

# Un modello DAMA in stile *agile*: la proposta di Paesaggiaperti.org

Luca Corsato<sup>1</sup>, Alfonso Crisci<sup>2</sup>, Massimiliano Magli<sup>3</sup>, Teodoro Georgiadis<sup>4</sup>

<sup>1</sup> opensensorsdata srl (OSD), Via D. Manin 60, Venezia, Italia  
luca@osd.tools

<sup>2</sup> Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via G. Caproni 8, Firenze, Italia  
a.crisci@ibimet.cnr.it

<sup>3</sup> Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Piero Gobetti 101, Bologna ,  
Italia  
m.magli@ibimet.cnr.it

<sup>4</sup> Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Piero Gobetti 101, Bologna ,  
Italia  
t.georgiadis@ibimet.cnr.it

## Sommario

Con l'avvento degli strumenti digitali, all'attuale prassi scientifica nei vari domini della conoscenza ha corrisposto una grande produzione di dati e/o di prodotti intermedi di ricerca. Questi dati sono beni, che possono diventare bene comune, e si ottengono dai normali processi di inferenza come le simulazioni modellistiche, le misurazioni da monitoraggi e da esperimenti. Tali dati sono la matrice su cui poi il ricercatore o il tecnologo finalizza il suo lavoro con una pubblicazione. Le dinamiche offerte dal rilascio di dati aperti mutuati dai progetti collaborativi aperti - generalmente considerati fuori dagli ambiti scientifici canonici che seguono il panorama delle licenze aperte (Creative Commons, GPL ed altre ancora) - producono un ulteriore grado di libertà alla produttività di dati digitali scientifici. Il problema del riconoscimento a chi produce i dati e l'attivazione al riuso di queste risorse è oggi oggetto di dibattito. Questa comunicazione vuole descrivere la proposta operativa per un impianto di Data Management Plan denominata "Paesaggiaperti.org" frutto di una partnership tra ricerca e impresa, che utilizza risorse allo stato dell'arte del mondo "open-source". Tale impianto è applicato al vasto dominio di conoscenza afferente al paesaggio, permettendone il rafforzamento dell'informazione qualificata come punto di resilienza per fronteggiare i recenti cambiamenti antropici e climatici.

## 1 Premessa

La ricerca scientifica si basa sulla formulazione di un'idea a cui segue la raccolta di dati [1]. Il dato è un bene che aggregandosi costituisce una base informativa. Dalla base informativa il processo deduttivo deriva delle tesi che entrano nel dibattito scientifico sotto forma di articolo. Paesaggiaperti.org è un modello per aggregare i dati di ricerca relativi ai temi del paesaggio, della conoscenza territoriale e dei parametri che li descrivono distribuendoli come pacchetti informativi applicabili a molteplici contesti.

## 2 Impianto del progetto Paesaggiaperti.org

Il progetto Paesaggiaperti.org nasce da una convenzione non onerosa stipulata tra l'Istituto di Biometeorologia Ibimet CNR e opensensorsdata (OSD) srl che ne sono proponenti unici. Lo scopo è di esplorare e analizzare le soluzioni ottimali per il rilascio, la valorizzazione tramite

citazione, e l'accesso con licenze aperte di dati territoriali utili per l'analisi di paesaggio, come ad esempio quelli relativi alle tematiche di pianificazione territoriale e ai monitoraggi climatologici/meteorologici urbani e rurali. La collaborazione è stata focalizzata su un impianto di Data Management Plan [2] che consenta al soggetto pubblico (es: IBIMET CNR) di poter valorizzare i propri dati e distribuirli, con costi e tempi molto ridotti, e al soggetto privato (es: opensensorsdata OSD) di poter accedere in autonomia ad una disponibilità di dati potendone valutare il potenziale valore per i propri prodotti e/o servizi. Questo impianto consente un'immediata collaborazione tra ricerca e impresa perché permette all'istituto di valorizzare i ricercatori, attraverso le competenze rappresentate dai dataset distribuiti. Contestualmente l'impresa ha una immediata visione del capitale scientifico, potendo riferirsi direttamente al ricercatore competente facendosi accompagnare nelle applicazioni.

### 3 Elementi di peculiarità del Data Management Plan di Paesaggiaperti.org

Paesaggiaperti.org è un'applicazione web a cui si accede su invito. Quattro elementi peculiari costituiscono il Data Management Plan:

1. un open business model per la sostenibilità dell'intero progetto
2. un'applicazione web riusabile che distribuisce e aggrega i dataset con metodologia *agile*
3. la certificazione e tracciabilità con l'attribuzione di DOI ai dataset
4. uno schema di licenze di distribuzione dei dati

#### 3.1 Open Business Model

L'open business model [3, 4, 5, 6] serve a definire il valore generato dalla distribuzione di capitale informativo in termini di benefici sociali ed economici per soggetti esterni all'ambito di origine della ricerca, siano essi istituzionali, pubblici e/o privati.

#### 3.2 Software Git

L'applicazione web - sviluppata da una comunità esterna e completamente disponibile senza costi e aperta - è basata su Git [7, 8], che provvede a renderla utilizzabile senza competenze di programmazione. Git è un software libero scritto da Linus Torvalds [9] dotato di un controllo di versione (*versioning*) crittografato, e altri strumenti utili per la derivazione di dataset (*forking*) e altri per la confluenza di altri dataset in uno unico (*merging*). Questa dinamica consente a Paesaggiaperti.org di tracciare ogni attività, senza perdere né l'*authoring* (con l'attribuzione automatica di una chiave unica *hash* per ogni corpus di edit) e nemmeno la sua tipologia (*commit*), dando valore ad ogni attività, sia essa creativa che di controllo come la correzione degli errori.

#### 3.3 Certificazione e tracciabilità con l'attribuzione di DOI ai dataset

La relazione tra chi distribuisce un dataset e un potenziale utilizzatore si stabilisce più facilmente tanto più il dataset è distribuito in forma aperta, perché riduce la complessità legale ed

amministrativa. La scelta di una distribuzione aperta è definita dallo spostamento della produzione di valore sulla relazione con chi ne può beneficiare dalla distribuzione, dando la possibilità di riferirsi direttamente con l'autore (o l'*estrattore*) di quei dati per risolverne la complessità. Per questo la tracciabilità di un dataset è garanzia per chi ha necessità di riutilizzo, e per chi ha distribuito genera valore in termini di citazione e attrazione di fondi. La tracciabilità e la citazione sono date da un identificativo digitale unico (DOI) rilasciato da Datacite [10], un'organizzazione no-profit focalizzata sull'attribuzione di DOI a dataset. I DOI saranno messi a disposizione di chiunque distribuirà dataset attraverso Paesaggiaperti.org, per recuperare basi informative che altrimenti resterebbero negli archivi, catalogandole e rendendole disponibili per ulteriori ricerche in qualsiasi ambito (accademico, imprenditoriale, sociale).

### 3.4 Schema di licenze di distribuzione

L'attribuzione della licenza di distribuzione è una questione oggetto di dibattito recente. Paesaggiaperti.org si propone di armonizzare la maggiore distribuzione possibile con la massima tutela e preservazione del capitale informativo della ricerca e dei ricercatori. Questo significa proporre, a chiunque vorrà distribuire i dati, una serie di licenze il cui scopo non il "condividere a tutti i costi" ma una valutazione tra impatto e benefici derivati dalla distribuzione ma sempre con il beneficio del riconoscimento personale di autore. Per questo l'open business model sarà modellato sulle istanze che hanno prodotto i dati e le fonti da cui provengono, bilanciandole con i costi, i benefici sociali e la capacità di attrarre ulteriori fonti di finanziamento, proprio grazie alla distribuzione del capitale informativo.

## 4 Conclusioni

Il progetto è in itinere e si focalizza principalmente su piccoli e medi dataset e dalla presentazione entrerà ufficialmente nella fase di sperimentazione per poterne monitorare limiti e potenzialità.

## Riferimenti bibliografici

- [1] William K. Michener. Ten simple rules for creating a good data management plan. *PLOS Computational Biology*, 11(10):1–9, 10 2015.
- [2] DAMA International. *The DAMA Dictionary of Data Management, 2Nd Edition: Over 2,000 Terms Defined for IT and Business Professionals*. Technics Publications, LLC, USA, 2nd edition, 2011.
- [3] Henry Chesbrough. Business model innovation: opportunities and barriers. *Long range planning*, 43(2):354–363, 2010.
- [4] Henry Chesbrough. Business model innovation: it's not just about technology anymore. *Strategy & leadership*, 35(6):12–17, 2007.
- [5] Henry William Chesbrough. *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press, 2006.
- [6] Paul Stacey. What is an open business model and how can you generate revenue? [medium.com/made-with-creative-commons/what-is-an-open-business-model-and-how-can-you-generate-revenue-5854d2659b151](https://medium.com/made-with-creative-commons/what-is-an-open-business-model-and-how-can-you-generate-revenue-5854d2659b151), 2016 (accessed May, 2017).
- [7] Scott Chacon and Ben Straub. *Pro Git*. Apress, Berkely, CA, USA, 2nd edition, 2014.
- [8] Yu Wu, Jessica Kropczynski, Patrick C Shih, and John M Carroll. *Exploring the ecosystem of software developers on GitHub and other platforms*. 2014.

- [9] Git. A short history of git. [git-scm.com/book/en/v2/Getting-Started-A-Short-History-of-Git](https://git-scm.com/book/en/v2/Getting-Started-A-Short-History-of-Git), 2017 (accessed September, 2017).
- [10] DataCite. DataCite mission. [www.datacite.org/mission.html](http://www.datacite.org/mission.html), 2016 (accessed September 14, 2017).