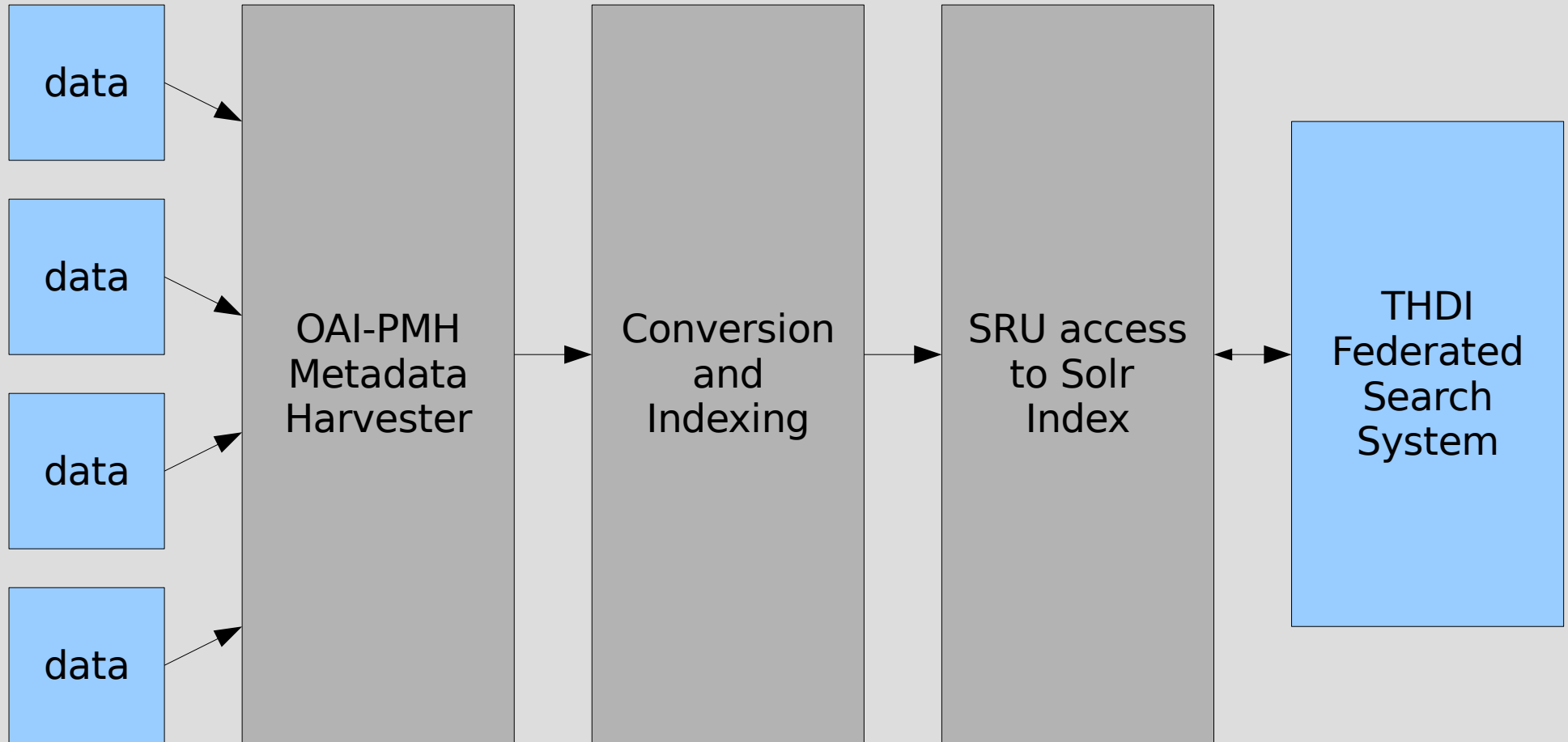




# Facet Forward: Faceted Navigation of Federated Search Results for Cultural Heritage Materials

Mark E. Phillips  
LITA National Forum 2007  
Denver, Co

# OAI-PMH



# OAI-PMH

data-provider - (institutions)  
service-provider – (UNT Libraries')

Harvester  
OAI-PMH Repository

6 verbs

GetRecord

Identify

ListIdentifiers

ListMetadataFormats

ListRecords

ListSets

# Getting Metadata Records

Harvesters:

Many available at [openarchives.org](http://openarchives.org)

Customized version of OCLC's one page harvester.

# Harvested Records

```
<?xml version="1.0"?>
<repository>
  <record>
    <header>
      <datestamp>2007-08-30T00:00:00Z</datestamp>
      <identifier>oai:thc.state.tx.us:4300000079</identifier>
      <setSpec>THCCH</setSpec>
      <setSpec>institution:THC</setSpec>
    </header>
    <metadata>
      <identifier>http://atlas.thc.state.tx.us/common/viewform.asp?atlas_num=4300000079</identifier>
      <relation>ispartof: Texas Historical Commission County Courthouses</relation>
      <subject>County Courthouse</subject>
      <creator>Townes & Funk</creator>
      <subject>Moderne</subject>
      <description>4-story limestone building with jail facilities on fourth floor. Aluminum frame windows set in vertical relief emphasize the vertical. Entrance loggia features modern fluted columns(no capitals). The overall building is made of smaller rectilinear par</description>
      <date>1939</date>
      <title>Childress County Courthouse</title>
      <coverage>United States - Texas - Childress County - Childress</coverage>
    </metadata>
  </record>
  <record>
    ...
  </record>
</repository>
```

□

# Indexing OAI Records

Python conversion scripts

Convert from harvested temp format to  
Solr format

Send to Solr for indexing

Solr Index

- Dublin Core based document structure

- MD5 hash of OAI Identifier as doc id

- Fulltext and several sorting fields

# SRU Support

mod\_python based SRU access  
CQL conversion to Lucene query was  
challenging  
Less than 1 second from 10 simultaneous  
requests.

# Challenges

Thumbnail support  
Simultaneous SRU requests.  
Getting repositories to support OAI-PMH  
Customization of OAI-PMH output