

Breakout Session 3

Metasearch

Authentication and Authorization

What is “Metasearching” ?

- When we talk about “Metasearching”, we mean searching the content of multiple repositories at once
- 2 approaches:
 - “Federated” search by harvesting all searchable content and indexing it in one place
 - “Parallel” (distributed) search by letting the repositories do the indexing and searching on request

Pros and Cons

- “Federated” approach
 - It's usually faster
 - It's more reliable (no network needed)
 - No implementation needed on repository side
- “Parallel” approach
 - Better for unstable collections
 - No transport of the resources' data needed

Possible Solutions

- Existing standard: Z39.50,
But: very heterogenous configurations possible, so lots of standardization would be needed here
- Extending OAI-PMH:
Con: Some implementations of OAI-PMH just can't be extended (commercial, not self-written)
Pro: If it can be extended, lots of code can be reused

Authentication & Authorization

- Why is this needed with “open” archives?
 - to prevent abuse
 - to offer individual responses
 - closed groups working with (not yet) open available documents

Existing and possible solutions

- IP-based access control
 - Problems with proxies
 - impossible to see behind the IP
 - But: simple implementation
- Other solutions:
 - Centralized Authentication services (Shibboleth, MS Passport, ...)
 - Certificate (chains), password, Athens, PKI
- Hard to integrate in the OAI-PMH