

Qualität ist Teamwork

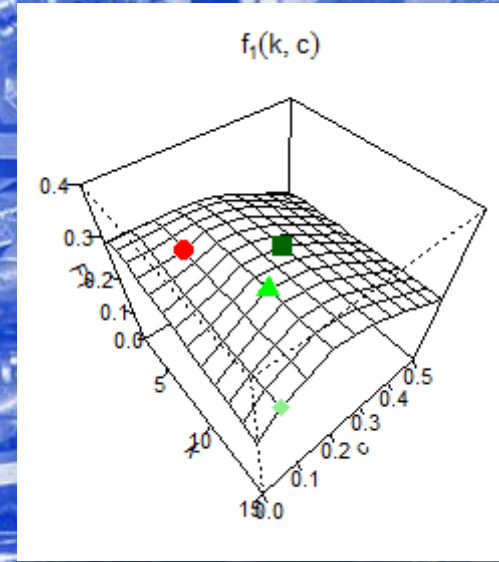
Maschinelles Lernen,
Softwareprojektmanagement
und gemeinsame Verantwortung
für den Erfolg automatischer Indexierung

Martin Toepfer, ZBW – Leibniz-Informationszentrum Wirtschaft, Kiel

Impulsvortrag beim Hands-on Lab: Bewertung maschineller Indexierung - "Qualität ist kein Zufall!",
107. Deutscher Bibliothekartag 2018, Berlin, 14.06.2018

Methoden sind flexibel

Datengetriebenes Vorgehen ist sinnvoll. Module sollten nach Funktionalitätskriterien ausgewählt und entwickelt werden



Künstliche Intelligenz:
Maschinelles Lernen = Optimierung
→ **Optimierungskriterien** festlegen!

Qualität erfordert Balance!

precision

**viele Dokumente
erschließen**

spezifische Inhaltsdarstellung

Verarbeitung ohne Volltext

...

recall

**ausschließlich qualitativ
hochwertige Indexate**

kompakte Inhaltsdarstellung

Vollständige Inhaltsbeschreibung

...

Q i Z
u s u
a t f
l k l
i k l
t e l
ä i !
t n

Automatische Indexierung =
Softwareprojektmanagement!

► Qualität ist eine **Anforderung**



Planen

Qualität: Welche Kriterien sind relevant?
Wie werden sie gemessen?

Ressourcen: Personal, etc.

Zeit: Was soll zuerst umgesetzt werden?

Projektkontext

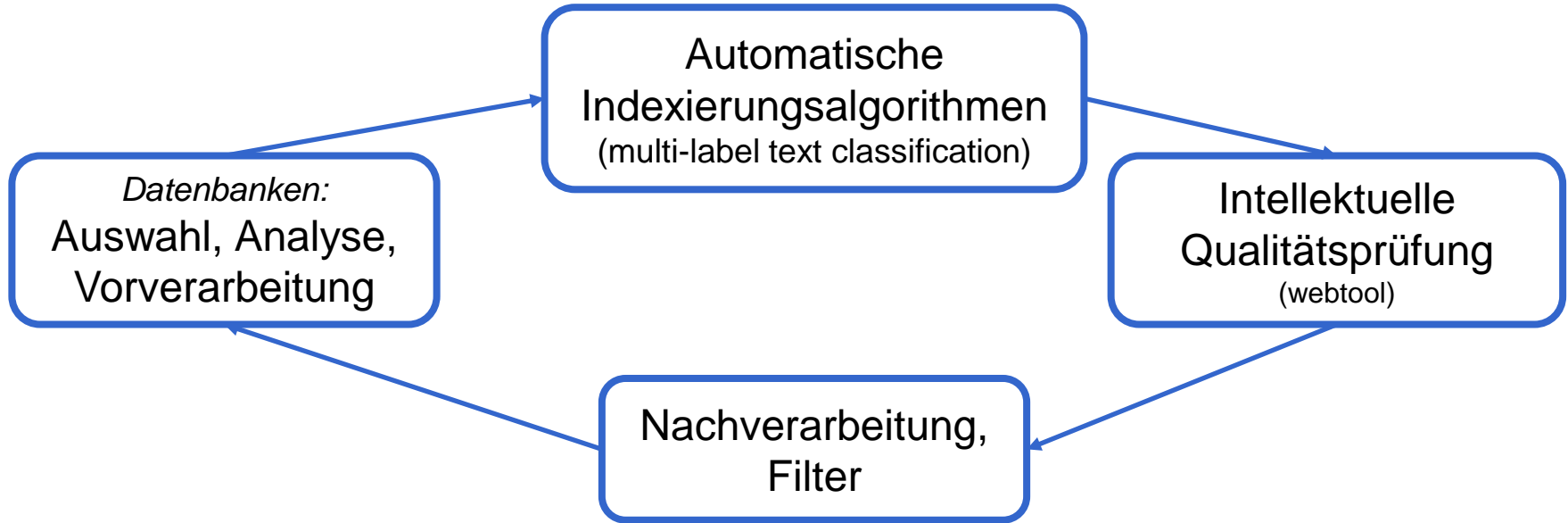
seit 2014 Aufbau einer eigenen
Infrastruktur für die automatische
Indexierung

Etappenziel erreicht

Seit 2017 >130.000 automatisch
erzeugte Indexate in **ECONBIZ**

Find Economic Literature.

Überall ist... **Software**



Überall ist... Qualitätsprüfung

Softwaretests
(unit tests)

Intellektuelle
Prüfung
(Stichproben,
ausführliche
Bewertung)



Automatische
Evaluation
(precision, recall,
F1, etc.)

Für sehr große Datenmengen:
Neu: **automatische
Qualitätsschätzung**

Siehe *Toepfer & Seifert 2018**,
will be presented at **TPDL-2018**,
Porto, Portugal, September 10-13

Wir bauen aufeinander auf!

Sprechen Sie mich gerne an:
m.toepfer@zbw.eu

für Einblicke in Software & Verfahren:

- statistische Lernverfahren
- Qualitätssicherung (releasetool)
- Wörterbuchverfahren (stwfsa)

z.B.: <https://github.com/zbw/stwfsa>



Qualität ist (auch) Softwareprojektmanagement

Forschung ▶ ▶ Praxis



▶ **Teamarbeit** ist gefragt!
Wo soll die Reise hingehen?

m.toepfer@zbw.eu

Bildquellen (CC0-Lizenzen)

<https://pixabay.com/photo-261208/> (Folie 2: „Werkzeuge“)

<https://pixabay.com/photo-504140/> (Folie 3: „Balancieren“)

<https://pixabay.com/photo-3431327/> (Folie 4: „Baustelle“)

<https://pixabay.com/photo-3301648/> (Folie 4: „Würfel“)

<https://pixabay.com/photo-620397/> (Folie 5: „Sanduhr“)

<https://pixabay.com/photo-79593/> (Folie 6: „Schwimmer“)

<https://pixabay.com/photo-1080585/> (Folie 8: „Sicherungskasten“)

<https://pixabay.com/photo-987345/> (Folie 8: „Prüfanzeige“)

<https://pixabay.com/photo-2026366/> (Folie 8: „Sommelier“)

<https://pixabay.com/photo-1336120/> (Folie 9: „Menschenpyramide“)

<https://pixabay.com/photo-3014142/> (Folie 10: „Chemikerin“)

<https://pixabay.com/photo-2791671/> (Folie 10: „Maschine“)

<https://pixabay.com/photo-3152585/> (Folie 10: „Hände“)

BACKUP SLIDES