

Technische Aspekte der Systemarchitektur von PROBADO



Dr. Harald Krottmaier, TU Graz

Agenda

- Warum PROBADO?
- Wie?
- Aktueller Projektstand
- Ausblick
- Zusammenfassung

Warum PROBADO?

- Aktuelle Probleme (Auszug...)
 - Immer mehr mediale Formate
 - Multimedia Objekte
 - Schwer zu beschreiben...
 - Wie danach suchen?
 - Metadaten?
 - Integration in bestehende Prozessketten

Projektparameter

- DFG-Leistungszentrum für Forschungsinformation
 - Themenfeld „Digitale Text- und Datenzentren“
- *Entwicklung von Werkzeugen für den prototypischen Betrieb eines Dienstes zur automatischen Erschließung, Lagerung und Bereitstellung von nicht-textuellen Dokumenttypen*
- Projektstart: 1.2.2006
- 6 Partner
- Projektlaufzeit: 5 Jahre

Partner

- OFFIS
 - Prof. Appelrath
- TU Graz
 - Prof. Fellner
- Uni Bonn, Informatik III
 - Prof. Clausen
- Uni Bonn, Informatik II
 - Prof. Klein



Partner

- Technische Informationsbibliothek
Universitätsbibliothek Hannover
 - Dir. Rosemann, Dr. Sens



- Bayerische Staatsbibliothek München
 - Dr. Griebel, Scholz



Weiterer Partner (ohne Förderung):

- TU Darmstadt /
Fraunhofer IGD



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Fraunhofer
Institut
Graphische
Datenverarbeitung



Wie?

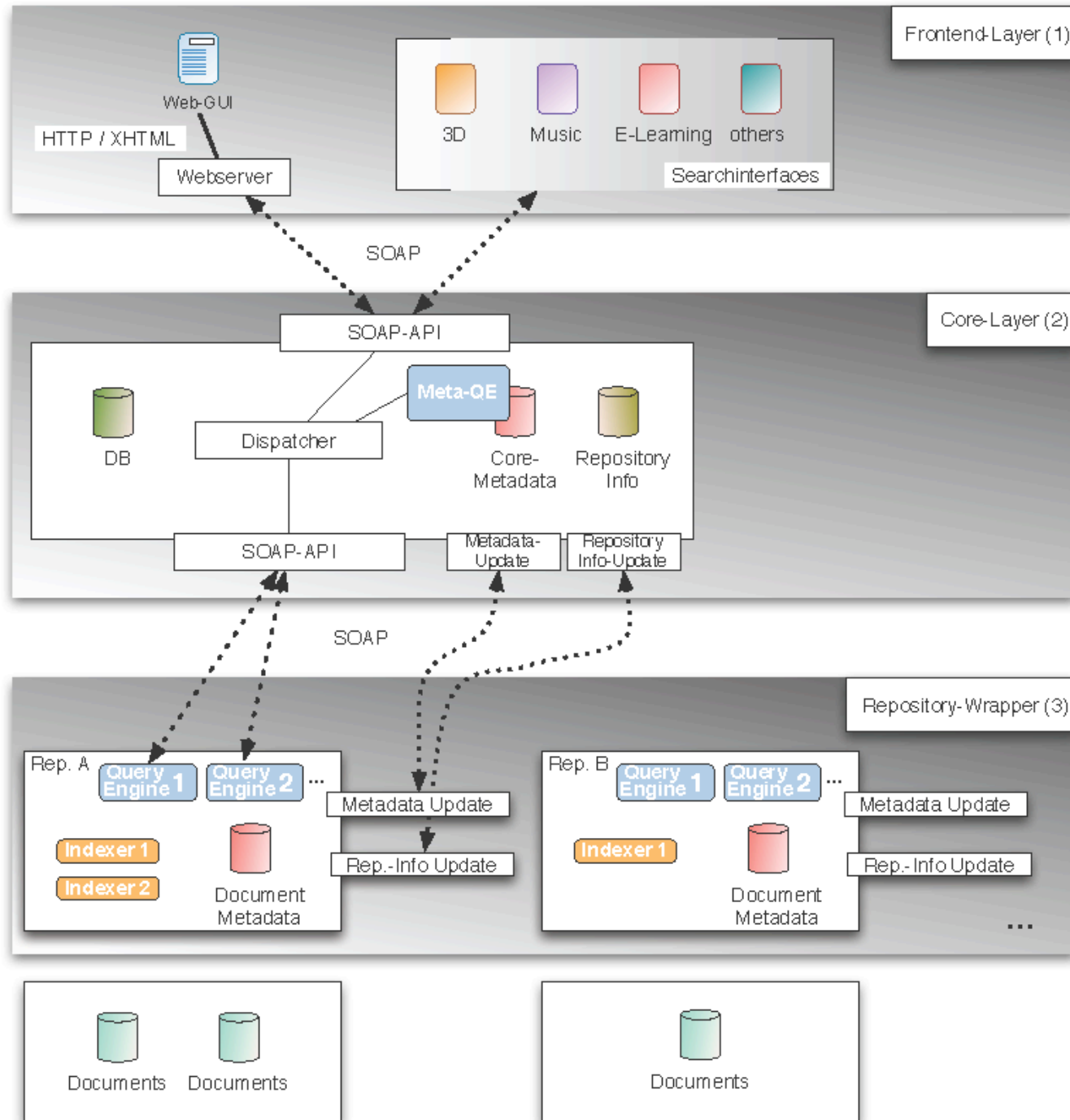
- Framework
- Bestehende Welten verbinden...
 - Bestehende Algorithmen für die Praxis
- Skalierbarkeit
- Einfache
 - Erweiterbarkeit
 - Integration
 - Benutzbarkeit

Wie?

- Suchverläufe
- Verwalten von Metadaten
- Systeminformation
- Anbindung von
 - Diensten
 - bestehenden Repositorien

Systemarchitektur

- drei Schichten
 - Benutzerschnittstelle
 - PROBADO-Systemschicht
 - Repositorien-Anbindung
- Kommunikation via Webservices
 - freie Systemscheidung
 - „nur“ Implementation der API



Aktueller Projektstand

- Metadatenschemata
- Relationale Datenbanken
- Protokoll zwischen den Schichten
 - Bekannte/Funktionierende Modelle
 - z.B. Schicht 2-3
 - Anlehnung an OAI-PMH
 - teilweise bereits implementiert

Anfragearten

- 5 Typen
 - 2 x im Core
 - 3 x im Repository

Anfrage...

- z.B. Suche in Metadaten

```
<xsd:complexType name="ProbadoMetadataSearchType">
<xsd:sequence>
<xsd:element name="Metadata" minOccurs="1"
  maxOccurs="unbounded">
...
<xsd:element name="SortBy" minOccurs="0"
  maxOccurs="unbounded">
...
<xsd:element name="CoreSearch" ...
<xsd:element name="RepositoryId" ...
<xsd:element name="StartIndex" ...
<xsd:element name="Count" ...
<xsd:element name="SessionId" ...
...
```

...Antwort

- Suchresultat

```
<xsd:complexType name="ProbadoMetadataSearchResultType">
  <xsd:sequence>
    <xsd:element name="TotalResultsCount" type="xsd:int" minOccurs="0"
      maxOccurs="1"/>
    <xsd:element name="ResultElement" minOccurs="0" maxOccurs="unbounded">
      <xsd:complexType>
        ...
        <xsd:element name="DocumentId" ...
        <xsd:element name="Ranking" ...
        <xsd:element name="Title" ...
        <xsd:element name="Accessible" ...
        <xsd:element name="ContextInfo" ...
        <xsd:element name="LinkToDocument" ...
        <xsd:element name="LinkToPreview" ...
        <xsd:element name="LinkToThumbnail" ...
        <xsd:element name="RepositoryId" ...
        <xsd:element name="Metadata" minOccurs="0" maxOccurs="unbounded">
          <xsd:complexType>
            ...
```

div. Entscheidungen...

- Framework: Cocoon
- Kommunikation: SOAP
- Inhalte über XML-Ströme
- REST vs. SOAP
- Relationale Datenbanken

Ausblick

- Dokumentation der bisherigen Arbeiten
- Weiterentwicklung der Suchanfragen
- Einpflegeinterfaces
- Prototyp in Endversion überführen
 - Pareto: 80:20...
- Tests mit Echtdaten

Zusammenfassung

- Warum gibt's PROBADO?
 - Wie wird es umgesetzt?
 - Wo stehen wir im Moment?
 - Wohin gehen wir?
-
- Wann erreichen wir unsere Ziele?

Kontakt

- Dr. Harald Krottmaier
- Institut für Computergraphik und Wissensvisualisierung
- TU Graz
- harald.krottmaier@tugraz.at