



DATENPROJEKTE
ERGON

Automatisierte Dokumentenverarbeitung mit KI im Mittelstand

18.2.2025 - Marc Wilken - ARIC Brown Bag Session

Unsere Expertise

BI, Analytics und Reporting



Wir entwickeln aussagekräftige, zuverlässige Datenplattformen



Holen Sie das Beste aus Ihren Daten heraus mit modernen Analysetools und KI

Beispiele:

- Interfaces
- Cubes
- Data Warehouses
- ETL date routes
- Dashboards
- Reports
- Data Lakes



Azure Data
Factory



Azure Synapse
Analytics



Databricks



PowerBI



SQL Server
(SSIS, SSAS, SSRS)



Microsoft
Fabric

Unsere Expertise

Anwendungen



Digitalisierte Prozesse zur Steigerung der Effektivität und Verbesserung der Zuverlässigkeit



Moderne, benutzerfreundliche und skalierbare Software durch eigene oder gemischte Teams

Beispiele:

- Cloud-native Apps
- Cloud-Migration
- Apps/Webentwicklung und –Services
- DevOps



Azure Cloud



C#



Kubernetes



Docker



Vue.js



Svelte



ASP.net



SonarQube



Azure DevOps

Der Kunde

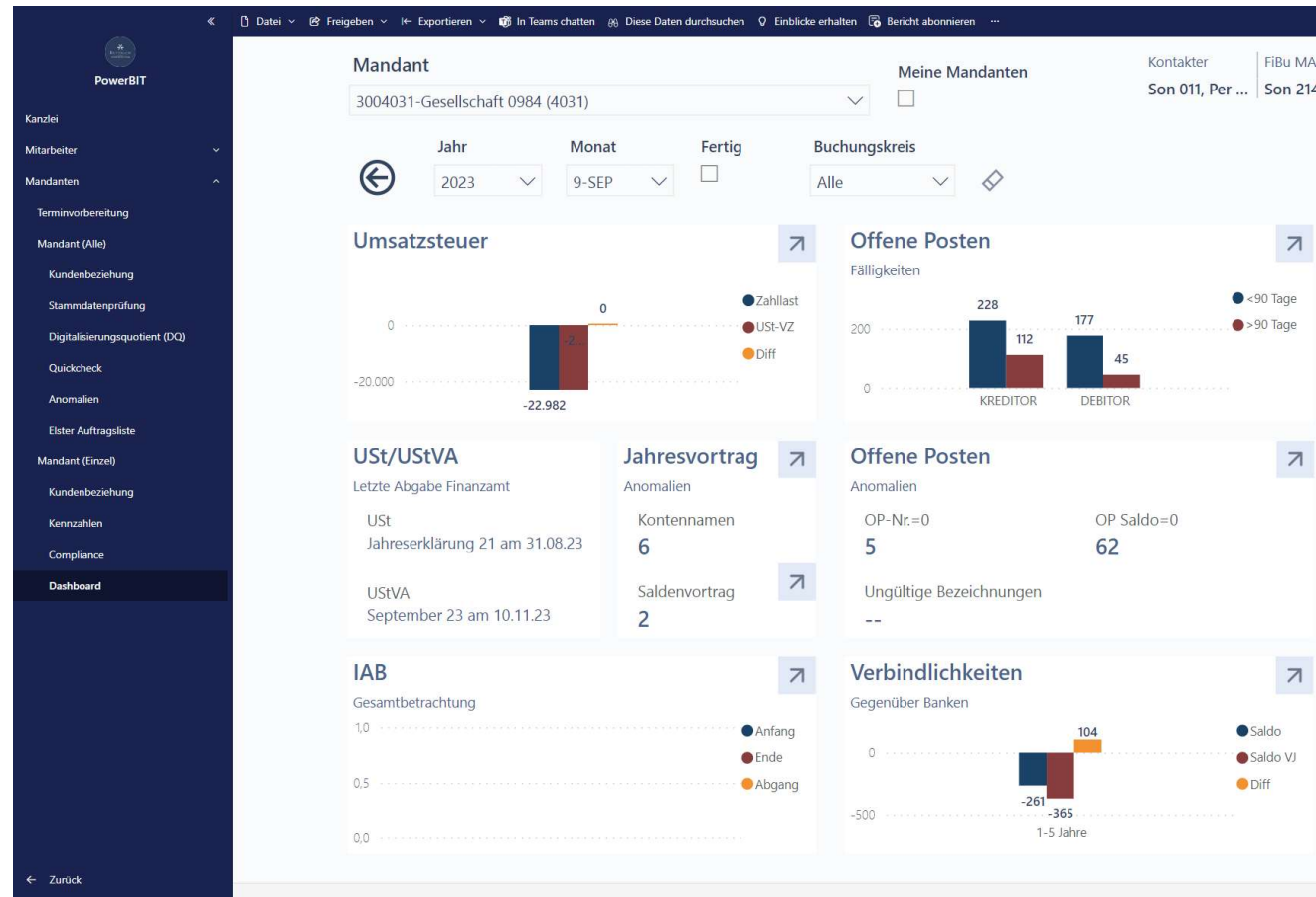


Die Einführung eines Reporting-Systems gemeinsam mit der Ergon Datenprojekte hat dazu geführt, dass wir auf all unsere Daten von überall sicher zugreifen können. Ein kurzer Blick vor einem Mandantentermin auf unser Reporting Tool zeigt uns z.B. alle Informationen und Entwicklungen zu dem jeweiligen Mandanten.



Holger Bittrich

Geschäftsführender Gesellschafter der
Bittrich & Bittrich Steuerberatungsgesellschaft mbH



Problemstellung: Steuerbescheidprüfung

Kanzlei erfasst Finanzdaten von Mandanten

... und berechnet Steuern

.. und erstellt Steuererklärung

Finanzamt prüft Erklärung

.. und versendet Bescheid

Kanzlei scannt Bescheid

.. und gleicht Bescheidaten mit Berechnung ab

.. und informiert Mandant

The image displays several overlapping German tax forms (Steuerbescheide) from the Finanzamt Dresden and Finanzamt Lüneburg. The forms contain detailed tax calculations, including income tax (Einkommensteuer), trade tax (Gewerbesteuer), and solidarity surcharge (Solidaritätszuschlag). They also feature official stamps, including the 'EINGANG' (received) stamp from the Finanzamt Dresden dated 21.08.2024. The forms are filled out with specific data for various taxpayers, including names, addresses, and tax amounts.

ca. 12.000 Bescheide / Jahr

Lösungsansatz: Automatische Erkennung Bescheiddaten

Document Intelligence Service (DIS)

- Erkennung von Daten über Training eines neuronalen Netzes
- Neuronales Netz ist vortrainiert, erkennt zum Beispiel Tabellen in einem Dokument
- Zum Training müssen Dokumente (mind. 5) gelabelt werden

The image shows a German tax assessment document (Bescheid) from the Finanzamt Lüneburg. The document is for the tax year 2021 and relates to the taxpayer's income tax (Einkommensteuer) and solidarity surcharge (Solidaritätszuschlag). The document includes a barcode, a table of tax amounts, and a section for prepayments.

Structured Data Extraction:

Field	Value
TaxAmount	880,00
TaxYear	2021
SendAt	21.10.2022
TaxableIncome	24.786
TaxAmountPrePayment	664
PrePaymentPeriods	1
ReceivedAt	25. Okt. 2022
TaxNumber	33/118/00100
PaymentType	zahlen
PaymentDueAt	14.11.2022
PaymentAmount	880,00

GPT4 Vision (GPT4V)

- Generatives Sprachmodell, das aus Bildern Informationen auslesen kann
- Erfordert kein spezielles Training
- Daten, die ausgelesen werden sollen, können mit natürlicher Sprache beschrieben werden

The image shows a GPT4 Vision interface. On the right, there is a thumbnail of a German tax document (Bescheid) from the Finanzamt Lüneburg. Below the thumbnail, a query is entered: "Wie lautet die Steuernummer?". The interface also shows a response box with the text: "Die Steuernummer lautet 24/346/89534".

DIS oder GPT4V ?

Unsere Empfehlung für den Kunden:

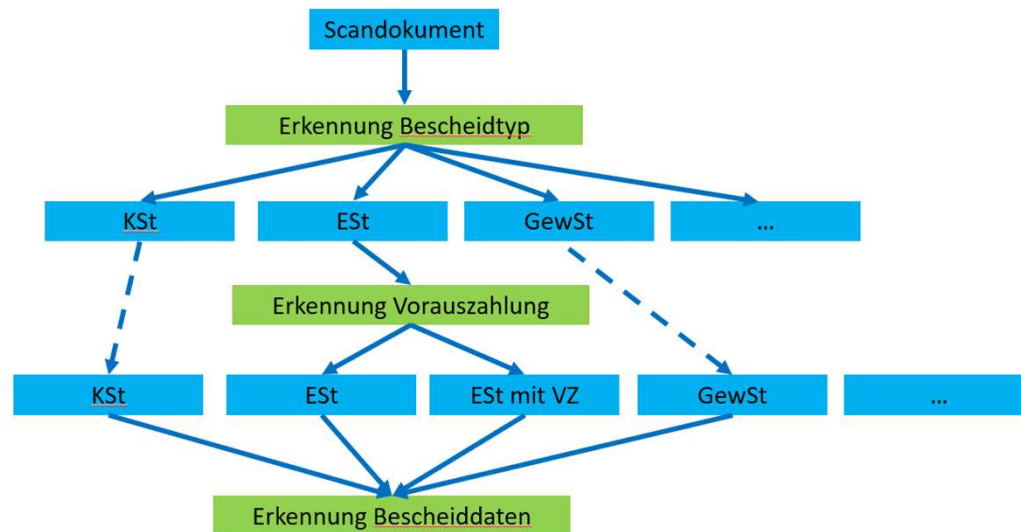
Neuronales Netz für Dokumentenerkennung (Document Intelligence Service)

Warum?

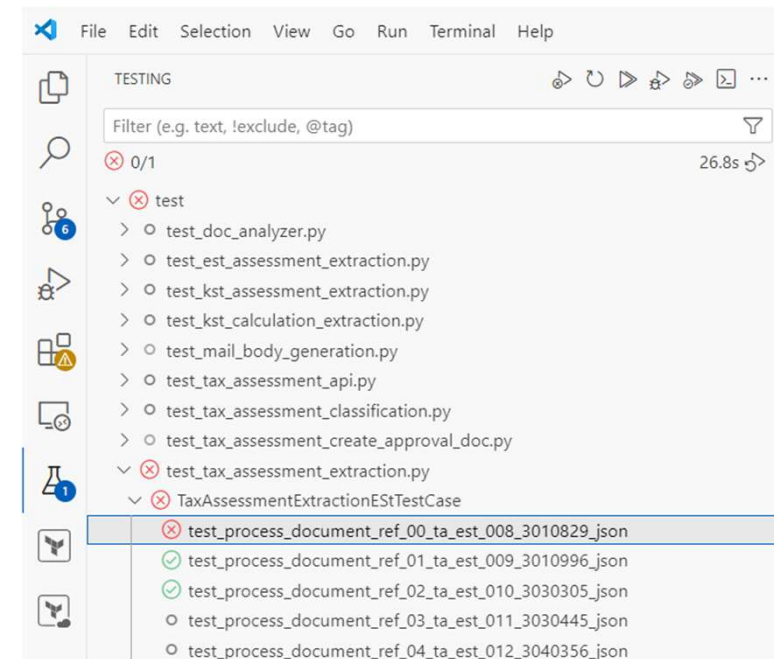
- Im Bescheid sind sämtliche relevanten Informationen enthalten, es wird keine inhaltliche Information außerhalb dieses „Informationsraumes“ benötigt (z.B. aus anderen Datenquellen)
- Erkennungsmodelle können explizit und einfach auf spezielle Anwendungsfälle/Bescheide trainiert werden
- Nachvollziehbarkeit: Ergebnisse der Erkennung können erklärt werden (und daraus Schlüsse zur Verbesserung abgeleitet werden)

Umsetzung Erkennung

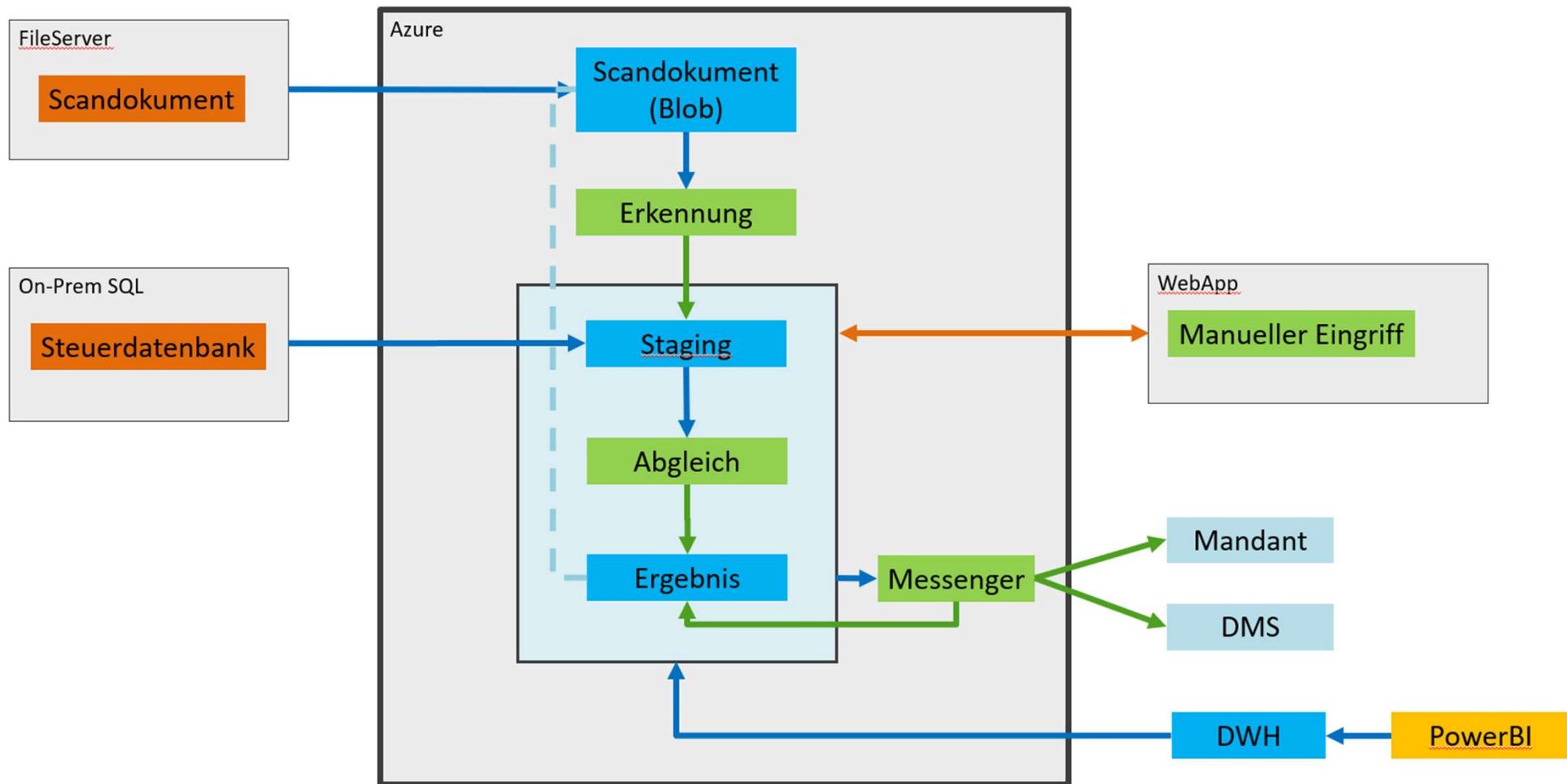
Modelldifferenzierung



Regressionstests



Gesamtsystemaufbau




Backend

The screenshot displays the Microsoft Azure Data Factory portal interface. The top navigation bar shows 'Microsoft Azure | Data Factory | adf-ta-bittrich'. The left sidebar contains 'Factory Resources' with a search bar and a tree view showing 'Pipelines' (11 items) and 'Datasets' (29 items). The main area shows the 'Check Tax Assessment' pipeline with a sequence of activities: 'Update Entry' (Stored procedure), 'Load Tax Assessment Data' (Execute Pipeline), 'Create Result Documents' (Execute Pipeline), and 'Create and Send Messages' (Execute Pipeline). The pipeline is in a 'Validated' state. Below the pipeline, the 'Output' tab shows the 'Pipeline run ID: 616fa3e1-54b7-45d2-9128-423919f599d7' and a table of activity results.

Activity name	Activity status	Activity t...	Run start
Create and Send Messages	✓ Succeeded	Execute Pipeline	1/24/2025, 11:17
Create Result Documents	✓ Succeeded	Execute Pipeline	1/24/2025, 11:15

The screenshot shows the 'Load Tax Assessment Data' pipeline in the Microsoft Azure Data Factory portal. The pipeline is in a 'Validated' state and contains a 'ForEachFile' trigger. The activities are: 'Classify Document' (Web), 'Extract Tax Assessment Data' (Web), and 'Import Tax Assessment Data' (Copy data). The pipeline is in a 'Validated' state.

Frontend



BITTRICH
STEUERBERATUNG

Steuerbescheidprüfungen

Mandantennummer
Suche erfolgt ab 4 Ziffern, z.B. 1234

☐ Nicht erledigte ☒ Alle

Mandantennummer	Jahr	Periode	FKB Nummer	Fristende	Steuerart	Bescheidtyp
3030989	2024	Ganzjährig	00000095	17.01.2025	Gewerbesteuermessbetrag	Festsetzung
3030989	2023	Ganzjährig	00000096	17.01.2025	Gewerbesteuermessbetrag	Festsetzung
3030989	2022	Ganzjährig	00000098	17.01.2025	Gewerbesteuer	Besteuerungsgrundlage
3030989	2023	Ganzjährig	00000099	17.01.2025	Gewerbesteuer	Besteuerungsgrundlage
		Ganzjährig	00000102	10.01.2025	Unbekannt	Unbekannt
		Ganzjährig	00000103	30.12.2024	Unbekannt	Festsetzung
3011008	2023	Ganzjährig	00000108	03.01.2025	Gewerbesteuer	Verlustvortrag
3020358	2023	Ganzjährig	00000110	31.01.2025	Körperschaftsteuer	Feststellung
3020358	2023	Ganzjährig	00000111	31.01.2025	Körperschaftsteuer	Unbekannt
	2022	Ganzjährig	00000112	31.01.2025	Gewerbesteuermessbetrag	Unbekannt


BITTRICH
STEUERBERATUNG

FKB Eintrag bearbeiten

Mandantennummer
Suche erfolgt ab 4 Ziffern, z.B. 1234

Veranlagungsjahr

Bescheiddatum

Fristdatum

Eingangsdatum

Steuerart

Bescheidtyp

Festgesetzte Steuer

Berechnete Steuer

Festgesetzter Zahlbetrag

Berechneter Zahlbetrag

Tatsächlicher Zahlbetrag

Zahlungsart

Kommentar

Erledigungsart

Finanzamt Lüneburg
Steuernummer: 33/
(Bitte bei Rückfr

Finanzamt, 21332 Lüneburg

Bittrich & Bittrich
Steuerberatungsgesellschaft
Köllestr. 42-44
21335 Lüneburg

Dieser Bescheid ergiebt sich aus dem

Festsetzung und Abrechnung
Art der Festsetzung
Der Bescheid ist nach
Festsetzung

Festgesetzt werden
Abrechnung
(Stichtag: 15.11.2024)
Abzurechnen sind
Bereits getilgt
Noch zu zahlen

Bitte zahlen
spätestens am 30.12.2024

Besteuerungsgrundlage
Berechnung des zu veranlagenden Einkommens
Einkünfte aus Gewerbebetrieb
Steuerlicher Jahresüberschuss
Investitionsabzugsbetrag
Aufwendungen nach § 9
Körperschaftsteuer
Solidaritätszuschlag
Gewerbesteuer (Gr. 0)
Sonstige nichtabziehbare
Summe der Einkünfte
Gesamtbetrag der Einkünfte

Konten des Finanzamts:

Zusammenfassung

Umsetzung

- Anwendung KI mit überschaubarem Aufwand möglich
- Erklärbarkeit der Erkennung schafft Vertrauen
- Regressionstests sind (insbesondere hier) wichtig
- Möglichkeit des manuellen Eingreifens und Überwachung essenziell
- (Prozessabbildung war der arbeitsreiche Teil)

Mehrwert

- Zeitersparnis : Prüfprozess -50%, Verwaltungsprozess: -80% (geschätzt)
- Detaillierte Auswertungen über Steuerbescheiddaten möglich

Wenn ausstellende Behörden und empfangende Systeme Daten direkt digital verarbeiten würden, bräuchte man an dieser Stelle keine KI :-)

Vielen Dank!

Noch Fragen?