KI mit Augenmaß – Bericht zum aktuellen Stand der Automatisierung der Inhaltserschließung (AutoSE) an der ZBW

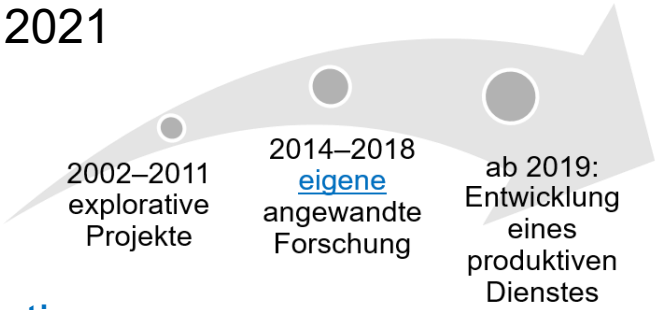
Dr. Argie Kasprzik [ˈkasp.ʃɪk] *(Pronomen en/ens/em/en oder er/sein/ihm/ihn)*
ZBW – Leibniz-Informationszentrum Wirtschaft
FIS-Tagung am DIPF, Frankfurt am Main, 23.04.2026

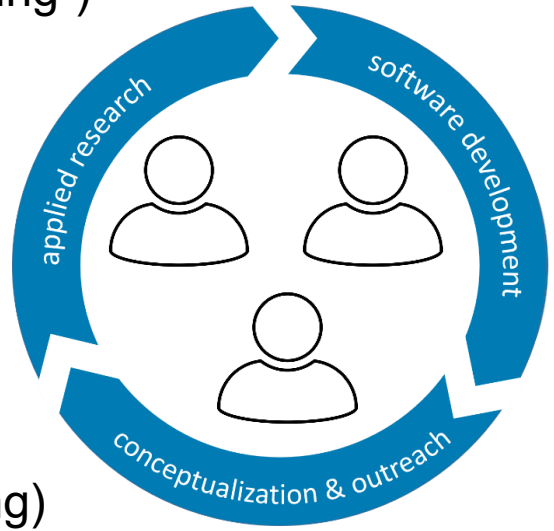
ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Die ZBW ist Mitglied der Leibniz-Gemeinschaft.

AutoSE: Der Weg zum verstetigten Forschungstransfer

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

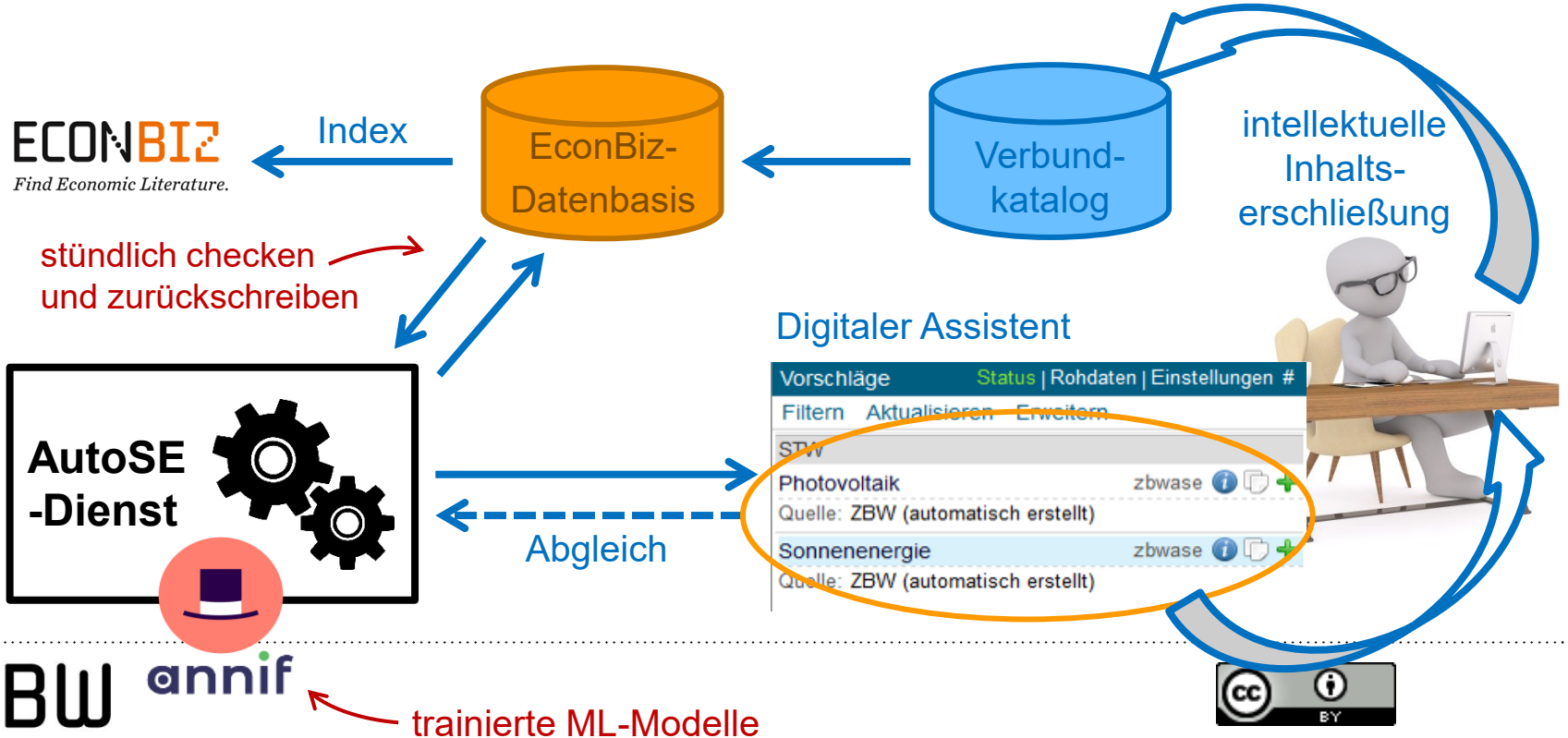




- fortlaufend:
  Evaluierung und Integration
  neuer Entwicklungen aus der KI-Forschung
  (einschließlich unserer eigenen angewandten Forschung)

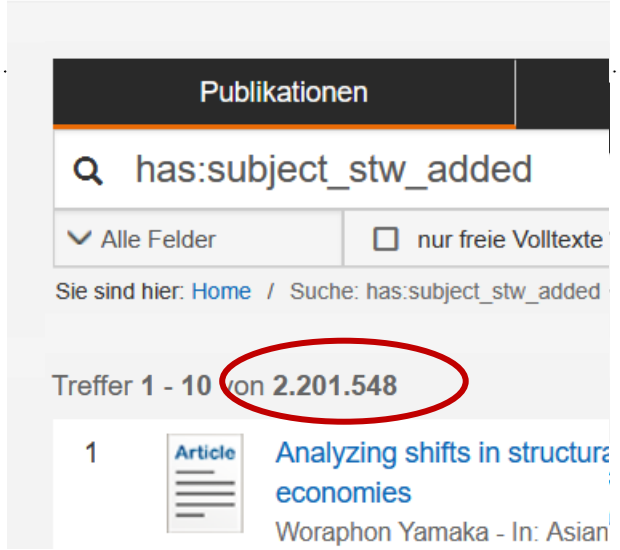
CC BY

Automatisierte und maschinengestützte Inhaltserschließung



Fakten und Zahlen

- wir prüfen die EconBiz-Datenbank stündlich und verschlagworten direkt
- aktuell nur für Publikationssprache „Englisch“ (>Hälfte des Bestandes)
- aktuell verwenden wir nur Titel und, wenn vorhanden, Autor:enkeywords (Abstracts: Experimentalphase)



Stand April 2026: ca. 2,2 Mio. Datensätze angereichert – ca. 1/3 des Bestandes und ca. 55% des englischsprachigen Bestandes

CC BY

Anzeige von AutoSE-Vorschlägen im Digitalen Assistenten (DA-3)

Anzeige im DA-3, einer Plattform für intellekt. Inhaltserschließung, umgesetzt seit 2021 ↓



https://services.eurospider.com/da3/edit?q=1762949687

Intellektuelle Bewertungen im Digitalen Assistenten (DA-3)





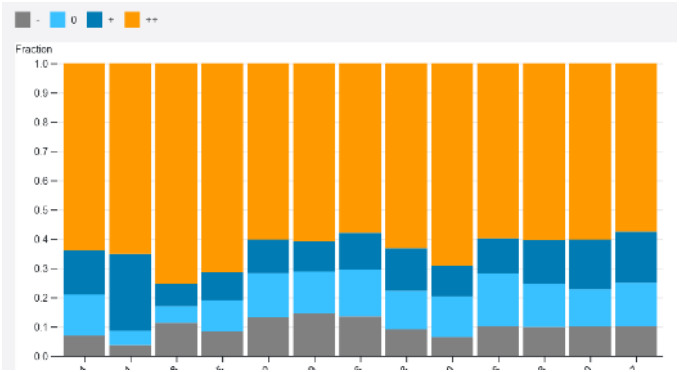
CC BY

Wissenschaftliche Methode bisher

- Verschlagwortung: Kombination von bewährten, leistungsfähigen Open-Source-Machine-Learning-Algorithmen inkl. einer Eigenentwicklung (*stwfsa* *)
- vorab separat: Parameteroptimierung
- nachgeschaltet: Regeln und Filter
- automatisierte Qualitätskontrolle mittels selbstentwickeltem Machine-Learning-Modell *qualle* **
- Gesamtperformanz aktuell: F1-Wert ~0,6 (in Anbetracht unser Trainingsdaten: gut!)

omikuji
parabel *bonsai*
fastText





CC BY

* https://github.com/zbw/stwfsapy ** https://github.com/zbw/qualle

Was kann alles den Automatisierungserfolg beeinflussen?

- Wahl der Methoden (klar), aber auch
- Trainingsdaten – Menge, Eigenschaften, …
- Design, Struktur und andere Eigenschaften des Zielvokabulars
- verfügbare Ressourcen – von Rechen- bis zu Personalressourcen (Expertise)
- Möglichkeiten zur Verstetigung, insbesondere
- Commitment der Beteiligten! – einschließlich der Leitung 😬,
  der intellektuell Erschließenden, der angewandt Forschenden …

CC BY

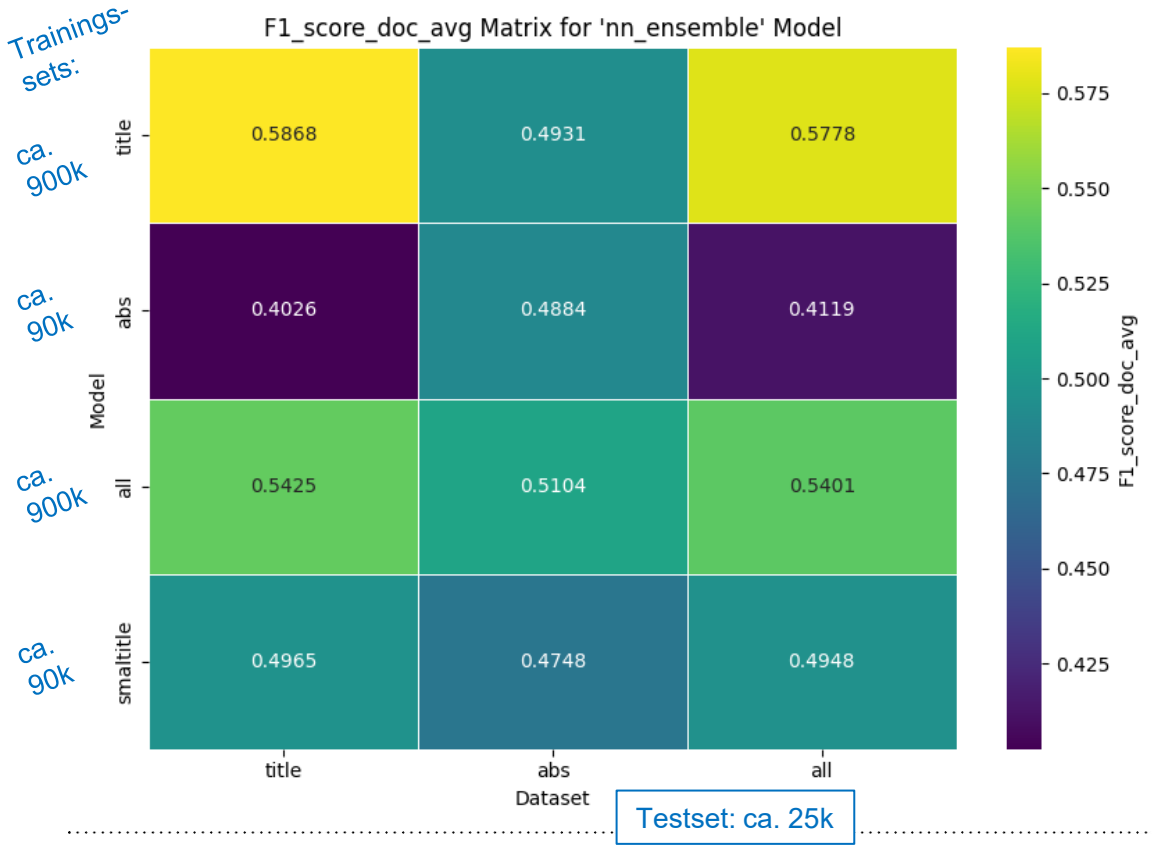
Was kann alles miteinander interagieren?

z.B.:

- Modelle mit Datenbestand
- Modelle mit Zielvokabular
- Modelle mit vorhandenen Rechenressourcen
- Modelle mit vorhandener Infrastruktur / Softwarearchitektur
- Output mit Qualitätsempfinden der intellektuell Erschließenden
- geplanter Dienst mit Workflow der intellektuell Erschließenden
- …

CC BY

Abstracts – für uns an der ZBW aktuell gar nicht so nützlich?



Testset: ca. 25k

Analyse:

- auf Titeln und ggf. AKWs trainiertes Modell performt insgesamt am besten in unserer Realsituation ('title')
- Abstracts in Trainingsdaten scheinen sogar zu stören – ?
- wir haben zu wenig Abstracts in unseren Trainingsdaten (weniger als 10%)! und einige haben die falsche Sprache
- entscheiden: Abstracts überhaupt in den Input?

CC BY

Graphik: Danke an Lakshmi Rajendram Bashyam (ZBW)

Analyzing interaction with our controlled vocabulary (STW)

Tail Gap

~40%

labels with <100 training examples account for a disproportionate share of all misses

Facet Omission

93.4%

model catches one facet (e.g. "Economics") but entirely misses another (e.g. "Geography")

Specific → General

4.1%

model predicts a broader parent category, failing to pinpoint the precise sub-topic

Silent Failures

3.1%

model predicts nothing at all for a valid, labeled document

CC BY

Folie: Danke an Lakshmi Rajendram Bashyam (ZBW)

Analyzing interaction with our controlled vocabulary (STW)

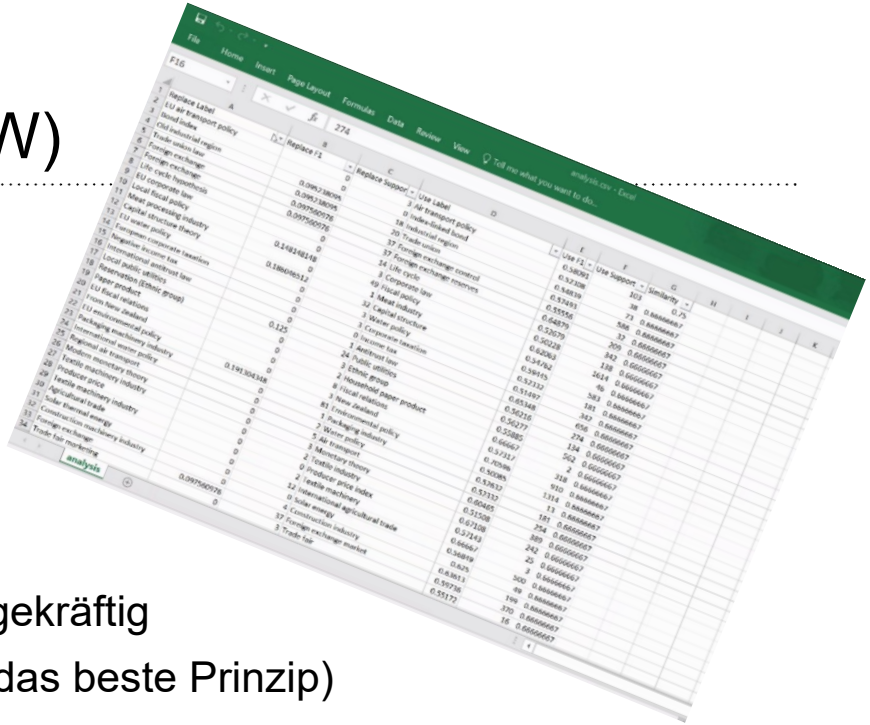
hierarchical near misses

- 12.5% of false positives are hierarchical near-misses
  (sibling or parent labels predicted instead of the correct one)
- in 4.1% of false negatives, a broader parent category is predicted,
  failing to pinpoint the precise sub-topic
- high semantic similarity between confused labels

CC BY

Folie: Danke an Lakshmi Rajendram Bashyam (ZBW)

Analysen zum Zielvokabular (STW)

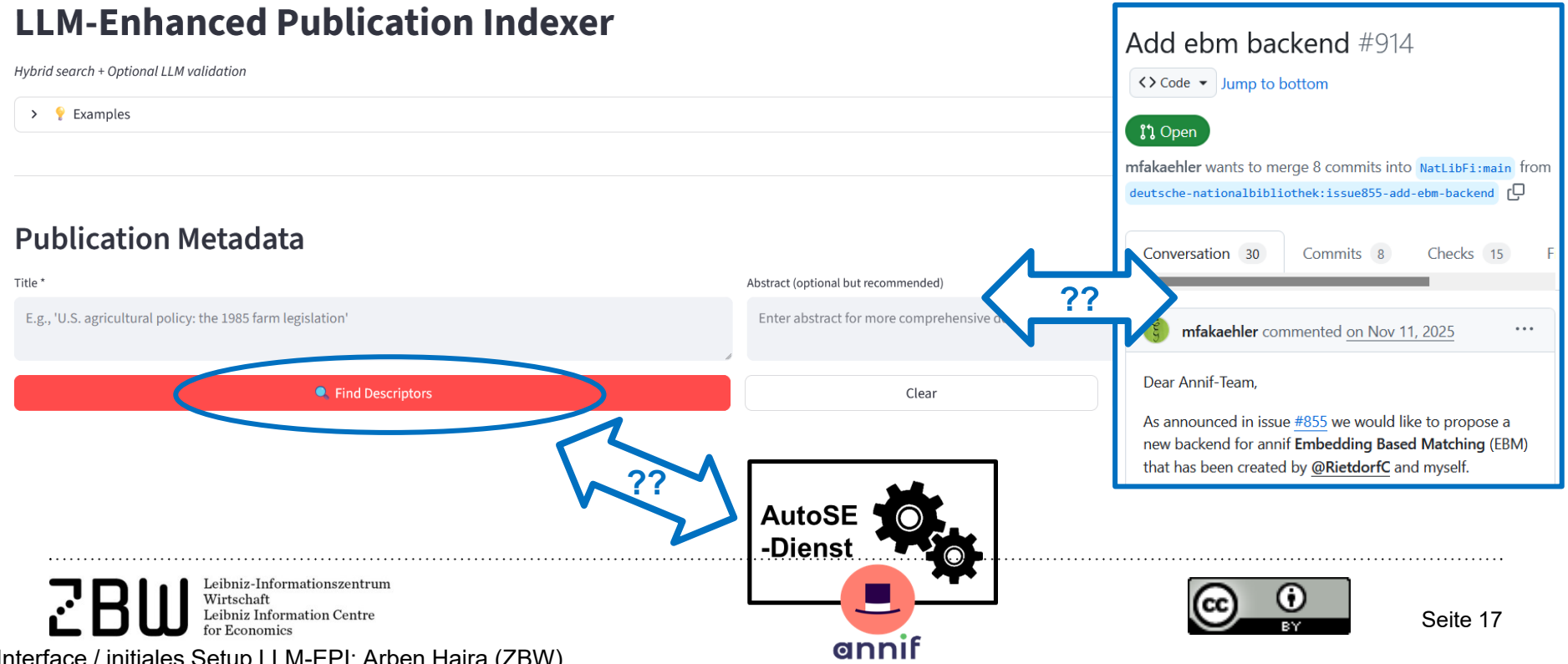


verschiedene vorläufige Erkenntnisse, z.B.:

- AutoSE-Modelle sagen oft den allgemeineren Begriff voraus ("Air transport" statt "EU air transport"), weil der spezifischere in den Trainingsdaten viel seltener vorkommt
- entsprechend ist auch F1-Wert nicht so aussagekräftig ("knapp daneben ist auch vorbei" ist evt. nicht das beste Prinzip)
- Lösungen ggf.: Kontextinformationen in Verwendungshinweisen? Kindknoten im Training gegenüber Elternknoten boosten? etc.

CC BY

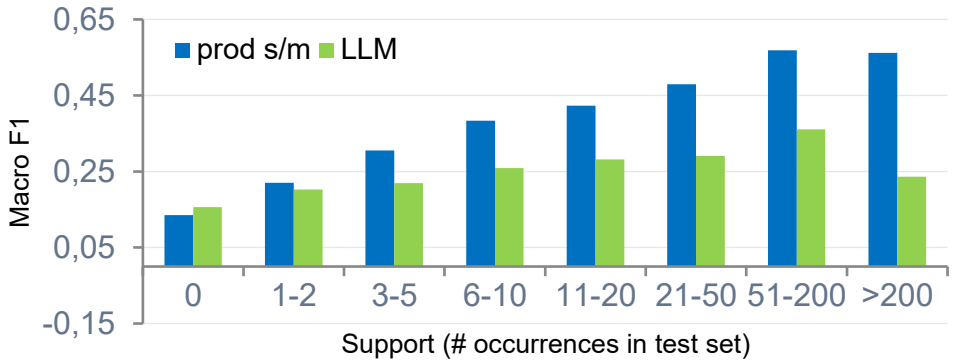
Embedding-basierte Modelle / Generative LLMs (in progress)



Preliminary experiments with LLMs in subject indexing / retrieval

overall our productive model wins

LLM trades precision for recall
– but only on tail labels



Segment	N	Prod s/m F1	LLM F1	ΔF1	Prod s/m Prec	LLM Prec	Prod s/m Rec	LLM Rec
Tail ≤5	2475	0.2114	0.1875	-0.0240	0.299	0.217	0.1799	0.2202
Tail ≤2	1650	0.1644	0.1716	+0.0072	0.19	0.1737	0.1527	0.2145
Support = 0	11	0	0	0.0000	0	0	0	0
Head >50	110	0.5678	0.3505	-0.2173	0.8779	0.5963	0.4387	0.3085

in cooperation with Arben Hajra, ZBW

Folie: Danke an Lakshmi Rajendram Bashyam (ZBW)

Research Roadmap

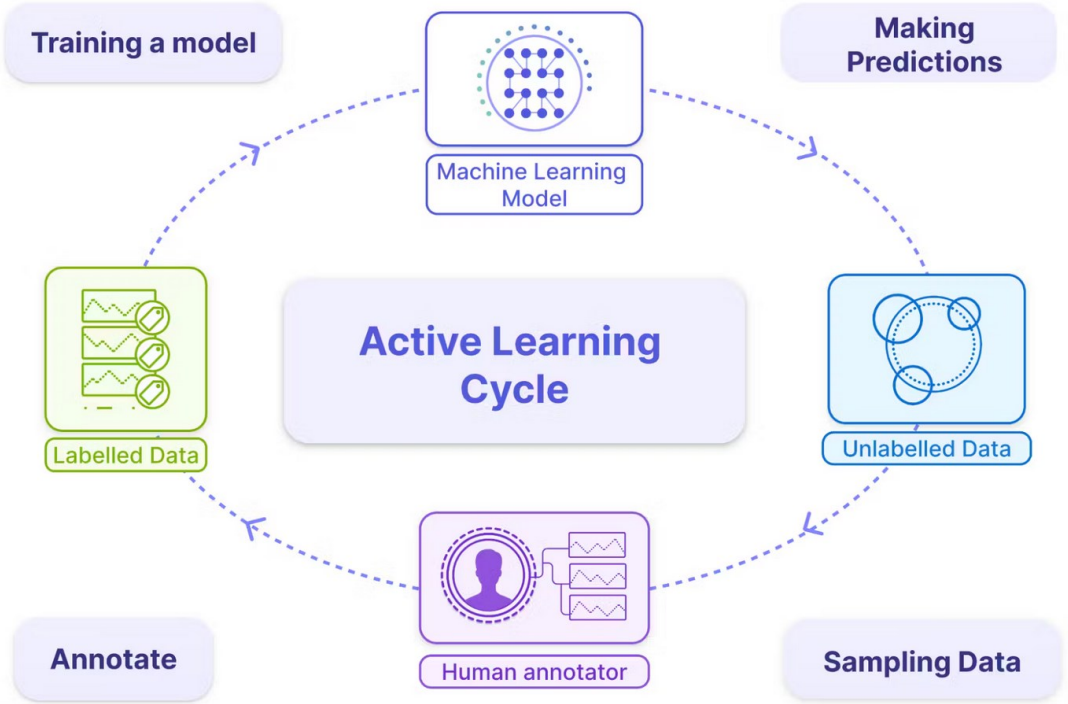
- Experimentieren mit neuem Annif-Feature: *subject inclusion/exclusion*, separate Modelle trainieren und kombinieren (z.B. für Geo)
- insbesondere für *tail labels*:
  - Experimentieren mit LLMs
  - Evaluierung verschiedener Modelle neuerer Generation für AutoSE, u.a. X-Transformer und das *embedding*-basierte EBM (DNB)
  - Data Augmentation
- Active Learning

EBM: https://github.com/NatLibFi/Annif/pull/914
X-Transformer: https://github.com/NatLibFi/Annif/pull/798
https://github.com/NatLibFi/Annif/wiki/Subject-exclusion-and-inclusion

CC BY

Active Learning

- Active Learning: System entscheidet selbst, welche Datensätze das Modell am besten weiterbringen würden
- effizienterer Umgang mit menschlichen Erschließenden als Ressource



- im Produktivbetrieb würde das System diese Datensätze in eine Queue bzw. einen Pool schieben und darauf warten, dass Menschen sie annotieren

CC BY

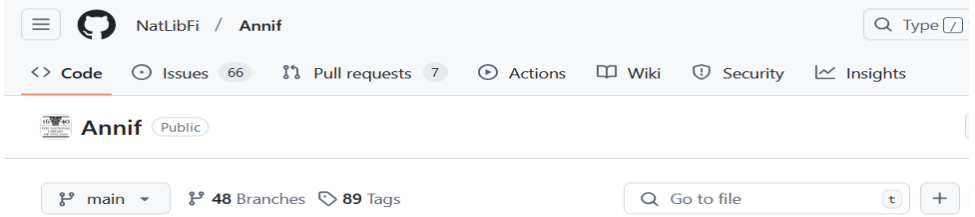
https://encord.com/blog/active-learning-machine-learning-guide/

Kommerziell oder Open Source?

- kommerzielles Produkt oder Open Source?
  - kommerziell → bitte wenden Sie sich an Ihre*n Arzt* oder Apothek*er
  - Open Source: erfordert
    - modularen Aufbau
    - gewisse Inhouse-Expertise
    - Recherche- und Einarbeitungsphase zu Voraussetzungen, erforderliche Ressourcen, Modellen, …



CC BY



Annif

- Annif – entwickelt von der finnischen Nationalbibliothek: ein (vergleichsweise) niederschwelliges Toolkit für die Automatisierung der Inhaltserschließung, bietet diverse Verfahren (Backends) an
- perspektivisch zu prüfen, ob die Anforderungen von "Power Users" mit produktiven Systemen an die Funktionalitäten in AutoSECore weiterhin mit dem Anspruch der Niederschwelligkeit von Annif kompatibel sind
- nächstes Annif-Tutorial (NLF+ZBW+DNB):
  Open Repositories 8. bis 11. Juni 2026 online
  https://github.com/NatLibFi/Annif-tutorial (auch zum Selbststudium)

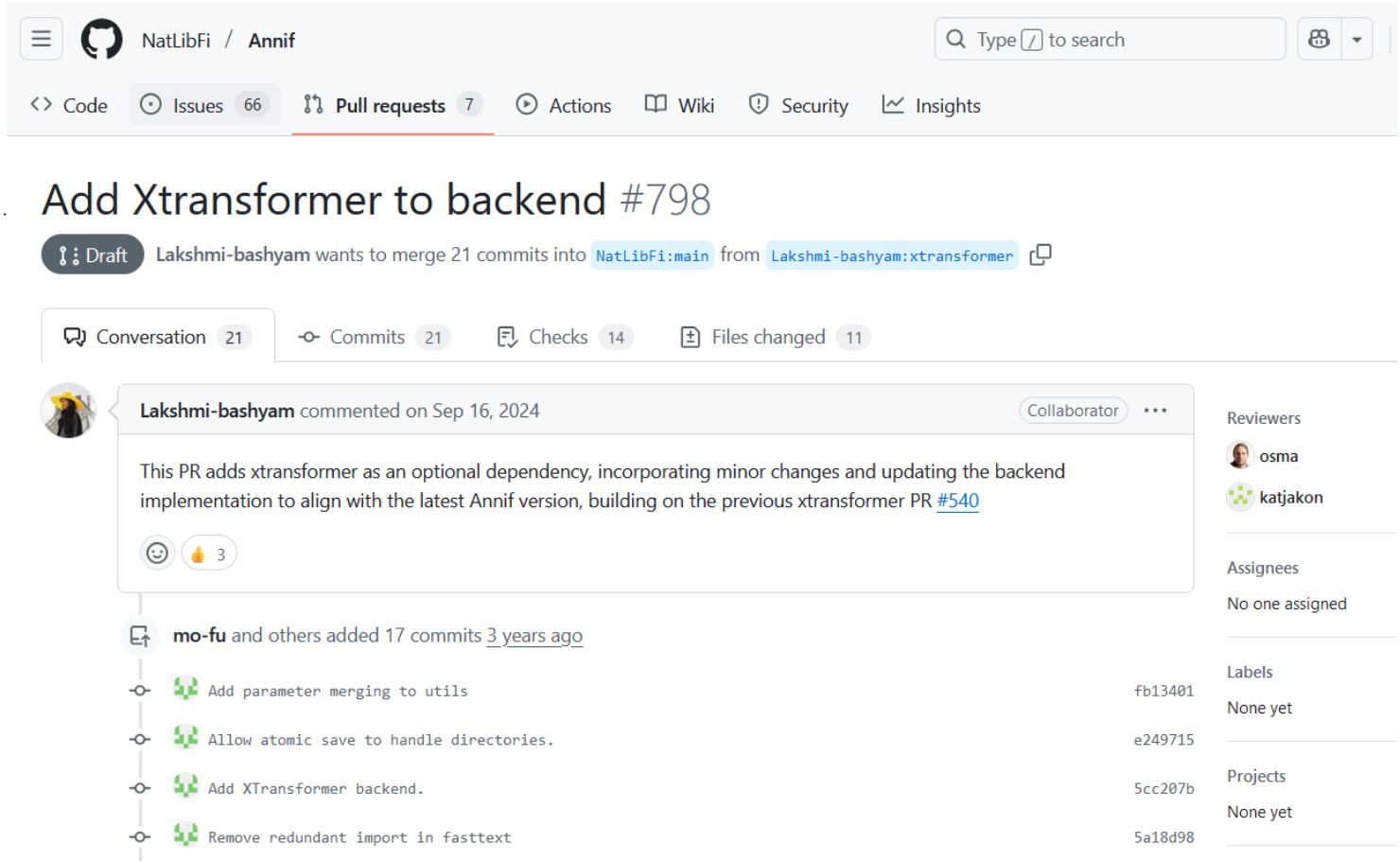
Herausforderung Open Source Community

- Annif ist eine Open-Source-Plattform – auch nachgenutzte externe Komponenten und angebotene Modelle sind Open Source
- Herausforderung: Projekte werden öfters plötzlich nicht mehr gepflegt
  - aktuelles Beispiel: neues Modell X-Transformer → Pecos-Bibliothek
  - aber auch fastText, omikuji, also "Bestandsmodelle"!
- vs. dauerhafter Produktivbetrieb – ?
- erfordert kooperative Anstrengung der Community!



CC BY

Pull Request auf GitHub von AutoSE-Team, um Transformer-modelle in Annif zu integrieren



CC BY

OpenAI *et al. ante portas* – was ist so problematisch an kommerzieller Large AI (LAI)?



CC BY

LAI schadet Umwelt und Klima



STROM



LUFT

WASSER



Das Training und vor allem
auch der dauerhafte Betrieb von
LAI-Anwendungen verbraucht
gigantische Mengen von Ressourcen,
oft mehr als kleine Staaten, und der
Trend zeigt immer noch nach oben





ROHSTOFFE

E-SCHROTT

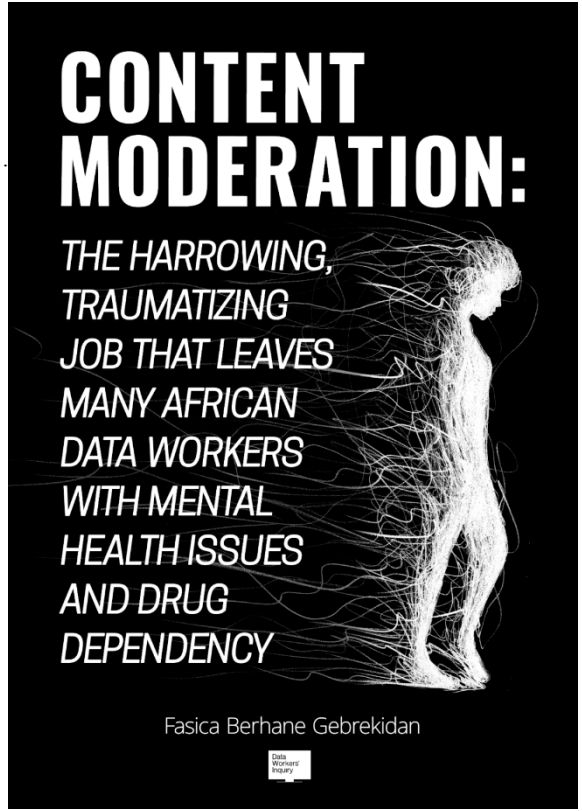
ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

CC BY

Bilder: https://icon-icons.com/de/symbol/Stecker/39479 ; https://ccnull.de/foto/dry-white-long-rice-in-wooden-spoon/1038560 ; https://www.pexels.com/de-de/foto/industrie-nebel-smog-pflanze-7001364/ ; https://commons.wikimedia.org/wiki/File:Elektroschrott.jpg ; https://ccnull.de/foto/computerchip/1006620

Ausbeutung

- Large AI stützt sich weiterhin auf die Sichtung, Bereinigung und Annotation von Daten durch Menschen
- diese Arbeit wird von Big Tech oft in weniger privilegierte Länder outgesourcet, unter ausbeuterischen Bedingungen – Stundenlöhne unter 2 USD (z.B. Kenia, Argentinien), kein Arbeitsschutz
- viele Menschen, die diese Arbeit verrichten, leiden unter Gesundheitsproblemen und posttraumatischen Störungen



CC BY

Screenshot: https://data-workers.org/

Gatekeeping und unzulässige Einflussnahme

Wissen ist Macht – Asymmetrie verschärft sich

- Bezahlmodelle ("Premium") → Zugang zu (angeblich) hoch-qualitativen Informationsquellen und Tools werden zum Luxusgut
- wer kein Geld hat, zahlt mit Daten und Privatsphäre
- Big-Tech-Firmen bauen Monopole (Blase!) und werden zu Gatekeepern – kontrollieren Zugänge, beuten Nachnutzende aus
- Big-Tech-Firmen erlangen unzulässige Möglichkeiten zur Einflussnahme auf Use Cases im öffentlichen Interesse (etwa Forschung, Bildung, Infrastruktur)



Bild: CC-BY vectorportal.com

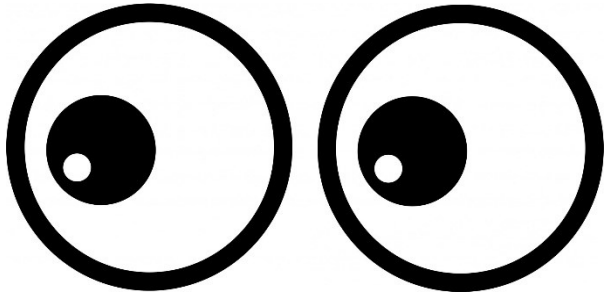
CC BY

siehe auch von Thun, Max und Hanley, Daniel, Stopping Big Tech from Becoming Big AI: A Roadmap for Using Competition Policy to Keep Artificial Intelligence Open for All. Open Markets Institute 2024. Online unter http://dx.doi.org/10.2139/ssrn.4990780

Datenschutz, Infrastruktur – zwei Beispiele

oft ungefragte, invasive Nutzung von Daten ohne Rücksicht auf Persönlichkeits-rechte, Urheberrecht, Lizenzbedingungen, KI-Ausschlussvorkehrungen

- Bsp. Experimente durch OpenAI / Forschende der Uni Zürich in Online-Community (Reddit)
- Bsp. Belastung der Open-Access-Infrastruktur durch Crawling / Bots

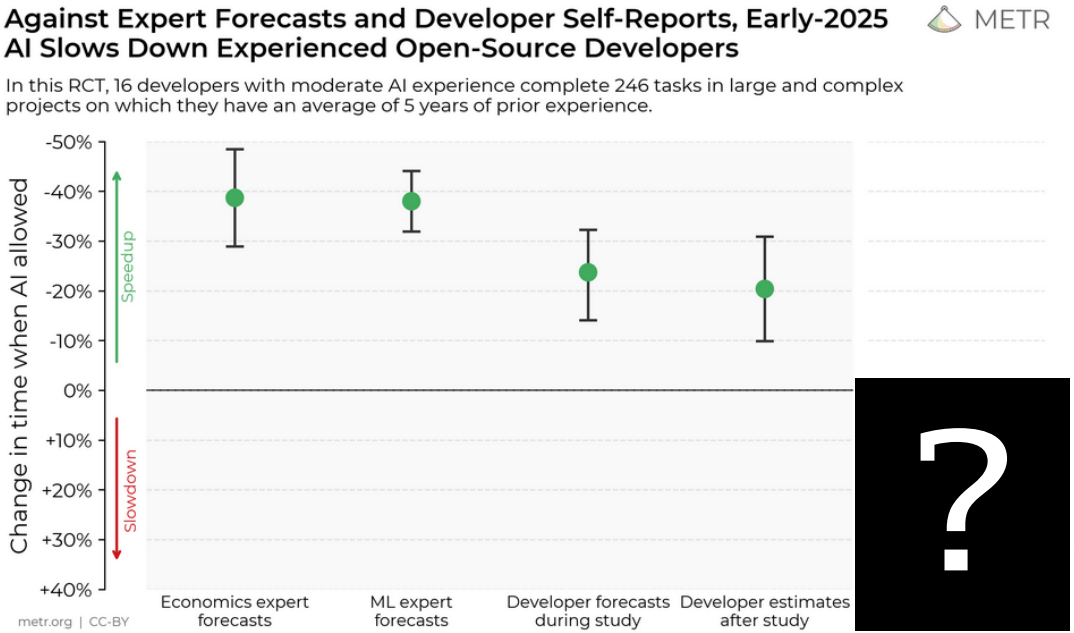


https://opentools.ai/news/openai-turns-rchangemyview-into-ais-persuasion-playground
https://retractionwatch.com/2025/04/29/ethics-committee-ai-llm-reddit-changemyview-university-zurich/
https://arstechnica.com/ai/2025/03/devs-say-ai-crawlers-dominate-traffic-forcing-blocks-on-entire-countries/
https://coar-repositories.org/news-updates/open-repositories-are-being-profoundly-impacted-by-ai-bots-and-other-crawlers-results-of-a-coar-survey

CC BY

Change Management

- Berichte aus der Softwareentwicklungscommunity: immer mehr Manager erzwingen den Gebrauch von KI-Tools
- potentiell negativer Einfluss von fehlgeleitetem Change Management auf Produktivität (!), Psyche, Motivation des Personals ↓ →



?

Tech Industry 1w · Follow
Uber · vibesuxx

I hate vibecoding, change my mind

Vibecoding sounded cool in the beginning but now it's driving me crazy. The junior on my team keeps churning out AI-generated CRs like a machine.

I waste hours reviewing the same lazy, copy-paste AI junk. Pretty sure they're pasting code straight from ChatGPT or whatever, tossing feedback back into it, and resubmitting without even bothering to look.

Meanwhile, my manager won't shut up about AI adoption, so now it's basically a free pass to flood the repo with garbage. No care for the codebase, zero respect for anyone's time.

Anyone else stuck in this mess? I'm seriously done with this crap.

CC BY

Screenshot: https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/

Screenshot: https://mapstodon.space/@abel/114975709443152718



Timnit Gebru (she/her)
@timnitGebru@dair-community.social

Friends, for something to be open source, we need to see

1. The data it was trained and evaluated on

2. The code

3. The model architecture

4. The model weights.

DeepSeek only gives 3, 4. And I'll see the day that anyone gives us #1 without being forced to do so, because all of them are stealing data.

LAI und Open Source

Screenshot: https://dair-community.social/@timnitGebru/113909880610412733

siehe auch https://opensource.org/ai/open-source-ai-definition und https://www.forkable.io/p/metas-new-llama-4-ai-models-arent#%C2%A7openness-is-binary ;

CC BY

Greenwashing / Open Weight vs. ECHTES Open Source

- Firmen wie OpenAI und Meta behaupten regelmäßig "Open-Source-Modelle" zu veröffentlichen (LLaMa, GPT-OSS)



- kriegen aber Widerspruch von der Open-Source-Community – die sind nur Open Weight
- … es gibt leider noch recht wenig volle Open-Source-Modelle



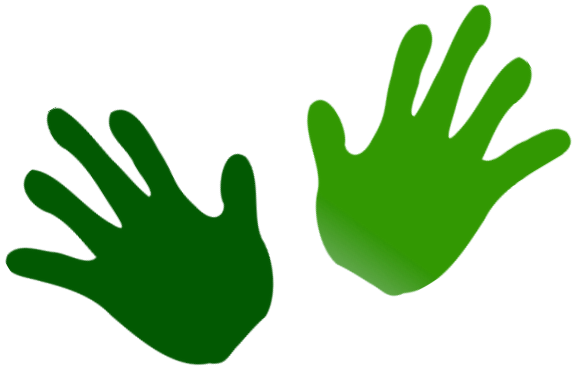
- aber: SwissAI mit ihrem LLM Apertus (2025) kommt nah ran!
  - trainiert auf 15 Mrd. Token über mehr als 1000 Sprachen
  - 40% der Daten sind nicht-englisch
  - "the entire development process, including its architecture, model weights, and training data and recipes, is openly accessible and fully documented"

siehe auch https://www.forkable.io/p/metas-new-llama-4-ai-models-arent#%C2%A7openness-is-binary ;
Swiss-AI Apertus: https://www.swiss-ai.org/apertus

CC BY

Warum Open Source?

- wichtigste Botschaft immer noch: von Fall zu Fall prüfen –
  was ist die sparsamste auf den Use Case zugeschnittene Lösung?
- aber: wenn KI-Methoden, dann sollte vollkommen transparent sein,
  - wie sie entstanden sind
  - wem sie zugute kommen (Gemeinwohl!)
  - was dafür erforderlich ist, sie nachzunutzen
  - wieviel sie verbrauchen



… und das geht nur mit Open Source.

CC BY

Civilized AI

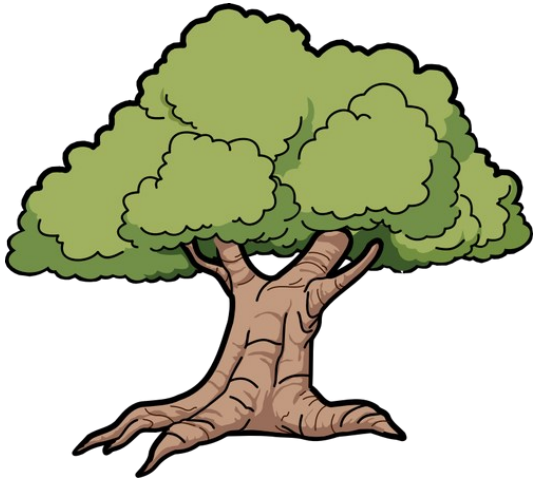
1. “use AI to make the world better”
   (automation)
2. “use the smallest AI that works”
3. “don’t depend on corporate AI”
4. “evaluate & create data sets”
5. “be open and transparent”

perfect task for libraries!



Solarpunk-Logo: LySioS-wkp, CC BY-SA 4.0, https://commons.m.wikimedia.org/wiki/File:SolarPunk_logo-flat.svg

Wie können sich Institutionen engagieren?



- sensibilisieren (und selbst sensibel agieren)
- ✊ **Aktivismus !**
  - **Openness** pushen (Open Data, Open Source)
  - Verbesserung der Rechtslage pushen
    (→ *fair use* – für Zwecke der Bildung, Forschung, Informationsversorgung, NICHT für kommerzielle Zwecke)
- 🤝 **Kooperation !**
  - "evaluate & create data sets" !
  - ressourcenschonende Modelle / Dienste nachnutzen, möglichst lokal betreiben (lassen) und auf garantierten Datenschutz achten
- für Kompetenzzentren mit entsprechenden Möglichkeiten:
  - ressourcenschonende Modelle / Dienste entwickeln, Prozess transparent machen, lokal und datenschutzbewusst betreiben, offen zur Nachnutzung zur Verfügung stellen
  - eigene Software / Hardware CO2-optimieren (parametereffizientes Finetuning etc.)

CC BY

Bild: https://publicdomainvectors.org

Beispiele für Initiativen:

Library and Archives 101: AI and the False Promise of Control

By Dave Hansen / March 16, 2026

AUTHORS ALLIANCE

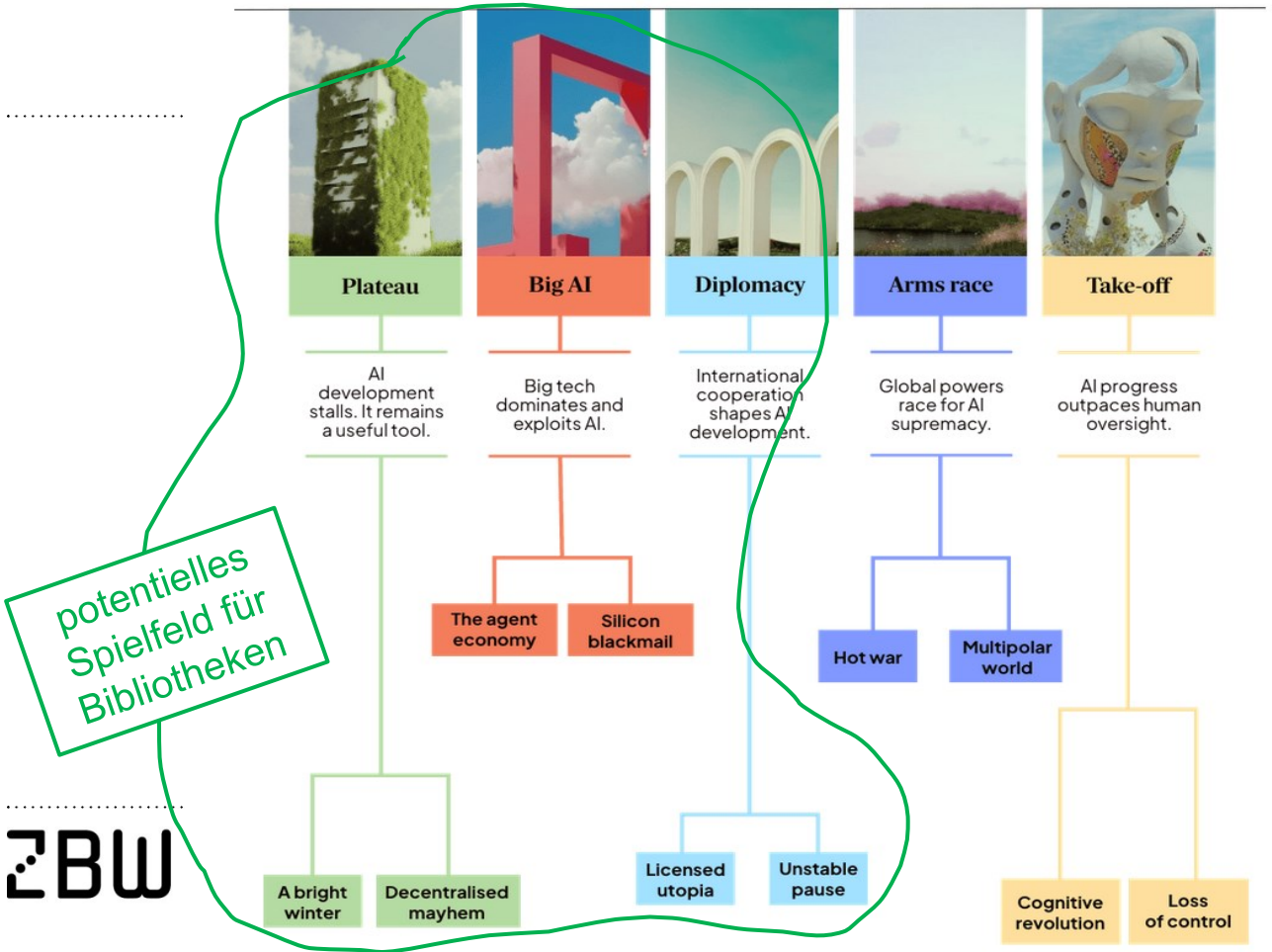
So what can we do?

This is not a call toward passive acceptance of AI company dominance. I understand the anxieties that motivate efforts to exert more control, and they are responding to real problems. But the answer to those anxieties cannot be to remake the library in the image of Elsevier or Disney. Asserting control, demanding compensation, and conditioning access on the institution's ability to dictate the terms of downstream use—that is not what libraries and archives are for.

Libraries are not going to out-monopoly the monopolists by asserting control over their collections. But they can do something far more important: they can ensure that institutions, researchers, and communities they serve have a seat at the table—not by withholding access, but by building the infrastructure that supports open, accountable research of every kind.

CC BY

Quelle: https://www.authorsalliance.org/2026/03/16/library-and-archives-101-ai-and-the-false-promise-of-control/



https://cfg.eu/advanced-ai-possible-futures/

Meine beiden Botschaften





Quelle links: https://todon.eu/@autonomysolidarity/115994031631358355 ;
Quelle rechts: https://mastodon.online/@streetartutopia/115186042059851789

CC BY

Neuer Sondertatbestand MetaFor an der ZBW

- Metadaten wertvoll für Entwicklung und Evaluierung von KI-Verfahren
- liegen oft verteilt, schwer zugänglich, nicht "AI-ready"
- ZBW baut interoperable Infrastruktur MetaFor zur fortlaufenden Automatisierung ihrer Metadatenworkflows und befähigt sich damit,
- als Data Provider maschinenlesbare Metadaten für KI-Forschung und -Entwicklung über offene Schnittstellen bereitzustellen
- Vorhaben wird mit 7,5 Mio. € gefördert
- offizieller Beginn 2027, Pilotarbeiten laufen bereits

CC BY

Danke!

Kontakt: a.kasprzik@zbw.eu

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

CC BY