



UNIONE EUROPEA
Fondo Sociale Europeo
Fondo Europeo di Sviluppo Regionale



FSC
Fondo per lo Sviluppo
e la Coesione

Consiglio Nazionale delle Ricerche
Dip. Scienze del Sistema Terra e Tecnologie per l'Ambiente



MARINE HAZARD

Progetto
Marine Hazard – Sviluppo di tecnologie innovative per
l'identificazione, monitoraggio e mitigazione di
fenomeni di contaminazione naturale e antropica

Documento per la definizione di una politica nella
gestione e utilizzo dei dati e dei prodotti resi
disponibili nell'ambito del progetto *Marine Hazard*

Regole condivise per l'accesso e l'utilizzo dei dati e
delle informazioni



CNR
IAS
ISTITUTO PER LO STUDIO
DEGLI IMPATTI ANTROPICI
E SOSTENIBILITÀ
IN AMBIENTE MARINO



CNR
ISMAR
ISTITUTO
DI SCIENZE
MARINE



Agenzia nazionale per le nuove tecnologie,
l'energia e lo sviluppo economico sostenibile



CoNISMa
Consorzio Nazionale
Interuniversitario
per le Scienze del Mare



MATER
MAnagement
Training
Education
Research



L'attività descritta nella presente pubblicazione è stata finanziata dal Progetto ***Marine Hazard – Sviluppo di tecnologie innovative per l'identificazione, monitoraggio e mitigazione di fenomeni di contaminazione naturale e antropica*** - Coordinato dal Consiglio Nazionale delle Ricerche e finanziato dal Ministero dell'Università e della Ricerca nell'ambito del Piano Stralcio «Ricerca e innovazione 2015-2017».

Autori: Gruppo di Lavoro sulla data policy del Progetto Marine Hazard

Roberto De Michele, Alcide di Sarra, Laura La Gattuta, Manfredi Longo, Nadia Sabatino, Mattia Vallefucio

Questo documento mutua i principi e si basa largamente sulla data policy del Progetto RITMARE (Ribera l'Alcalà et al., 2014; <http://www.ritmare.it/area-download?download=187:data-policy-new>). Il progetto RITMARE (<http://www.ritmare.it/>), coordinato dal CNR, ha riunito la comunità scientifica italiana, costituendo un ampio partenariato di Enti pubblici di Ricerca coinvolti in attività di ricerca su temi marini e marittimi. Gran parte del partenariato pubblico coinvolto in RITMARE partecipa al progetto Marine Hazard. Le linee guida sulla politica di gestione dei dati di RITMARE, già adeguatamente approfondite e condivise dagli Enti partecipanti, sono applicabili alla comunità di Marine Hazard. Pertanto, la data policy di RITMARE è stata qui riproposta aggiornata alle esigenze del Progetto Marine Hazard.

Marzo 2021

INDICE

Sintesi	2
Executive summary	2
1. Introduzione	3
2. Termini dell'accordo di data policy	5
2.1 Premessa	5
2.2 Scopo della data policy di Marine Hazard	5
2.3 Tipologie di dati	6
2.4 Regole vincolanti per l'utilizzo dei dati/prodotti	7
2.5 Buona pratica	8
2.6 Governance della data policy	8
2.7 Entrata in vigore, durata e termine degli accordi sulla data policy	9
2.8 Norma Transitoria	9
2.9 Attività prodromica all'accordo sulla data policy	9
2.10 Indicazione finale	9
3. Glossario	10
Appendice	12
Riferimenti	13

Sintesi

Questo documento descrive le regole individuate dal gruppo di lavoro sulla *data policy*, e proposte per l'approvazione da parte dallo *Steering Committee*, per l'accesso e l'utilizzo dei dati prodotti dal progetto **Marine Hazard – Sviluppo di tecnologie innovative per l'identificazione, monitoraggio e mitigazione di fenomeni di contaminazione naturale e antropica**.

La proposta è il più possibile vicina ad un accesso/uso di tipo aperto a dati e prodotti, ma al tempo stesso considera le esigenze dei generatori di dati e prodotti di veder riconosciuta la loro attività e di avere un tempo sufficiente per poterne comunicare prima di altri le implicazioni scientifiche.

Nella definizione delle modalità di utilizzo si è tenuto conto che diverse tipologie di dati possono essere prodotte dal Progetto; inoltre, il Progetto utilizza e gestisce anche dati provenienti da altre fonti o prodotti precedentemente all'inizio del Progetto, e sono state definite anche in questi casi le modalità di citazione ed utilizzo.

Executive summary

This document contains the data policy rules defined by the data policy Working Group of the **Marine Hazard - Development of innovative technologies for identification, monitoring and mitigation of natural and anthropic contamination processes** Project. The document is submitted by the Working group to the approval by the Project Steering Committee, and is intended to regulate access and use of data produced by the Project.

The proposed rules address two different needs: an open and free access to data and products, and acknowledgement of data producers and generators, with the definition of a timeframe for elaboration and communication of scientific results.

Different data types and stages have been taken into consideration in the definition of the rules. The rules on the use and citation of data originated by different sources, or produced prior to the Project, have also been set.

1. Introduzione

Il testo che segue stabilisce le regole sottoposte per l'approvazione allo *Steering Committee* del Progetto **Marine Hazard – Sviluppo di tecnologie innovative per l'identificazione, monitoraggio e mitigazione di fenomeni di contaminazione naturale e antropica**; queste regole definiscono la *data policy* del progetto.

Esiste una generale tendenza, a livello sia degli Stati e delle organizzazioni interstatali come l'UE, sia delle iniziative organizzate in programmi, associazioni etc. della comunità scientifica di favorire o imporre la libera circolazione dell'informazione scientifica, socioeconomica ed ambientale, soprattutto se prodotta con il supporto di fondi pubblici. Un recente esempio è il Decreto Legge dell'8 Agosto 2013 che afferma chiaramente nell'art. 4 che il finanziamento pubblico dell'attività di ricerca implica l'accessibilità in tempi brevi e senza costi dei suoi risultati per tutti gli interessati.

L'attuale quadro normativo non è sufficientemente dettagliato per poter essere immediatamente applicabile a tutti i tipi di dati e di prodotti che verranno generati in Marine Hazard. Tuttavia il principio generale è che la mole di dati generati in anni recenti e quelli generabili nel prossimo futuro, in virtù del rapidissimo sviluppo di tecnologie mirate all'osservazione di processi naturali su tutte le scale, da quella molecolare a quella astronomica, può essere valorizzata a pieno solo favorendo la sua esplorazione ad opera di una comunità molto più vasta di quella che li ha prodotti. In parallelo vi è anche la convinzione, dimostrata dai fatti, che vi è una ricaduta positiva, sia a livello di conoscenze che di visibilità, per chi ha prodotto i dati se essi vengono condivisi.

Esistono vari elementi che suggeriscono una politica di accesso il più possibile aperto a dati e prodotti. In particolare, relativamente a dati e prodotti del Progetto Marine Hazard, è necessario tenere conto che:

1. sono dati prodotti con fondi pubblici e pertanto ne va favorito un uso pubblico anche oltre la comunità Marine Hazard;
2. la condivisione aperta permette un avanzamento più rapido delle conoscenze con un effetto positivo sui risultati del progetto ed in generale della ricerca. Come con altre piattaforme di condivisione, questo consente -tenuto conto delle regole concordate- una migliore organizzazione del lavoro scientifico e di coordinamento tra i vari SP di Marine Hazard;
3. la condivisione dei dati, contrariamente ad una percezione diffusa, favorisce anche la visibilità e la produzione scientifica dei singoli che li hanno generati;
4. la condivisione permette anche il riutilizzo dei dati da parte di altri che possono avere obiettivi parzialmente diversi dagli "acquisitori" (portando a un maggiore sfruttamento dei dati acquisiti) e/o possono sviluppare meglio ricerche successive alla luce di tali dati (es. disponibilità di una batimetria in fase di progetto di una nuova ricerca biologica, oceanografica, geologica).

Dati e prodotti

Il gruppo di lavoro ha elaborato una proposta il più possibile vicina ad un accesso/uso di tipo aperto a dati e prodotti, ma ha anche considerato le esigenze dei generatori di dati e prodotti di veder riconosciuta la loro attività e di avere un tempo sufficiente per poterne comunicare prima di altri le implicazioni scientifiche. Questo documento ha mutuato impostazione e ambiti di applicazione identificati nella *data policy* elaborata dalla Comunità Italiana che ha lavorato sul Progetto RITMARE (Ribera l'Alcalà et al., 2014; <http://www.ritmare.it/area-download?download=187:data-policy-new>). Il progetto RITMARE (<http://www.ritmare.it/>), coordinato dal CNR, ha riunito la comunità scientifica italiana, costituendo un ampio partenariato di Enti pubblici di Ricerca coinvolti in attività di ricerca su temi marini e marittimi. Gran parte del partenariato pubblico coinvolto in RITMARE partecipa oggi

al progetto Marine Hazard. Le linee guida sulla politica di gestione dei dati di RITMARE, già adeguatamente approfondite e condivise dagli Enti partecipanti, sono applicabili alla comunità di Marine Hazard. Pertanto, la data policy di RITMARE è stata qui riproposta aggiornata alle esigenze del Progetto Marine Hazard.

Seguendo le linee tracciate dalla *data policy* del Progetto RITMARE, è stato anche evidenziato il ruolo che, nella condivisione di dati e prodotti dell'attività intellettuale, hanno le Licenze. Nel testo si fa riferimento ai possibili tipi di licenze, già codificati, che possono essere utilizzati per la condivisione di dati e prodotti. Il gruppo di lavoro ritiene che poter fare riferimento ad un insieme di regole chiaramente definito nelle specifiche di una Licenza è senz'altro un obiettivo da raggiungere, ma ritiene altresì che sia prioritario convergere verso regole condivise indipendentemente dal fatto che si conformino ad un tipo di Licenza piuttosto che un altro. D'altra parte questa è la modalità più comunemente adottata nella maggior parte delle comunità scientifiche e dei progetti internazionali. Ne deriva quindi che in questa proposta le licenze, così come formalizzate, vengono applicate solo al termine di un periodo di tempo, variabile in relazione al tipo di dati o prodotti, in cui la circolazione dei dati avviene secondo regole specifiche di Marine Hazard.

D'altra parte le regole specifiche di Marine Hazard costituiscono il numero massimo di vincoli accettati per la circolazione di dati/prodotti e ciascun ricercatore è libero di renderli meno restrittivi o eliminarli del tutto, come viene precisato nel seguito.

Quanto proposto non preclude minimamente la possibilità di aggiornare, durante lo svolgimento del programma, la data policy in modo da conformarla ai tipi di licenze, esistenti o di nuova introduzione, più rispondenti alle caratteristiche dei dati e dei prodotti di Marine Hazard.

Considerazioni analoghe valgono per i metadati. Nonostante esistano standard riconosciuti per la struttura e l'informazione contenuta nei metadati, si ritiene preferibile in questa fase iniziale stabilire e concordare caso per caso il tipo di metadati da fornire, rimandando ad una fase successiva l'adeguamento agli standard esistenti.

Il gruppo di lavoro evidenzia inoltre come si vada consolidando una tendenza, ormai irreversibile a livello internazionale, per favorire anche normativamente la libera circolazione dell'informazione. Questa tendenza ha prodotto in alcuni casi direttive ed indicazioni dettagliate, con regole di condivisione che prima o poi saranno vincolanti per tutti gli operatori che producono dati o generano prodotti con il supporto di fondi pubblici. Si è però scelto di non imporre d'ufficio e da subito questi vincoli, ma di individuare invece una modalità che, pur mantenendo tra gli obiettivi primari di Marine Hazard il miglioramento del livello di integrazione e di produzione scientifica, lo favorisse nella prima fase, attraverso il coinvolgimento diretto della comunità marina italiana in elaborazioni indipendenti dei generatori dei dati/prodotti.

Poiché esiste un'ampia gamma di dati e prodotti, che richiedono impegni assai differenti per la loro generazione, è stato fatto uno sforzo preliminare di classificazione per poter calibrare le regole della data policy in relazione al tipo di dato o prodotto. Raggiungere un accordo anche sulla classificazione dei dati e prodotti è un elemento essenziale per varare la data policy di Marine Hazard.

Pubblicazioni scientifiche

La UE e molte istituzioni scientifiche internazionali suggeriscono di rendere accessibili le pubblicazioni scientifiche derivate da attività di ricerca sostenute con fondi pubblici entro sei mesi dalla loro finalizzazione. Questa indicazione non è sempre ottemperabile, almeno in linea di principio, se la scelta della rivista segue principalmente la pertinenza, la visibilità ed il prestigio, dal momento che solo una parte delle riviste scientifiche hanno politiche editoriali che consentono l'accesso libero alla consultazione, alcune dietro pagamento da parte degli autori. Anche la valutazione della ricerca, recentemente effettuata dall'ANVUR, è stata basata su criteri bibliometrici che spesso vedono le

riviste con accesso libero collocate più in basso nelle graduatorie rispetto a riviste di più lunga tradizione e prestigio che solitamente non permettono un accesso libero ai loro contenuti. Ma questo scenario potrebbe cambiare nei prossimi anni, per cui risulta al momento difficile dettare criteri su questo specifico aspetto. Esistono varie soluzioni per conciliare, in larga misura, la spinta dei ricercatori a pubblicare i loro lavori sulle riviste più importanti e l'indicazione a rendere pubblici i risultati delle loro ricerche.

Individuare delle riviste di qualità che garantiscano l'accesso libero, giovandosi anche del fatto che l'editoria scientifica è in grosso fermento dovendo affrontare i pro e i contro dell'accesso libero, cosa che sta portando ad un aumento delle riviste che optano per quella politica.

Pubblicare su riviste che richiedono un pagamento aggiuntivo per concedere l'accesso libero, programmando ovviamente la conseguente spesa aggiuntiva, che dovrebbe essere tenuta presente nel finanziamento dei programmi di ricerca.

Mettere a disposizione (in repository istituzionali o di settore) una versione nel formato di bozza del lavoro o comunque non nel formato stampa che, in quanto preliminare e non corrispondente esattamente alla pubblicazione finale, spesso non viene considerata infrangere il copyright della rivista (il cosiddetto Green Open Access per differenziarlo dai due precedenti che rientrano nel Gold Open Access).

Il gruppo di lavoro ritiene che, pur essendo l'utilizzo di riviste a Gold Open Access la soluzione probabilmente ideale per la circolazione dell'informazione, non sia possibile imporlo come regola obbligatoria, almeno fino a che non vengano raggiunti accordi internazionali che creino condizioni uguali per tutta la comunità scientifica. Può essere però richiesto che il contenuto dei lavori pubblicati venga reso accessibile a tutti senza costi utilizzando le forme del Green Open Access, col doppio vantaggio di una migliore diffusione e di un aumento di visibilità dei risultati ottenuti.

2. Termini dell'accordo di data policy

2.1 Premessa

Marine Hazard propone un insieme di regole per l'accesso e l'utilizzo dei dati, delle elaborazioni e dei prodotti generati nel corso dell'attività del progetto, che contemplino altresì un giusto riconoscimento allo sforzo condotto a produrli. Queste regole, mutate dalla data policy di RITMARE, sono generalmente compatibili con la normativa corrente o con i tipi di Licenze esistenti, e non sono automaticamente riconducibili ad un singolo tipo di Licenza o di norma. Le regole previste dai vari tipi di Licenza esistenti sono consultabili su <http://opendefinition.org/licenses/>.

2.2 Scopo della data policy di Marine Hazard

La data policy di Marine Hazard è finalizzata a favorire la sinergia verso una piena integrazione della comunità scientifica marina italiana attraverso il trasferimento e la condivisione di dati e di conoscenza in modo da migliorare la qualità della ricerca prodotta, la sua visibilità e la sua competitività a livello internazionale.

Essa deve quindi favorire la realizzazione di un quadro di conoscenza condiviso, "aperto", per ottimizzare:

- a) l'attività di ricerca;
- b) i processi decisionali;

- c) il coinvolgimento dei cittadini nelle problematiche ambientali e di gestione del territorio;
- d) un accesso "trasparente" e semplificato alle risorse da parte degli utenti.

La data policy di Marine Hazard non si occupa invece dei diritti della Proprietà Intellettuale (IPR) legata alle innovazioni tecnologiche, invenzioni, nuove procedure sperimentali, software di visualizzazione, etc. che sono collegate alle attività di trasferimento tecnologico di cui si occupa un diverso Gruppo di Lavoro afferente al Progetto.

2.3 Tipologie di dati

In relazione alle attività del programma i dati si possono distinguere nei seguenti modi:

- Dati che fanno parte del Background (B);
- Dati che fanno parte del Foreground (F).

Inoltre i dati si possono classificare in base al grado di elaborazione:

- dati grezzi (rawD): dati ambientali di parametri standard grezzi con dati ancillari, dati di sequenza, etc.;
- dati pre-elaborati (srawD): dati ambientali calibrati e QC, dati tassonomici, dati di traccianti chimici non standard, dati prodotti da esperimenti in situ o in laboratorio, dati risultanti da simulazioni numeriche di routine;
- prodotti (proD): modelli, dati interpretati, rappresentazioni grafiche sintetiche, carte tematiche, etc.

o in relazione alle procedure:

- Dati osservativi. Dati generati con misure od osservazioni di variabili marine, atmosferiche, biologiche etc., a contatto con il sistema o per via remota, da singoli o gruppi, come semplici registrazioni di osservazioni umane dirette o come dati strumentali che esaminano il sistema senza perturbarlo e senza vincolare alcuna delle variabili che ne regolano il funzionamento. Dati osservativi sono ad esempio quelli generati da satelliti, programmi di monitoraggio, survey esplorativi, conteggi microscopici di campioni di comunità naturali, sequenze genomiche di organismi, etc.
- Dati computazionali o di modelli. Dati generati da simulazioni modellistiche o da analisi statistiche che forniscono una rappresentazione virtuale del funzionamento o catturano alcuni pattern caratteristici. Appartengono a questa categoria i dati prodotti dai modelli di circolazione atmosferica per le previsioni del tempo, le statistiche sulle maree o sulle catture, etc.
- Dati sperimentali ottenuti da esperimenti in situ o in laboratorio. Osservazioni derivate a seguito di una ragionata perturbazione di un sistema allo scopo di verificare una o più ipotesi e di esplorare aspetti del suo funzionamento.

2.4 Regole vincolanti per l'utilizzo dei dati/prodotti

Le regole descritte nel seguito sono specifiche del progetto Marine Hazard e sono differenti da quelle previste dalle Licenze prese a riferimento (<http://opendefinition.org/licenses/>).

In quanto più restrittive delle seconde, le regole della data policy valgono per il periodo di tempo limitato e definito per ciascun tipo di dato/prodotto (vedi Regola 4).

Sono esclusi da queste regole i Dati Sensibili (v. glossario).

Eventuali eccezioni a queste regole per casi particolari (dati potenzialmente sensibili per motivazioni socio – economiche) possono essere discusse con lo *Steering Committee* del Progetto Marine Hazard.

Regola 1. Le regole per l'utilizzo dei dati/prodotti si applicano senza differenze sia ai partecipanti al programma che al resto dei potenziali utenti al di fuori di esso.

Regola 2. In tutti i casi in cui vengano utilizzati dati/prodotti va obbligatoriamente citato il generatore e, se diverso, il proprietario.

Regola 3. Entro i primi 2 anni dalla data di generazione dei dati/prodotti (creazione di un data base di raw data), chiunque ne faccia uso per pubblicazioni scientifiche, abstract, rapporti tecnici, etc. è obbligato a verificare l'interesse del generatore (e/o proprietario) a partecipare alla loro elaborazione come co-autore.

Regola 4. Per ogni tipo di dati/prodotti viene stabilito un periodo di tempo (embargo) che parte dalla data di loro generazione¹, durante il quale i dati/prodotti restano nella disponibilità del generatore (proprietario) che ne decide liberamente le modalità di utilizzo.

Nel rispetto della regola 3, per i dati di Foreground si applicano i seguenti periodi di embargo:

Embargo di mesi 6 (o del tempo minimo necessario per trasferire i dati/prodotti in un formato appropriato per l'utilizzo di terzi)

Dati ambientali di parametri standard grezzi (rawD) nei formati tipici in cui sono generati all'origine.

Embargo di mesi 12

Dati ambientali calibrati e QC (srawD); dati derivati da esperimenti numerici (proD); Modelli numerici (modP).

Embargo di mesi 18

Dati prodotti da esperimenti in situ o in laboratorio (srawD)

Dati interpretativi elaborati a partire da serie di osservazioni ambientali o esperimenti in situ o in laboratorio.

Regola 5. I dati grezzi (rawD) devono essere accompagnati da tutti i dati ancillari necessari per la derivazione, indipendentemente dai valori dei parametri cui la raccolta dei dati grezzi era finalizzata (file di calibrazione, specifiche dello strumento, etc.).

Regola 6. Per i dati/prodotti di Background, qualora immediatamente disponibili, si applicano le Licenze o le regole d'uso a loro associati.

¹ Viene identificato come istante di "generazione" del dato la data di consegna del relativo deliverable di Marine Hazard (dove il dato viene presentato, descritto, ecc.).

Regola 7. Al termine del periodo di embargo (di durata variabile dai 6 ai 18 mesi, a seconda della tipologia dei dati), ai dati/prodotti di Foreground viene automaticamente associato uno dei tipi di Licenza riportata in Appendice, scelta dal generatore o dal proprietario dei dati/prodotti.

E' necessario evidenziare che, pur essendo state considerate come opzioni valide all'interno del progetto Marine Hazard le licenze che limitano il riutilizzo del dato per scopi solamente non commerciali (CC-BY-NC, CC-BY-NC-SA, CC-BY-NC-ND) esse non sono considerate compatibili con i concetti di open access e open data secondo la Open Definition (<http://opendefinition.org>) e le principali indicazioni di livello nazionale ed internazionale; per questo motivo, viene suggerito l'utilizzo di licenze che richiedano solamente l'attribuzione, con al massimo la richiesta di distribuzione allo stesso modo (CC-BY, CC-BY-SA, IODL, ODBL).

In particolare si incoraggia l'utilizzo della licenza CC-BY, essendo quella che maggiormente viene utilizzata sia per la pubblicazione di articoli in Open Access, sia per la pubblicazione di molte serie di dati e che massimizza il possibile riutilizzo dei dati e la compatibilità tra le licenze.

Si auspica che, ove esistano dati o banche dati non espressamente disponibili, ma potenzialmente utili al mondo di Marine Hazard, venga resa disponibile almeno la loro descrizione (attraverso la compilazione dei metadati), indicando se i dati sono sensibili o derivanti da enti e istituzioni che hanno beneficiato di finanziamento pubblico.

Si auspica inoltre che un catalogo di questi metadati e, in caso di disponibilità potenziale, il contatto con il generatore o gruppo generatore, o proprietario del dato o banca dati sia disponibile per tutta la comunità, dentro e fuori Marine Hazard. Questo ne favorirebbe il recupero nell'ambito di Marine Hazard, con la collaborazione e il lavoro dei ricercatori "fruitori" interessati, sempre rispettando le regole della d.p.

2.5 Buona pratica

I partner si impegnano a rendere accessibile sul sito pubblico del Programma Marine Hazard una descrizione sintetica delle attività in corso finalizzate alla generazione di dati o prodotti e della loro durata prevista, al fine di far conoscere il tipo di dati/prodotti che si renderà disponibile al termine dell'attività a tutti potenziali utenti dentro e fuori Marine Hazard.

I Partner si impegnano a rilasciare i metadati (v. glossario) relativi ai dati prodotti entro tre mesi dal completamento delle attività o dal completamento dei singoli moduli in cui sono strutturate, anche se non previste specificamente come deliverable dell'attività.

2.6 Governance della data policy

Lo *Steering Committee* di Marine Hazard verificherà l'applicazione delle regole della data policy contenute nel presente accordo, il loro rispetto e la loro piena attuazione. Tale organismo informerà tutti partecipanti ed il coordinamento del progetto di ogni eventuale problema o ostacolo alla corretta applicazione della data policy, incluse le difficoltà che potrebbero richiedere precisazioni, rettifiche e/o suoi aggiornamenti.

Lo *Steering Committee* potrà consultare rappresentanti della comunità scientifica marina relativamente ad aspetti specifici o casi particolari di applicazione della data policy.

2.7 Entrata in vigore, durata e termine degli accordi sulla data policy

La data policy illustrata nel presente documento entra in vigore all'atto della firma da parte di Partner del progetto ed ha, prelieve eventuali precisazioni, valore retroattivo dalla data di avvio del Progetto. La data policy si estende solamente a dati e prodotti ottenuti con le risorse del Progetto.

Considerando i termini sopra illustrati, i vincoli imposti dalla data policy si estendono indefinitamente oltre il termine del programma per tutti i dati e prodotti generati con le risorse determinanti di Marine Hazard. In questo contesto, il termine "determinanti" si riferisce al salario, compenso, borsa al generatore dei dati, il supporto parziale o totale all'acquisto e gestione della strumentazione ed alla logistica, il supporto alle spese di funzionamento, il supporto per eventuali consulenze.

2.8 Norma Transitoria

Il Coordinamento del programma, prima dell'approvazione definitiva, stabilisce un tempo congruo di monitoraggio del processo di governance (punto 2.6), e verifica di efficacia ed efficienza delle regole stabilite per garantire l'attuazione della data policy. Nondimeno, persistendo oltre le difficoltà della sua corretta applicazione, si potrà eventualmente agire per la sottrazione delle risorse determinanti impiegate, qualora il Comitato dovesse giudicare insufficiente l'attuazione della data policy.

2.9 Attività prodromica all'accordo sulla data policy

Ciascuna unità operativa si impegna a classificare i prodotti che sono stati o che verranno generati o messi a disposizione nel corso del Programma, in relazione alla tipologia, alla proprietà e al generatore dei dati, e ad indicare se prevede di applicare regole meno restrittive di quelle descritte nella sezione 2.4, ovvero di ridurre la durata del periodo di embargo o di eliminare il vincolo di coinvolgimento nell'analisi. Ciascuna unità operativa si impegna inoltre a precisare il tipo di Licenza che verrà associata ai dati/prodotti allo scadere del periodo di embargo. Tutte queste informazioni dovranno parte degli elementi essenziali dei Metadati.

2.10 Indicazione finale

Ogni dato/prodotto reso disponibile in Marine Hazard sia come Background sia come Foreground (v. Glossario) nel periodo di embargo a partire dalla sua generazione, deve contenere una descrizione delle regole d'uso, secondo lo schema previsto dal presente Accordo, o del tipo di Licenza ad esso associato (<http://opendefinition.org/licenses/>). Allo scadere dei due anni l'associazione della Licenza tra quelle individuate è obbligatoria per i dati di Foreground, e va incoraggiata la loro applicazione, previo accordo con i generatori originali, anche ai dati di background.

3. Glossario

- Background.** Dati in possesso di partecipanti al programma che sono stati prodotti prima dell'inizio del programma o in parallelo alle attività del programma, ma senza il suo supporto, che possono o devono essere utilizzati per le sue finalità.
- Dati ancillari.** Tutti i dati necessari ad ottenere dai dati grezzi prodotti da osservazioni, specie strumentali, le grandezze fisiche di interesse nelle loro unità di misura, e.g., dati per la calibrazione, funzioni per la conversione, dati di ground truth, etc.
- Dati grezzi.** Dati acquisiti che non hanno subito alcuna modifica, manipolazione, correzione o aggregazione (i dati grezzi devono includere i dati ancillari (v. sopra) per renderne possibile l'utilizzo).
- Dati pre-elaborati.** Dati modificati, manipolati, corretti o aggregati al puro fine di migliorarne la qualità.
- Dati sensibili.** Dati che forniscono informazioni legate alla sicurezza nazionale, a rischi di varia natura per la popolazione ed il territorio, che richiedono un processo di valutazione e di decisione a carico delle Autorità Pubbliche prima della loro divulgazione o, infine, dati personali.
- Dati.** Qualunque record, o insiemi di record, in formato testuale, numerico, acustico, d'immagine, o richiedente l'uso di sistemi informatici per essere utilizzato, come ad esempio spettri, sequenze polinucleotidiche, etc.; dati provenienti da microscopi elettronici o con focali, etc.; dati derivati da osservazioni satellitari, geospaziali, etc. utilizzabili come base di partenza per l'analisi e la ricostruzione di processi dell'atmosfera, dell'oceano e delle attività umane ad essi connesse e che, in definitiva, sono gli elementi necessari per comprovare i risultati della ricerca. Un insieme di dati costituisce una sistematica, se pure parziale, rappresentazione del sistema investigato (16).
- Foreground.** Dati o prodotti in possesso di partecipanti al programma che derivano direttamente dalle attività del programma e sono quindi prodotti con il suo supporto determinante.
- Generatore (di dati).** I soggetti, come singoli o organizzati in gruppi di lavoro, che conducono le osservazioni e/o gli esperimenti e generano i dati, i metadati e i dati ancillari e, direttamente o dopo processamento, li organizzano in record tali da poter essere utilizzati da terzi.
- Invito a partecipare.** Obbligo da parte del licenziatario di invito al licenziante a partecipare ad attività derivate dall'utilizzo dei dati rilasciati per un periodo limitato dalla disponibilità del dato stesso. Nel caso di dato sottoposto ad un periodo di embargo tale periodo inizia alla fine del periodo di embargo.
- Licenza.** Regole imposte dal proprietario o dal fornitore dei dati o dei prodotti per accedervi ed utilizzarli. La licenza viene concessa dal generatore o fornitore dei dati o prodotti e può prevedere anche il trasferimento libero dei dati e dei prodotti a terzi che non abbiano a loro volta ottenuto direttamente la licenza, può essere concessa senza costi o a pagamento, e può richiedere altre condizioni relativamente alle modalità d'uso dei dati e prodotti.
- Metadati.** Informazioni relative a dati o insiemi di dati che ne arricchiscono e completano la comprensione e conoscenza e che possono riguardare la loro tipologia, la geolocalizzazione, l'intervallo temporale a cui si riferiscono, le unità di misura in cui sono espressi, la loro accuratezza e precisione, il metodo utilizzato per la loro generazione, la struttura della loro organizzazione, la loro provenienza, la proprietà etc..
- Periodo di embargo.** Periodo di tempo a disposizione del generatore di dati o prodotti per procedere alla loro elaborazione prima di renderli pubblici e quindi condivisibili per la comunità.
- Prodotti.** Risultati dell'applicazione di un qualunque tipo di elaborazione, ovvero rappresentazioni di un sistema o di processi ad esso relativi ottenute attraverso il trattamento o l'interpretazione di dati anche combinandoli tra di loro guidati da un'elaborazione intellettuale. Essi sono quindi il risultato di un'attività che produce un valore aggiunto di conoscenza (ad es. previsioni di marea, simulazioni numeriche, metanalisi, etc.).
- Proprietario (di dati).** I soggetti che, o per essere generatori di dati senza altri vincoli, o per aver finanziato la generazione di dati con precisi vincoli contrattuali, ne detengono la proprietà e quindi possono decidere di limitare o impedire l'accesso e l'utilizzo dei dati stessi.

QC ed indici relativi. Informazioni relative alle procedure di calibrazione, validazione e quantificazione della qualità dei dati o dei prodotti in relazione a standard accettati dalla comunità scientifica, di solito forniti insieme agli altri Metadati, e valori di indici (flag) che quantificano la qualità ed affidabilità del dato, di solito associati ai dati stessi.

Appendice – Licenze

Considerate le definizioni delle tipologie di dati introdotte in precedenza, testi di riferimento sono il documento “Linee Guida per l’interoperabilità semantica attraverso i Linked Open Data” (http://www.digitpa.gov.it/sites/default/files/allegati_tec/CdC-SPC-GdL6-InteroperabilitaSemOpenData_v2.0_0.pdf) (14) redatto nell’ambito dell’attività dell’Agenzia per l’Italia Digitale nella definizione dei servizi di governance per l’interoperabilità semantica ed il documento “Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico (secondo semestre 2013)” (http://www.digitpa.gov.it/sites/default/files/allegati_tec/LG_Val_PSI_v1.0.pdf) (20). I servizi sono uno degli obiettivi delle infrastrutture condivise del Sistema Pubblico di Connettività (SPC) (<http://www.digitpa.gov.it/spc>), come definite dal Codice dell’Amministrazione Digitale (CAD).

Sulla base di tali documenti, le licenze previste per il progetto Marine Hazard sono definite da Creative Commons (CC) () e le licenze tipo IODL (Italian Open Data License) e OdbL (Open Data Commons Open Database License). CC è un’organizzazione non a fini di lucro che nasce con l’intenzione di armonizzare l’articolato mondo del diritto d’autore (in Italia regolato dalla legge n. 633 del 22 aprile 1941). Gli altri due tipi sono Licenze direttamente orientate alla condivisione di dati.

Nel seguito vengono riportate sinteticamente le operazioni possibili che ciascuna licenza concede, allo scopo di fornire elementi di orientamento per la concessione di Licenze legate a prodotti e dati.

CREATIVE COMMONS ZERO (CC0): libera utilizzazione gratuita, anche per fini commerciali e con finalità di lucro.

Apponendo su un documento la dichiarazione CC0 si rinuncia a tutti i diritti su dati e prodotti, nella misura massima possibile prevista dalla legge. La Creative Commons Zero deve essere preceduta da una dichiarazione relativa alla provenienza del dato del seguente tenore:

“Il riutilizzo del/dei dato/i prodotto/i nell’ambito del progetto Marine Hazard è concesso da <nome ENTE>, che autorizza, pertanto, la libera e gratuita consultazione, estrazione, riproduzione, modifica e riutilizzo da parte di chiunque vi abbia interesse per qualunque fine secondo i termini della Dichiarazione Creative Commons - CC0 1.0 Universal.”

CREATIVE COMMONS ATTRIBUZIONE (CC-BY): permette al soggetto utilizzatore di riprodurre, distribuire, comunicare, esporre, rappresentare, nonché di modificare e usare un insieme di dati anche a fini commerciali con il solo obbligo di attribuire la paternità del dato.

La Creative Commons Attribuzione deve essere preceduta da una dichiarazione relativa alla provenienza del dato e alla forma con cui attribuire la paternità del dato del seguente tenore:

“La titolarità piena ed esclusiva del/dei dato/i prodotto/i nell’ambito del progetto Marine Hazard [DENOMINAZIONE E DESCRIZIONE SINTETICA DEL DOCUMENTO] è di <nome ENTE>, che ne autorizza la libera e gratuita consultazione, estrazione, riproduzione e modifica da parte di chiunque (Licenziatario) vi abbia interesse per qualunque fine, purché nel rispetto dei termini della licenza Creative Commons. L’attribuzione prevista dalla licenza dovrà avvenire nella seguente forma: [INSERIRE NOTA PER L’ATTRIBUZIONE]”.

Riferimento: <http://creativecommons.org/licenses/by/4.0>

La licenza CC-BY può essere ulteriormente declinata nelle seguenti, che aggiungono altre condizioni da rispettare, oltre all’attribuzione:

LICENZA CREATIVE COMMONS ATTRIBUZIONE Share Alike (CC-BY-SA): obbliga i dati derivati a essere licenziati con la stessa licenza del dato originale

Riferimento: <http://creativecommons.org/licenses/by-sa/4.0>

LICENZA CREATIVE COMMONS ATTRIBUZIONE Non Commercial (CC-BY-NC): consente la copia, la distribuzione e l’uso del dato solo per scopi non commerciali

Riferimento: <http://creativecommons.org/licenses/by-nc/4.0>

LICENZA CREATIVE COMMONS ATTRIBUZIONE Non-Commercial SHARE ALIKE (CC-BY-NC-SA): consente la copia, la distribuzione e l'uso del dato solo per scopi non commerciali, con l'obbligo di licenziare i dati con la stessa licenza del dato originale

Riferimento: <http://creativecommons.org/licenses/by-nc-sa/4.0>

LICENZA CREATIVE COMMONS ATTRIBUZIONE Non-Commercial-NoDerivative (CC-BY-NC-ND): consente la copia, la distribuzione e l'uso del dato solo per scopi non commerciali, ma non consente la distribuzione di prodotti derivati

Riferimento: <http://creativecommons.org/licenses/by-nc-nd/4.0>

LICENZA ITALIAN OPEN DATA 2.0 (IODL 2.0)

La IODL 2.0 prevede che l'utente possa liberamente:

- 1) consultare, estrarre, scaricare, copiare, pubblicare, distribuire e trasmettere le informazioni;
- 2) creare un lavoro derivato, per esempio attraverso la combinazione con altre informazioni (c.d. mashup), includendole in un prodotto o sviluppando un'applicazione informatica che le utilizzi come base dati.

In cambio, all'utente è chiesto solo di indicare la fonte delle informazioni e il nome del soggetto che fornisce il dato, includendo, se possibile, un link alla licenza.

Rispetto alla IODL 1.0, la Italian Open Data Licence 2.0 non prevede l'obbligo dell'utente di pubblicare e condividere gli eventuali lavori derivati (ad esempio, applicazioni) con la stessa licenza o con altra licenza aperta, ritenuta compatibile.

Riferimento: <http://www.dati.gov.it/iodl/2.0/>

LICENZA OPEN DATA COMMONS OPEN DATA BASE 1.0 (ODbL 1.0)

La OdbL 1.0 prevede che l'utente possa liberamente:

- 1) consultare, estrarre, scaricare, copiare, pubblicare, distribuire e trasmettere le informazioni;
- 2) creare un lavoro derivato, per esempio attraverso la combinazione con altre informazioni.

All'utente è chiesto solo di citare il nome del Data base di origine e di estendere gli stessi termini di Licenza alla nuova elaborazione.

Riferimenti

Licenze sull'utilizzo di dati e prodotti: <http://opendefinition.org/licenses/>

Ribera l'Alcalà, M., A. Basoni, P. Carrara, R. d'Adamo, G. Manzella, S. Menegon, G. Paziienza, A. Rampini, e A Sarretta, Documento per la definizione di una politica nella gestione e utilizzo dei dati e dei prodotti resi disponibili nell'ambito del progetto RITMARE, 2014; <http://www.ritmare.it/area-download?download=187:data-policy-new>