



Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

Accesso aperto e riutilizzo dei dati della ricerca scientifica: spunti dal mondo Open Access e Open (Government) Data e nuove sfide

Federico Morando

Conferenza ASITA 2014, 16 Ottobre 2014

Firenze

Open Definition (2.0)

- *“Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness.”* (<http://opendefinition.org/od/>)
 - Open License
 - pubblico dominio (es. CC0)
 - attribuzione (es. CC BY)
 - share-alike (es. CC BY-SA)
 - Open Access
 - ai costi di riproduzione o gratis
 - Open Format
 - machine-readable
 - specifiche liberamente disponibili

Open Access

- movimento inizialmente noto anche come “Free Online Scholarship”
 - “libero”, non solo “aperto”, non solo “accesso”
- Budapest Open Access Initiative:
 - *“free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or **use them for any other lawful purpose, without financial, legal, or technical barriers** other than those inseparable from gaining access to the internet itself”*
 - limiti ammessi per integrità e citazione appropriata

scienza → Open Source

- ma il “full text” non è la “full story”
- Ince, D. C., Hatton, L., & Graham-Cumming, J. (2012). *The case for open computer programs*. *Nature*, 482(7386), 485-488:
 - “evidence cannot be entirely included in publications”
 - “it is now accepted that data should be made available on request”
 - “the rise of computational science has added a new layer of inaccessibility”
 - “the current regulations regarding the availability of software are inconsistent”
 - “with some exceptions, anything less than the release of source programs is intolerable for results that depend on computation”

accesso, codice, dati (linked) → riproducibilità

- Peng, R. D. (2011). *Reproducible research in computational science*. *Science* (New York, Ny), 334(6060), 1226.

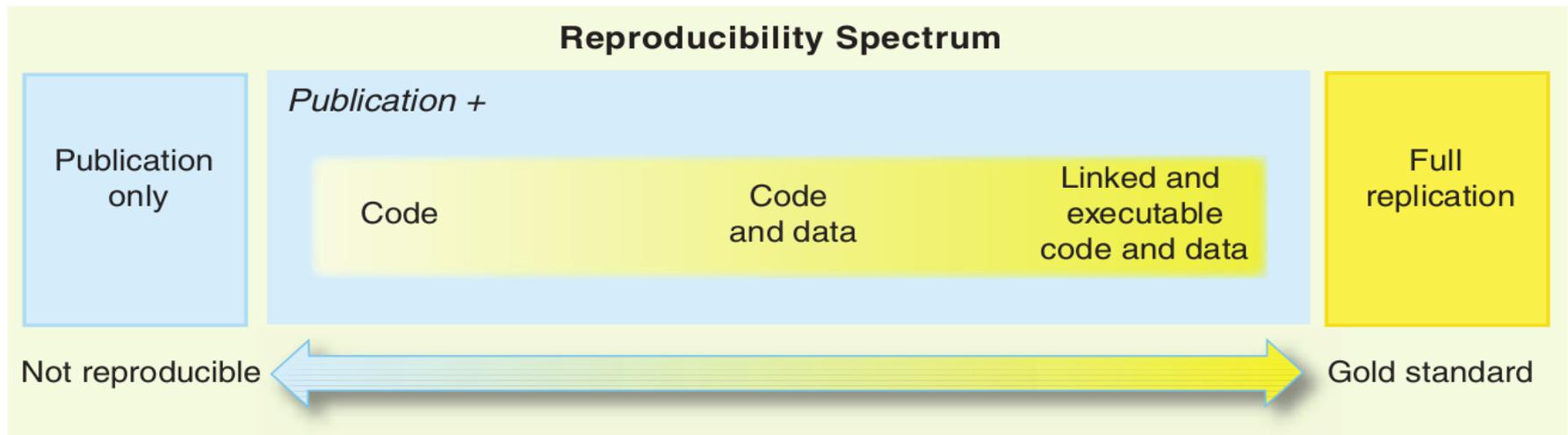
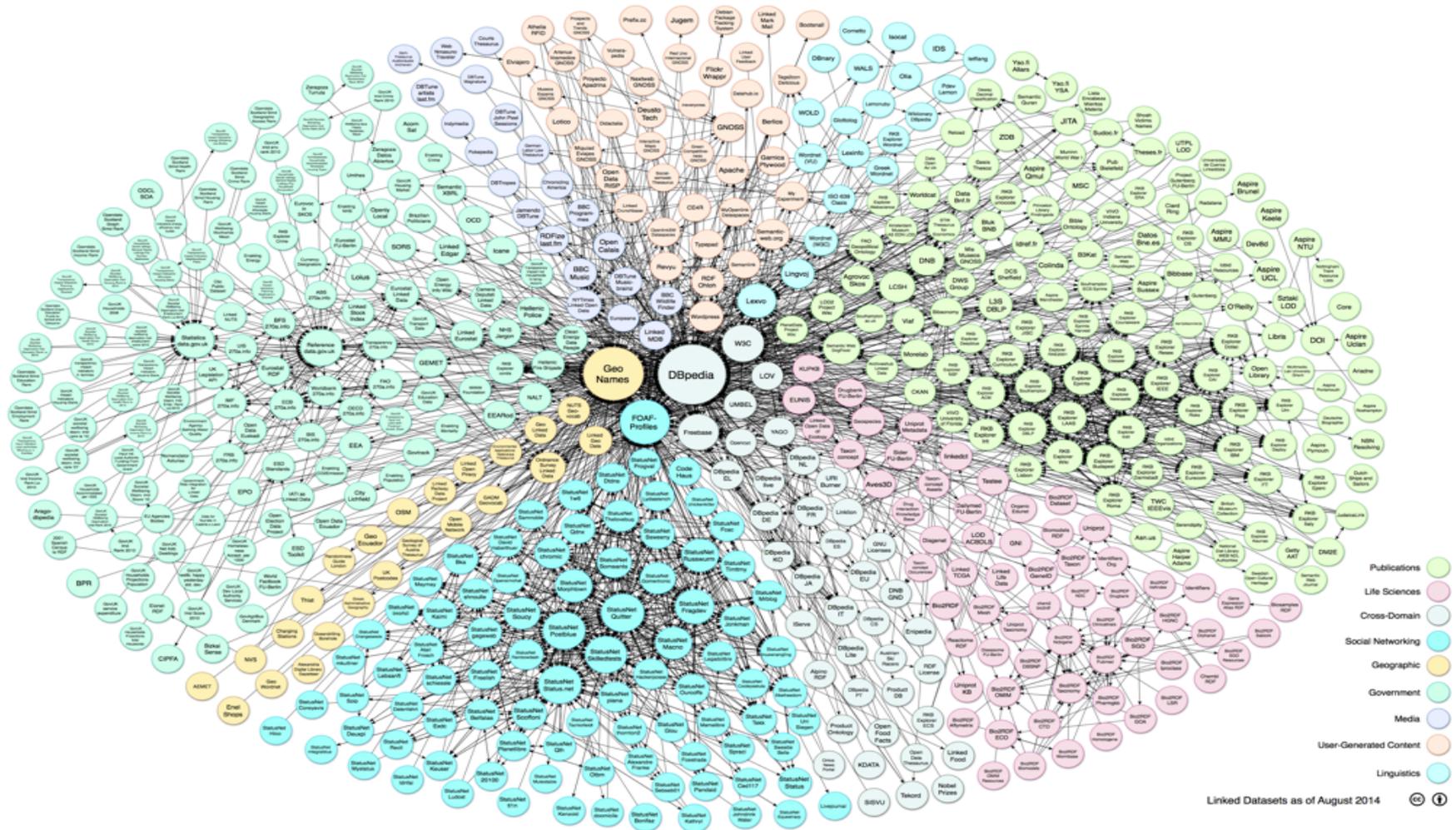


Fig. 1. The spectrum of reproducibility.

metadati come Linked Open Data



ostacoli

- incentivi accademici
 - sistema delle riviste/citazioni premia il paper, non il dataset
- elementi culturali
 - contributo allo stato dell'arte visto soprattutto in ottica di originalità, meno di “sporco lavoro”
- mancanza di consapevolezza (ma cresce) tra finanziatori della ricerca
- questioni giuridiche
 - in fondo non a caso (anche se mi pagano da vivere ;-))

lezioni dal mondo Open Data

- una lezione dagli antipodi: Laurence Millar (NZ)
 - con 3/4 di questi fattori si può avere successo:
 - leadership
 - anche/principalmente politica
 - crisi
 - crisi finanziaria; tagli di budget
 - eroi (nel governo/amministrazione)
 - gente che scommetta e ci metta la faccia
 - pressione da parte dei cittadini
- anche per i dati scientifici, si può procedere con un misto di leadership istituzionale, cultura tra i ricercatori, e uso in positivo dei budget sempre in diminuzione

in pratica

- per molte discipline, siamo ancora (speriamo per poco) nella fase in cui cominciare è già encomiabile
- e basta poco per cominciare
 - anche condividere i dati solo dopo averli “spremuti” è meglio di niente
 - più utili nelle mani di studenti di master e tesisti che a “prender polvere” in un hardisk che presto o tardi si romperà ;-)
- in caso I dubbi siano giuridici, si possono trovare interlocutori
 - es. <http://selili.polito.it/>
- inoltre...

c'è chi può aiutare su OA (formazione/policy)

<http://www.fosteropenscience.eu>

<http://www.pasteur4oa.eu>



si può fare

- esempi in questo stesso panel
- esempi in Italia
 - Consiglio Nazionale delle Ricerche
 - <http://data.cnr.it/>
- esempi molti strutturati e con policy articolate all'estero
 - Australian National Data Service
 - <http://www.ands.org.au/>

e ogni tanto si deve fare (1/3)

- Horizon2020
 - Pilot on Open Research Data
 - progetti interessati per 3 miliardi di Euro nel 2014-2015
 - “tutti” i dati necessari a validare i risultati delle pubblicazioni
 - per altri ricercatori, aziende, cittadini
 - il “pilota” OA di FP7 si è ora esteso a tutto H2020
 - tutte le pubblicazioni peer-reviewed in OA
 - (se il pilota avrà ragionevole successo, questa è la strada anche per i dati)

e ogni tanto si deve fare (2/3)

- Decreto Valore Cultura (D.L. 8 agosto 2013, n. 91, convertito con L. 7 ottobre 2013, n. 112), Art. 4.2:
 - “Le pubblicazioni che documentano i risultati di ricerche **finanziate per una quota pari o superiore al cinquanta per cento con fondi pubblici** [...] devono essere depositate, non oltre sei mesi dalla pubblicazione, in archivi elettronici istituzionali o di settore, predisposti in modo tale da garantire l'accesso aperto, libero e gratuito”

e ogni tanto si deve fare (3/3)

- sperando che prima o poi i bandi SIR siano valutati ;-)
- mandato OA del Bando SIR (Decreto Dir. 23 gennaio 2014 n. 197, art. 9)
 - "garantire l'accesso aperto (accesso gratuito on-line per qualsiasi utente) a tutte le pubblicazioni scientifiche 'peer-reviewed'"
 - "depositare i dati necessari per validare i risultati presentati nelle pubblicazioni"
 - eccezioni per dati personali, ecc.
 - serve nota "che espliciti i motivi alla base della mancata messa a disposizione di parti dei dati della ricerca"

e spesso conviene (o potrebbe)

- Belter CW (2014) Measuring the Value of Research Data: A Citation Analysis of Oceanographic Data Sets. PLoS ONE 9(3): e92590. doi:10.1371/journal.pone.0092590
 - “demonstrate the bibliometric impact of properly curated and openly accessible data set”
 - citation counts in most cases higher than 99% of all the journal articles published in Oceanography during the same years
 - BUT “methods of citing and referring to these data sets in scientific publications are highly inconsistent, despite the fact that a formal citation format is suggested for each data set”

Conclusione

- per risultati su ampia scala servono le istituzioni
 - “such an effort would probably need government coordination and support” (Peng, 2011)
 - gli enti finanziatori (anche non pubblici) hanno il loro ruolo
- ma...
 - le istituzioni sono fatte di persone
 - ognuno dei presenti può far la differenza nella sua!
 - dentro aree disciplinari in cui il dato è importante, si possono creare “bolle” in cui la condivisione dei dati è normale (e reciproca)
 - val la pena tentare (e si fanno esperienze utili per Horizon2020)

grazie

federico.morando@polito.it

Si ringrazia il progetto PASTEUR4OA per il supporto:
Open Access Policy Alignment Strategies for European Union Research funded by the
Seventh Framework Programme for Research and Development (FP7)

Questa presentazione può circolare liberamente ed è riutilizzabile ai sensi della licenza:
Creative Commons Attribution 4.0 International license

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