



A Practical Workflow for an Open Scientific Lifecycle Project: EcoNAOS

Annalisa Minelli, Alessandro Sarretta, Alessandro Oggioni, Caterina Bergami, Alessandra Pugnetti

IRCDL 2019 - 15th Italian Conference on Digital Libraries
Pisa - January 31st, February 1st, 2019

Summary

- Open Science applied to LTER data
- EcoNAOS project: studied zone and the database
- EcoNAOS as a workflow:
 - The Open Scientific Lifecycle model
 - Work done: data harmonisation, ancillary data collection, data interoperability
 - Ongoing work: sharing and dissemination
 - ToDos: dynamic dataset citation, guidelines preparation
- Conclusions and future perspectives

EcoNAOS: Open Science applied to LTER data



Sites' network:
study of ecosystems on multidecadal scale, support to environmental management



EcoNAOS
Ecological
North Adriatic
Open Science
Observatory System



Italian flagship project:
maritime technologies, fishing, marine spatial planning, observatories, interoperable research data infrastructure

EcoNAOS: Open Science applied to LTER data



IT-12: NAS

Created on Inkatlas. © OpenStreetMap contributors (openstreetmap.org). Map data Oct 27, 2017. 1:2000000

<https://deims.org/92fd6fad-99cd-4972-93bd-c491f0be1301>

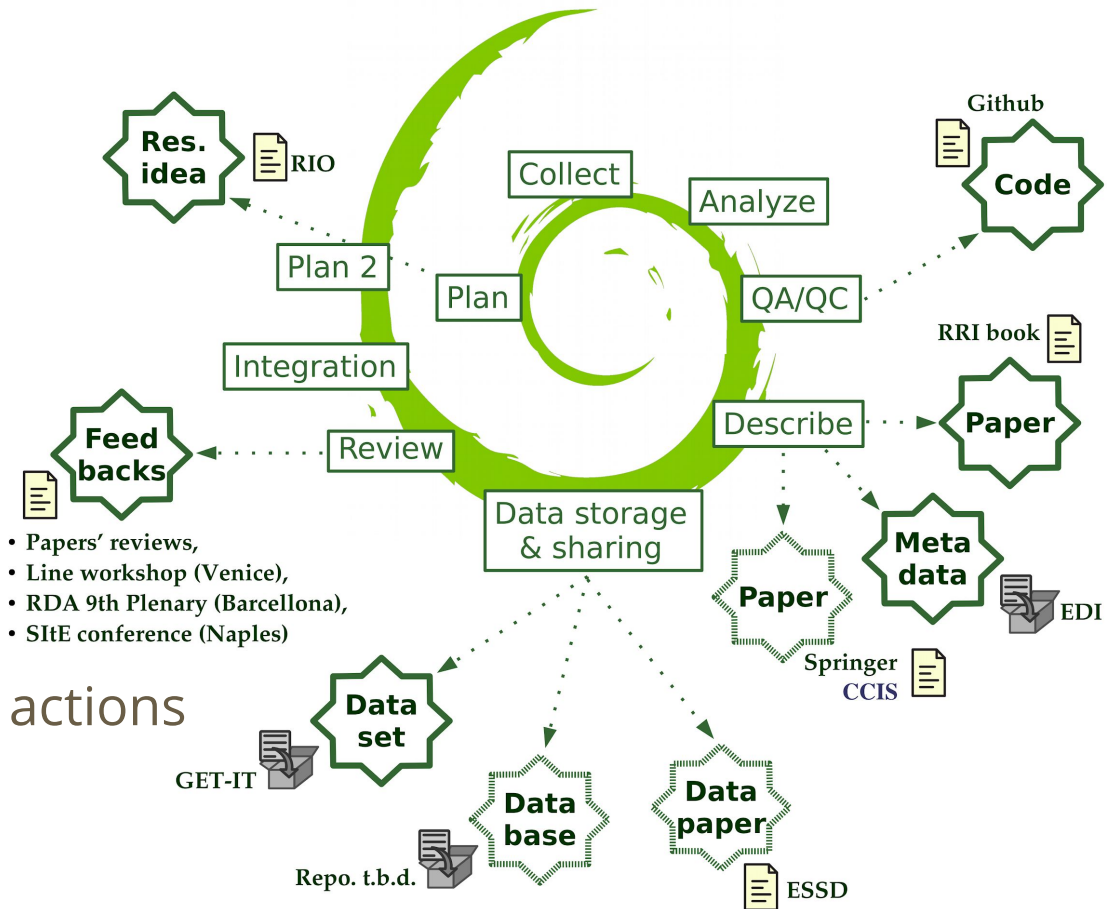
The database:

- 22 parameters for 50 years of recordings
- >110000 observations
- abiotic, phyto and zooplankton data
- data from fixed sensors and cruises

Long	Lat	Depth	Station	Station updated_name	Cruise	Ship	dd/mm/yy	hh:mm	Bot. Depth [m]	Meteo	Secchi (m)	Temp_sensor	Temp
12.68	45.33	0.50	B C04		PP/1	VERCELLI	12/04/1965	09:33	23.00	Y	3.70	Termometri a rovesciamento	13.12
12.68	45.33	5.00	B C04		PP/1	VERCELLI	12/04/1965	09:33	23.00	Y	3.70	Termometri a rovesciamento	12.35
12.68	45.33	10.00	B C04		PP/1	VERCELLI	12/04/1965	09:33	23.00	Y	3.70	Termometri a rovesciamento	12.45
12.68	45.33	20.00	B C04		PP/1	VERCELLI	12/04/1965	09:33	23.00	Y	3.70	Termometri a rovesciamento	12.14
12.86	45.28	0.50	C		PP/1	VERCELLI	12/04/1965	12:20	29.00	Y	6.00	Termometri a rovesciamento	12.25
12.86	45.28	5.00	C		PP/1	VERCELLI	12/04/1965	12:20	29.00	Y	6.00	Termometri a rovesciamento	12.24
12.86	45.28	10.00	C		PP/1	VERCELLI	12/04/1965	12:20	29.00	Y	6.00	Termometri a rovesciamento	11.16
12.86	45.28	20.00	C		PP/1	VERCELLI	12/04/1965	12:20	29.00	Y	6.00	Termometri a rovesciamento	12.30
12.48	45.40	0.50	A		PP/2	VERCELLI	28/04/1965	06:42	16.40	Y	3.50	Termometri a rovesciamento	12.27
12.48	45.40	1.00	A		PP/2	VERCELLI	28/04/1965	06:42	16.40	Y	3.50	Termometri a rovesciamento	12.37
12.48	45.40	5.00	A		PP/2	VERCELLI	28/04/1965	06:42	16.40	Y	3.50	Termometri a rovesciamento	12.44
12.48	45.40	10.00	A		PP/2	VERCELLI	28/04/1965	06:42	16.40	Y	3.50	Termometri a rovesciamento	12.23
12.68	45.33	0.50	B C04		PP/2	VERCELLI	28/04/1965	09:10	22.30	Y	3.50	Termometri a rovesciamento	12.49
12.68	45.33	5.00	B C04		PP/2	VERCELLI	28/04/1965	09:10	22.30	Y	3.50	Termometri a rovesciamento	12.43
12.68	45.33	10.00	B C04		PP/2	VERCELLI	28/04/1965	09:10	22.30	Y	3.50	Termometri a rovesciamento	11.92
12.68	45.33	20.00	B C04		PP/2	VERCELLI	28/04/1965	09:10	22.30	Y	3.50	Termometri a rovesciamento	10.50
12.86	45.28	0.50	C		PP/2	VERCELLI	28/04/1965	11:20	31.00	Y	3.80	Termometri a rovesciamento	12.40
12.86	45.28	5.00	C		PP/2	VERCELLI	28/04/1965	11:20	31.00	Y	3.80	Termometri a rovesciamento	12.09
12.86	45.28	8.00	C		PP/2	VERCELLI	28/04/1965	11:20	31.00	Y	3.80	Termometri a rovesciamento	11.50
12.86	45.28	20.00	C		PP/2	VERCELLI	28/04/1965	11:20	31.00	Y	3.80	Termometri a rovesciamento	10.42
12.48	45.40	0.50	A		PP/3	VERCELLI	13/05/1965	06:47	16.00	Y	2.30	Termometri a rovesciamento	15.92
12.48	45.40	1.00	A		PP/3	VERCELLI	13/05/1965	06:47	16.00	Y	2.30	Termometri a rovesciamento	15.80
12.48	45.40	5.00	A		PP/3	VERCELLI	13/05/1965	06:47	16.00	Y	2.30	Termometri a rovesciamento	14.92
12.48	45.40	10.00	A		PP/3	VERCELLI	13/05/1965	06:47	16.00	Y	2.30	Termometri a rovesciamento	11.34
12.68	45.33	0.50	B C04		PP/3	VERCELLI	13/05/1965	09:23	21.00	Y	6.00	Termometri a rovesciamento	17.40
12.68	45.33	5.00	B C04		PP/3	VERCELLI	13/05/1965	09:23	21.00	Y	6.00	Termometri a rovesciamento	15.66
12.68	45.33	10.00	B C04		PP/3	VERCELLI	13/05/1965	09:23	21.00	Y	6.00	Termometri a rovesciamento	13.64
12.68	45.33	20.00	B C04		PP/3	VERCELLI	13/05/1965	09:23	21.00	Y	6.00	Termometri a rovesciamento	11.83
12.86	45.28	0.50	C		PP/3	VERCELLI	13/05/1965	12:15	31.00	Y	4.00	Termometri a rovesciamento	18.03
12.86	45.28	2.00	C		PP/3	VERCELLI	13/05/1965	12:15	31.00	Y	4.00	Termometri a rovesciamento	16.51
12.86	45.28	8.00	C		PP/3	VERCELLI	13/05/1965	12:15	31.00	Y	4.00	Termometri a rovesciamento	13.31

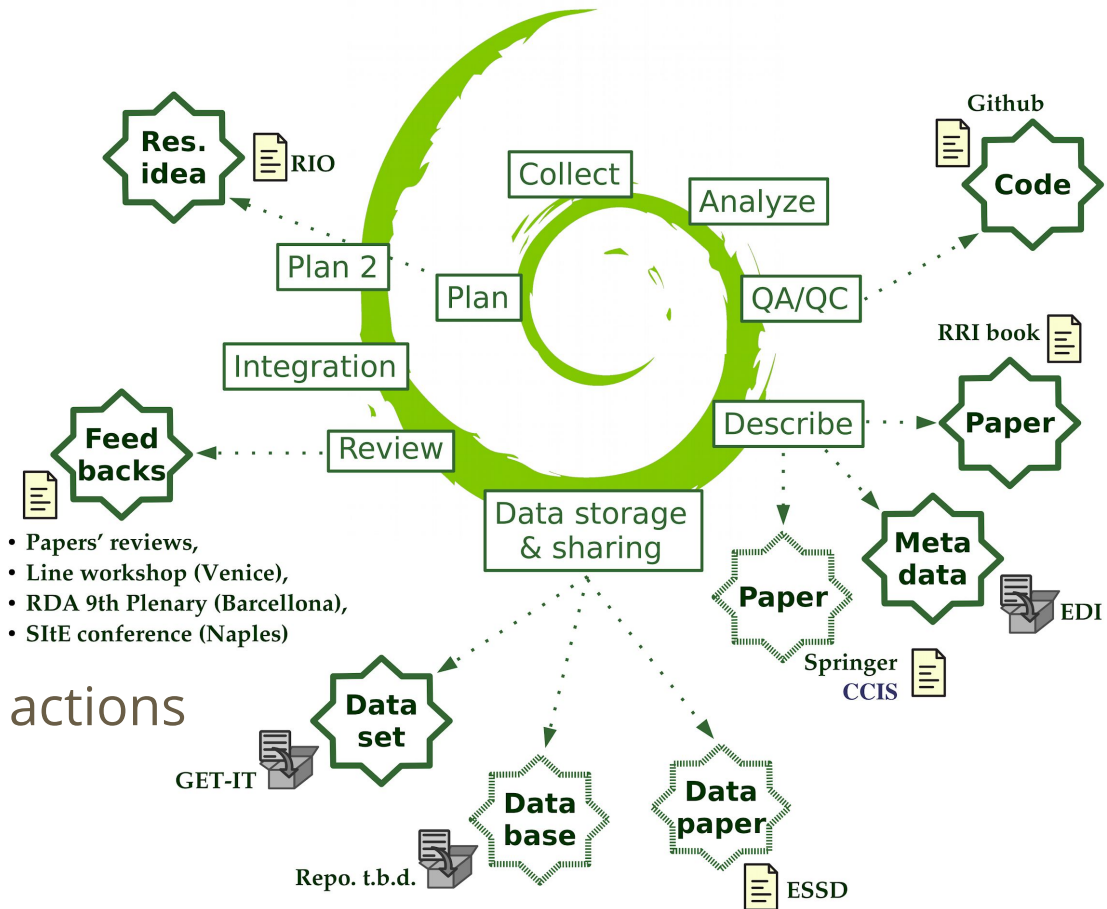
The workflow: Open Scientific Lifecycle Model

1. data harmonisation
2. ancillary data collection
3. data interoperability
4. sharing and dissemination actions
5. dynamic dataset citation
6. guidelines preparation



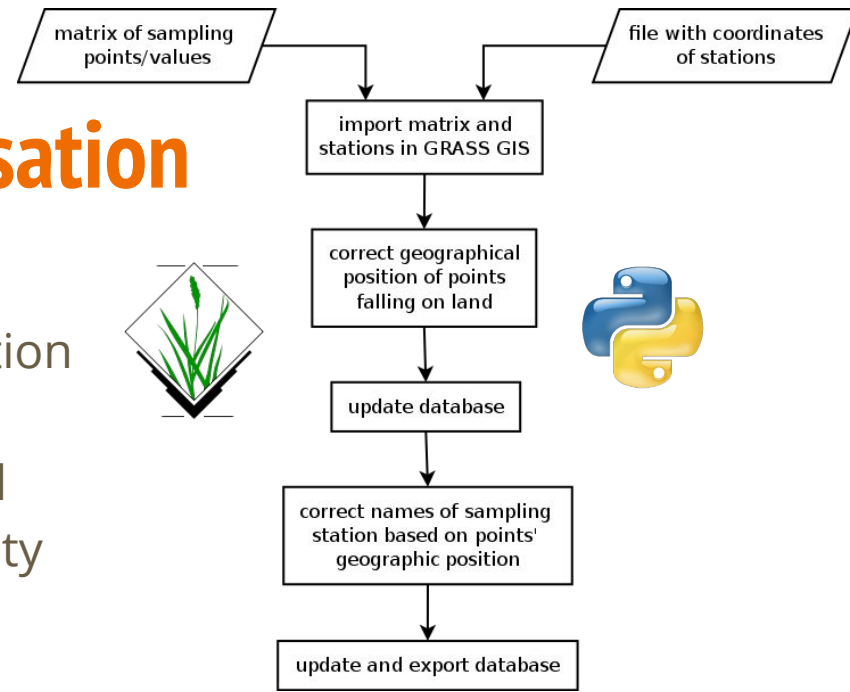
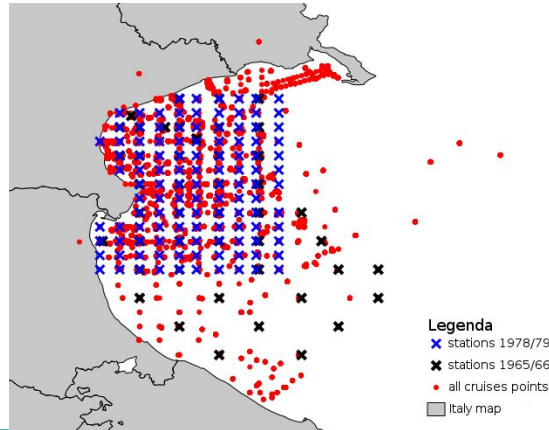
The workflow: Open Scientific Lifecycle Model

1. data harmonisation
2. ancillary data collection
3. data interoperability
4. sharing and dissemination actions
5. dynamic dataset citation
6. guidelines preparation



The workflow: data harmonisation

- points falling on land → geographical position correction
- recurrent stations → individuation of usual sampling points and study of their variability in time
- heterogeneous station naming → semantic homogenisation



long	lat	name	periods
12:29E	45:24N	A	summer_1965,autumn_1965
12:52E	45:16N	C	summer_1965,autumn_1965
12:46E	45:30N	1/1	summer_1965,autumn_1965
13:14E	45:20N	3/2	summer_1965,autumn_1965
12:46E	45:20N	Y/2	autumn_1965
12:46E	45:25N	04/2A	1978,1979
13:21E	44:45N	09/2E	1978,1979
13:13E	44:30N	08/0G	1978,1979

The workflow: ancillary data collection

About methods variation in time for the same parameter

redaction of three technical reports about early cruises in the NAS (years 1965-1980)

integration of the rest of database (years 1981-2015) based on expert judgment

improvement of data:

Readability

Reliability

Consistency

The workflow: data interoperability



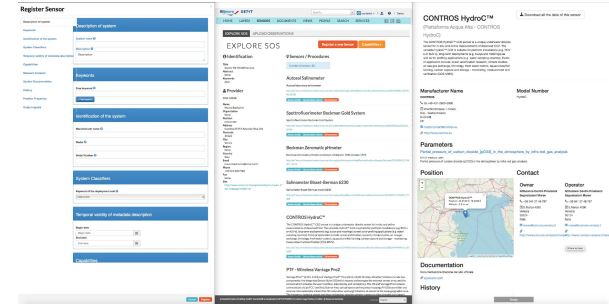
creation of a **SOS** (OGC) for data and metadata related to sensors and observations through an SDI



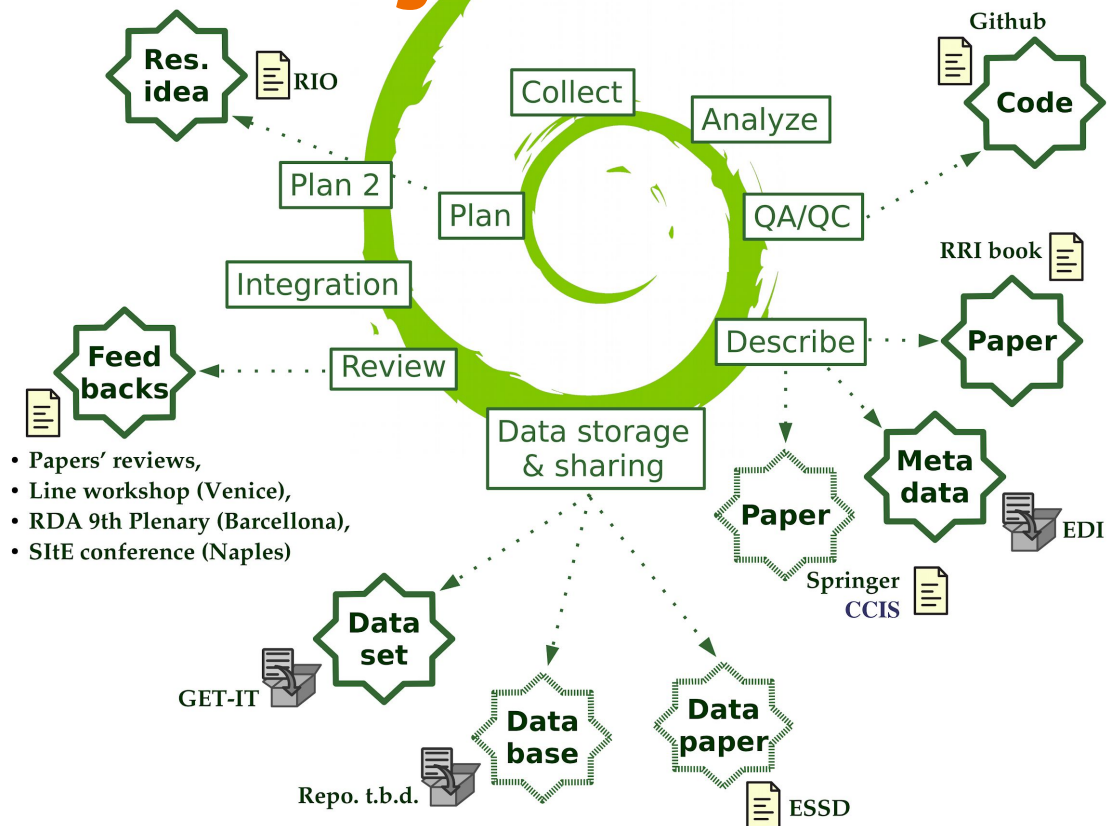
open source platform
GET-IT for visualization
and sharing of more
than 15000 observation

**Interoperability
&
FAIR data
management**

insertion of new sensors
by **EDI** interface (INSPIRE,
RNDT and SensorML
compliant)



The workflow: sharing and dissemination actions



The workflow: dynamic dataset citation

- how can we cite dynamic data?
- how to cite aggregated data and portion of data?
- persistency of data citation
- how to update references?
- which is the right threshold of data variation for reference variation?



1. identify and cite databases or portions of it
2. be interoperable
3. cite and retrieve data also from past versions of the database



**Follow some
pilot projects**

The workflow: dynamic dataset citation

- how can we cite dynamic data?
- how to cite aggregated data and portion of data?
- persistency of data citation
- how to update references?
- which is the right threshold of data variation for reference variation?

rules shared with data journals:



- data persistency
- liberal license
- DOI assignment
- **living data mode**



**Update of the
datapaper**

The workflow: guidelines preparation

Aim:

1. facilitate
 2. guide
- 
- the application of Open Science principles to research projects

The workflow: guidelines preparation

Aim:

1. facilitate
 2. guide
- 
- the application of Open Science principles to research projects

2 types of actions:

**choice of open peer
review journals**

**release of
source code**

operating on the whole process

**dissemination
actions**

**drafting of
data policy**

operating on data

The workflow: guidelines preparation

Aim:

1. facilitate
 2. guide
- 
- the application of Open Science principles to research projects

2 types of actions:

operating on the whole process

**semantic
harmonisation**

**create interoperable
instruments**

operating on data

**choice of
repository**

**data
maintenance**

Conclusions and future perspectives

- EcoNAOS is specifically conceived around marine science and LTER data but the application of these practices can be extended to projects in **any research topic** and involving **any kind of data**
- EcoNAOS is an occasion to deepen into **physical and psychological barriers** to Open Science principles application
- EcoNAOS is **not an exhaustive model** but it is open to necessities and specific requirements

In the near future..

- deepen into data citation task (wide argument, few experiences in oceanography with dynamic databases)

Thank you for your attention

Some links:

Open Science def.: <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>

IT-12 NAS site: <https://deims.org/92fd6fad-99cd-4972-93bd-c491f0be1301>

LTER Italia: <http://www.lteritalia.it/>

RITMARE project: <http://www.ritmare.it/>

Harmonisation code: <https://github.com/CNR-ISMAR/econaos>

SOS OGC standard: <http://www.opengeospatial.org/standards/sos>

GET-IT: <http://www.get-it.it/>

EDI references: <http://edidemo.get-it.it>, DOI: [10.5334/jors.106](https://doi.org/10.5334/jors.106)

Annalisa.Minelli@ve.ismar.cnr.it

