

Verso l'Open Science in ISMAR: iniziative, progetti e prospettive per la condivisione e l'apertura delle informazioni sulla ricerca marina

Alessandro Sarretta, Alessandra Pugnetti, Stefano Menegon
CNR-ISMAR



Perché Open Science

La pubblicazione e la condivisione del dato hanno diverse **ragioni** e **motivazioni**

- Passaggio da “printing-age” a “**information-age**”
- Motivazione **etiche** (dati finanziati da fondi pubblici)
- Trasparenza e **riproducibilità**
- Migliore **programmazione** delle attività sperimentali e della raccolta di dati nuovi
- Creazione di **nuove conoscenze**, ipotesi, teorie, visioni

La pubblicazione e la condivisione del dato fanno parte integrante della cultura di alcune discipline scientifiche, e.g.: fisica, astronomia, climatologia, biologia molecolare.



Chi sta spingendo verso l'apertura

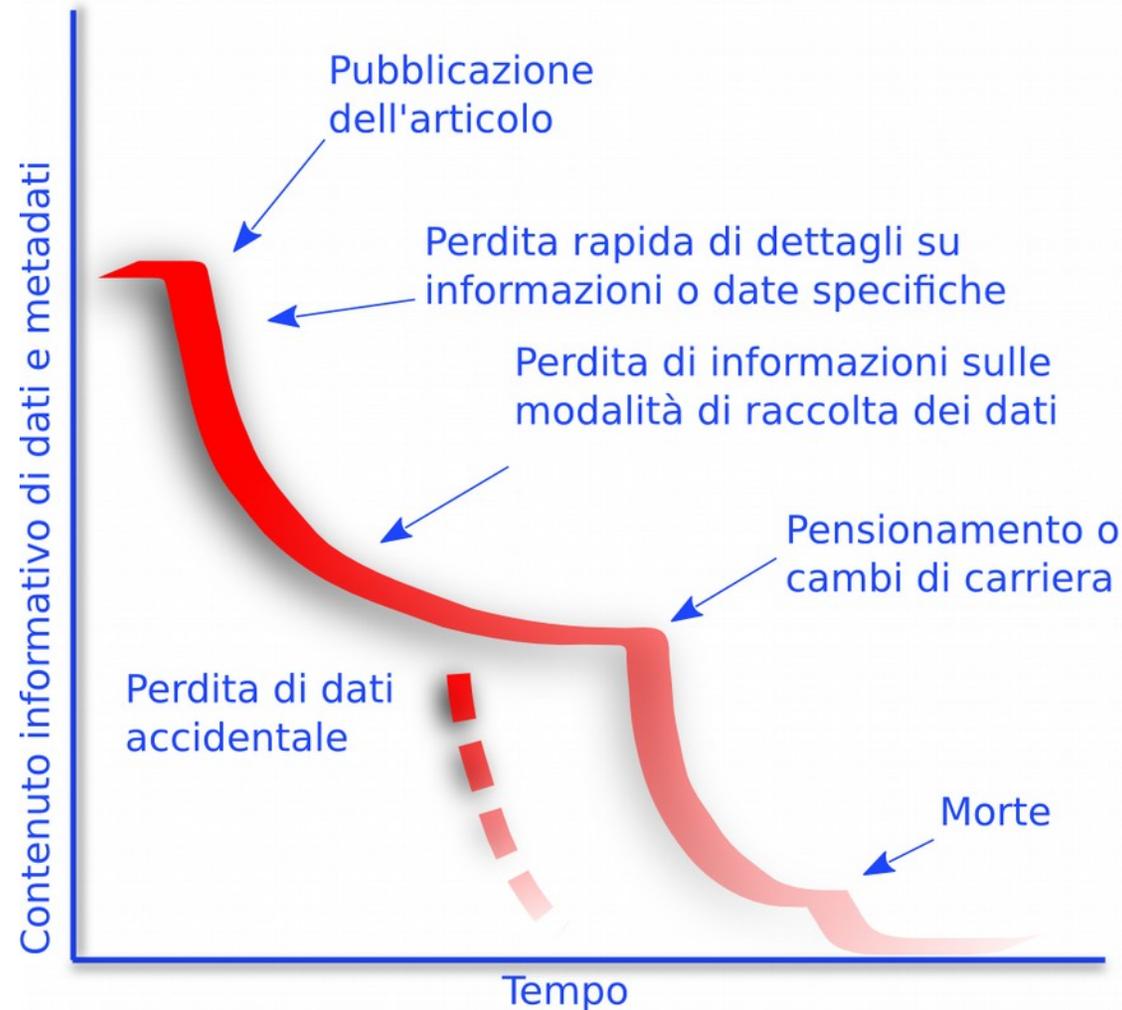
- Molte agenzie/iniziative internazionali hanno sviluppato raccomandazioni, linee guida, direttive volte a archiviare i dati ai fini di renderli disponibili. e.g.: OECD, ICSU, IOC, GBIF, ERC, NSF, NIH.
- Alcune riviste “tradizionali” e non open access, incluse Nature e Science, richiedono esplicitamente che i dati siano resi accessibili, indicando quali possano essere le modalità e gli archivi più idonei.

L'open access alle pubblicazioni e ai dati viene esplicitamente richiesto in ambito H2020:

- Precise linee guida su open access di riviste e dati e su data management
- **“Open access to scientific peer reviewed publications”**
- **“Open Research Data Pilot”** which aims to improve and maximise access to and re-use of research data generated by projects.



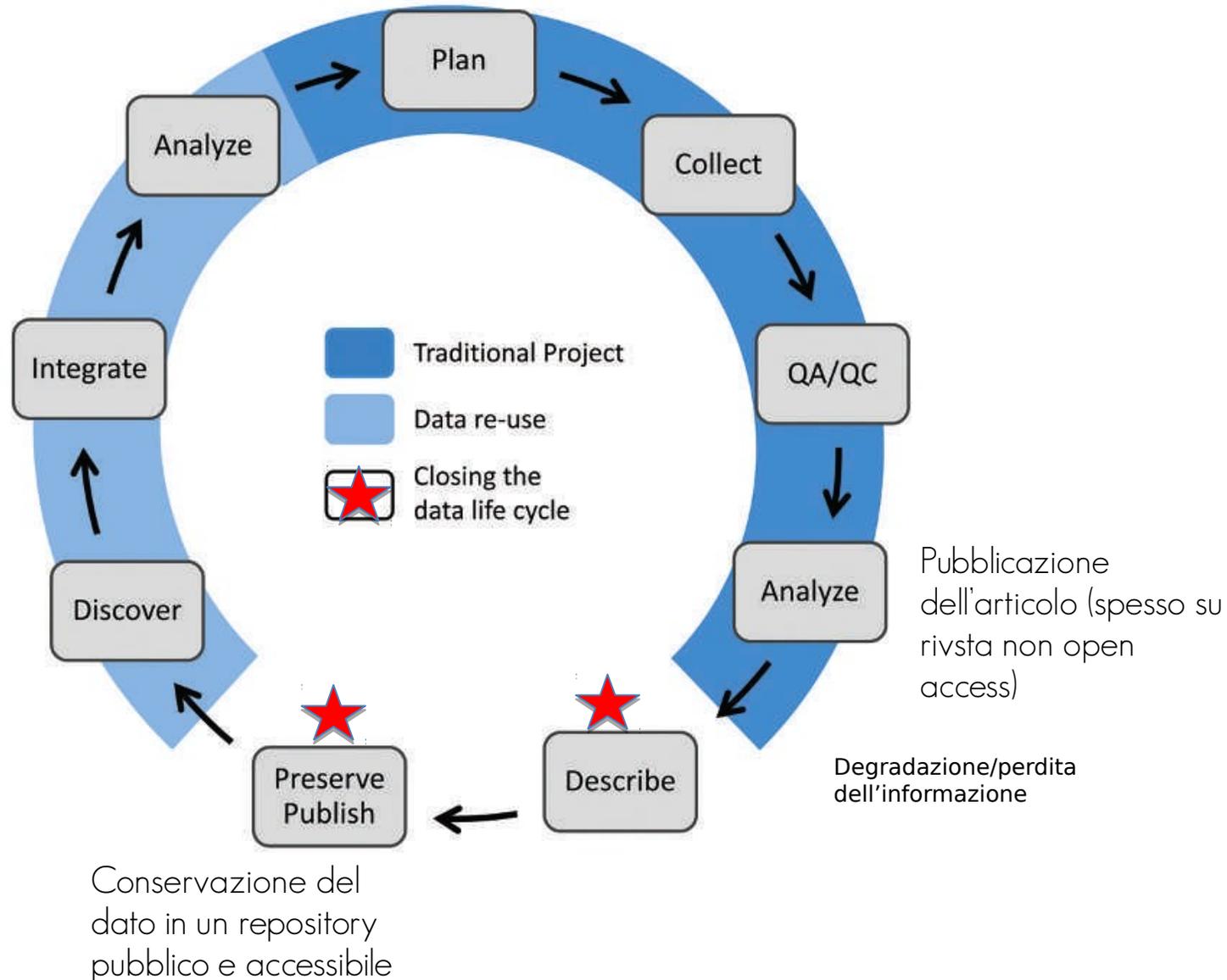
“Degradazione” dell’informazione contenuta nei dati



Example of the normal degradation in information content associated with data and metadata over time ("information entropy"). Accidents or changes in storage technology may eliminate access to remaining raw data and metadata at any time Michener et al. 1997 «Nongeospatial metadata for the ecological sciences», *Ecological Applications*, 7, 1, p. 330-342

Ciclo vitale completo del dato

Il contenuto informativo del dato e dei metadati aumenta



Iniziative nella ricerca marina

- L'importanza della disponibilità e dell'accessibilità ai dati è globalmente riconosciuta anche nella ricerca marina
- In questo ambito si sono sviluppate negli ultimi due decenni varie iniziative, progetti, infrastrutture che si occupano della gestione dell'informazione e letteratura dedicata all'argomento



GBIF



Ricerca marina: livello tecnologico

- Alcune comunità concordano e sono armonizzate (formati dati e standard di interoperabilità, software e tool consolidati), e.g.: **oceanografia fisica, sistemi osservativi, modellistica**
- Altre (e.g. **ecologi e biologi**) sono caratterizzate da una grande eterogeneità nella tipologia dei dati e da difficoltà di armonizzazione e standardizzazione.
- Inoltre, dominano attività condotte dai singoli ricercatori (o piccoli gruppi), su piccole scale spaziali e temporali e con sistemi di finanziamento che non prevedono la gestione, la cura, la pubblicazione e la condivisione del dato



Livello culturale

La **riluttanza** alla condivisione dipende da vari fattori:

- Sistema di **valutazione** basato esclusivamente sulla pubblicazione in riviste (ancora troppo spesso ad accesso chiuso; SISTEMA PUBLISH OR PERISH, troppo legato all'IF penalizzante per le riviste OA)
- Assenza di **riconoscimento** per la pubblicazione e disseminazione dei dati
- Condivisione del dato vista (e temuta) solo in termini di **competizione** e “proprietà” del dato

In alcune comunità (biologi ed ecologi)

- Assenza del **tempo** e delle **risorse** finanziarie necessarie per la cura e la pubblicazione del dato (sul lungo termine)
- Mancanza di conoscenze ICT adeguate e di una **cultura ICT** che tenga in considerazione le reali esigenze delle comunità ecologica

Possibili soluzioni e percorsi

- Avviare un cambiamento nel sistema di valutazione, estendendo la valutazione della ricerca a tutti i suoi **prodotti** (pubblicazioni di articoli e di datasets)
- Estendere le citazioni anche ai datasets oltre che alle pubblicazioni (e.g. DOI)
- Attivare **data policy** che prevedano adeguati tempi di embargo, licenze aperte, modalità di citazione del dato
- Proporre progetti in cui l'information management sia condiviso fra **ICT e Ricercatori** (problematiche, linguaggi): piani realistici di data management e di comunicazione fra le due comunità
- Ruolo dei **Journals**: rendere norma la pubblicazione dei dati assieme al lavoro
- Ruolo delle **istituzioni** : linee guida chiare, finanziamenti e mantenimento long-term, percorsi formativi
- Ruolo culturale delle **società scientifiche**

Data policy **Ritmare**



1. Universalità

→ Applicazione uniforme all'interno e all'esterno del progetto

2. Attribuzione

→ Obbligo di citazione del generatore o proprietario del dato

3. Collaborazione scientifica

→ Obbligo di richiesta di collaborazione scientifica nei primi 2 anni

4. Periodo di moratorium

→ Periodo 6, 12, 18 mesi nei quali i dati rimangono in gestione esclusiva

5. Metadattazione

→ I dati devono essere accompagnati da dati ancillari e metadati

6. Rispetto policy per dati pregressi

→ Ai dati non raccolti all'interno del progetto si applicano licenze a loro associate

7. Applicazione licenza

→ Al termine del moratorium si applicano Licenze (preferibilmente aperte → CC-BY)

Le regole sono da considerare i vincoli massimi applicabili ai dati.
E' data la libertà di applicare regole meno restrittive, ovvero diminuire o eliminare i periodi di moratorium e/o di collaborazione scientifica

Licenze

- La scelta della licenza da applicare ai prodotti della ricerca ne condiziona il possibile **riutilizzo**
- Ci sono chiare indicazioni a livello nazionale, europeo e internazionale ad adottare **licenze standard**:
 - European Commission Guidelines on recommended standard licences, datasets and charging for the re-use of documents (2014)
 - OpenAIRE: Safe to be open: Study on the protection of research data and recommendations for access and usage

Licenze Creative Commons

(<http://creativecommons.org/>)

Organizzazione non a fini di lucro che nasce con l'intenzione di armonizzare l'articolato mondo del diritto d'autore.
Definisce varie licenze con diversi gradi di libertà.



No Rights Reserved CC0



Attribution CC BY



Attribution-ShareAlike CC BY-SA



Attribution-NoDerivs CC-BY-ND



Attribution-NonCommercial CC BY-NC



Attribution-NonCommercial-ShareAlike CC BY-NC-SA



Attribution-NonCommercial-NoDerivs CC BY-NC-ND

This work is licensed under a
Creative Commons Attribution 4.0 International License.



OPEN  **International
ACCESS WEEK**