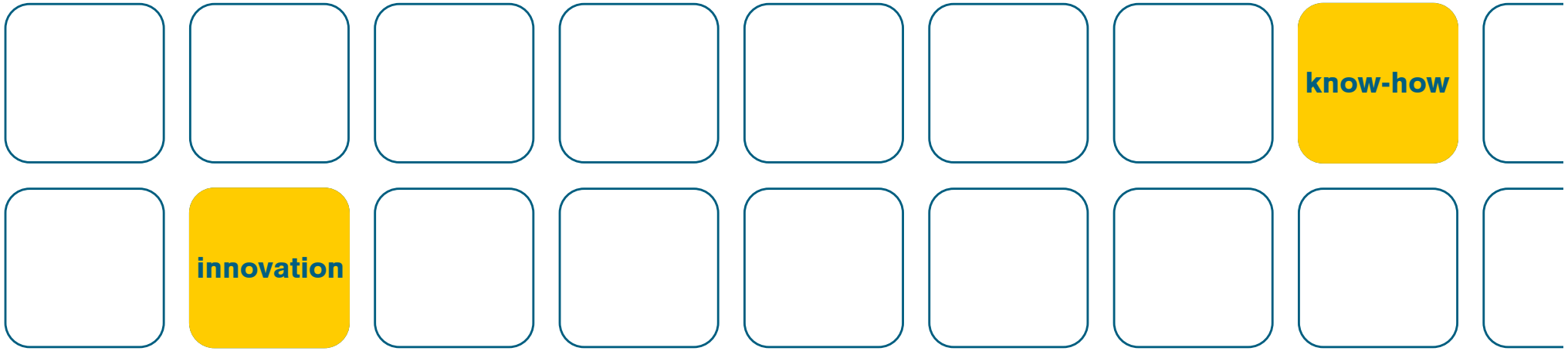


Mögliche zukünftige Formen der Sacherschliessung

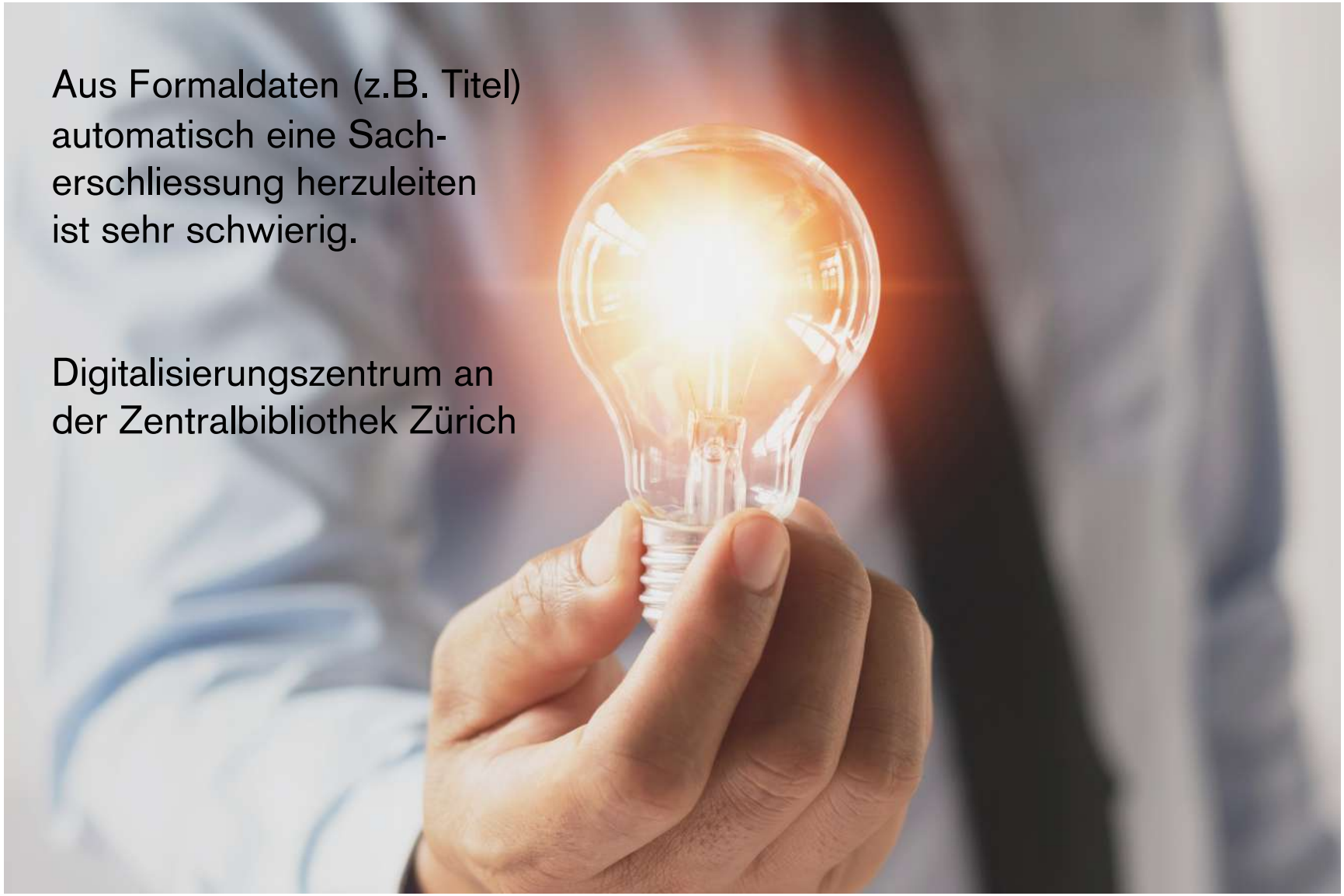
Dritter Stuttgarter Workshop Computerunterstützte Inhaltserschließung

7.11.2019, Peter Schäuble

eurospider
relevancy retrieval

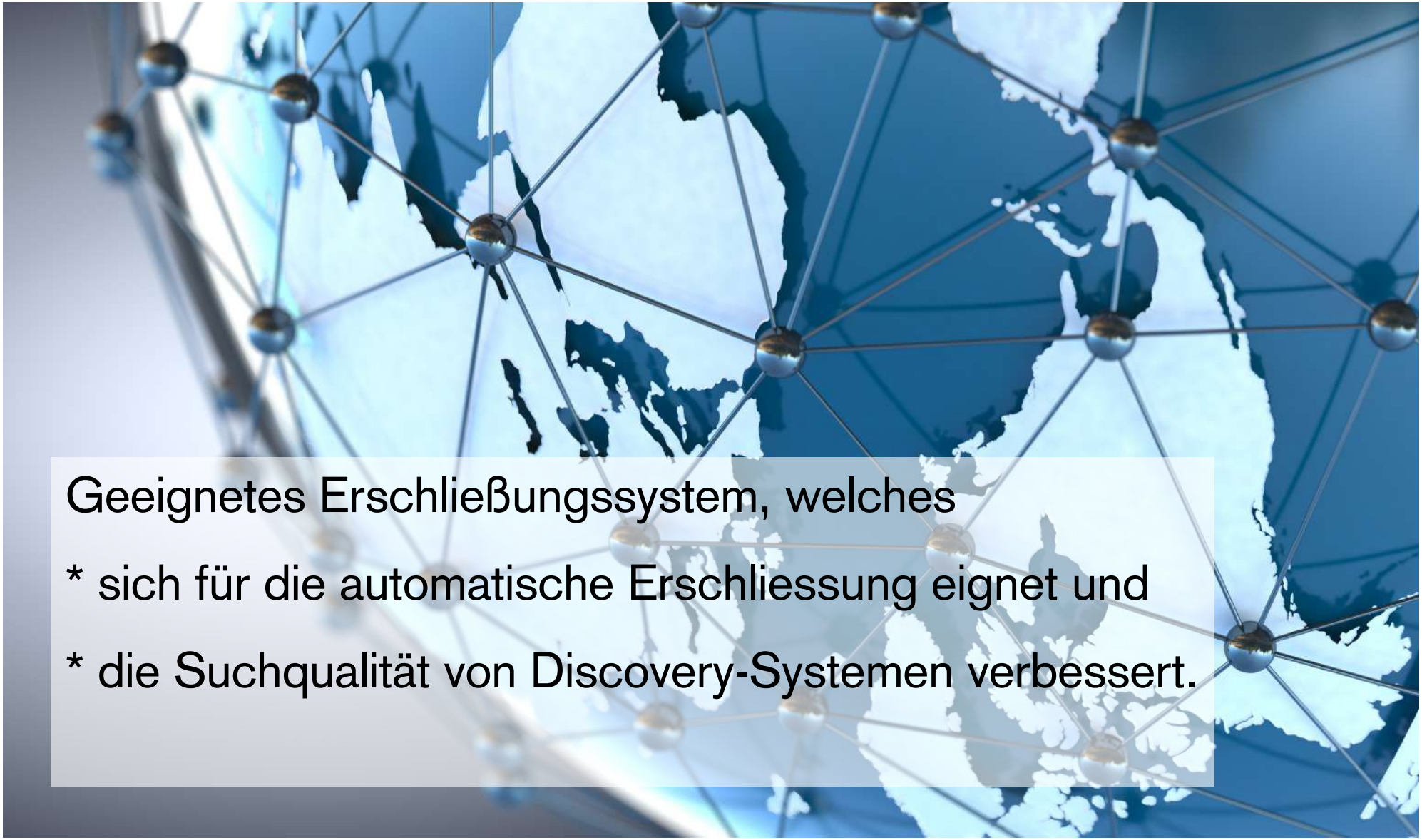
solution

Idee der automatisierten Sacherschliessung vor 10 Jahren am 8.5.2009




Vermutete Erfolgs-
faktoren, welche die
Sacherschliessung
weiterbringen
Erschliessungssystem
Trainingsdaten
Organisationsformen

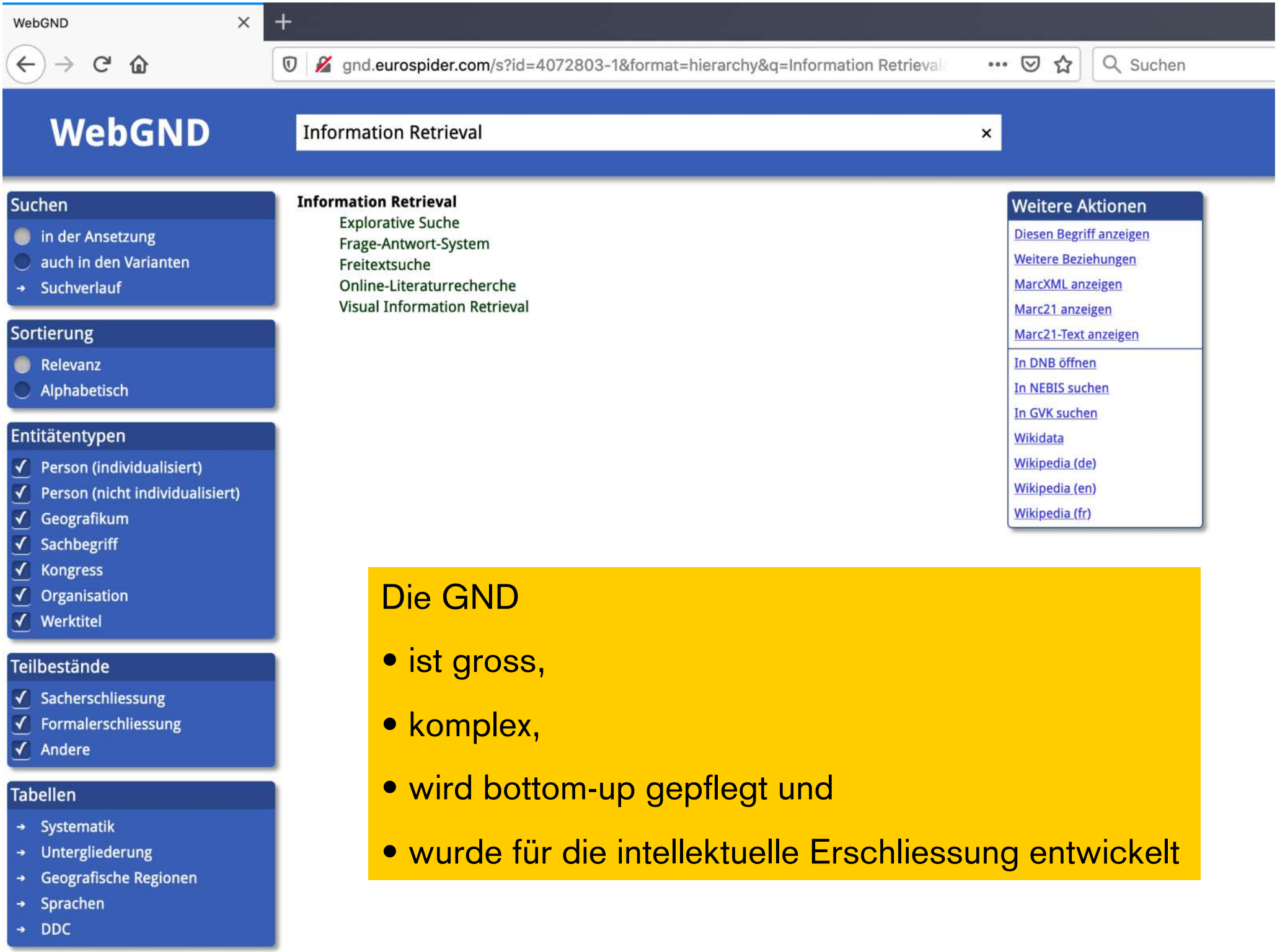
Ein neues Erschliessungssystem?



Geeignetes Erschließungssystem, welches

* sich für die automatische Erschliessung eignet und

* die Suchqualität von Discovery-Systemen verbessert.

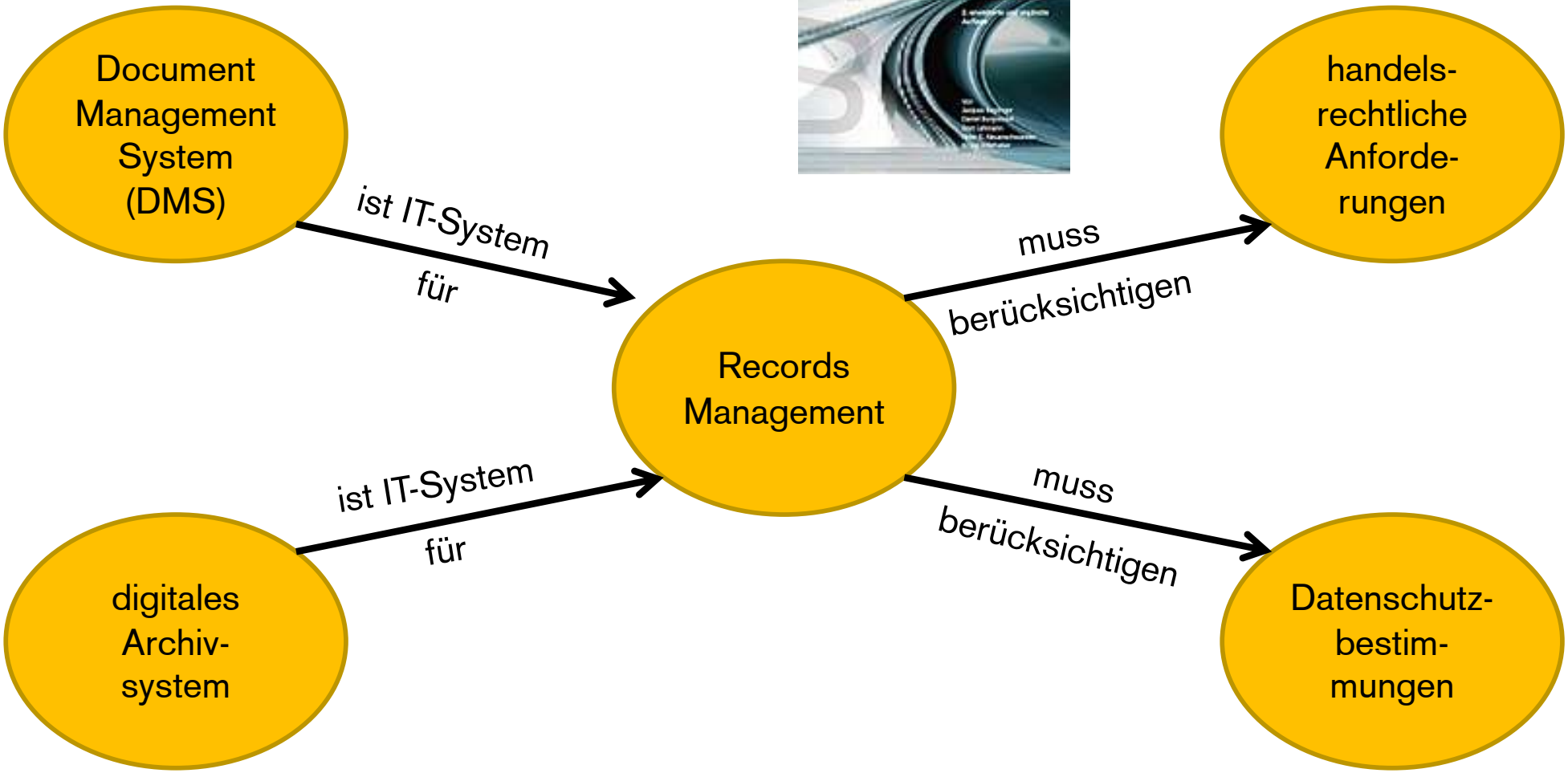


Die GND

- ist gross,
- komplex,
- wird bottom-up gepflegt und
- wurde für die intellektuelle Erschliessung entwickelt

Resource Description Framework (RDF) mit S-P-O

Die vielen Freiheitsgrade und die hohe Komplexität erinnern an das *5th Generation Project* in Japan (1982) und an das amerikanische Gegenprojekt CYC



https://en.wikipedia.org/wiki/Fifth_generation_computer

https://en.wikipedia.org/wiki/Cyc

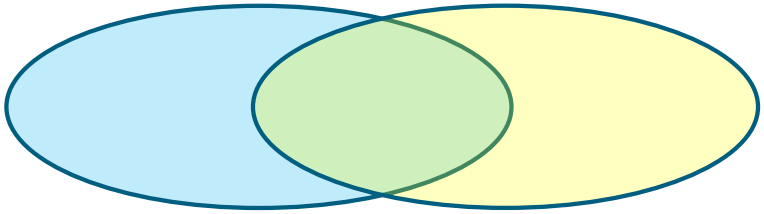
DDC Kurznotationen als neues Erschliessungssystem?

DDC Kurznotationen (DNB Jahresbericht 2018):

- 2018 sind fast 32.000 Monografien und 251.000 Artikel mit DDC Kurznotationen erschlossen worden
- Die klassifikatorische Erschließung wird jetzt um DDC Kurznotationen erweitert, also ausgewählte DDC-Notationen, die differenzierter sind als die DDC-Sachgruppen und gewisse statistische Eigenschaften haben (Dokumentenhäufigkeit)

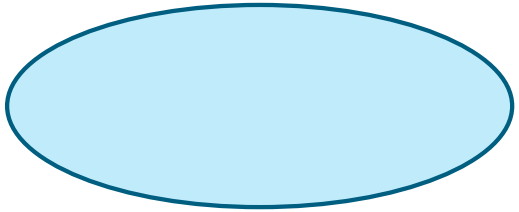
Genügen ungeordnete Mengen mit Notationen ohne Operatoren?

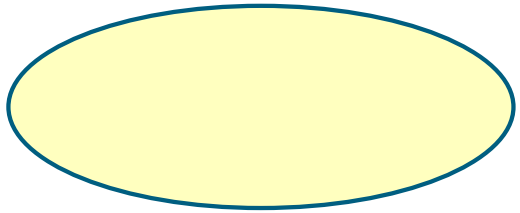
- Dokumentenmanagementsysteme um handelsrechtliche Anforderungen zu erfüllen



Hohe Dokumentenhäufigkeiten alleine garantieren weder Erschliessungs- noch Retrieval-Qualität

- Statistik und Algebra für Ingenieure



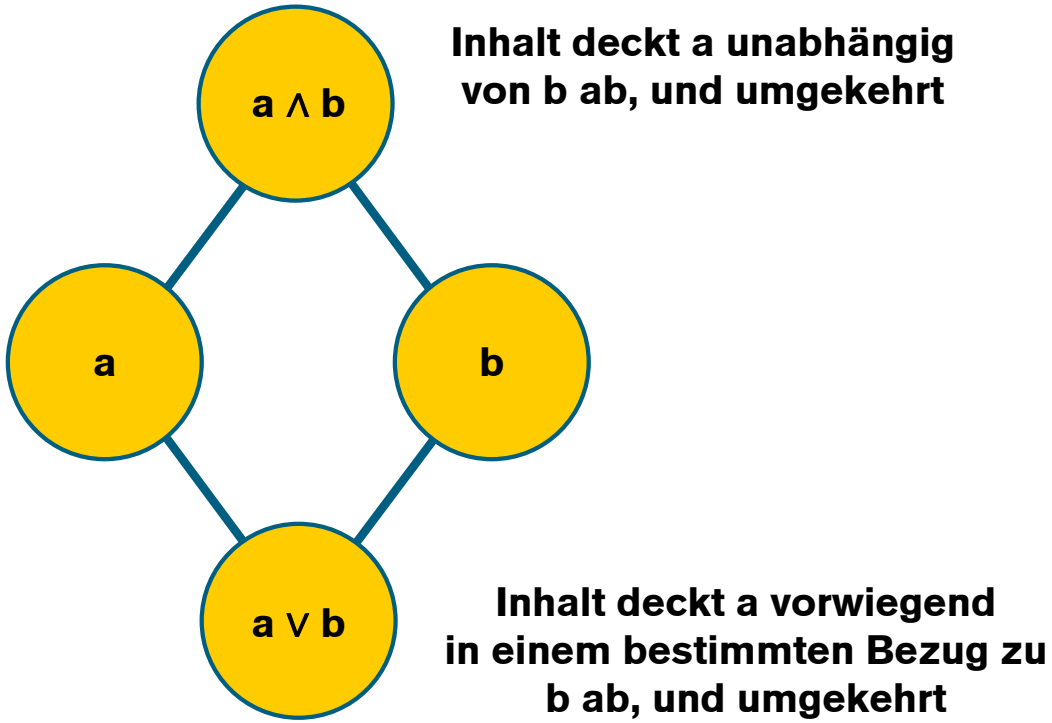


Verbände (Wikipedia)

Ein Verband ist in der Mathematik eine Struktur, die sowohl als Ordnungsstruktur als auch als algebraische Struktur vollständig beschrieben werden kann.

Als Ordnungsstruktur ist ein Verband dadurch gekennzeichnet, dass es zu je zwei Elementen a und b ein **Supremum a ∨ b** gibt, d. h. ein eindeutig bestimmtes kleinstes Element, das größer oder gleich a und b ist,

und umgekehrt ein **Infimum a ∧ b**, ein größtes Element, das kleiner oder gleich a und b ist.



7 Conclusions

The following conclusions are reached from this study:

1. The performance of a classifier depends strongly on the choice of data used for evaluation. Using a seriously problematic collection[8], comparing categorization methods without analyzing collection differences[1], and drawing conclusion bas on the results of flawed experiments[2] raise questions about the validity of some published evaluation se problems need to be addressed to clarify of the confusions among researchers, and to prevent th on of similar mistakes. Providing information and analysis on these problems is a major effort in

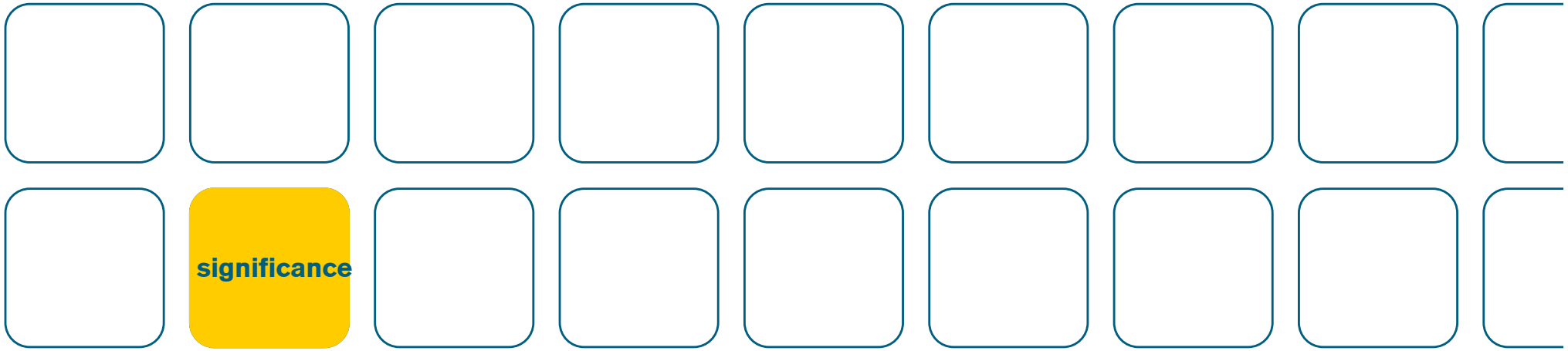
aluations into a
e baseline classi
baseline classifie
iers. Such an in
involving two o
t to standardiz

top perform
had a relative
outperformed WORD, the non-learning m
not as large as previously claimed[1, 2
better than linear models, or that m
on the strengths and weaknesses
sticated
ifferent models requires further research.

4. Scalability of a classifier when the problem size grows by several magnitudes, or when the category space becomes a hundred times denser, has been rarely examined in text categorization evaluations. KNN is the only learning method evaluated on the full set of the OHSUMED categories. Its robustness in scaling up and dealing with harder problems, and its computational efficiency make it the method of choice for approaching very large and noisy categorization problems.

The Performance of a classifier depends strongly on the choice of data used for evaluation.

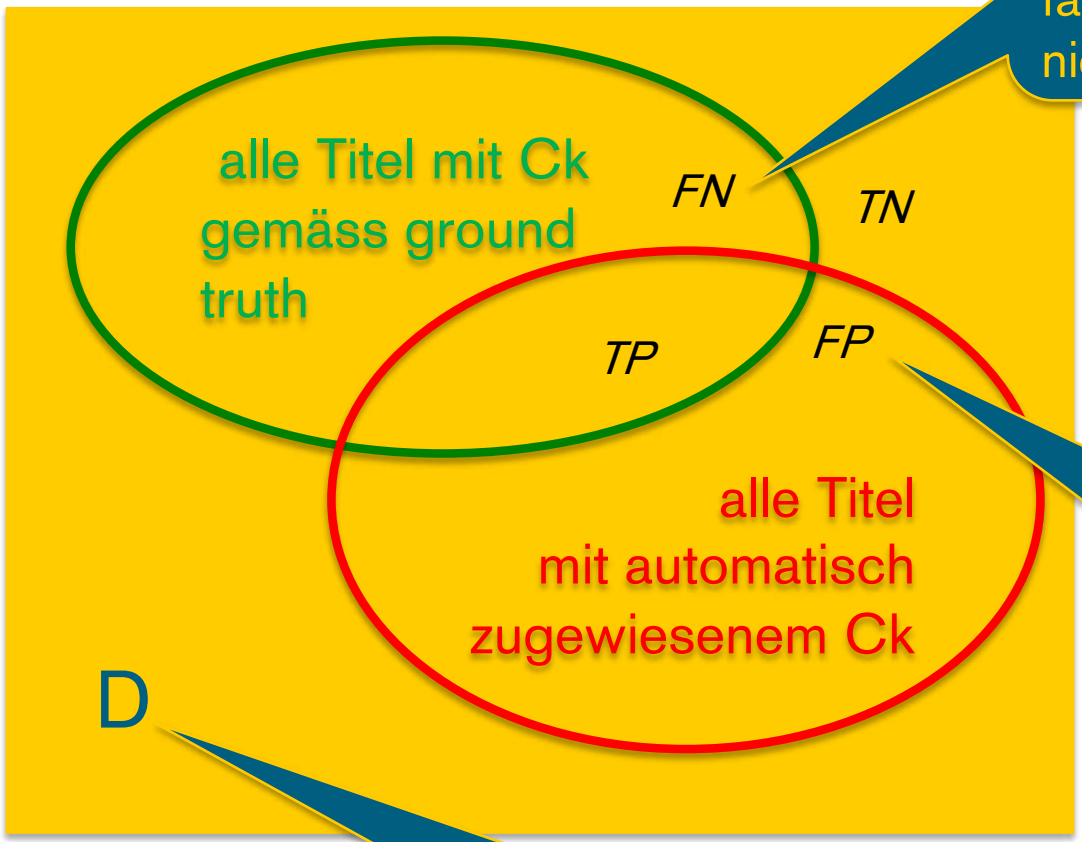
Scalability of a classifier when the problem size grows by several magnitudes, or when the category space becomes hundred times denser, has been rarely examined in text catgeorization evaluation.



Trainings- und Testkollektion für automatische Kategorisierung

eurospider
relevancy retrieval

Evaluation automatischer Zuweisungen von Ck (SW, Notation, etc.)



False Negatives (FN), d.h. Titel, denen fälschlicherweise das SW nicht zugewiesen wurde.

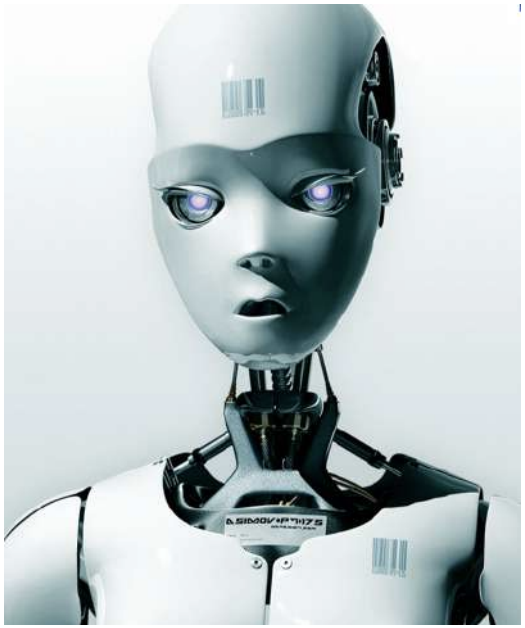
False Positives (FP), d.h. Titel, denen fälschlicherweise das SW unnötig zugewiesen wurde.

Menge der für die Evaluation eingesetzter Titel.

Merkmalsvektor

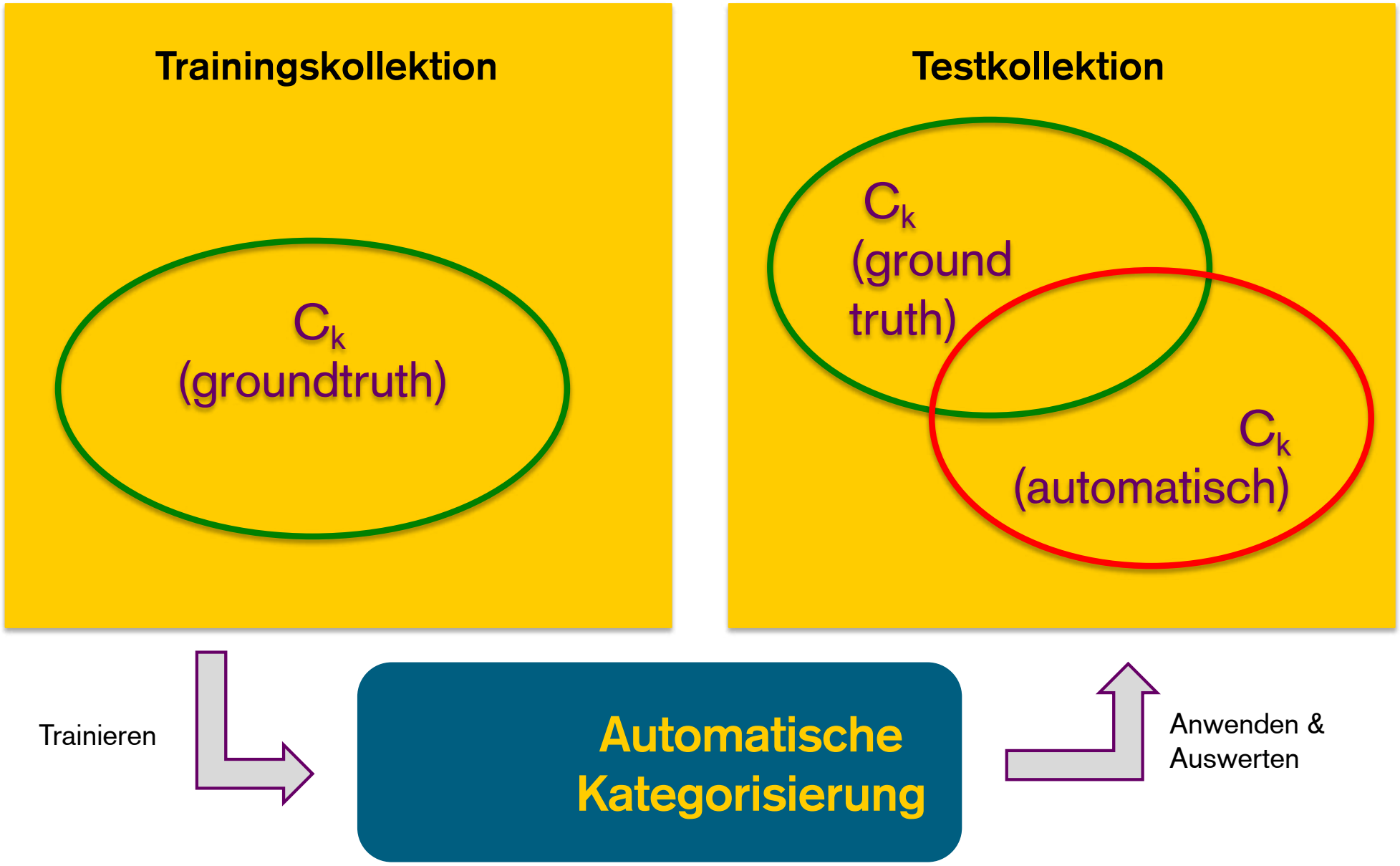
Notation

Merkmalsvektor



Notation

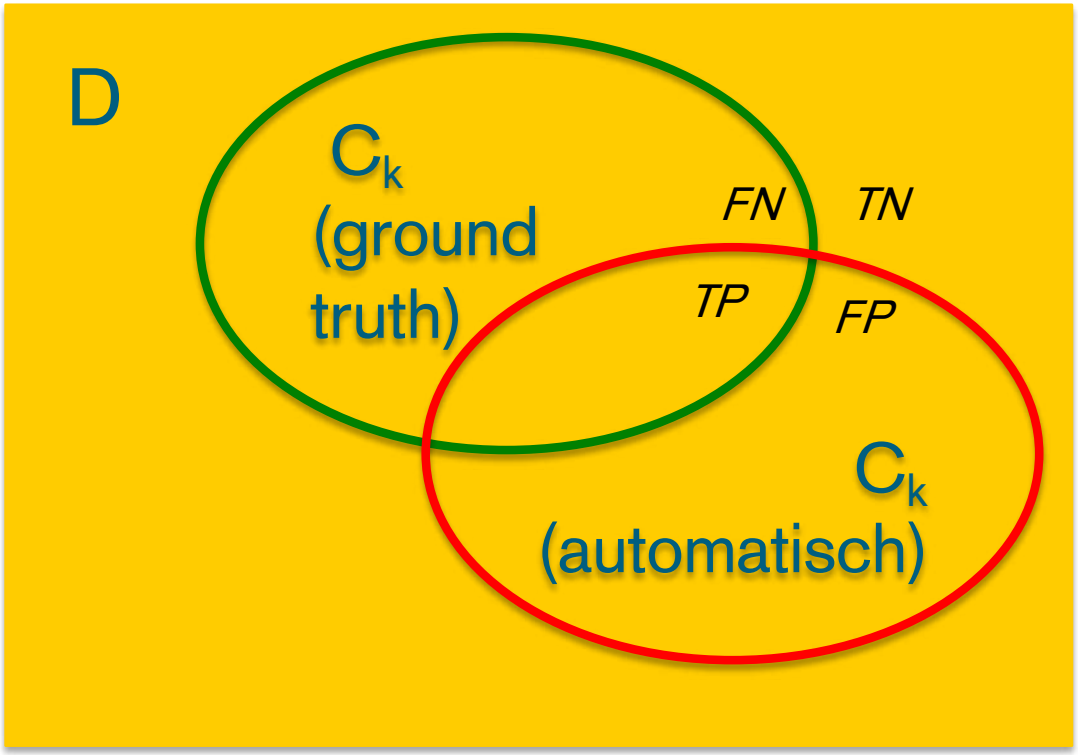
Test- und Trainingskollektion



Qualität der Kategorisierung ermitteln

Die Kategorisierungsqualität kann mit Ausbeute, Präzision, F1, etc. quantifiziert werden.

- Ausbeute (ρ= recall): Wieviele relevante wurden automatisch gefunden?
  ρ = TP / (TP + FN)
- Präzision (π = precision): Wieviele der automatisch gefundenen sind relevant?
  π = TP / (TP + FP)
- F1: Harmonisches Mittel
  F1 = 2 *ρ* π / (ρ+ π)
  F1 = 2 / (1/ ρ + 1/ π)



Automatische Sachgruppenvergabe Ergebnisse Testfall II/3

DEUTSCHE NATIONAL BIBLIOTHEK

	Syst. A	System B	System C	System D
F-Measure$_1$*	0,75	0,68	0,71	0,65
Recall$_1$*	0,75	0,70	0,68	0,63
Precision$_1$*	0.77	0.70	0.76	0.86

	Syst. A	System B	System C	System D
F-Measure$_2$*	0,61	0,54	0,58	0,55
Recall$_2$*	0,88	0,83	0,85	0,79
Precision$_2$*	0,50	0,43	0,46	0,45

	Syst. A	System B	System C	System D
F-Measure$_3$*	0,54	0,48	0,53	0,51
Recall$_3$*	0,91	0,87	0,90	0,86
Precision$_3$*	0,43	0,37	0,41	0,40

*Gewichtetes Mittel über alle SG

know-how

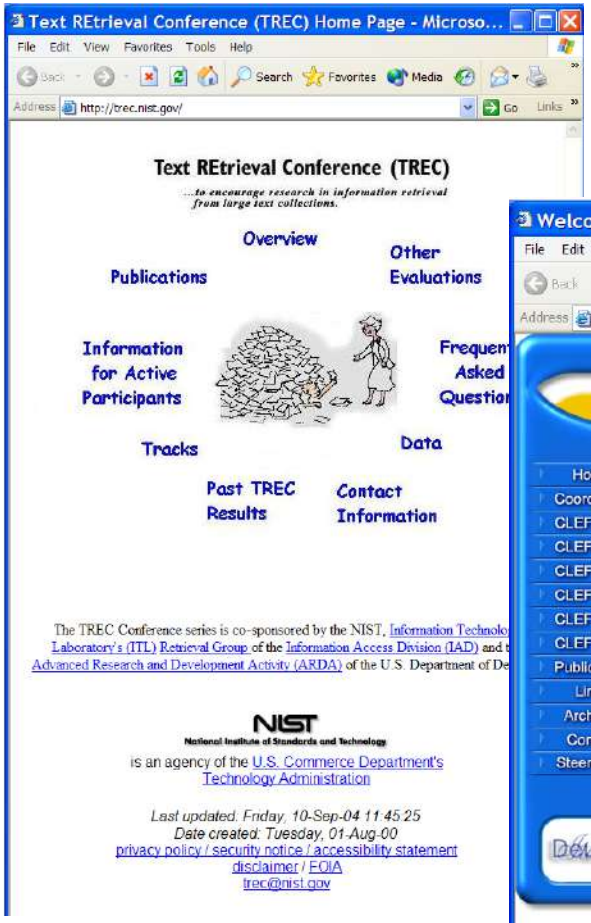
innovation

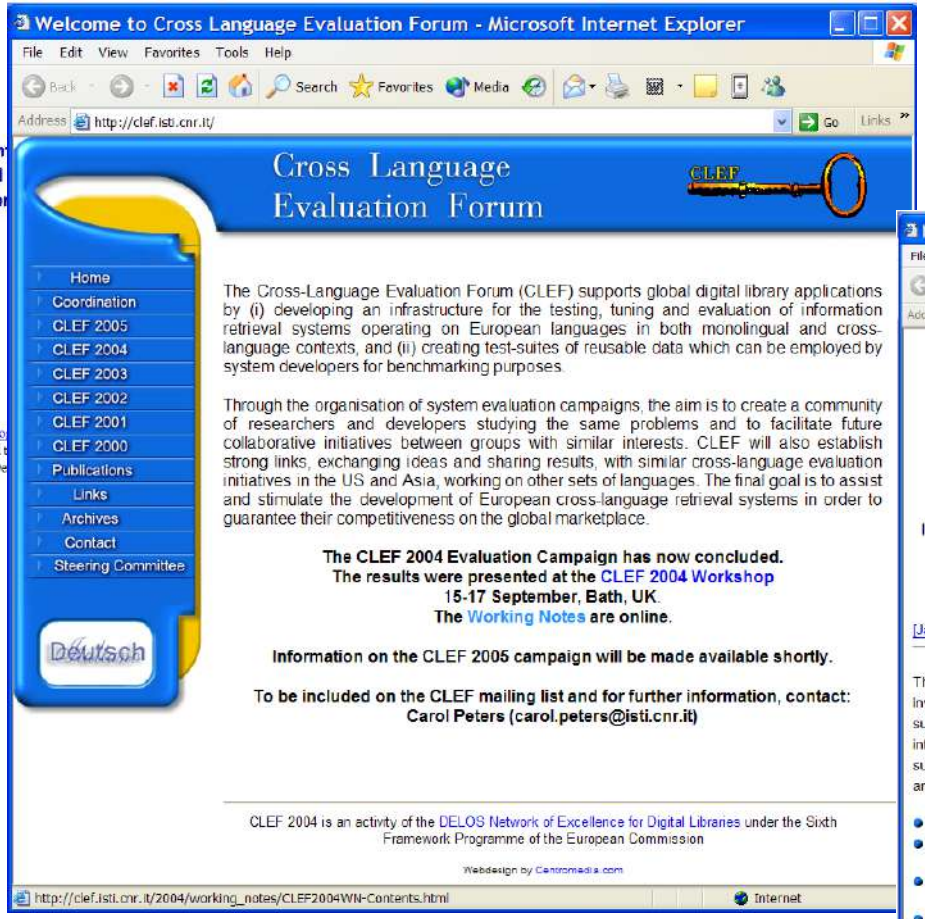
Retrieval Effektivität

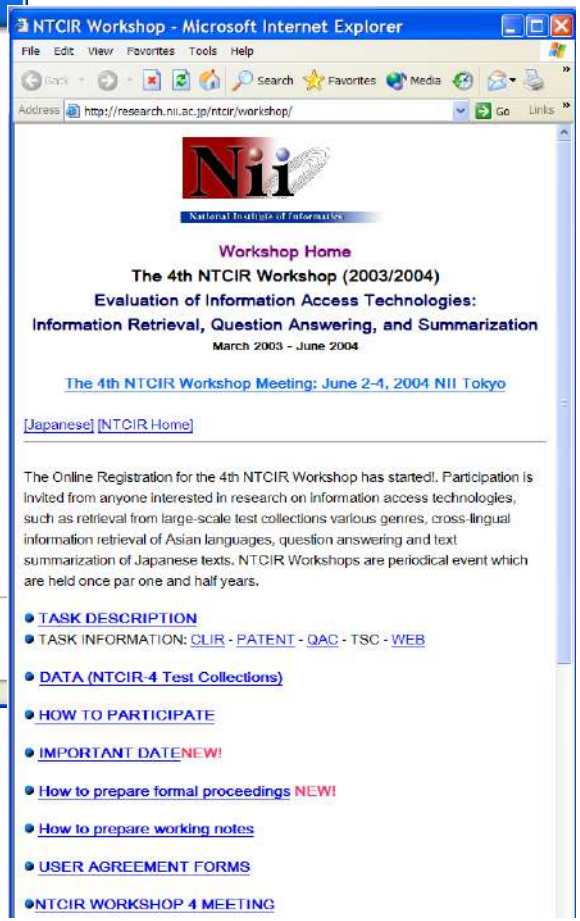
Wird relevante Information effektiv gefunden?

solution

Internationale Evaluationsforen: TREC, CLEF, NTCIR







Nicht angemeldet Diskussionsseite Beiträge Benutzerkonto erstellen Anmelden

WIKIPEDIA
Die freie Enzyklopädie

Artikel Diskussion Lesen Bearbeiten Quelltext bearbeiten Mehr Wikipedia durchsuchen

Hauptseite
Themenportale
Zufälliger Artikel

Mitmachen
Artikel verbessern
Neuen Artikel anlegen
Autorenportal
Hilfe
Letzte Änderungen
Kontakt
Spenden

Werkzeuge
Links auf diese Seite
Änderungen an verlinkten Seiten
Spezialseiten

Cross-Language Evaluation Forum

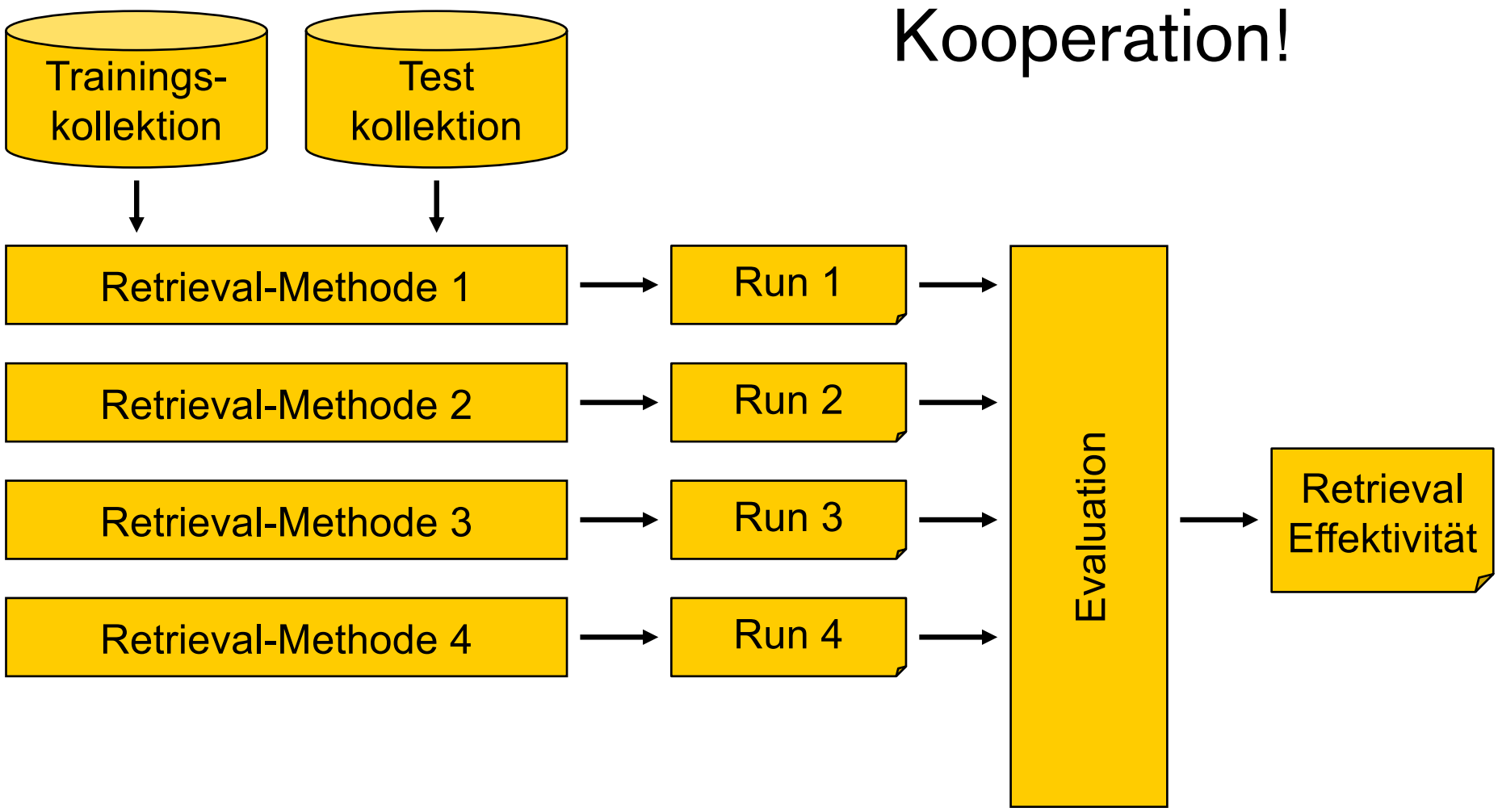
Das **Cross-Language Evaluation Forum** (*kurz*: **CLEF**) ist aus der TREC-Aufgabe Cross-Language Information Retrieval (CLIR) entstanden, welches sich hauptsächlich mit dem Cross-Language Information Retrieval europäischer Sprachen befasste.

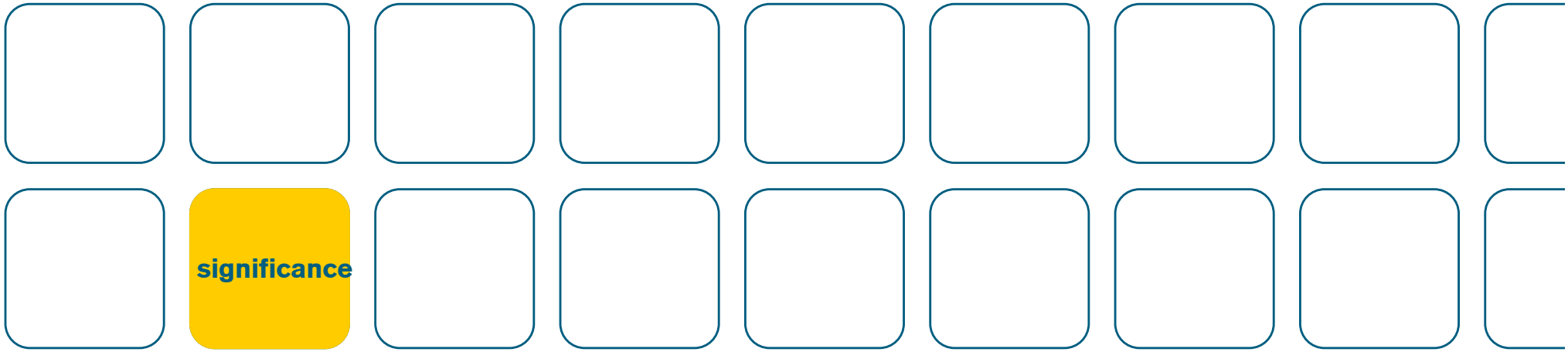
CLEF ist mittlerweile ein eigenständiges EU-Projekt und bietet eine Plattform zur Evaluierung und Verbesserung von Information-Retrieval-Systemen für europäische Sprachen.

Die seit 2000 von CLEF jährlich organisierten System-Evaluations-Kampagnen sollen die Zusammenarbeit von Forschern und Entwicklern fördern und somit zukünftige Initiativen zur Zusammenarbeit von Gruppen mit ähnlichen Interessen vereinfachen und fördern. Es geht hierbei darum, Nutzeranfragen, die in einer beliebigen europäischen Sprache gestellt werden, in beliebigsprachigen Dokumentmengen abzuarbeiten und eine nach Relevanz geordnete Ergebnismenge zu erhalten, die auf diese Frage eine Antwort darstellt. Auch einsprachiges Information Retrieval stellt einen Schwerpunkt der Evaluierung dar, ist jedoch vor allem für Teams vorgesehen, die das erste Mal an der Kampagne teilnehmen. Es bestehen auch Kooperationen mit ähnlichen, anderssprachigen Initiativen aus den USA und Asien.

Das eigentliche Ziel ist, die Entwicklung der europäischen Cross-Language Retrieval Systeme zu unterstützen und anzuregen, damit ihre Wettbewerbsfähigkeit auf dem Weltmarkt gesichert ist.

Wie funktioniert ein Evaluationsforum?





Zusammenfassung

eurospider
relevancy retrieval

Sacherschliessung in 10 Jahren

Aufgrund der gemachten Erfahrungen bei der Entwicklung des Digitalen Assistenten und den sich stets weiter entwickelnden Anforderungswünschen an das System haben wir in einigen Gesprächen überlegt, wie wir insgesamt die Situation, um bessere Ergebnisse bei der automatisierten Inhaltserschliessung zu bekommen, verbessern können.

Wir haben dabei zunächst drei Erfolgsfaktoren identifiziert.

1. Innovative Organisationsformen, um Skaleneffekte zu erreichen (DINI Vortrag)
2. geeignetes Erschließungssystem, welches
   - für die automatische Erschliessung geeignet ist (weniger komplex als die GND) und
   - die Suchqualität von Discovery-Systemen signifikant verbessert (also detaillierter als die DNB Sachgruppen ist)
3. qualitativ und quantitativ geeignete Trainingsdaten, um die automatische Erschliessung und die Suchqualität von Discovery-Systemen zu evaluieren.

Tradition ermöglicht erst und hemmt dann wieder die Entwickelung der Gedanken

A. Einstein 1924.

Über die spezielle und die allgemeine Relativitätstheorie

(Gemeinverständlich)

Von

A. EINSTEIN