



www.bodc.ac.uk

Introduction to BODC and how to submit data

KATIE GOWERS, BRITISH OCEANOGRAPHIC DATA CENTRE
BANGOR UNIVERSITY OPEN ACCESS WEEK, OCTOBER 2015



**National
Oceanography Centre**
NATURAL ENVIRONMENT RESEARCH COUNCIL

noc.ac.uk

NERC SCIENCE OF THE
ENVIRONMENT



British Oceanographic Data Centre
Part of National Marine Facilities (NMF) at
the National Oceanography Centre (NOC)
which is a
Natural Environment Research Council (NERC)
Research Centre

The **British Oceanographic Data Centre (BODC)** is a national facility for looking after and distributing data concerning the marine environment.

What this means in practice

- Ingestion, long term storage and delivery of marine data
- Processing and delivery of near real-time data streams
- Data management for UK, European and International marine projects and operational science activities
- Maintaining and publishing Vocab
- On ship data management support for scientists
- Providing advice to other organisations
- Software development
- Data management training
- And many other things!

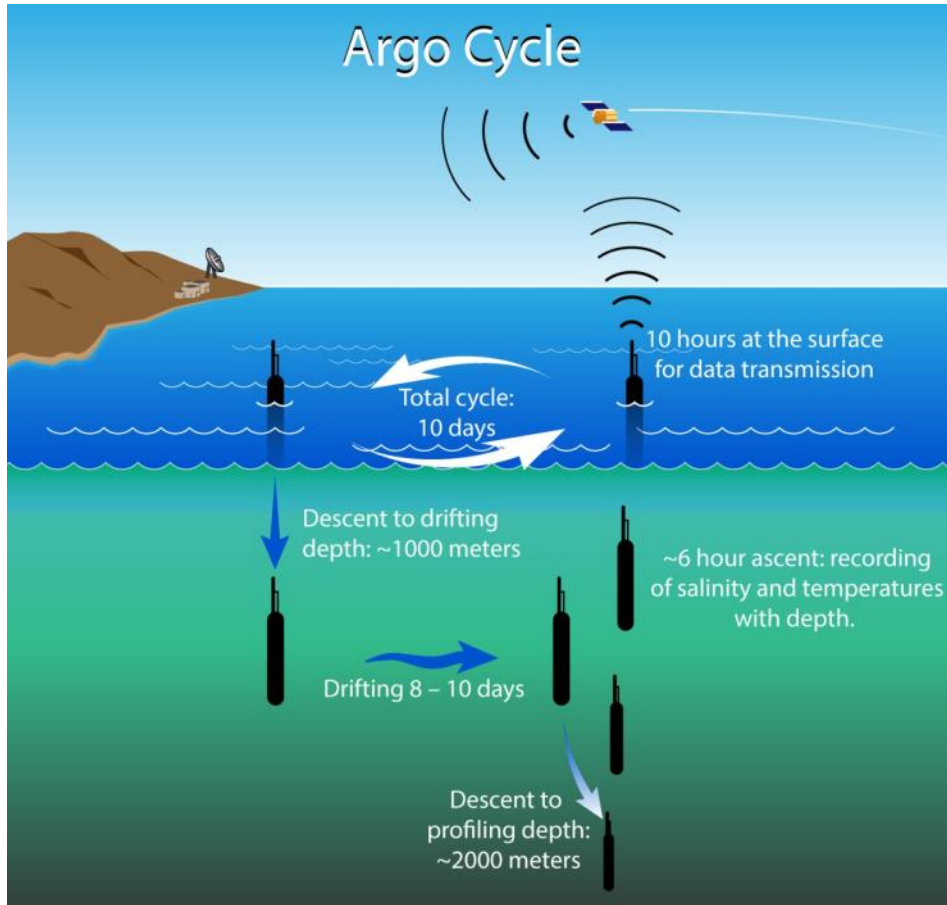
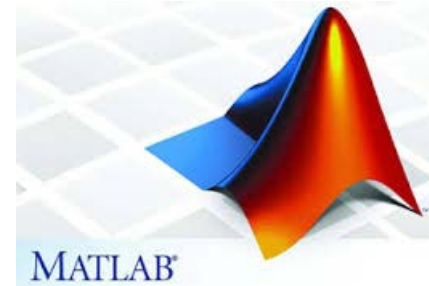
All activities need to be agreed in advance and funded





My day job

Oracle
SQL Developer



What happens to your data?



Data delivered
to BODC



We check the data
for completeness



We archive a copy
(in future proof format)



We record all processing
steps and decisions

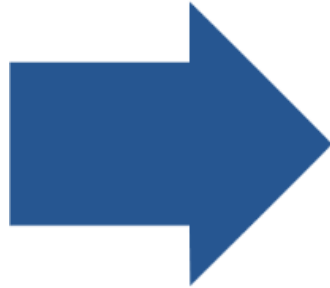


We research the
context



We document
details of the dataset

What happens to your data?



We convert the data to a standard format



We review the quality control of the data



We tag with metadata from controlled vocabs



We check our processing



We bank the final version



We expose the dataset online



Data management plans

- Planning data management should start before the project is even funded
- NERC grants require a data management plan to have been approved with a data centre
- This is good practise for all projects
- BODC are here to help! Please get in touch if you think you need a data management plan. We can:
 - Work out if you need a data management plan
 - Help you to identify the most appropriate data centre
 - Facilitate cross data centre planning (for multi-disciplinary projects)
 - Give advice about what's needed, costings etc.



Data formats



BODC accept data in any standard format for example text, CSV, Microsoft Excel, MATLAB .MAT, Seabird data files.

We are generally looking for:

- Fully processed
- Calibrated (if appropriate)
- “Science ready” data

We will retain a copy of the files supplied and create a future-proof copy.

BODC transfer the data to a standard NetCDF format for our processing and long-term storage.

What?

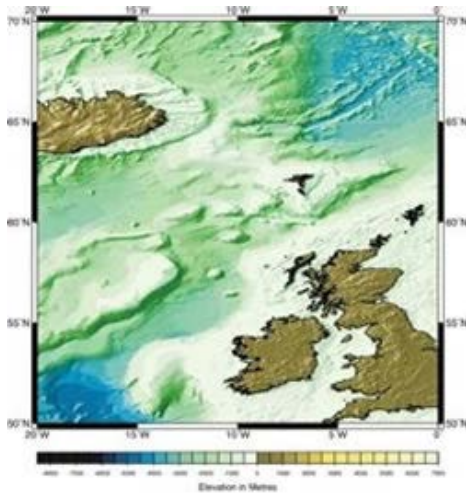
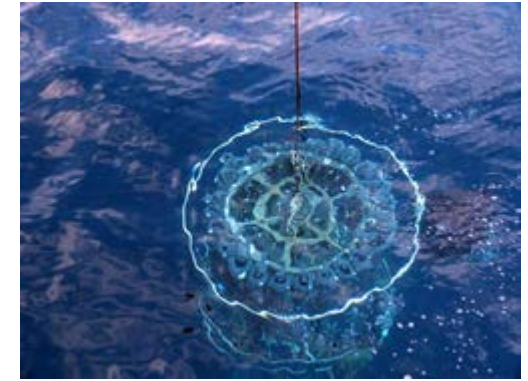
When?

Where?

How?

Why?

Who?



Without a context environmental data are just numbers ...

It starts on ship

EVENT	CTD	JDAY	GMT	EVENT	CTD	JDAY	GMT
092	CTD083	166	15:55				
093	CTD084	166	17:43				
094	CTD085	166	19:20				
001	CTD001	153	09:14	040	CTD0036	159	14:56
002	CTD002	153	21:29	041	CTD0037	159	16:52
003	CTD003	153	23:40	042	CTD0038	159	22:30
004	CTD004	154	02:42	043	CTD0039	160	00:59
005	CTD005	154	06:37	044	CTD0040	160	03:13
006	CTD006	154	10:11	045	CTD0041	160	06:06
007	GLIDER DEPLOYMENT	154	12:36	046	CTD0042	160	09:16
008	CTD 007	154	15:30	047	CTD0043	160	13:05
009	CTD 008	154	20:01	048	CTD0044	160	17:24
010	CTD 009	154	23:35	049	CTD0045	160	21:51
011	CTD 010	155	04:16	050	CTD0046	161	11:41
012	CTD 011	155	08:53	051	CTD0047	161	15:21
013	CTD 012 FAILED	155	13:41	052	CTD0048	161	15:14
014	CTD 013	155	22:58	053	CTD0049	161	16:46
015	CTD014	156	02:06	054	CTD0050	161	19:10
016	CTD015	156	07:06	055	CTD0051	161	22:43
017	Argo Float	156	09:37	056	CTD0052	162	01:57
018	CTD016	156	11:20	057	CTD0053	162	04:22
019	CTD 017 FAILED	156	15:35	058	CTD0054	163	08:18
020	CTD 018	157	02:02	059	CTD0055	163	23:15
021	CTD 019	157	08:29	060	CTD0056	164	01:04
022	CTD020	157	12:50	061	CTD0057	164	03:52
023	CTD021	157	18:23	062	CTD0058	164	06:04
024	CTD022	157	21:56	063	CTD0059	164	09:18
025	Argo float 2	158	00:00	064	CTD0060	164	21:25
026	CTD023	158	01:46	065	CTD0061	165	00:37
027	CTD024	158	04:52	066	CTD0062	165	02:49
028	CTD025	158	08:31	067	CTD0063	165	05:09
029	CTD026	158	10:56	068	CTD0064	165	06:19
030	CTD027	158	13:49	069	CTD0065	165	08:10
031	CTD028	158	16:32	070	CTD0066	165	09:12
032	CTD029	158	19:36	071	CTD0067	165	10:22
033	CTD030	158	22:18	072	CTD0068	165	11:55
034	CTD031	159	01:34	073	CTD0069	165	13:28
035	CTD032	159	04:18	074	CTD0070	165	15:15
036	CTD033	159	07:05	075	CTD0071	165	16:32
037	CTD034	159	09:35	076	CTD0072	165	17:49
038	GLIDER DEPLOYMENT	159	12:30?	077	CTD0073	165	19:22
039	CTD035	159	14:33	078	CTD0074	163	20:32
				079	CTD0075	166	22:37
				080	CTD0076	166	00:29
				081	CTD0077	166	02:41
				082	CTD0078	166	04:52
				083	CTD0079	166	07:09
				084	CTD0080	166	09:13
				085	CTD0081	166	11:30
				086	CTD0082	166	13:32

Event logs

It starts on ship

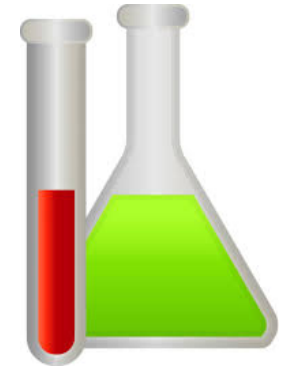
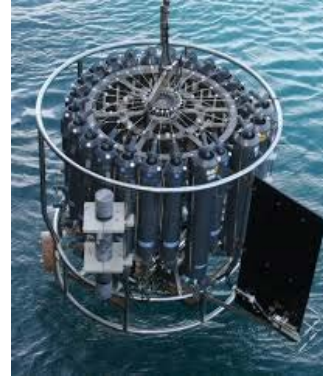
Labelling
samples



Metadata - Methodology

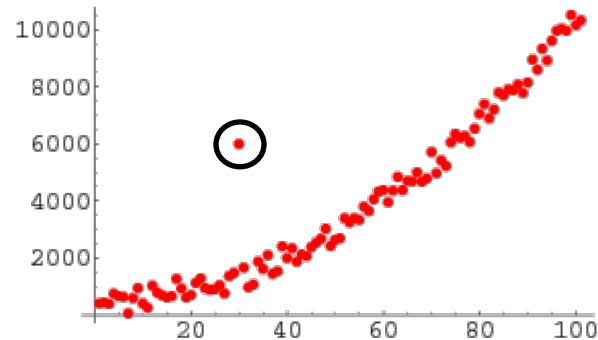


How was the data collected?



What quality control/
Calibration has been done?

What instruments were
used?



Have outliers been excluded
and why?

How were the data
analysed?



Is the data of
good quality?

Metadata - Instruments

Scientific instruments are identified by their make, model and sub-model (version) like cars ...



Ford



Focus



Zetec



Titanium



ST-1

We want to know all three bits of information + the serial number!

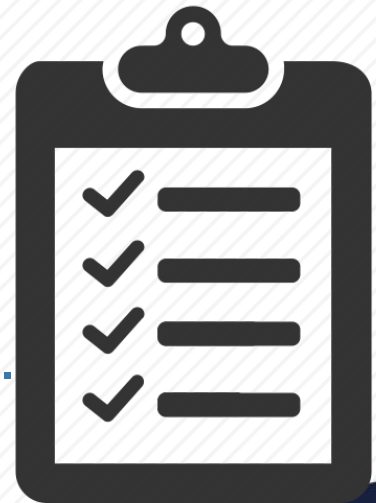
Submission templates

Our website (www.bodc.ac.uk) has comprehensive information about how to submit data to us and what we need.

We also have submission templates to guide you through what information we are looking for.

We can not ingest or bank processed data without full metadata records.

Get in touch with us early in the process to understand what we need (and how we can help).



Getting the most from BODC's resources

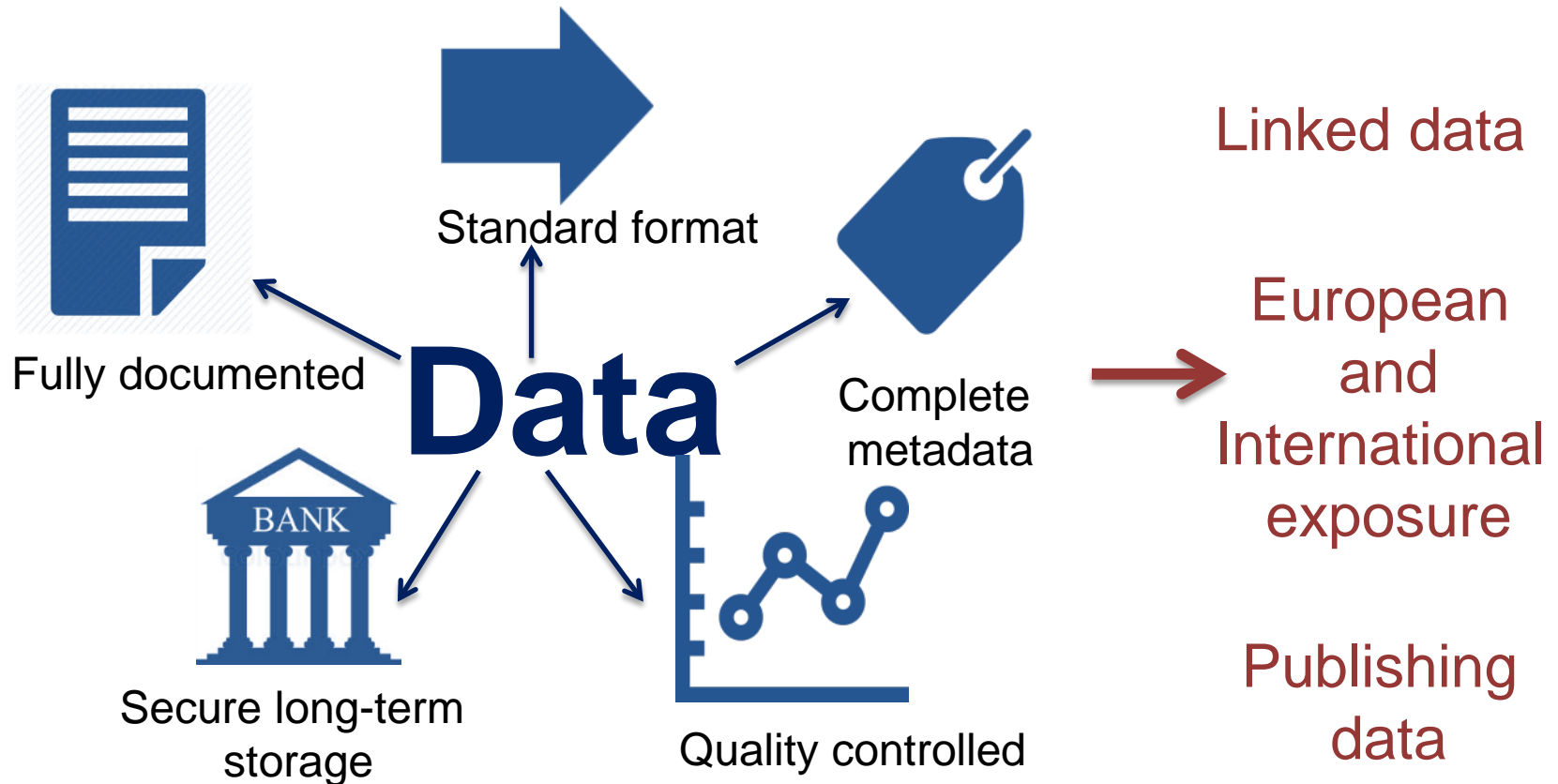
BODC is funded to manage data for specific projects. This funding has limits and usually has a deadline!

You can get the most out of the resource we have by:

- Providing final version of data as early as possible
- Providing all metadata required
- Providing a complete and detailed cruise report, including event logs
- Responding to emails in a timely fashion where possible

Please don't leave data submission to the last possible moment!

The end result



Any questions?

**“In God we trust.
All others must bring data.”**

W. Edwards Deming
Statistician



National
Oceanography Centre
NATURAL ENVIRONMENT RESEARCH COUNCIL

noc.ac.uk

NERC SCIENCE OF THE
ENVIRONMENT