

human in the loop: intellektuelle Bewertung von automatisiert erzeugter Verschlagwortung

*im Kontext der AutoSE – Automatisierung der Inhaberschlüsselung
mit Machine-Learning-Methoden an der ZBW*

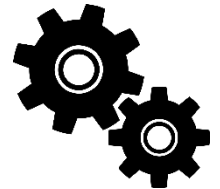
Dr. Anna Kasprzik

ZBW – Leibniz-Informationzentrum Wirtschaft

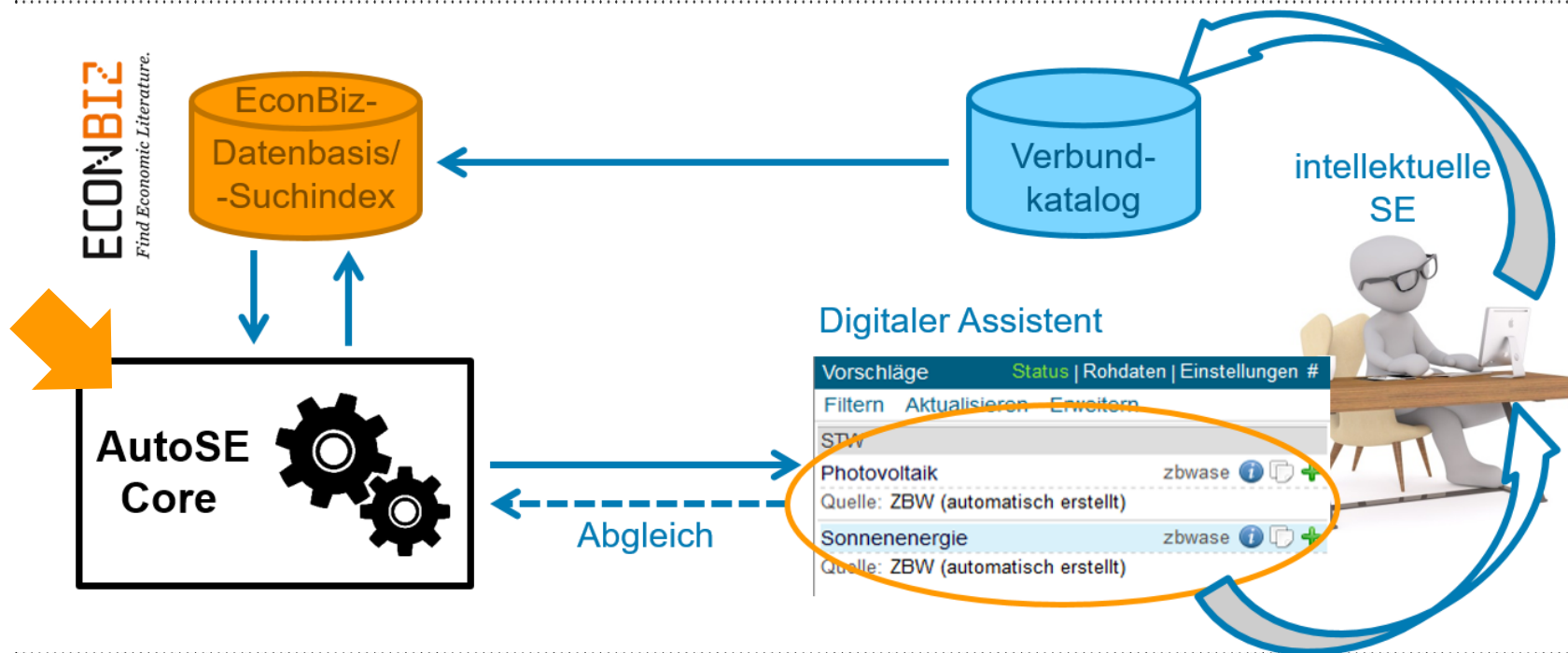
Workshop „Computerunterstützte Inhaberschlüsselung“, online, 16. und 17.11.2022

Kontext

- AutoSE: ein produktiver Dienst an der ZBW, der permanent durch Ergebnisse aus der hauseigenen angewandten Forschung verbessert wird
- Methoden: Annif * als Steckrahmen, ergänzt durch ein ZBW-eigenes Modell und ZBW-eigene Software-Komponenten –
 - für die Weiterentwicklung: wissenschaftliches Experimentieren, Parameteroptimierung, Qualitätskontrolle
 - für den Produktivbetrieb: Anschluss an Erschließungsworkflows, Monitoring, Deployment, Continuous Integration, ...

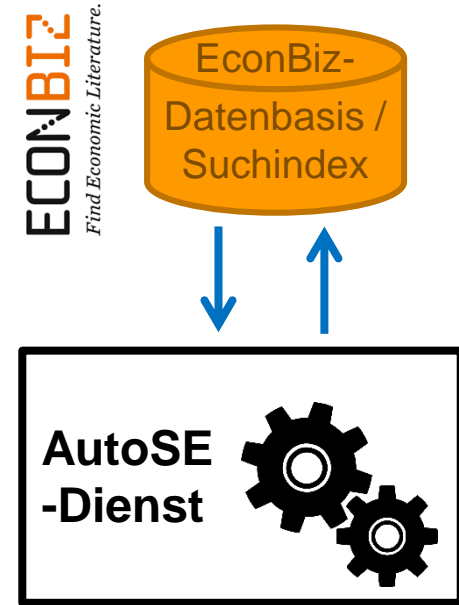


Datenflüsse: Interaktion der verschiedenen Produktivsysteme



Anbindung an EconBiz-Datenbasis

- wir prüfen EconBiz-Datenbasis **stündlich** auf neue Publikationen im Bestand der ZBW und verschlagworten sofort
- aktuell nur Publikationen mit Publikationssprache „**englisch**“
- aktuell nur auf der Basis des **Titels**, und, wenn vorhanden, der **Autoren-Keywords** (perspektivisch: Abstracts, ToCs, ...)
- Stand September 2022: **1,4. Mio Datensätze** mit AutoSE, das sind ~25% des ZBW-Bestandes



Anzeige von Vorschlägen im DA-3

Kurztitel	#
Nummer: 1032536500	
Titel: Signature experience : art and science of customer engagement for fashion and luxury companies / edited by Stefania Saviolo	
Personen: Saviolo, Stefania [HerausgeberIn]	
Ausgabe: First edition	
Publ.: Milano : BUP, August 2018	
ISBN: 978-88-99902-31-5, 978-88-85486-59-1	
Sprache: Englisch [text]	
Weitere Daten	

Vorschläge	Status	Rohdaten	Einstellungen	#
STW				
Beziehungsmarketing			zbwase	
Quelle: ZBW (automatisch erstellt)				
Konsumentenverhalten			zbwase	
Luxusgüter			zbwase	
Markenführung			zbwase	
Mode			zbwase	
GND				
Beziehungsmarketing [Sach]			@stw-exact	
Luxusout [Sach]			@stw-exact	

Qualitätskontrolle – *human in the loop*

- Aufgabe: Ausarbeitung eines konsistenten Qualitätssicherungskonzeptes, um nur qualitätsgeprüfte Daten herauszugeben
- wichtiger Baustein: *human in the loop – ways for humans and machine learning algorithms to interact to solve problems*
- Spektrum von Umsetzungen:
 - intellektuell annotierte Trainingsdaten
 - intellektuell gepflegte Wissensorganisationssysteme und Mappings
 - maschinengestützte Erschließung → DA-3
 - intellektuelle Bewertung des Outputs, Identifizieren systematischer Abweichungen
 - Online-Learning, Active Learning



Das *releasetool*

Title: **Improved calendar time approach for measuring long-run anomalies**

Keywords:

Abstract: Although a large number of recent studies employ the buy-and-hold abnormal return (BHAR) methodology and the calendar time portfolio approach to investigate the long-run anomalies, each of the methods is a subject to criticisms. In this paper, we show that a recently introduced calendar time methodology, known as Standardized Calendar Time Approach (SCTA), controls well for heteroscedasticity problem which occurs in calendar time methodology due to varying portfolio compositions. In addition, we document that SCTA has higher power than the BHAR methodology and the Fama-French three-factor model while detecting the long-run abnormal stock returns. Moreover, when investigating the long-term performance of Canadian initial public offerings, we report that the market period (i.e. the hot and cold period markets) does not have any significant impact on calendar time abnormal returns based on SCTA.

Collection: [BRLR, fsta no-min2](#)

Document: 10011449859

Links:

Navigation:

Actions:

Progress: 0 / 200

ca. 1000 Dokumente
pro Review geprüft

Automatically Assigned Subjects

[\(explain\)](#)

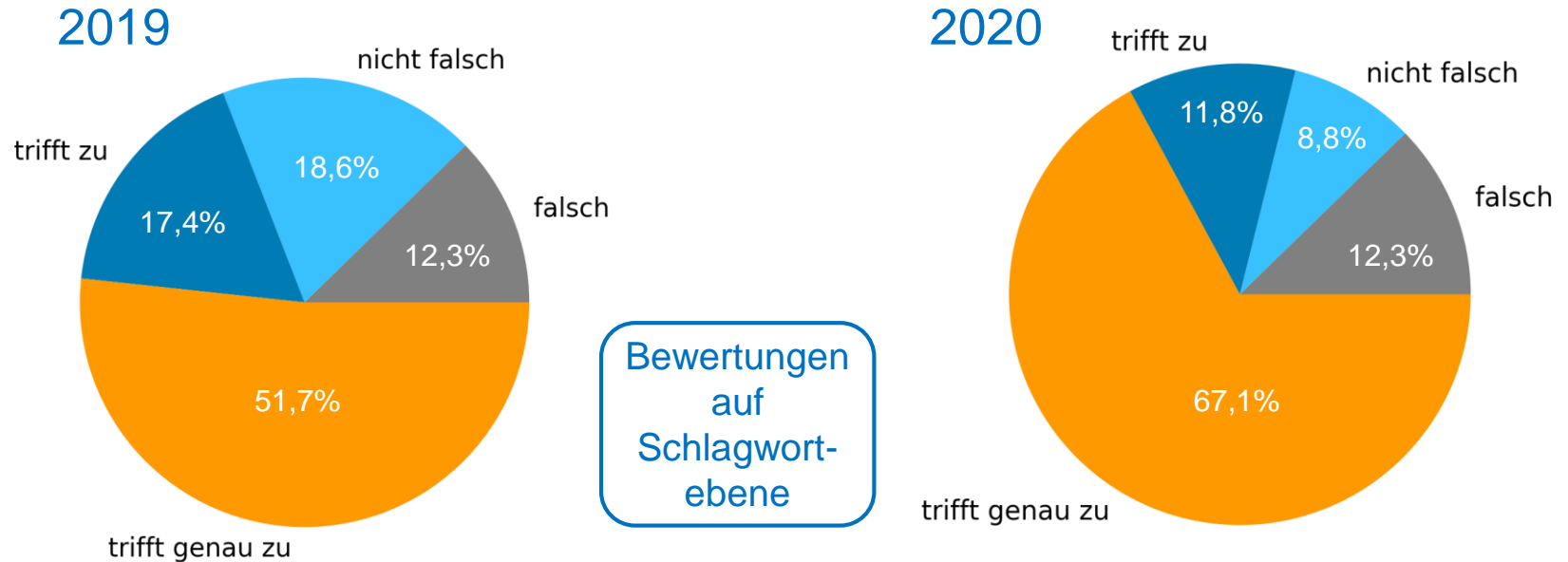
Rating	Subject	Categories
-- 0 + ++		
<input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Power	<input checked="" type="checkbox"/> N
<input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/>	Time	<input checked="" type="checkbox"/> V <input checked="" type="checkbox"/> N
<input type="radio"/> <input type="radio"/> <input type="radio"/> <input checked="" type="radio"/>	Capital market returns	<input checked="" type="checkbox"/> V

Document-level Quality

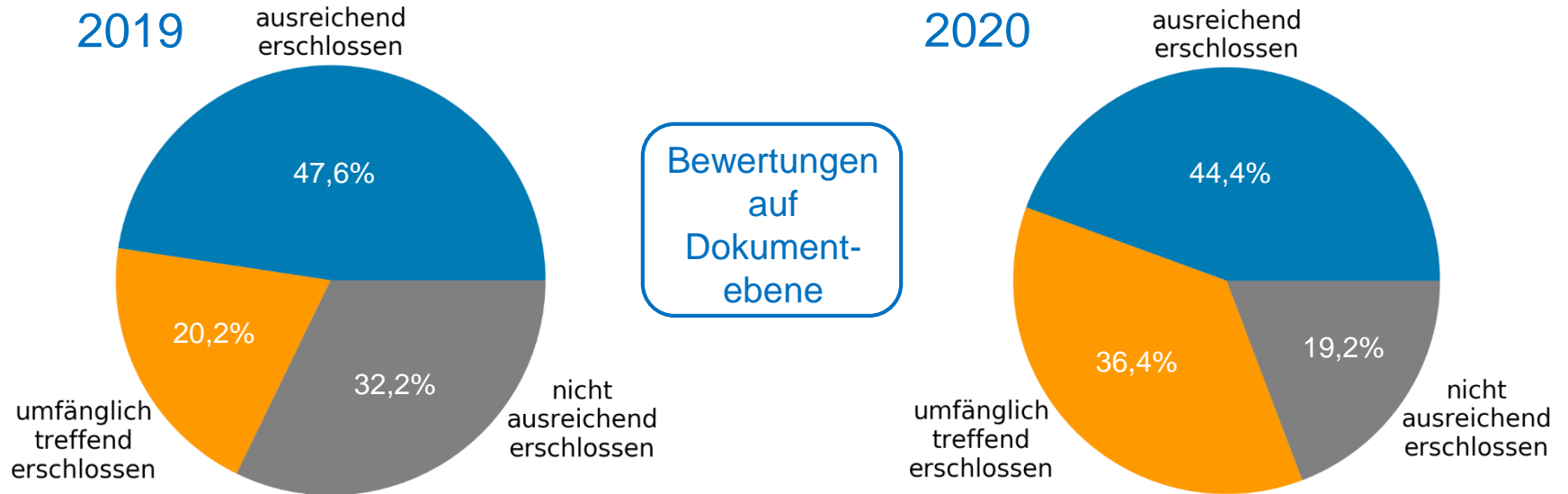
- good
- fair
- reject
- skip

Missing Subjects

Entwicklung vorletztes auf letztes Review – Deskriptoren



Entwicklung vorletztes auf letztes Review – Dokumente



2 Beispiele für spezifische Erkenntnisse aus Review 2020

- Erschließende: AutoSE schlägt zu oft fälschlicherweise „Theorie“ und „USA“ vor
- Erklärung: „Theorie“ (27%) und „USA“ (16%) sind auch die **intellektuell** – und damit in den Trainingsdaten – **am häufigsten vergebenen Konzepte!**

wie beheben wir das? zwei neue Filter:

- „USA“ wird **geblockt**, außer „USA“ („US“, „United States“) erscheint **explizit**
- Erschließende liefern eine **Liste mit Konzepten** für spezifische Theorien, die „Theorie“ **blockieren** sollen

Wunschliste für ein Bewertungstool

zwei Personae:

- SEB: sacherschließende Person mit dem Auftrag (Auto)SE zu bewerten
- LBM: Person, die die Bewertung von (Auto)SE leitet



Wunschliste für ein Bewertungstool – für SEB (Prio 1)

- SEB kann den Zeitpunkt der Bewertung frei wählen
- SEB kann für die SE relevante Textelemente parallel sehen
- SEB kann jedes einzelne vorgeschlagene Schlagwort (SW) bewerten
- SEB kann pro Quelle die Gesamtmenge der von dieser Quelle für eine Ressource vorgeschlagenen SW bewerten
- SEB kann fehlende SW ergänzen
- SEB kann nach Bewertung diese noch verändern und neu abspeichern
- wenn SEB die Maske verlässt ohne Bewertung zu speichern → Fehlermeldung
- nicht alle SW oder Gesamtmenge nicht bewertet → Fehlermeldung

Wunschliste für ein Bewertungstool – für LBM (Planung)

- LBM kann mehrere Ziel-Vokabulare spezifizieren
- LBM kann spezifizieren, ob die Auswahl der zu bewertenden Ressourcen kontinuierlich und zufällig erfolgt oder anhand einer vorgegebenen Liste
- LBM kann eine solche Liste erstellen und zuweisen oder zufällig generieren lassen und zuweisen oder zufällig generieren und zufällig zuweisen lassen
- LBM kann zu jedem Bewertungsdurchlauf [*Vokabular, Liste, SEB-Nutzergruppe, ...*] individuell Bewertungsstufen definieren
- LBM kann Anzahl und Bezeichnung der Bewertungsstufen spezifizieren
- LBM kann spezifizieren, ob auch fehlende SW angegeben werden können
- LBM kann spezifizieren, ob fehlende Aspekte als freie Eingaben möglich sind

Wunschliste für ein Bewertungstool – für LBM (Durchlauf)

- LBM kann Gesamtbearbeitungsstand eines Bewertungsdurchlaufs einsehen (nicht für individuelle SEB)
- LBM kann Stichtag der Datenübertragung in Hauptzielsystem festlegen (*push*)
- LBM kann jederzeit Datenübertragung in Hauptzielsystem veranlassen (*pull*)
- bei der Datenübertragung werden ID der Ressource, vorgeschlagene SW, Bewertungen, Datum der Bewertung & ergänzte SW übermittelt
- LBM kann verschiedene Zielsysteme für verschiedene Elemente der Datenübertragung spezifizieren
[z.B. Bewertungsauswertung, Inhaberschlüssel, Thesaurusmanagement]

Umsetzung eines Teils davon mit Eurospider im DA-3

Kurztitel

Nummer: 1745269002

Titel: **Impact of employee job attitudes on ecological green behavior in hospitality sector / Muhammad**

Vorschläge Status | Rohdaten | Einstellungen #

Filtern Aktualisieren Erweitern

STW

Arbeitsverhalten	zbwase			
Arbeitszufriedenheit	zbwase			
Mitarbeiterbindung	zbwase			
Umweltbewusstsein	zbwase			
Umweltmanagement	zbwase			
Verhalten in Organisationen	zbwase			

GND

Arbeitsverhalten [Sach] @stw-exact

DA-3-Profil: „k10plus“

Tools > Bewertung Einstellungen #

Bewertung abschicken 7/7

Gesamtbewertung

Quelle zbwase

STW

Arbeitsverhalten	zbwase					
Arbeitszufriedenheit	zbwase					
Mitarbeiterbindung	zbwase					
Umweltbewusstsein	zbwase					
Umweltmanagement	zbwase					
Verhalten in Organisationen	zbwase					

Zwischenstand

nach ein paar Testphasen zur Benutzbarkeit:

- seit Frühjahr 2022 sind die Sacherschließenden der ZBW angehalten, nach der Erschließung jedes Dokuments, zu dem es AutoSE-Vorschläge gibt, diese auch über das Bewertungstool im DA-3 zu bewerten
 - → (tendenziell) für alles, was laut Erschließungskonzept der ZBW noch intellektuell erschlossen wird und was AutoSE verschlagworten kann, bekommen wir auch Bewertungen
- AutoSE-Team greift diese Daten ab, aggregiert sie und wertet sie aus

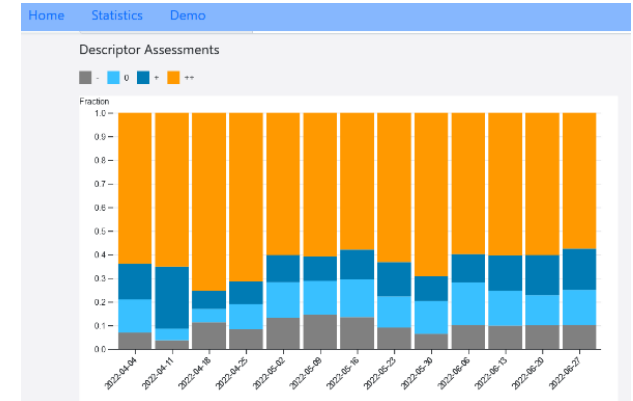


Abgreifen und Auswerten der Bewertungen

Tools > Bewertung		Einstellungen #
Bewertung abschicken		7/7
Gesamtbewertung		
Quelle zbwase		++ + o - X
STW		
Arbeitsverhalten	zbwase	++ + o - X
Arbeitszufriedenheit	zbwase	++ + o - X
Mitarbeiterbindung	zbwase	++ + o - X
Umweltbewusstsein	zbwase	++ + o - X
Umweltmanagement	zbwase	++ + o - X
Verhalten in Organisationen	zbwase	++ + o - X

Bewertungen filtern:
`export-id="DE-206"`
`source="zbwase"`

Erschließung filtern:
`system="stw"`
`type="keyword"`
`df.type in`
`("650", "651", "653")`



DA-3

XML

JSON



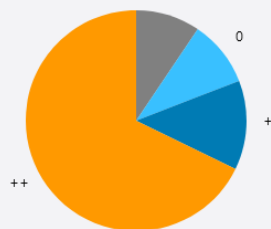
AutoSE
Key-Value
Store

Darstellung
in Web-UI

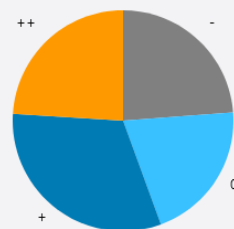
Willkommen bei AutoSE

Bewertungen der letzten 30 Tage

Deskriptorbewertungen

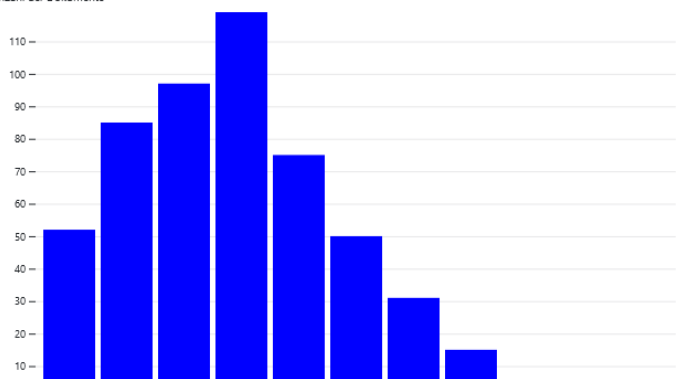


Dokumentbewertungen



Hinzugefügte Deskriptoren

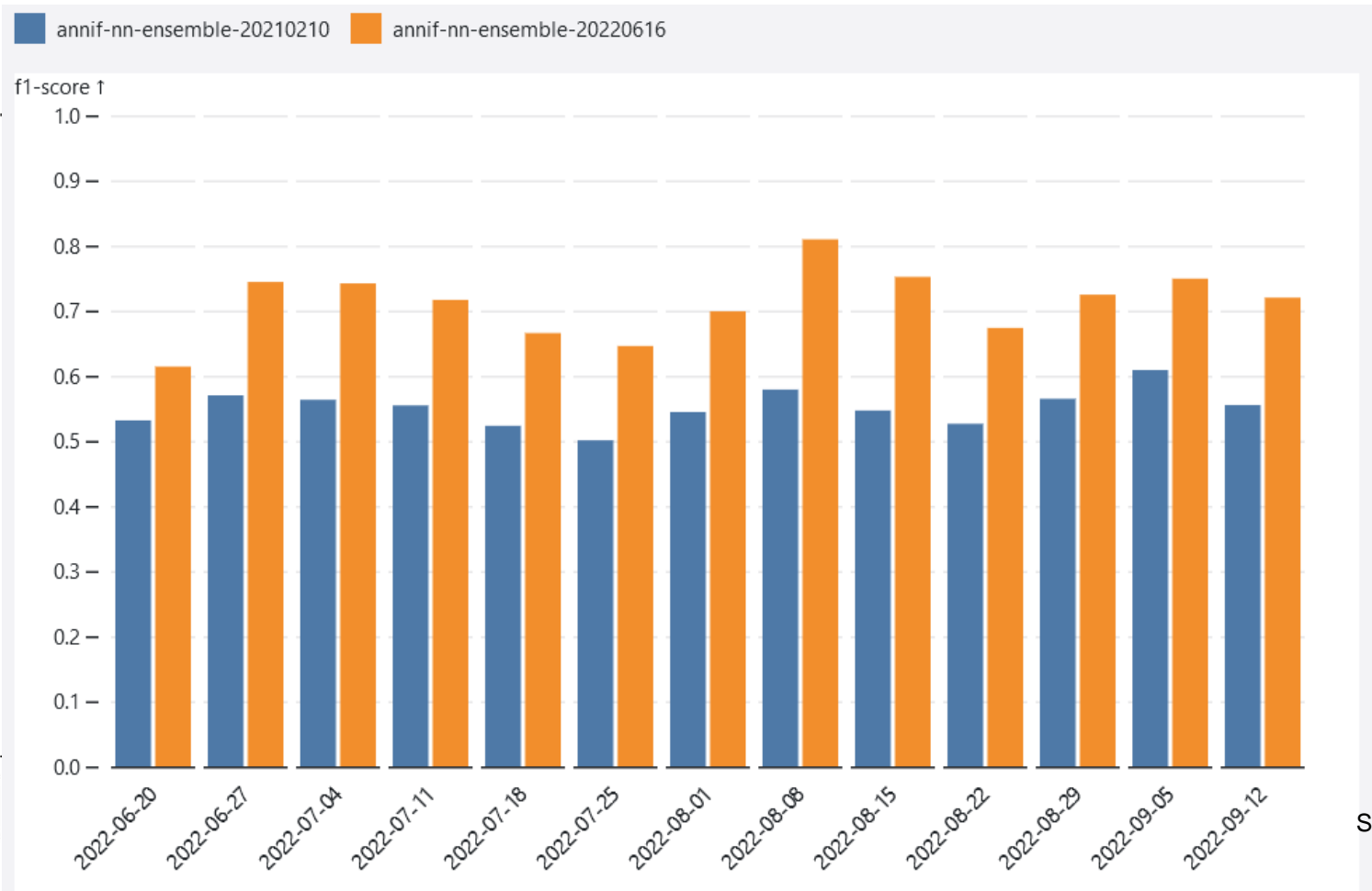
Anzahl der Dokumente



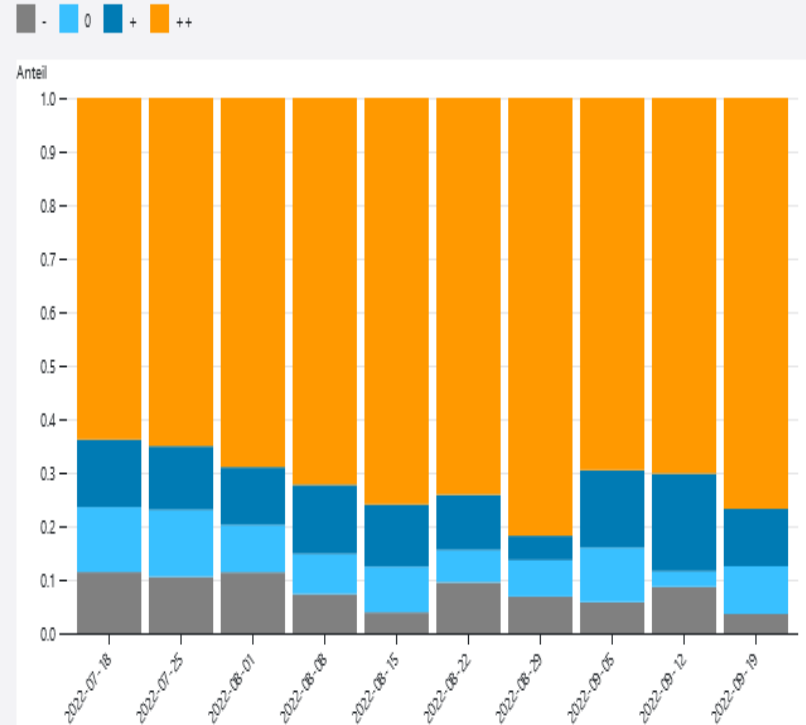
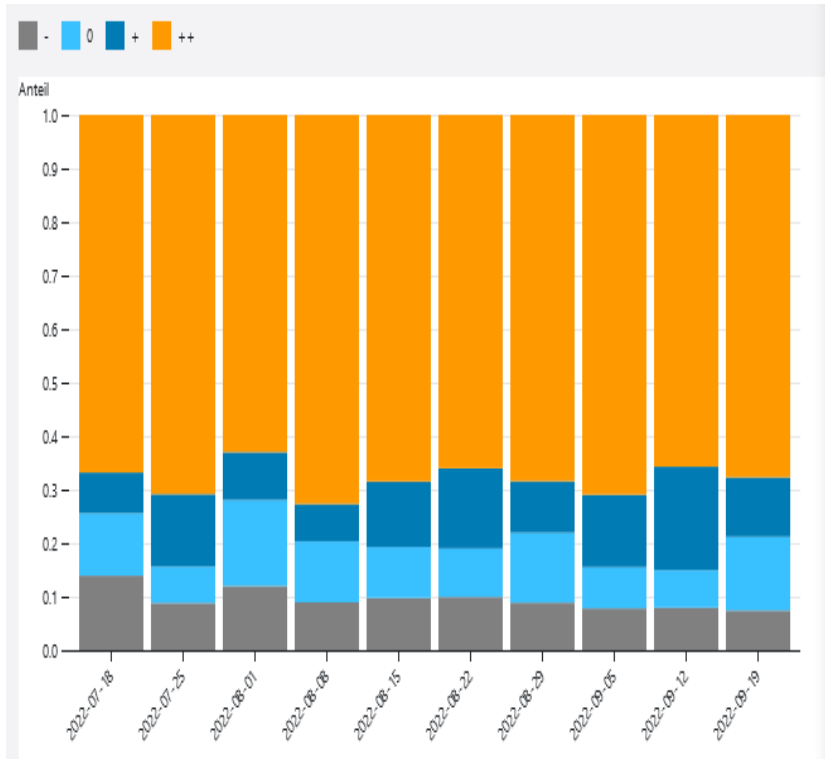
Auswerten der Bewertungen

- Demo Web-UI: [Link \(intern\)](#)
(wenn Zeit und wenn gewünscht, sonst siehe Screenshots)
- eine erste Beobachtung aus dem letzten A/B-Test:
 - „F1-Wert hat sich spektakulär verbessert, gestuftes Feedback ist aber nicht soviel besser geworden – Tendenz setzt sich fort: Deskriptorebene verbessert sich, Dokumentenebene weniger“
 - „zukünftige Entwicklungsrichtung könnte sein: nicht auf F1 optimieren, sondern *recall* höher gewichten (z.B. 1,5fach)“

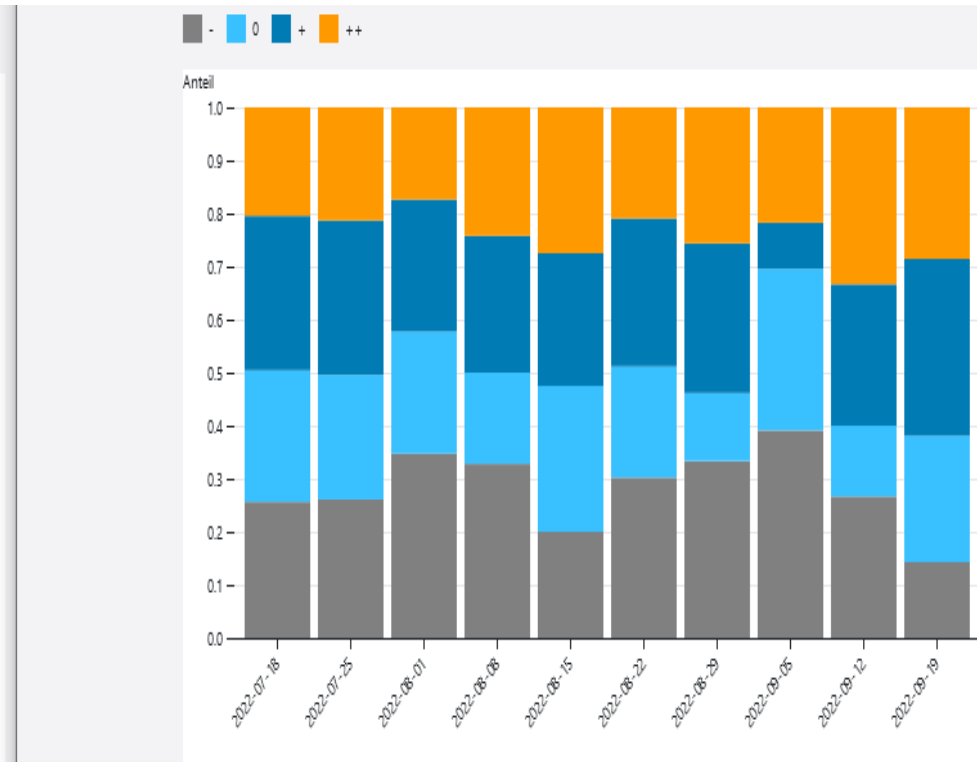
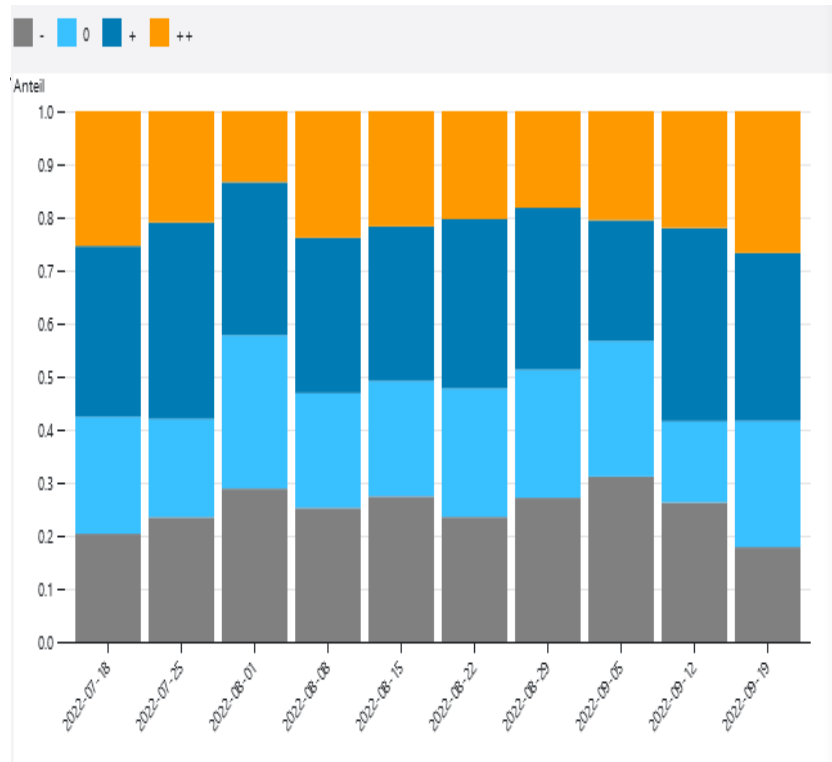
F1-Werte aus binärem Feedback über DA-3 (A/B-Test)



Gestuftes Feedback Deskriptoren über DA-3 (A/B-Test *)



Gestuftes Feedback Deskriptoren über DA-3 (A/B-Test *)



Fazit

- keine „Comfort“-Features wie etwa
 - SEB: Warnungen/Fehlermeldungen; Tastaturkürzel; Kommentare; Überblick über bisher Bewertetes; Wiederaufnahme unvollständiger Bewertungen bei späteren Aufrufen; etc.
 - LBM: variable Anzahl Stufen; Hilfs- und Anweisungstexte; Steuerung Zuweisung zu bewertender Dokumente – „Experiment-Modus“; etc.
- gezielte (A/B-)Experimente nur schwer umzusetzen – ggf. tlw. möglich über ausgefeiltere Steuerung im eigenen Backend (Zuweisung PPN-Listen, nachträgliches Filtern, etc.)
- → eher geeignet für Standortbestimmung durch Einsammeln im Produktivbetrieb
- mit den vorhandenen Basisfeatures lassen sich bereits interessante Daten sammeln



Herzlichen Dank!

Weitere Vorträge und Publikationen zu AutoSE:

siehe Hinweise unten auf der Seite

<https://www.zbw.eu/de/ueber-uns/arbeitsschwerpunkte/automatisierung-der-erschliessung/>

Kontakt: {a.kasprzik,c.bartz,autose}@zbw.eu
