Dateinamen nach einer Konvertierung von Sonder, Accent

und Leerzeichen weiterverwendet. Die Konvertierung ist notwendig weil der erlaubte Zeichensatz eingeschränkt ist.

2. Prozessorauftrag zusammenstellen und starten

Die Liste der hochgeladenen Dateispeicherorte und die Vorgaben wie Dokumenttyp, Sprache und Erkennungs-Features wie Textextraktion, Tabellenextraktion, Dokumentklassifizierung, Schlüsselwertextraktion werden als Auftrag an den Prozessor gesendet.

3. Abwarten der Verarbeitung

Der die Konfiguration **Asynchron ausführen** aktiviert ist, wird eine Scheduler Job gestartet der den Prozessor-Job Status aktualisiert. Der Prozessor-Job Status wird in der Benutzeroberfläche alle 2 Sekunden abgerufen und ein Dialogfenster informiert während der Ausführung über den Status und die erledigten Prozente des Auftrags. Wenn der Auftrag erledigt ist wird die Verarbeitung fortgesetzt.

4. Aufbereitung der Ergebnisse

Nach dem erfolgreichen Abschluss eines Prozessor-Jobs, werden die erkannte Sprache und Dokumenttyp die Feldbeschriftungen mit Feldinhalte, Einzelposten und Tabellen aus dem Analyse-Ergebnis der Prozessor-Jobs extrahiert. Die Konvertierung von Zahlen und Datumsangaben wird gemäß länderspezifischen Formate durchgeführt weil der AI Prozessor nur amerikanische Formate kennt.

Für werden die länderspezifischen Datumsformate, Währungssymbole, Währungscode und das Dezimal -und Gruppenzeichen der Sprachen des Prozessor Auftrag und der gefundene Sprache verwendet. Außerdem spezielle Datumsformate die für Feldbeschriftungen und Einzelposten in den Prozessorergebnissen verwendet werden.

Bei der Validierung der Konvertierung werden erweiterte Bedienungen überprüft um zu verhindern, das z.B. 51.234,50€ nicht als 51,2345 und 11.04.23 nicht als 04. November 2023 erkannt wird.

Für Zahlen darf das Dezimalzeichen nur einmal und nach den Gruppenzeichen vorkommen und das Gruppenzeichen muss mindestens 3 Zeichen vor dem Ende vorkommen und der Abstand zum Dezimalzeichen muss ein Vielfaches von 3 sein. Passende Währungssymbole, Währungscode und das Gruppenzeichen können optional enthalten sein. Deshalb wird für die Zahlenkonvertierung eine dynamische Formatmaske verwendet.

Der Pfad zu hochgeladenen Dateien oder durchsuchbare Pdf Dateien wird in der Datenbank so abgespeichert das ein Oracle Text Index für die Dateien erstellt werden kann. Weil in den erkannten Texten die Akzentzeichen und Umlaute fehlen muss das durchsuchen der Dokumente als unscharfe Suche durchgeführt werden.

Die Feldinhalte werden verdichtet, konvertiert und zusammengefasst und dann in Datenbank-Tabellen gespeichert. Diese stehen danach als Listen zum Download zur Verfügung.

5. Algorithmus zur Analyse von Textfragmenten

Die Document Understanding AI liefert als Ergebnis eines Document Scans in einer JSON Datei neben den Key-Value Paaren (für Formular-Felder mit Inhalten) auch Textfragmente mit X/Y-Koordinaten für Textzeilen und Tabellen-Zellen.

Die Key-Value Ergebnisse der Document AI sind nicht ausreichend für den Import in eine Buchhaltung. Deshalb wurde ein Algorithmus entwickelt der die Key-Value Ergebnisse verbessert.

Der Algorithmus verwendet Alias-Namen, Datentypen Validierung und relative Positionen um gültige Key-Value-Paare zu bilden.

1. Problem: Die im Analyse-Ergebnis enthaltenen Key-Value-Paare für Rechnungen verfügen über zu wenige Felder, um eine eindeutige Zuordnung zu Mandanten/Rechnungsstellern herzustellen. Für die Erkennung des Rechnungsstellers stehen lediglich die Felder VendorName, VendorAddress und VendorTaxId zur Verfügung. Eine Volltextsuche mit Fuzzy Search liefert keine brauchbaren Adresszuordnungen zu einer bestehenden Adressliste.

1.1. Erweiterung der Feldliste

Es wurden zusätzliche Felder wie **VendorEmail**, **VendorPhone**, **VendorMobile**, **BankBIC** und **BankIBAN** für die Adresszuordnung hinzugefügt.

1.2. Ergänzung der Adressen

Die Adressen der Projektbeteiligten wurden so ergänzt, dass Anschrift, UST-ID Nummer und IBAN sowie SWIFT-BIC in einer Bankverbindung vorhanden sind.

1.3. Die Suche nach den Adressen der Rechnungssteller umfasst nun die zusätzlichen Felder, wodurch eine eindeutige und zuverlässige Zuordnung ermöglicht wird.

2. Problem: Die gefundenen Key-Value-Paare für Rechnungen sind unvollständig oder falsch zugeordnet.

2.1. Filterung anhand von Kontextwerten

Wenn für den Vorgang ein Kontext (Projekt) festgelegt wurde, können die Feldinhalte wie VendorName, VendorEmail, VendorPhone, VendorTaxId und BankIBAN mit den Werten des Kontexts (Client_Name, Client_Email, Client_Phone, Client_Tax_Id, Client_IBAN) verglichen werden. Wenn Übereinstimmungen gefunden werden, können Zuordnungen zu den Vendor-Feldern als falsch erkannt und ausgeschlossen werden.

2.2. Verbesserte Suche nach Feldbezeichnungen

Eine neue Tabelle DOCUMENT_SCAN_AI_FIELD_ALIAS mit Alias-Namen für Feldbezeichnungen der Key-Value-Paare ermöglicht eine präzisere und umfassendere

Feldzuordnung. Für deutsche und englische Rechnungen wurden spezifische Alias-Namen eingerichtet.

Für deutsche Rechnungen wurden folgende Feld Alias-Namen eingerichtet:

AmountDue: Zahlungsbetrag, Zu zahlender Betrag, Zu zahlender Betrag in Euro, berweisungsbetrag

BankBIC: BIC, BIC (SWIFT-Code) der Empfängerbank, BIC(SWIFT-Code) der Empfängerbank, BIC(SWIFT-Code) der Empfängerbank, BIC-Code, BIC., SWIFT, SWIFT Code, SWIFT/BIC, Swift, Swift Code

BankIBAN: IBAN, IBAN (EUR), IBAN Code, IBAN Nummer, IBAN code, IBANEmpfängerin

CustomerId: Ihre Kundennummer, Kunden-Nr, Kunden-Nr., Kundennr, Kundennr., Kundennummer

DueDate: Fällig am, Fälligkeit, Fälligkeitsdatum, Zahlungstermin

InvoiceDate: Datum, Leistungsdatum, Rechnungsdatum

InvoiceId: Belegnummer, Re.- Nr., Rechn.-Nr., Rechnung, Rechnung Nr., Rechnungs-Nr., Rechnungsnr., Rechnungsnummer, Zahlungsreferenz

InvoicePaidDate: BEZAHLT

InvoiceReceiptDate: EINGANG, EINGEGANGEN, Eingang

InvoiceTotal: Endbetrag EUR (brutto), Gesamt, Gesamtbetrag, Gesamtbetrag (inkl. USt), Gesamtbetrag brutto, Gesamtsumme, Rechnungsbetrag, Rechnungsbetrag in EUR, Summe, Summe Brutto (EUR)

PaymentTerm: Zahlungsbedingung, Zahlungsbedingungen, Zahlungsziel

SubTotal: Betrag Netto, Betrag netto, Netto, Netto Gesamtbetrag, Nettobetrag, Nettosumme, Rechnungsbetrag netto

TotalTax: Summe USt (EUR)

TotalVAT: 20 % MWST, 20 % USt, Betrag Ust. 20%, MwSt, MwSt.

VendorEmail: E-MAIL, E-Mail, E-Mail-Adresse, E-mail, Email

VendorMobil: Mobil

VendorName: Abs, Absender, Account Owner, Empfängerin Name/firma, EmpfängerIn, EmpfängerIn Name/Firma, Name des Empfängers

VendorPhone: TEL., Tel., Telefon

VendorTaxId: UID, UID Nr, UID Nr., UID-Nr., UID-Nummer, UST-ID NR., USt.-ID, Umsatzsteuer-ID, Unsere UID, Unsere UID- Nummer, Unsere UID-Nummer

Für englische Rechnungen wurden folgende Feld Alias-Namen eingerichtet:

AmountDue: Amount due, Amount payable, Total Amount due, Total amount due

BankBIC: BIC, Swift Code

BankIBAN: IBAN, IBAN code

CustomerId: Customer number

DueDate: Date due, Due date, Payment Due

InvoiceDate: Date of issue, Invoicing date

InvoiceId: Invoice Number, Invoice number

InvoiceTotal: TOTAL, Total, Total amount

SubTotal: Subtotal
TotalVAT: 20% VAT
VendorEmail: E-mail, Email
VendorName: Account Name
VendorPhone: Phone
VendorTaxId: Our VAT ID, Tax number

In den Ergebnissen des AI-Scans sind Textfragmente mit genauen Koordinaten für erkannte Texte vorhanden. Durch das Finden von Alias-Namen und die Suche nach Textfragmenten, die sich relativ rechts oder unterhalb befinden, werden Key-Value-Paare gebildet.

Mit einer Mustererkennung werden Alias-Namen und Feldwerte ermittelt:

- Wenn nach einem Alias-Name ein Wert in einem Textfragment folgt, wird der Feldwert rechts vom Doppelpunkt extrahiert.
- Zusammengesetzte Textfragmente die mehrere Alias-Namen und Werte enthalten werden bei Semikolon, senkrechtem Strich, Punkt oder Komma aufgetrennt und die Namen und Werte extrahiert.
- Wenn ein Alias-Name in einem Textfragment und in einem zweiten Textfragment rechts neben dem Alias-Namen in derselben Zeile erscheint, wird der letzte Wert (bei Zahlen) oder der nächste Wert (sonst) verwendet.
- Wenn ein Alias-Name in einem Textfragment und in einem zweiten Textfragment unter dem ersten in der selben Spalte, links- oder rechtsbündig, erscheint, erfolgt die Extraktion.
- Die potenziellen Feldinhalte werden auf ihre Übereinstimmung mit den Datentypen (EMAIL, IBAN, SWIFTBIC, PHONE, INITCAP, NUMERIC, ALPHANUM) überprüft, und unpassende Daten werden herausgefiltert.

3. Problem: Die gefundenen Key-Value-Paare für NUMBER und DATE werden nicht richtig gemäß der Sprache und Region konvertiert.

3.1 Anhand der voreingestellten Sprache für den Scan-Job und der von der Document AI ermittelten Sprache werden gültige Formate für Datums- und Zahlenwerte ermittelt und validiert. Bei mehreren gültigen Treffern für einen Feldtyp wird der größte Wert bevorzugt.

4. Ergänzung der Key-Value-Paare

Der Algorithmus wurde als Oracle View implementiert. Diese View liefert je Importjob und Dokument mögliche Treffer für die Suche nach Key-Value-Paaren mit gültigen Aliasnamen, Datentypen und Validierungen für Beträge und Datumsangaben. Die besten Treffer pro Feld werden aus den möglichen Treffern ausgewählt und der Key-Value-Liste hinzugefügt. Basierend auf dieser Key-Value-Liste werden die Datenzeilen für den Export in die Buchhaltung erstellt.

5. Fortlaufende Verbesserung der Ergebnisse

5.1 Vor dem Export der Rechnungsdatenzeilen kann eine Liste von neuen Aliasnamen geprüft werden. Das macht insbesondere dann Sinn, wenn die Rechnungs-Scans von einem bestimmten Rechnungssteller das erste Mal importiert werden und die angezeigten Rechnungsdaten unvollständig erscheinen. Es werden automatisch neue Feldbezeichnungen gefunden, die mit einem Doppelpunkt enden. Außerdem werden mögliche Feldbezeichnungen aufgelistet die vor oder über einem gültigen Datums oder Zahlenwert stehen.

Klicken Sie auf der Seite 'Details zum Prozessorauftrag' auf den Button 'Schlüsselalias zuweisen'. Wenn in der eingeblendeten Liste Bezeichnungen gefunden werden, die sich einem bekannten Feldtyp zuordnen lassen, dann kann diese Zuordnung mit einem Klick auf den Button 'Aliase zuweisen' für ausgewählte Zeilen und den ausgewählten Feldtyp durchgeführt werden. Uninteressante Bezeichnungen können ausgeblendet werden, indem Sie den Feldtyp *Ignorieren wählen.

5.2 Nach der Durchführung eines Import-Jobs mit PDF-Dokumenten wird in der Liste Rechnungen in der Spalte 'Adresse des verknüpften Anbieters' das zugeordnete Personenkonto angezeigt. Prüfen Sie die Zuordnungen. Beachten Sie, dass die Zuordnung eines Personenkontos in der Datensammlung am besten gelingt, wenn in den Adressdaten der Personenkonten auch E-Mail, Telefon, UST-ID Nummer, IBAN und Swift-BIC eingetragen wurden.

Wenn Sie die Adressdaten oder Aliasnamen-Zuordnung ergänzt haben und eine Neuordnung durchführen wollen, klicken Sie auf 'Adressen neu verknüpfen' auf der Seite 'Details zum Prozessorauftrag'. Nach der Neuordnung werden die Datenzeilen für die Liste der Rechnungen aktualisiert, um die neuen Felder einzubeziehen.

5. Bereinigung der Prozessor-Job Ergebnisse.

Die Ausgabedateien eines Prozessor-Job werden gelöscht nachdem die Inhalte in der Datenbank gespeichert wurden.

Die vom Prozessor-Job erzeugten durchsuchbare Pdf Dateien werden in dem dafür konfigurieren Verzeichnis (pdf_location_prefix) in einem Unterverzeichnis mit Job-Nummer abgelegt..

Die im Objekt Store abgelegten Originaldateien werden gelöscht, wenn das so konfiguriert wurde und eine durchsuchbare Pdf Dateien vom Prozessor-Job erzeugt wurde.

6. Verbindungen zur Adressen-Tabelle herstellen.

Die Feldinhalte für erkannte Rechnungen und Quittungen werden verwendet um automatisch passende Einträge in einer Adressenliste zu finden und den Dokumentattributen zuzuordnen. Mit der Prozedur Document_Scan_Ai_Pkg.Link_Document_Addresses wird die Suche nach passende Einträge in einer Adressenliste durchgeführt und in den Dokumentattributen CustomerAddress_ID, VendorAddress_ID, BillingAddress_ID, ShippingAddress_ID, ServiceAddress_ID und RemittanceAddress_ID eingetragen. Weil in den erkannten Texten die Akzentzeichen und Umlaute fehlen muss die Suche mit den Feldinhalten nach passenden Adressen unscharf (FUZZY) durchgeführt werden.