International Standards for National Data Management Excellence

Anna Luff, Data Officer
Who are we?

- UK Data Archive Centre
- Core funded by the UK and Scottish Government
- Hub for accessing, safeguarding, managing and disseminating data
- Support for monitoring and reporting frameworks
- Committed to applying the FAIR Data Principles

Keith Hiscock
The Role of DASSH

National role
- Accredited MEDIN Data Archive Centre since 2005
- Marine Node of the National Biodiversity Network
- Data provider to the Atlas of Living Scotland

EU role
- EMODnet Biology partner

International role
- UK Node of OBIS
- Accredited Data Unit of IODE
1) How does the work at DASSH contribute to the development of a global marine biodiversity observation network?

2) How do automated QA and Ingestion tools streamline data at a National and Global scale?
Data Dissemination

DASSH
The archive for marine species and habitats data

MEDIN
UK MEDIN
DASSH Portal
WMS/WFS
GeoServer
PostgreSQL
DASSH

Darwin Core
OBIS/GBIF/EMODnet
DwC-A
IPT

MEDIN
marine environmental data & information network
Data Standardisation

- Data are manually standardised to increase interoperability and re-use of marine data
- Use of MEDIN (Marine Environmental Data and Information Network) guidelines to standardise and input data
### General Metadata Form

<table>
<thead>
<tr>
<th>Field Title</th>
<th>Your Data and Information</th>
<th>Add additional columns as required</th>
</tr>
</thead>
<tbody>
<tr>
<td>surveyName</td>
<td>2020 Plymouth Sound Data Survey</td>
<td></td>
</tr>
<tr>
<td>surveyType</td>
<td>Benthic Data Survey</td>
<td></td>
</tr>
<tr>
<td>surveyObject</td>
<td>A grab survey site conducted by the Marine Biological Association of the United Kingdom (MBA) in Plymouth Sound to survey benthos.</td>
<td></td>
</tr>
<tr>
<td>surveyCode</td>
<td>MBS110000054</td>
<td></td>
</tr>
<tr>
<td>orgName</td>
<td>Marine Biological Association of the United Kingdom (MBA)</td>
<td></td>
</tr>
<tr>
<td>timeline</td>
<td>22/01/2004-14/08/04</td>
<td></td>
</tr>
<tr>
<td>uniqueId</td>
<td>MD08M-01955-4523</td>
<td></td>
</tr>
</tbody>
</table>

### Sample Event Form

<table>
<thead>
<tr>
<th>Field Title</th>
<th>sample/surveyID</th>
<th>replicate</th>
<th>taxonID</th>
<th>taxonName</th>
<th>alphaID</th>
<th>qualifier</th>
<th>abundance</th>
<th>abundanceUnits</th>
<th>determin</th>
<th>sampleSize</th>
<th>sampleLocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DAISHD010000055_SE_001</td>
<td></td>
<td></td>
<td>Corbula arenaria</td>
<td>00121</td>
<td>1</td>
<td>100</td>
<td>Presence/Absence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Species Form

<table>
<thead>
<tr>
<th>Field Title</th>
<th>sample/surveyID</th>
<th>replicate</th>
<th>taxonID</th>
<th>taxonName</th>
<th>alphaID</th>
<th>qualifier</th>
<th>abundance</th>
<th>abundanceUnits</th>
<th>determin</th>
<th>sampleSize</th>
<th>sampleLocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DAISHD010000055_SE_001</td>
<td></td>
<td></td>
<td>Corbula arenaria</td>
<td>00121</td>
<td>1</td>
<td>100</td>
<td>Presence/Absence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dashboard

Create

Enter the title
1992 Centre for Environment, Fisheries and Aquaculture Science (Cefas) North Sea

What type of resource do you want to describe
dataset

I am happy for my metadata to be passed to data.gov.uk

Update

Choose a record to update

Export

Search for a record

Diagram:

Dataset Received → Transformed → MEDIN Format Data

No transform required
Manual Quality Assurance check of data and metadata

✓ Abstract, title, start/end dates
✓ Do we have the necessary Detailed Metadata? Is it correct?
✓ Check of species data
✓ QA of Discovery Metadata on MEDIN - are all of the mandatory fields completed? Does the metadata match the raw data?
Validation Tool used to check data in MEDIN guidelines

Benefits?

✓ Speeds up data processing
✓ Increased accuracy
✓ Development of validation tool as a web-service
✓ Potential to be embedded in other applications

Validation

The archive for marine species and habitats data
Ingestion Tool

- Takes each MEDIN guideline sheet in turn, and ingests into the database.
- Extracts tables, reads and unpivots matrices
- Data quickly made available via GeoServer
1) How does the work at DASSH contribute to the development of a global marine biodiversity observation network?

- Data flow feeds into European and Global marine biological data infrastructure
- Contribute to data products on a global scale
- Influential in development of UK standards in accordance to INSPIRE specifications
2) How do automated QA and Ingestion tools streamline data at a National and Global scale?

- Data published faster- freeing up staff time
- Increased process transparency, and user confidence
- Tools can be used by other organisations
- Accept data in wide range of formats
Thank you for listening, do you have any questions?

Anna Luff, Data Officer  annluf@mba.ac.uk
DASSH enquiries  dassh.enquiries@mba.ac.uk