

PID Group INGV: a bridge towards Open Data

**Agata Sangianantoni,
Adriano Azzarone,
Mario Locati
e il PID Group**

Ponte sul fiume Sile (2001), Treviso, arch. Paolo Portoghesi

Premessa

Nel 2005 è stato creato “Earth-Prints” (archivio istituzionale), dal 2010 “Annals of Geophysics” (rivista internazionale) pubblica solo ad accesso aperto, queste sono due **esperienze editoriali** ad accesso aperto INGV e alla base della firma del Position Statement del 21 marzo 2013.

L'importanza dei **dati** è cresciuta nell'ultimo decennio registrando un incremento esponenziale di:

- volume e sorgenti di dati
- tipologie di utilizzatori
- complessità delle richieste da soddisfare
- complessità delle elaborazioni necessarie
- quantità delle richieste di dati

Temi da affrontare

- Varietà nella tipologia di “*digital objects*”
 - Necessità di categorizzare i prodotti e servizi
- Predisposizione di una “*data policy*”
 - Principi ispiratori di riferimento
 - Conformità alla normativa vigente
 - Condivisione di una terminologia comune
 - Regole di accesso, utilizzo e relativa responsabilità
- Affidabilità scientifica
 - Tracciabilità del dato
 - Analisi dell'utilizzo dei dati
- Interoperabilità delle banche dati

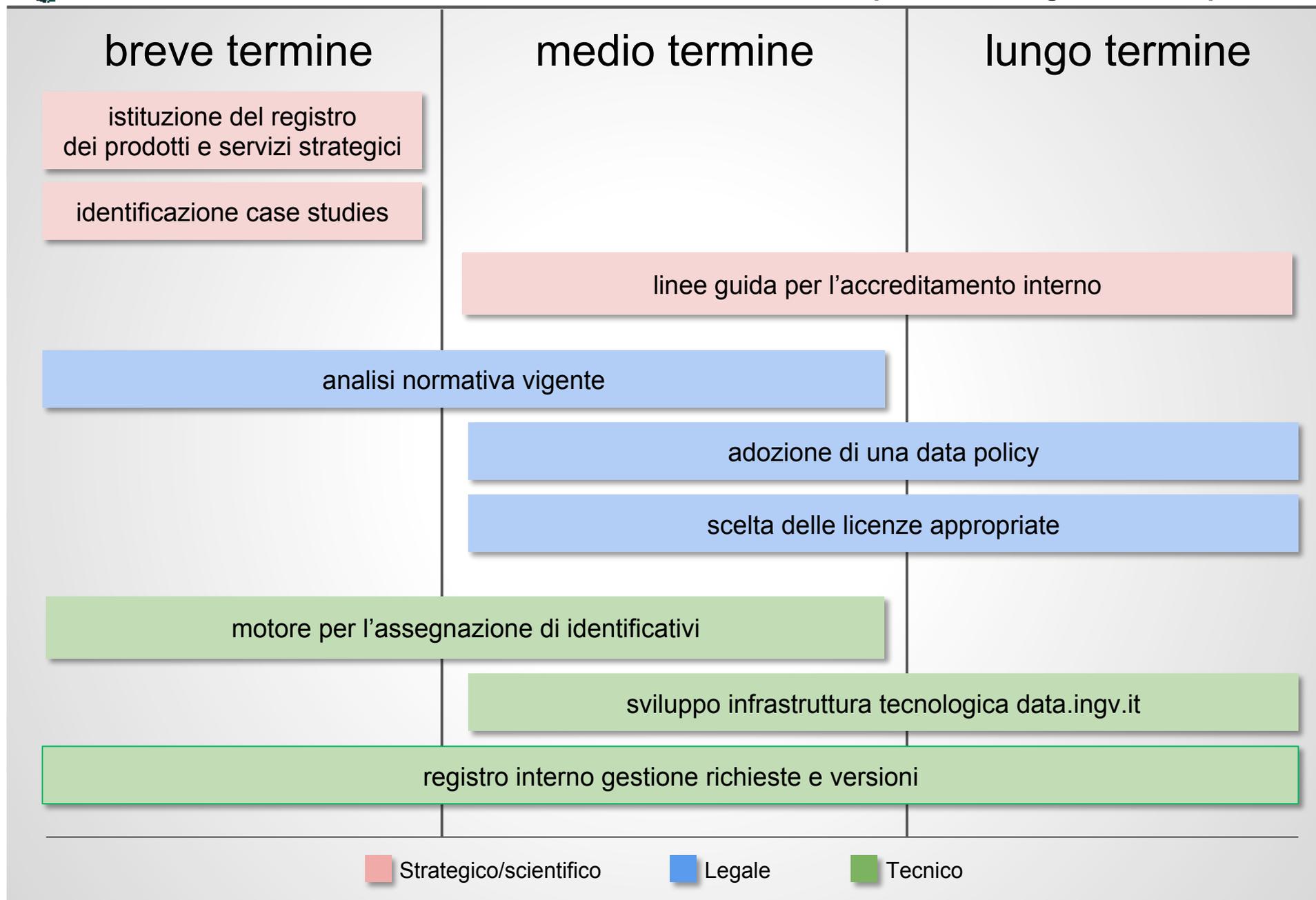


PID Group

L'INGV sta costituendo un gruppo di lavoro multidisciplinare per l'assegnazione di identificativi persistenti alle proprie banche dati.

Struttura del gruppo

- Strategico / Scientifico
- Legale
- Tecnico





Il contesto europeo e globale

Political framework (Europe 2020)

- Innovation Union
- Digital Agenda
- Raccomandazione sull'accesso all'informazione scientifica e sulla sua conservazione

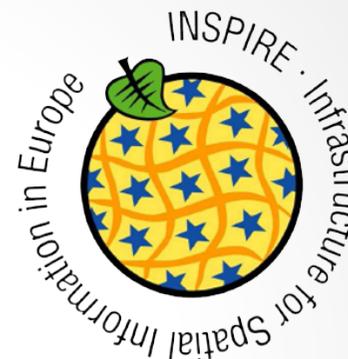
OECD Principles and Guidelines for Access to
Research Data from Public Funding

UNESCO Report to Scientific Information

Il contesto normativo europeo

Legal Framework

- PSI Directive *2003/98/CE*
- PSI amendments *2013/37/UE*
- Inspire Directive *2007/2/EC*
 - Inspire Metadata Regulation *EC 1205/2008*
 - Commission Regulation *(EU) n. 268/2010*
 - Commission Regulation *(EU) n. 102/2011*





Una licenza si applica a qualsiasi opera tutelata dal diritto d'autore.

Il presupposto fondamentale per poter concedere un'opera tramite Creative Commons è di avere la legittimazione giuridica.

Per cui:

- 1) occorre essere titolari dei diritti concessi con la licenza;
- 2) avere ottenuto una esplicita autorizzazione dal titolare dei diritti.

Linked Open Data

I Linked Open Data sono dati collegati rilasciati con licenza aperta, che ne permette il loro libero utilizzo.

Creative Commons CC-BY è un esempio di licenza aperta.

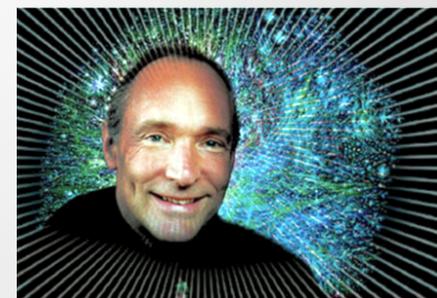
Linked Open Data è un movimento atto a rendere disponibili i dati esistenti in un comune formato “machine-readable”.

Grazie all’adozione di questi standard gli utenti finali possono combinare fra loro i dati favorendo la nascita di nuove interpretazioni.

A Web of Data

“If HTML and the World Wide Web made all the online documents look like one huge book, [Linked Data] will make all the data in the world look like one huge database”

Sir Tim Berners-Lee, W3C, Weaving the Web (1999).



Identificativi persistenti (PID)

Un identificativo persistente è un riferimento duraturo nel tempo verso una risorsa digitale, utilizzabile tramite servizi dedicati.

I PID dovrebbero essere assegnati solo alle risorse che saranno preservate per molto tempo, anche dopo varie generazioni hardware e software.

Utilizzati frequentemente per pubblicazioni e dataset.
I servizi sono fortemente basati su tecnologie web.

Vantaggi principali:

- Identificazione univoca stabile nel tempo di una risorsa digitale
- Maggiore interoperabilità “machine to machine”
- Citazione più facile per chi utilizza la risorsa
- Metadati associati arricchiscono le informazioni sulla risorsa

Diversi sistemi PID esistenti: Handle (DOI, EPIC), PURL, URN, ARK, ORCID,

.... **quale scegliamo?**

TUTTI! POSSIAMO INTEGRARE DIVERSI STANDARD!





Study Case INGV & DOI DataCite

INGV è accreditato presso CRUI per DOI (Digital Object Identifier) DataCite.

Una serie di metadati accompagna il DOI secondo specifica DataCite, arricchendo il PID di informazioni e migliorandone la interoperabilità.

Esempio DOI per un Dataset

DOI: 10.6092/INGV/BANCADATI/12345

CITAZIONE: Cognome, N. et al. (2013). Nome del dataset da un particolare strumento / modello / elaborazione. doi:10.6092/INGV/BANCADATI/12345

RISOLUZIONE: <http://dx.doi.org/10.6092/INGV/BANCADATI/12345>
=> <http://bancadati.ingv.it/...../12345>

Che fare se il DOI è troppo lungo?

ShortDOI.org rende più corto un DOI, pur rimanendo risolvibile come un normale DOI. Esempio: 10/pdh

I dati gestiti da INGV

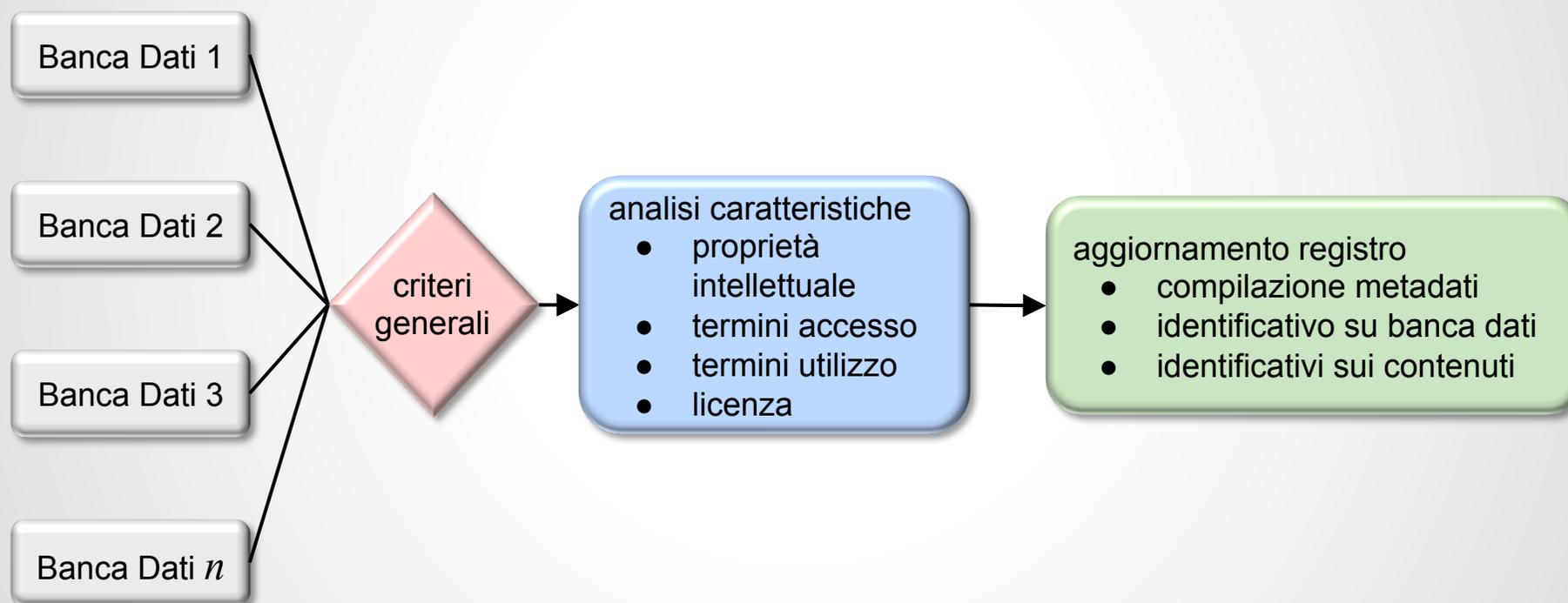
2) Banche dati (database)

Nome	Struttura di riferimento	Settore di riferimento	Linee di attività	Scalabilità finanziaria	Proiezioni ESFR	Strategie delle attività
Banca dati geologica del territorio						Banca dati geologica del territorio
dati da strumenti a terra, satellitari e in mare						dati da strumenti a terra, satellitari e in mare
flussi di dati in tempo reale						flussi di dati in tempo reale
dati elaborati ad aggiornamento periodico						dati elaborati ad aggiornamento periodico
dati da esperimenti di laboratorio						dati da esperimenti di laboratorio
dati storici dell'ultimo millennio						dati storici dell'ultimo millennio
dati sensibili per la Protezione Civile						dati sensibili per la Protezione Civile
banca dati delle pubblicazioni						banca dati delle pubblicazioni
dati da rilevamenti di campagna						dati da rilevamenti di campagna
dati finalizzati alla divulgazione						dati finalizzati alla divulgazione

Tabella delle banche dati gestite da INGV tratta dal Piano Triennale di Attività Scientifica 2013-2015

Una cinquantina di banche dati estremamente eterogenee, sia nella natura del dato, sia nei metodi e tempi di aggiornamento con grado di riservatezza variabile, realizzate in ambito locale, nazionale, europeo e mondiale.

Processo per l'accreditamento interno e assegnazione identificativi persistenti

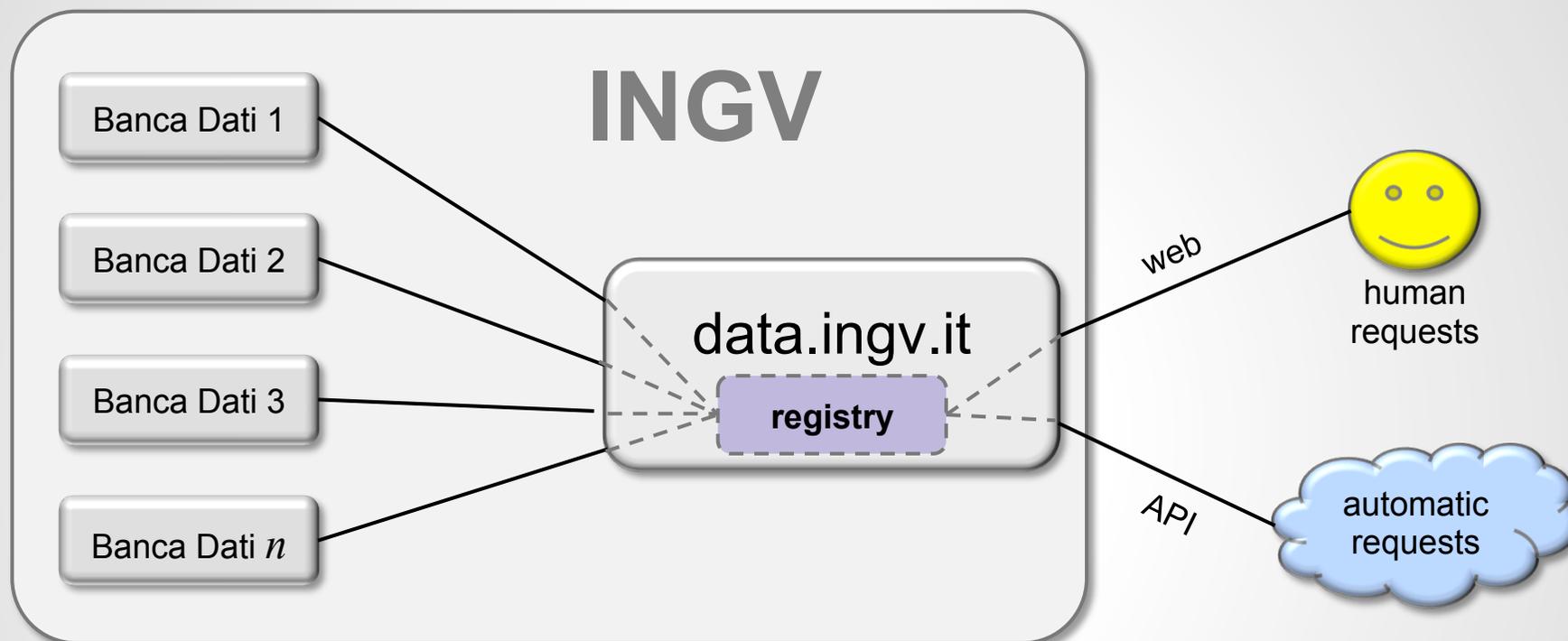


 Strategico/scientifico

 Legale

 Tecnico

Portale unico per l'accesso ai dati



Un portale unico per la ricerca e l'accesso ai dati prodotti.

Il portale non sostituirà le attuali banche dati, ma ne uniformerà i metadati permettendone l'aggregazione rendendoli disponibili a utenti e servizi web.

Apertura prevista: Q4 2014

Grazie!

pidgroup@ingv.it



Quest'opera è distribuita con Licenza
Creative Commons Attribuzione 3.0 Unported.