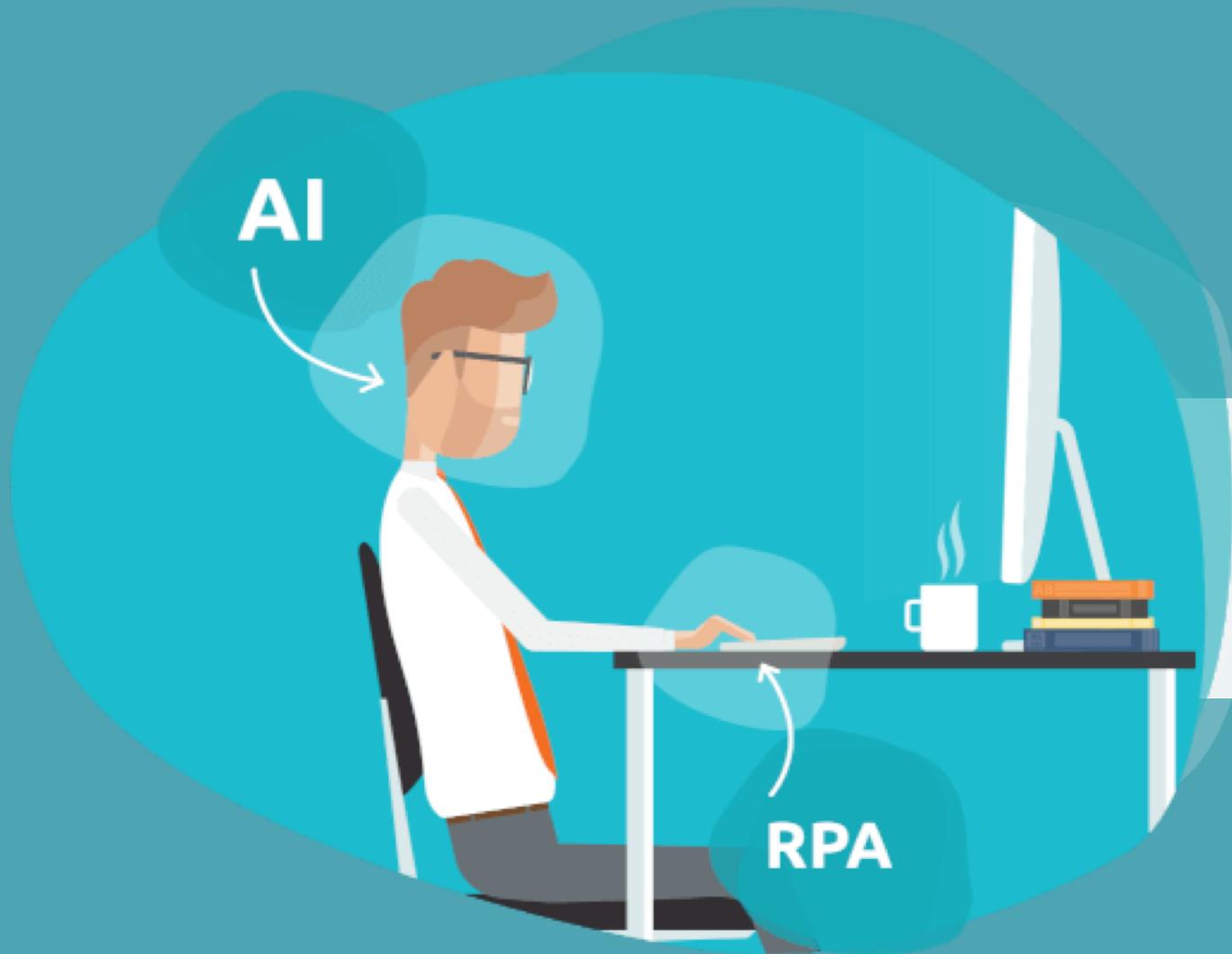




RPA UND KI KOMBINIEREN

# WIE AUTOMATISIERT MAN PROZESSE?





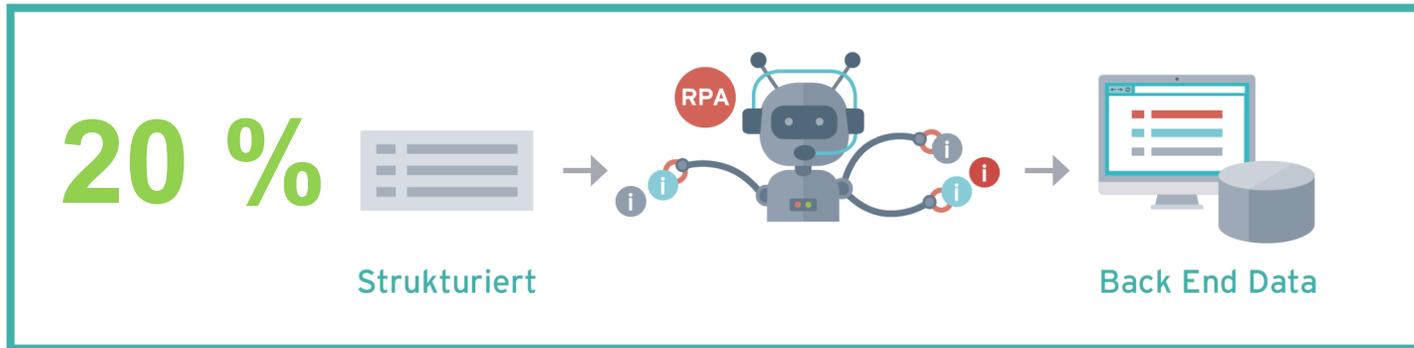
Das RPA-Dilemma:

Strukturierte Daten

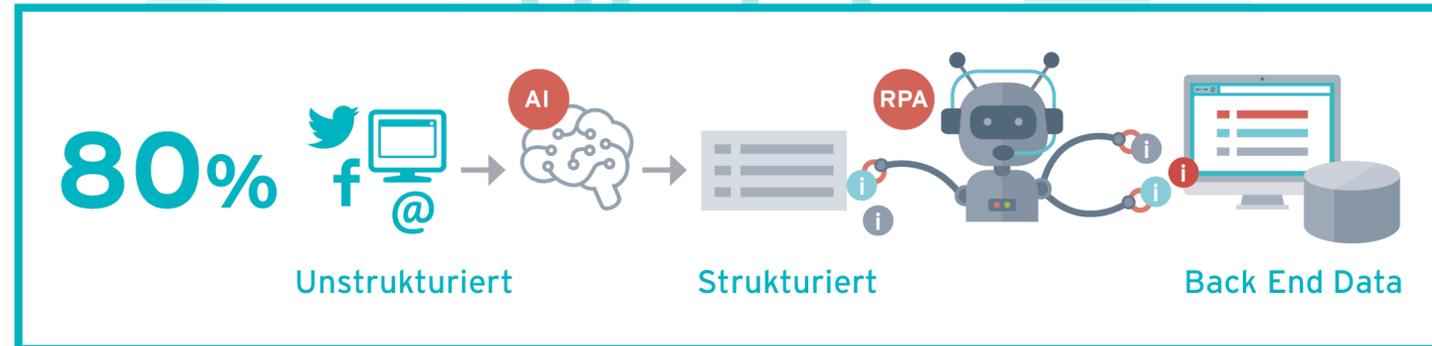


## Welche Prozesse sind grundsätzlich für RPA geeignet?





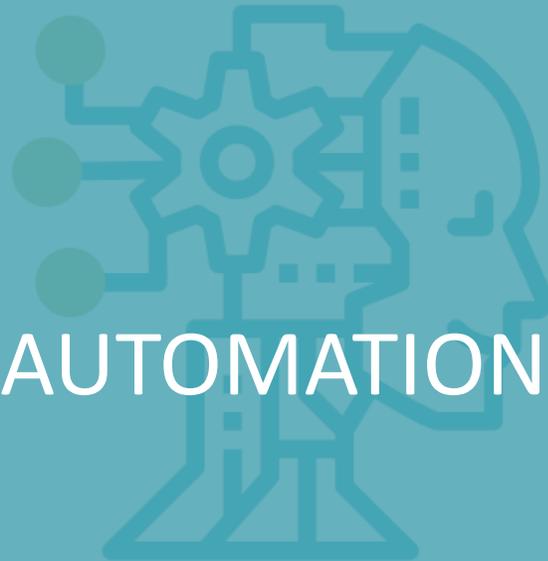
... aber die meisten Daten im Unternehmen sind unstrukturiert. Sie müssen erkannt und in eine strukturierte Form überführt werden.

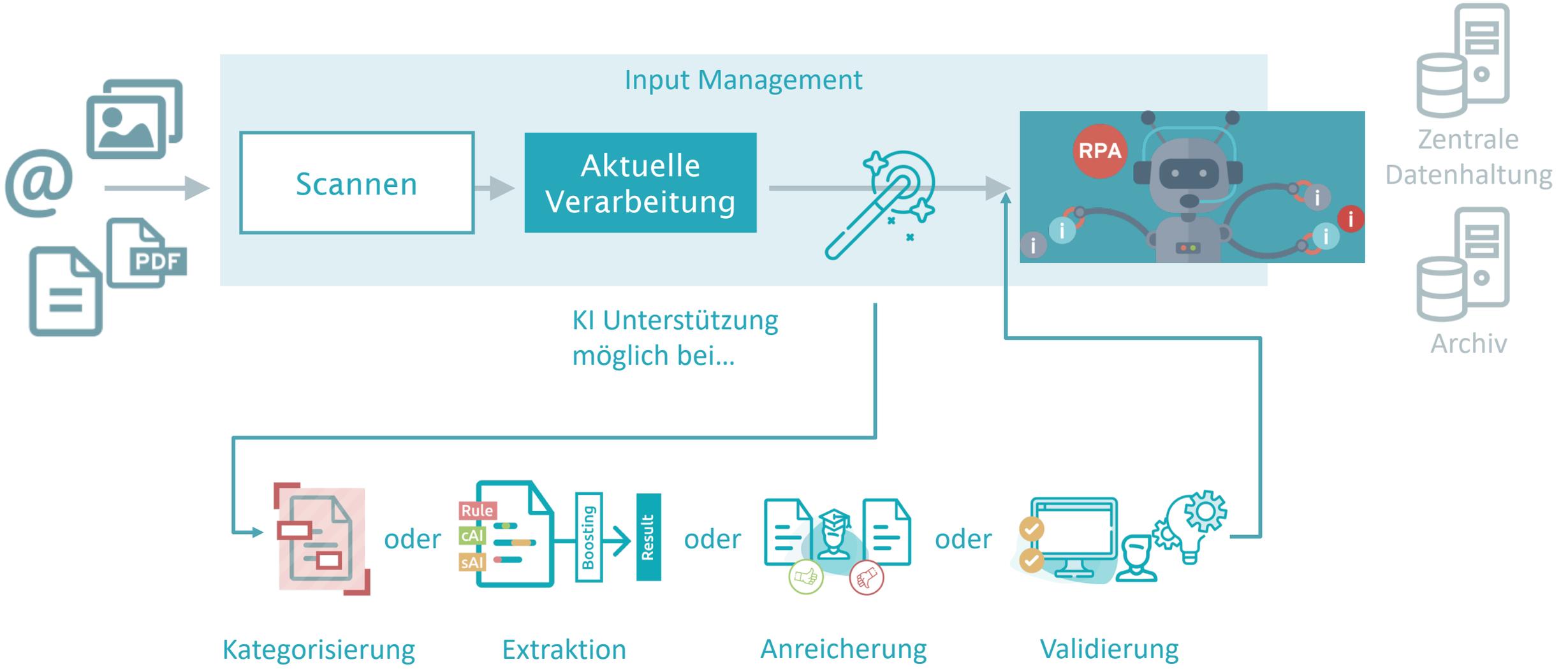


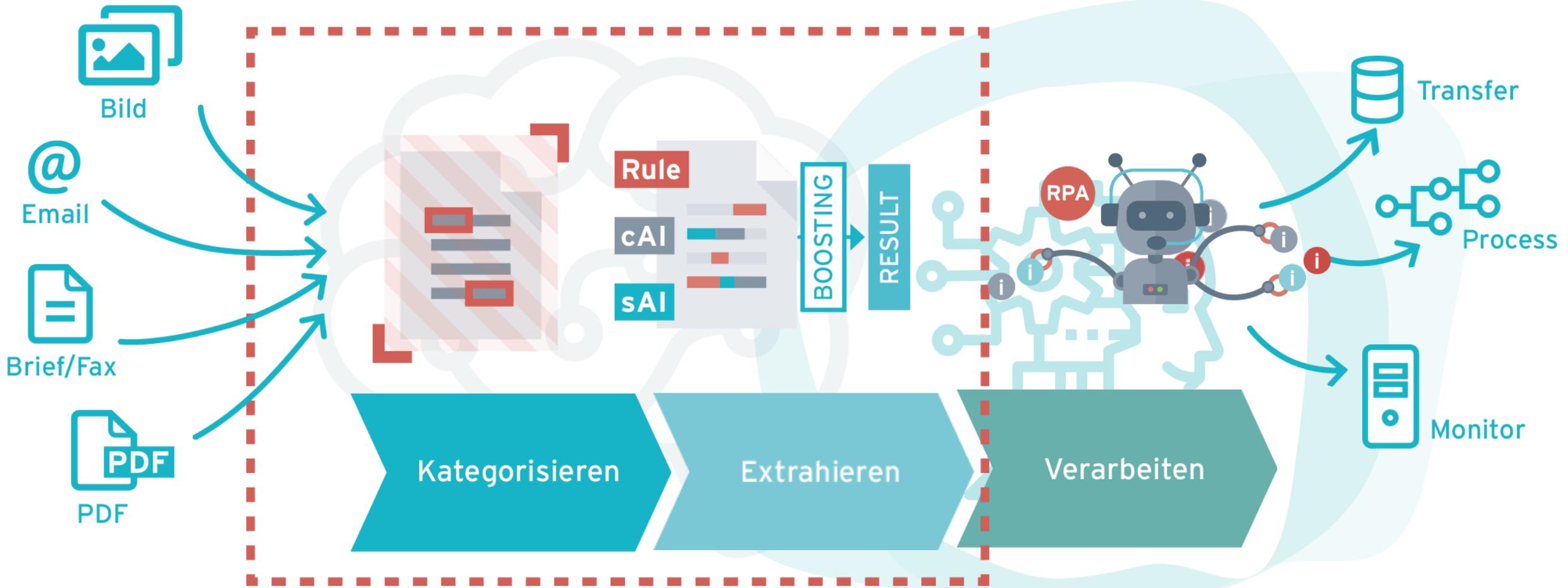


RPA UND KI KOMBINIEREN

# INTELLIGENT PROCESS AUTOMATION (IPA)







An: kundenservice@stadtwerke-karlsruhe.de  
Von: max.mustermann@eigentuemergemmm.de  
Betreff: Vertragskonto 210736562

Sehr geehrte Damen und Herren,

die zu diesem Vertragskonto gehörende Wohnung/Verbrauchsstelle wurde am 27.09.2020 neu vermietet.

Hier die End- bzw. Anfangsstände der Zähler:

Strom	Zählernummer 1ITR0055230469	Zählerstand 704
Gas	Zählernummer 233014	Zählerstand 12288,254

Die neuen Mieter für die Verbrauchsstelle in 76227 Karlsruhe-Durlach, Grenzstraße 14 sind seit 27.09.2020

Lars Müller-Lüdenscheid und Nils Holgerson

Die Endrechnung für die knapp 2 Monate Leerstand bitte an:

Eigentümergeinschaft  
Max und Moritz Mustermann  
Ziegelstraße 23  
69190 Walldorf

Den fälligen Betrag buchen sie bitte von dem Konto der Eigentümergeinschaft ab:

IBAN: DE98 6725 0020 0003 2046 75

BIC: SOLADES1HDB

Sparkasse Heidelberg

Kontoinhaber: Max und Moritz Mustermann

Auf dieses Konto überweisen Sie bitte ebenfalls die Gutschrift in Höhe von 1067,23 € bzgl. der Gesamtrechnung des Objektes

Kirchenstr. 39  
76676 Graben-Neudorf  
VK 210693571

Sollten Sie noch Fragen haben, stehe ich jederzeit zur Verfügung.

Liebe Grüße aus Walldorf  
Max Mustermann

---

Max Mustermann  
Hauptstraße 54  
69190 Walldorf

Tel.+Fax: 06227-4711  
mobil: 0172-70815612

PTME Power-Text-Mining-Engine (ITyX)

Neural Mind (ITyX)

Relevance Matcher (ITyX)

Hidden Markov Chains

SVM Support Vector Machines

KNN K-Nearest Neighbor

Deep Field Categorizer (Themen)

RegEx

Deep Learning

TCC Topic Clustering Categorizer

*Auf Basis von Computerlinguistik*



PTME Power-Text-Mining-Engine (ITyX)

Neural Mind (ITyX)

Relevance Matcher (ITyX)

Hidden Markov Chains

SVM Support Vector Machines

KNN K-Nearest Neighbor

Deep Field Categorizer (Themen)

RegEx

Deep Learning

TCC Topic Clustering Categorizer

Mustererkennung

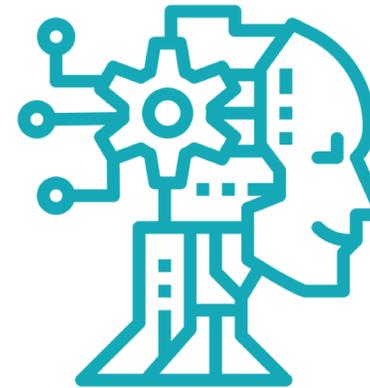
Wortbezüge

Themen

Wortfolgen

Komplexe Regelmäßigkeiten

Worthäufigkeit



ITyX verwendet  
mehrere  
Klassifikations-  
methoden

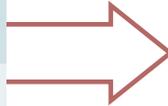
Betreff **wichtige Info**  
 An info@somewhere.local★

Ich ziehe demnächst um,  
 in die Gartenstrasse 15, 51149 Köln. Meine Kundennummer bei Ihnen ist 987654.  
 Dabei ziehe ich mit meinem Freund zusammen.

MfG  
 Elke Breitbach

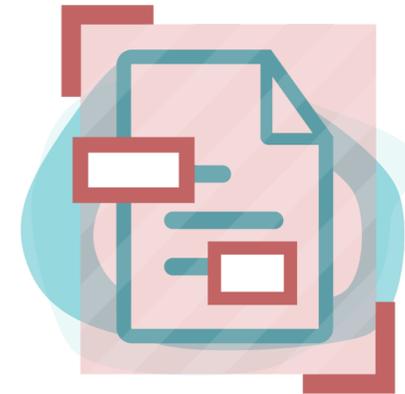
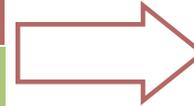
## Basis Trainingsmenge: E-Mail der Kategorie Adressänderung

Kategorisierer	Kategorie
A	Adresse
B	Adresse
C	BLZ
D	Kündigung
E	Adresse
F	BLZ
G	Adresse
H	BLZ



IBOOSTER	
Kategorisierer	Kategorie
A	Adresse
B	Adresse
C	BLZ
D	Kündigung
E	Adresse
F	BLZ
G	Adresse
H	BLZ

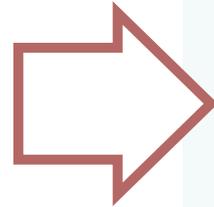
- richtig
- richtig
- falsch
- falsch
- richtig
- falsch
- richtig
- falsch



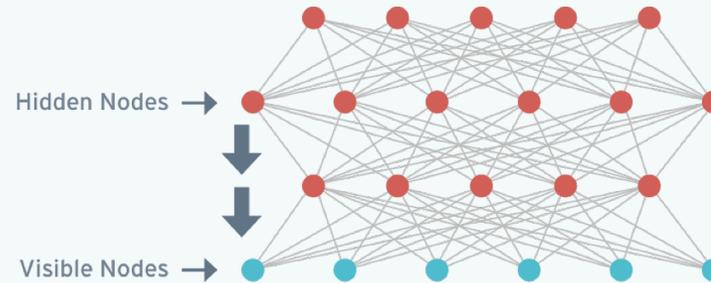
Der ITyX Categorizer



Categorizer	Ergebnisse
PTME Power-Text-Mining-Engine (ITyX)	Kat1
Neural Mind (ITyX)	Kat2
Relevance Matcher (ITyX)	Kat3
Hidden Markov Chains	Kat4
SVM Support Vector Machines	Kat5
KNN K-Nearest Neighbor	Kat6
Deep Field Categorizer (Themen)	Kat7
RegEx	Kat8
Deep Learning	Kat9
TCC Topic Clustering Categorizer	Kat10



## iBooster analysiert

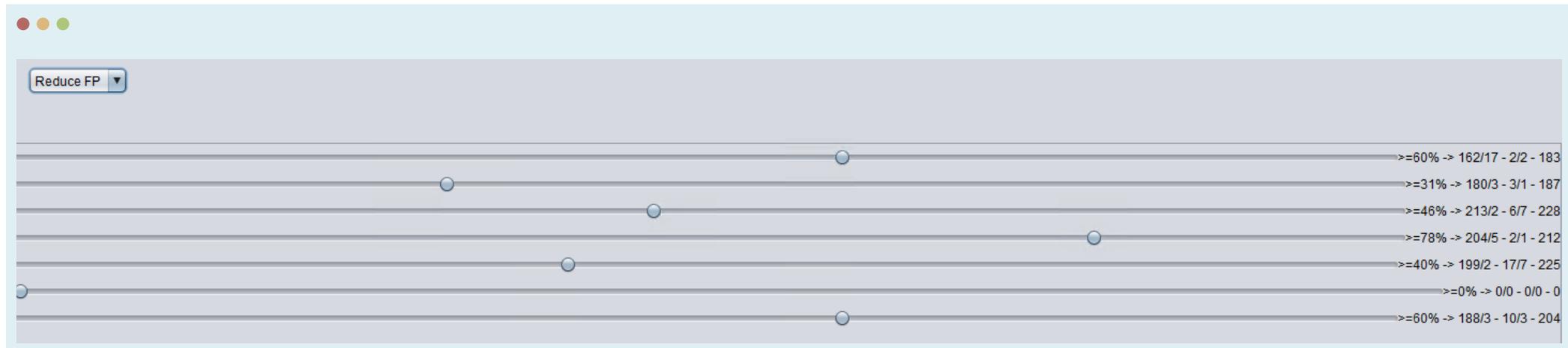


Korrigiert gegebenenfalls die Fehler der Einzel-Categorizer.

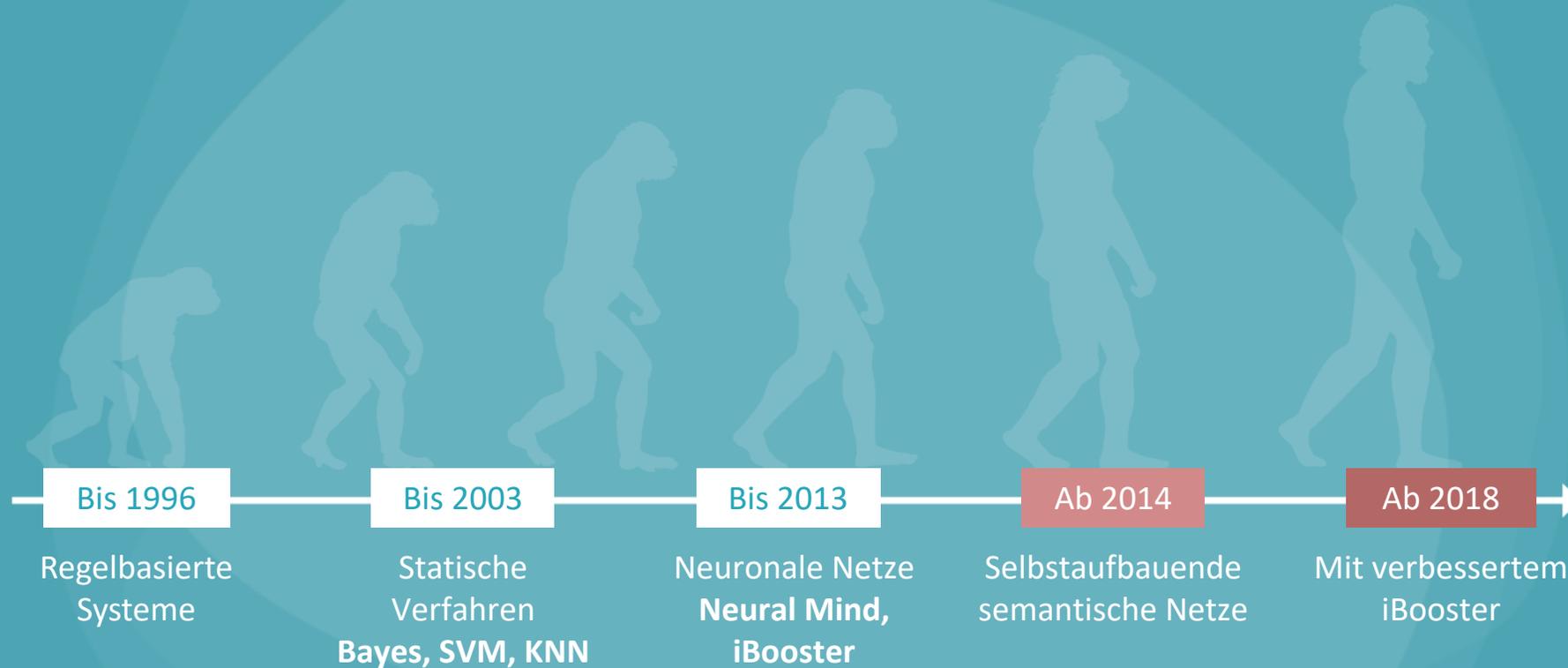
Im Einzelfall ergibt sich durch den iBooster eine andere Kategorie als alle Einzel-Categorizer ermittelt haben.

Durch den iBooster werden die Kategorisierungsergebnisse gegenüber dem Ergebnis eines einzelnen Categorizers erheblich verbessert.

Über die Reliability ermittelt der Categorizer automatisch je Kategorie Grenzwerte die pro Dokument mindestens erreicht werden müssen. Andernfalls werden die Dokumente als „unclassified“ deklariert um „False Positive“ zu minimieren.



# ENTWICKLUNG DER KI DIE EVOLUTION DER METHODEN



## Basisextraktion

- ✓ Suche nach identifizierenden Nummern  
Kundennr, Vertragsnr, etc.
- ✓ Bestandsabgleich
- ✓ Absender ermittelt
- ✓ Eindeutig  
Richtig oder falsch

- Kategorisierung
- Regel-/Validierungsunabhängig

## Qualifizierte Extraktion

- Komplexe Dokumente wie Verträge, Bestellungen, etc.
- Suche nach fachlichen Inhalten, Kaufpreis, Leasingrate, Vertragsbeginn, 3. Parteien (Vermieter, Verkäufer)
- Keine Validierungsmöglichkeit
- Höhere fachliche Qualifikation notwendig

## Mehrere Extraktionsmethoden

Inhaltliche KI (content)

Formate (regelbasiert)

Strukturelle KI (structure)

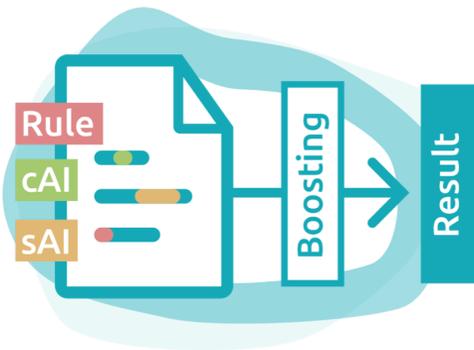
Evaluation



## Boosting

Kombinierendes Boosting

Strukturelles Boosting



Die ITyX  
Extraktion

The screenshot displays the ITyX software interface. On the left, a list of documents is shown with checkboxes for 'Annotated'. The main window shows a document with extracted fields like 'Versand Testfirma GmbH & Co. KG' and 'Mustermanns Firma'. A 'Regex Tag' window is open, showing a 'Regex Definition' and a 'Match Type' table.

Match Type	Title	Test	# TagMatches	# marked true
?	075_rechnungsvorlage-kleinunternehmer.jpg	<input type="checkbox"/>	71	2
?	Beispielrechnung.pdf	<input type="checkbox"/>	45	0
?	Rechnungsvorlage.jpg	<input type="checkbox"/>	80	1
?	muster-rechnung-verschiedene-mehrwertsteuersaetze.jpg	<input type="checkbox"/>	71	1
?	rechnung-ausdruck.jpg	<input type="checkbox"/>	68	1

Below the table, a document preview shows extracted data for 'Versand Testfirma GmbH & Co. KG' and 'Rechnung Nr. RG2009110'. A list of extracted items is shown on the right, including 'Versand Testfirma GmbH & Co. KG', 'Mustermanns Firma', 'professional', 'Co. KG', 'Teststraße 123', 'DEU1235 Testoft', 'pany', 'Meyer Corporation GmbH', 'Meyer Hanz', 'Kunden Nr.', 'Teststraße 27', 'Rechnungsdatum 11.12.2008', '12345 Musterort', 'Kontoinhaber Hanz Meyer', 'Bankname', and 'Deutsche Tes Bank'.

```
Testumgebung/Import/umzuege/Umzuege_Muster.txt</SourceURI>
```

```
<MimeType>text/plain</MimeType>
```

```
<DocumentType/>
```

```
<FormType>gutschrift</FormType>
```

```
<Language/>
```

```
<TagMatch identifier="IBAN">
```

```
<Value>DE98 6725 0020 0003 2046 75</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="BIC">
```

```
<Value>SOLADES1HDB</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="ZaehlerNr">
```

```
<Value>11TR0055230469</Value>
```

```
<Confidence>1.5</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="ZaehlerNr">
```

```
<Value>233014</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="ZaehlerStand">
```

```
<Value>12288,254</Value>
```

```
<Confidence>1.5</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="ZaehlerStand">
```

```
<Value>704</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="Sparte">
```

```
<Value>Strom</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="Sparte">
```

```
<Value>Gas</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="VK">
```

```
<Value>210736562</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
<TagMatch identifier="VK">
```

```
<Value>210693571</Value>
```

```
<Confidence>1.0</Confidence>
```

```
</TagMatch>
```

```
</Tags>
```

```
<Notes>
```

```
<Note key="cats">
```

```
[/gutschrift = 0.4655157239764514,
```

```
/bankdaten = 0.2310654916566026,
```

```
/zaehlerstand = 0.07734208319086064,
```

```
/umzuege = 0.1522225372210828,
```

```
/abschlagszahlung = 0.07385416395500295]
```

```
</Note>
```

<b>Kaufverträge</b>	<ul style="list-style-type: none"> <li>Abteilung 2 <b>83%</b></li> <li>Abteilung 3 <b>94%</b></li> <li>Fälligkeitsmitteilung <b>78%</b></li> </ul>
<b>Internationale Rechnungen</b>	<ul style="list-style-type: none"> <li>PO-number <b>94%</b></li> <li>Invoice amount <b>100%</b></li> <li>Invoice number <b>94%</b></li> </ul>
<b>Deutsche Rechnung</b>	<ul style="list-style-type: none"> <li>Rechnungssteller <b>86%</b></li> <li>Bruttobetrag <b>84%</b></li> <li>MwSt <b>93%</b></li> <li>Rechnungsdatum <b>96%</b></li> </ul>
<b>Kunden Emails</b>	<ul style="list-style-type: none"> <li>Kundennummer <b>97%</b></li> <li>IBAN <b>97%</b></li> </ul>

## Einsparpotential



Gute Basis für  
Ablaufoptimierung