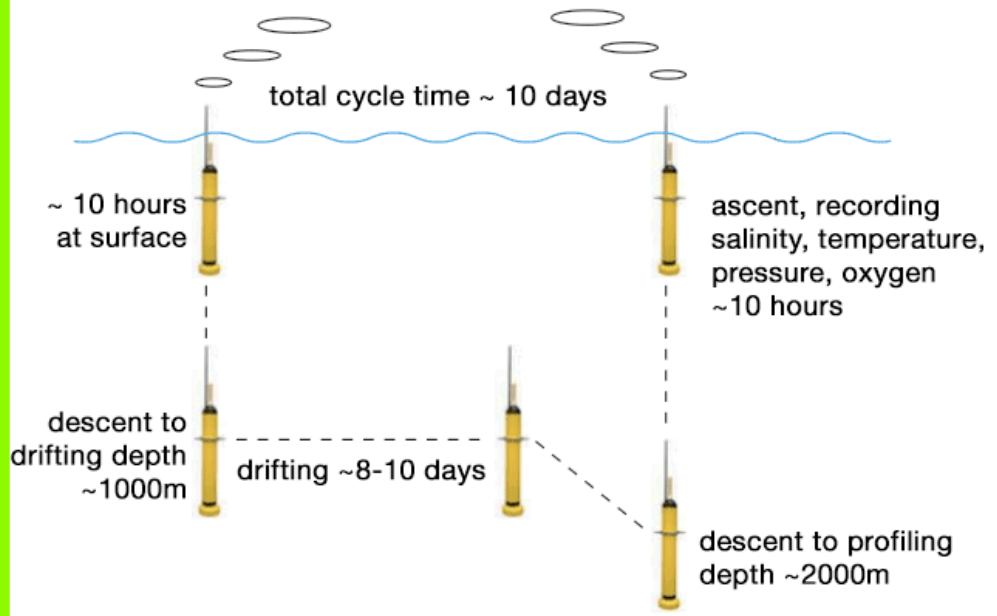
INFORMATION & DATA MINING TOOL FOR NORTH ATLANTIC ARGO DATA

Guillaume Maze, Ifremer, LPO

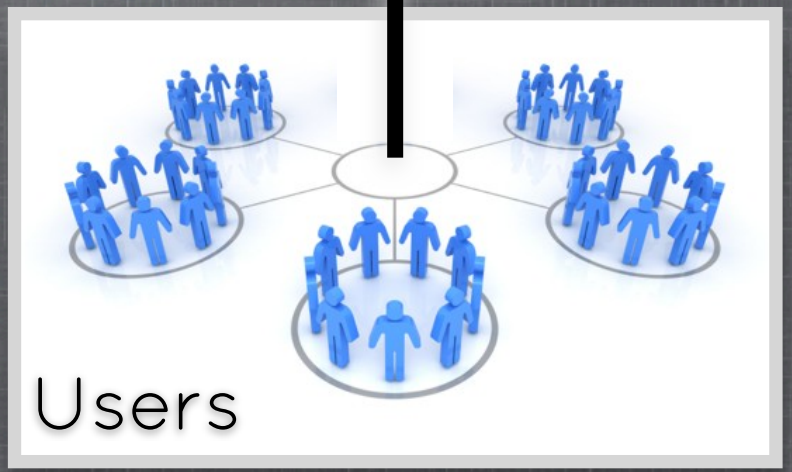
Jan. 20, 2012

THE ARGO ARRAY

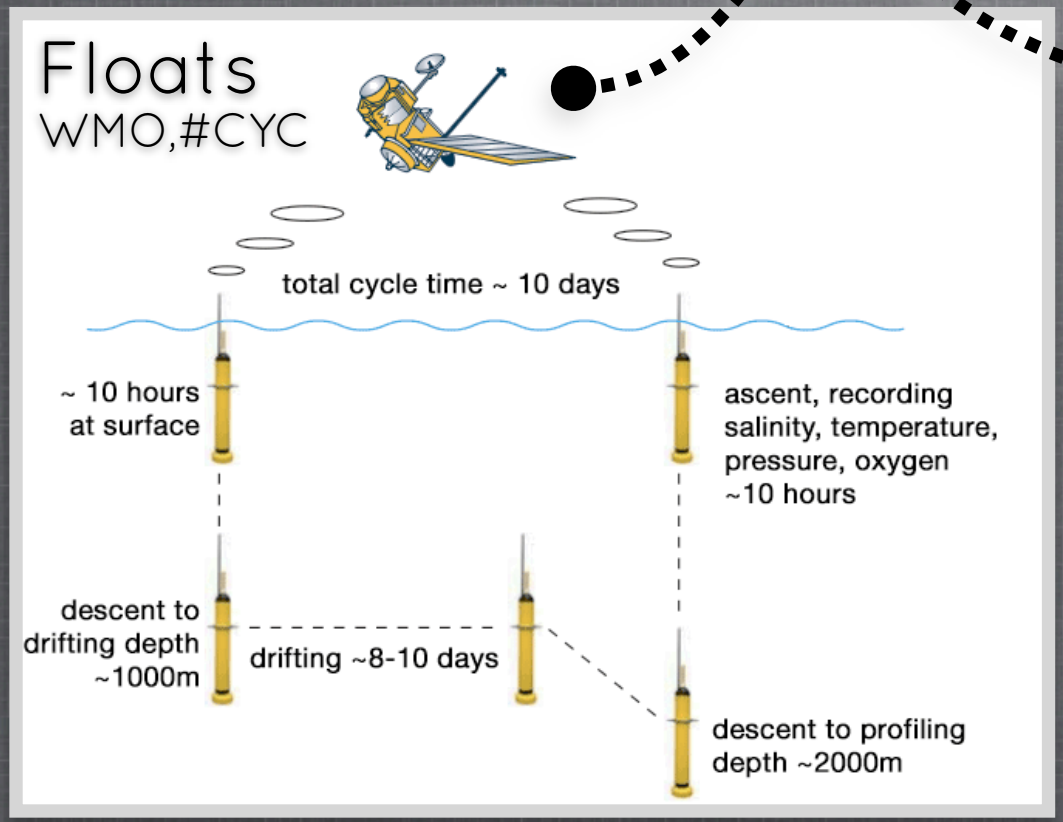
Floats
WMO,#CYC



Data Assembly Center (DAC)

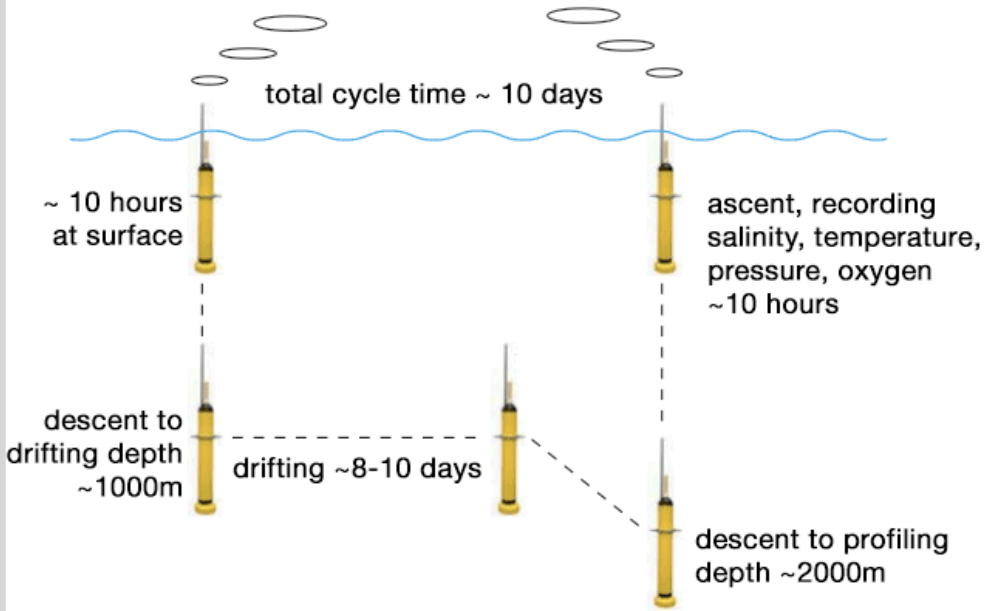


THE ARGO ARRAY



THE ARGO ARRAY

Floats
WMO,#CYC



Data Assembly Center (DAC)

THE ARGO ARRAY

PIs

Data Assembly Center (DAC)

Regional Centers

North-Atlantic Argo Regional Center: NA-ARC, OSU-IUEM

3 Data "Mode"

- Real Time (R)
- Delayed (D)
- Adjusted Real Time (A)

Users

HOW TO ENGAGE WITH THE DATA ?

Parameters ?

Data mode ?

Geography ?

DAC ?

Time ?

Float issues ?

Quality ?

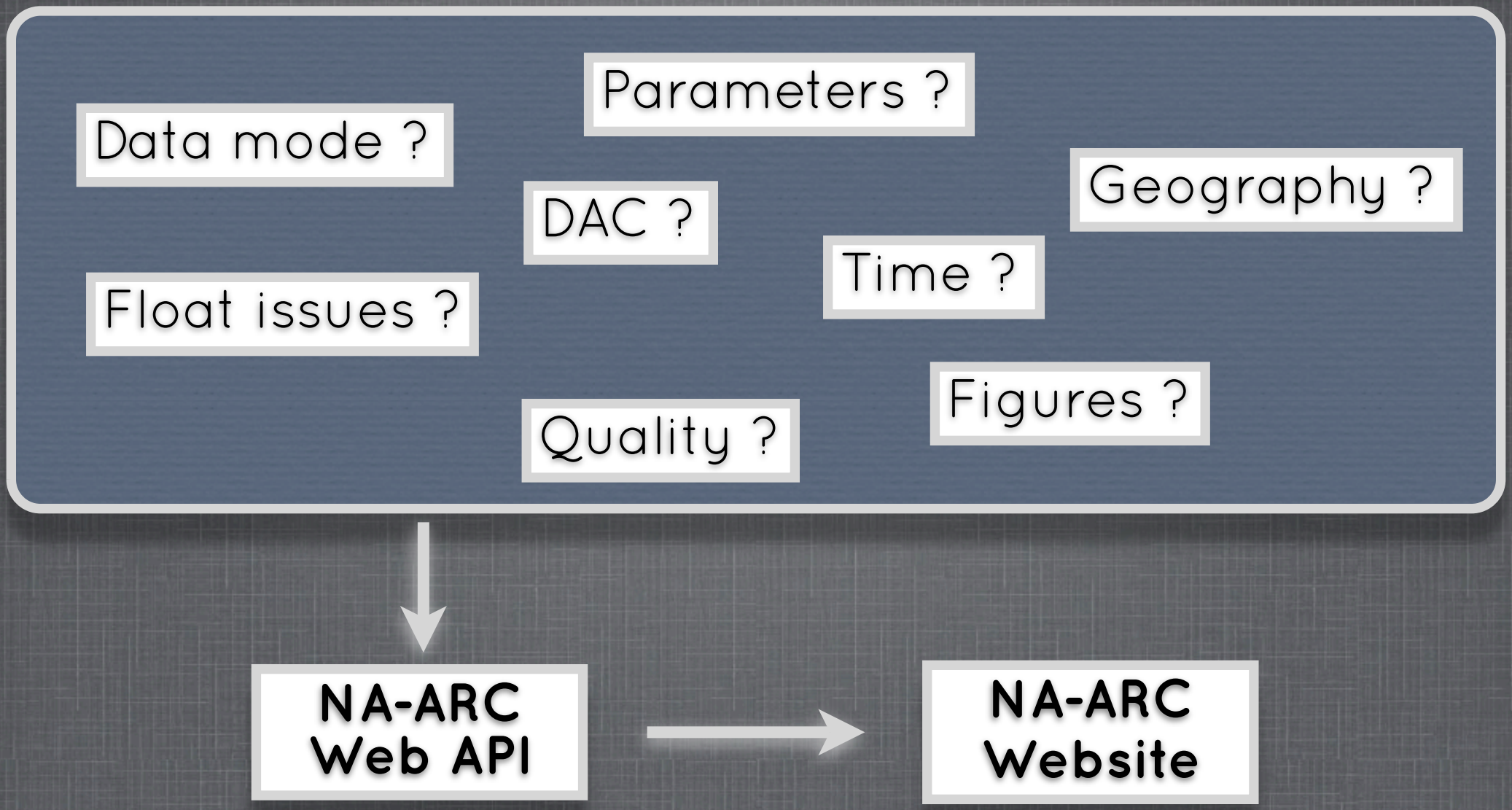
Figures ?

SERVICE PROVIDERS

Argo Information Center: argo.jcommops.org
Coriolis: coriolis.eu.org
FTP: you're on your own !

NOTHING SPECIFIC TO NORTH ATLANTIC DATA

Wouldn't be nice to have a frontdesk, where you could checkin your requirements and obtain answers ?



NEW

INFORMATION & DATA MINING TOOL FOR NORTH ATLANTIC ARGO DATA

NA-ARC DATABASE CONTENT

- ONLY FOR POSITION_QC = [1,2,5,8]
- Float/cycle standard informations: wmo, cycle #, DAC, netcdf file location on the ftp server.
- Informations grabbed from netcdf profiles files: date, latitude, longitude, data mode, parameters and profiles QC.
- Area code: Mediterranean Sea (1), Nordic Seas (2) and 0 otherwise, ie North Atlantic north of 20S.
- Issues: a boolean parameter (ticket) is set to 1 if a profile/float is reported as having an issue on the:
  - ftp greylist,
  - LPO/Argo database,
  - CLS/Coriolis altimetry QC test results
- Links to figures or extra files.

- Deployment plan

Update starts every night at 0h00
New version online around 4h00

WHAT IS A WEB API ?

The simpliest link between a database and an user

The link is done through http

`http://www.google.com/` → *Home page*

`http://www.google.com/search?q=geostrophy` → *Search results*

`http://en.wikipedia.org/w/api.php?action=query&prop=revisions&rvprop=content&titles=Geostrophy&redirects` → *Raw Wikipedia content for "geostrophy"*

Makes the difference between the *old* and the *new* web !
Split the content from the interface:
Allows for mobile and mashup applications

NA-ARC WEB API

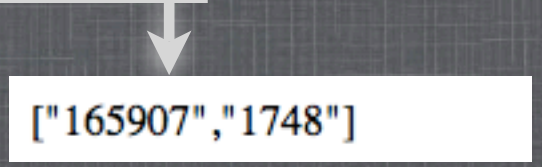
`http://www.ifremer.fr/lpo/naarc/api/v1/?`

Parameters used to formulate a **query** fall into three categories:

1. parameters selecting what we gonna call a **service** (get, plan, file, qwmo, doc)
2. parameters defining the service's **function** and options you want to use (n,list,coord,...),
3. parameters defining **restrictions** on profiles properties (year,area,...)

`http://www.ifremer.fr/lpo/naarc/api/v1/?get=version`

`http://www.ifremer.fr/lpo/naarc/api/v1/?get=n`



NA-ARC WEB API: SERVICE "GET"

Function	Returns ...	Option Name	Option Values	Output Format
version	Current database version	-	-	json, matlab
n,np,nf	# of profiles / floats	by	dac,dmode,year,month,wmo,ticket, area,<params>,<params>_qc	json, matlab, csv
list	List of unique properties	of	dac,dmode,year,month,wmo,ticket, area,<params>,<params>_qc,file	json, matlab, csv
coord	[time,lat,lon] of profiles	-	-	json, matlab, csv, kml
fullcoord	[time,lat,lon,ticket,wmo,cyc,area, doxy,dac] of profiles	-	-	json, matlab, csv, kml
ftplist	a shell script to download netcdf files	-	-	csh
ticket	description of tickets	wmo, qwmo	a WMO string or substring	json

http://www.ifremer.fr/lpo/naarc/api/v1/?get=np&by=year

{"1997":["247"],"1998":["595"],"1999":["634"],"2000":["1717"],"2001":["3704"],"2002":["6702"],"2003":["8107"],"2004":["10098"],"2005":["13736"],"2006":["17878"],"2007":["19881"],"2008":["20078"],"2009":["19615"],"2010":["19660"],"2011":["22177"],"2012":["1078"]}

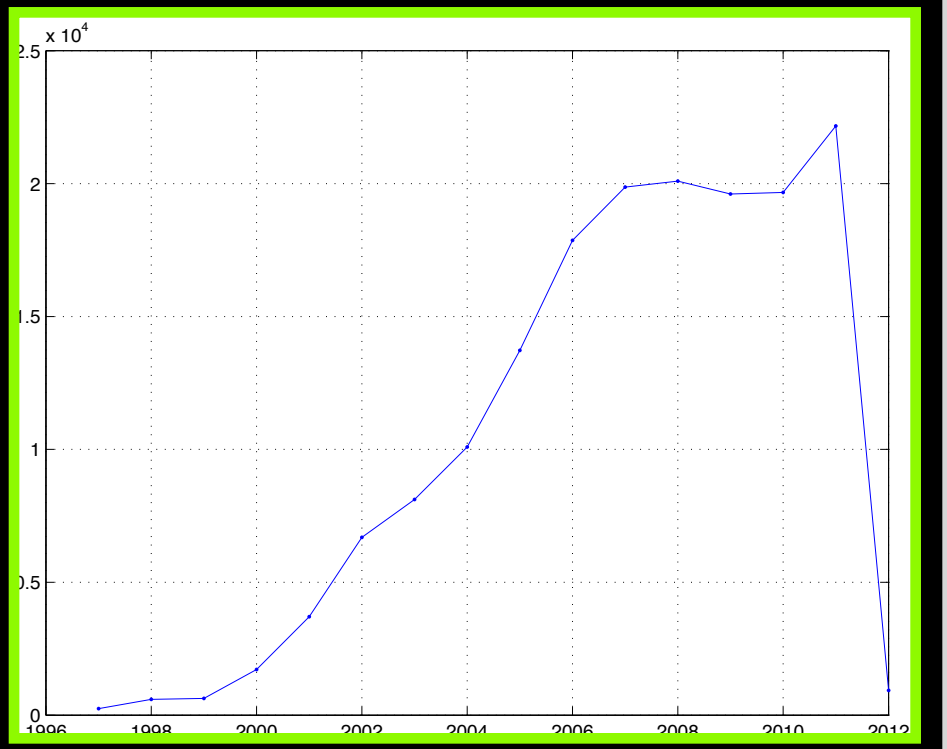
NA-ARC WEB API: RESTRICTIONS

"Meta"	`dac`	any DAC name
	`wmo, cyc`	any WMO id or cycle number
	`par, nopar`	any parameter in the list: pres (p), temp (t), psal (s), doxy (o)
Space/ time	`year, yearmin, yearmax`	any year after 1997 in the form YYYY
	`month, monthmin, monthmax`	any month number between 1 and 12
	`lr`	profile age in days (between 2 and 60)
	`from, to`	minimum/maximum profile date: YYYYMMDD
	`area, box`	bassin code (0,1,2) or rectangular box (x,y,dx,dy)
Data type/ quality	`dmode`	any data mode in: 'R','D','A'
	`<params>_qc`	using any parameter, any profile QC flag: 'A','B','C','D','E','F' and none ''
	`ticket`	0 or 1

```
http: ... /?get=<any_function>&lr=30
http: ... /?get=<any_function>&dmode=D,A
http: ... /?get=<any_function>&dac=coriolis&year=2002
http: ... /?get=<any_function>&from=20111225&to=201112311
http: ... /?get=<any_function>&box=-35,60,8,4&month=12,1,2
http: ... /?get=<any_function>&pres_qc=A,B
```

PROGRAMMATICALY ACCESS DATA FROM MATLAB

```
>> query = 'http://www.ifremer.fr/lpo/naarc/api/v1/?get=np&by=year&format=matlab';
>> urlread(query)

ans =

[1997,246;1998,596;1999,632;2000,1718;2001,3703;2002,6693;2003,8113;2004,10091;2005,13
727;2006,17868;2007,19868;2008,20097;2009,19608;2010,19668;2011,22170;2012,932;]

>> C = eval(urlread(query))

C =

        1997          246
        1998          596
        1999          632
        2000         1718
        2001         3703
        2002         6693
        2003         8113
        2004        10091
        2005        13727
        2006        17868
        2007        19868
        2008        20097
        2009        19608
        2010        19668
        2011        22170
        2012          932

>> figure;plot(C(:,1),C(:,2),'.-');grid on;
```



PROGRAMMATICALY ACCESS DATA FROM PYTHON

```
>>> import urllib
>>> import json
>>> url = "http://www.ifremer.fr/lpo/naarc/api/v1/?get=n&by=dac"
>>> ans = urllib.urlopen(url)
>>> N   = json.load(ans)
>>> for dac in N:
...     np = int(N[dac][0])
...     nf = int(N[dac][1])
...     print dac, np, nf
...
coriolis 75730 827
meds 13397 143
bodc 5940 89
aoml 70718 689
>>>
```

DATA CAN BE ACCESSED BY ANY SOFTWARE HANDLING HTTP REQUEST

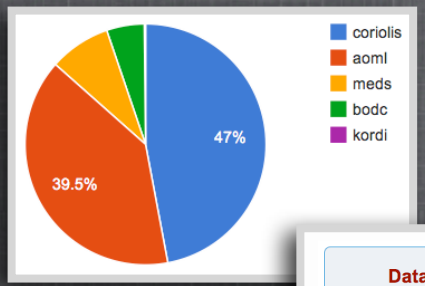
WEB API → WEBSITE

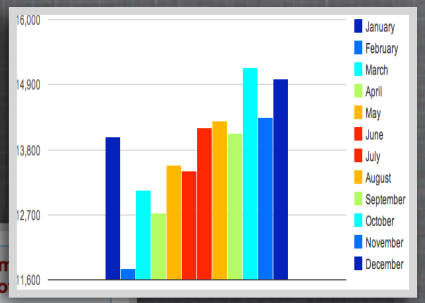
http://www.ifremer.fr/lpo/naarc

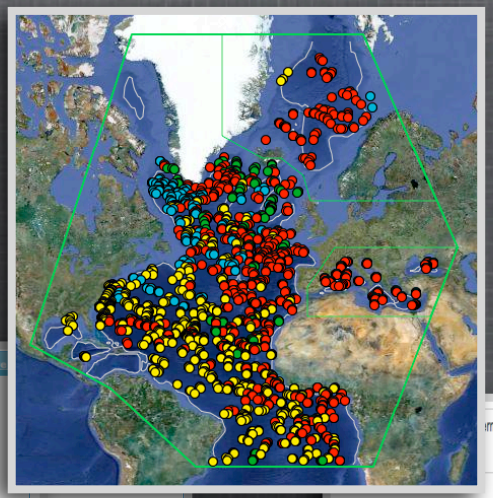
charts & maps (interactive)

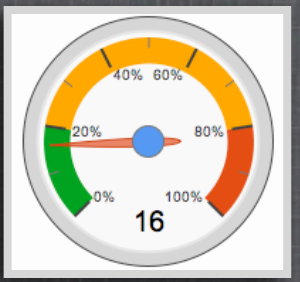
see data (figures, issues)

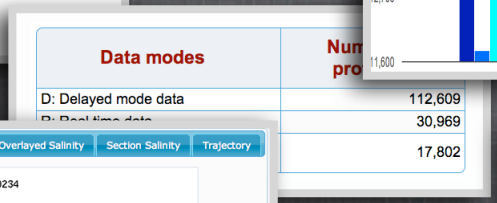
shop for your profiles (wizard)

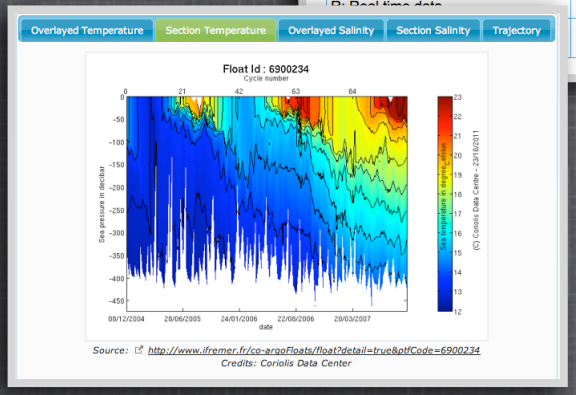


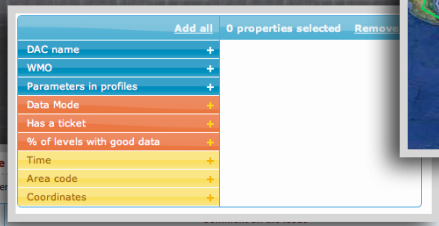


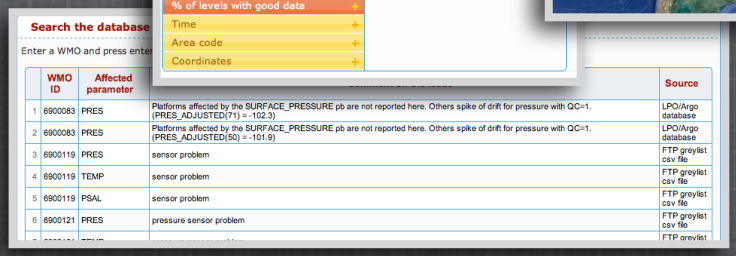


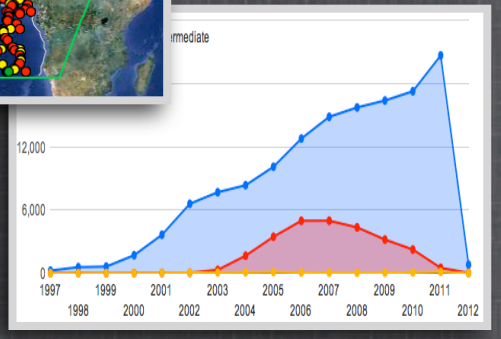












Another restriction parameter:

around=WMO,CYC,RADIUS(km),LAG(days)

Eg: `around=6900678,2,200,60`

around=LON,LAT,RADIUS(km)

Eg: `around=-55,40,300`