

The British Atmospheric Data Centre and the NERC DataGrid (for)



Bryan Lawrence on behalf of
BADC, BODC, CCLRC, PML and SOC





Outline



- Introduction to the BADC
 - (NDG Motivation)
- NDG1 Component Status
- Introduction to NDG2 Goals
- Issues

National
SNOW & ICE
Data Center

JET2000



Data at the BADC

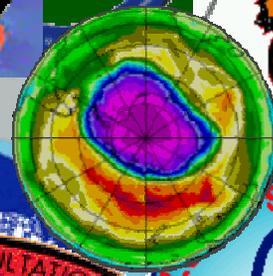


URGENT



The Met. Office
GLOBAL WEATHER SERVICES FROM THE UK

MIT
JPI



NSSDC

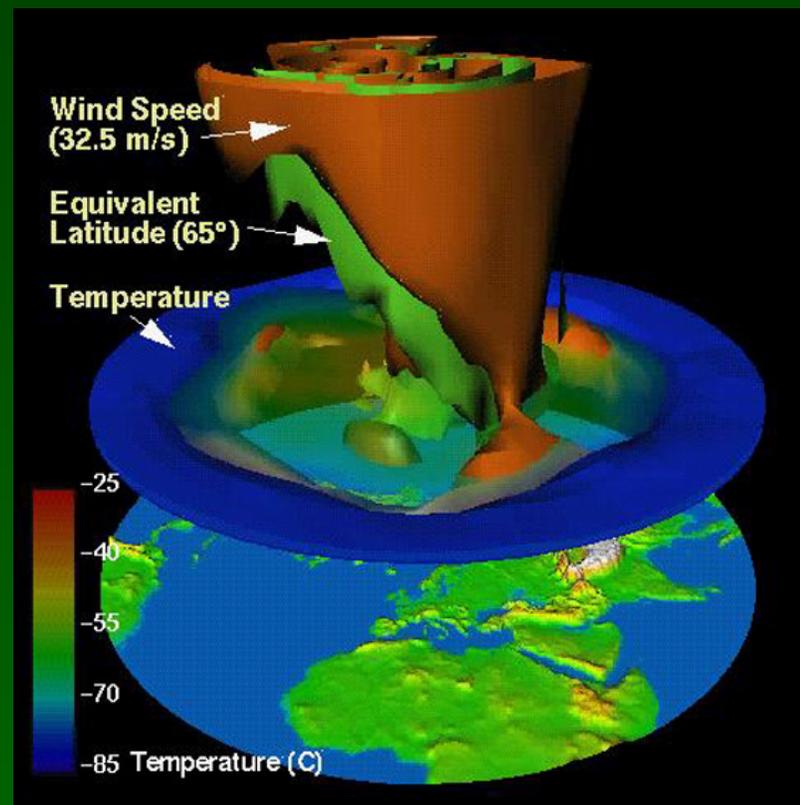
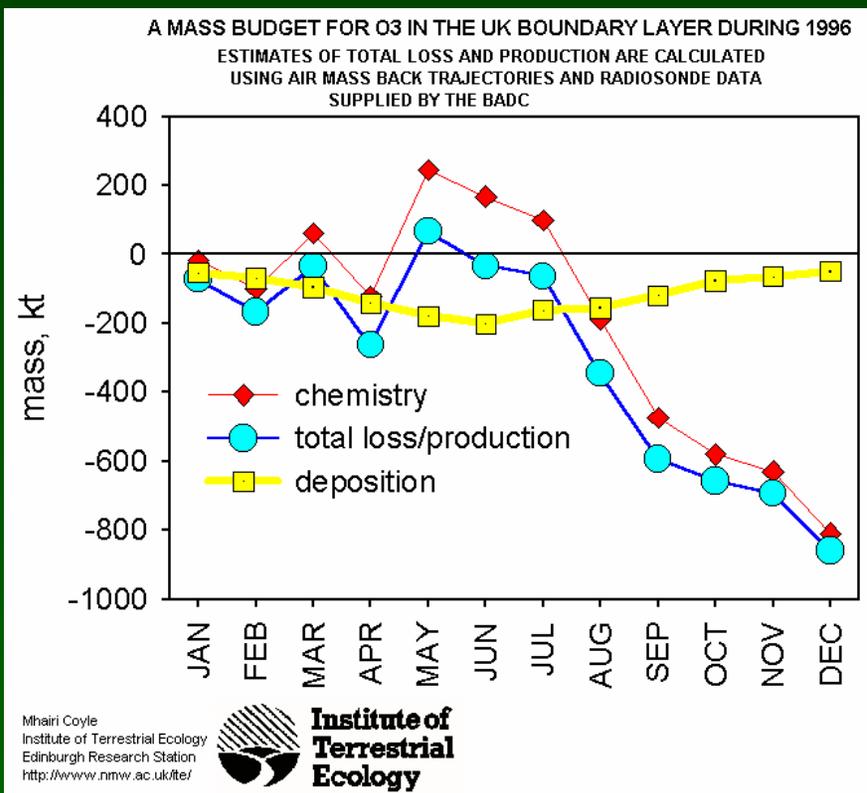
Space
Science
Data
Center



GO-ESSP, June 2005



User examples



- Pollution chemistry measurement campaigns.

- Atmospheric chemistry models.



User examples



- Bird feeding habits.





User examples



- Radio communication modelling.

- Wind power research.



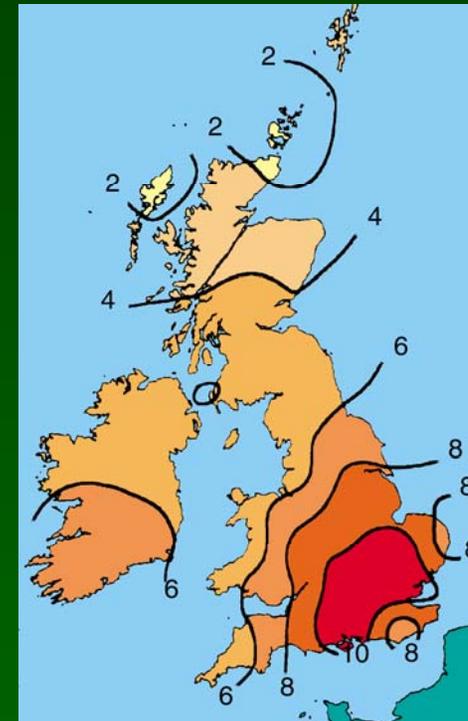
- A & E influenza cases.



User examples



- Castle mortar decay.



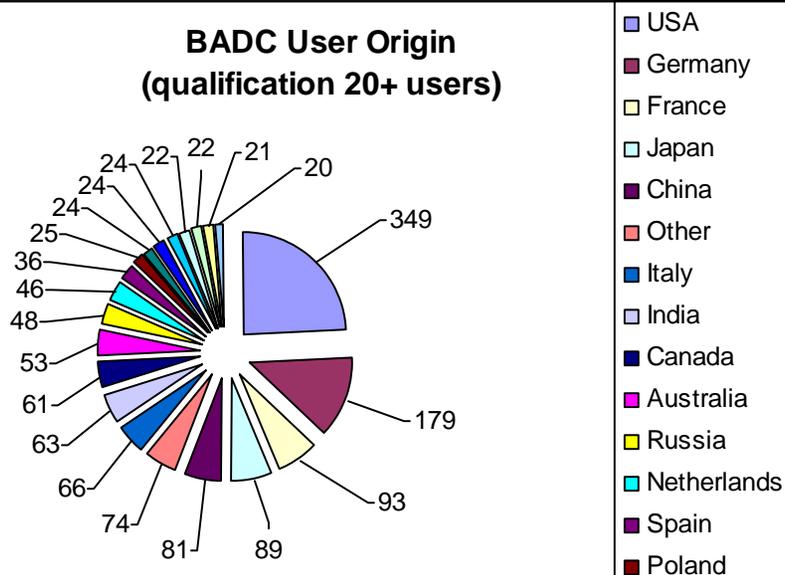
- Discomfort indices.



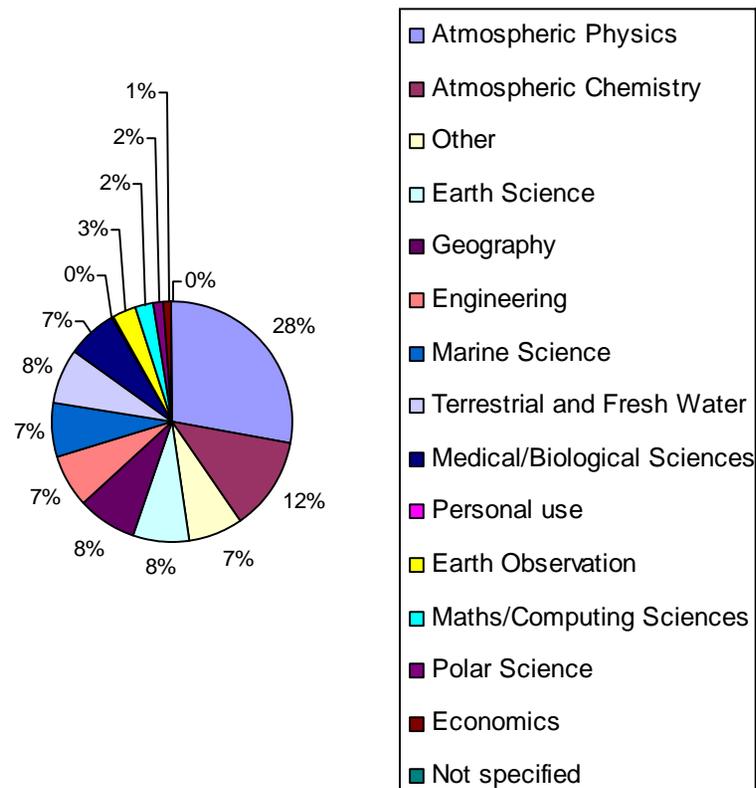
BADC Users (2004)



BADC User Origin (qualification 20+ users)



Registered Users by discipline



~ 6000 registered users

~ 1500 active last year



UTLS Ozone



- Requirements Capture
- Data Ingestion
 - Metadata Conventions
- 3rd Party Data Acquisition
 - Met Office Unified Model Analyses (with core)
 - ENVISAT
 - SLIMCAT,
 - THESEO, CARIBIC
- Trajectory Service Modifications (with core)

The screenshot shows a web browser window displaying the BADC dataset catalog. The address bar shows the URL: http://badc.nerc.ac.uk/data/list_all_datasets.html?source=data. The page lists several datasets:

- UGAMP Ozone Climatology**
Start date: 01-jan-1985 End date: 31-dec-1989
[Catalogue record](#) [Dataset web page](#) [Get data](#)
- Universities Facility for Atmospheric Measurement (UFAM)**
Start date: 01-sep-2002 End date: Not defined
[Catalogue record](#) [Dataset web page](#) [Get data](#) [Apply for access](#)
- Urban Regeneration and the Environment (URGENT)**
Start date: 01-jan-1997 End date: Not defined
[Catalogue record](#) [Dataset web page](#) [Get data](#) [Apply for access](#)
- UTLS-OZONE**
Start date: 01-jan-1999 End date: Not defined
[Catalogue record](#) [Dataset web page](#) [Get data](#) [Apply for access](#)
- VIRTEM Validation of IASI Radiative Transfer: Experiments and Modelling**
Start date: 01-jan-1999 End date: Not defined
[Catalogue record](#) [Dataset web page](#) [Get data](#) [Apply for access](#)

Archive Example



Location Edit View Go Bookmarks Tools Settings Window Help

Location: http://badc.nerc.a

BADC Testbeds Google

Home My BADC Data Search Community Help

Get Data Access Rules Submit Data Dataset Index

Home My BADC Data Search Community Help

Home My BADC Data Search Community Help

Location Edit View Go Bookmarks Tools Settings Window Help

Get Data

Username:
 Current dir:

Errata
he000
he000
he000
he000
he000
he000
he000
theseo

Home

/badc/utls/data/act/c-130/he000503.o36

OZONE MIXING RATIO (PPB)

Elapsed UTs from 0hrs on 03-MAY-00

Plot: OZONE MIXING RATIO (PPB) against Elapsed UTs from 0hrs on 03-MAY-00 Go

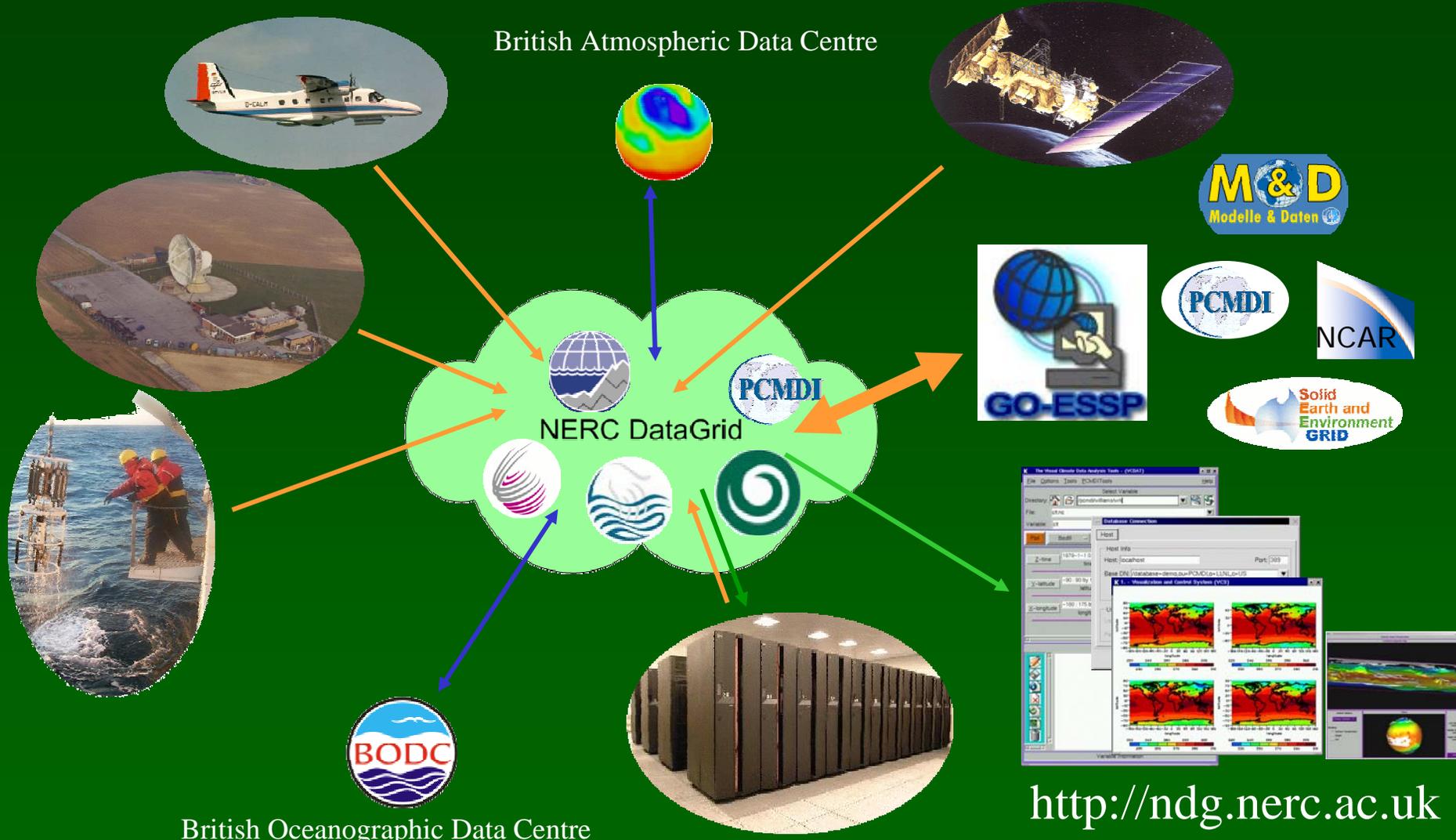
Omit points: Where: OZONE MIXING RATIO (PPB) =

Scale Ymin: Ymax: Xmin: Xmax: Reset

Plot Plus sign Asterisk Period Diamond X Connect points?

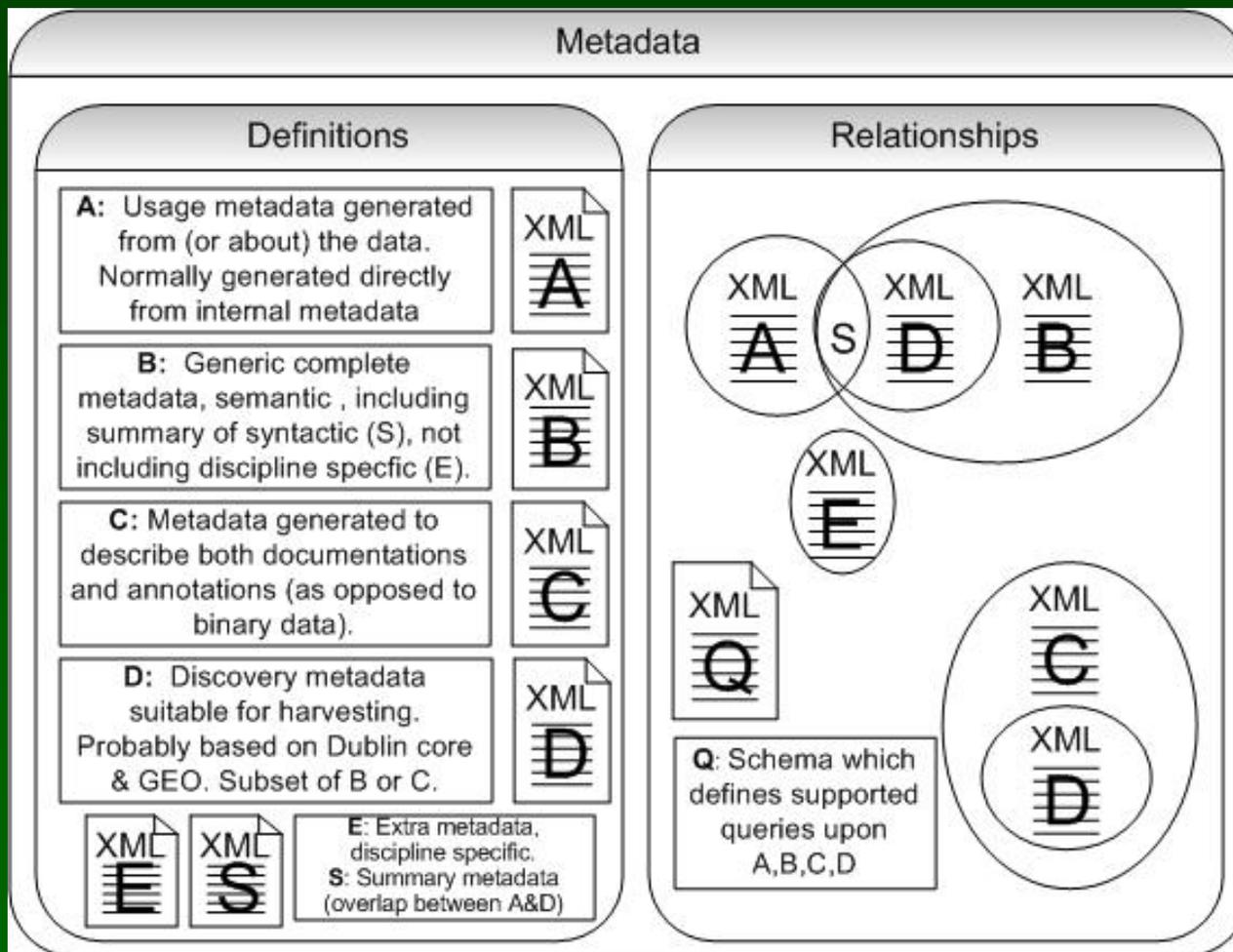


Complexity + Volume + Remote Access = Grid Challenge





NDG Metadata Taxonomy





Taxonomy Examples



A:

- NCML, CDML, ESML, **CSML**
(Climate Science **Modelling** Language)

B:

- THREDDS, **MOLES**
(Metadata Objects for Linking Environmental Sciences)

C:

- Citation, Trackback (blogging)
(See CLADDIER Project starting soon)

D:

- GEO, DIF, ISO19115, Dublin Core



NDG1 Key Aims



- Establish methodology for handling data from discovery to usage (done)
 - “compare and contrast an extensive range of data from within one context”
 - (methodology established)
- Establish size of metadata problems with vocabularies (done, large)
- Build tools to support key data types
 - (some: cdunif-pp, NAPPY, DataExtractor, CSML)
- Build Prototype
 - (done, but limited functionality)



NDG1 Key Components



1) CSML

- V1.0 Schema and Documentation released in January.
 - Including GML modifications to support met-ocean grids vertical coordinates etc.
 - Python based GML/CSML parser in preparation
 - Will form the heart of NDG2 as we will migrate to operational A-services (Currently ad hoc + CDML) to CSML.
- Considerable work on Standards Compliance
 - AUKEGGS project (of which more later)
 - OGC Membership
 - INSPIRE
 - WMS port for NetCDF underway, with migration to CSML planned.



NDG-A (CSML) in MarineXML (1)



The screenshot displays the SeeMyDENC software interface. On the left, a menu lists various actions like 'Load SENC Cells', 'Import S-57 Cells', and 'Add NDG XML Overlay'. The main window shows a map of the North Sea with depth contours and a data table. The table lists chart names, EDTN, and UPDN values.

Chart:	EDTN	UPDN
7CMXML01.000	1	0
7CMXML06.000	1	0
7CMXML07.000	1	0
7CMXML22.000	1	0
M0ADAILY_2004_091_5m	1	0
MERIS_20040401_095208	1	0

Chart:	EDTN	UPDN
7CMXML01.000	1	0
7CMXML06.000	1	0
7CMXML07.000	1	0
7CMXML22.000	1	0
depth_ndggml	1	0



CSML in MarineXML (2)



MarineXML Draft Final report:

“... there is a momentum from organisations such as IHO and WMO to adopt consistent approaches for the vocabulary of their data along the reference implementation of ISO Standards prescribed by the [Open Geospatial Consortium]...”

“The NDG format proved a robust recipient for the data from each community. It produced economical files with few redundant elements, striking about the right balance between weak and strong typing.”



NDG1 Key Components



2) MOLES

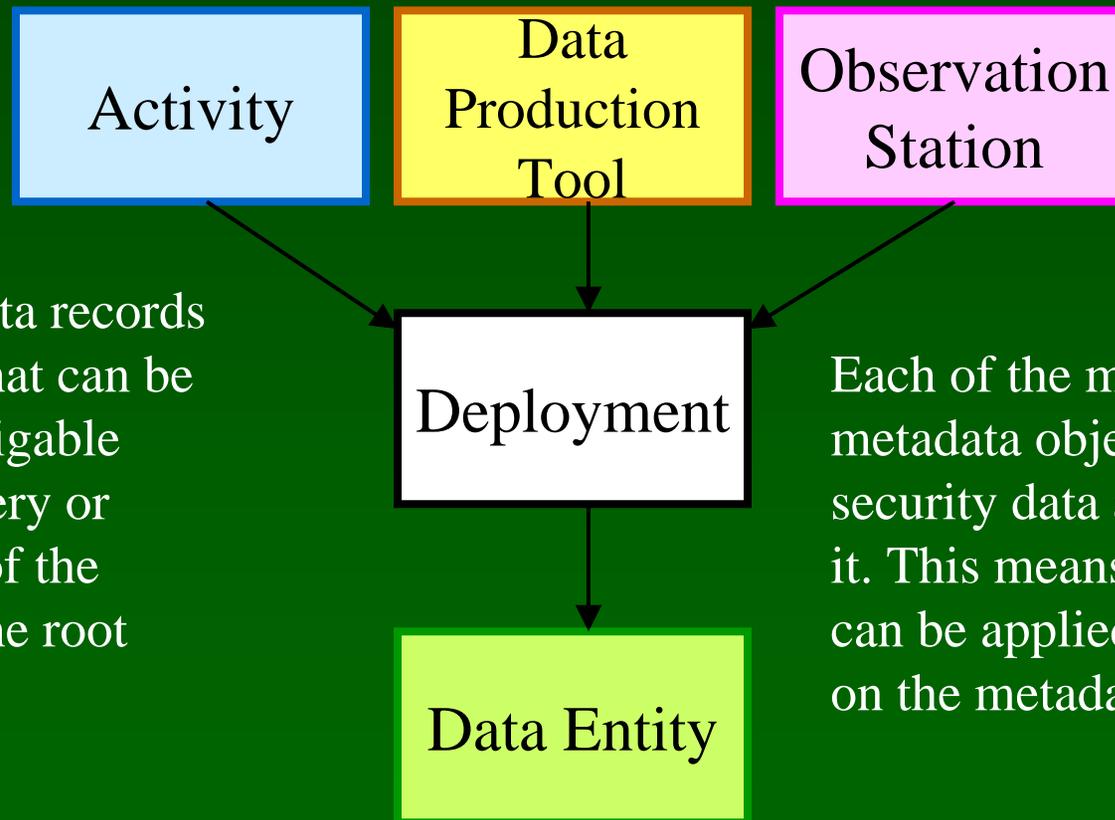
- V0.2 NDG released last week included data provider documentation for MOLES
- Code for
 - MOLES to DIF
 - MOLES to DC
 - MOLES based xqueries exists and is in testing ...
- Concepts for MOLES-Browsing GUI under discussion
- Relational Schema implementation due shortly
- Web Service interface to XMLDB and relational-DB to follow.



MOLES Implementation



Core linking concept is the **deployment** of a **Data Production Tool** at an **Observation Station** on behalf of an **Activity** that produces a **Data Entity**



Links the metadata records into a structure that can be turned into a navigable XML using Xquery or XSLT with any of the record types as the root element.

Each of the main metadata objects has security data attached to it. This means that this can be applied to queries on the metadata



NDG1: Key Components



3) Discovery

- Deployed exist XML database
- Deployed GUI interface
- Core datasets loaded from BADC, BODC
- Geographical search interface implemented
- Web Service Interface to text query implemented
 - Full geographical and secure interface underway
- OAI based experiments with NCAR and European partners underway



NDG Advanced Search Interface



**NDG - Discovery Service**

Advance Search

Text search

Full Text

Wildcards *, ? Character ranges [a-zA-Z] support

Fields

Temporal Coverage

Start Date

End Date

Date Input (yyyy-mm-dd)

Spatial Coverage

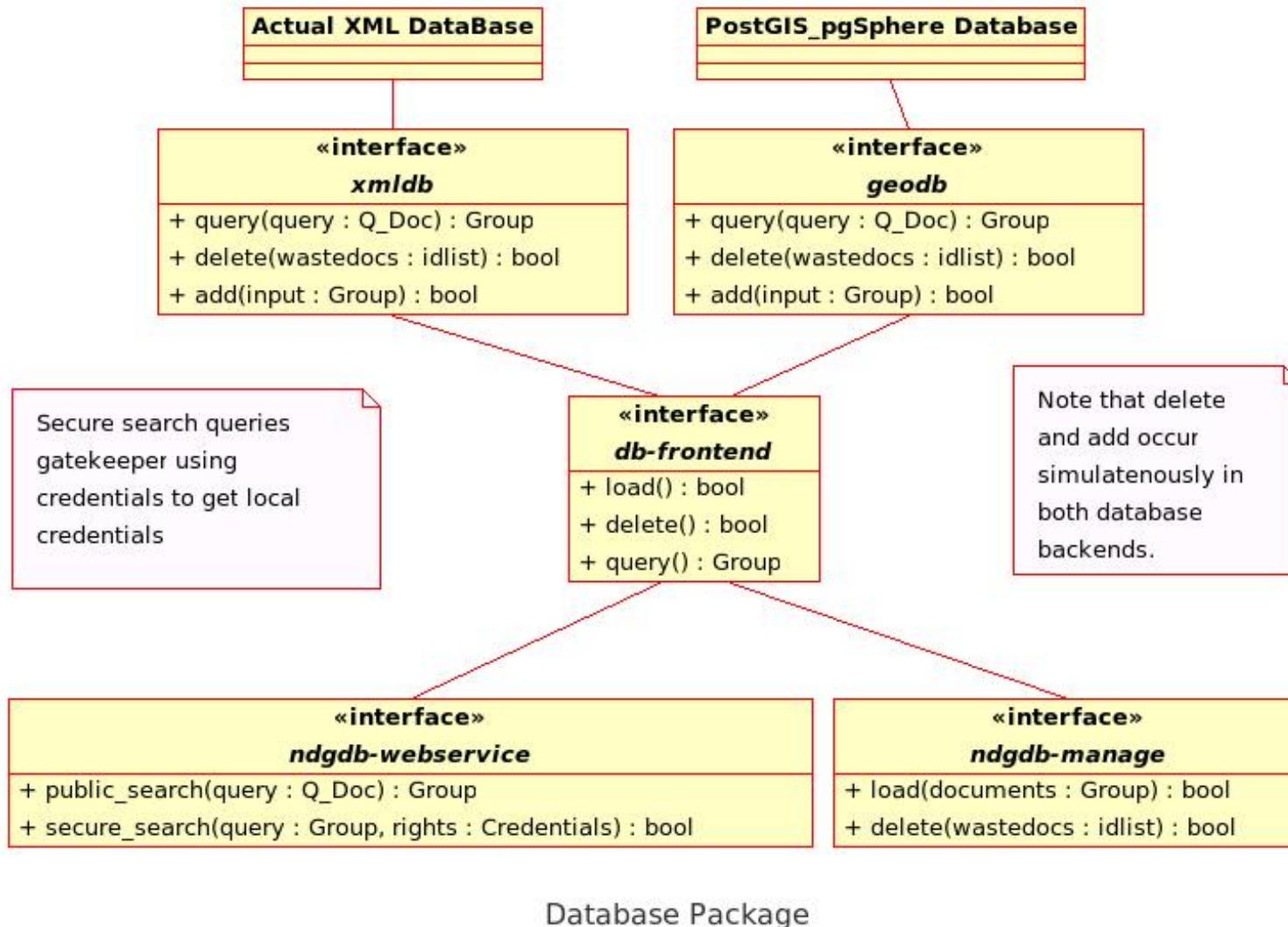
West East

North

South



Schematic: Secure access to metadata





NDG Discovery

outside NDG

Data Presentation

User Interface

Data Objects Bindings

Data Objects

WS Client API:
•Java
•Python
•Perl
•php

Data Resources

Lat – long / temp

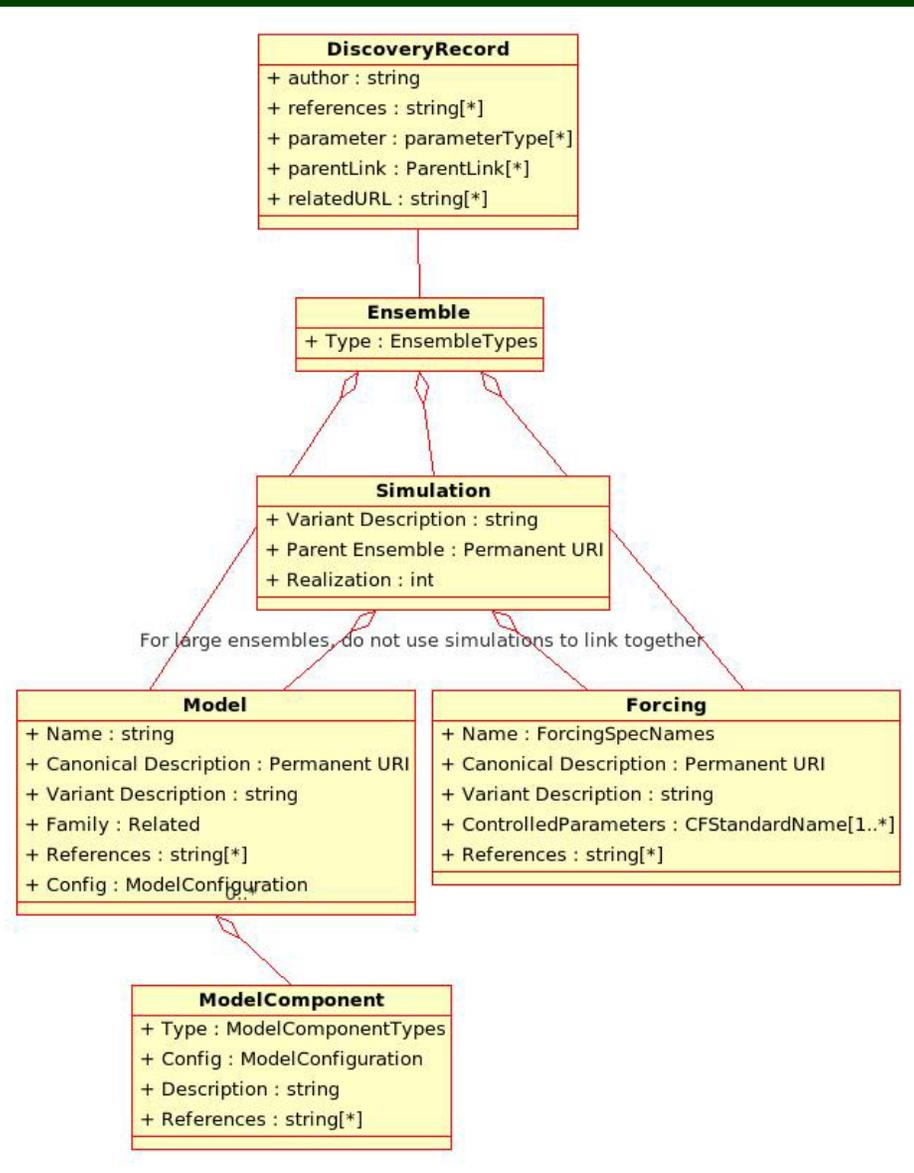
XML Discovery

SQL

SOAP: Xquery



Model Discovery





... more controlled vocab ...



«enum» ModelComponentTypes
Atmosphere
Ocean
Cryosphere
Land Surface
Chemistry
Aerosol
Trajectory

«enum» ForcingSpecNames
IPCC SRES A1
IPCC SRES A2
IPCC SRES B1
IPCC SRES B2
IS92A
preindustrial
present day
double preindustrial
double present day
background
historical
1% year increase
4% year increase
north atlantic

«enum» ModelConfiguration
GCM
StandAlone
Idealised
Energy Balance
Assimilation
Beta-Plane
Aqua-Planet
Column Mean
Slab

«enum» EnsembleTypes
Initial Condition
Perturbed Physics
Multiple Forcing
Arbitrary



NDG1: Key Components



4) Tools:

- Secure (Python/CDAT) Data Extractor deployed at BADC, utilising
 - PPI[O] (CDAT interface to met office PP data)
 - NAPPY (NASA Ames Python Processor)
 - CDML (for now)
- (Python/CDAT) GEOSPLAT deployed at BADC
- Web Service interfaces nearing completion



NDG1: Security

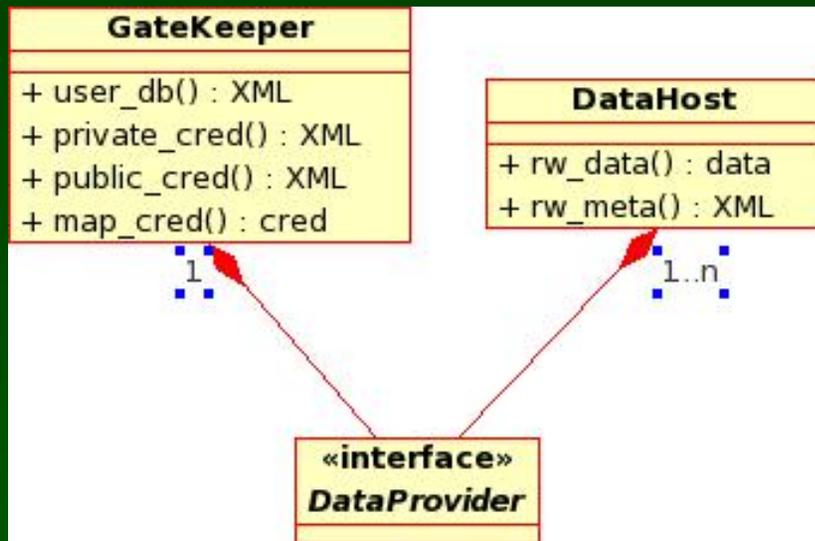


(Security=Authentication, Authorisation, Access)

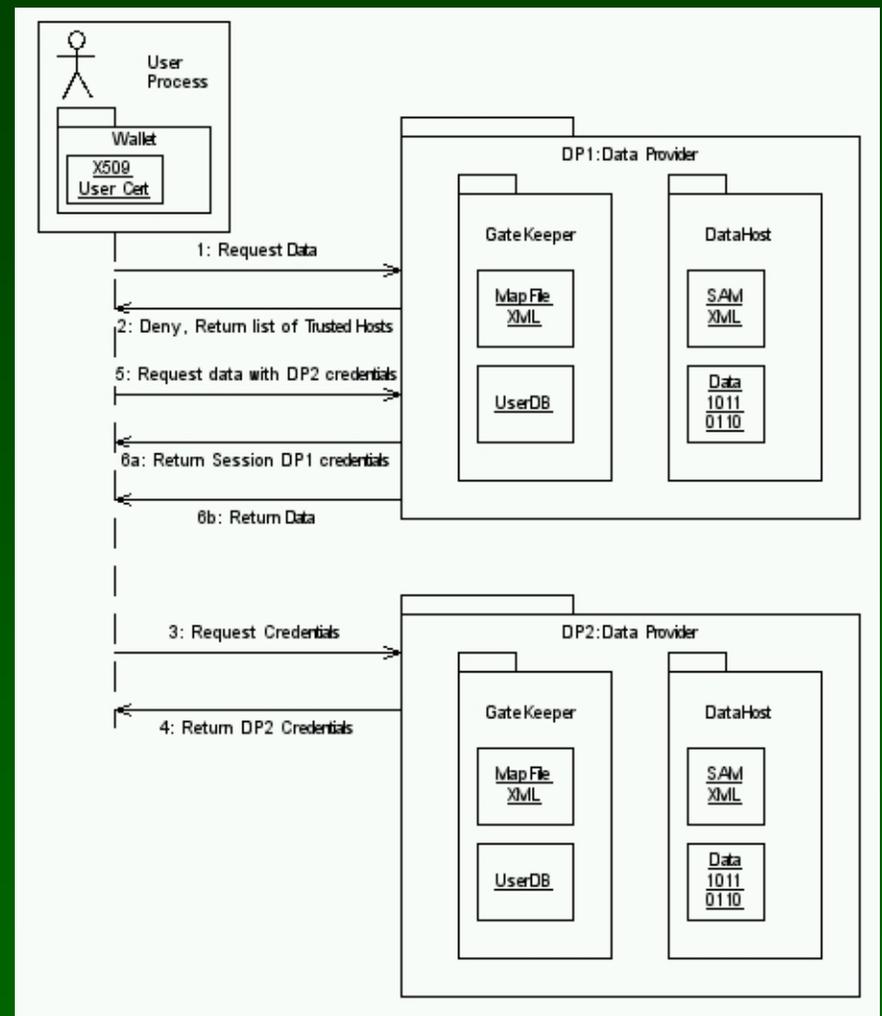
- Attribute Authority (web service) prototype complete
 - to produce authorisation tokens based on xml-signature
- Currently attempting to replicate PURSE functionality for deployment at BADC
 - to provide X509 certificates for existing users
- Metadata Structures included in MOLES and CSML to support security.



NDG Security



Certificate based, pass encrypted credentials between user and gatekeeper.





NDG2 Key Aims



- Development and Deployment of NDG1
 - Needs to be much easier to become a data provider, and to use the tools ... much work to be done ...
- Deployment of NDG in
 - HIGEM, BDAN, RAPID, EcoGRID, DEWS (DTI funded), MOTIIVE
 - Possibly: QUEST, e-Jules and other NERC initiatives under discussion.
- NDG Peering with Earth System Grid (US), SeeGrid (Aussie), Hamburg MPI (WDC-A Climate), GO-ESSP
- Become more involved in OGC/ISO community to get CSML subsumed in GML.
- Content, content, content ...
- Concluding with
 - Evaluating the requirements and sustainability of long term deployment of the grid infrastructure
 - Splitting the software and schema development and maintenance from grid infrastructure management.



Other Key Relationships



- Earth Simulator project depending on NDG information management ...
- NCAS Big Data Analysis Network depending on NDG information management
- Met Office, Delivering Environmental Web Services (DEWS, DTI Funded).
- Commercial Geographic Information Systems via Open Geospatial Consortium membership
- AUKEGGS:



AUKEGGS

Australia-UK collaboration on
Exploitation of Grid and Geospatial Standards



NDG2 Partner Roles



CCLRC e-science

- Schema Development, Schema Tools, Grid Infrastructure

BADC

- Look and Feel, Project Management, Atmospheric Vocabularies, Security, Metadata Population, Testing, Command Line and Browser Tools

BODC

- Oceanographic Vocabulary, Oceanographic Schema Issues, Metadata Population, Testing

PML

- Remote Sensing Implementation, Metadata Population, Testing

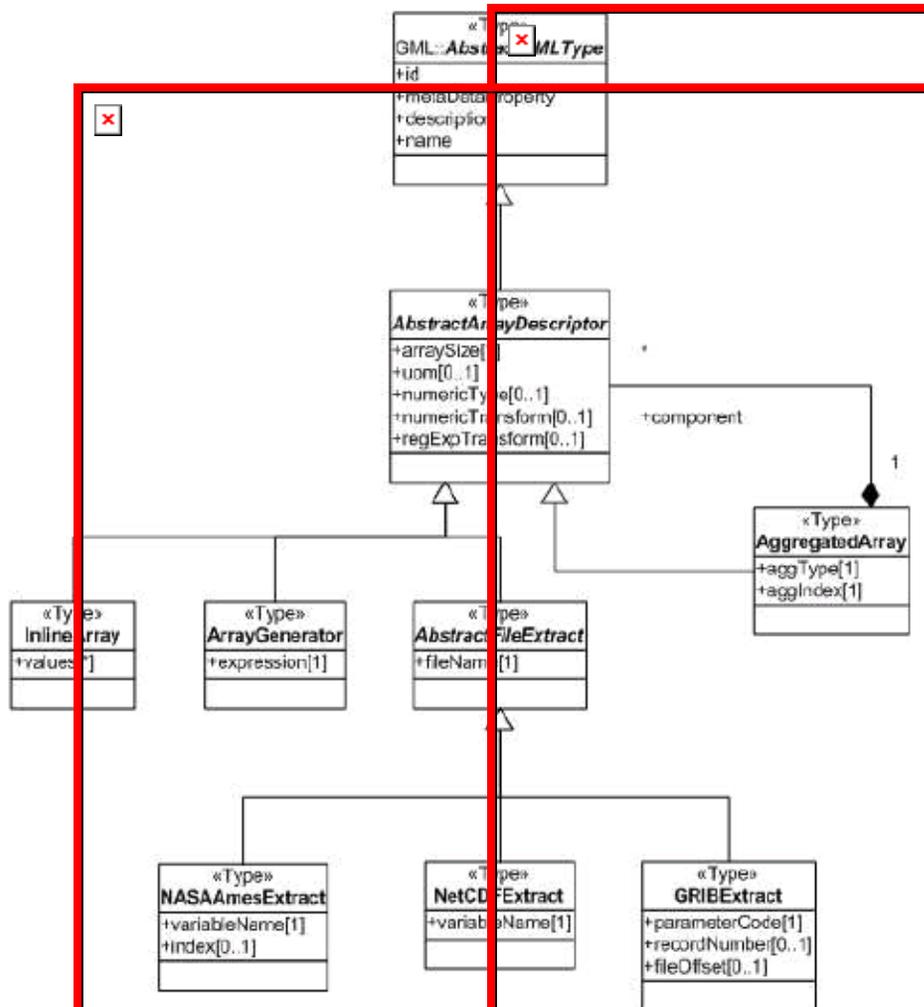
NOC(S)

- Link NDG to NERC funded Archive Developments, Complex Relationships between data (e.g. Seismic and Cruise) Metadata Population, Testing



Interoperability

- Vocabularies (CF + What? How?)
- Model Metadata
 - Detailed
 - Discovery
- OGC Metadata
- Agreed D for OAI
 - Schema (DIF -> ISO19115 variant)
 - Namespaces (need to support local extensions)
- Access Control where appropriate



Position,
Trajectory,
Orientated Position,
Orientated Trajectory

Figure 1: CSML numeric array definition