

# Zwischen Hype und heiligen Kühen

---

## Künstliche Intelligenz an Bibliotheken am Beispiel der Inhaltserschließung

*Dr. Anna Kasprzik,  
ZBW – Leibniz-Informationszentrum Wirtschaft  
Vortrag für die DINI-KIM-AG, online, 15.04.2024*

# Alter Wein ...

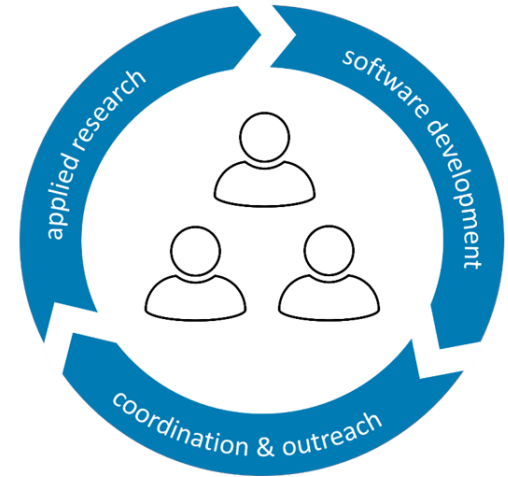
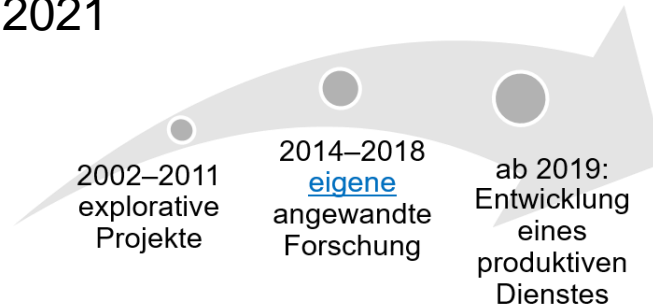
---

- Ideen zu Künstlicher Intelligenz sind nicht neu – sie sind Jahrzehnte oder sogar Jahrhunderte alt
- in Bibliotheken sind Automatisierungsbestrebungen rund um Metadaten ebenfalls nicht neu: quasi seit Aufkommen der EDV, verstärkt ab 2000
  - bis ca. 2012: einige Projekte\* mit computerlinguistischen Verfahren für die Inhaltserschließung – **keine Machine-Learning-Verfahren**
  - ab ca. 2012: implementierte **Machine-Learning-Methoden**, auch und gerade Open Source, werden **BENUTZBAR**
    - technisch und (zu diskutieren) in Bezug auf den Output



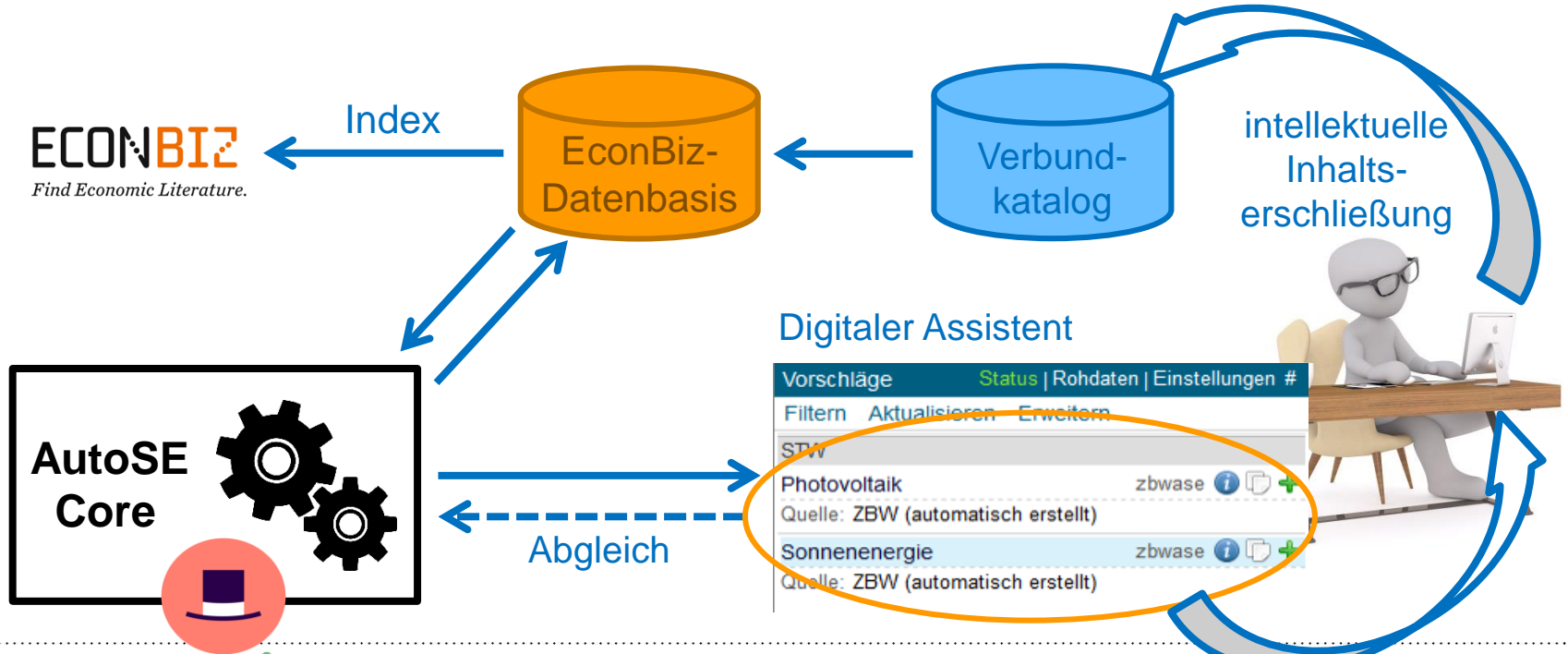
# AutoSE: Der Weg zum verstetigten Forschungstransfer

- Vorläufer: Projekt AutoIndex (bis 2018) – angewandte Forschung & Prototypen
- ab 2019: **AutoSE** („Automatisierung der Sacherschließung“)
- **produktiver Dienst** seit 2021



 Meilenstein „**Statusänderung vom Projekt zur Daueraufgabe**“: 

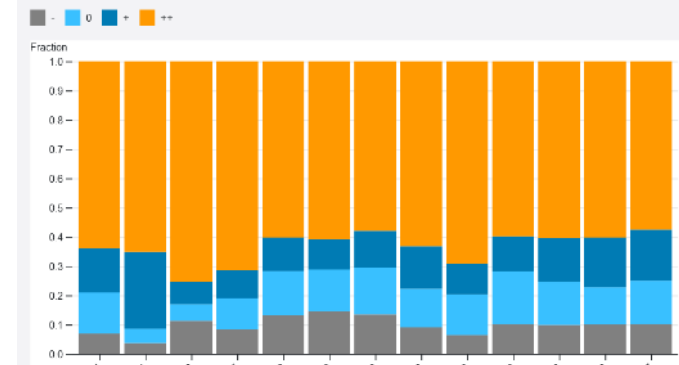
# Automatisierte und maschinengestützte Inhaltsererschließung



# Wissenschaftliche Methode (2019–2022)

- Kombination von *state-of-the-art-Algorithmen* aus dem klassischen Machine Learning inkl. maßgeschneiderter Eigenentwicklung (*stwfsa* \*)
- vorab separat durchgeführte **Parameteroptimierung**
- nachgeschaltet **angewendete Regeln und Filter**
- automatisierte **Qualitätskontrolle** (*qualle* \*\*, machine-learning-basiert, selbstentwickelt), intellektuelle **Qualitätskontrolle**
- Gesamtperformanz aktuell: F1-Wert ~0,6 (in Anbetracht unser Trainingsdaten: gut!)

*omikuji*  
*parabel bonsai*  
*fastText*



## Zwischenfazit: *lessons learned* bezüglich Produktivbetrieb

- ein Produktivbetrieb braucht ein verlässliches Commitment der Institution, d.h. **permanente Ressourcen – Daueraufgabe** statt einer Serie von Projektchen
- es gibt (noch) **keine regalfertige Lösung** – für den Aufbau einer entsprechenden Architektur werden verschiedene **Expertisen im Haus benötigt**
  - **zu besetzende Rollen:** Koordination, angewandte Forschung, Entwicklung Softwarearchitektur und IT-Administration
- Herausforderung: durchdeklinierte **Forschungstransferkette** inkl. entsprechender Anreize für Forschende / für die Institution
- Herausforderung: **TDM-Rechte ...**



---

# Neue Entwicklungen: Large Language Models und die Rolle von Metadaten

# seit 2012 ... neuronale Netze, Deep Learning

- 2012: Google trainiert ein neuronales Netz darauf, Katzen in Bildern zu erkennen
- 2017: Facebook trainiert zwei Chatbots darauf, sich zu unterhalten – nach einer Weile sprechen sie eine eigene gemeinsame Sprache statt Englisch
- 2018: ein von Alibaba (China) trainiertes Modell schlägt menschliches Intellekt in einem Textverständnisstest der Stanford University
- 2018: Google stellt BERT für Sprachverarbeitung vor – Deep Learning
- 2020: Start der Beta-Tests von GPT-3 (OpenAI) – Deep Learning – GPT-3 generiert Code, Gedichte, ...
- 2021: DALL-E (OpenAI) – generiert Bilder aus Prompts →





# Generative Künstliche Intelligenz (*generative AI*)

---

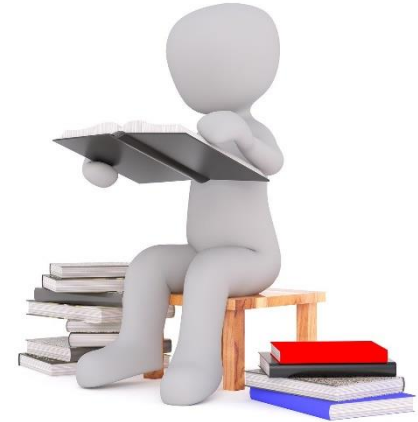
- **generative KI**: Entwicklung von Modellen, die **neue Inhalte** wie Geschichten, Bilder, Videos und Musik erzeugen können
- **Prompt**: natürlichsprachliche Beschreibung (z.B. per Chat- oder per Spracheingabe), was das Modell erzeugen soll
- **ChatGPT** ist ein Chatbot, **über den** man mit einem solchen Modell (GPT-3.5, GPT-4) **interagieren kann** (ChatGPT ist das Interface, nicht das Modell selbst!)



# Large Language Models

---

- sind auf riesigen Mengen von verschiedenen Textarten aus dem World Wide Web trainiert worden
- Use Cases aus dem Natural Language Processing:
  - klassifizieren, clustern, ranken
  - generativ



# LLMs *ad portas!* Bibliothekswesen, *quo vadimus?*

---

Warum bekommen LLMs (auch in Bibliotheken) so viel Aufmerksamkeit?

u.a.: Potential, dass ein funktionierendes **Frage-Antwort-System** diesmal wirklich Realität werden könnte ... ?

ChatGPT: „let me GPT that for you ...“

„steile These“:

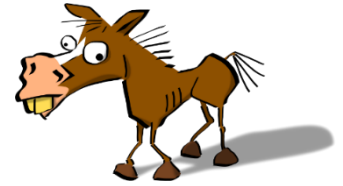
LLMs werden die **klassische Suche** über ein Rechercheportal **obsolet machen!**

Das bedeutet dann ggf. aber auch ...

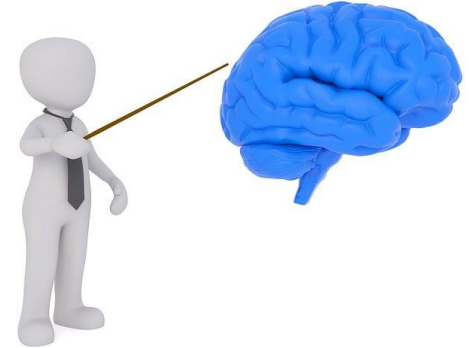
- **keine intellektuelle Inhaltserschließung mehr?**
- **keine Inhaltserschließung mehr?**
- **keine Metadaten mehr??**

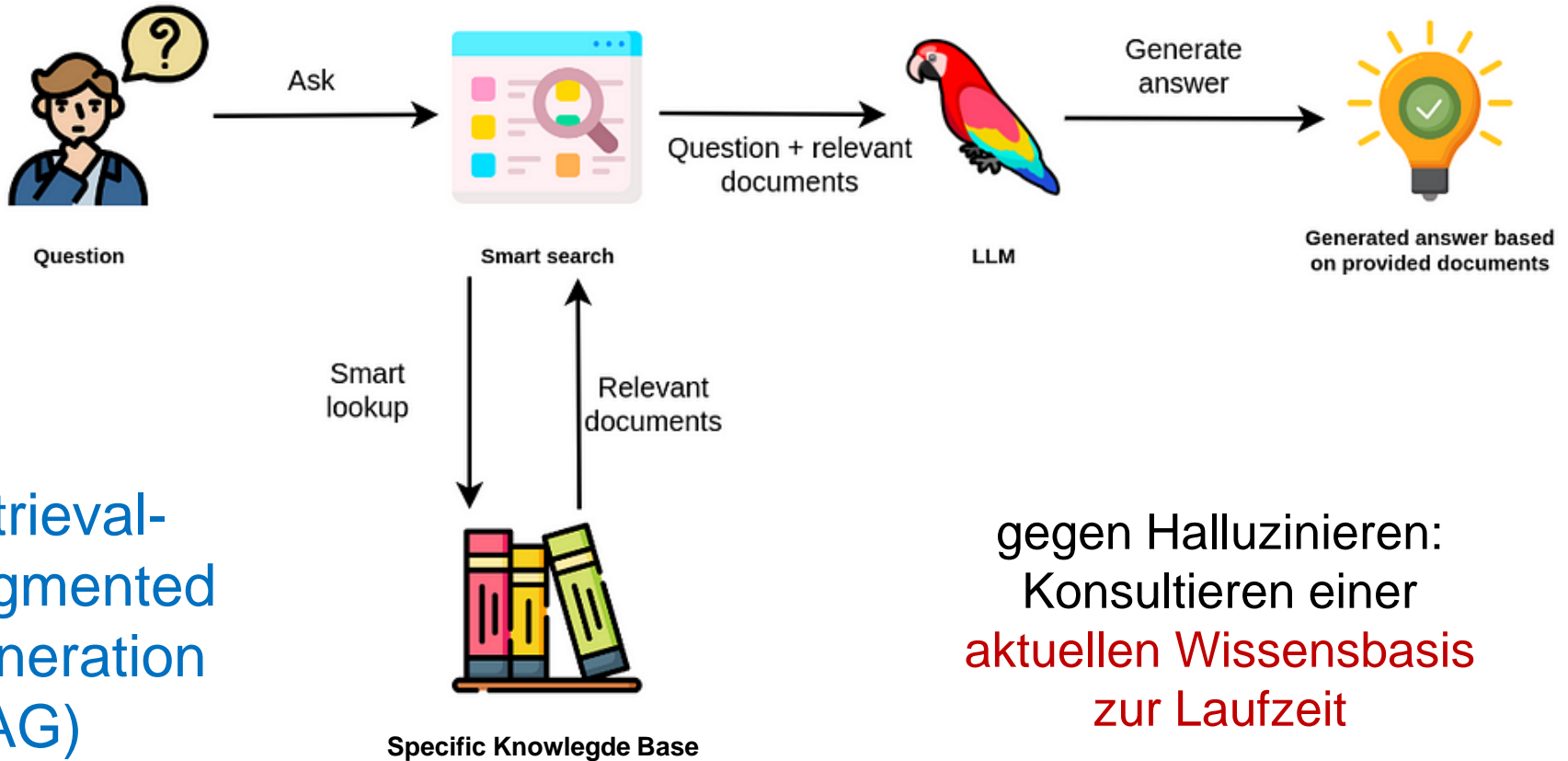


# LLMs als trojanisches Pferd? *hold your horses!*



- Der Output von LLMs kann mit intelligent gewählten Prompts gesteuert werden – aber auch das erfordert Übung!  
→ neue Art der Expertensuche
- Menschen erwarten häufig, dass der Output mit **etablierten Fakten** abgeglichen ist – wenn sie Diskrepanzen feststellen, nennen sie das „**Halluzinieren**“
- Abgleich muss implementiert und mit Prompts erzwungen werden!  
→ **Integration symbolisch repräsentierter Information**, etwa aus Wissensgraphen  
→ **auch Ressourcenfrage!** Anfrage an LLM >>> Anfrage an Wissensgraphen





# Retrieval-Augmented Generation (RAG)

gegen Halluzinieren:  
Konsultieren einer  
**aktuellen Wissensbasis**  
**zur Laufzeit**

# Option: LLMs + Retrieval-Augmented Generation (RAG)

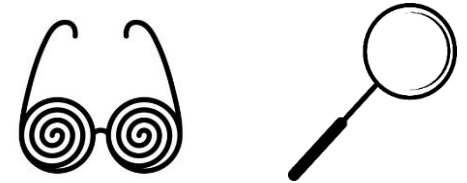
---

- **Preprocessing** für RAG

- aktuellen Dokumentbestand aufsplitten, Datensäuberung
- Generieren von Embeddings für Inhalt, Metadatenextraktion
- Speichern beider Informationsarten (Embeddings, Metadaten) in **Wissensbasis**

- **Retrieval**

- Anfrage in Embeddings übersetzen
- ähnliche Embeddings in **Wissensbasis** suchen
- Ranking der Ergebnisse



- **Reasoning**: LLM beantwortet die Anfrage **im Kontext dieser Ergebnisse**

---

# Neuer Trend: LLMs + Retrieval-Augmented Generation (RAG)

---

- **Preprocessing:** „An important but often overlooked next step is **!!** to extract **metadata** from the [...] input documents.

This could involve steps like generating summaries of each [part], extracting references, headings, document metadata (like timestamps, etc.), and more.“  
„[...] Lastly, we **store the metadata** in a relevant metadata store.“

- **Retrieval:** “Another important step that is often performed during retrieval **!!** is **metadata filtering**. This includes matching certain names, keywords, dates, etc. in the user query and **using the previously extracted metadata** [...] in order to filter out irrelevant pieces of information.” **!!**



## Inhaltserschließung und Retrieval an Bibliotheken – „it's complicated“

---

- Erschließungs- und Retrievalpraktiken bis heute **nicht gut verzahnt**
- **Retrieval-Augmented Generation** als neue Chance?
- Bibliotheken müssen ihre Kernaufgabe neu denken: **METADATEN!**  
jetzt in **neuer Form**, für Interaktion mit LLMs → **weg von Altformaten!**



### zusammenfassend:

- **Option 1:** nur noch LLMs benutzen – problematisch wegen Halluzinieren, fehlender Anbindung an Domänenwissen und wegen Ressourcenverbrauch
- **Option 2:** weiterhin Metadaten in irgendeiner Form einsetzen und mit LLMs kombinieren – erfordert **umfassende Innovation!**



# Aktuelle Forschungsroadmap für AutoSE an der ZBW

---

**Ansatz:** LLMs zur Metadaten- und also Wissensgenerierung einsetzen

- verschiedene **LLMs für die (multilinguale) Inhaltserschließung evaluieren**
- identifizieren, wo diese Modelle Schwierigkeiten mit unseren Daten haben
- Abmildern durch Kombination mit **explizitem Wissen** ausloten
- Abmildern durch Mensch-Maschine-Interaktion ausloten  
(**human in the loop**)



**Ausgang offen** – „steile These“ vom Untergang der Metadaten bleibt zu prüfen

---

# Herzlichen Dank!

**AutoSE** (inkl. weiterer Vortragsfolien und Publikationen):

<https://www.zbw.eu/de/ueber-uns/wissensorganisation/automatisierung-der-erschliessung-mit-methoden-der-kuenstlichen-intelligenz/>

**Kontakt:** a.kasprzik@zbw-online.eu