

10-20-2008

Integrating Dissertation Metadata and INDURE Using Z39.50

Vijendra Purohit

Purdue University, vpurohit@purdue.edu

Michael Witt

Purdue University, mwitt@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_research

Purohit, Vijendra and Witt, Michael, "Integrating Dissertation Metadata and INDURE Using Z39.50" (2008). *Libraries Research Publications*. Paper 97.

http://docs.lib.purdue.edu/lib_research/97

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

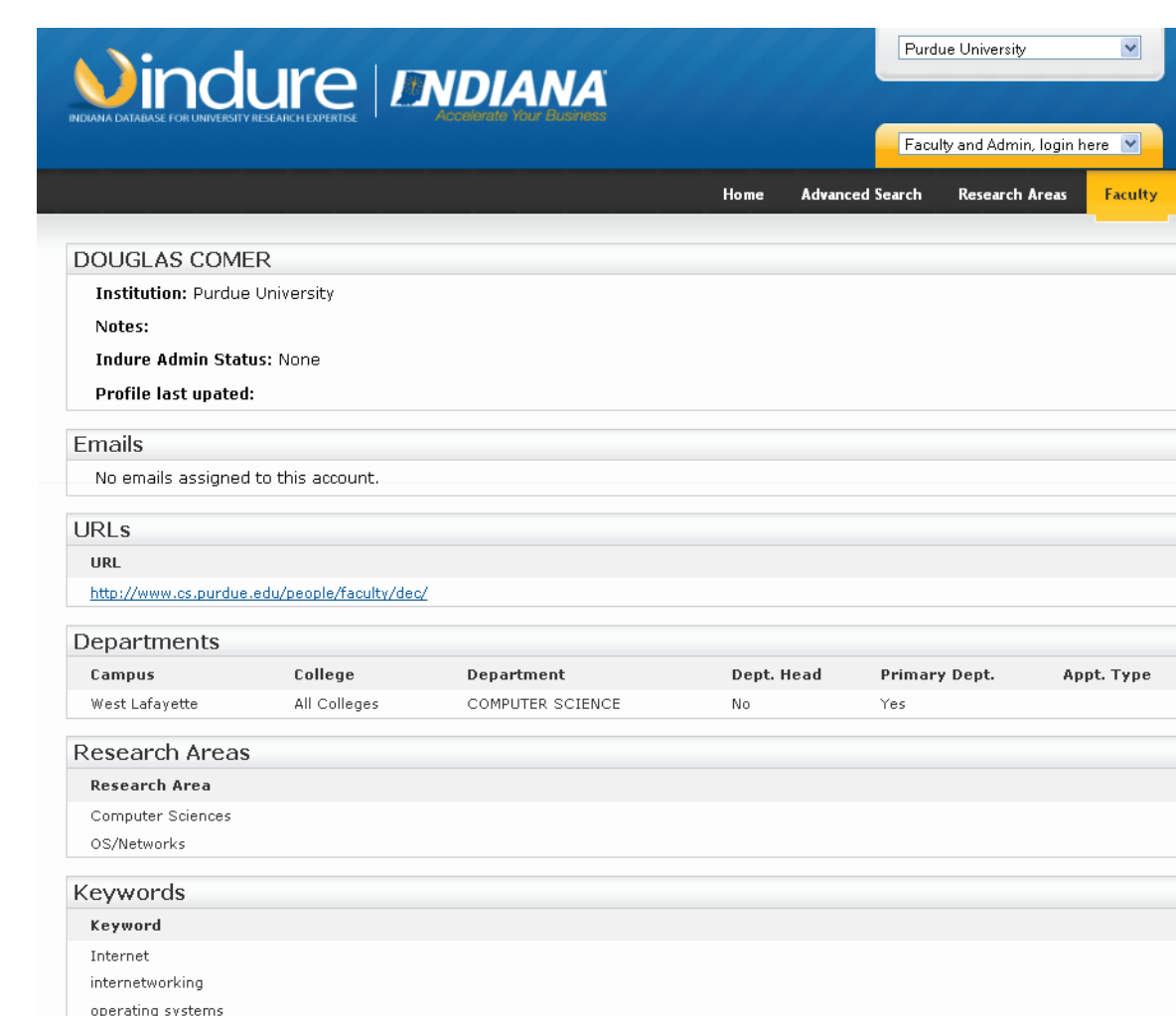
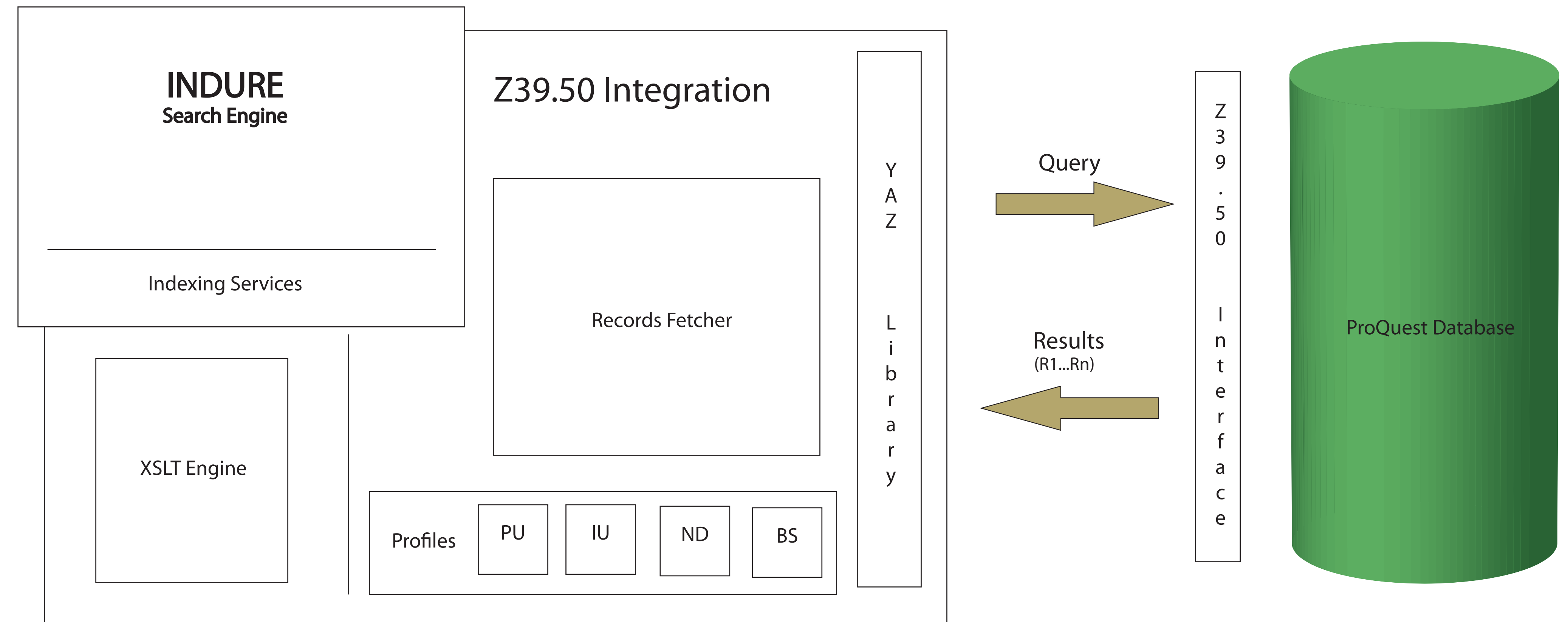
The Indiana Directory of University Research Expertise (INDURE) [1] is an easy-to-access database of experts, intellectual property, and on-going sponsored research projects at academic institutions across the state of Indiana. Researchers from Purdue, Ball State, Indiana Universities and the University of Notre Dame can be located by searching INDURE by their individual research interests, expertise, and collaborations.

The majority of these researchers are faculty members. One strong descriptor of faculty's expertise can be found by analyzing the dissertations published by the doctoral students whom they advised. Through a database service licensed from ProQuest LLC, the Purdue University Libraries has access to the dissertations from these four institutions.

The purpose of this project is to query and harvest detailed, structured metadata describing these dissertations using the Z39.50 protocol [2] and to include this information in INDURE to improve the accuracy of its search engine.

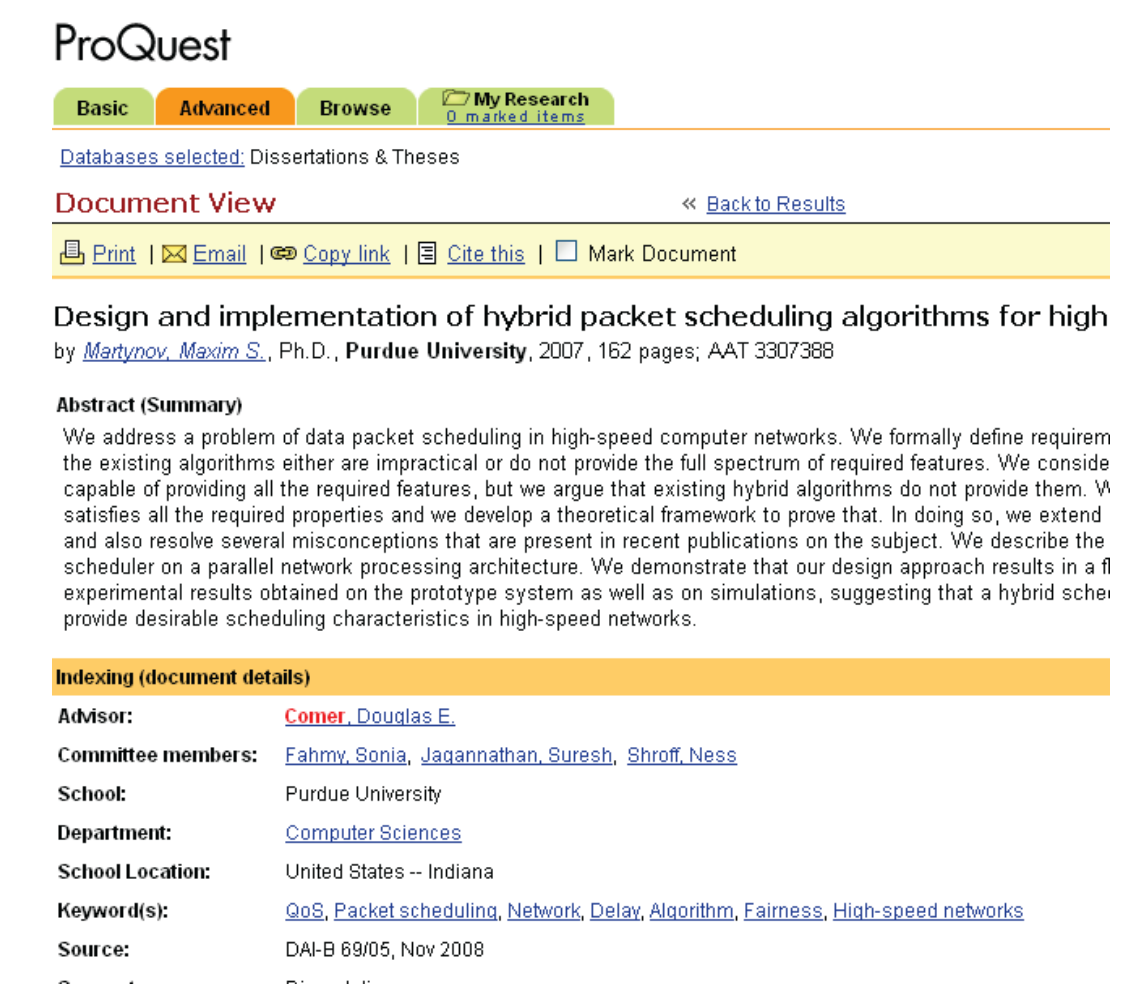
The integration module that we are developing utilizes the YAZ client with connection profiles for the four different institutions. These are used to fetch batches of MARC [3] records from a Z39.50 interface that is exposed by the ProQuest server. The records contain the dissertations' titles, authors, advisors, and abstracts. These records are converted to qualified Dublin Core [4] using the oai_dc schema and an XSLT engine. The oai_dc records are then indexed by the INDURE search engine.

The INDURE project is led by Professor Aditya Mathur, Department of Computer Science at Purdue University, with support from the Indiana Economic Development Corporation.



```

- <collection xsi:schemaLocation="http://www.loc.gov/MARC21/dim http://www.loc.gov/standards/marcxml/schema/MARC21.xsd" >
  <record>
    <header>01853cam 2200133 4500</header>
    <controlfield tag="001" value="000000"></controlfield>
    <datafield tag="035" ind1="1" ind2="1">
      <subfield code="a">1537026941</subfield>
    </datafield>
    <datafield tag="500" ind1="1" ind2="1">
      <subfield code="a">Text available - WC</subfield>
    </datafield>
    <datafield tag="245" ind1="1" ind2="1">
      <subfield code="a">
        Design and implementation of hybrid packet scheduling algorithms for high speed networks
      </subfield>
    </datafield>
    <datafield tag="100" ind1="1" ind2="1">
      <subfield code="a">Martynov, Maxim S.</subfield>
    </datafield>
    <datafield tag="510" ind1="1" ind2="1">
      <subfield code="a">ProQuest Dissertations and Theses</subfield>
    </datafield>
    <datafield tag="9" ind1="1" ind2="1">
      <subfield code="a">2007</subfield>
    </datafield>
    <datafield tag="9" ind1="1" ind2="1">
      <subfield code="a">162 pages</subfield>
    </datafield>
  </record>
</collection>
- <oai_dc:dc Collection xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc/" >
  <dc:title>
    Design and implementation of hybrid packet scheduling algorithms for high speed networks
  </dc:title>
  <dc:creator> Martynov, Maxim S. </dc:creator>
  <dc:type>text</dc:type>
  <dc:language> </dc:language>
  <dc:description>Text available - WC</dc:description>
  <dc:description>ProQuest Dissertations and Theses</dc:description>
  </dc:description>
  We address a problem of data packet scheduling in high-speed computer networks. We form that the existing algorithms either are impractical or do not provide the full spectrum of required features, but we argue that existing hybrid algorithms satisfies all the required properties and we develop a theoretical framework to prove that. In doing so, we extend and also resolve several misconceptions that are present in recent publications on the subject. We describe the scheduler on a parallel network processing architecture. We demonstrate that our design approach results in a desirable scheduling characteristics in high-speed networks.
  </dc:description>
  <dc:description>DISSERTATION</dc:description>
  <dc:subject>Computer science</dc:subject>
</oai_dc:dc>
</oai_dc:Collection>
  
```



[1] INDURE, <http://www.indure.org>

[2] Z39.50, <http://www.loc.gov/z3950/agency>

[3] MARC, <http://www.loc.gov/marc>

[4] Dublin Core, <http://dublincore.org>

Integrating Dissertation Metadata and INDURE Using Z39.50

Vijendra Purohit (vpurohit@purdue.edu) & Michael Witt (mwitt@purdue.edu), Purdue University, West Lafayette, Indiana, USA