



Info_Net Passau

**Netzgestützter Wissenstransfer von
Hochschulen in Unternehmen**

Passau, 21.03.2012

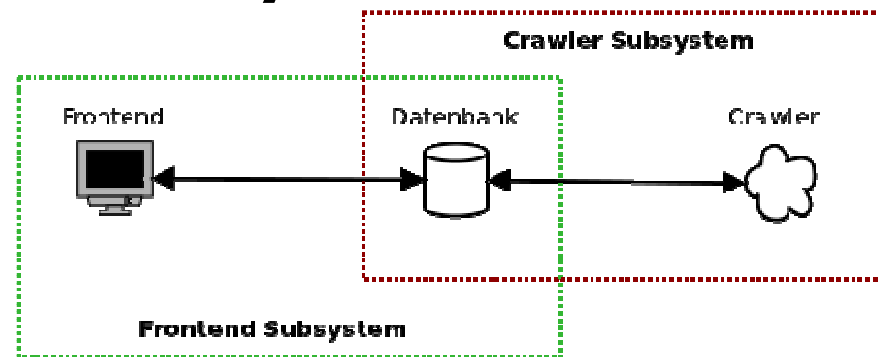


EUROPÄISCHE UNION
Europäischer Sozialfonds



virtuelle
hochschule
bayern

Systembild



Apache, Perl, Postgres, Javascript, Lucene

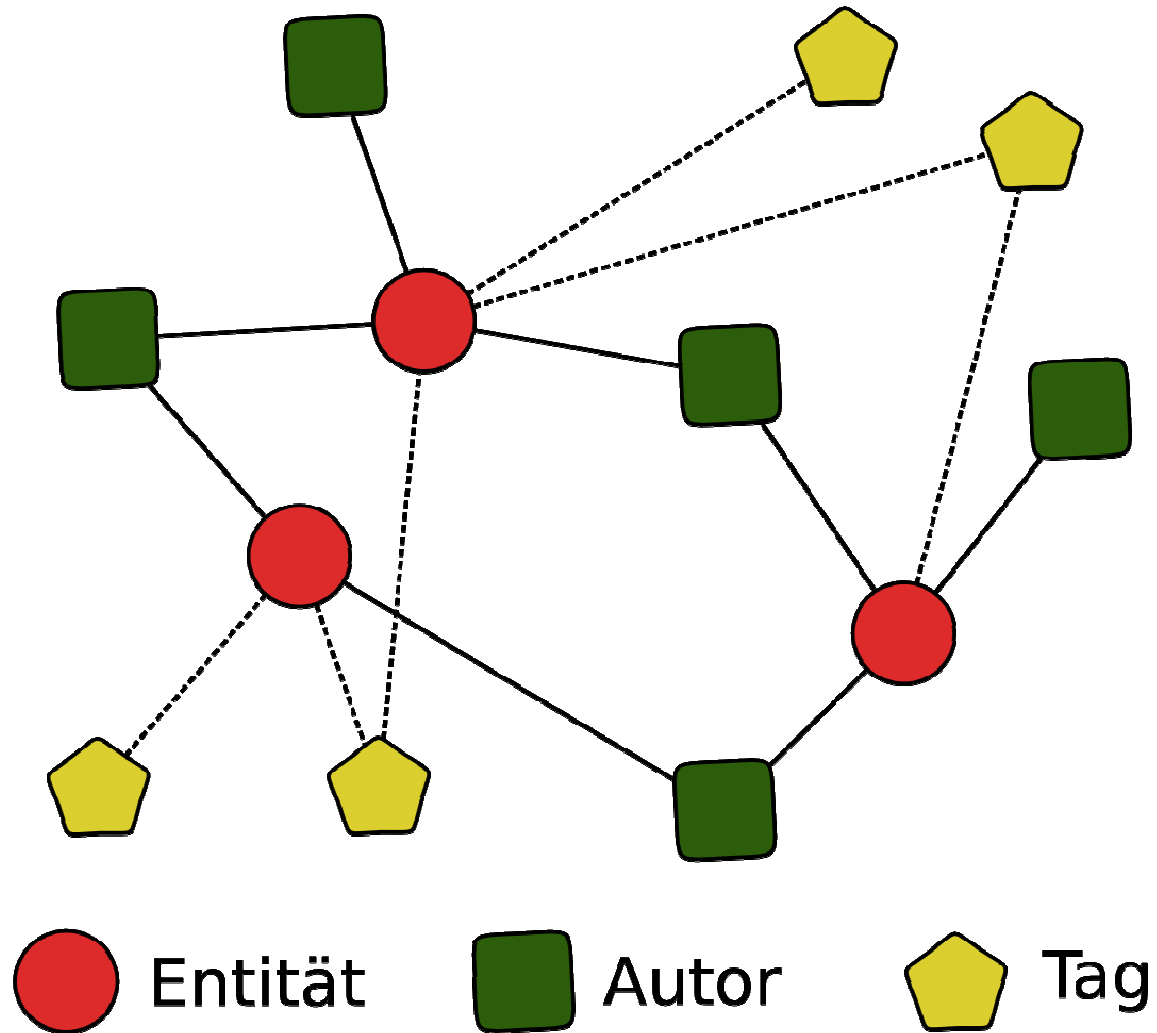
Crawler:

Input/Output: Daten aus verschiedenen Quellen, Web/DB

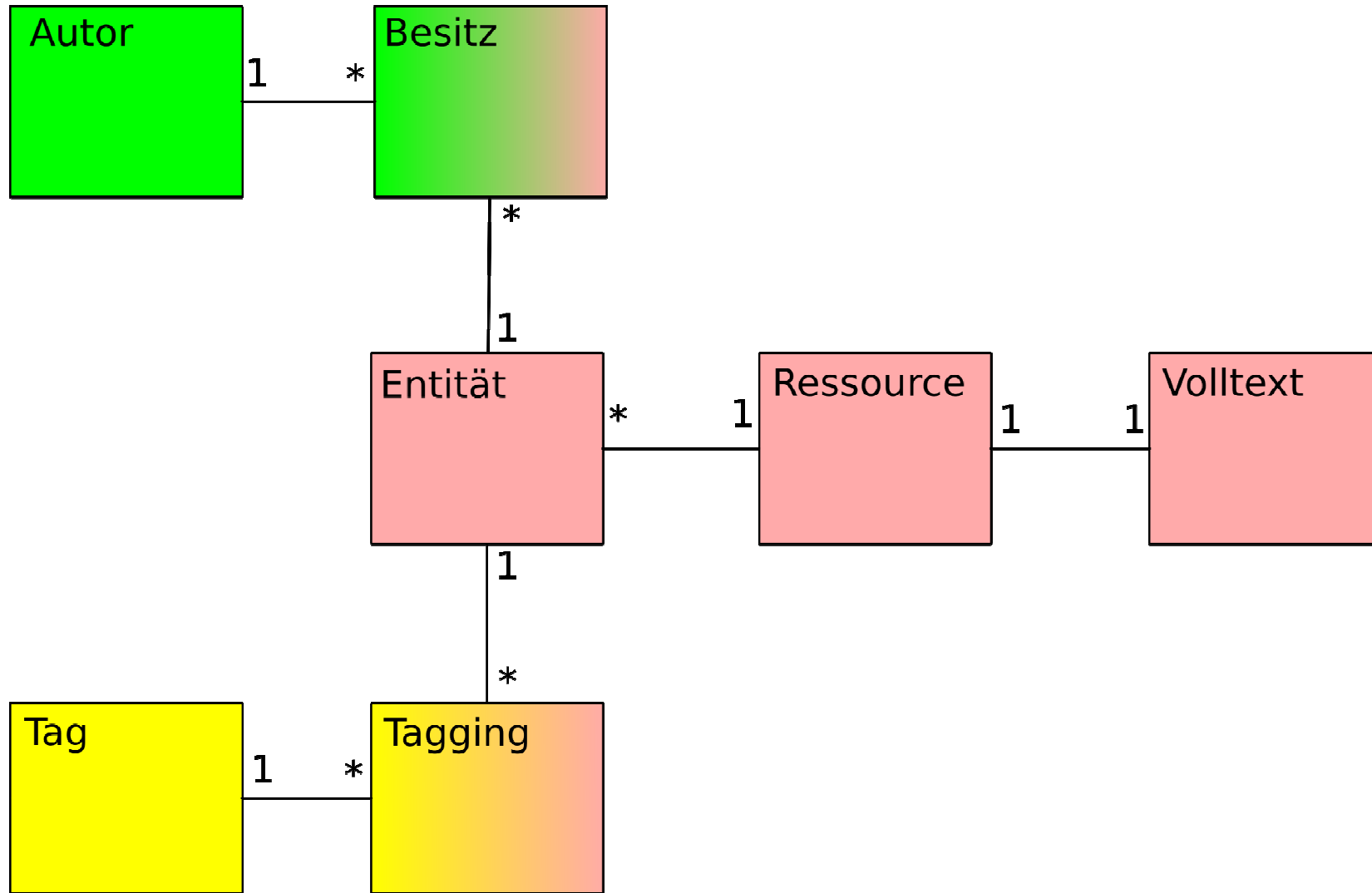
Frontend:

Input/Output: Suchanfrage/entsprechende Ergebnisliste

Das Netz



Datenbank



Crawler

- fixe Struktur mit modularer Schnittstelle
- Ziel: Untersuchung des gesamten Netzes
 - periodisch
 - minimale Zeit und Auslastung, keine Redundanz

1. Arbiter

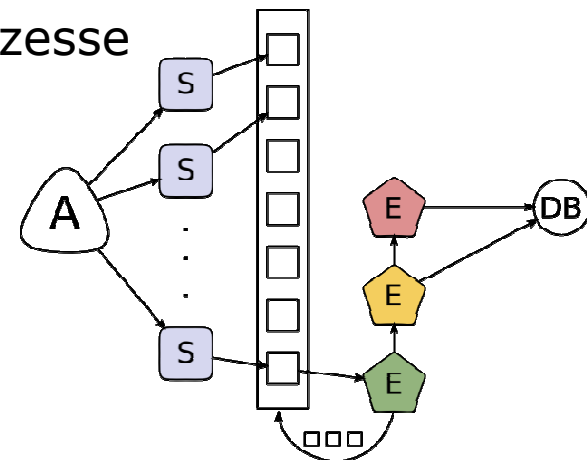
Verwaltung und Verteilung der Scanner-Prozesse
(1 zentraler Prozess)

2. Scanner

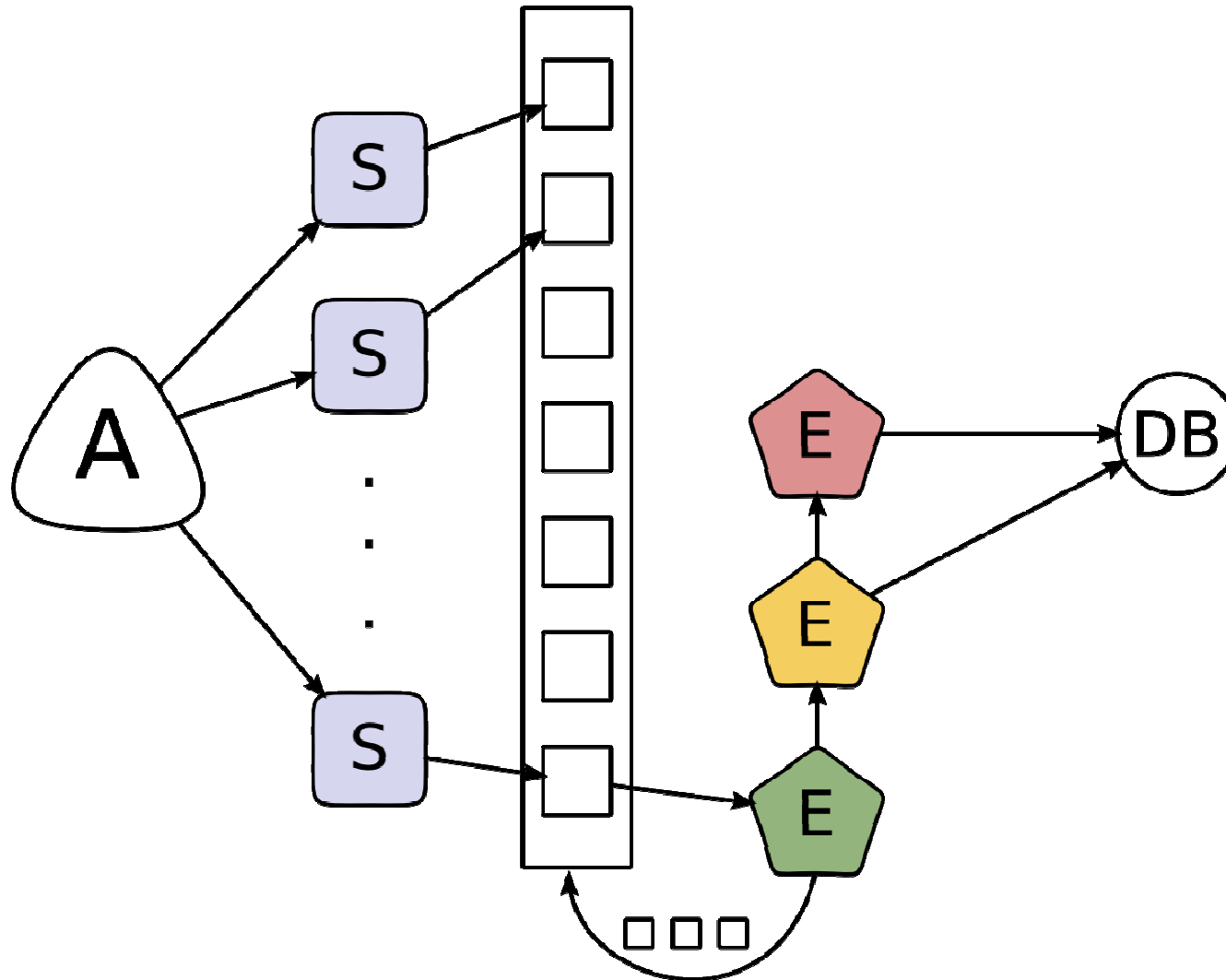
Transfer und grobe Quellenauswertung
(mehrfach geklonter Prozess)

3. Extractor

zielorientierte Datenanalyse
(mehrfache unterschiedliche Prozesse)



Crawler: Schema



Beispiel: Metadaten-Extractor



```
185
186 <!--
187 infonet
188 author:Ondrej Havel
189 name:Klapache
190 info:Tiny webserver written in C++
191 tags:software,maintained,private
192 -->
193 <h2>Klapache 1.2 features</h2>
194 <ul>
195 <li>partial HTTP/1.1 support, partial file download support
```

- eingebettete Beschreibungen
- imperative Eingabe von Seitenbesitzern

Frontend

- Suche nach Namen und Beschreibungen
- dynamische Suche unter Tags, Autoren und im Volltext



Suchel

[Tags ausblenden](#)

[crawler](#) [wissenstransfer](#) [vhb](#) [projekt](#) [video](#) [interactive](#) [esf](#) [kmu](#) [elearning](#) [private](#) [suchmaschine](#)
[maintained](#) [sprachen](#) [spanisch](#) [software](#) [finished](#) [russisch](#) [mature](#)

[Autoren ausblenden](#)

[Franziska Zacher](#) [Britta Meixner](#) [Prof. Lehner](#) [Harald Kosch](#) [Verena Dorner](#) [Klaus Kandlbinder](#)
[Ondrej Havel](#) [Prof. Hohlfeld](#) [Peter Schultes](#) [Franz Lehner](#) [Beate Siegel](#)

Keine Einträge gefunden.