Web Services interfaces to Discovery Service GO-ESSP 2005

Marta Gutierrez

Ray Cramer, Philip Kershaw, Kerstin Kleese, Siva Kondapalli, Sue Latham, Bryan Lawrence, Dominic Lowe, Roy Lowry, Kevin O'Neill, Ag Stephens, Andrew Woolf

Outline

- Discovery Service: Metadata involved
- How we store/access discovery metadata
- How we present discovery metadata
- How you can access our discovery metadata

...from the A,B,C,D taxonomy

D Discovery Metadata: Suitable for harvesting, Public Access Domain.

Dublin Core, DIF, ISO 19115 (XML Format)

NDG discovery repository:

- BADC (British Atmospheric Data Centre, UK)
- BODC (British Oceanographic Data Centre, UK)
- SOC (Southampton Oceanographic Centre, UK)
- NCAR (National Centre Atmospheric Research, US)

// Other NERC data centres on the way

- PML(Plymouth Marine Laboratory), EcoGrid

NDG - OAI:

- Easy way to bring metadata records together
- Good metadata quality filter
- Maintain updated metadata

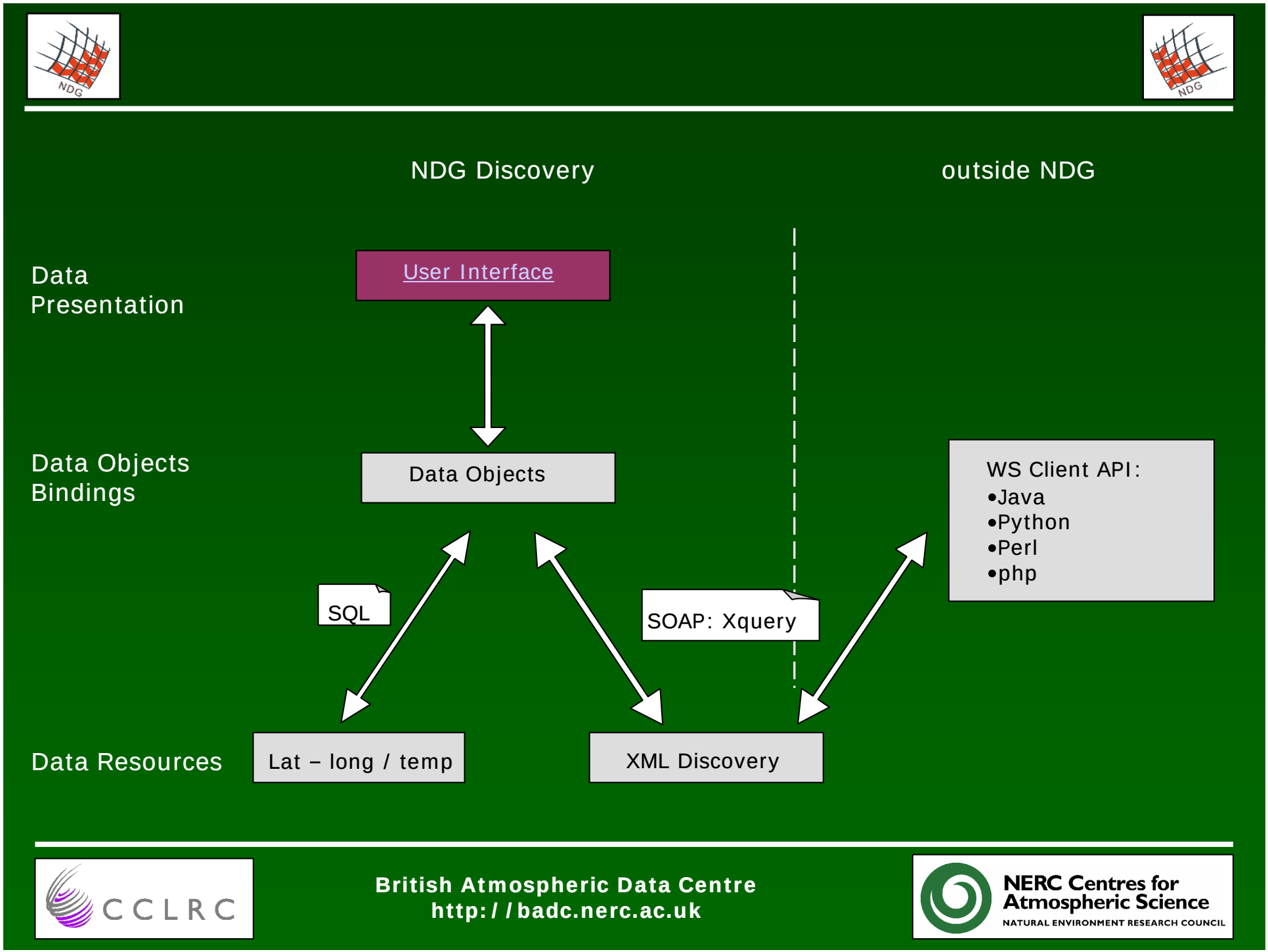
Store / Access

eXist:

- Native xmldb
- Xquery (W3C Standard)
- Full Text
- Web Services: SOAP, WSDL (W3C Standards)
- Allow for indexes on xml path elements

Postgres:

- Lat-Long & temp range queries
- Spherical coordinates (PG-Sphere)


NDG Discovery
outside NDG
Data Presentation
User Interface
Data Objects Bindings
Data Objects
WS Client API:
•Java
•Python
•Perl
•php
SQL
SOAP: Xquery
Data Resources
Lat – long / temp
XML Discovery

WSDL

Published interface: http://ndg.nerc.ac.uk/discovery/services

Demo1 (java client)

connect

xquery

retrieve

disconnect

Demo2 (python client)

connect

getResource

disconnect

Client Code (java)

```
1 QueryService service = new QueryServiceLocator();
2 Query query = service.getQuery(new URL("http://localhost:8080/exist/services/Query"));
3 String sessionId = query.connect("user", "passw");

4 String xquery = "//DIF/Title";
5 byte[] queryData = xquery.getBytes("UTF-8");
6 QueryResponse resp = query.xquery( sessionId, queryData );

7 String [] data =    new String[1];
8 data = query.retrieve(sessionId, 1, 1, true, false, "none");
9 System.out.println(data[0]);

10 query.disconnect(sessionId);
```

Client Code (python)

```
1 from ZSI import ServiceProxy

2 wsdl = 'http://localhost/exist/services/Query?WSDL'
3 query = ServiceProxy(wsdl, use_wsdl=True)

4 connectResp = query.connect(in0='ndg', in1='ndguser')
5 sessID = connectResp['connectReturn']
6 result = query.getResource(in0=sessID, in1=u'/db/dif/badc/AAOE.xml', in2=1, in3=0)
7 print result

8 query.disconnect(in0=sessId);
```

… issues

1. eXist: Scale 10^6 records? (It does 10^4),
   - Berkley DB, Xquery
   - NGS (National Grid Service) Oracle clusters XML modules.

2. SOAP: It is not a good practice to send big envelopes,
   - SOAP Attachments.

3. WS-I vs WSRF (state modelling + resources)