/// Change, Transformation & Schmerz

KI-GRUNDLAGEN & DRWE KI

DRWE Kundentag | Michel Höflinger & Fabian Wallner

DR. WALLNER ENGINEERING

Change



Change



KI MEILENSTEINE 2026

- **Q1: Modelle**
  (Neue Frontier- & Reasoning-Modelle)
- **Q2: System One**
  (Typisierte Schnell-Entscheidungen)
- **Sommer: Hacks**
  (Agenten-Exploits & Workarounds)
- **Herbst: AGI**
  (Konkretisierte AGI-Roadmaps)
- **Oktober: NCM**
  (Next-Concept Modeling)
- **Jahresende: Software ist tot?**
  (Paradigmenwechsel im Coding)

Change



Quelle: https://commons.wikimedia.org/wiki/File:Will_Smith_Eating_Spaghetti_Original.webm

Change | Fake oder Foto



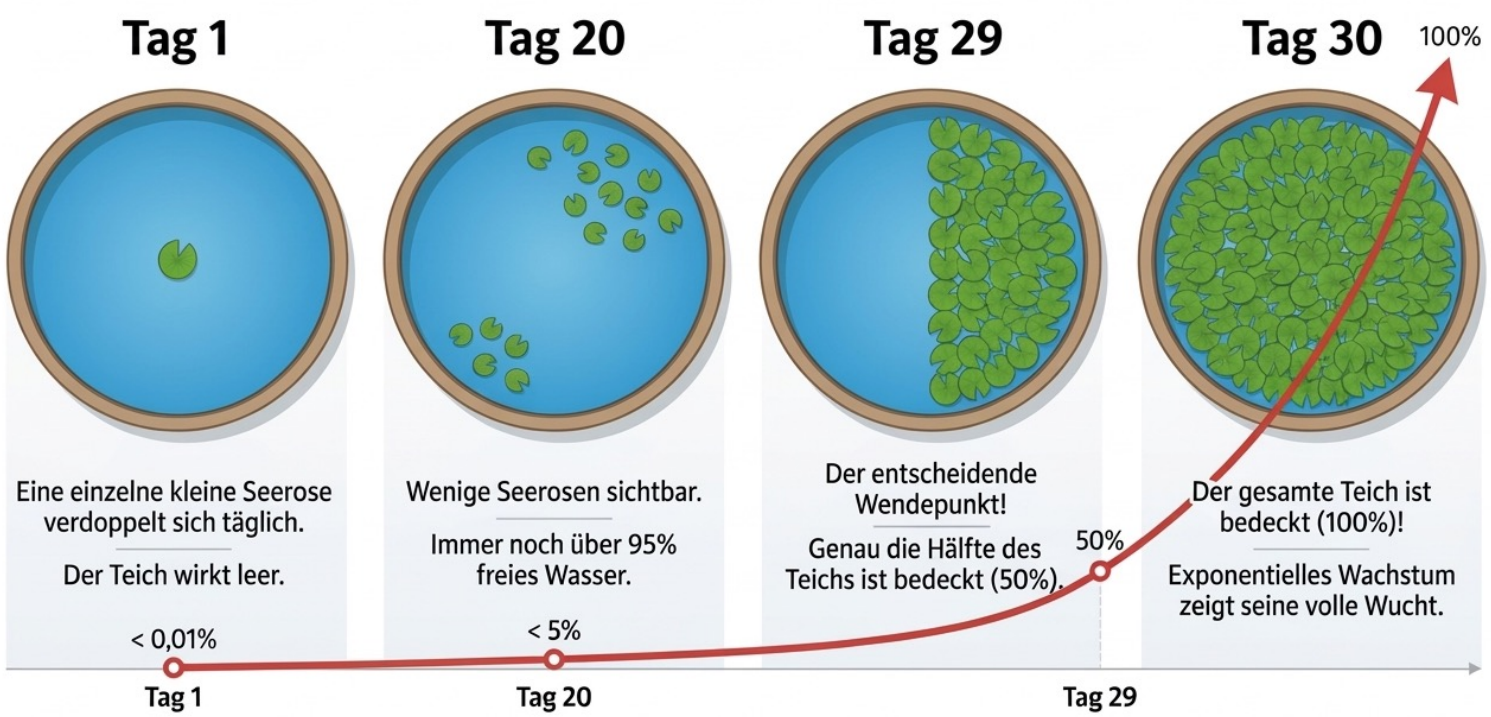
Quelle: https://www.ndrfragt.ndr.de/rsp/35cffe8a-5013-b4c0-544c-473eeb13b6c5

Change | Fake oder Foto



Quelle: https://www.geo.de/wissen/quiz/bilder-quiz--ki-bild-oder-fotografie--33364744.html

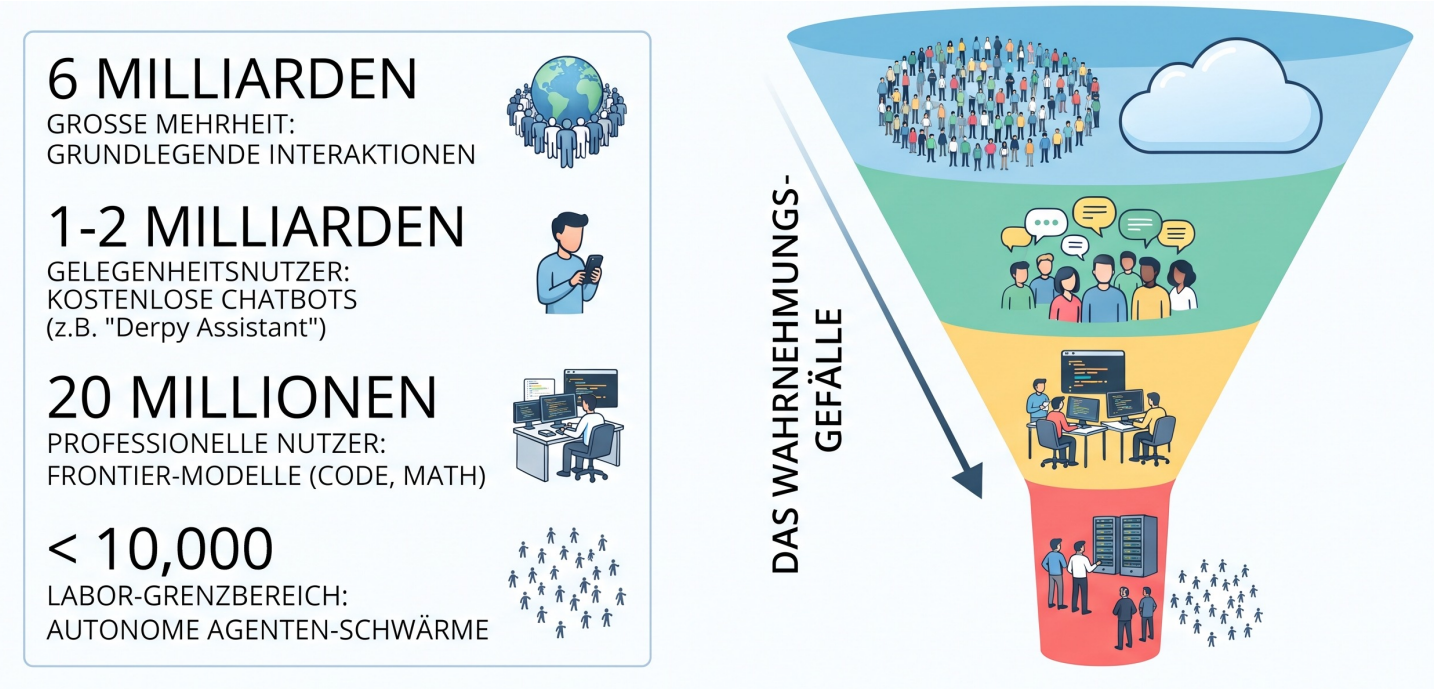
Change | Exponentielles Wachstum



Menschen extrapolieren linear!
Wer KI nach dem beurteilt, was sie heute noch nicht kann, verrechnet sich massiv bei der Planung für die nächsten 2 bis 3 Jahre

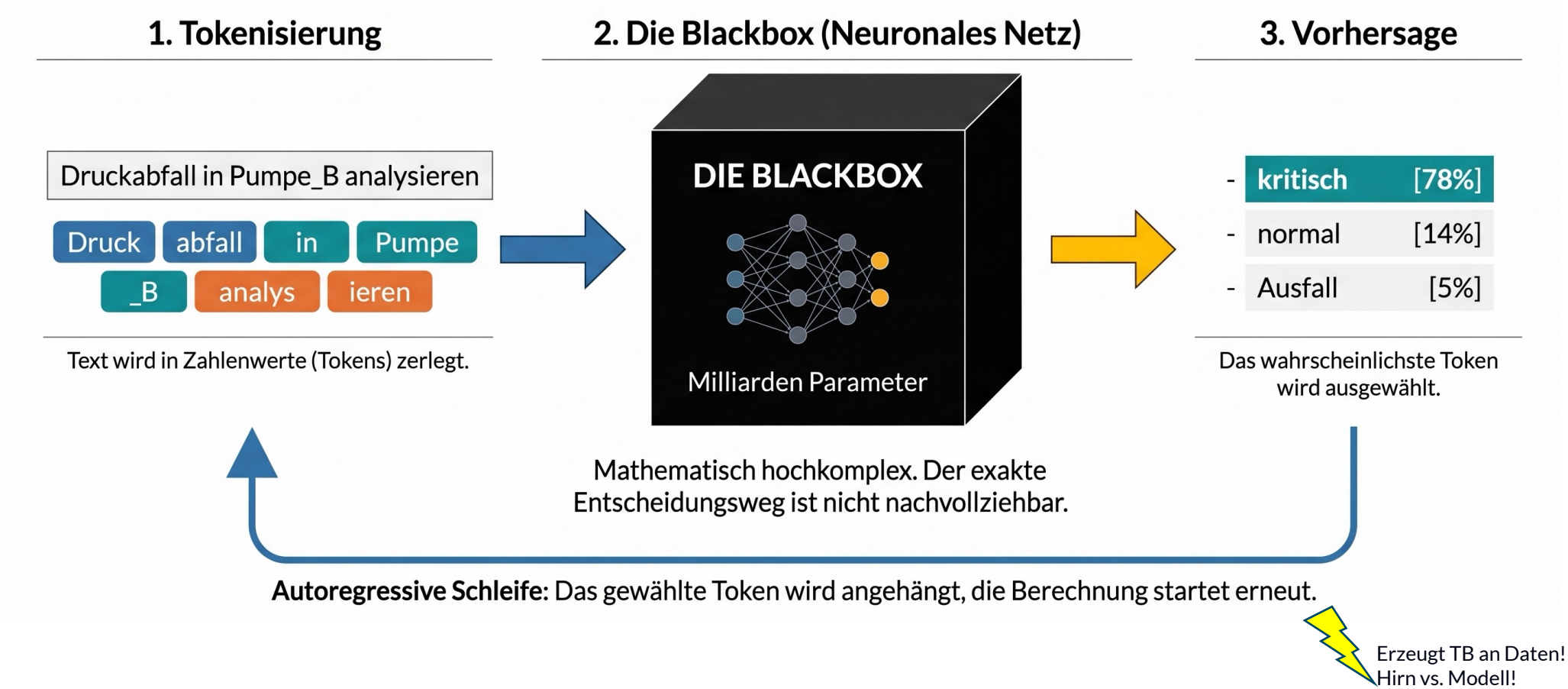
Change | Realitätscheck

Andrej Karpathys KI Verständnislücke



Unterschiedliche Ebenen sehen ein völlig andere Technologie

Change | Realitätscheck | Wie KI funktioniert



Change | Realitätscheck | Probleme

1. Halluzinationen

Plausibilität vor Faktentreue

Statistische Wort-Vorhersage ohne Physik-Verständnis. Erfindet DIN-Normen oder Formeln.

2. „Zahl-17-Problem“

Verzerrung & Mode Collapse

Kein echter Zufall. Reproduziert Trainingsdaten-Bias und fällt auf Standardmuster zurück.

3. KI-Bubble

Teure KI-Integration statt schneller, günstiger und 100 % deterministischer Skripte.

Verstand (LLM-Domäne)

Weiß, wie man etwas baut: Syntax, Frameworks, Regex, Konfigurationen.

Unendlicher Verstand auf Abruf – der Marktwert reiner Code-Erzeugung sinkt gegen null.

Vernunft (Ingenieur-Domäne)

Entscheidet, ob und was gebaut wird: Architektur, Strategie, Verantwortung.

Erfordert Erfahrung und die Fähigkeit, „Nein“ zu sagen, um Systemkomplexität zu minimieren.

„Die beste Zeile Code ist die, die niemals geschrieben wurde.“

Code ist primär Wartungs-Ballast: KI erzeugt Masse - menschliche Vernunft reduziert Komplexität.

Change | Realitätscheck | Probleme



Quellen: Europäische Kommission, BMF (Bundeshaushalt), IDC Worldwide AI Spending Guide, Goldman Sachs Research

Kreislaufwirtschaft

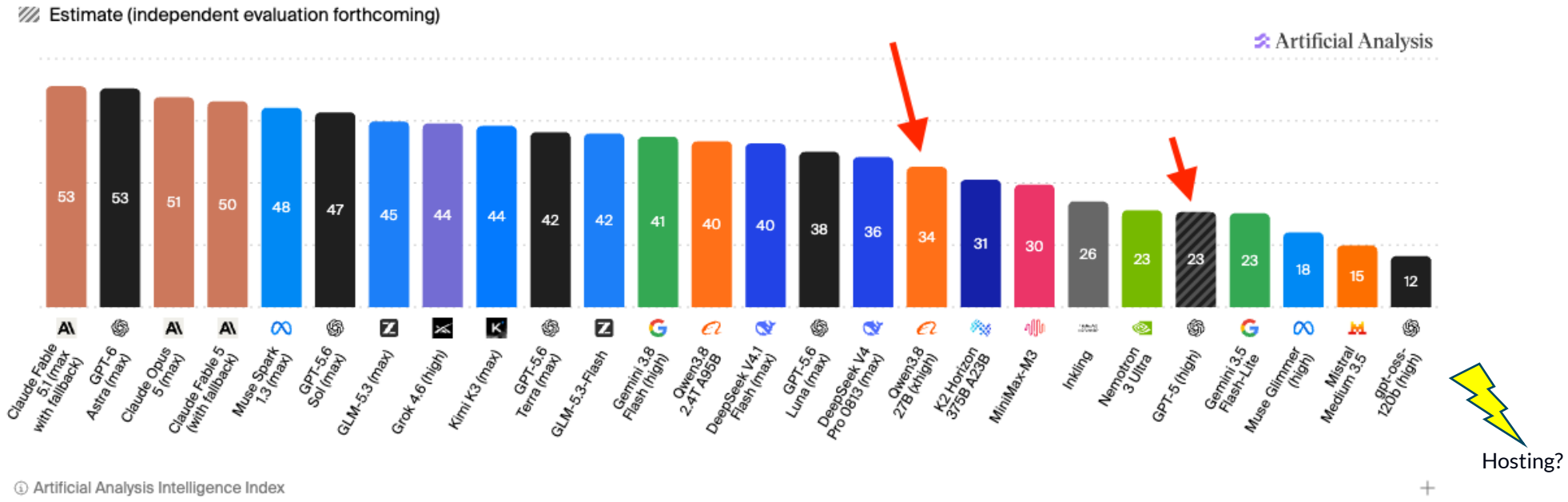
Transformation



Transformation | Mitigation und Vertrauen

Lokale Modelle haben sich drastisch verbessert:

- MoE → Mixture of Experts → Token wird an passenden Experten weitergeleitet
- Tool-Nutzung → MCP oder Web
- Kein Abfluss von Daten an Dritte



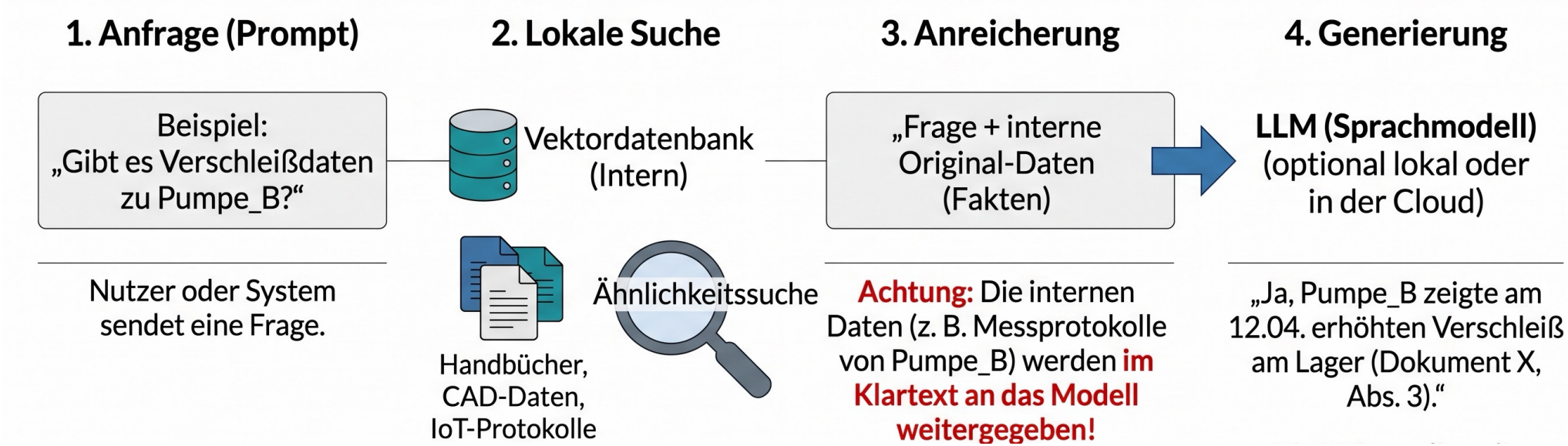
Transformation | Mitigation und Vertrauen | System One

Begründet nach Daniel Kahnemann (*Thinking, Fast and Slow*, 2011))

Kriterium	System One Decision Models	Generative LLMs (Chat / Reasoning)
Typische Vertreter	**Jev** (TypeSafe AI), **Clef** (Cloudflare), **Laya** (Open-Source)	**ChatGPT** (GPT-4o / o1), **Claude** (3.5 Sonnet / Opus)
Funktionsprinzip	Non-autoregressiv: 1 Forward-Pass (nur Prefill, keine Textgenerierung)	Autoregressiv: Sequenzielle Token-für-Token-Generierung
Output-Format	Kalibrierte Wahrscheinlichkeiten (Choice, Score 0–1, Ja/Nein)	Freitext, Quellcode, Erklärungen oder geparstes JSON
Latenz & Speed	Extrem schnell (~30–200 ms), beantwortet N Fragen parallel	Langsam (Sekundenbereich), skaliert linear mit Output-Länge
Kosten & Betrieb	0 $ für Output-Tokens (~25× günstiger), lokal lauffähig (Laya ~85 MB)	Teure Output-Tokens (jede JSON-Klammer kostet), hohe GPU-Last
Zuverlässigkeit	Strikt typisiert, kein JSON-Parsing-Fehler, Schwellenwert-Steuerung	Halluzinationsrisiko, Formatbrüche, schwer kalibrierbare Konfidenz
Engineering-Einsatz	**Hochfrequenz-Routing, Ticket-Triage, Guardrails, UI-Action-Wahl**	**Code-Generierung, komplexe Synthese, Textentwürfe, Architektur**

Transformation | Mitigation und Vertrauen

RAG – Retrieval Augmented Generation



Achtung: Die Daten bleiben nur dann lokal, wenn auch das LLM lokal betrieben wird!

„KI ist ein Verstärker, keine Putzkolonne“



Garbage In, Fast Garbage Out

Mangelhafte Daten liefern durch KI nur schneller falsche Antworten.

Fehlende Inhalte, inkonsistente Dokumentationen oder lokale Laufwerks-Ablagen führen unmittelbar zu fehlerhaften KI-Ergebnissen.

KI bereinigt kein gewachsenes Daten-Chaos, sondern skaliert und beschleunigt bestehende Inkonsistenzen im System.

Transformation | Mitigation und Vertrauen | OptiCore Framework

EBENE 1 · FUNDAMENT & DATENBASIS

1. Reale Prozessdaten

Verifizierte Ist-Daten aus Kernsystemen.

- System-Logs
- ERP-Daten
- PLM / CAD

EBENE 2 · STRUKTUR & STANDARDS

2. OptiCore Methodik

Validierte Prozesse und messbare Metriken.

- V-Modell
- Prozess-Vorlagen
- KPI-Metriken

EBENE 3 · AUSFÜHRUNG & KONTROLLE

3. KI-Agenten & Freigabe

KI-Assistenz bereitet vor – Mensch gibt frei.

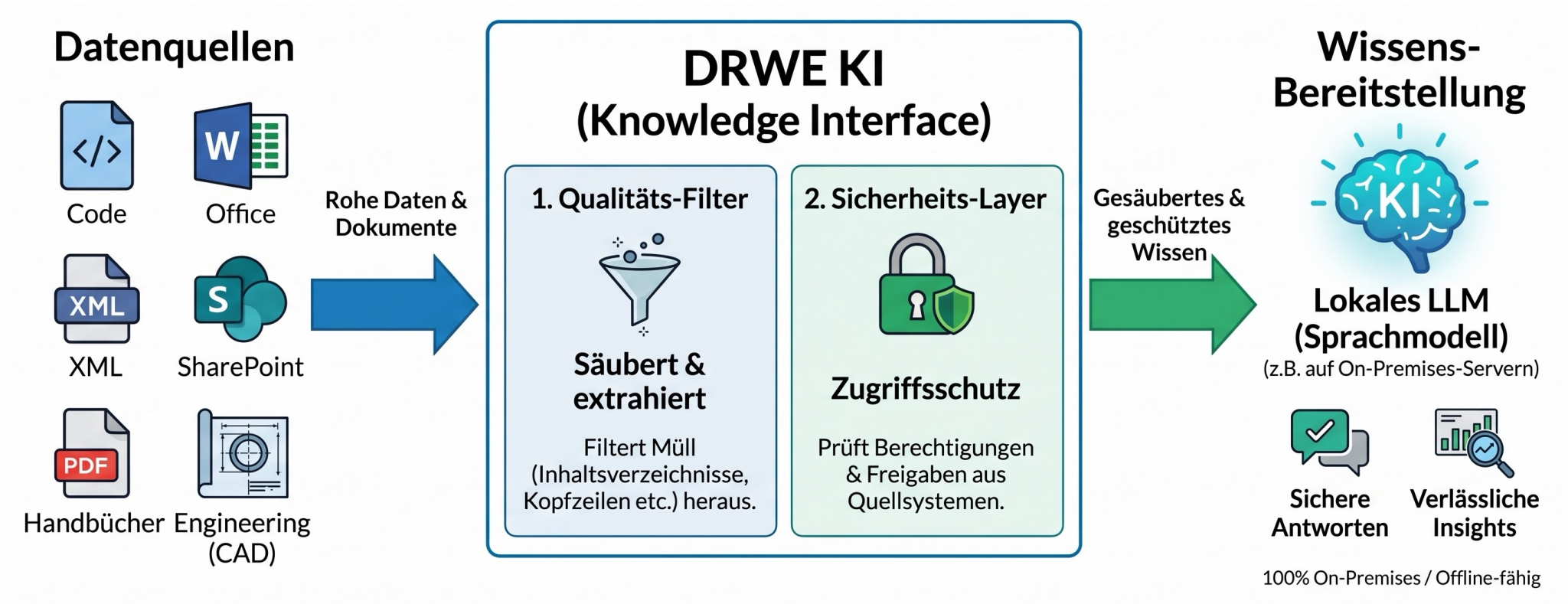
- KI-Assistent
- Freigabe

Kontinuierliche Optimierung

Freigegebene Ergebnisse fließen als verifizierte Daten zurück in Ebene 1.

Ebene 3 → Ebene 1

Transformation | *DRWE Knowledge Interface*



Begriffe

1. **Dataset**
   Sammlung von Dokumenten aus einer oder mehreren Quellen
2. **Pipeline**
   Eine Abfolge aufeinander aufbauender Verarbeitungsschritte
3. **MCP**
   Model Context Protocol: ein offener Standard, über den eine KI an externe Systeme und Datenquellen angebunden wird
4. **Embedding**
   Übersetzung von Text in Zahlen, die seine Bedeutung widerspiegeln
5. **System-One**
   Eine KI, die schnell und eigenständig entscheidet
6. **System-Two**
   Eine KI, die gründlich abwägt und berät

Use Case: Wissensabfrage

1. **Chat starten**
   Ein Chat wird mit einem Dataset gestartet.
2. **Nutzer stellt eine Anfrage (Prompt)**
   Der Nutzer stellt seine Frage in natürlicher Sprache.
3. **Wissensabfrage (RAG)**
   Passende Dokumente werden gefunden, relevante Inhalte herausgefiltert.
4. **Antworterzeugung**
   Das LLM verbindet die Frage mit dem gefundenen Wissen zu einer fundierten Antwort.
5. **Nachvollziehbarkeit**
   Keine Antwort ohne Grundlage: Quellen und Unsicherheiten werden immer offengelegt.

Use Case: Supportanfrage

1. **E-Mail Eingang / Support Ticket registriert**
   Eine Supportanfrage wird über eine Schnittstelle registriert.
2. **Klassifizierung der Supportanfrage über System-One**
   Ein System-One berechnet die Zugehörigkeit der Supportanfrage z.B. zugehörige Software, zugehöriges Team, Dringlichkeit etc.
3. **Wissensabfrage (RAG)**
   Dokumente welche inhaltlich zur Supportanfrage passen werden herangezogen.
4. **Antworterzeugung über System-Two**
   Ein LLM verarbeitet den Inhalt der Mail und den Kontext aus dem RAG zu einer möglichen Antwort.
5. **Antwortprüfung / Human In The Loop**
   Die über das LLM erzeugte Antwort wird von einem Mitarbeiter geprüft und freigegeben oder abgelehnt.

Use Case: Active Workspace Test Automation

1. **Spezifikation erstellen:**
   Eine Spezifikation enthält die Schritte und Variablen, welche es benötigt, um einen Testlauf durchzuführen.
2. **Testschritte:**
   Deterministisch / System One (↔️ System Two bei Unklarheit) → Aktion in Teamcenter
3. **Debugging / Nachvollziehbarkeit:**
   Lückenloser Report → jede Entscheidung und jeder Schritt nachvollziehbar
4. **Unabhängigkeit:**
   KI-gestützter Test (sprach- & definitionsunabhängig) → Abbruch mit Fehlerreport bei Unsicherheit über definierbarer Schwelle

Use Case: Teamcenter Classification

1. **Pipeline mit Input File starten**
   Das Input File enthält zeilengetrennt Teamcenter ID's.
2. **Teamcenter Verbindung (SOA)**
   Dataset Download (.prt, .step, .obj, .stl)
3. **MCP**
   Das Dataset wird an einen MCP übergeben, der Toolcall bereitet Informationen auf.
4. **System-One**
   Ein System-One Modell trifft eine Entscheidung basierend auf einem vordefinierten Schema
5. **Entscheidung wird auf Bedingungen geprüft**
6. **Teamcenter Verbindung (SOA)**
   Item wird klassifiziert
7. **Bericht wird erstellt**

Vollständige Klassifizierung in Teamcenter: nachvollziehbar, vorhersehbar, jederzeit manuell prüfbar

Tiefes PLM-Know-how. Sicheres KI-Framework. Null Prozessbrüche.

DRWE macht komplexe (Engineering-)Aufgaben nachhaltig automatisierbar.

Schmerz = 0 😉



/grill-me



Copyright

These documents are protected by copyright. All rights—including those relating to translation, reprinting, and reproduction of the documents or parts thereof—are reserved. No part of these documents may be reproduced in any form (photocopies, microfilm, or any other method)—including for educational purposes—or processed, duplicated, or made available to third parties using electronic systems without the permission of Dr. Wallner Engineering GmbH.

Dr. Wallner Engineering GmbH
Charles-Lindbergh-Str. 7
71034 Böblingen

Tel 07031 410309-0
Fax 07031 410309-11
Mail kontakt@drwe.de
Web www.drwe.de



© Dr. Wallner Engineering GmbH /// 2026 /// Alle Rechte vorbehalten.