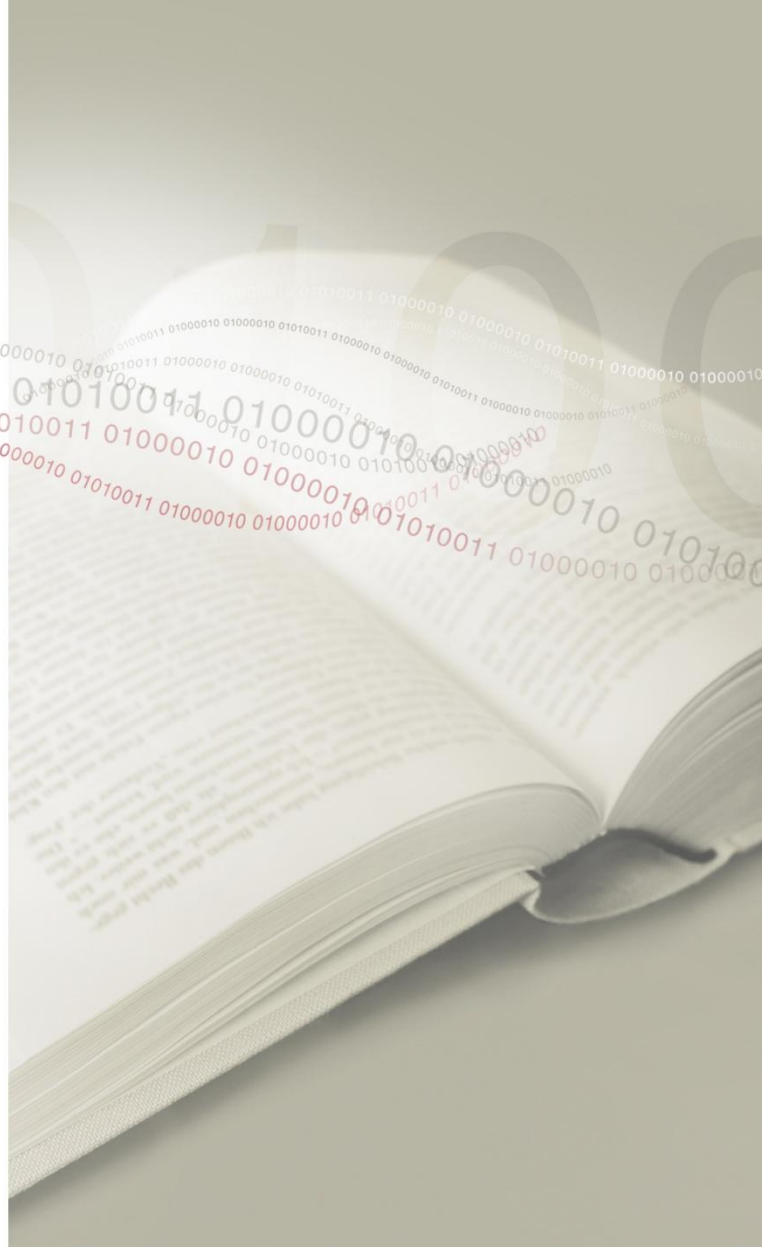


OPACplus auf neuer Plattform

TouchPoint & Solr Lucene



Motivation

- ▶ FAST ESP läuft aus
 - kein Support
 - keine weitere Entwicklung
- ▶ Neue Entwicklungen auf neuer Plattform
 - unabhängig von Bibliotheksanwendung
 - direkte Servicefunktionen versus Verfügbarkeitsprüfung
 - Einbindung Discovery Service
 - Integration der Volltextsuche
 - hierarchische Facette & Browsing–Sucheinstieg mittels Checkboxes







Neue Komponenten

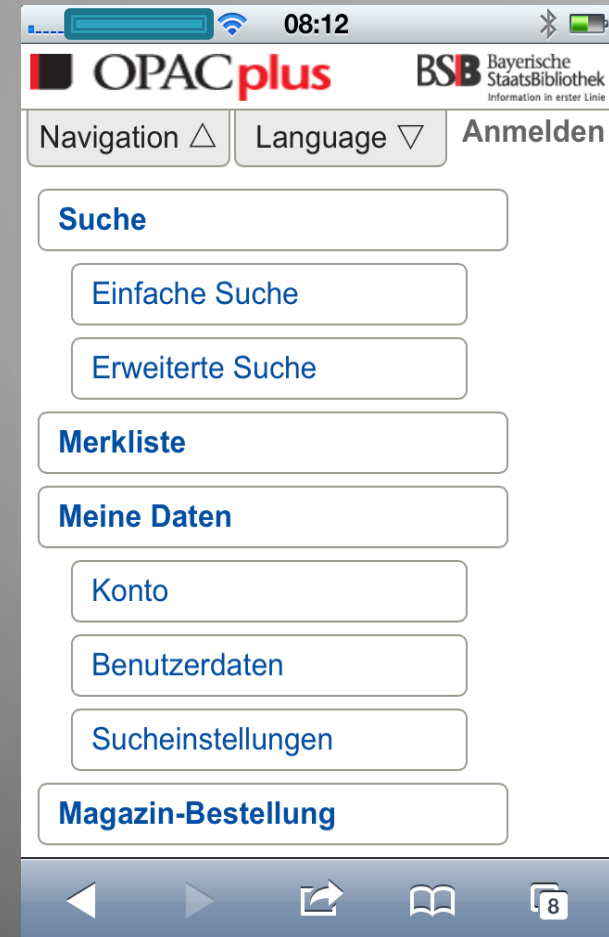
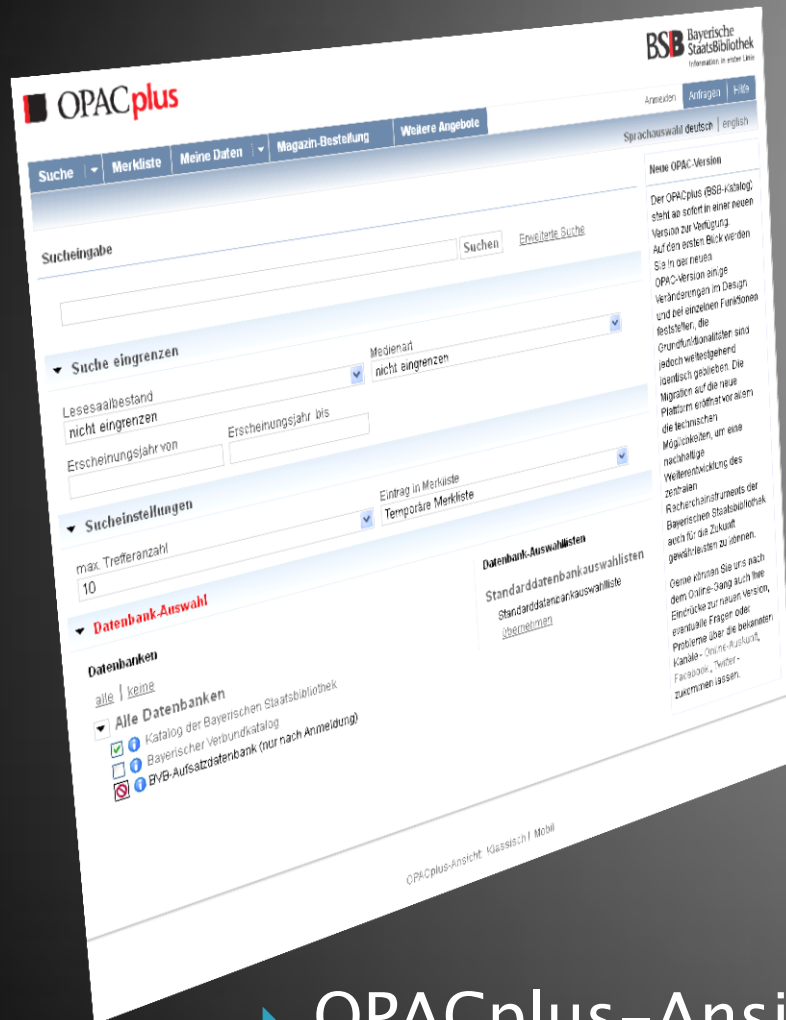
- ▶ Verarbeitungsformat MARCXML
- ▶ Anzeige erfolgt jetzt komplett aus Suchmaschinenindex
- ▶ NCIP–Schnittstelle zum Ausleihsystem
- ▶ Suchmaschine: Solr
- ▶ Suchapplikation und User–Frontend: TouchPoint
- ▶ TP–Admin Datenbank

Änderungen/Verbesserungen

- ▶ Service-Funktionen direkt in der Trefferliste
- ▶ mobil/klassisch
- ▶ Neue Icons für Medientypen
- ▶ Zusammenführung der zahlreichen Medientypen aus Navigatoren-Feld
- ▶ Navigation in und zur Trefferliste optimiert
- ▶ bessere Sichtbarkeit und direkterer Zugriff auf Online-Ressourcen

► Service direkt

- ☐ 1.  **Solr 1.4 Enterprise Search Server : enhance your search with faceted navigation, result highlighting, fuzzy queries, ranked scoring, and more** (Birmingham [u.a.], 1. publ. 2009) VI, 317 S. : Ill.
Smiley, David
Buch [Vormerken \(€\)](#)
- ☐ 2.  **Jahresbericht / Universitätsbibliothek Erlangen-Nürnberg** (Erlangen) Online-Ressource
[e Online lesen](#) [Kopie bestellen \(€\)](#)
Zeitschrift
- ☐ 3.  **Library automation and OPAC 2.0 : information access and services in the 2.0 landscape** (Hershey, PA 2013) XVI, 228 S. : Ill, graph. Darst.
Tramullas, Jesús
Buch [Vormerken \(€\)](#)
- ☐ 4.  **En kulturhistorisk epoke i bøgernes verden : en beretning om bibliotekerne i Tune Herred 1805 - 1970** (Egå 2009) 100 S., [5] Bl. : Ill., Kt. ;
Kjølhede, Karen
Buch [Ausleihen](#)
- ☐ 5.  **Liguria sconosciuta : itinerari insoliti e curiosi** (Milano 2002) 240, XXIV S. : zahlr. Ill. ;
Ricchetti, Mauro
Buch [Ausleihen](#)
- ☐ 6.  **Mittelniederdeutsches Lesebuch** (Hamburg 1921) 148 S.
Stammler, Wolfgang
Buch [Ausleihen](#)



► OPACplus-Ansicht: Klassisch | Mobil

- ▶ Neue Icons und Zusammenführung der Medientypen




- ▶ Navigation in und zur Trefferliste optimiert



- ▶ bessere Sichtbarkeit und direkter Zugriff auf Online-Ressourcen

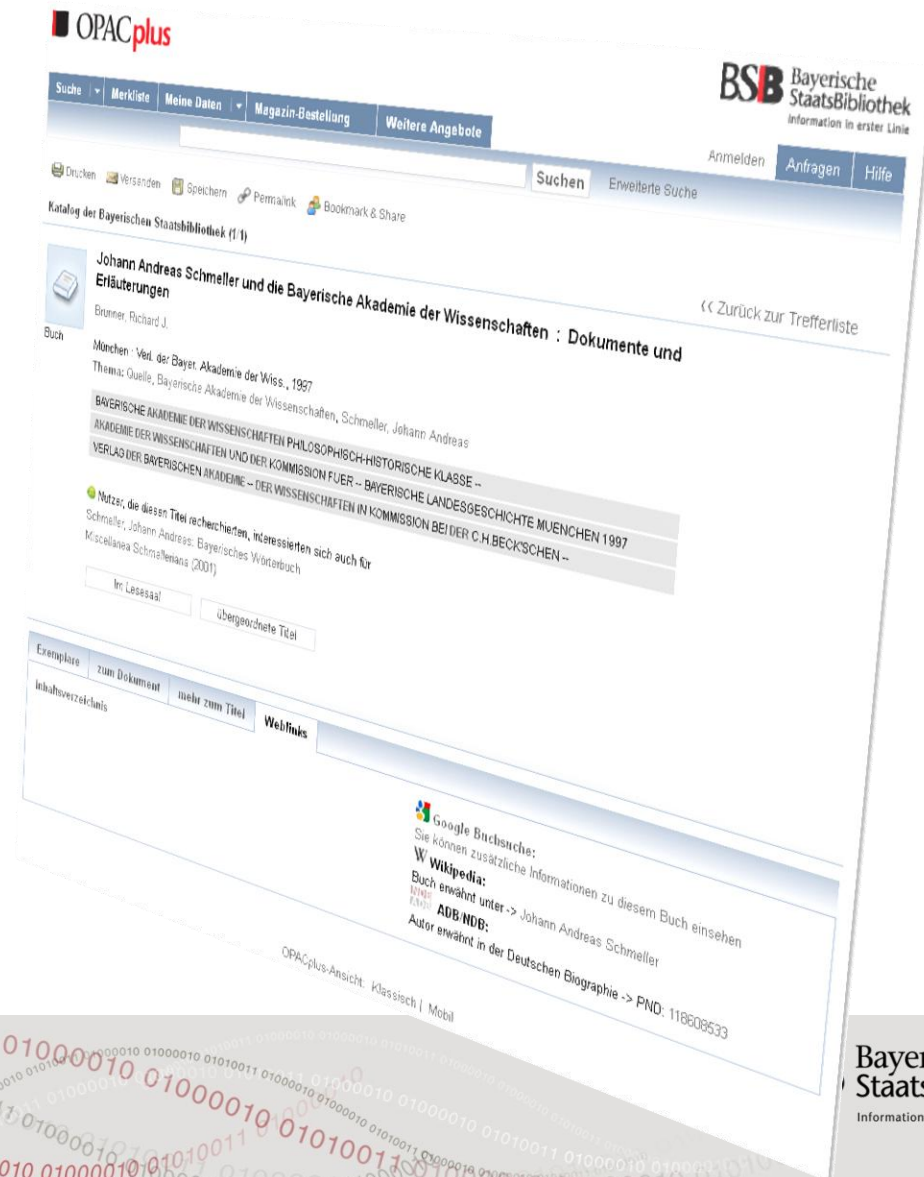
Treffer Katalog der Bayerischen Staatsbibliothek (551) |« < 1 **2** 3 > »|

Titel auswählen: alle | keine
 Speichern in: Temporäre Merkliste Speichern
 Sortieren nach: Relevanz Sortieren
 max. Trefferanzahl: 10 Anzeigen

| | | |
|------------------------------|---|--|
| <input type="checkbox"/> 11. |  Sonstiges | Multimodal approaches in human brain mapping (2012) IX, 81 Bl. : Ill., graph. Darst. Online-Ressource Kirsch, Valerie e Online lesen Ausleihen |
| <input type="checkbox"/> 12. |  Buch | Media, communication and development : three approaches (New Delhi c2012) xx, 258 p. Manyozo, Linje e Online lesen |

Bewährtes in neuem Gewand

- ▶ dyn. Textteaser
- ▶ BibTip
- ▶ Wikipedia und Google MashUp
- ▶ Service direkt
- ▶ Permalink
- ▶ Query Language
- ▶ Bookmark & Share
- ▶ Weblinks



Spezial

- ▶ Solr MARC XML Schnittstelle mit Http Content Negotiation

<https://opacplus.bsb-muenchen.de/title/BV021693819> (BVNummer / Katkey)

Suche

Merkmale

Meine Daten

Magazin-Bestellung

Weitere Angebote

Zurück zur Trefferliste

java Insel

Drucken

Versenden

Speichern

Permalink

Bookmark & Share

Katalog der Bayerischen Staatsbibliothek (1/115)

Karte

Quelle

Politische Karte der Insel Java

Ca. 1:500 000 - [Berlin] : [Selbstverl.], [1863]

Thema: Karte, Grenze, Java

| Exemplare | zum Dokument | mehr zum Titel | Weblinks |
|--|--------------|----------------|----------|
| <div> <div>Titel:</div> <div>Thema:</div> <div>Ausgabe:</div> <div>Ort:</div> <div>Verlag:</div> <div>Erscheinungsjahr:</div> <div>Sprache:</div> <div>Umfang:</div> <div>Medienart:</div> <div>Medienart:</div> <div>Identifier:</div> </div> <div> <div>Politische Karte der Insel Java</div> <div>Karte / Grenze / Java</div> <div>Ca. 1:500 000</div> <div>[Berlin]</div> <div>[Selbstverl.]</div> <div>[1863]</div> <div>ger</div> <div>1 Kt. : Lithographie ;</div> <div>Karte</div> <div>Druckschrift</div> <div>9466908</div> </div> | | | |

```
<mx:record id="9466908">  
  <mx:leader>01309nea a2200000 i4500</mx:leader>  
  <mx:controlfield tag="001">9466908</mx:controlfield>  
  <mx:controlfield tag="005">20121117</mx:controlfield>  
  <mx:controlfield tag="006">a|||||||u|||||||</mx:controlfield>  
  <mx:controlfield tag="007">a|||||||</mx:controlfield>  
  <mx:controlfield tag="008">060811s1863uuuwg d|| | ||||||||ger||</mx:controlfield>  
- <mx:datafield tag="035" ind1=" " ind2=" " >  
  <mx:subfield code="a">(CoLC)162363356</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="035" ind1=" " ind2=" " >  
  <mx:subfield code="a">(DE-604)BV021693819</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="040" ind1=" " ind2=" " >  
  <mx:subfield code="e">rakwb</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="245" ind1="0" ind2="0">  
  <mx:subfield code="a">Politische Karte der Insel Java</mx:subfield>  
  <mx:subfield code="9">  
    PolitischeKarte0der0Insel0Java0000000000000000000000000000000000000000000000000  
  </mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="255" ind1=" " ind2=" " >  
  <mx:subfield code="a">Ca. 1:500 000</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="260" ind1=" " ind2=" " >  
  <mx:subfield code="a">[Berlin]</mx:subfield>  
  <mx:subfield code="b">[Selbstver.]</mx:subfield>  
  <mx:subfield code="c">[1863]</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="300" ind1=" " ind2=" " >  
  <mx:subfield code="a">1 Kt. : </mx:subfield>  
  <mx:subfield code="b">Lithographie ; </mx:subfield>  
  <mx:subfield code="c">8 x 21 cm, gefaltet</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="650" ind1="0" ind2="7">  
  <mx:subfield code="a">Grenze</mx:subfield>  
  <mx:subfield code="0">(DE-588)J4130793-8</mx:subfield>  
  <mx:subfield code="0">(I2)J54027</mx:subfield>  
  <mx:subfield code="2">swd</mx:subfield>  
</mx:datafield>  
- <mx:datafield tag="651" ind1=" " ind2="7">  
  <mx:subfield code="a">java</mx:subfield>  
  <mx:subfield code="0">(DE-588)Q428527-3</mx:subfield>  
  <mx:subfield code="0">(I2)I10205</mx:subfield>
```

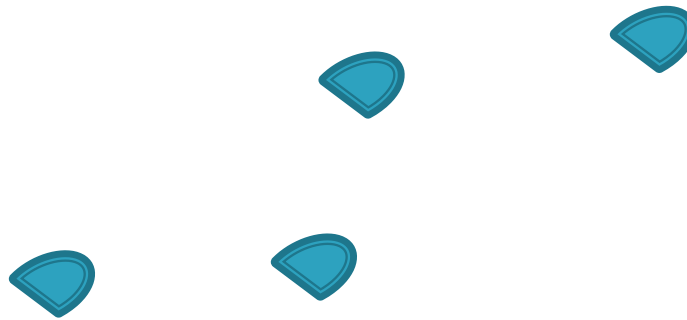
<https://opacplus.bsb-muenchen.de/title/BV021693819/marc>

BSB–Erweiterungen

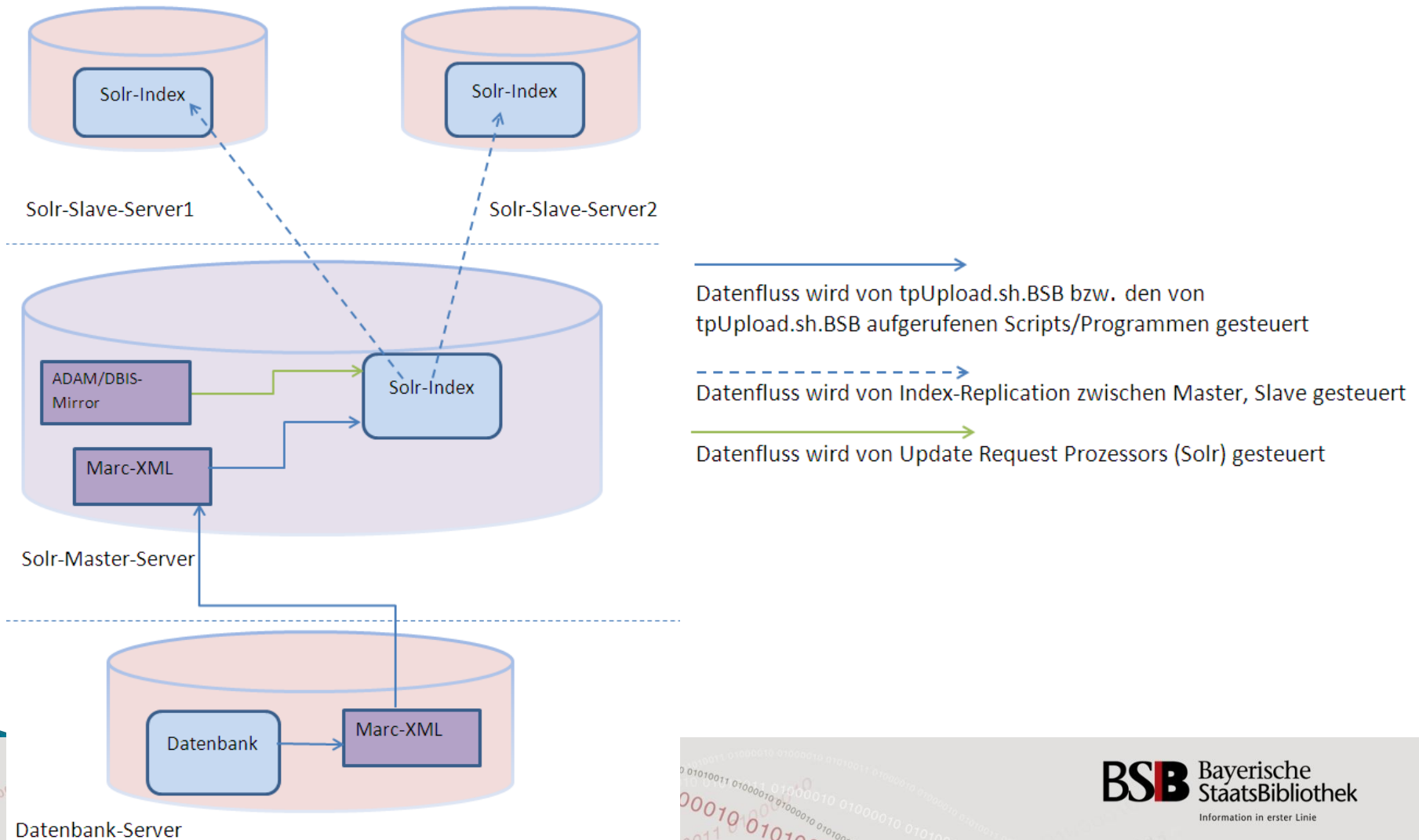
- ▶ Service–Funktionen: JSP & Taglibs
- ▶ Neue Icons für Medientypen, Zusammenführung Medientypen aus Medientyp–Facette:
JSP & Taglibs
 - Mapping d. Medientyp–Facette aus Solr auf das mit Icons abbildbare Subset
 - Prioritätenreihenfolge bei mehreren Medientypen
- ▶ MarcXml–Schnittstelle: JSP & separate Java–Webapp
- ▶ QueryLanguage: Servlet in TouchPointClient
- ▶ Weblink–Verbalisierung: JSP & Modul im Solr–GenericTarget
 - Analyse des MarcXml konfigurierbar mit XSLT
 - konfigurierbares Modul für HAN–Proxy–URLs
 - Priorisierung der Volltext–Links f. Online–Lesen–Button
 - mehrsprachige Verbalisierung

Noch zu tun von OCLC

- ▶ Behebung von Fehlern
- ▶ TouchPoint Konfiguration dateibasiert
- ▶ Autocompletion Verbesserung



Datenprocessing für Solr



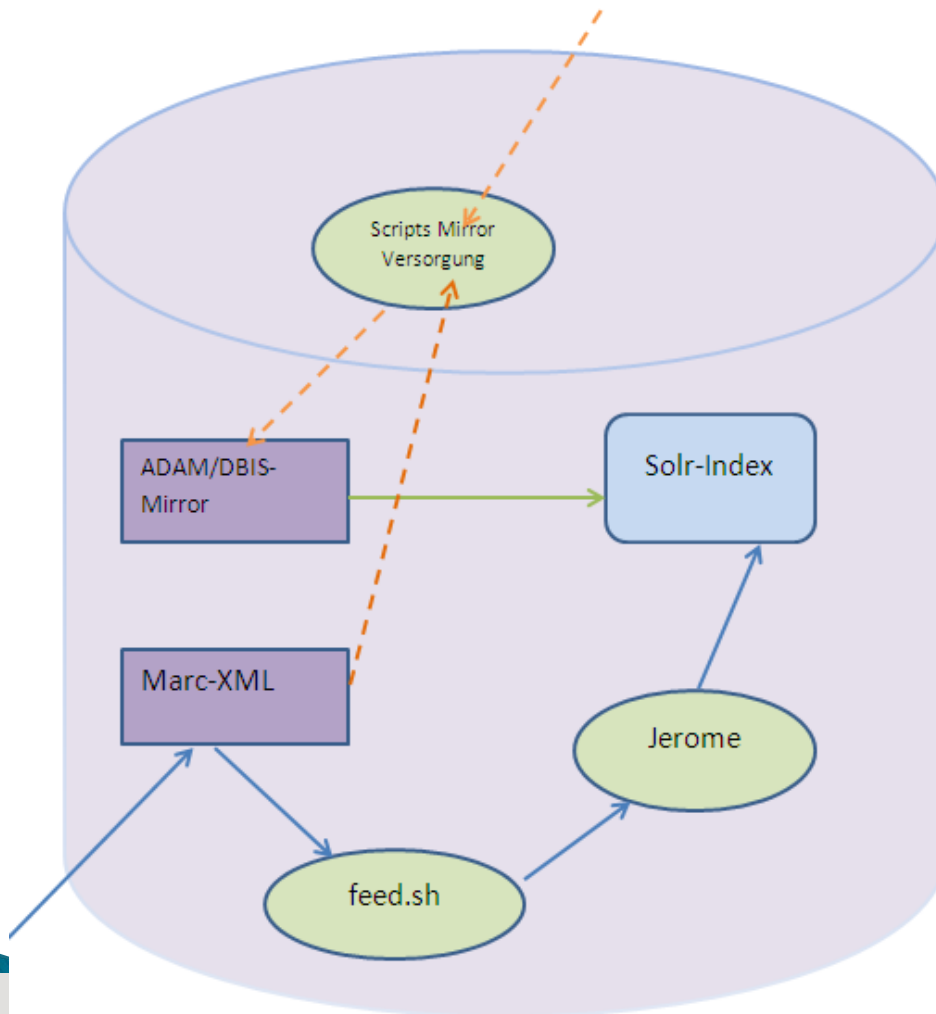
Solr-Masterserver

Das Marc-XML wird nach dem Entladen auf den Solr-Server kopiert.

Die Scripts AdamMirror.py, DbisMirror.py prüfen das Marc-XML und holen ggf. neue ADAM/DBIS-Objekte aus der ADAM-/DBIS-Schnittstelle.

feed.sh holt das Marc-XML und schickt es an Jerome. Dort wird das Marc-XML bearbeitet und an Solr geschickt.

Die ADAM/DBIS-Objekte im ADAM-/DBIS-Mirror werden via Solr-Update Request Processor im Solr-Index angereichert.



Datenbearbeitung via Jerome

- ▶ BSB-Anpassungen in XSLT-Node und XPATH-Node

- ▶ Beispiel: XPATH-Node für Feldmapping:

`xpath.author=//marc:datafield[@tag='100' or @tag='700'[@ind1='1']]/marc:subfield[@code='a']`

- ▶ Beispiel: XSLT-Node für Medientyperkennung:

```
<xsl:variable name="leader" select="marc:leader"/>
<xsl:variable name="leader7" select="substring($leader,8,1)"/>
<xsl:variable name="controlField008_21" select="substring($controlField008,22,1)"/>
<xsl:text>|</xsl:text>
<xsl:text>index:media_type=</xsl:text>
<!--Regel 1 -->
<xsl:if test="$leader7='s' and $controlField008_21='m'">
<xsl:text>Schriftenreihe</xsl:text>
<xsl:text>,</xsl:text>
</xsl:if> ...
```

..via Solr-UpdateRequestProcessor

- ▶ UpdateRequestProcessor ermöglicht die Verarbeitung der Feldinhalte mit Java. Für den BSB-Index wurden bislang 6 UpdateRequestProcessor entwickelt: Kataloganreicherung der ADAM/DBIS-Objekte, Mehrfachindexierung der ISBNs und ...
- ▶ Beispiel Einbindung UpdateRequestProcessor in solrconfig.xml:

```
<updateRequestProcessorChain>
  <processor class="de.bsb_muenchen.solr.isbn.BsblsbnExpanderProcessorFactory"/>
  <processor class="de.bsb_muenchen.solr.adam.BSBAadamExtractionProcessorFactory">
    <str name="urlField">content_url</str>
    <str name="outField">content_ce</str>
    <str name="tagName">Inhaltsverzeichnis</str>
  </processor>
  .....
</updateRequestProcessorChain>
```


Ranking

Das Ranking basiert auf Analyse der Treffermenge anhand folgender Parameter, die im Zusammenspiel miteinander die Relevanz jedes Dokuments messen. Die Konfiguration erfolgt in solrconfig.xml:

- ▶ 1. *Gewichtung der Indexfelder:*

```
<str name="qf">free_search title^1 author^0.8 subjects^0.8 content_ce^0.7 index_ce^0.7 dbis_ce^0.7 </str>
```

- ▶ 2. *Proximität:*

```
<str name="pf">title^0.5 author^0.5 subjects^0.5 content_ce^0.5 index_ce^0.5 dbis_ce^0.5 </str>
```

- ▶ 3. *Aktualität:*

```
<str name="bf">recip(ms(NOW/YEAR,freshness),3.16e-11,1,1)^10.0</str>
```

- ▶ 4. *Filterquery*

```
<str name="bq">statusband:false ^50 </str>
```

Performance/Verfügbarkeit

- ▶ Master/Slave Solr-Server
- ▶ Loadbalancer
- ▶ Einrichten Monitorquery mit TouchPoint-Syntax in Loadbalancer
- ▶ Indexfetch auf die beiden Slaves via crontab asynchron einstellen
- ▶ Warming Queries

Kenndaten Solr

- ▶ **Serverdaten**

VMware

4 Kerne

32 GB RAM

700 GB Festplatte

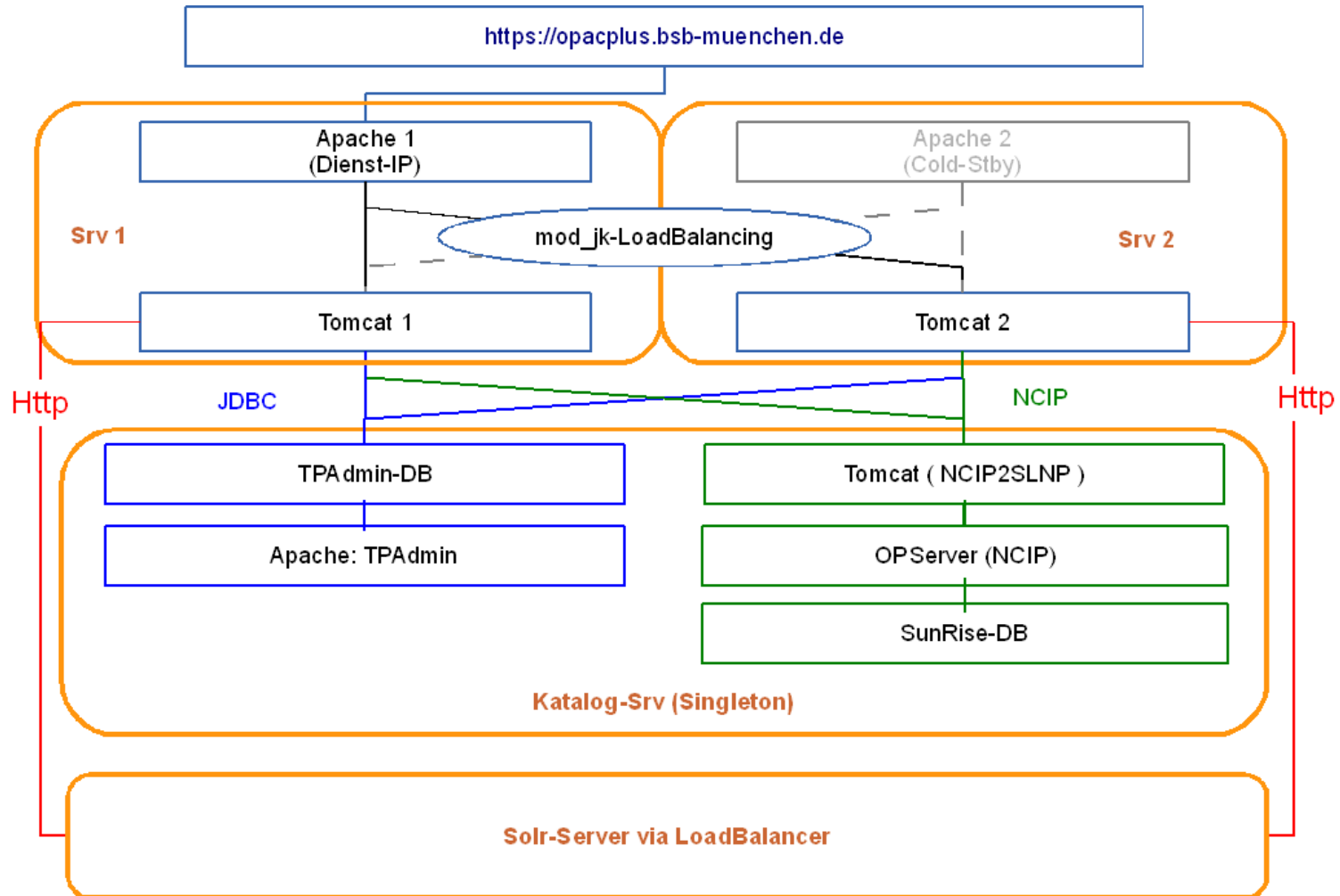
- ▶ **Indexdaten**

ca. 12 Millionen Dokumente + 700 TSD Adam-Obj

Indexgröße ca. 97 G

Indexierungszeit für Kompletindex ca. 18 Stunden

TouchPoint-Architektur



Kenndaten TP

▶ Tomcat-Server:

VMware

4 Kerne

8 GB RAM

30 GB Festplatte

Ausblick

- ▶ Einbindung des Discovery Service
- ▶ Integration der Volltextsuche
Beta = <http://opac-plus-volltext.bsb-muenchen.de>
- ▶ Hierarchische Facette – Systematische Suche
– Checkboxes
- ▶ Indexierung Daten aus Fremdquellen (PDA)

Danke für Ihre Aufmerksamkeit