

Implementing Federated Search Across Educational Digital Libraries

Martin Koning-Bastiaan, MERLOT

Andy Dong, SMETE.ORG/UC Berkeley

Alyssa LaLanne, MERLOT

Brandon Muramatsu, MERLOT & SMETE.ORG/UC Berkeley

Sam Shamseldin, MERLOT



www.smete.org

CONNECTING
Science
Mathematics
Engineering
Technology
Education



Originally Published 2002. Republished 2013. This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License (<http://creativecommons.org/licenses/by-nc-sa/3.0/us/>)

Workshop Agenda

- 2:30-2:35 Introductions
- 2:35-2:45 Goals of the workshop and Introduction to Federated Search
- 2:45-3:00 Technology and Specifications
- 3:00-3:15 Demos of MERLOT and SMETE.ORG
- 3:15-3:30 Break
- 3:30-4:00 Writing a Client
- 4:00-4:30 Discussion: Issues in Federated Search

What is “federated search?”

- **Search across multiple collections, whose item-level metadata have been “federated” into a “union catalog”**
- **Strategies**
 - Metadata harvesting: Build a centrally-located union catalog by “harvesting” item-level metadata from multiple remote collections
 - Open Archives Initiative Protocol for Metadata Harvesting
 - Distributed search: Build a “virtual union collection” by querying multiple remote repositories based on user-initiated action
 - Z39.50

Advantages and Disadvantages

- **Metadata harvesting**

- Advantages

- Strong guarantee of “quality of service” to end-users

- Disadvantages

- Requires central repository to physically catalog and store item-level metadata
 - Timeliness of item-level metadata at central repository dependent on reliability and frequency of harvesting

- **Distributed search**

- Advantages

- Remote collection retains tighter control over re-distribution of item-level metadata
 - Real-time changes to item-level metadata reflected immediately

- Disadvantages

- Potentially lower quality of service (time-outs, Internet outages)

Apache AXIS

- **Java Servlet Implementation**
- **Low overhead to create clients**
- **WSDL**
 - <http://soap.smete.org/axis/services/urn:SmeteSearch?wsdl>
 - <http://dev.merlot.org:9090/searchservice/services/MerlotSearch?wsdl>
- <http://xml.apache.org/axis/>

PERL SOAP::Lite

- **PERL-based SOAP Server and Client**

- **Simple!**

```
#!/usr/local/bin/perl -w
use SOAP::Lite;
print SOAP::Lite
    ->uri("urn:SmeteSearchAction")
    ->proxy("http://soap.smete.org:80/axis/servlet/AxisServlet")
    ->doSmeteSearch(SOAP::Data->name("key" => "2222"),
        SOAP::Data->name("q" => "volcano"),
        SOAP::Data->name("start" => 1),
        SOAP::Data->name("maxResults" => 10),
        SOAP::Data->name("lr" => "en-US"))
```

- **Resources**

- <http://www.soaplite.com/>

- **Limitations**

- SOAP::Lite not able to handle complex types

Contact Info

Martin Koning-Bastiaan, mjkb@merlot.org

Andy Dong, adong@smete.org

Brandon Muramatsu, mura@smete.org

