SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

key challenges in reconciling needs and requirements of data providers and data systems versus those of users

T. Loubrieu, G.Maudire (IFREMER)

Data providers issues with users

- Research
- Private sectors
- National environment agencies, member states
- National, thematic Data Centres

***Research** vs users*

- **Heterogeneity** of contents (experimental observations) → high cost for integration
- Embargo/**moratorium** on observation data up to scientific paper publication: researchers are asked to provide data, provide contextual information (metadata) late after they gave up interest on it.
- When infrastructure is not critical (e.g coastal deployment, lab analysis): **lack of coordination** for observation and data management

SME, private observation *vs users*

- Specific data policies for big industry, however when client are local authorities no data policy.
- Lack of collaboration with “public” sector at data management level (but existing collaboration at science and observation system development level).

Environment agencies, member states *vs users*

- Member states manage observation and provide indicators for the environment status at national level (e.g. water quality, fisheries).
- They are more likely to deliver indicators to European Agencies (e.g. EAA) than raw observations.

Data centres *vs users*

- Data centres have also issues with sustainable funding, fear for data aspirators
- Data centres counterproductive competition
- Data centres are reluctant to open their data to private partners who would make money from it
- Commercially valuable observations (e.g. seismic) are not free

Data centres *vs users*

- Specific data policies, e.g. observations in territorial waters of non European countries (Mediterranean sea).
- Data management, quality controlled results is often opposed to free circulation of datasets (duplicate management, mis-used, observation quality)

Possible solutions, organization

- **Sustainable funding** for data centres, promote self-confidence of data centres as primary source for data (e.g. Copernicus).
- **Data against fundings**: when data from previous experiment is not delivered to sustainable data centres, don't fund next projects (e.g NSF)
- **Services for data**: Associate partners, organize consortium together with e-infrastructures or research infrastructures where services, organization support are provided against free data policy (e.g. ARGO, H2020 project data management plan)

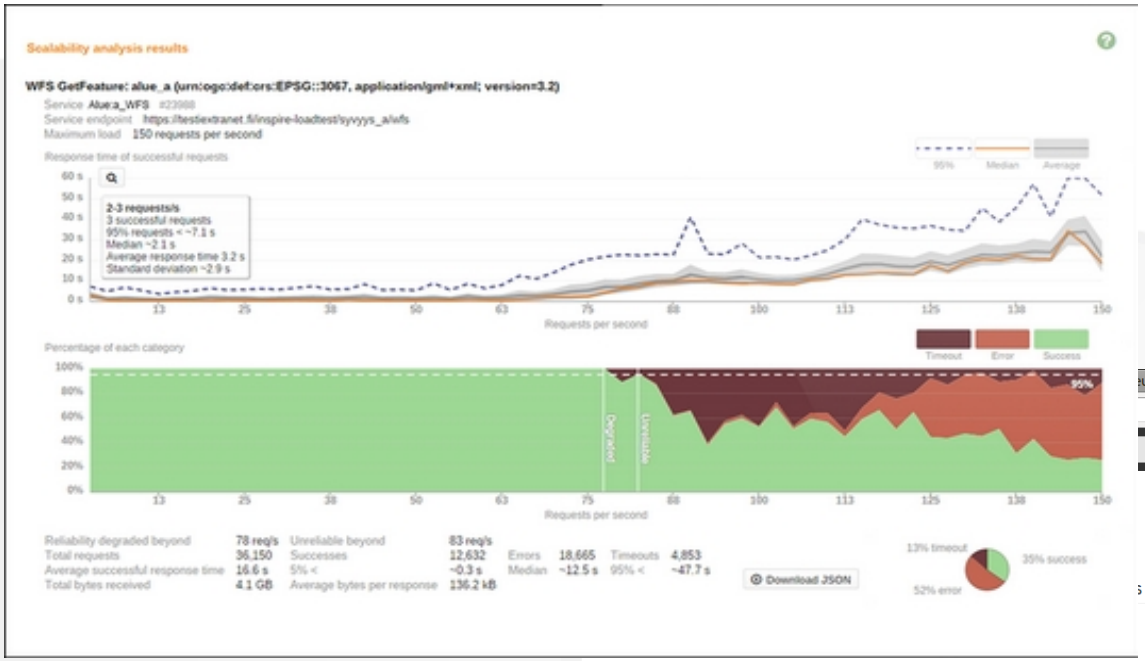
Possible solutions, organization

- Don't forget to **associate key data providers in projects**: for private sector (e.g. ingestion system DG MARE tender) or non European countries (e.g. SeaDataNet)
- **Close, near-by, friendly, non-competing** proxy data centres (e.g. ROOS for Copernicus).
- **Viral Data license**: when open data is compiled in a product, the product must be open as well (Creative Common Attribution-ShareAlike 4.0 International)

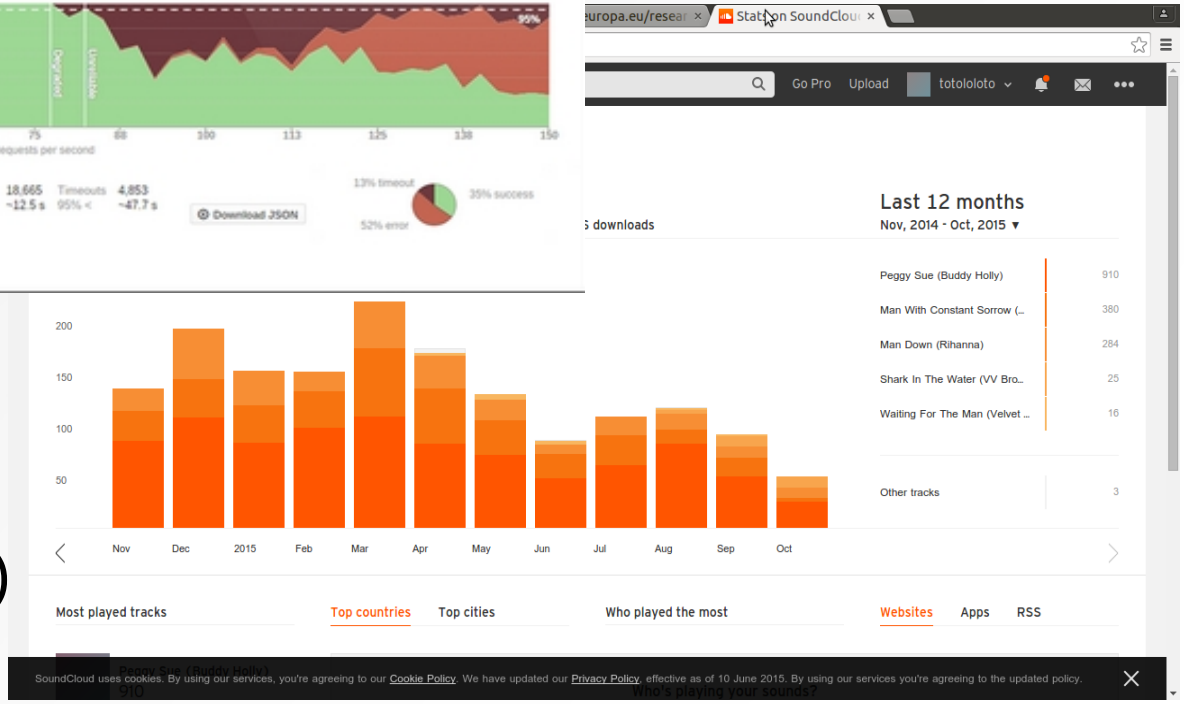
Possible solutions, tools

- **Usage traceability**: open data does not mean anonymous usage
  - Identify users, download transactions to motivate fundings
- Enable **promotion of data contributions**:
  - **DOI**s on dataset: to measure contribution of data to science and knowledge
  - Standardize **data citation** by publishers
  - Provide **feedback, usage statistics** on data (e.g. contribution to big compilations : world ocean data base, gebco, …) for visibility of individual (research) and organization (data centres, private sectors, ...)

Feedback Statistics



(spatineo.com)



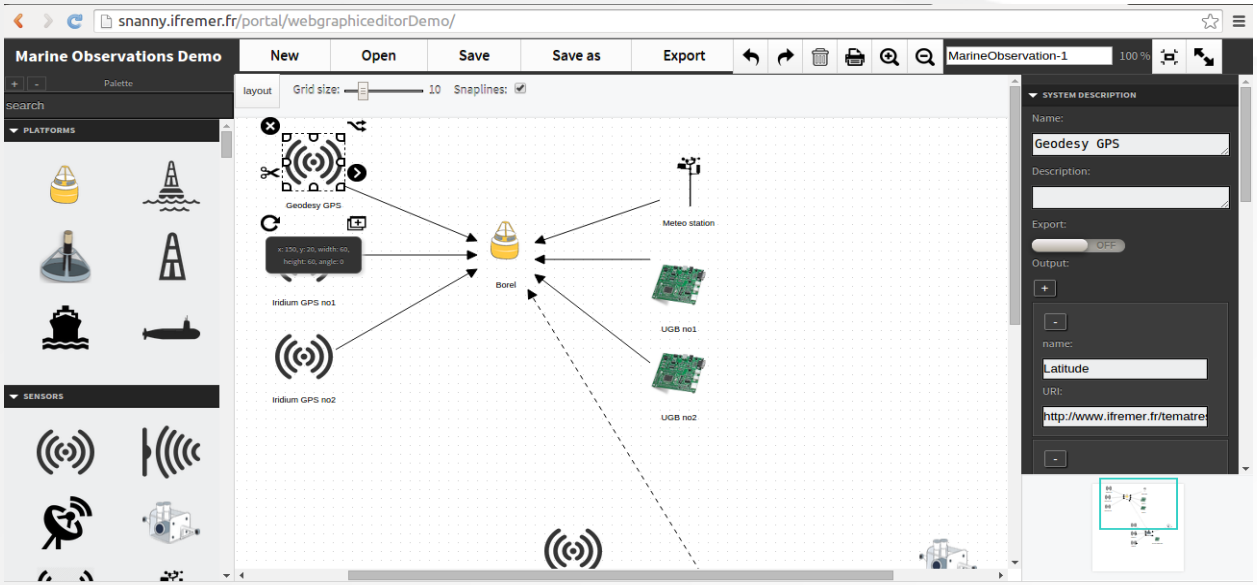
(soundcloud.com)

Possible solutions, tools

- Cloud paradigm "**free services against information**" applied to marine observations: Provide data management tools, collaborative environments as support for observation and research (e.g. VRE, google business model)
- Framework for **observation data distribution**, the "best copy" approach has limitations: replication without creating a messy environment where similar but different data are distributed everywhere:
  - Properly identify origin: Unique platform identifiers
  - Properly identify, describe quality assessment, processing, versions

Collaborative environment, support for observation

Deployment documentation (IFREMER, sensor nanny editor)



+ Alfred Wegener Institut, Ana Macario et al. ,
automated QC and alert on observations:
http://meetingorganizer.copernicus.org/EGU2015/EGU2015-3395.pdf