



# Search-Driven Applications

Florian Hopf, freiberuflicher Softwareentwickler  
Tobias Kraft, exensio GmbH



# Agenda

Motivation

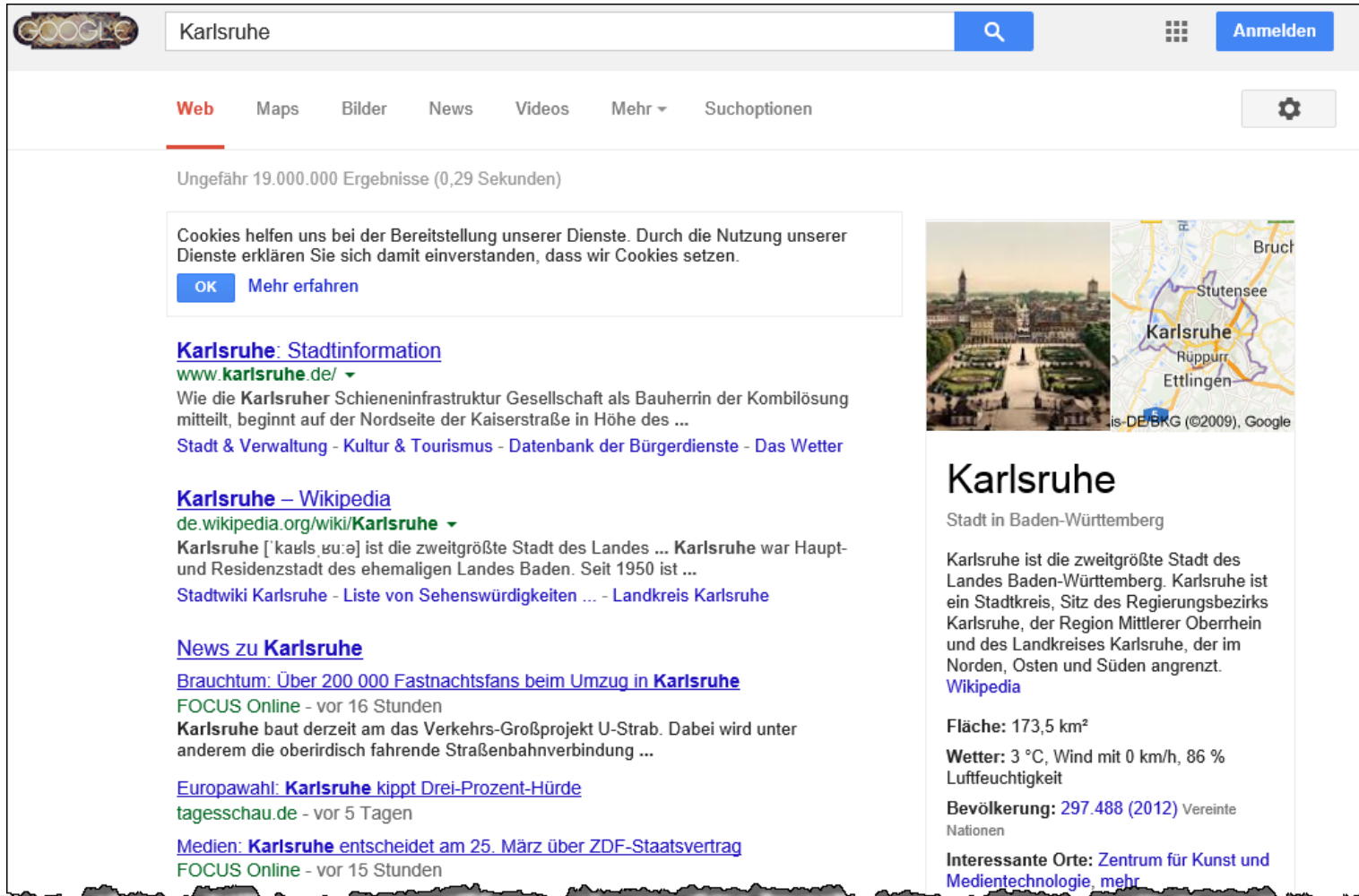
Aufbau der Such-Datenstruktur

Anwendungsfälle

Fallstricke



# Was ist Suche?



The screenshot shows a Google search interface with the search term 'Karlsruhe'. The search results are displayed in a list format. The first result is from 'www.karlsruhe.de', followed by a Wikipedia entry. Below these are news snippets from 'FOCUS Online' and 'tagesschau.de'. On the right side, there is a knowledge panel for 'Karlsruhe' with a map, a description, and key statistics.

**GOOGLE** Karlsruhe

**Web** Maps Bilder News Videos Mehr ▾ Suchoptionen

Ungefähr 19.000.000 Ergebnisse (0,29 Sekunden)

Cookies helfen uns bei der Bereitstellung unserer Dienste. Durch die Nutzung unserer Dienste erklären Sie sich damit einverstanden, dass wir Cookies setzen.  
 [Mehr erfahren](#)

**[Karlsruhe: Stadtinformation](#)**  
[www.karlsruhe.de/](http://www.karlsruhe.de/) ▾  
Wie die **Karlsruher** Schieneninfrastruktur Gesellschaft als Bauherrin der Kombilösung mitteilt, beginnt auf der Nordseite der Kaiserstraße in Höhe des ...  
[Stadt & Verwaltung](#) - [Kultur & Tourismus](#) - [Datenbank der Bürgerdienste](#) - [Das Wetter](#)


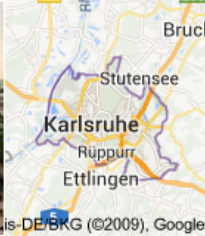
**[Karlsruhe – Wikipedia](#)**  
[de.wikipedia.org/wiki/Karlsruhe](https://de.wikipedia.org/wiki/Karlsruhe) ▾  
**Karlsruhe** [ˈkaʁls, ʁuːə] ist die zweitgrößte Stadt des Landes ... **Karlsruhe** war Haupt- und Residenzstadt des ehemaligen Landes Baden. Seit 1950 ist ...  
[Stadtwiki Karlsruhe](#) - [Liste von Sehenswürdigkeiten](#) ... - [Landkreis Karlsruhe](#)

**[News zu Karlsruhe](#)**

**[Brauchtum: Über 200 000 Fastnachtsfans beim Umzug in Karlsruhe](#)**  
[FOCUS Online](#) - vor 16 Stunden  
**Karlsruhe** baut derzeit am das Verkehrs-Großprojekt U-Strab. Dabei wird unter anderem die oberirdisch fahrende Straßenbahnverbindung ...

**[Europawahl: Karlsruhe kippt Drei-Prozent-Hürde](#)**  
[tagesschau.de](#) - vor 5 Tagen

**[Medien: Karlsruhe entscheidet am 25. März über ZDF-Staatsvertrag](#)**  
[FOCUS Online](#) - vor 15 Stunden

   
is-DE/BKG (©2009), Google

## Karlsruhe

Stadt in Baden-Württemberg

Karlsruhe ist die zweitgrößte Stadt des Landes Baden-Württemberg. Karlsruhe ist ein Stadtkreis, Sitz des Regierungsbezirks Karlsruhe, der Region Mittlerer Oberrhein und des Landkreises Karlsruhe, der im Norden, Osten und Süden angrenzt.  
[Wikipedia](#)

**Fläche:** 173,5 km<sup>2</sup>  
**Wetter:** 3 °C, Wind mit 0 km/h, 86 % Luftfeuchtigkeit  
**Bevölkerung:** 297.488 (2012) [Vereinte Nationen](#)  
**Interessante Orte:** [Zentrum für Kunst und Medientechnologie](#), [mehr](#)



# Was wollen wir in diesem Vortrag zeigen

- Suche ist mehr als Volltextsuche
- Suche funktioniert im Kleinen wie im Großen
- Suche kann auch bei klassischen Anwendungsfällen unterstützen



<http://www.morguefile.com/archive/display/861760>



# Strukturierte vs. unstrukturierte Daten

## Strukturiert

- Datenbanken
- LDAP
- Spreadsheets
- (teilweise)  
Applikations- /  
API-Daten

## Unstrukturiert

- Web Seiten
- Email
- Text- und Office-  
Dokumente
- PDF-Dokumente
- Log-Files



# Probleme klassischer Applikationen

- Daten-Silos  
→ Konsolidierung Daten
- Daten-Menge steigt immer weiter  
→ Skalierbarkeit
- Häufige Änderungen  
→ Flexibles Schema und Abfragen



# Suchserver vs. Datenbank

Eigenschaft	Suchserver	Datenbank
Semantisches Modell	Dokumenten-Modell	Relationales Modell
Speicherstruktur	Index	Tabelle
Ablage der Daten	Denormalisiert	Normalisiert



# Search-Driven Applications

Suche ist Kernbestandteil

Suchserver als zentraler Lesespeicher





# Agenda

Motivation

Aufbau der Such-Datenstruktur

Anwendungsfälle

Fallstricke



# Vom Inhalt zum Index

## Document

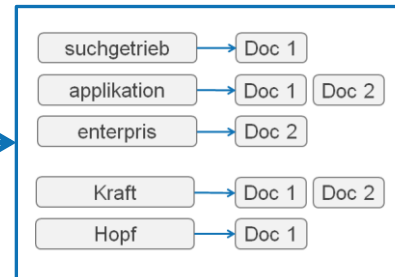
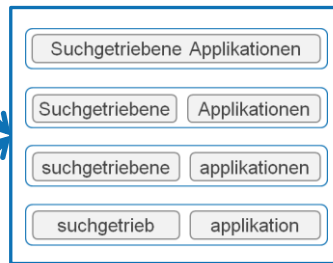
## Analyzing

## Struktur

## Aufbau

```
{
  "conference": "Berlin Expert Days",
  "title": "Suchgetriebene Applikationen",
  "speakers": [
    "Hopf, Florian",
    "Kraft, Tobias"
  ],
  "date": "04-03-2014",
  "location": "Berlin"
}
```

```
{
  "conference": "JUG SAXONY DAY",
  "title": "Grails für Enterprise-Applikationen?",
  "speakers": [
    "Kraft, Tobias"
  ],
  "date": "04-04-2014",
  "location": "Dresden"
}
```

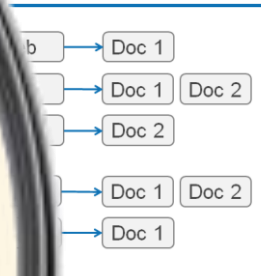


```
1 PUT /bedcon/talk/_mapping
2 {
3   "properties": {
4     "conference": {
5       "date": {
6         "type": "date",
7         "format": "dateOptionalTime"
8       },
9     },
10  },
11  "location": {
12    "type": "string",
13    "analyzer": "single_char_analyzer"
14  },
15  "speakers": {
16    "type": "multi_field",
17    "fields": {
18      "speakers": {
19        "type": "string"
20      },
21      "letter": {
22        "type": "string",
23        "analyzer": "single_char_analyzer"
24      }
25    }
26  },
27  "title": {
28    "type": "string",
29    "analyzer": "german"
30  }
31 }
```



# Input-Dokumente

```
{  
  "conference": "Berlin Expert Days",  
  "title": "Suchgetriebene Applikationen",  
  "speakers": [  
    "Hopf, Florian",  
    "Kraft, Tobias"  
  ],  
  "date": "04-03-2014",  
  "location": "Berlin"  
}  
  
{  
  "conference": "JUG SAXONY DAY",  
  "title": "Grails für Enterprise-Applikationen?",  
  "speakers": [  
    "Kraft, Tobias"  
  ],  
  "date": "04-04-2014",  
  "location": "Dresden"  
}
```



```
1 PUT /bedcon/talk/_mapping  
2 {  
3   "properties": {  
4     "conference": {  
5       "type": "string",  
6       "format": "dateOptionalTime"  
7     },  
8     "date": {  
9       "type": "date",  
10      "format": "dateOptionalTime"  
11    },  
12    "location": {  
13      "type": "string"  
14    },  
15    "speakers": {  
16      "type": "multi_field",  
17      "fields": {  
18        "speakers": {  
19          "type": "string"  
20        },  
21        "letter": {  
22          "type": "string",  
23          "analyzer": "single_char_analyzer"  
24        }  
25      }  
26    },  
27    "title": {  
28      "type": "string",  
29      "analyzer": "german"  
30    }  
31  }  
32 }
```

# Text-Analyzing

- Verarbeitung der Inputdokumente
- Preprocessing vor der Ablage
- Entfernen, modifizieren, hinzufügen von Termen

```
{
  "conference": "B...",
  "title": "Suchg...",
  "speakers": [
    "Hopf, F...",
    "Knaft, ..."
  ],
  "date": "0...",
  "location": "..."
}
```

```
{
  "co...",
  "tl...",
  "sp...",
  ],
  "da...",
  "lo..."
}
```

```
1 PUT /bedcon/talk/_mapping
2 {
3   "properties": {
4     "conference": ( ),
7     "date": {
8       "type": "date",
9       "format": "dateOptionalTime"
10    },
11    "location": ( ),
14    "speakers": {
15      "type": "multi_field",
16      "fields": {
17        "speakers": {
18          "type": "string"
19        },
20        "letter": {
21          "type": "string",
22          "analyzer": "single_char_analyzer"
23        }
24      }
25    },
26    "title": {
27      "type": "string",
28      "analyzer": "german"
29    }
30  }
}
```



# Beispiel Text-Analyzing

```
{  
  "conference": "Berlin Expert",  
  "title": "Suchgetriebene Applikationen",  
  "speakers": [  
    "Hopf, Florian",  
    "Kraft, Tobias"  
  ],  
  "date": "04-03-2014",  
  "location": "Berlin"  
}
```

```
{  
  "conference": "Berlin Expert",  
  "title": "Grail",  
  "speakers": [  
    "Kraft, Tobias"  
  ],  
  "date": "04-04-2014",  
  "location": "Dresden"  
}
```

Suchgetriebene Applikationen

Suchgetriebene Applikationen

suchgetriebene applikationen

suchgetrieb applikation

Source

Standard  
Tokenizer

Lowercase  
Filter

German-Light-  
Stem-Filter



# Der Indexaufbau

```
{
  "conference": "Berlin Expert Days",
  "title": "Suchgetriebene Applikationen",
  "speakers": [
    "Hopf, Florian",
    "Kraft, Tobias"
  ],
  "date": "04-03-2014",
  "location": "Berlin"
}
```

```
{
  "conference": "JUG SAXONY DAY",
  "title": "Grails für Enterprise-Applikationen",
  "speakers": [
    "Kraft, Tobias"
  ],
  "date": "04-04-2014",
  "location": "Dresden"
}
```

- Invertierter Index
- Inputfeld kann auf unterschiedliche Weise abgelegt werden
- Suchterm zu Dokument zuordnen

Doc 2

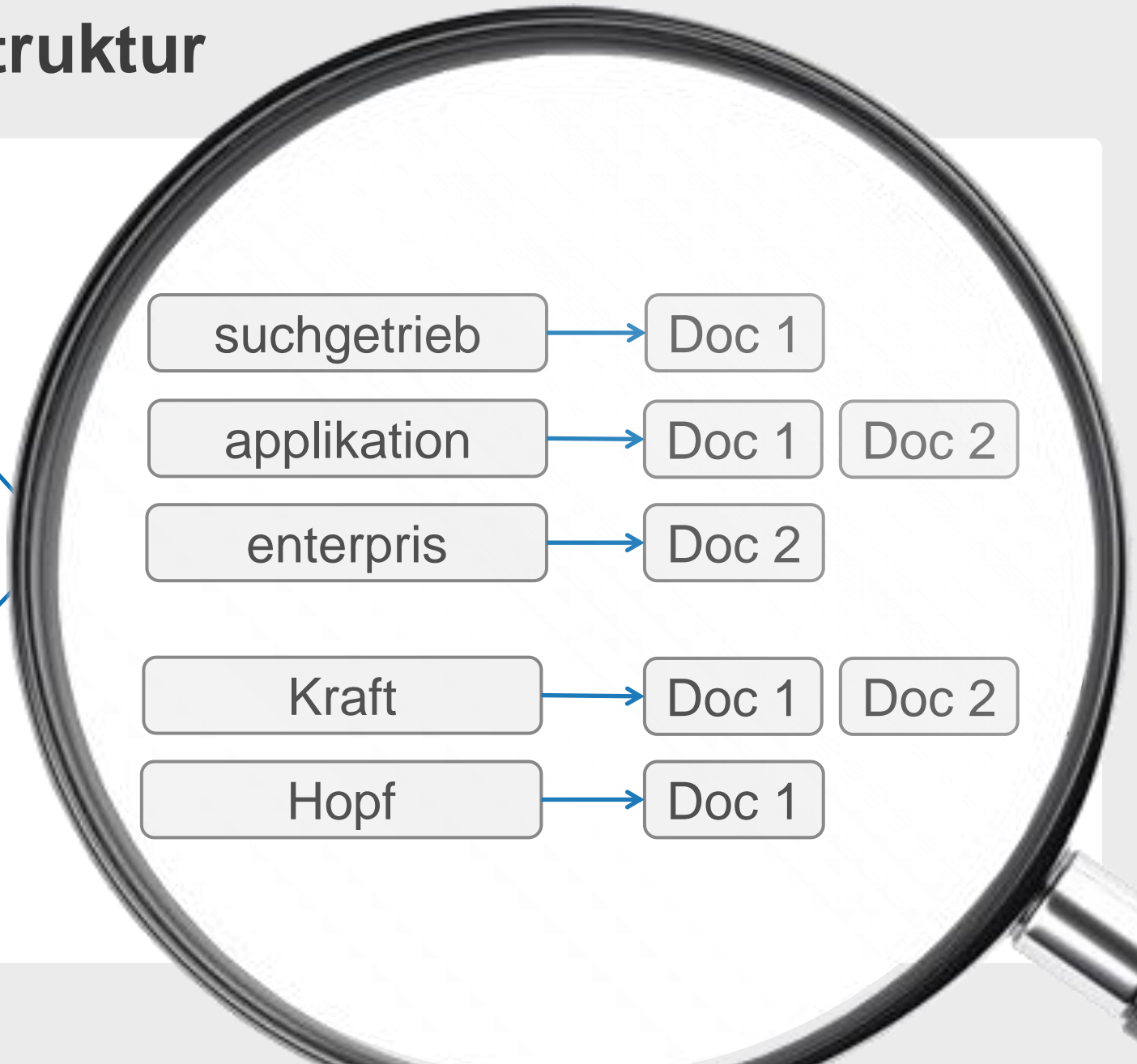
Doc 2



# Die Indexstruktur

```
{  
  "conference": "Berlin Expert Days",  
  "title": "Suchgetriebene Applikationen",  
  "speakers": [  
    "Hopf, Florian",  
    "Kraft, Tobias"  
  ],  
  "date": "04-03-2014",  
  "location": "Berlin"  
}
```

```
{  
  "conference": "JUG SAXONY DAY",  
  "title": "Grails für Enterprise-Applikationen?",  
  "speakers": [  
    "Kraft, Tobias"  
  ],  
  "date": "04-04-2014",  
  "location": "Dresden"  
}
```



# Das Mapping

```
{
  "conference": "Berlin Expert Days",
  "title": "Suchgetriebene Applikationen",
  "speakers": [
    "Hopf, Florian",
    "Kraft, Tobias"
  ],
  "date": "04-03-2014",
  "location": "Berlin"
}
```

```
{
  "conference": "JUG SAXONY DAY",
  "title": "Grails für Enterprise-Applikationen?",
  "speakers": [
    "Kraft, Tobias"
  ],
  "date": "04-04-2014",
  "location": "Dresden"
}
```



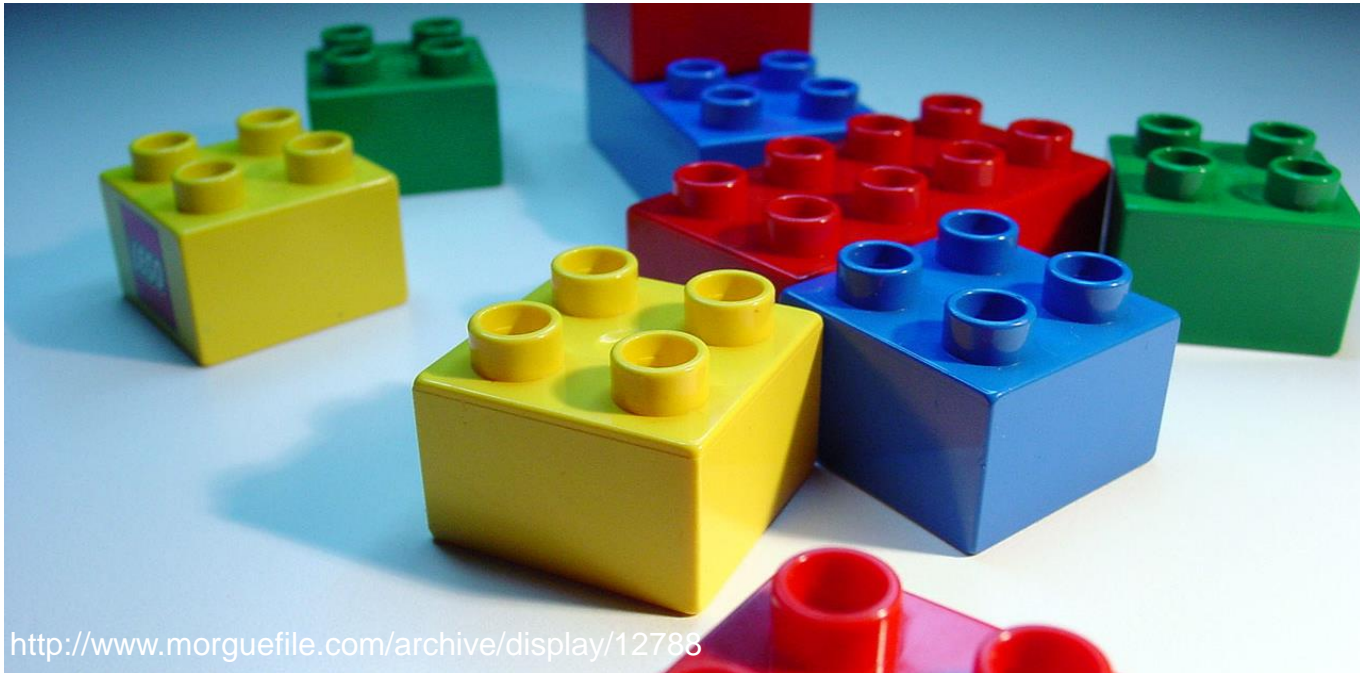
```
/bedcon/talk/_mapping {
  "properties": {
    3   "date": {
    4     "type": "date",
    5     "format": "dateOptionalTime"
    6   },
    7   "speakers": {
    8     "type": "multi_field",
    9     "fields": {
   10       "speakers" : { "type": "string" },
   11       "letter" : {
   12         "type": "string",
   13         "analyzer": "single_char_analyzer"
   14       }
   15     }
   16   },
   17   "title": {
   18     "type": "string",
   19     "analyzer": "german"
   }
  }
  ....
}
```





# Modellierung des Index

- Dokumentenmodell
- Datenablage anhand geplanter Abfragen
- Duplizierte Datenablage



# Near Realtime

- Inkrementell Indexieren
- Änderungen auch im Cluster schnell verfügbar
- Onlineabfragen für Facetting / Aggregationen



# Agenda

Motivation

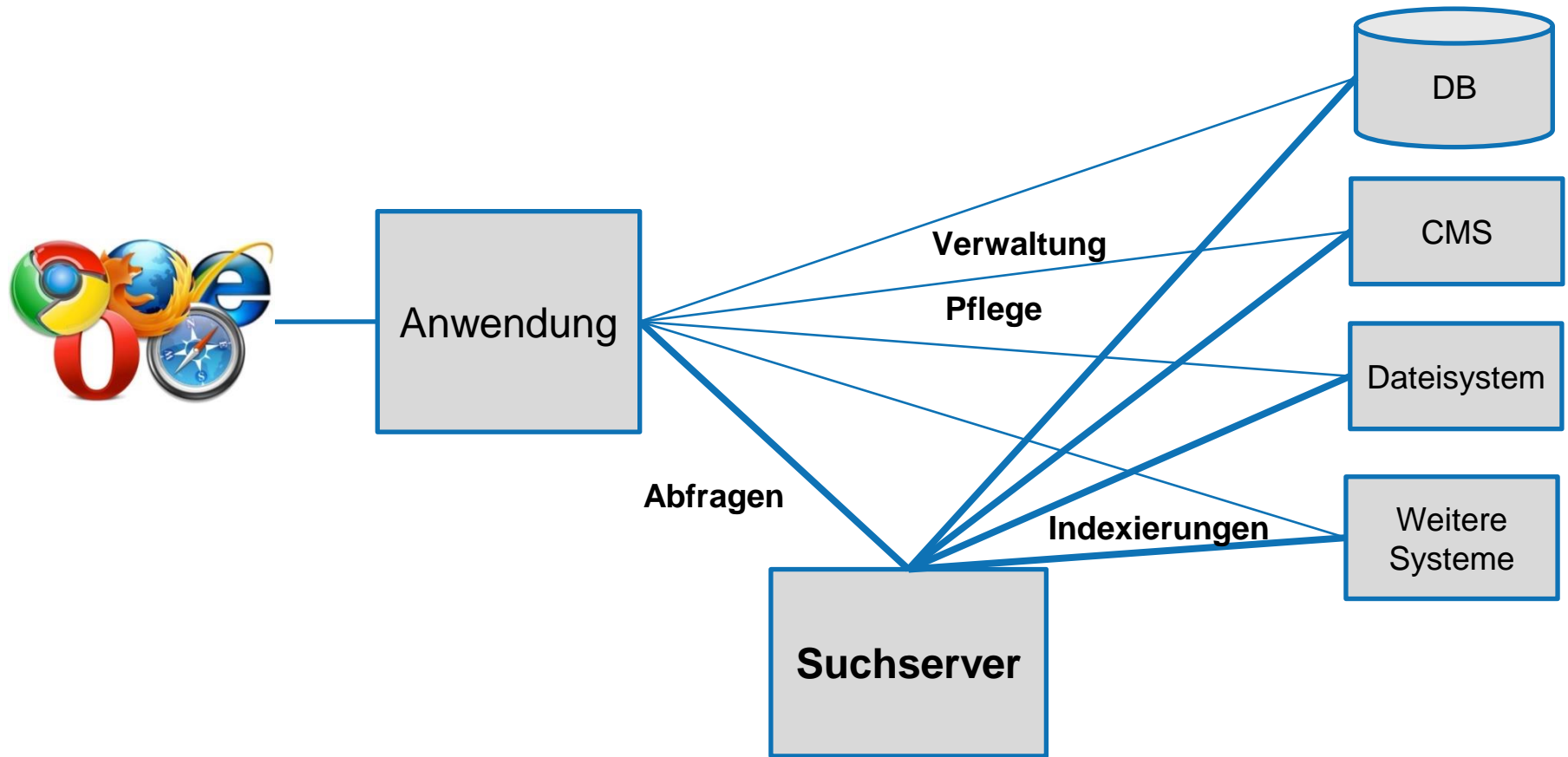
Aufbau der Such-Datenstruktur

Anwendungsfälle

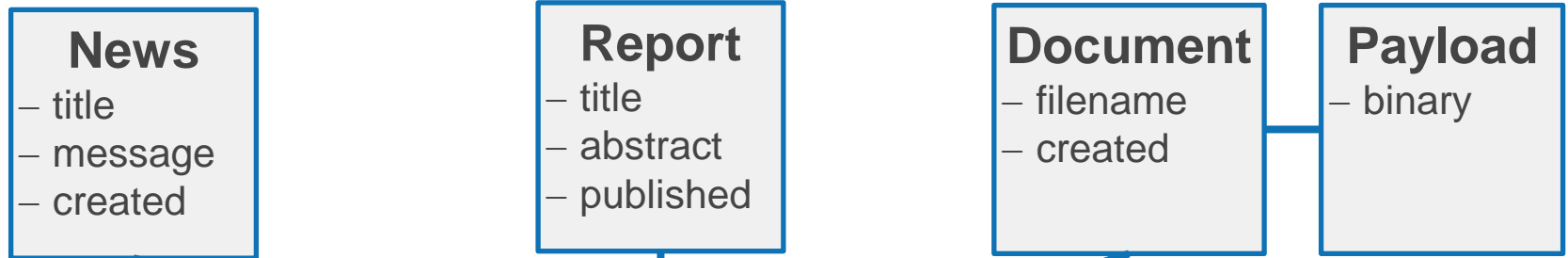
Fallstricke



# Suche ist zentraler Bestandteil der Applikation



# Vereinheitlichung von verschiedenen Quellen



## Relationale DB

- Vereinheitlichung in einer Ansicht/Liste ist schwierig

## Suchengine

- Vereinheitlichung über Index
- Entsprechendes Mapping Index-Prozess



# Listen für CRUD-Dialoge

Mitteilungen Übersicht

+ Mitteilung anlegen

Geschäft Vertrieb

Suchen

Mitteilung	Erstellt	Text	Aktionen
Neues Vertriebskonzept für unsere Kunden in BW	03.02.14 / 15:32	Mit dem Vertrieb steht und fällt das Geschäftsmodell eines	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Jahresbericht	31.01.14 / 16:43		

```
1 select * from wdb_news n
2 where
3 (n.title like '%Geschäft%'
4  or n.text like '%Geschäft%')
5 and
6 (n.title like '%Vertrieb%'
7  or n.text like '%Vertrieb%')
```

```
1  POST /_all/_search?pretty
2  {
3    "query": {
4      "match": {
5        "_all": {
6          "query": "Vertrieb Geschäft",
7          "operator": "and"
8        }
9      }
10   }
11 }
```



# Listen für CRUD-Dialoge

Mitteilungen Übersicht + Mitteilung anlegen

Geschäft Vertrieb Aussendienst Q Suchen

Mitteilung	Erstellt	Text	Aktionen
------------	----------	------	----------

```
1 select * from wdb_news n
2 inner join ex_taglink tl
3   on n.id = tl.object_id
4 inner join ex_tag t
5   on tl.tag_id = t.id
6 where
7 (n.title like '%Geschäft%'
8  or n.text like '%Geschäft%'
9  or t.name like '%Geschäft%')
10 and
11 (n.title like '%Vertrieb%'
12  or n.text like '%Vertrieb%'
13  or t.name like '%Vertrieb%')
14 and
15 (n.title like '%Aussendienst%'
16  or n.text like '%Aussendienst%'
17  or t.name like '%Aussendienst%')
```

```
1 POST /_all/_search?pretty
2 {
3   "query": {
4     "match": {
5       "_all": {
6         "query": "Vertrieb Geschäft Aussendienst",
7         "operator": "and"
8       }
9     }
10  }
11 }
```

# Facetten für den Drilldown von Events

► Customize your menu

You are here: [Home](#) > [Trainings](#) > CONGRESSES & MEETINGS

NEWS & PRESS   CONGRESSES & MEETINGS   LIBRARY   CONTACTS

## Congresses & Meetings

**Filter**

[Reset](#) [Apply](#)

**By category** ▼



**By indication** ▼





**By location** ▲

- Amsterdam (2)
- Amsterdam, Netherlands (1)
- Atlanta, Georgia (1)
- Barcelona (3)
- Beirut, Lebanon (1)
- Beijing, China (1)
- Berlin (1)

**By date** ▼

**By country** ▲

-  Corporate (34)
-  Netherlands (1)

Date	Title	Location	Owner	Country
07-Sep-2013 – 20-Nov-2013	<b>SAAS (Software as a Service): Train the Trainer</b>	<i>Bogota, Mainz, Vienna, Singapore</i>	<a href="#">Peter Soth</a>	
31-Aug-2013	<b>HAM Amsterdam August 2013: Introduction of Version 1.0</b>	<i>Amsterdam</i>	<a href="#">Irving Tschepke</a>	
31-Aug-2013 – 04-Sep-2013	<b>Best Practices for Functional Testing: Geb rules the world</b>	<i>Amsterdam</i>	<a href="#">Roland Rickborn</a>	
28-Feb-2013 – 01-Mar-2013	<b>MAM Prague February/March 2013</b>	<i>Hilton Prague</i>	<a href="#">Florian Mutter</a>	



# 7 SQL-Abfragen für den Aufbau der Facetten ...

```
1 SELECT location,COUNT(*) FROM (SELECT e.location,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((e.location IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY e.location,e.id) GROUP BY location ORDER BY location ASC
2
3 SELECT path,COUNT(*) FROM (SELECT category.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((category.path IS NOT NULL) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY category.path,e.id) GROUP BY path ORDER BY path ASC
4
5 SELECT path,COUNT(*) FROM (SELECT country.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) LEFT OUTER JOIN BR_EVENT_TAG eventCOUNTRY ON (e.id = eventCOUNTRY.eventid) LEFT OUTER JOIN BR_TAG COUNTRY ON ((eventCOUNTRY.tagid = COUNTRY.id) AND (COUNTRY.type = 'COUNTRY')) WHERE ((country.path IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY country.path,e.id) GROUP BY path ORDER BY path ASC
6
7 SELECT path,COUNT(*) FROM (SELECT indication.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((indication.path IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY indication.path,e.id) GROUP BY path ORDER BY path ASC
```

```
SELECT country.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON
```

```
8
9 SELECT userid,COUNT(*) FROM (SELECT u.userid,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((u.userid IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY u.userid,e.id) GROUP BY userid ORDER BY userid ASC
10
11 SELECT path,COUNT(*) FROM (SELECT brand.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventBRAND ON (e.id = eventBRAND.eventid) LEFT OUTER JOIN BR_TAG BRAND ON ((eventBRAND.tagid = BRAND.id) AND (BRAND.type = 'BRAND') AND (BRAND.path IN ('BR:PM/brandsfranchises/brands/actilyse','BR:PM/brandsfranchises/brands/aggrenox','BR:PM/brandsfranchises/franchises/hivaid/aptivus','BR:PM/brandsfranchises/franchises/diabetes/bi10773',NULL,NULL,'BR:PM/brandsfranchises/brands/metalyse','BR:PM/brandsfranchises/franchises/telmisartan/micardis','BR:PM/brandsfranchises/franchises/telmisartan/micardisplus','BR:PM/brandsfranchises/brands/pr','BR:PM/brandsfranchises/brands/sifrol','BR:PM/brandsfranchises/brands/spiriva','BR:PM/brandsfranchises/franchises/oncology/tomtovok','BR:PM/brandsfranchises/franchises/diabetes/trajenta','BR:PM/brandsfranchises/franchises/telmisartan/twynsta','BR:PM/brandsfranchises/franchise/s/oncology/vargatef','BR:PM/brandsfranchises/franchises/hivaidsviramune','BR:PM/brandsfranchises/franchises/oncology/volasertib') ) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((brand.path IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY brand.path,e.id) GROUP BY path ORDER BY path ASC
12
13 SELECT path,COUNT(*) FROM (SELECT franchise.path,e.id,COUNT(*) FROM BR_EVENT e INNER JOIN BR_USER u ON (e.initiator = u.id) LEFT OUTER JOIN BR_EVENT_TAG eventCATEGORY ON (e.id = eventCATEGORY.eventid) LEFT OUTER JOIN BR_TAG CATEGORY ON ((eventCATEGORY.tagid = CATEGORY.id) AND (CATEGORY.type = 'CATEGORY')) LEFT OUTER JOIN BR_EVENT_TAG eventFRANCHISE ON (e.id = eventFRANCHISE.eventid) LEFT OUTER JOIN BR_TAG FRANCHISE ON ((eventFRANCHISE.tagid = FRANCHISE.id) AND (FRANCHISE.type = 'BRAND') AND (FRANCHISE.path IN ('BR:PM/brandsfranchises/franchises/diabetes','BR:PM/brandsfranchises/franchises/hepatitisvirus','BR:PM/brandsfranchises/franchises/hivaid','BR:PM/brandsfranchises/franchises/oncology','BR:PM/brandsfranchises/franchises/telmisartan') ) LEFT OUTER JOIN BR_EVENT_TAG eventBRANDSELECTED ON (e.id = eventBRANDSELECTED.eventid) LEFT OUTER JOIN BR_TAG BRANDSELECTED ON ((eventBRANDSELECTED.tagid = BRANDSELECTED.id) AND (BRANDSELECTED.type = 'BRAND')) WHERE ((franchise.path IS NOT NULL) AND (category.path IN ('BR:PM/category/events/tags/corporatebrandteam') ) AND (brandselected.path IN ('BR:PM/brandsfranchises/brands/pr') ) AND (e.deleted = 0)) GROUP BY franchise.path,e id) GROUP BY path ORDER BY path ASC
```



# Facette mit Elasticsearch

```
1 POST /events/_search
2 {
3   "aggs" : {
4     "categories" : {
5       "indications" : {
6         "locations" : { "t
7         "countries" : { "t
8       "created_range": {
9         "date_range":
10        "field": "
11        "format":
12      "ranges":
13        { "fro
14        { "fro
15        { "fro
16      ]
17    }
18  }
19 }
20 }

15 "aggregations": {
16   "locations": {
17     "buckets": [
18       {
19         "key": "Berlin",
20         "doc_count": 626
21       },
22       {
23         "key": "Basel",
24         "doc_count": 93
25       },
26       .....
27     ]
28   },
29   "countries": {
30     "buckets": [
31       {
32         "key": "Germany",
33         "doc_count": 671
34       },
35     ]
36   }
37 }
```



# Anwendungsfälle für E-Commerce

The screenshot shows the Amazon.de website interface. At the top, the Amazon logo is followed by navigation links: 'Mein Amazon', 'Angebote', 'Gutscheine', 'Verkaufen', and 'Hilfe'. Below this is a search bar with 'Praktische Informatik' entered and a 'Los' button. A secondary navigation bar lists categories like 'Bücher', 'Erweiterte Suche', 'Stöbern', 'Bestseller', 'Neuheiten', 'Hörbücher', 'Fremdsprachige Bücher', 'Taschenbücher', 'Fachbücher', 'Schulbücher', 'Angebote', and 'Trade-In'. The main content area is divided into a left sidebar and a main product list. The sidebar includes sections for 'Neuheiten' (Last 30 days: 16, Last 90 days: 59, Next: 33), 'Kategorie' (Bücher, Fachbücher, Informatik), 'Praktische Informatik' (Datenbanken: 217, Grafische Datenverarbeitung: 31, Programmiersprachen & Compiler: 4.923, Programmiersprachen & Interpreter: 266, Software-Entwicklung: 1.605), 'Format' (Gebundene Ausgabe: 1.429, Taschenbuch: 4.909, Hörbuch: 16, Kindle eBook: 586), and 'Autor' (Helmut Erenkötter: 10, Jürgen Wolf: 16, Jon Erickson: 1, Thomas Theis: 11, Tilo Linz: 4, Andreas Spillner: 3, Gernot Starke: 10, and a 'Mehr...' link).

**Buch-Tipps für Sie**  
[Bücher bei Amazon.de](#)  
Entdecken Sie unsere große Auswahl an Büchern und erhalten Sie persönliche Empfehlungen im [Buch-Shop](#).

**Bücher > Fachbücher > Informatik > Praktische Informatik**

1-16 von 6.705 Ergebnissen

**Format**

**Blick ins Buch!**

**Programmieren lernen mit Java: Aktuell zu Java 7 - Keine Vorkenntnisse erforderlich** von Hans-Peter...

**EUR 19,90** Broschiert **Bestellen Sie in den nächsten 12 Stunden, um den Artikel am Dienstag, 18. März zu erhalten.**

Andere Angebote - Broschiert  
**EUR 18,95** neu (54 Angebote)  
**EUR 15,00** gebraucht (4 Angebote)

★★★★★ (40)  
**Bestseller Nr. 1** in [Java](#)  
Kostenlose Lieferung möglich.  
Tauschen Sie diesen Artikel gegen einen Amazon.de Geschenkzettel.

**Blick ins Buch!**

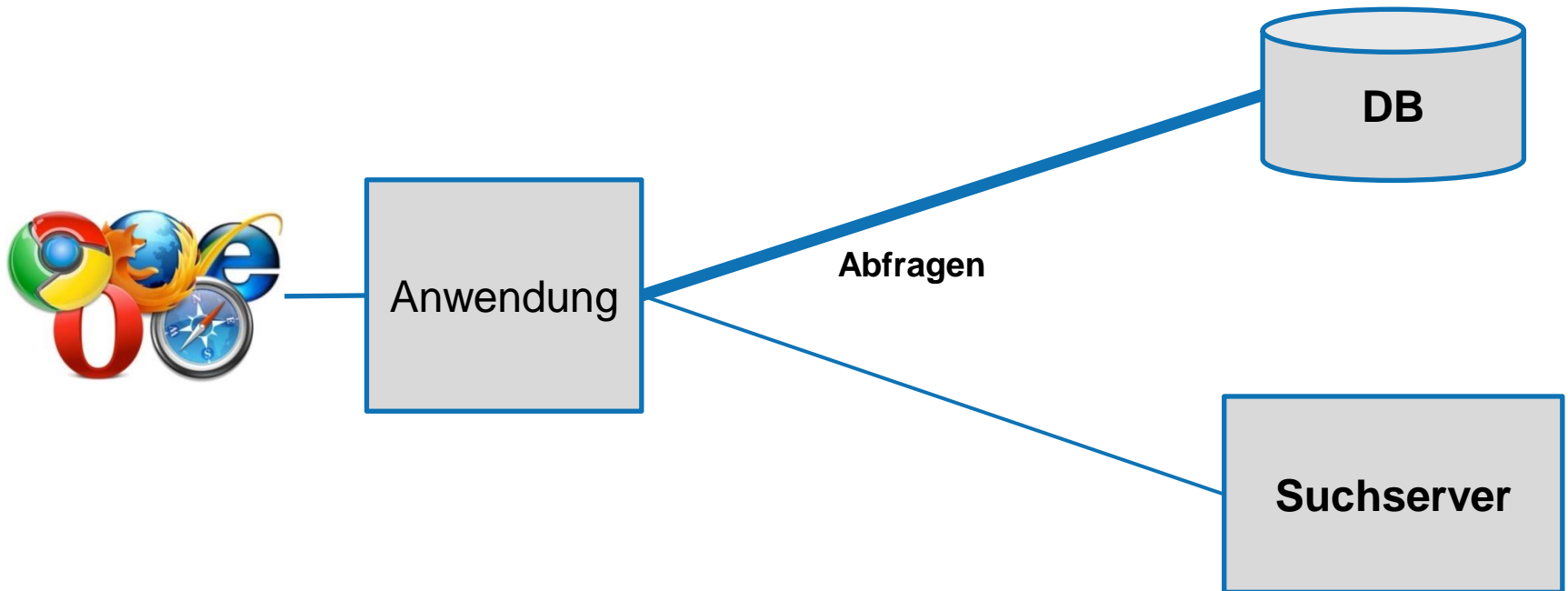
**Hacking: Die Kunst des Exploits (mit CD)** von Jon Erickson von dpunkt Verlag (1. September 2008)

**EUR 46,00** Broschiert **Bestellen Sie in den nächsten 12 Stunden, um den Artikel am Dienstag, 18. März zu erhalten.**

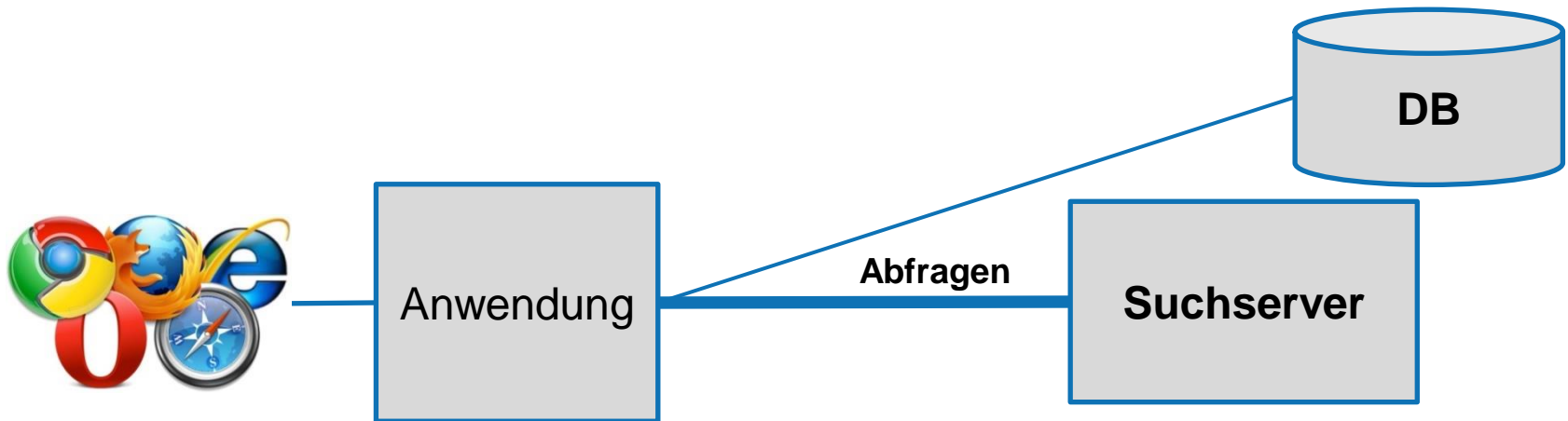
Andere Angebote - Broschiert  
**EUR 46,00** neu (77 Angebote)  
**EUR 34,60** gebraucht (7 Angebote)

★★★★★ (19)  
Kostenlose Lieferung möglich.  
Tauschen Sie diesen Artikel gegen einen Amazon.de Geschenkzettel.

# Der Suchserver rückt in den Mittelpunkt



# Der Suchserver rückt in den Mittelpunkt



# Kategorien-Navigation

## Neuheiten

Letzte 30 Tage (16)

Letzte 90 Tage (59)

Demnächst (33)

## Kategorie

◀ Bücher

◀ Fachbücher

◀ Informatik

◀ Grafische

Datenverarbeitung (31)

Programmiersprachen &  
Compiler (4.923)

Programmiersprachen &  
Interpreter (266)

Software-  
Entwicklung (1.605)

## Format

Gebundene Ausgabe (1.429)

Bücher



/Bücher

```
/select?q=*:*&facet=on&facet.field=category&rows=0
```



/Bücher



/Bücher/Fachbücher



# Produktlisting

Bücher > Fachbücher > Informatik > Praktische Informatik

1-16 von 6.705 Ergebnissen Details Bildansicht Sortieren nach Beliebtheit

**Format**

[Gebundene Ausgabe](#) [Taschenbuch](#) [Hörbuch](#) [Kindle eBook](#)

**Blick ins Buch!** **Programmieren lernen mit Java: Aktuell zu Java 7 - Keine Vorkenntnisse erforderlich** von Hans-Peter Habelitz von Galileo Computing (28. Mai 2012)  
**EUR 19,90** Broschiert ★★★★☆ (40)

**/select?q=\*&fq=category:/Bücher/Fachbücher/Informatik  
&sort=inStock desc, price asc**

**Blick ins Buch!** **Hacking**  
**EUR 46,00** Broschiert   
Bestellen Sie in den nächsten **12 Stunden**, um den Artikel am Dienstag, 18. März zu erhalten.  
Andere Angebote - Broschiert  
**EUR 46,00** neu (77 Angebote)  
**EUR 34,60** gebraucht (7 Angebote)  
★★★★☆ (19)  
Kostenlose Lieferung möglich.  
Tauschen Sie diesen Artikel gegen einen Amazon.de Geschenkgutschein ein

**Blick ins Buch!** **Raspberry Pi programmieren mit Python (mitp Professional)** von Michael Weigend von mitp (11. November 2013)



# Agenda

Motivation

Aufbau der Such-Datenstruktur

Anwendungsfälle

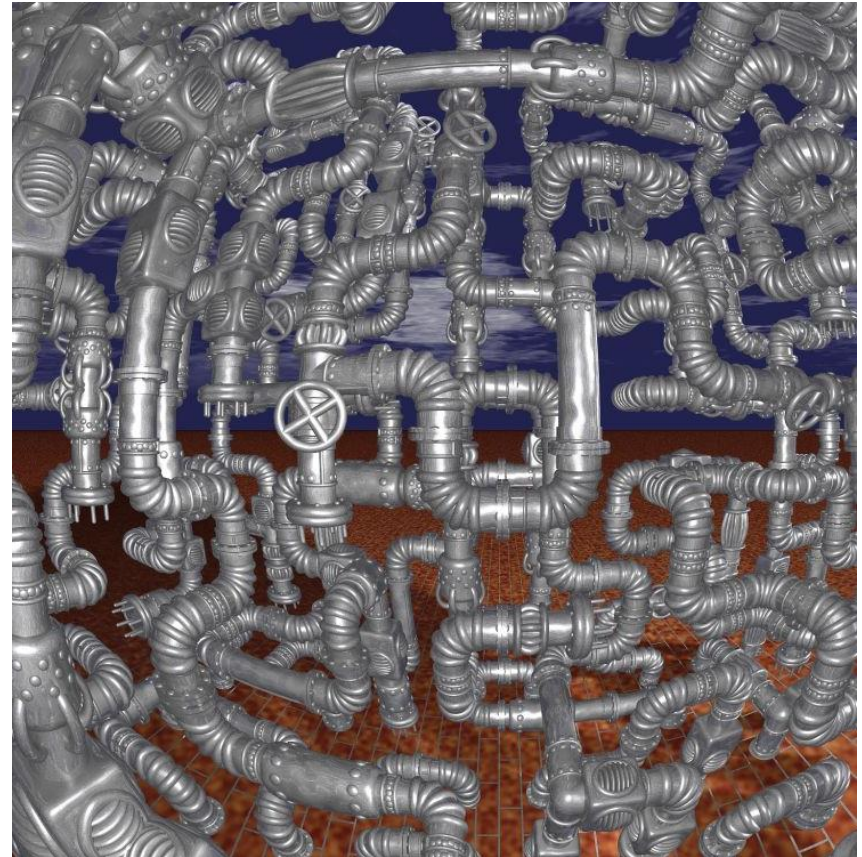
Fallstricke





# Komplexität der Architektur steigt

- Synchronisierung Systeme
- Fehlersuche
  - Mehrere `Verursacher`
  - Nachstellen Produktionsprobleme
- Entwicklungssetup




<http://www.morguefile.com/archive/display/591261>



# Transaktionen

```
if (study.validate() && study.validateTagHierarchy() && study.save(flush: true)) {  
    // index process  
    searchService.indexStudy(study)  
    result.message = tagLibService.validationTagLib.message(code: 'default.info.update')  
}  
else {  
    // reset  
    study.stepDone = stepBefore  
    study.discard() // necessary to discard changes on errors  
    return fail(code:"default.update.failure")  
}  
  
documentService.store(doc)  
  
// success  
return result
```



# Daten

- Analyzing ist verlustbehaftet
- Update einzelner Felder schwierig
- Originalinhalt für Anpassungen weiter zugreifbar



<http://www.morguefile.com/archive/display/183870>



# Testbarkeit

- Testunterstützung ist vorhanden
  - SolrtestCaseJ4
  - Elasticsearch Test Framework
- Relevanz ist subjektiv
- Analyzing-Prozess kann komplexer werden
- Rollback in Tests

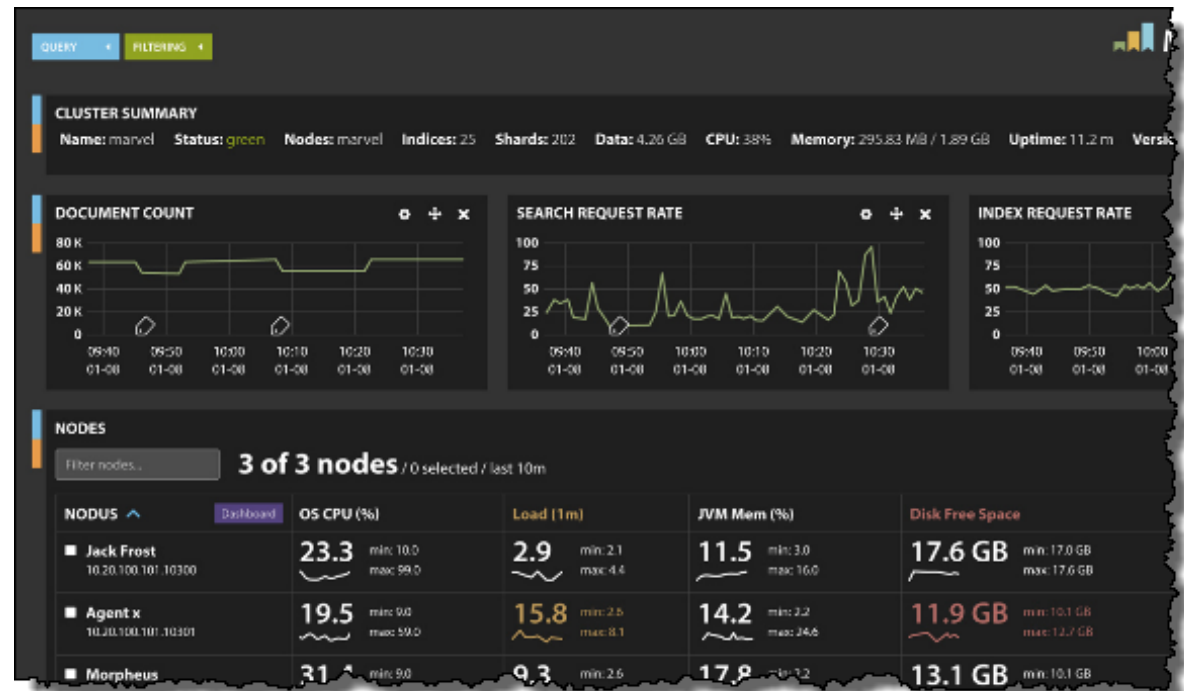


<http://www.morguefile.com/archive/display/87396>



# Betrieb

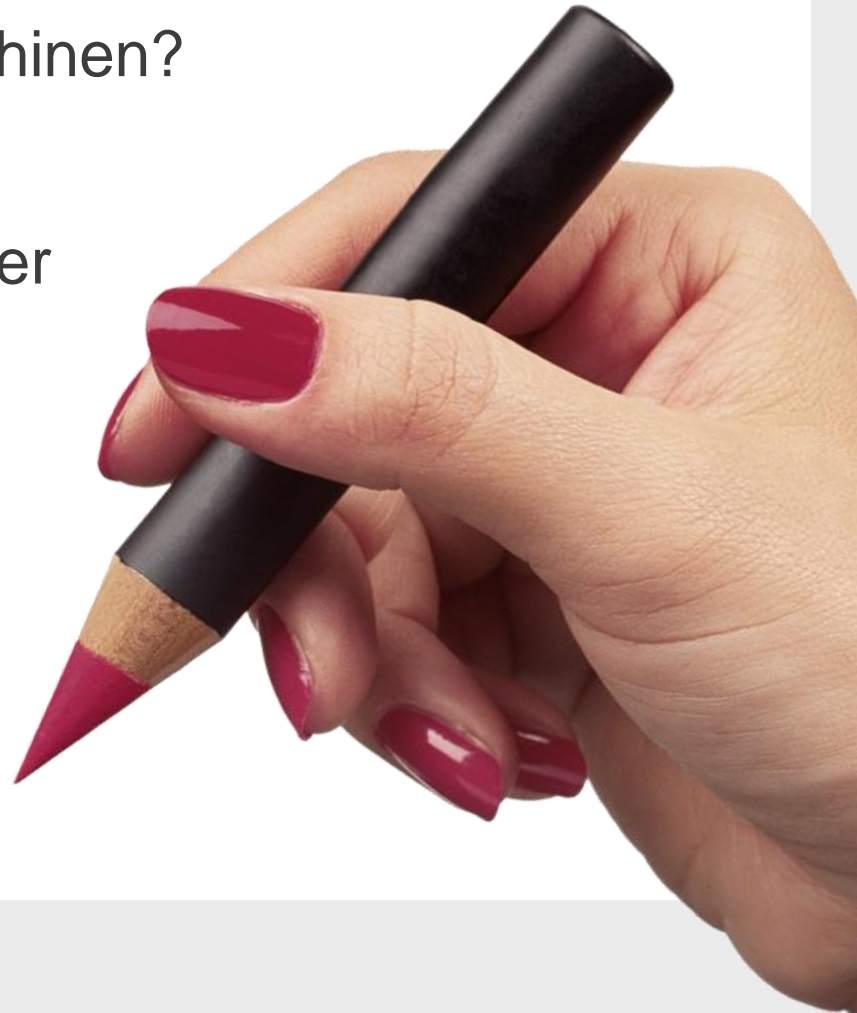
- Security
- Separate Installation
- Verfügbarkeit
- Verteilte Logfiles
- Monitoring



# Fazit

Was bringt die Umsetzung von Applikations-Logik mit Suchmaschinen?

- Abfragen sind teilweise einfacher
- Flexibler für Erweiterungen
- Performance-Vorteile
- Neue Anwendungsszenarien



# Mehr Interesse an Suche?



SEARCH  
MEETUP  
KARLSRUHE

@KASearch

**Vielen Dank!**

Florian Hopf, @fhopf

Tobias Kraft, @tokraft





ENTWICKLERTAG

meet the SPEAKER  
@speakerlounge



1. OG DIREKT ÜBER DEM EMPFANG