

Semantische Technologien in einem juristischen Verlag

 **Wolters Kluwer**
Deutschland

Christian Dirschl, M.A.
Content Architect



- Vorstellung Wolters Kluwer Deutschland
- Arbeitsdefinition „Semantische Technologien“
- Aufbau von Wissensstrukturen
- Automatische Klassifikation
- Linguistischer Support bei Nutzeranfragen
- Zusammenfassung



Wolters Kluwer Deutschland
23 Standorte in Deutschland
1.000 Mitarbeiter
155 Mio. Umsatz 2008



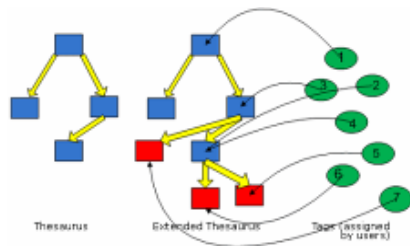
Wolters Kluwer n.v.
Mehr als 40 Länder weltweit
19.300 Mitarbeiter
3.425 Mrd. € Umsatz 2009

Technologien, die auf Basis von
maschinenlesbaren Informationen
aus der juristischen Domäne
automatisch
Wissen in die Kernprozesse
des Unternehmens oder des Kunden einsteuern

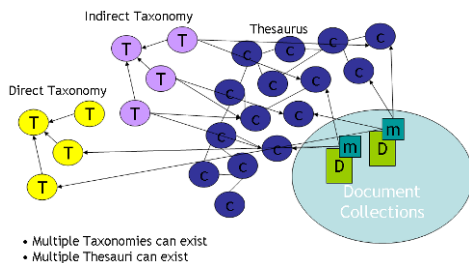
Aufbau von Wissensstrukturen



Zielgruppenspezifische Taxonomien
Navigation und Browsing
Post-Search-Filter
Bis zu 4 Ebenen
Bis zu 800 Einträge



Zielgruppenspezifische Thesauri
Contentanreicherung
Synonyme, Phrasen, etc.
Bis zu 1.500 Kernkonzepte
Bis zu 5.000 Relationen



Mapping Taxonomien und Thesauri

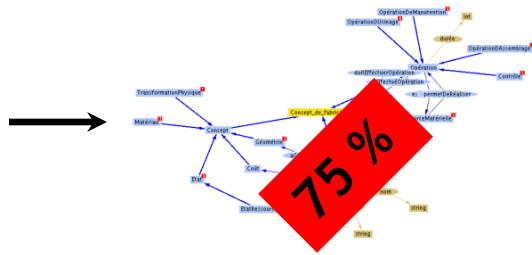
Automatische Klassifikation - Semantische Inhaltserschließung

```

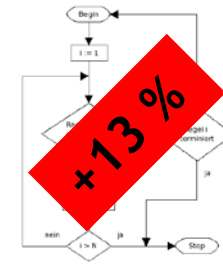
<?xml version="1.0" encoding="ISO-8859-1" ?>
<my_band>
  <bandname>Nero's Delight</bandname>
  <musicians>
    <musician>
      <name>Michael</name>
      <instrument>Trumpet</instrument>
    </musician>
    <musician>
      <name>Boon</name>
      <instrument>Trompete</instrument>
    </musician>
  </musicians>
</my_band>
  
```



XML Entscheidungen



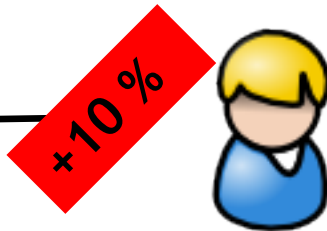
Semantische Klassifikation



Regelbasierter Post-Prozess



Daten klassifiziert in 20 Klassen



Trainingsmaterial

?

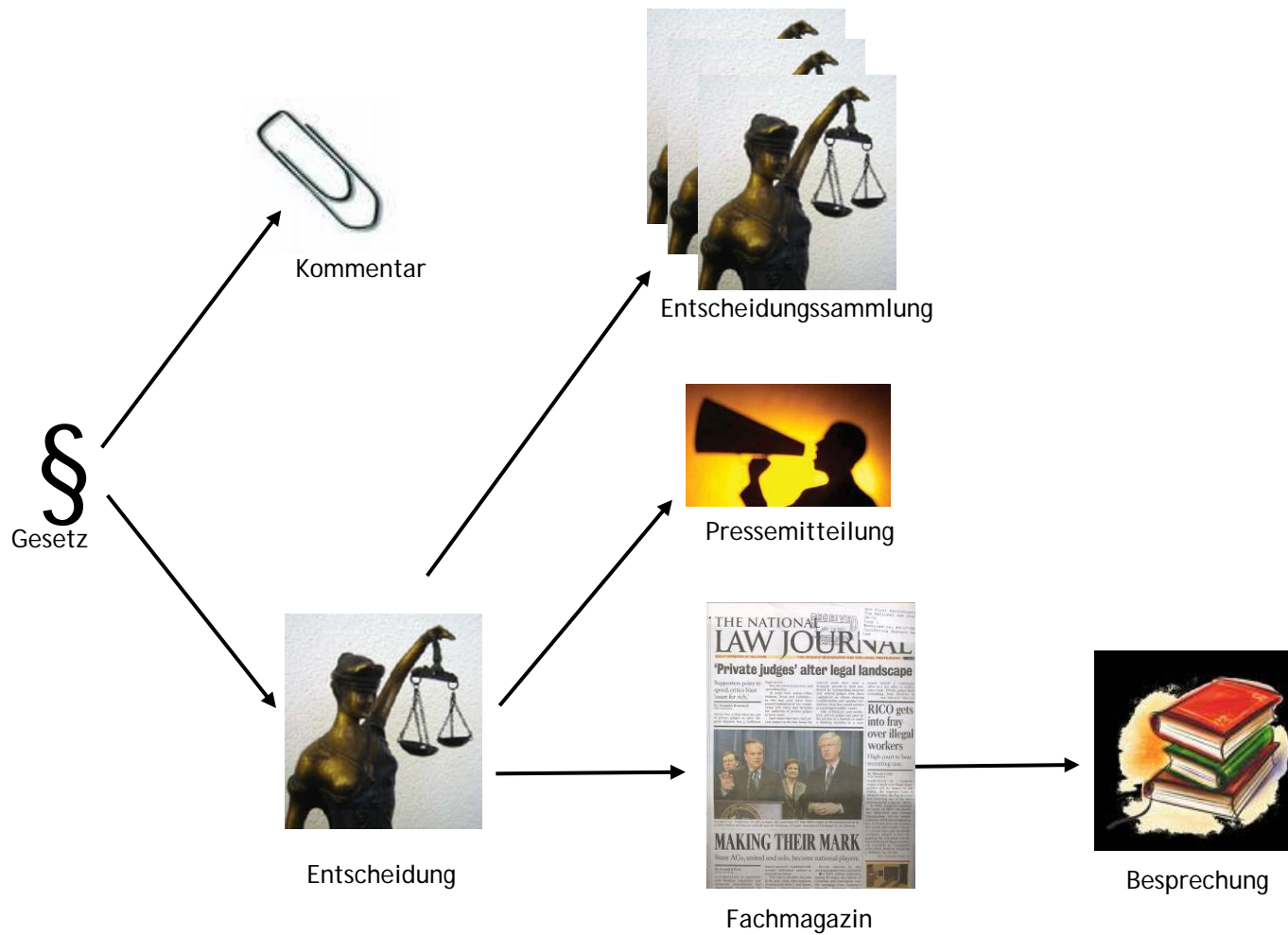
Nicht klassifizierte Daten

Analyse der Autoklassifikation

Klassifikation	Beispiel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

Report

Automatische Klassifikation - Transitive Klassifikation



Linguistische Analyse bei Nutzeranfragen

honor|

- Honorierung (Volltext)
- Verzicht auf Honorar (Volltext)
- Honorar für Übersetzungen (Volltext)
- Honorar für besondere Leistungen (Volltext)
- Prüfbare Honorarrechnung (Volltext)
- Schlüssigkeit der Honorarklage (Volltext)

Dokumente: 0

Rechtsprechung (0) Vorschriften (0) Au

Volltext

Suchargumente: Tabaksteuer

Probieren Sie es mal mit Verbrauchsteuer (2)

Dokumente: 504

Rechtsprechung (387) Vorschriften (19) Aufsätze (48)

Volltext

Suchargumente: "Anspruch auf Schadensersatz"

der Arbeiten für **Haus** 4 zu einem Zerwürfnis zwit
erbrachten Leistungen ab, und zwar nicht nach
Pauschalhonorar, sondern nach den Mindestsätzen d
lediglich noch um die Leistungen für die **Häuser**
Häusern 1 und 2 mit lediglich 12 Wohneinheiten und

Sortierung nach: Relevanz Entscheidungsdatum

Suchvorschläge (Word Wheel)

Semantische Erweiterung
um Oberbegriffe

Domänenspezifische Phrasensuche

Lemmatisierung

Relevance Ranking

- Semantische Technologien sind heute industriell einsetzbar
- Der Einsatz semantischer Technologien ist für jeden notwendig, der mit größeren Mengen Content (eigener oder Web-Content) zum Nutzen seiner Kunden strukturiert umgehen muss
- Eine Out-of-the-box-Lösung wird in den seltensten Fällen alle Requirements abdecken
- Eine intensive Auseinandersetzung mit den Möglichkeiten und den zu Grunde liegenden Konzepten jeglicher Anwendung ist eine Grundvoraussetzung für den Erfolg
- Der kombinierte Einsatz verschiedener Ansätze verspricht den größten operativen Nutzen



Vielen Dank für Ihre Aufmerksamkeit!

Diskussion!?