

Legal Revolution 2019

Complex Document Processing Services – effiziente, KI gestützte Verfahrensvorbereitung



Arvin Arora
(Geschäftsführer, Data Scientist)

aarora@inserve.de

www.inserve.de

Unternehmen / Team



Unternehmen:

- Gründung Ende 2015
- Enge Zusammenarbeit mit Uni Osnabrück
- Enge Zusammenarbeit mit versch. Kanzleien
- Schwesterfirma AIM: ML & DevOps
- Private Investoren

Team:

- 15 Mitarbeiter in 2019
- Wachstum auf 25 Mitarbeiter in 2020

**Arvin
Arora**



**Benjamin
Lobmüller**



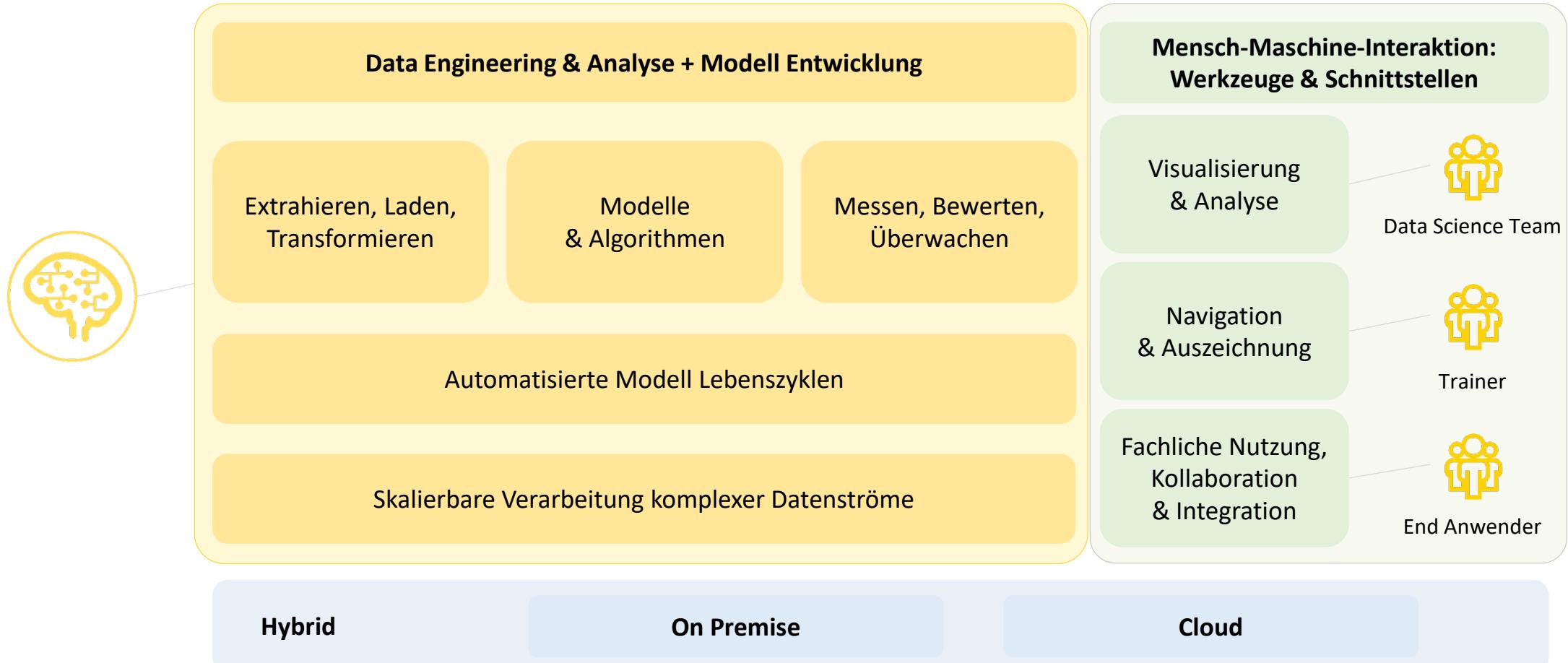
**Eric
Schloen**



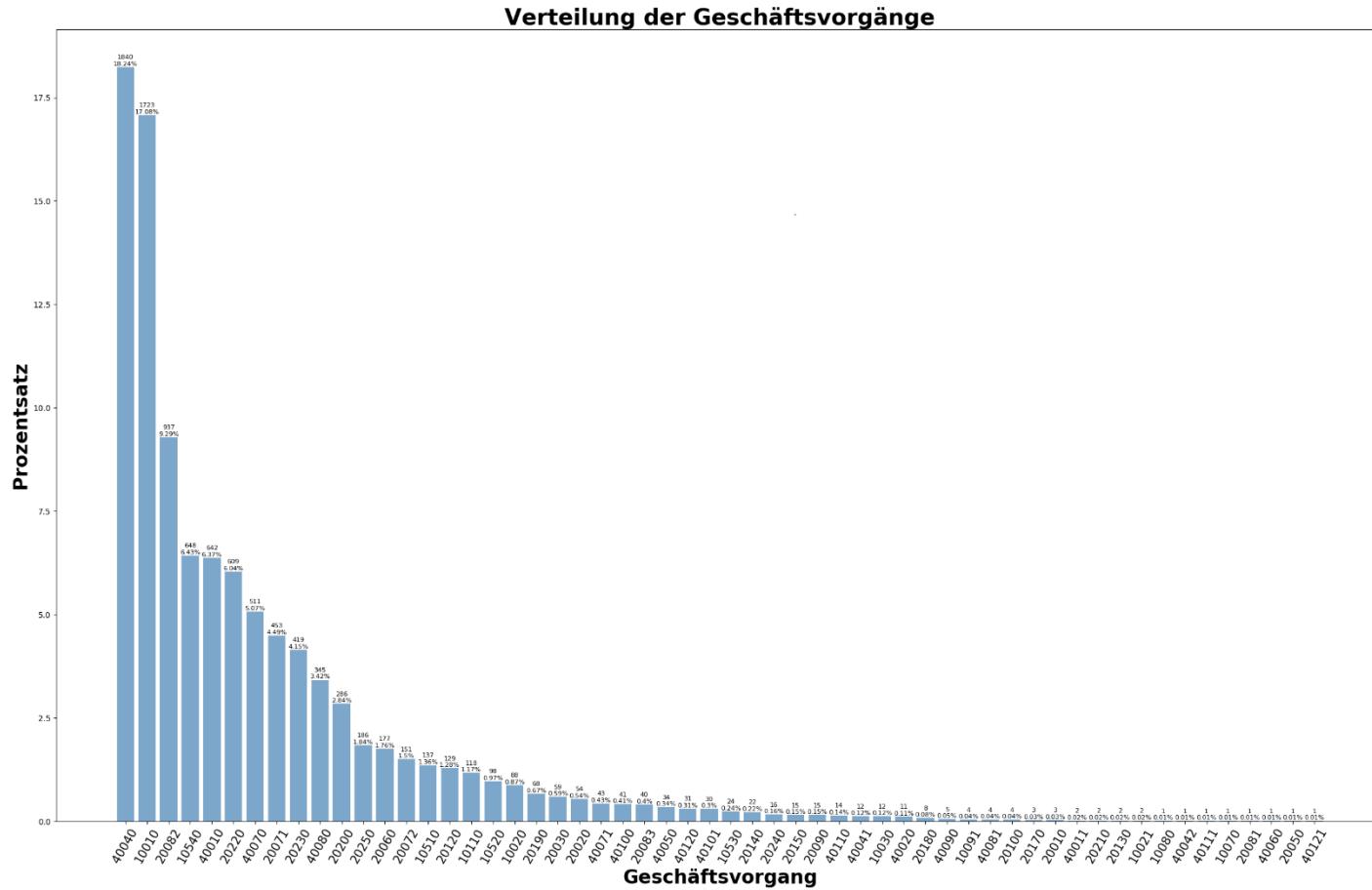
**Benjamin
Peters**



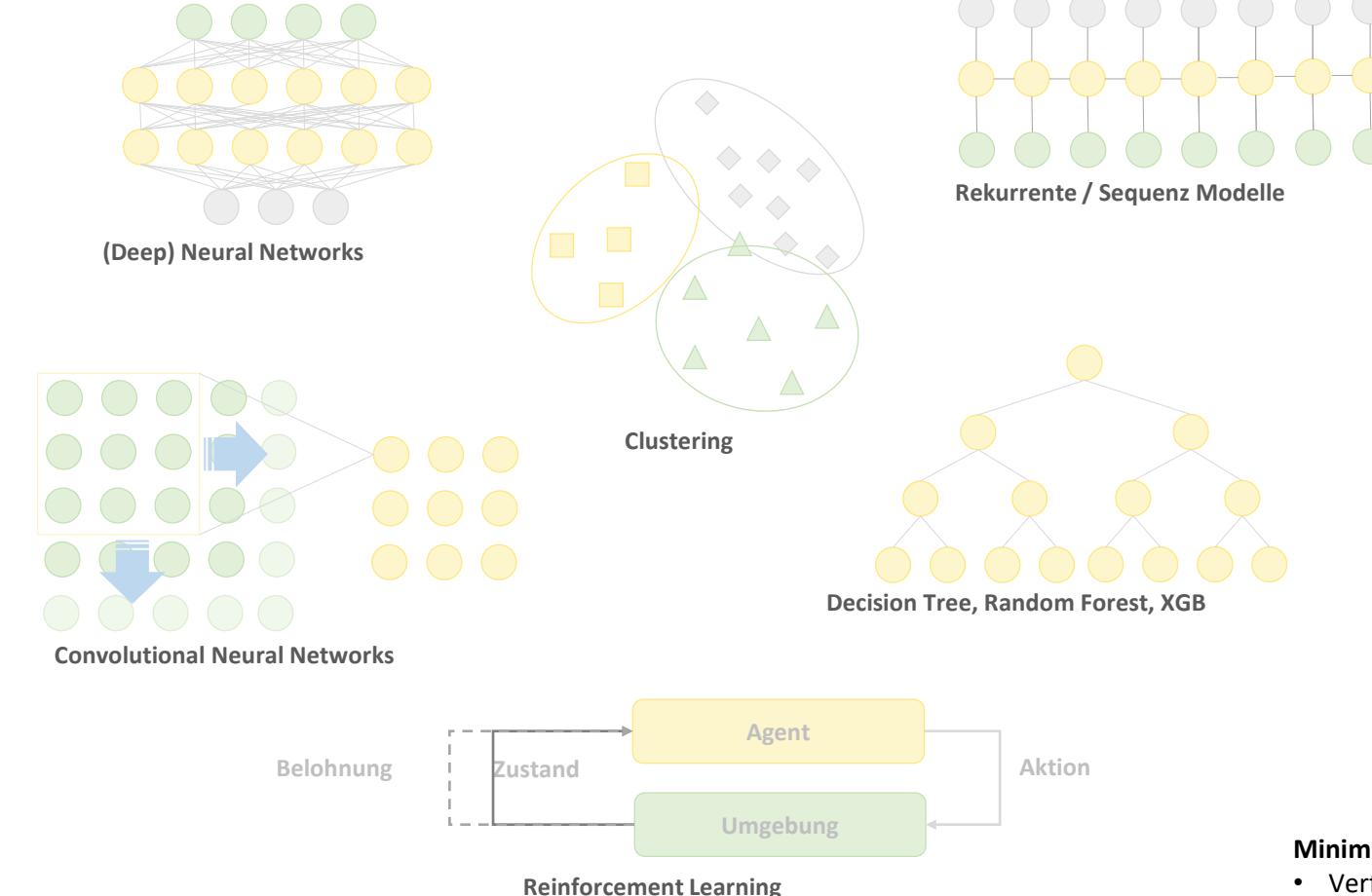
KI Erfolgsbausteine



KI Erfolgsbausteine



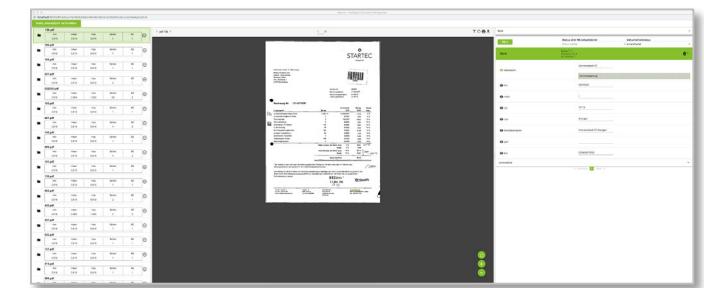
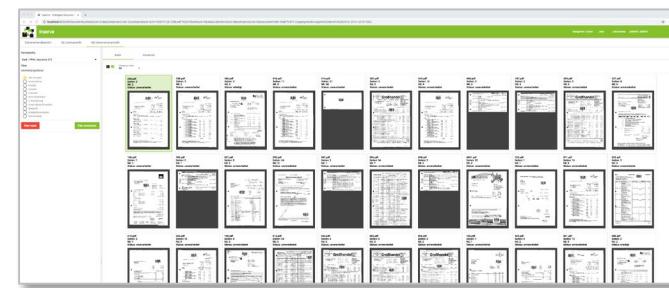
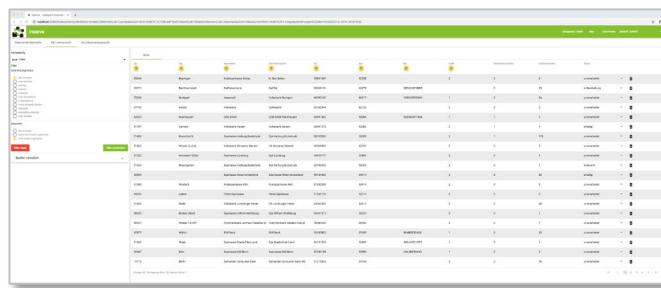
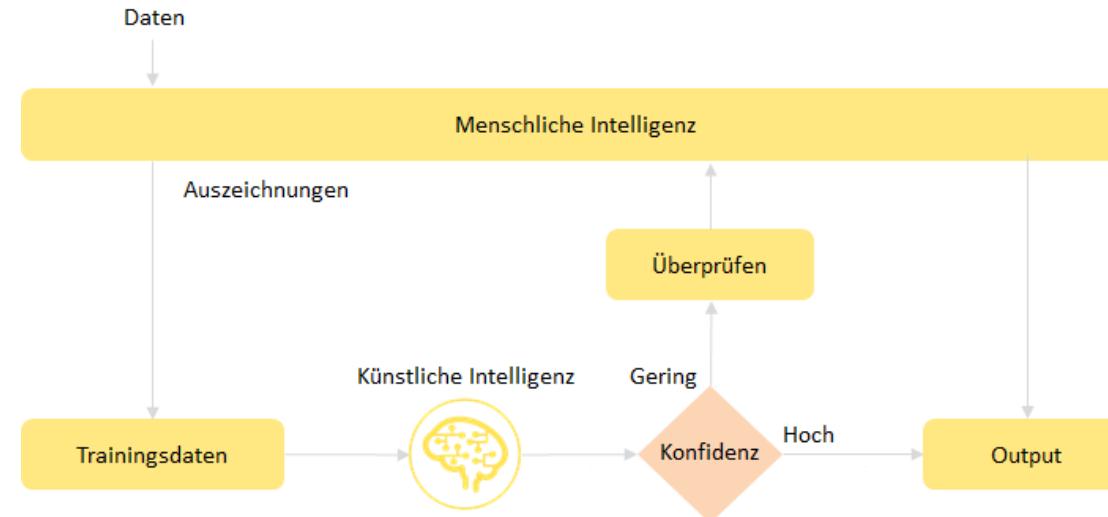
KI Erfolgsfaktoren



Minimum Viable Modell Strategie:

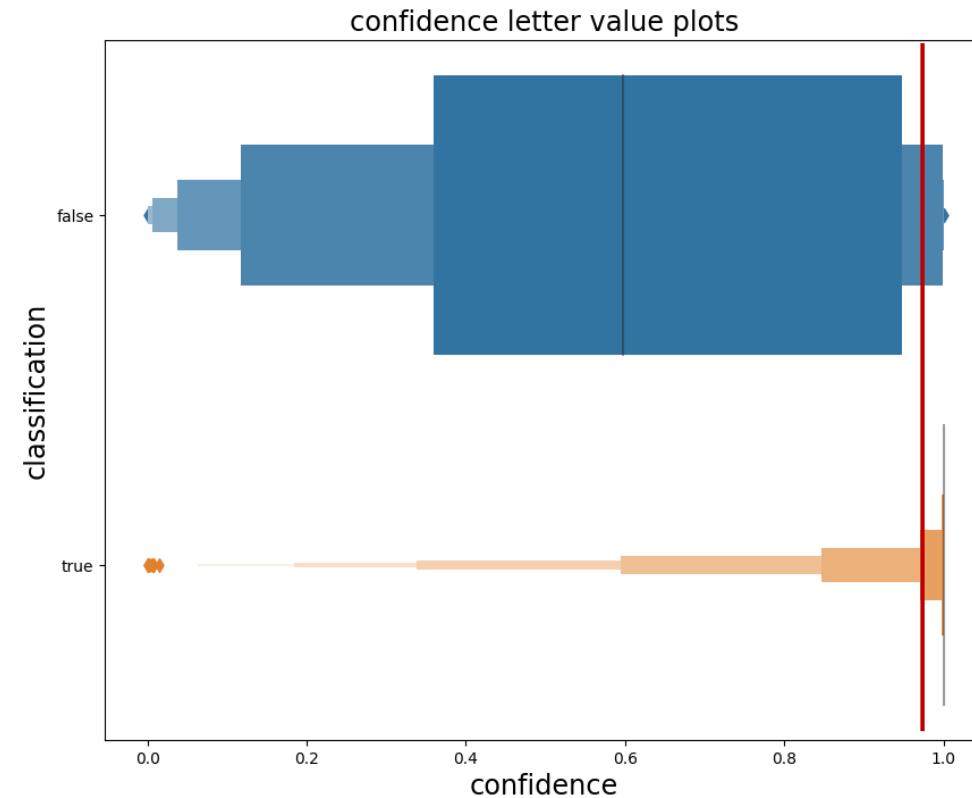
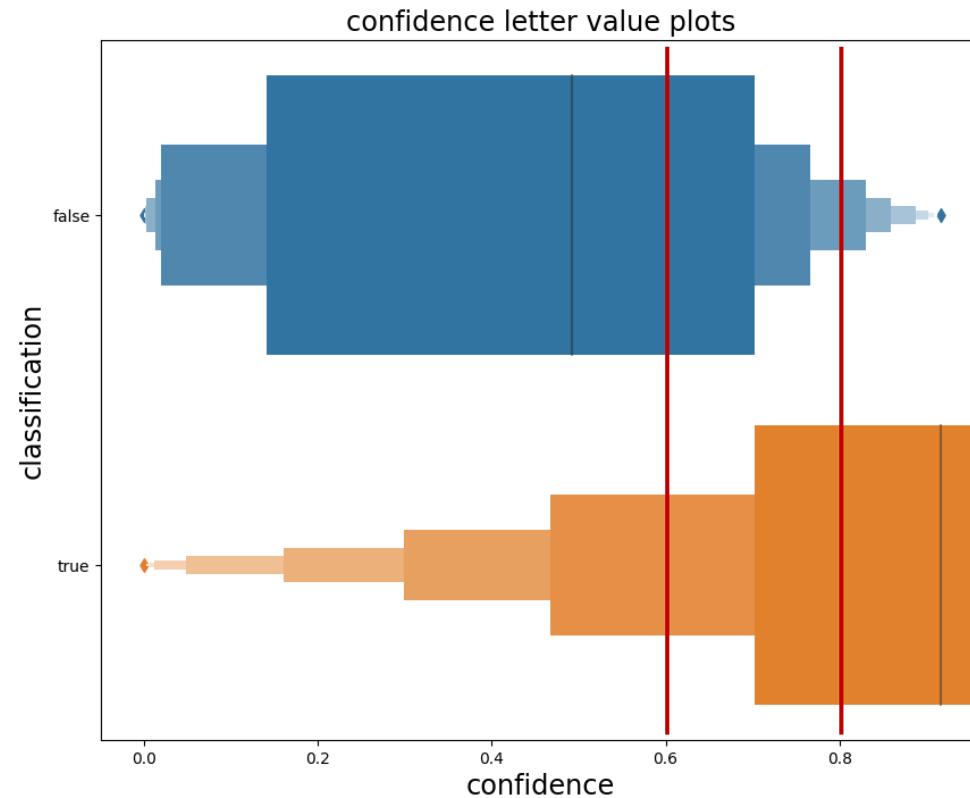
- Verfügbarkeit innovativer Modelle...
- ...aber Auswahl des einfachsten Modells

KI Erfolgsbausteine: HITL

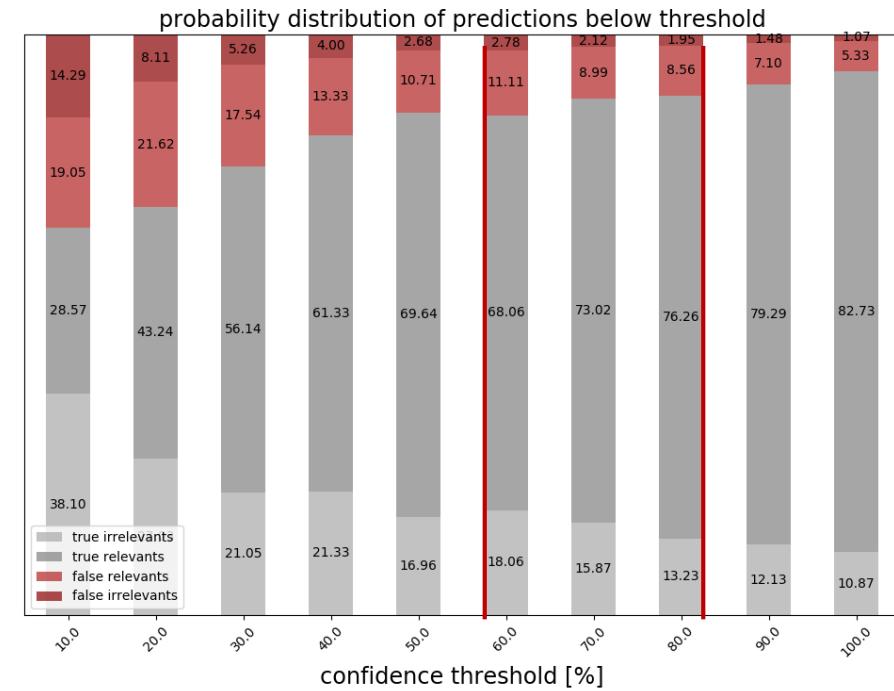
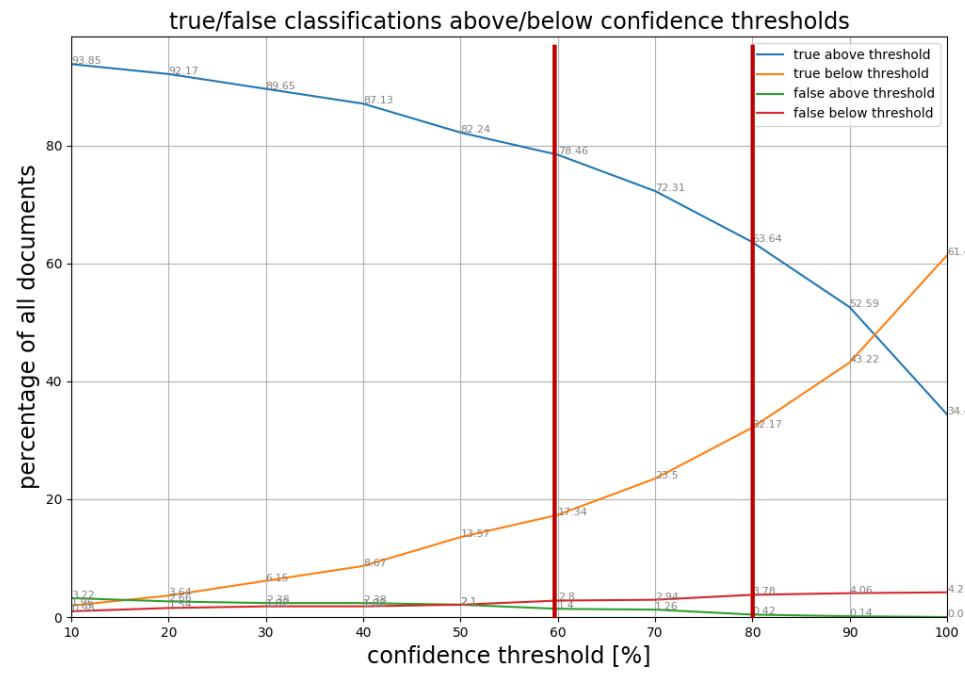


Human-in-the-Loop: kontinuierliche Verbesserung durch interaktiven Zyklus.

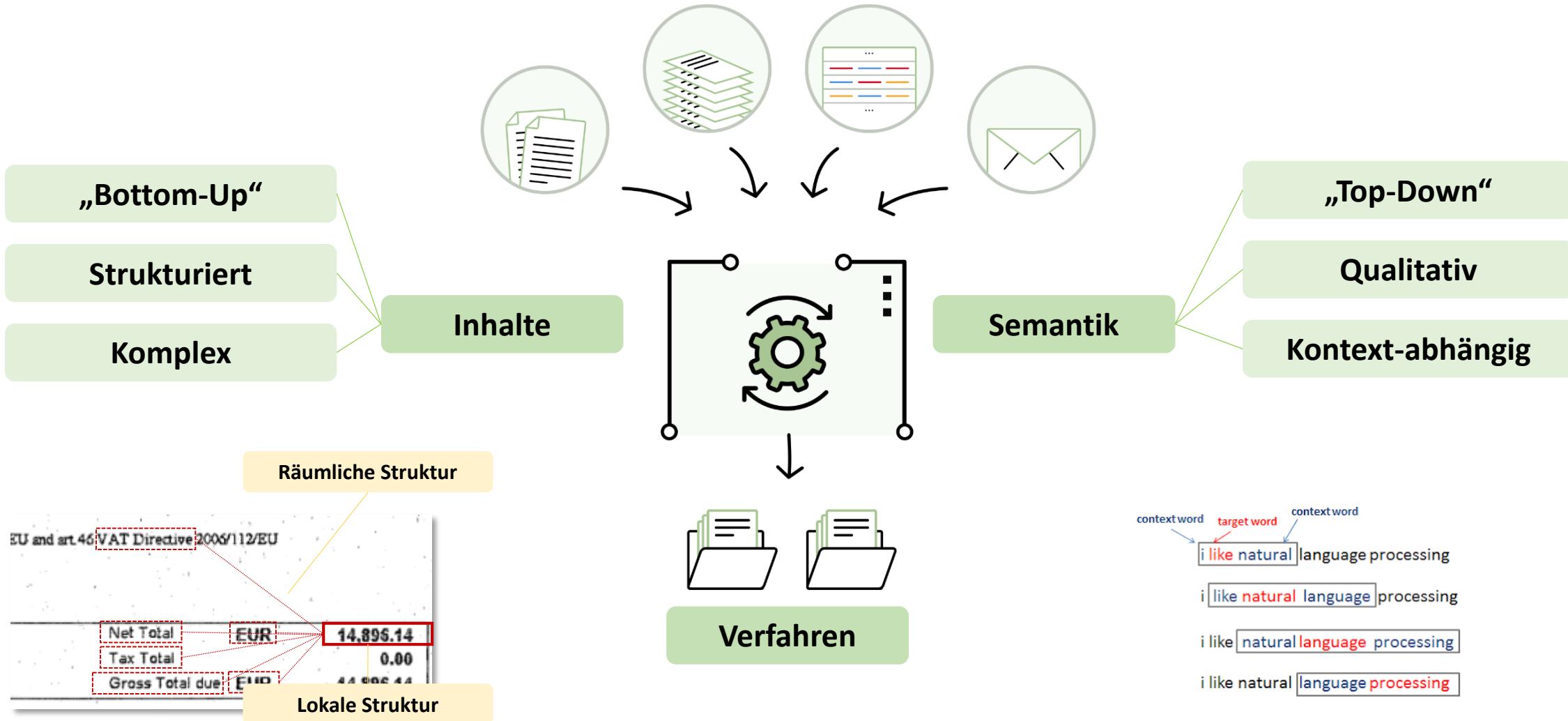
KI Erfolgsbausteine: HITL



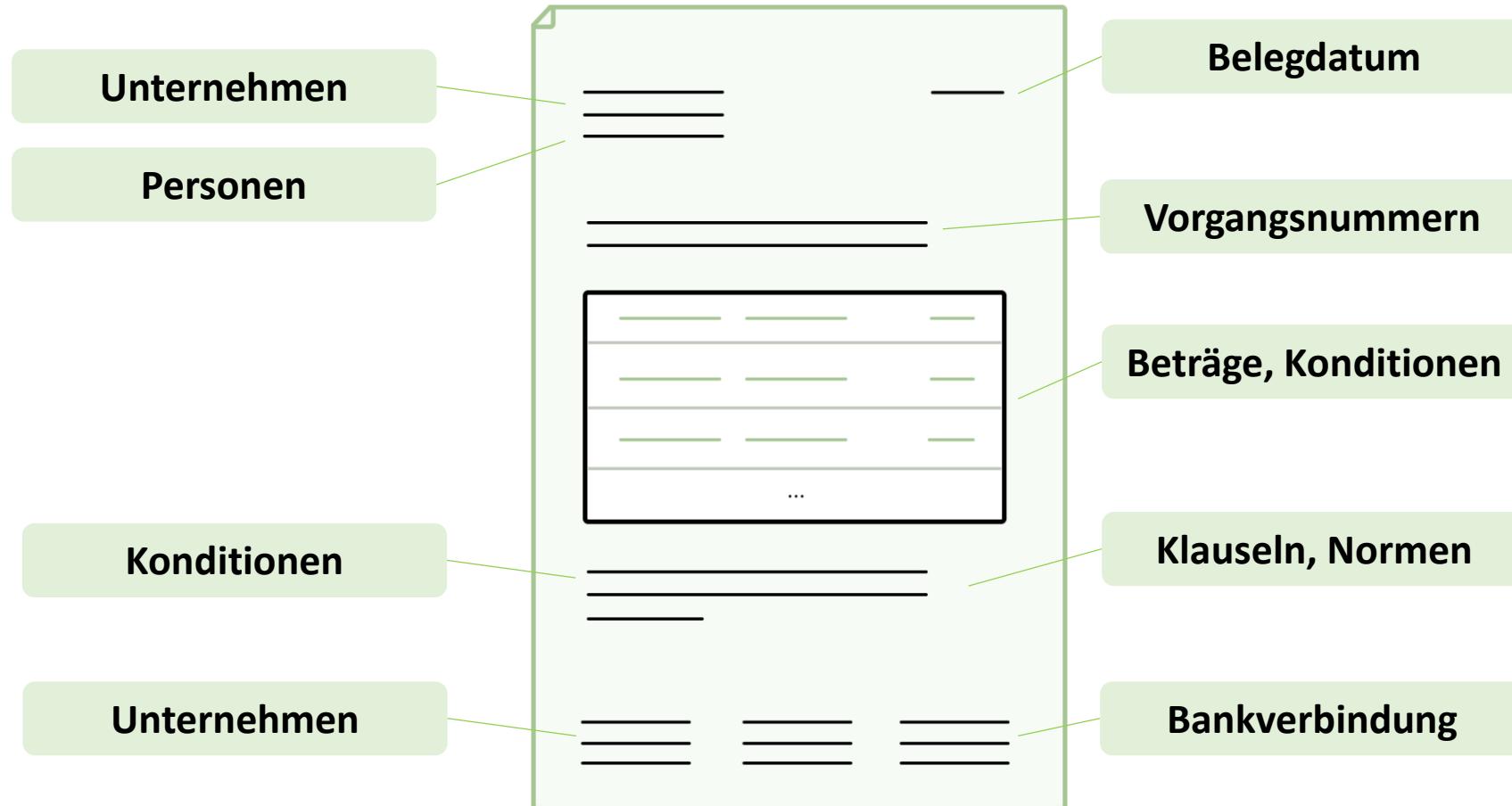
KI Erfolgsbausteine: HITL



Gesamtprozess: KI / Software



KI Bausteine: Extraktion



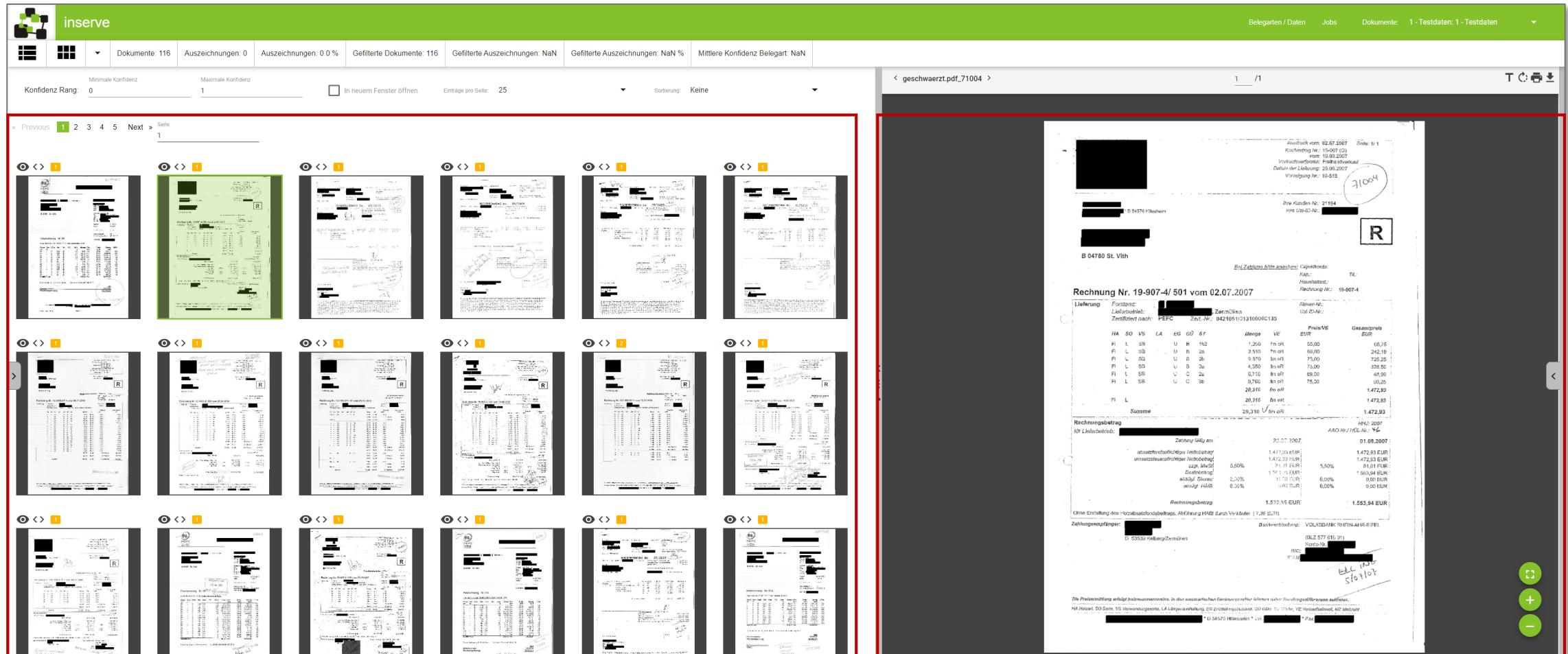
KI Bausteine: Extraktion

Maximale Nutzung der vorhandenen Informationen durch KI Modelle:

- Beispiel: Erkennung von Fußinformationen wie z.B. Belegnetto.
- Basierend auf Trainingsbeispielen:
 - Nutzung typischer Schlüsselwörter + deren Anordnung in der Umgebung.
 - Nutzung der lokalen Struktur z.B. XX,XXX.XX.
- Nutzung weiterer Information: Position, Bildinformation (z.B. Linien).
- Kombination dieser Features in einem lernfähigen KI Modell.
- Minimierung der benötigten Trainingsbeispiele + Maximierung der Erkennungsleistung.



Ausgangspunkt

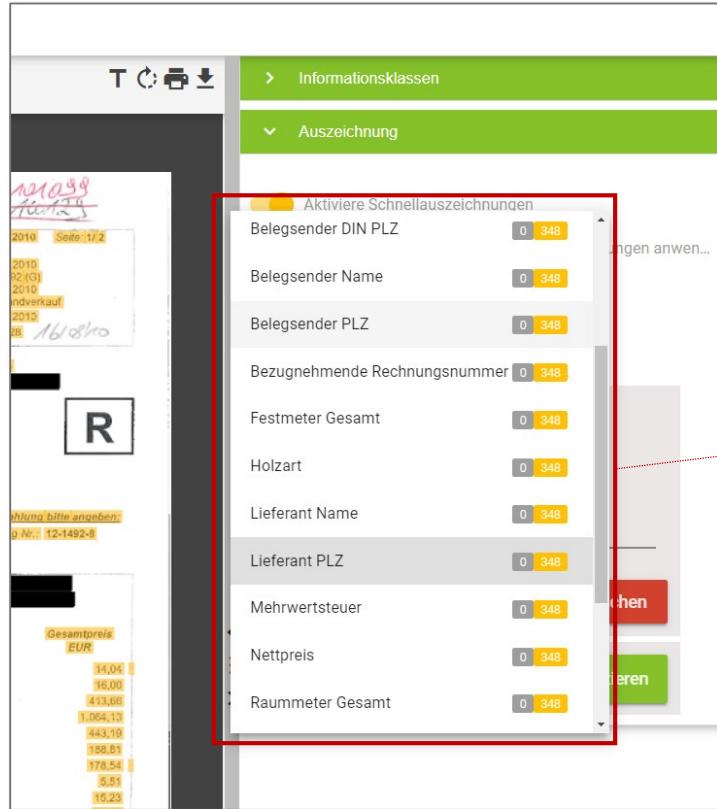


The screenshot displays the inserve document management system. On the left, a grid of 24 document thumbnails is shown, each with a red border. The thumbnails represent various scanned documents, likely invoices or receipts, with some containing handwritten text and stamps. On the right, a detailed view of a specific document is presented. The document header includes a date (02.07.2007), a reference number (19-907-4/501), and a name (B 04780 St. Vith). The main content is a table of a bill (Rechnung Nr. 19-907-4/501 vom 02.07.2007) with columns for description, quantity, unit, and price. The total amount is listed as 1.553,94 EUR. The bottom of the document contains a note about payment terms and a signature.

Übersicht aller Dokumente

Detailansicht Dokument

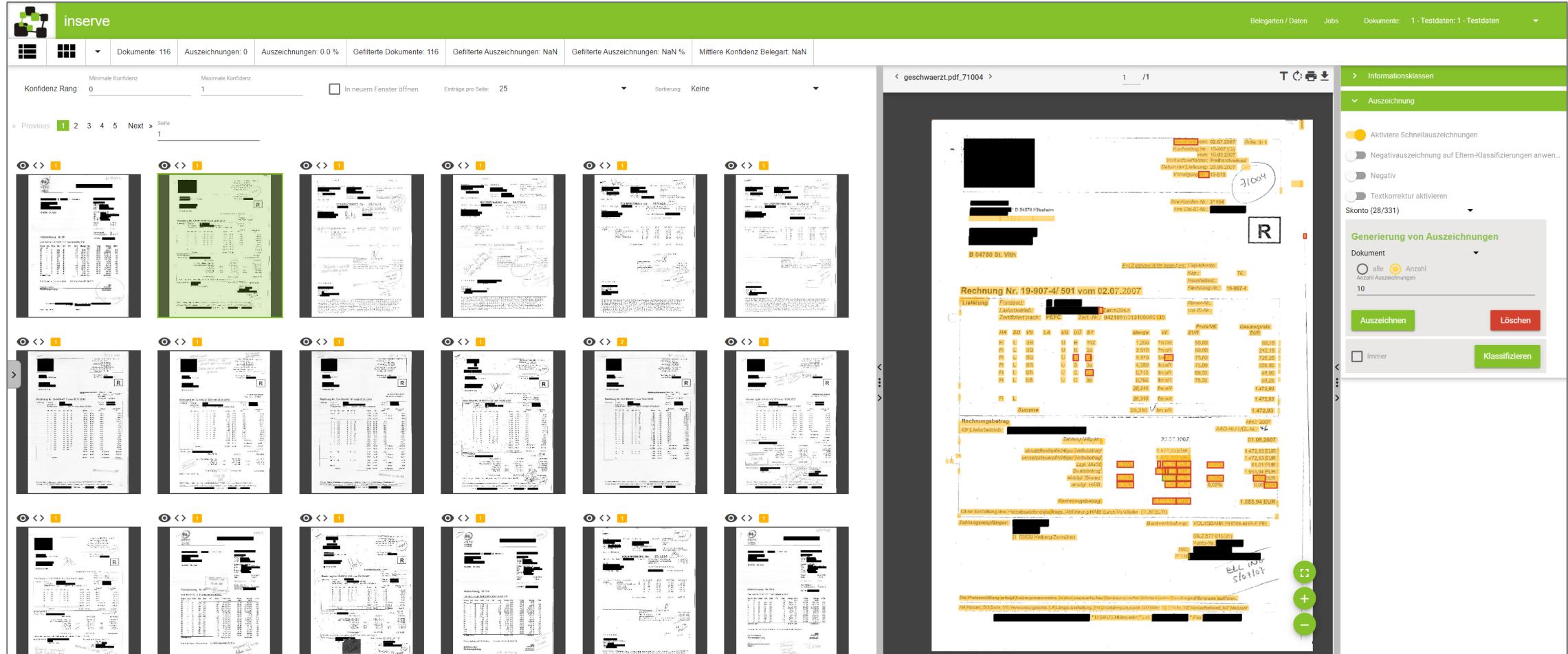
Definition Informationen



**Mit Kanzlei, Gutachter gemeinsam
definierte Informationen.**

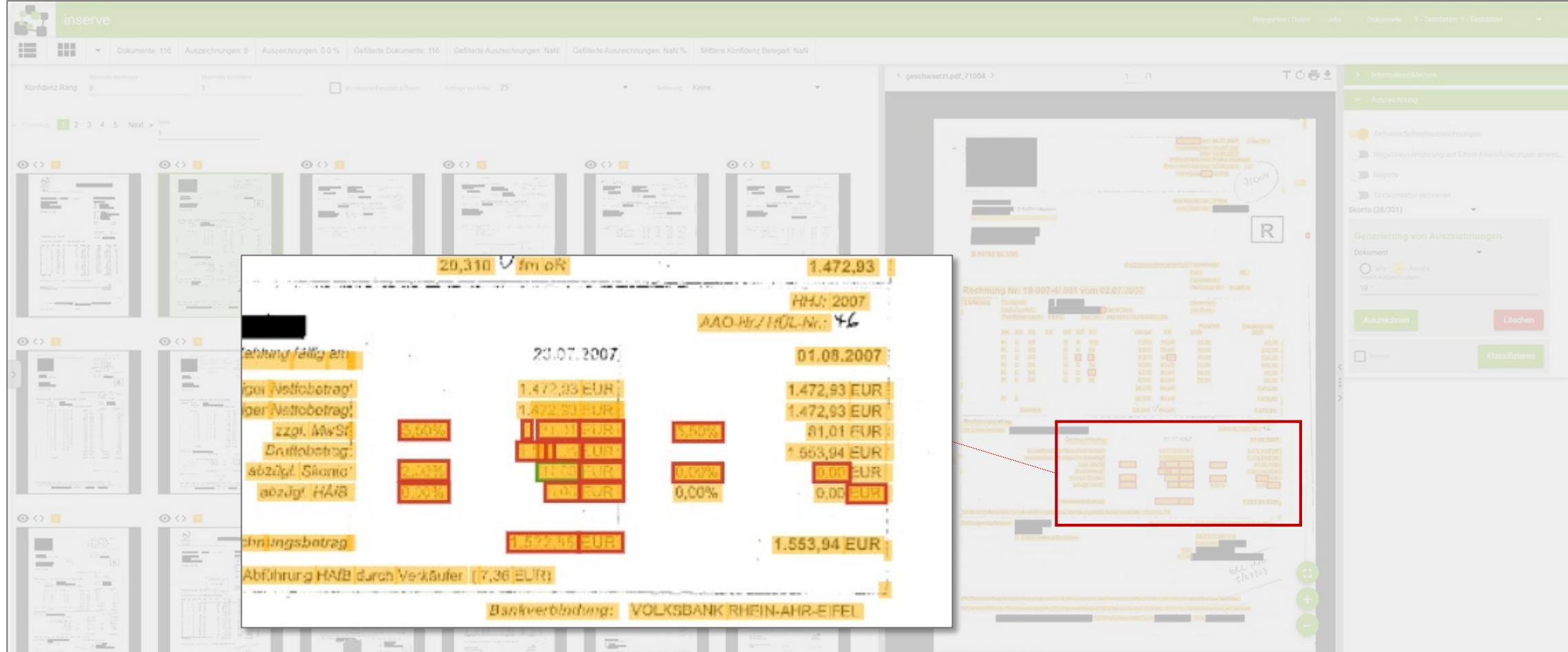
Erste Auszeichnungen

Erste Auszeichnungen



The screenshot displays the inserve software interface for document processing and annotation. The main area shows a grid of 24 document thumbnails, each with a red 'R' icon in the bottom right corner, indicating they are ready for review. The top navigation bar includes the inserve logo, document count (116), and various status indicators. The right sidebar contains configuration options for 'Informationsklassen' and 'Auszeichnung', including checkboxes for activating quick annotations, applying parent classification negative annotations, and enabling text correction. It also features a 'Generierung von Auszeichnungen' section with a 'Dokument' dropdown set to 'alle' and a 'Anzahl' input field set to 10, with 'Auszeichnen' and 'Löschen' buttons below. The bottom right of the interface shows a detailed view of a document titled 'geschwaerzt.pdf_71004', which is page 1 of 1. The document content is heavily redacted, but a table of charges and a total amount of 1.555,94 EUR are visible. The sidebar also includes a 'Klassifizieren' button.

Erste Auszeichnungen



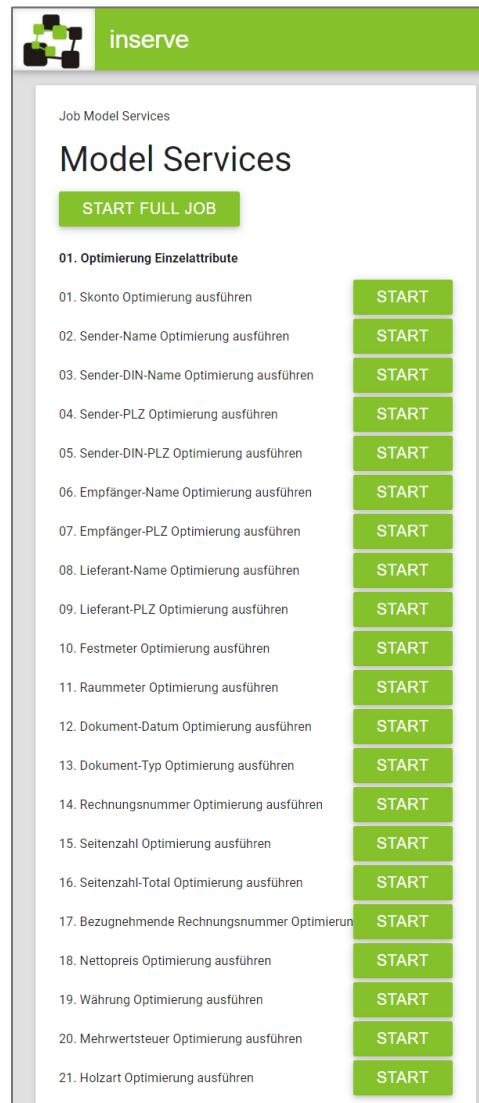
The screenshot shows the inserve software interface for document processing. The main window displays a document with several lines of text highlighted in yellow. The highlighted text includes:

- 20,310 fm OR
- 1.472,93
- HHJ: 2007
- AAO-Nr./HÖL-Nr.: 46
- 23.07.2007
- 01.09.2007
- 1.472,93 EUR
- 1.472,93 EUR
- 1.472,93 EUR
- 81,01 EUR
- 553,94 EUR
- 0,00 EUR
- 0,00 EUR
- 1.553,94 EUR
- Bankverbindung: VOLKSBANK RHEIN-AHR-EFFEL
- Abführung HAß durch Verkäufer (7,36 EUR)

A red box highlights a section of the document, specifically the line "Bankverbindung: VOLKSBANK RHEIN-AHR-EFFEL". The right side of the interface shows a sidebar with "Auszeichnung" (Annotation) settings and a "Generierung von Auszeichnungen" (Generation of Annotations) section. The "Auszeichnen" (Annotate) button is highlighted in green.

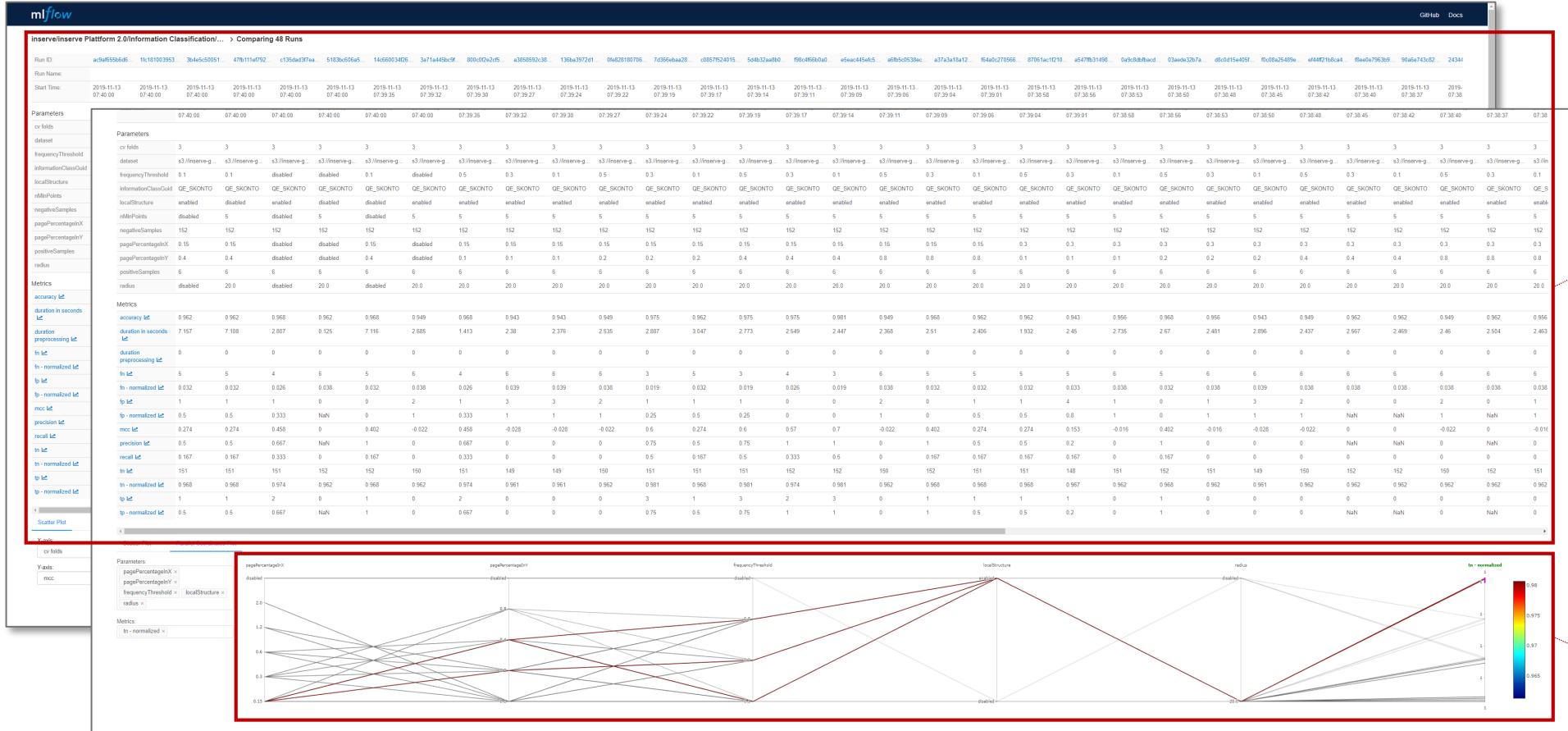
KI Modell = gelehriger Schüler: geschickte Auswahl von Trainingsbeispielen.

Modell trainieren / optimieren



**Starten von Trainings- und
Optmierungsläufen.**

Überprüfung der Modell Leistung



Optimierung = Erprobung verschiedener Modellkonfigurationen:
Auswahl genutzter Informationen, Größe der Umgebung (Kontext).

Visualisierung der Leistung des Modells für die (beste) Konfiguration(en).

Transparenz der Modell Leistung

| | |
|-------------------------------------|--------------|
| absatzfondspflichtiger Nettobetrag | 7.441,08 EUR |
| umsatzsteuerpflichtiger Nettobetrag | 7.441,08 EUR |
| zzgl. MwSt | 409,26 EUR |
| Bruttobetrag | 7.850,34 EUR |
| abzügl. Skonto | 2,00% |
| abzügl. HAFB | 157,31 EUR |
| | 0,00 EUR |

localStructure : 34.5%
- "letter-index 0" : 21.9%
- "letter-index 1" : 9.7%
- "letter-index 2" : 1.3%
- "letter-index 5" : 0.9%
- "letter-index 3" : 0.7%

Die lokale Struktur (Muster der Zeichenketten) von Skontoangaben ist für das beste Modell sehr wichtig.

neighborhoodDirections : 16.1%
- "skonto" : 6.8%
- "abzüglich" : 2.9%
- "mwst" : 2.6%
- "2" : 2.0%
- "netto" : 1.9%
neighborhoodEditDistance : 14.9%
- "skonto" : 11.6%
- "bis" : 1.0%
- "zahlung" : 0.8%
- "abzüglich" : 0.8%
- "%" : 0.6%

(Schlüssel) Worte in der Umgebung von Skontoangaben sind in Summe am wichtigsten. Dabei spielen Entfernung, Richtung und Fehlertoleranz ungefähr die gleiche Rolle.

neighborhoodDistance : 14.2%
- "skonto" : 7.6%
- "aus" : 1.8%
- "netto" : 1.8%
- ":" : 1.7%
- "mit" : 1.3%
positionLikelihood : 0.6%

Die Position der Skontoangaben auf der Seite spielt nur eine geringe Rolle (- da diese stark variiert).

Stichproben Überprüfung

The image displays three side-by-side screenshots of a software application, likely a document management or accounting system, with a green header bar and a sidebar on the right.

Header: Belegarten / Daten, Jobs, Dokumente: 1 - Testdaten; 1 - Testdaten

Screenshot 1 (Left): A document titled "geschwaerzt.pdf_90353" with handwritten notes. It includes a table of items with columns: HA, SO, VS, LA, EG, GU, ST, Menge, VE, Preis/VE, and Gesamtpreis EUR. The total sum is 7.441,08 EUR. A red box highlights a section of the text.

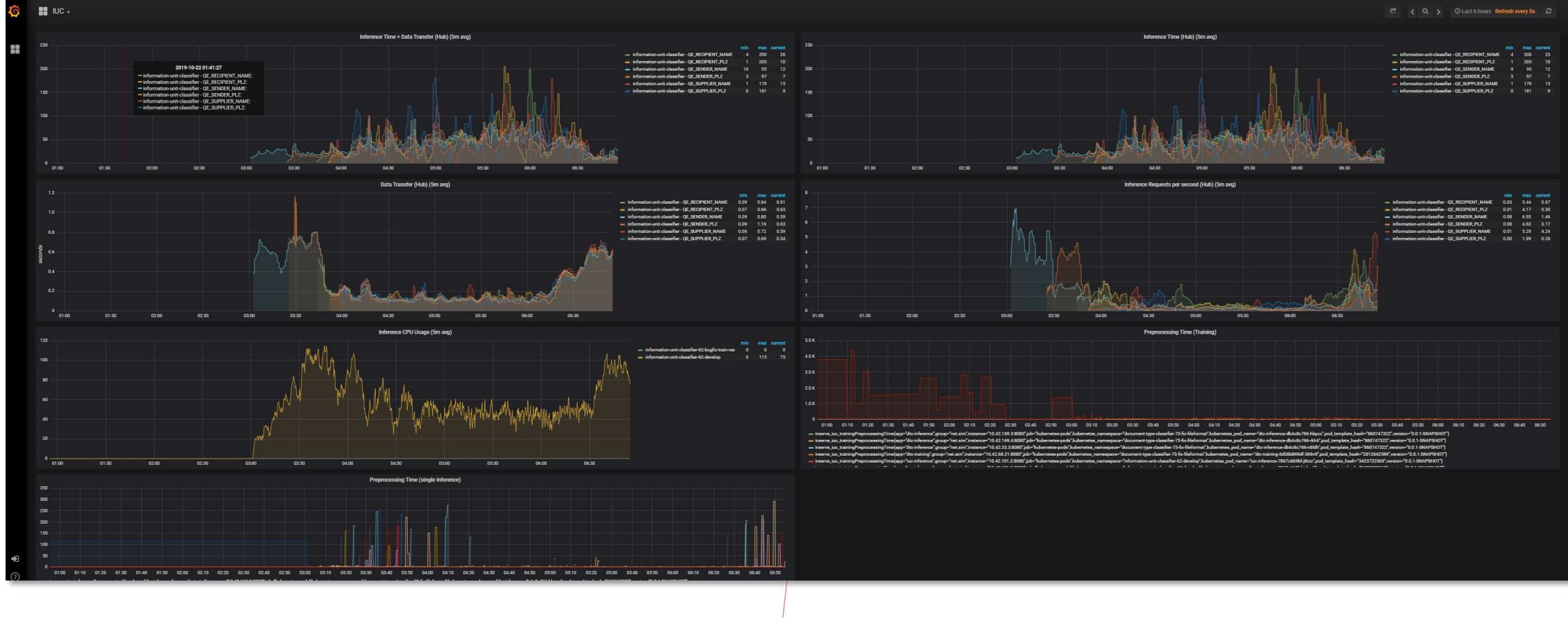
Screenshot 2 (Middle): A document titled "geschwaerzt.pdf_71912" with a table of items. The total sum is 256,96 EUR. A red box highlights a section of the text.

Screenshot 3 (Right): A document titled "geschwaerzt.pdf_110075" with a table of items. The total sum is 2.276,33 EUR. A red box highlights a section of the text.

Right Sidebar:

- Informationsklassen:** A list of classification types: Beide (yellow circle), Positiv (grey circle), and Negativ (white circle).
- Auszeichnungen:** A list of checkboxes: Immer (checked), Auszeichnung (grey), and Klassifizieren (green).
- Skonto (0):** A section showing the current discount percentage (0).
- Auszeichnung:** A section with a "Klassifizieren" button.

Gesamt Inferenz – Überwachung



Ein Rechenbeispiel zur Motivation von Skalierbarkeit, Stabilität und Überwachung:
300.000 Seiten x 300 Worte = 90.000.000 Millionen Inferenzen
x 20 Informationen = 1.8 Milliarden Inferenzen

KI Bausteine: Entstörung



Original

| Menge | Einzel-Preis | Gesamtbetrag ohne MwSt. | MwSt. % | MwSt. | Gesamtbetrag mit MwSt. |
|------------|--------------|-------------------------|---------|--------|------------------------|
| EUR | | | | | |
| N 14.720 T | 335,00 | 4931,20 | 7,00 | 345,18 | 5.276,38 |
| B 14.720 T | | | | | |
| 4.931,20 | 1,500% | 73,97 | 7,00 | 5,18 | 79,15- |

ABSCHLAG/DISCOUNTS -1.40 -140,00
SUBTOTAL: 5450,00

*45: 22.5 X 8.25; 12 R 22,5 152/148 L [2]
22.5 X 9.00; 13 R 22,5 154/150 K [2]
22.5 X 8.25; 275/70 R 22,5 148/145 L [2]
22.5 X 9.00; 295/60 R 22,5 150/147 K [2]
22.5 X 8.25; 305/70 R 22,5 152/148 L [2]
22.5 X 9.00; 315/60 R 22,5 152/148 L [2]
22.5 X 9.00; 315/70 R 22,5 154/150 L [2]
22.5 X 9.00; 315/80 R 22,5 154/150 M [2]

zum Preis von Euro
zuzüglich 20 % Mehrwertsteuer 457.838,85
Endbetrag 91.567,77
===== 549.406,62
Zur Information, Beträge in ATS: Netto: 6.299.999,93 MWSt: 1.259.999,99 Endbetrag: 7.559.999,91

Verbesserung

| Menge | Einzel-Preis | Gesamtbetrag ohne MwSt. | MwSt. % | MwSt. | Gesamtbetrag mit MwSt. |
|------------|--------------|-------------------------|---------|--------|------------------------|
| EUR | | | | | |
| N 14.720 T | 335,00 | 4931,20 | 7,00 | 345,18 | 5.276,38 |
| B 14.720 T | | | | | |
| 4.931,20 | 1,500% | 73,97 | 7,00 | 5,18 | 79,15- |

ABSCHLAG/DISCOUNTS -1.40 -140,00
SUBTOTAL: 5450,00

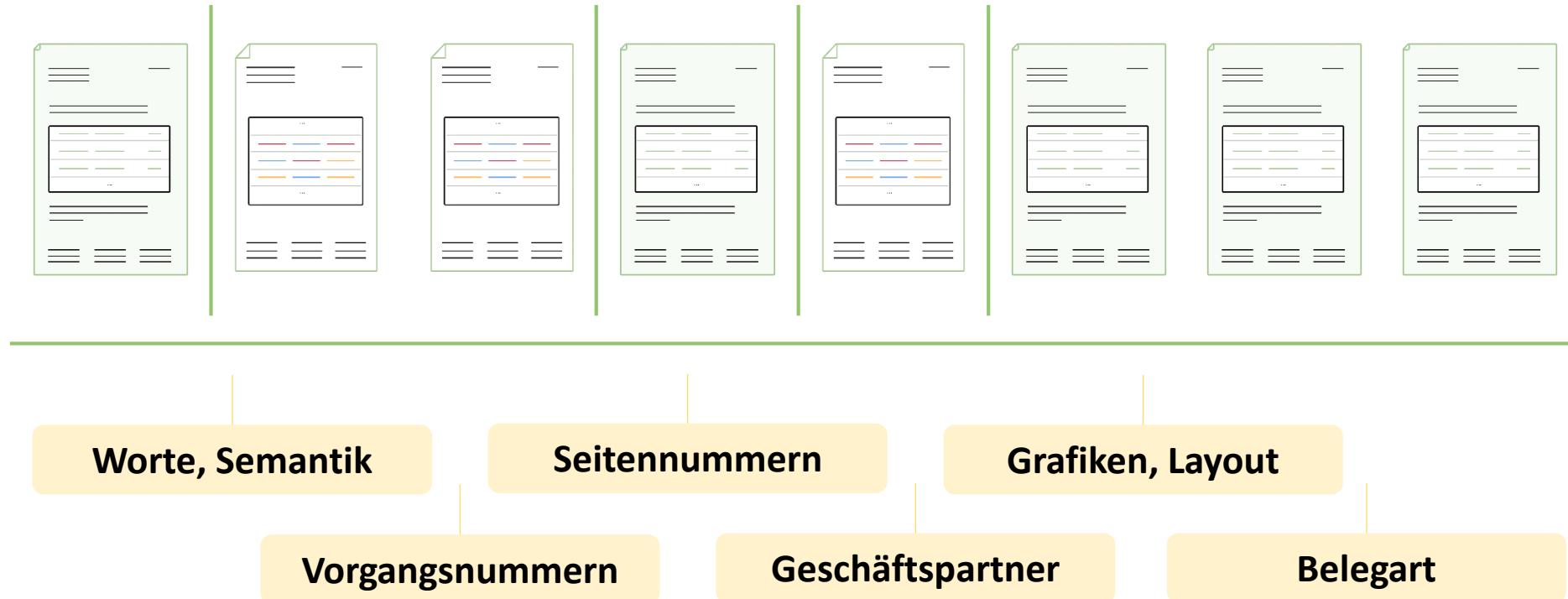
*45: 22.5 X 8.25; 12 R 22,5 152/148 L [2]
22.5 X 9.00; 13 R 22,5 154/150 K [2]
22.5 X 8.25; 275/70 R 22,5 148/145 L [2]
22.5 X 9.00; 295/60 R 22,5 150/147 K [2]
22.5 X 8.25; 305/70 R 22,5 152/148 L [2]
22.5 X 9.00; 315/60 R 22,5 152/148 L [2]
22.5 X 9.00; 315/70 R 22,5 154/150 L [2]
22.5 X 9.00; 315/80 R 22,5 154/150 M [2]

zum Preis von Euro
zuzüglich 20 % Mehrwertsteuer 457.838,85
Endbetrag 91.567,77
===== 549.406,62
Zur Information, Beträge in ATS: Netto: 6.299.999,93 MWSt: 1.259.999,99 Endbetrag: 7.559.999,91

KI Bausteine: Dokument Trennung

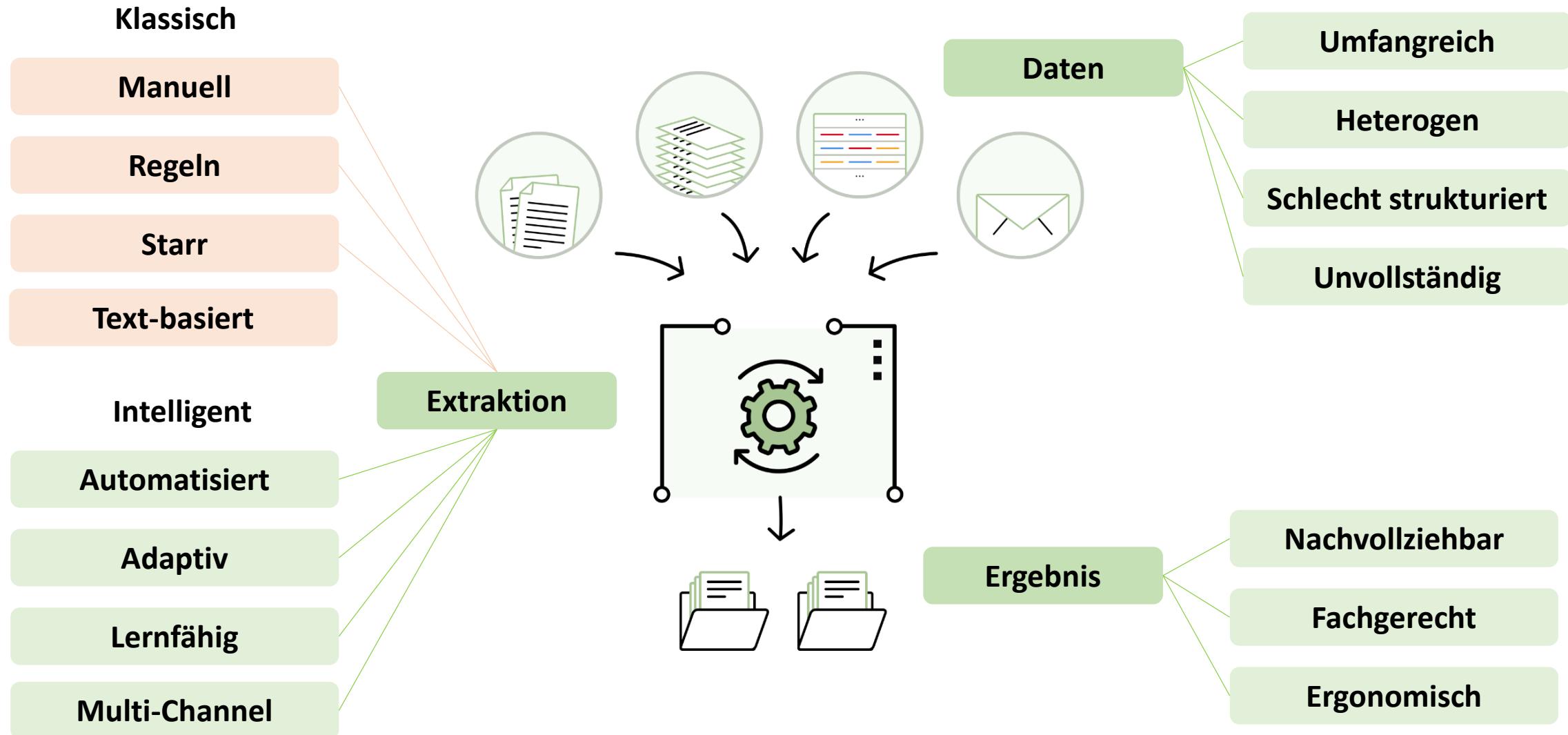


KI Bausteine: Dokument Trennung



KI Bausteine: Klassifikation

Gesamtprozess: KI / Software



Kartellrecht – Private Enforcement

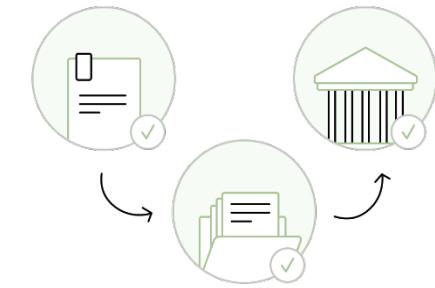
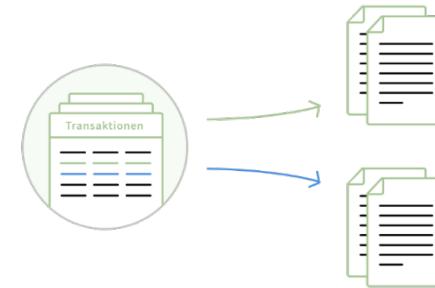
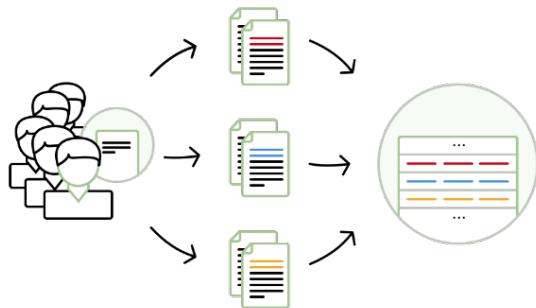
Akquisition

Transaktionslisten

Marktmodell

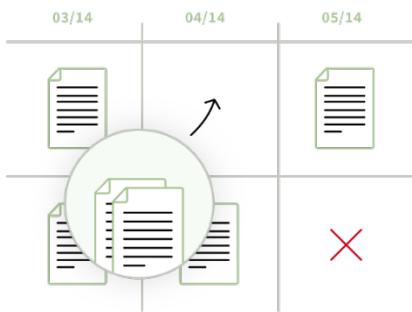
Nachweis

Vortrag / Replik

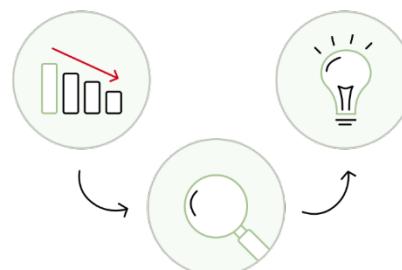


Informationsextraktion

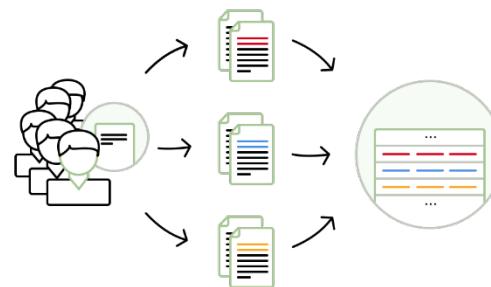
Gutachten



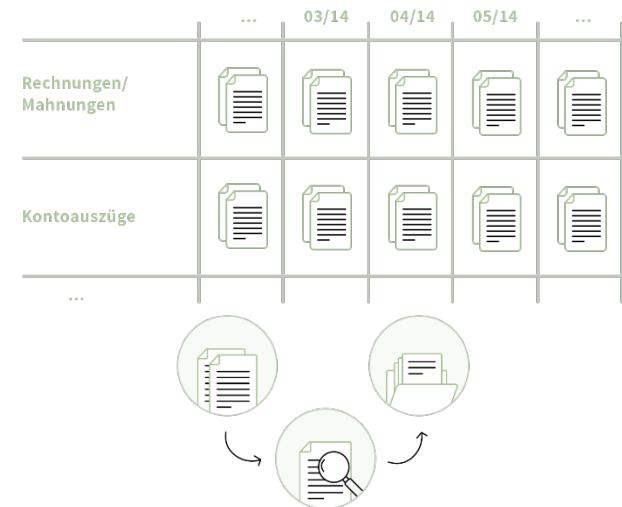
Ermittlung ZU



Tabellenführung

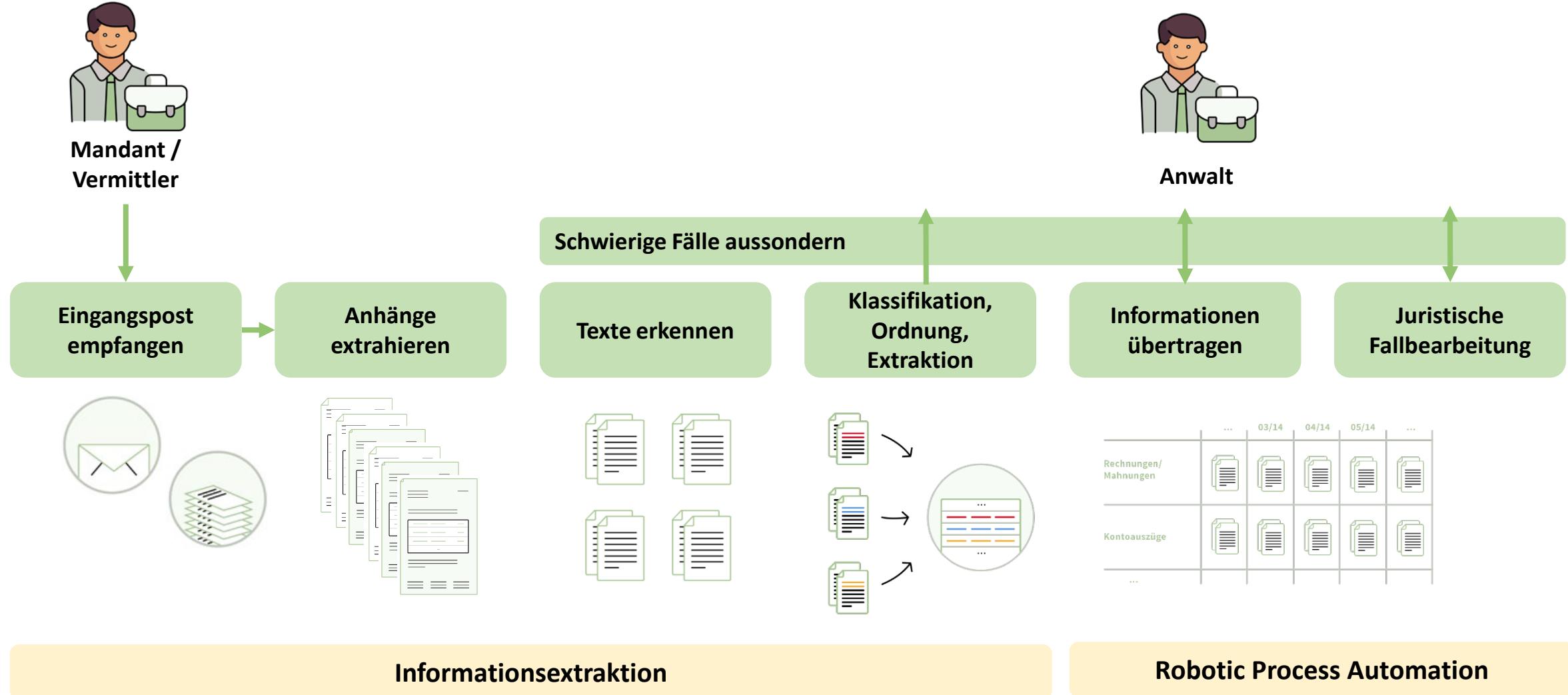


Anfechtung



Informationsextraktion

Kanzleiautomatisierung



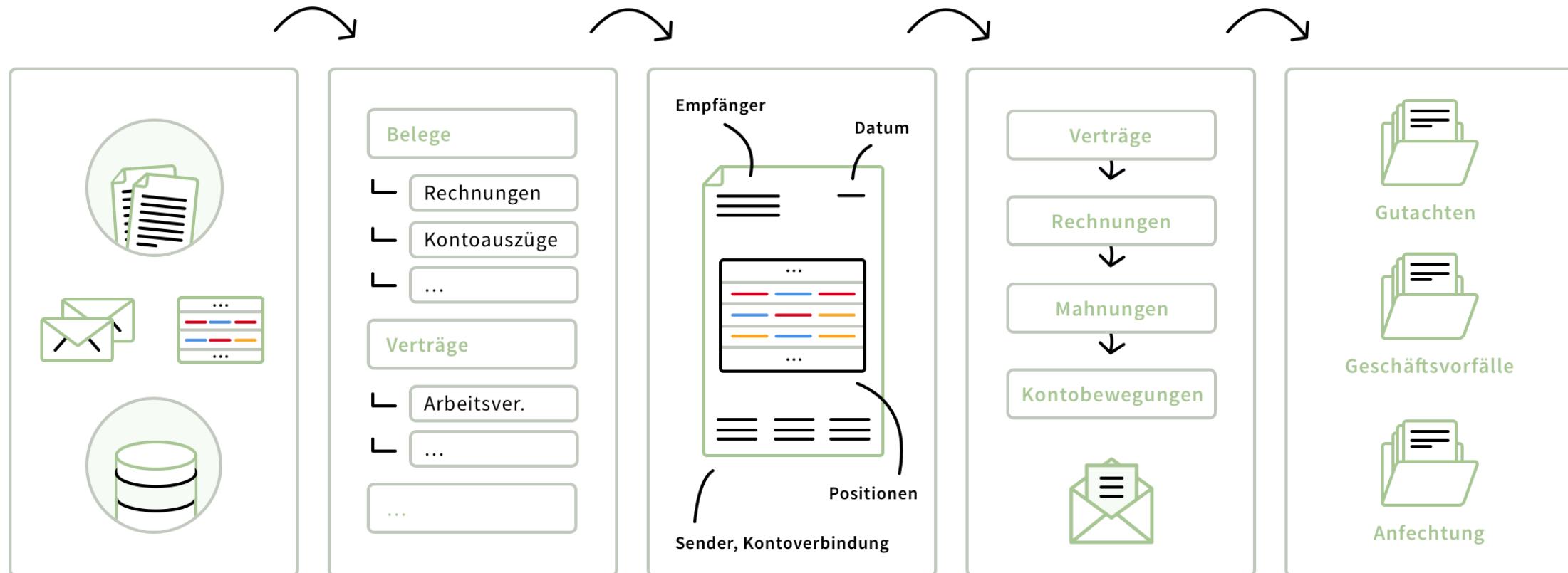
Fall Beispiele / Typische Parameter



| Fall | Mengengerüst |
|--|--|
| Kartellverfahren im Bereich Verschiffung | 1 Mandant, ~10 Lieferanten, ~300.000 Seiten |
| Kartellverfahren im Bereich „XXX“ | Mehrere 100 Mandanten, ~100-200.000 Seiten |
| Kartellverfahren Rundholz | 70 Mandanten, ~400 Lieferanten, 1 Mio. Seiten |
| Kanzlei Automatisierung | ~5.000 Fälle / Jahr |
| Insolvenzverfahren | ~300 Kreditoren, ~10-15.000 Seiten, FiBu Daten |
| Dokumenten Eingangsverarbeitung | ~8-10. Mio. Seiten / Jahr |

- Zwischen 5 und 300 Informationen pro Dokument
- Zwischen 12 Monaten (kontinuierliche Akquisition) und 4 Wochen Laufzeit
- Bandbreite von digitaler Anlieferung bis Abholung Aktenordner bei Mandanten

Gesamtprozess: KI / Software



Erkennung

Ordnung

Extraktion

Verknüpfung

Bereitstellung

Gesamtprozess / Anforderungen

| | Prozess Finanzierer | Gutachter | Rechts-anwälte | Mandanten | Dienstleister |
|----------------------------------|---------------------|-----------|----------------|-----------|---------------|
| Parameter festlegen & überwachen | ◆ | ◆ | ◆ | ◆ | ◆ |
| Dokumenten Logistik steuern | | | | ◆ | ◆ |
| Digitalisierung / Import | | | | | ◆ |
| Datenmodell abstimmen | ◆ | ◆ | | | ◆ |
| Klassifizierung & Extraktion | | | | | ◆ |
| Validierung & Qualitätssicherung | | | | | ◆ |
| Validierung & Abnahme | ◆ | ◆ | ◆ | ◆ | ◆ |
| Bereitstellung & Integration | ◆ | ◆ | ◆ | | ◆ |

- 1. Strukturiertes Datenhandling + Kollaboration**
- 2. Entlastung von Anwälten und Sekretariaten**
- 3. Mehrwerte für Mandanten**
- 4. Grundlage für weitere Automatisierung**
- 5. Bessere / frühere Risikobewertung von Ansprüchen**
- 6. Schneller und einfacher Datenzugriff**
- 7. Erhöhung Masse bzw. Erfolgspotential**
- 8. Substantiierte Klageschrift – weniger Repliken – schnellere Einigung**

Fragen & Antworten



- Vertiefung KI?
- Einstieg & Erprobung ?
- Ablauf ?
- Reaktions- & Laufzeit ?
- Datenaustausch ?
- Datenquellen ?
- Datenmengen ?
- Kosten ?
- ...

Wir freuen uns auf den weiteren Austausch.

Kontakt:

Arvin Arora, aarora@inserve.de

Ben Lobmüller, blobmueller@inserve.de