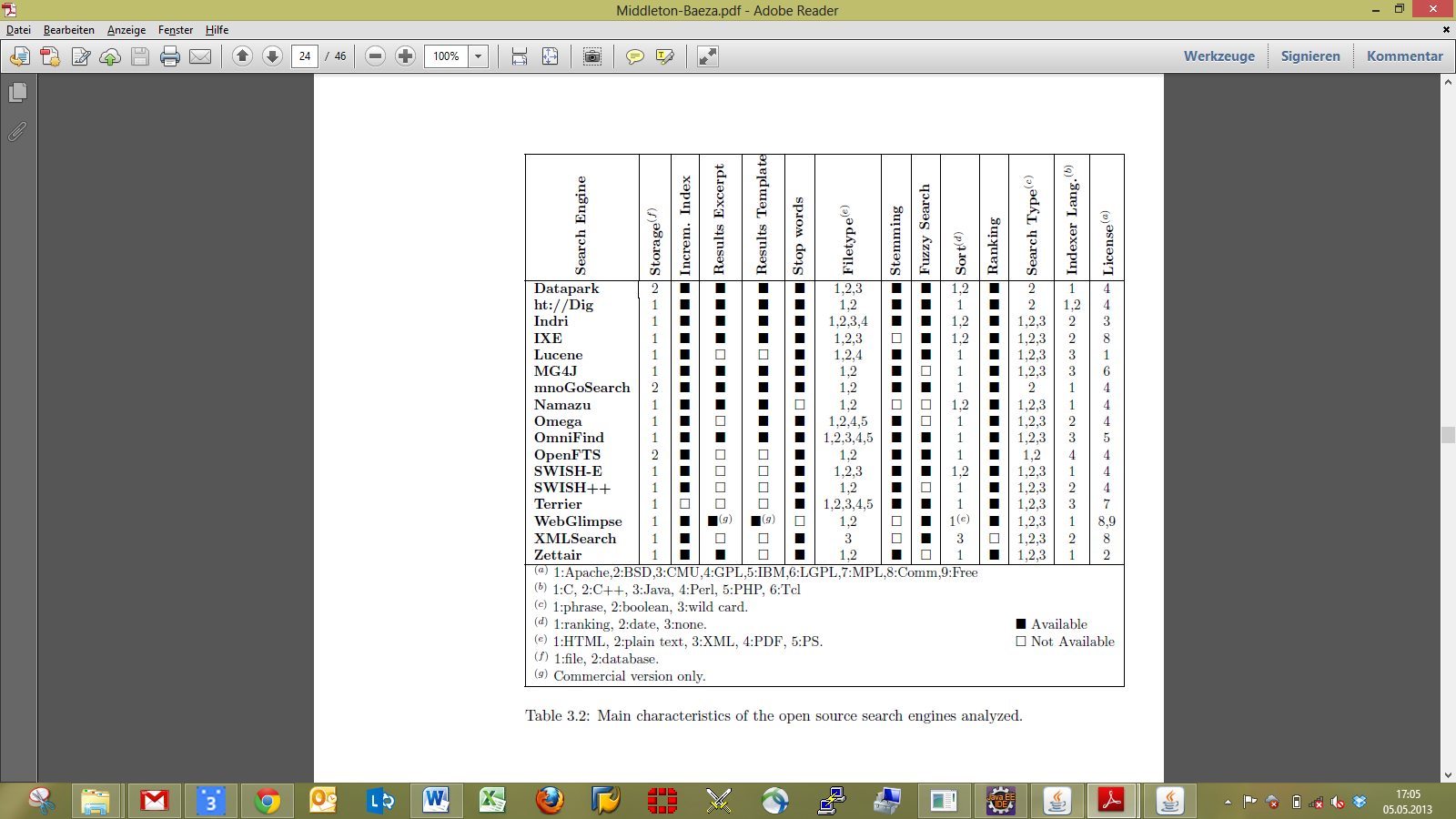
IRG SKB

Autoren: Sandro Scalco & Nadri Mamuti

Eingesetztes System:

Wir haben uns für die Opensource Variante Lucene der Apache foundation entschieden. Das dieses Framework sehr gut Dokumentiert ist.



Motivation:  
Untersuchung der Sprachendiversität. Bestes Resultat ist Durchschnitt von allen Sprachen?

Thanks to this international effort, many quality tools have now beenmade available to access information written in several language groups, includingthe Germanic (German, English, Dutch, and Swedish), Romance (Spanish, Portuguese, French, and Italian), Slavic (Russian, Bulgarian, and Czech) family groups, as well as members of the Uralic family (Finnish, Hungarian).

As described in this book, language diversity has generated an extensive variety of challenges to be solved by computer scientists. In German, for example, compound constructions are very frequent but the real concern is that the same concept might be expressed by two (or more) formulations, thus rendering it more difficult to find useful matches between search keywords and target items. The numerous grammatical cases found in Finnish grammar represent another example. Here the real problem involves irregularities imposed by vowel harmonies, which in turn generate complex morphological processing requirements. More research is required in order to promote better automatic processing of such linguistic constructions.

Aufbau Index:

Wie wurden die Anfragen an die Systeme gestellt (Manuell, automatisch, iterativ etc.)

Interpretation der Resultate