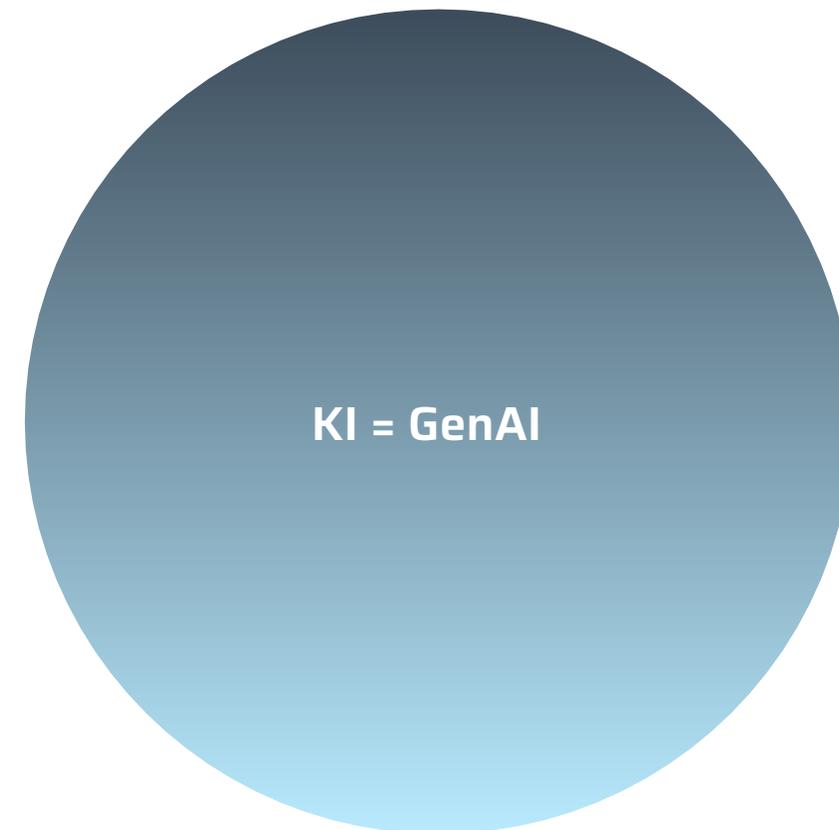
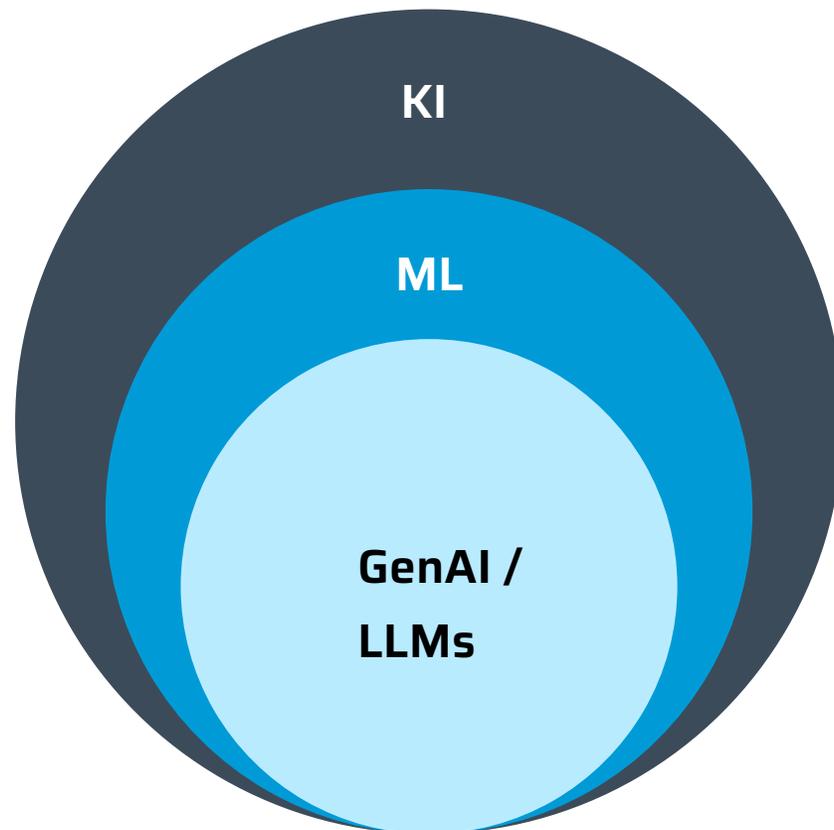




**ON-PREMISES KI FÜR
DATENSCHUTZGERECHTE
AUTOMATISIERUNG IM
DOKUMENTENMANAGEMENT**

Begrifflichkeiten - Realität vs. Wahrnehmung

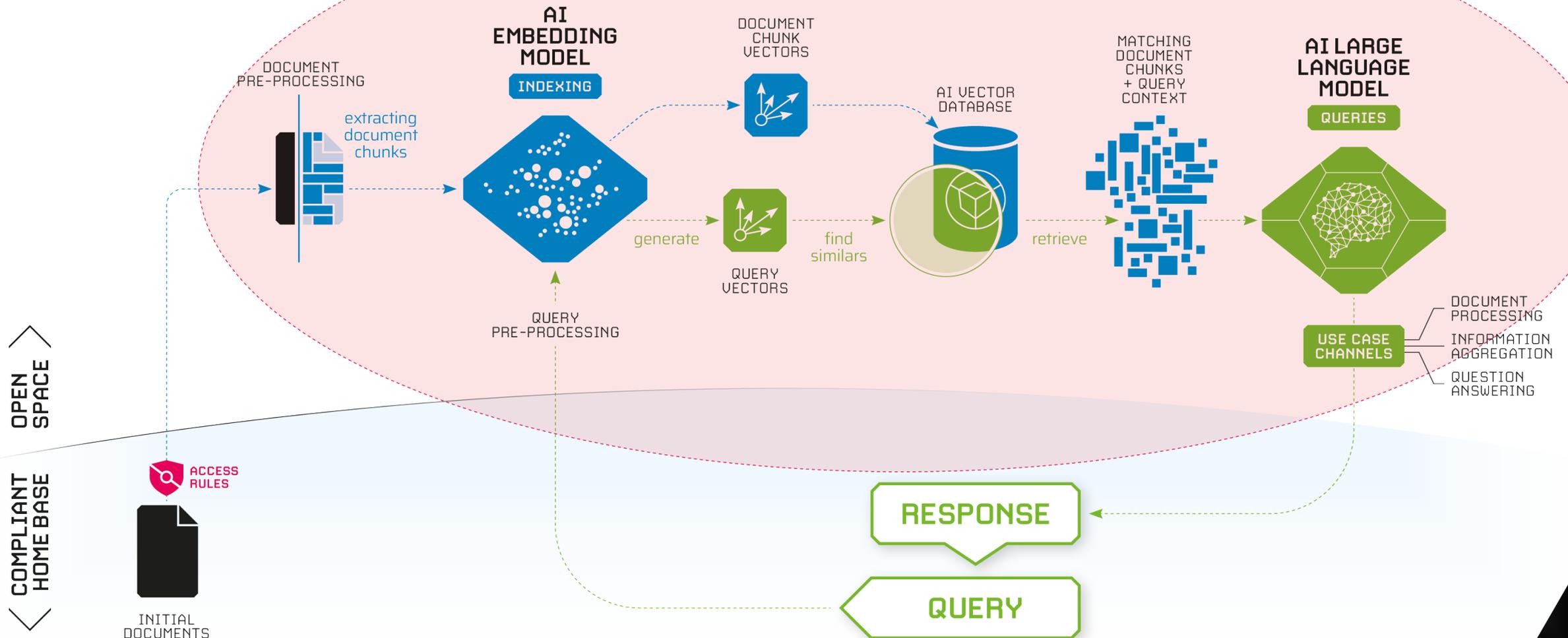


Große Sprachmodelle, große Verantwortung

Herausforderungen beim Einsatz von LLMs

- **Bereitstellung** - LLMs sind sehr ressourcenintensiv, so dass eigene Bereitstellung unwirtschaftlich ist und der Weg in die Cloud führt
 - **Kosten** - schlecht kalkulierbare Kosten / tokenbasierte Abrechnung
 - **Datenschutz** - Herausgabe von Daten an Dritte
- **Unklarheiten der Trainingsdaten** - Copyright und Bias
- **Problem-Solution Fit** - Ressourcen werden in eine Lösung investiert, die nicht das eigentliche Problem adressiert

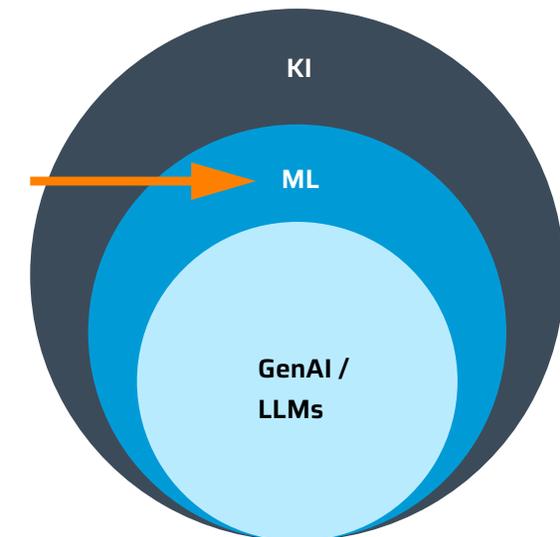
Beispiel: Aufbau einer RAG ChatBot Pipeline in der Cloud



Prozessautomatisierung abseits von LLMs

Automatisierung statt Chatten

- Nutzung etablierter Verfahren des maschinellen Lernens
- „Natural language processing“ - NLP
- Entwicklung von spezialisierten Modellen, die auf jeweiligen Anwendungsfall zugeschnitten sind
- Modelle werden auf unternehmenseigenen Daten trainiert
- Modelle sind klein und können on-premises gehostet werden
- Daten bleiben im Unternehmen



+ Sprachen-
erkennung

+ Anonymisierung
und Maskierung

+ Indizierung für
semantische Suche

+ Dokumenten-
klassifikation

+ Verschlag-
wortung

+ Inhaltsextraktion und
Vollständigkeitsprüfung

**Beispiele für
spezialisierte
NLP-basierte
Dokumentenlösungen**



BEISPIEL: DOKUMENTEN KLASSIFIKATION

Automatische Erkennung von Dokumententypen



- Manuelle Sortierung und Verarbeitung großer Dokumentenmengen aufwendig und ineffizient
- Dokumententypen-abhängige Zuweisung an verschiedene Nutzer oder Abteilungen möglich
- Dokumententypen-abhängige Folgeverarbeitung möglich, z.B. Inhaltsextraktion

Erkennung und Kategorisierung von Dokumententypen



Papierwaren Schulz
Auf dem Damm 49 | 48399 Erkeath

Herr
Konstantin Wlster
Neckarauer Str. 6
23456 Bremen

Papierwaren Schulz
Auf dem Damm 49
46789 Erkeath
Tel: 0123 5678-90
hallo@superdupersseite.de
www.superdupersseite.de

Lieferschein zu Bestellung #123-456-789

Rechnungsnummer: 1234 Kundennummer: 1234 Datum: 18.11.2030

Pos	Bezeichnung	Einzelpreis	Anzahl	Gesamtpreis
1	Buch "Supertolles Buch"	16,20 €	1	16,20 €
2	Buntstiftset Kunterbunt	23,50 €	1	23,50 €
3	Kunstpostkarte	4,00 €	1	4,00 €
Summe der Nettobeträge				43,70 €
zuzüglich 19% Umsatzsteuer				8,30 €
Gesamtbetrag (bereits bezahlt: Kreditkarte *0123)				52,00 €

Vielen Dank für Ihre Bestellung!



Paul Cheung

Paul Cheung | Heerhofstr. 8 | 52445 Freilberg

Paul Cheung
Meyerhofer-Allee 8
52445 Freilberg

Tel: 0123 5678-90
hallo@superdupersseite.de
www.superdupersseite.de

Angebot

Angebotsnummer: 1234 Kundennummer: 1234 Datum: 29.07.2030

Sehr geehrte Frau Hobner,
vielen Dank für Ihr Interesse. Gerne unterbreite ich Ihnen folgendes Angebot:

Pos	Bezeichnung	Einzelpreis	Anzahl	Gesamtpreis
1	Modellentwicklung 2D	2.500,00 €	1	2.500,00 €
2	Modellentwicklung 3D	4.000,00 €	1	4.000,00 €
Summe der Nettobeträge				6.500,00 €
zuzüglich 19% Umsatzsteuer				1.235,00 €
Gesamtbetrag				7.735,00 €



logoipsum

Max Müller
Firmenstraße 1
45678 Muster

Marta Muster
+49 123 456789
mail@example.com

Rechnungsnummer: 1234
Rechnungsdatum: 01.07.2020
Zahlbar bis: 14.07.2020

Rechnung

Sehr geehrter Herr Müller,
Vielen Dank für Ihren Auftrag und Ihr Vertrauen.
Für unsere Arbeit und Material stellen wir Ihnen folgende Summe in Rechnung:

Pos	Artikel	Anzahl	Einzelpreis	Summe netto
1	Beratung Website Neugestaltung Analyse der bestehenden Website und Aufschreibung der notwendigen Änderungen.	4	80,00 €	320,00 €
2	Entwicklung Website example.com	30	80,00 €	2400,00 €
3	Dokumentation	3	80,00 €	240,00 €
4	Projekübergabe	1	110,00 €	110,00 €
Summe netto				3070,00 €
USt 19%				583,30 €
Rechnungssumme				3653,30 €

Ich danke Ihnen für die gute Zusammenarbeit.
Mit freundlichen Grüßen
Marta Muster

Seite 1 von 1

Erkennung und Kategorisierung von Dokumententypen



Papierwaren Schulz
Auf dem Damm 49 | 48399 Erkrath

Herr
Konstantin Wlster
Neckarauer Str. 6
23456 Bremen

Papierwaren Schulz
Auf dem Damm 49
46789 Erkrath
Tel: 0123 5678-90
hallo@superdupersseite.de
www.superdupersseite.de

Lieferschein zu Bestellung #123-456-789

Rechnungsnummer: 1234 Kundennummer: 1234 Datum: 18.11.2030

Pos	Bezeichnung	Einzelpreis	Anzahl	Gesamtpreis
1	Buch "Supertolles Buch"	16,20 €	1	16,20 €
2	Buntstiftset Kunterbunt	23,50 €	1	23,50 €
3	Kunstpostkarte	4,00 €	1	4,00 €

Summe der Nettobeträge 43,70 €
zuzüglich 19% Umsatzsteuer 8,30 €

Gesamtbetrag (bereits bezahlt: Kreditkarte *0123) 52,00 €

Vielen Dank für Ihre Bestellung!

» Lieferschein «



Paul Cheung

Architekturbüro Eichholzer
Frau Ma Höbner
Zwinglistr. 67
54321 Münster

Paul Cheung
Mayerhofer-Allee 8
52455 Freilberg

Tel: 0123 5678-90
hallo@superdupersseite.de
www.superdupersseite.de

Angebot

Angebotsnummer: 1234 Kundennummer: 1234 Datum: 20.07.2030

Sehr geehrte Frau Höbner,
vielen Dank für Ihr Interesse. Gerne unterbreite ich Ihnen folgendes Angebot:

Pos	Bezeichnung	Einzelpreis	Anzahl	Gesamtpreis
1	Modellentwicklung 2D	2.500,00 €	1	2.500,00 €
2	Modellentwicklung 3D	4.000,00 €	1	4.000,00 €

Summe der Nettobeträge 6.500,00 €
zuzüglich 19% Umsatzsteuer 1.235,00 €

Gesamtbetrag 7.735,00 €

» Angebot «



logoipsum

Max Müller
Firmenstraße 1
45678 Muster

Marta Muster
+49 123 456789
mail@example.com

Rechnungsnummer: 1234
Rechnungsdatum: 01.07.2020
Zahlbar bis: 14.07.2020

Rechnung

Sehr geehrter Herr Müller,
Vielen Dank für Ihren Auftrag und Ihr Vertrauen.
Für unsere Arbeit und Material stellen wir Ihnen folgende Summe in Rechnung:

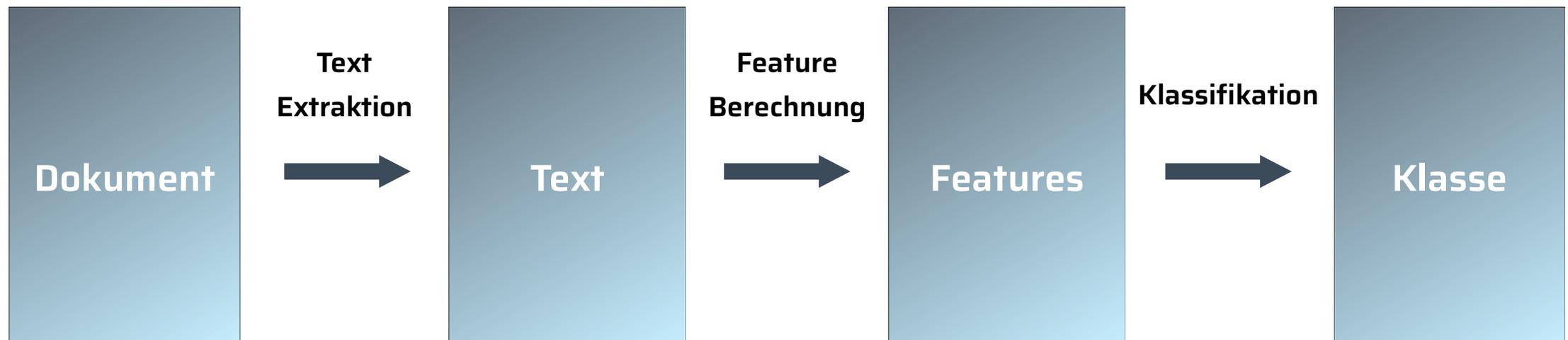
Pos	Artikel	Anzahl	Einzelpreis	Summe netto
1	Beratung Website Neugestaltung Analyse der bestehenden Website und Aufschreibung der notwendigen Änderungen.	4	80,00 €	320,00 €
2	Entwicklung Website example.com	30	80,00 €	2400,00 €
3	Dokumentation	3	80,00 €	240,00 €
4	Projekübergabe	1	110,00 €	110,00 €

Summe netto 3070,00 €
USt 19% 583,30 €
Rechnungssumme 3653,30 €

Ich danke Ihnen für die gute Zusammenarbeit.
Mit freundlichen Grüßen
Marta Muster

» Rechnung «

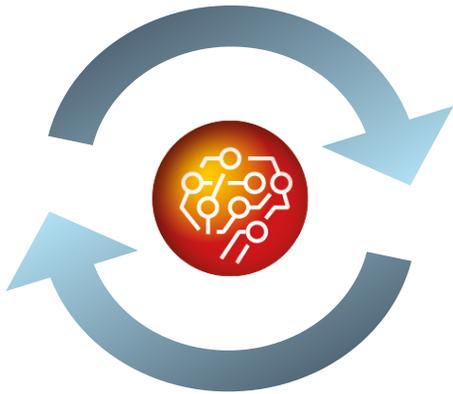
Wie wird der Dokumententyp ermittelt?





BEISPIEL: RETRAINING PIPELINES

Mitlernende KI

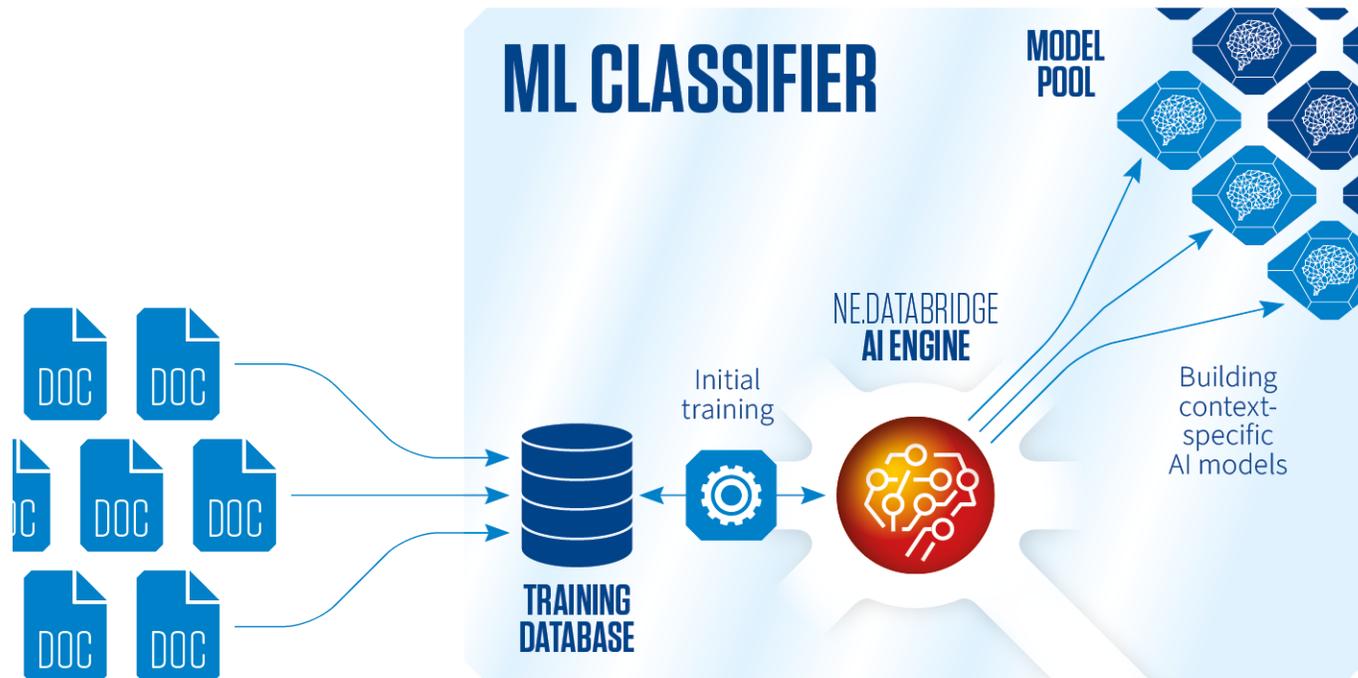


- Form und Inhalt von Dokumenttypen ändern sich
- Neue Dokumententypen kommen dazu
- Integration von Nutzerfeedback auf Basis der Konfidenz
- Bessere Nachvollziehbarkeit und höhere Akzeptanz durch Nutzer

Retraining Pipeline am Beispiel eines Classifiers



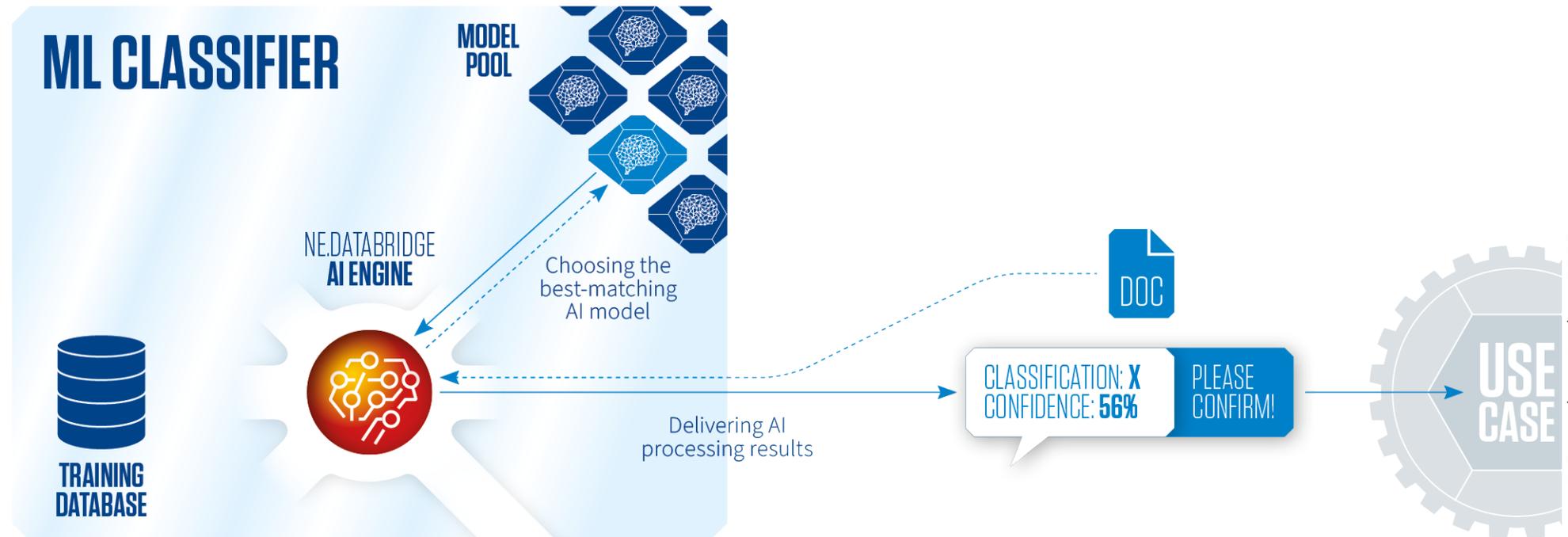
Initialtraining mit unternehmenseigenen Dokumenten



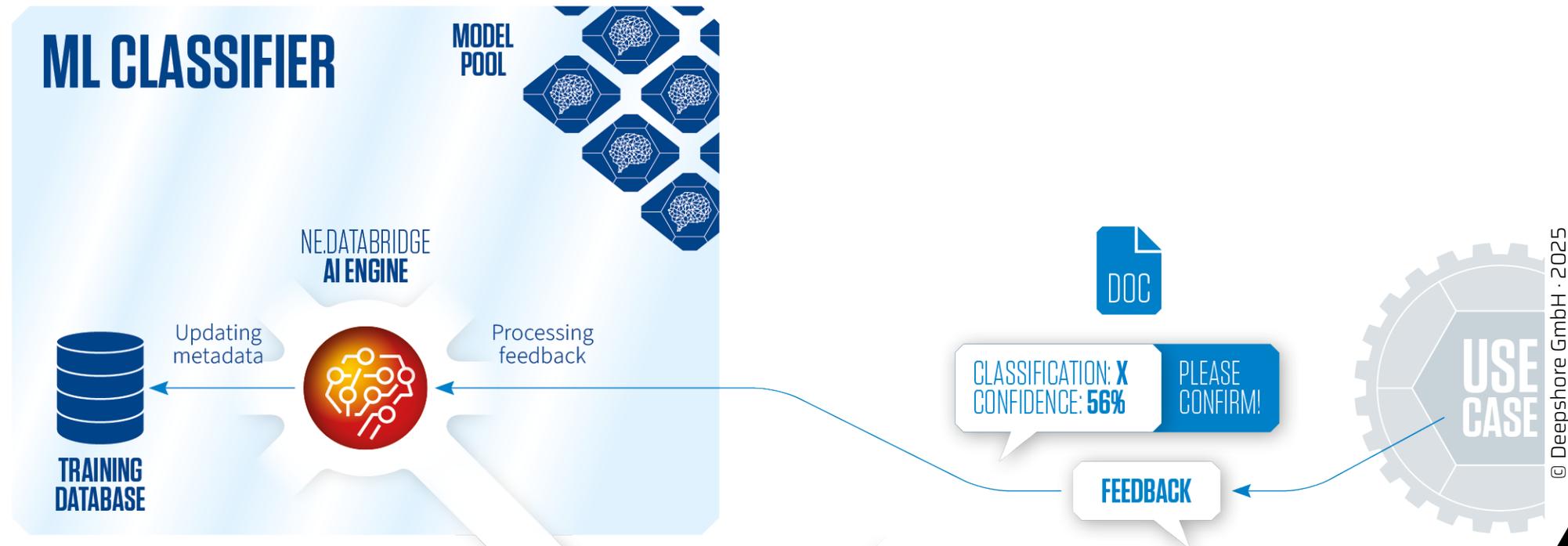
Importieren eines neuen Dokuments



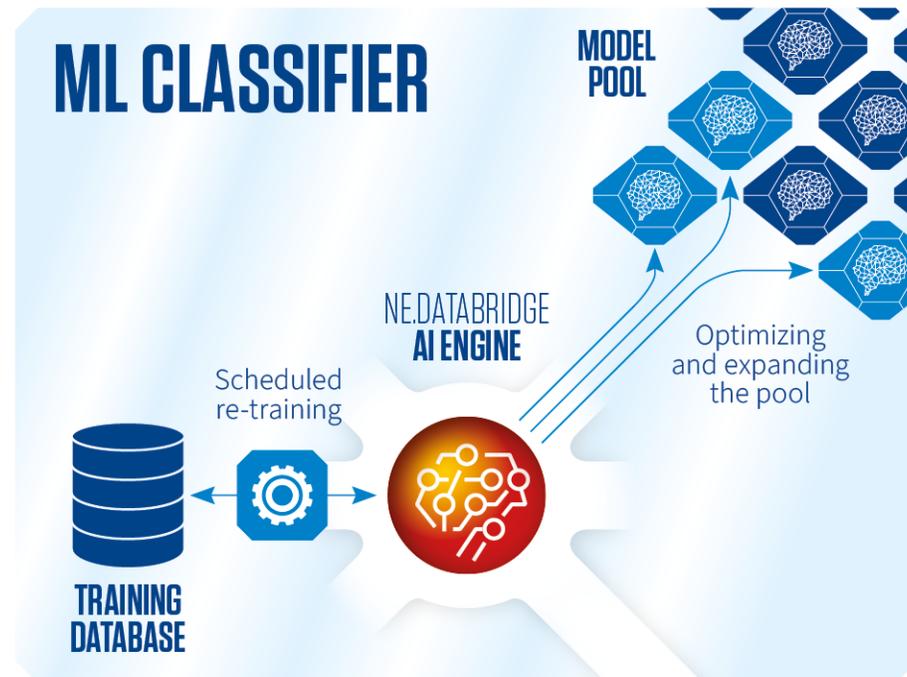
Modellauswahl und Klassifikation



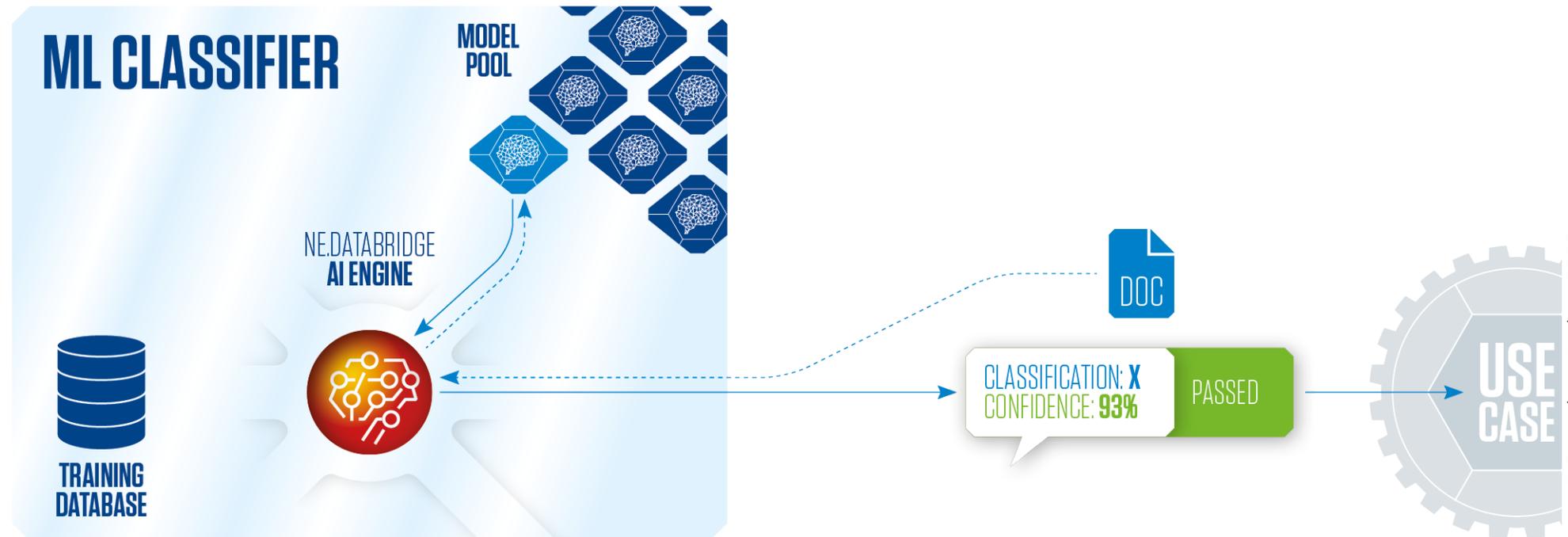
Feedback und Verarbeitung



Retraining und Modellpool-Erweiterung



Gesteigerte Konfidenz nach Retraining



VORTEILE EIGENER, ON PREMISES GEHOSTETER NLP MODELLE

Volle Compliance nach europäischen Standards

Erhöhte Transparenz
Klare und nachvollziehbare
Aufbereitung der
Ergebnisse

Erhöhte Automatisierung
Einleitung von bestimmten
Aktionen je nach ML-
basierter Entscheidung

Verbesserte Datenqualität
Standardisierte und
konsistente Ergebnisse

**Strategische
Unabhängigkeit von
Cloud-Anbietern**

Deepshore GmbH

Van-der-Smissen-Straße 9
22767 Hamburg

Tel +49 40 46664 296

Fax +49 40 46664 299

E-Mail info@deepshore.de

www.deepshore.de