Introduction to Digital Libraries

hussein suleman
uct cs honours 2003

“Web Services”

Definitions

- Web
  - = World Wide Web
  - = client-server hypermedia system layered over Internet.
- Web service
  - = Web-based Service (or service)
  - = service provided through the medium of the Web, beyond hypermedia.
- Web Service
  - = machine-to-machine communication based on interoperability standards defined by W3C.

Web-based Systems

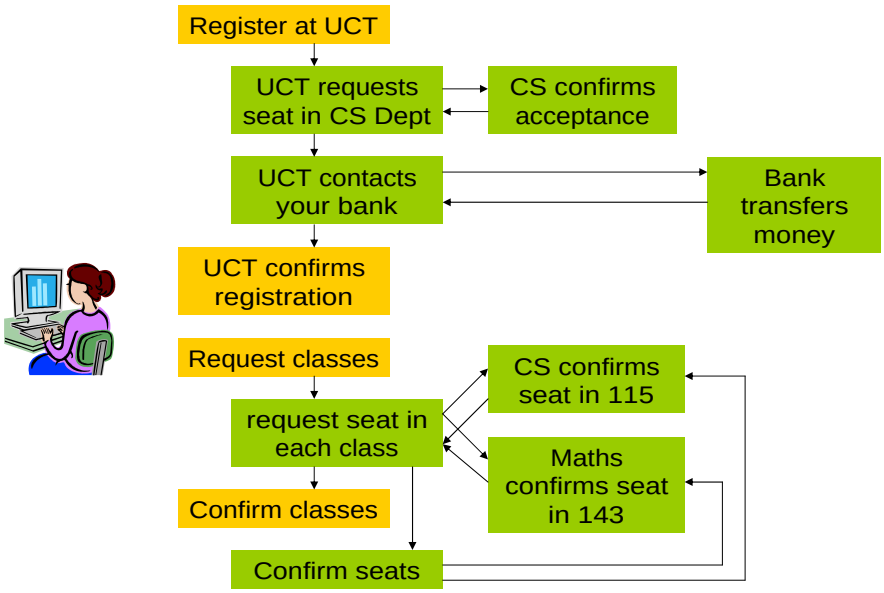
- Modern digital libraries are largely Web-based services.
  - They provide higher-level services beyond just document delivery, e.g., searching.
- Greenstone, ODL, EPrints, etc. are Web-based.
- Web-based architectures provide a machine-independent service interface (HTTP) and user interface (HTML).
- A network is NOT required (e.g., Greenstone on a single machine).

Example: Web-based System

Web Services

- Web Services are a specific set of standards for the description of machine-to-machine communication (possibly over the Web):
  - SOAP
  - WSDL
  - UDDI
  - WS-Choreography
  - WSFL
  - etc.

Use case: Registration



Mapping to Web Services

- SOAP
  - Transport protocol for message passing.
- WSDL
  - Service interface description for BANK, UCT, Dept1, Dept2, CS.
- UDDI
  - Central registry of all services e.g., CS's "did we accept you" service.
- WSFL, WS-Choreography, etc.
  - Description of sequence of operations to perform a task.
- Web-based services for user interaction!

SOAP

- Describes an XML format for messages to be exchanged among actors.
  - One-way information transfer.
  - Stateless.
  - Independent of lower-layer transport.
  - Specifies only syntax, not semantics.
- Used to be “Simple Object Access Protocol”.
- Currently the ONLY Web Services standard!

SOAP Example

```
<Envelope>
   <header>
      <origin>CS</origin>
   </header>
   <body>
      <requestSeat>
         <class>CSC115</class>
         <studentID>abc123</studentID>
      </requestSeat>
   </body>
</Envelope>
```

SOAP Actors and Roles

- Producer
  - Creates SOAP message.
- Consumer
  - Uses/interprets/understands SOAP message.
- Intermediary
  - Modifies message in transmission path between producer and consumer.
- Role attributes can specify if message block must be understood by next or final node.

SOAP headers and bodies

- Headers may contain control information that may be modified by intermediaries.
  - Can have multiple headers.
  - Headers can be modified, removed and inserted by intermediaries.
- Body contains application payload.
  - Only one body per message.
- Many simple use cases do not “need” headers.
- Both are defined using XML Schema.

SOAP RPC

- Remote Procedure Calls can be implemented using SOAP messages for the request and response.
- Procedures map to service names and parameters map to XML-encoded data structures.
- WSDL can be used to describe the interaction sequences more formally.

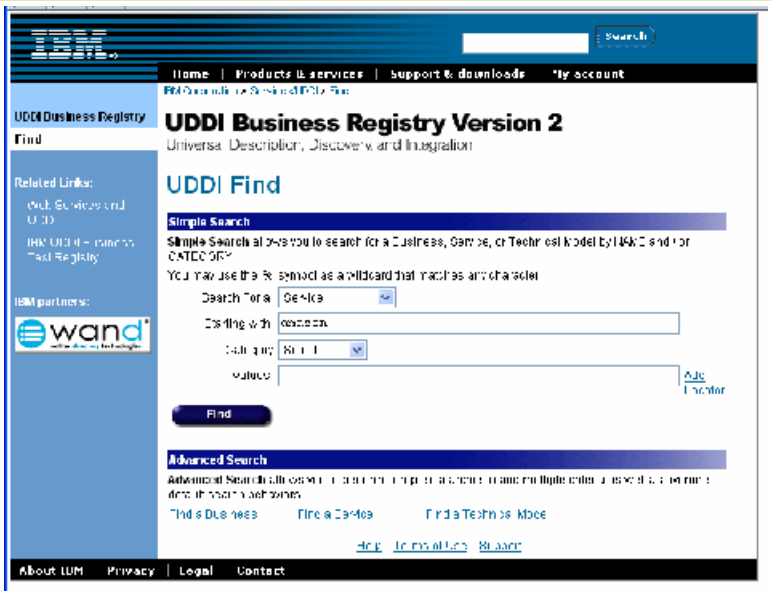
WSDL

- Web Services Description Language defines interactions among WS nodes in terms of:
  - corresponding message pairs or sequence of messages.
  - message types.
  - protocol binding (HTTP or SOAP).
  - network address of service.
- Using a WSDL definition, a client can determine how to use a Web Service (syntactically).

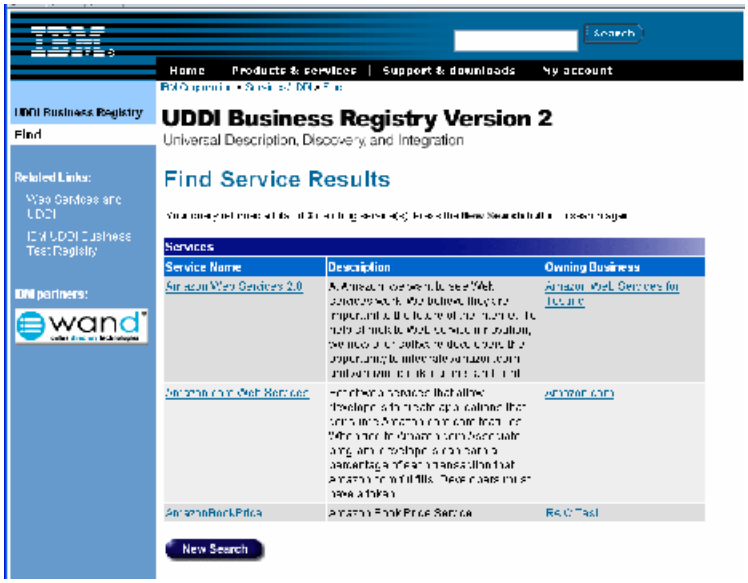
UDDI

- The Universal Description, Discovery and Integration of Web Services (UDDI) is a central/replicated clearinghouse for Web Services.
- UDDI is not machine-readable – needs human intervention to select services.

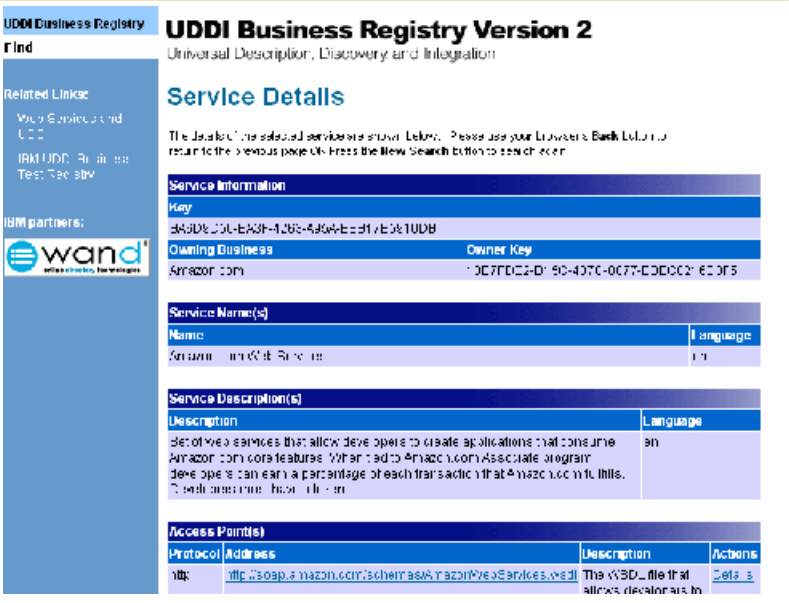
UDDI Example 1/4



UDDI Example 2/4



UDDI Example 3/4



UDDI Example 4/ 4 (WSDL)

```
<!
  interfaces are subject to
-->
<definitionsname="AmazonSearch" targetNamespace="urn:PI/DevCentral/SoapService">
<!--
 Data structures used in Amazon.com's Web Services calls and responses
 -->
 <types>
   <xsd:schematargetNamespace="urn:PI/DevCentral/SoapService">
   <xsd:complexTypename="ProductInfo">
     <xsd:all>
       <xsd:elementname="Details" type="typens:DetailsArray"/>
       </xsd:all>
     </xsd:complexType>
     <xsd:complexTypename="Details">
       <xsd:all>
       <xsd:elementname="Url" type="xsd:string"/>
       <xsd:elementname="Asin" type="xsd:string"/>
       <xsd:elementname="ProductName" type="xsd:string"/>
       <xsd:elementname="Catalog" type="xsd:string"/>
       <xsd:elementname="KeyPhrases" type="typens:KeyPhraseArray"/>
       <xsd:elementname="Artists" type="typens:ArtistArray"/>
       <xsd:elementname="Authors" type="typens:AuthorArray"/>
```

WSFL / WS-Choreography

- Web Services Flow Languages defines composition of services based on sequencing and aggregration.
  - specifies relationships and dependencies among services.
- WS-Choreography is a standards activity that is still very new!

Web Services and Digital Libraries

- Web Services only provide syntax – semantics are application-specific.
- Digital library applications will migrate to Web Services technology after standardisation. For example:
  - ZING (Z39.50 International Next Generation) is defining the SRW protocol using SOAP.
  - OAI-PMHv2.0 was designed to easily migrate to SOAP (SOAP became a standard on 24 June 2003).