

# Forme e gradi di apertura dei dati

ANTONELLA DE ROBBIO

Centro di ateneo per le biblioteche  
Università degli studi di Padova  
antonella.derobbio@unipd.it

---

## I nuovi alfabeti dell'Open Biblio tra scienza e società\*

---

*Data is the new oil?  
No. Data is the new soil  
(David McCandless)*

### Dati aperti tra scienza e società

I dati non sono soltanto il nuovo olio della conoscenza, ma sono come un nuovo terreno fertile da coltivare. Di Open Data si sta parlando da qualche tempo:<sup>1</sup> è indubbio che l'accesso al materiale informativo sia il motore propulsore per lo sviluppo delle attività umane, pratiche ed intellettuali. Restrizioni sui dati e sul loro riutilizzo limitano lo sviluppo delle comunità. I dati grezzi (*raw data*) prodotti dal mondo della ricerca sono un terreno fertile che crea ricchezza, non solo intellettuale, ma anche per il rilancio dell'economia. I dati pubblici sono realmente utili laddove essi sono grezzi, davvero aperti e connessi, in quanto il valore economico dei dati – esterno e interno – risiede nella loro apertura. Esistono diverse ragioni per le quali è necessario rendere i dati accessibili. In primo luogo perché i dati di interesse genetico, medico, ambientale, ma anche economico-sociale sono da considerarsi patrimonio dell'umanità e come tale devono essere disponibili non solo all'accesso, ma anche al loro riutilizzo. In secondo luogo perché la massa di dati finanziata dai contribuenti deve essere disponibile agli stessi. I dataset della pubblica amministrazione infatti sono stati creati grazie al sostegno della cittadinanza e al denaro pubblico che ne ha permesso la raccolta e la gestione.

Un rapporto del 2011 della società McKinsey<sup>2</sup> sui paesi dell'Unione Europea ha calcolato che gli Open Data potrebbero abbattere i costi della pubblica amministrazione del 20% creando valore fino a 300 miliardi di dollari annui tra riduzione di inefficienze, maggiori introiti fi-

scali e maggiore produttività. Seguendo la metodologia tipica dell'Open Government, molte istituzioni pubbliche stanno rilasciando dataset in formati aperti e riutilizzabili dalle comunità legate al mondo degli Open Data.<sup>3</sup> Gli esperti in materia si dividono nell'individuare le radici dell'Open Data entro il contesto dell'Open Science – laddove la comunicazione scientifica usa i dati grezzi per condividere dati – o piuttosto nel considerare l'Open Data come derivazione dell'Open Government, laddove le comunità di rete assumono nuove forme di socialità tra cittadinanza attiva e democrazia partecipata.

Tra questi due territori aperti, tra scienza e ambiente sociale, potremmo collocare un terzo luogo, che li collega entrambi e che riguarda la meta-informazione. Questa “terra di mezzo” viene definita Open Bibliographic Data o anche OpenBiblio e comprende i cataloghi delle biblioteche (OPAC), i database bibliografici, le basi di conoscenza, i tesauri e gli schemi di classificazione, gli archivi aperti e gli archivi di materiale didattico e-learning ecc. John Wilbanks<sup>4</sup> in un suo noto messaggio in lista afferma che buoni metadati, compatibili con i Principi Pantoni<sup>5</sup> – scritti per diffondere gli Open Data nel mondo e-science – sono attualmente la migliore medicina.

La scienza è basata sulla costruzione, il riuso e la possibilità di mettere in discussione apertamente il corpus di conoscenza scientifica già pubblicato. È essenziale che i dati scientifici siano resi aperti per fare in modo che la scienza sia più efficace e la società ottenga il massimo beneficio dalle ricerche scientifiche. Con dati aperti nella scienza intendiamo la possibilità per ogni utente di scaricare, copiare, analizzare, rielaborare, inviare a software o usare i dati per altri scopi gratuitamente senza barriere finanziarie, legali o tecniche se non quella dell'accesso alla stessa rete internet. A tal fine i dati relativi alla scienza già pubblicati dovrebbero essere esplicitamente resi di dominio pubblico.

Una definizione di principi che combacia perfettamente con il corpus delle definizioni di Open Access comunemente note come le BBB Definitions e corrispondenti alle tre dichiarazioni di Budapest (2002), Bethesda (2003) Berlino (2003). Nel 2007 l'European Science Foundation e la Deutsche Forschungsgemeinschaft, nell'ambito del Berlin5 tenutosi a Padova, organizzarono il workshop "Shared Responsibilities in Sharing Research Data", con lo scopo di aumentare la consapevolezza riguardo all'accesso aperto ai dati della ricerca e ad aprire un forum di discussione per le potenziali aree di cooperazione.

Ma ancora prima nell'ottobre 2005 sul sito di SPARC fu attivata una lista di discussione<sup>6</sup> dedicata agli Open Data, moderata da Peter Murray-Rust,<sup>7</sup> precursore degli OpenData già da quando mise a punto – nel 1995 – il *Chemical Markup Language* (CML) quale approccio utilizzato in chimica per la gestione delle informazioni riguardanti le molecole, usando XML e Java.

Nella scienza è fortemente raccomandato che i dati siano esplicitamente di dominio pubblico, soprattutto quando sono finanziati con denaro pubblico. I dati dovrebbero essere quindi rilasciati con una licenza di tipo *Public Domain Dedication Licence* (PDDL)<sup>8</sup> oppure *Creative Commons Zero Waiver*, o CC0.<sup>9</sup> Science Commons è il protocollo per l'implementazione dell'accesso aperto ai dati che indaga tra i meccanismi legali che governano le mere raccolte di dati online e individua i progetti di tecnologia per pubblicare le raccolte in modo da renderle più utili, questo perché nella scienza vi è la tendenza a parallelizzare la generazione di dati. In altri termini le macchine generano dati in grande quantità, mentre allo scienziato spetta interpretare i dati primari e confezionarli in informazione. Sono proprio queste interpretazioni che aggiungono valore ai dati, laddove dall'analisi di dati grezzi si deduce per esempio il meccanismo di azione con cui agisce il cancro, o si identifica la funzione di un gene o il processo di isolamento di un virus.

L'etica di fondo insita nella scienza si basa sulla condivisione della conoscenza che dovrebbe prevedere il riuso dei dati all'interno della comunità scientifica, in quanto la condivisione delle scoperte sta alla base della scienza pubblica, intesa come bene comune, considerato che gran parte della ricerca scientifica deriva da finanziamenti pubblici. Ma come evidenzia John Wilbanks è curioso come "proprio nel momento storico in cui le tecnologie consentono la disponibilità a livello globale dei sistemi di distribuzione dei dati scientifici che ci offri-

rebbero la possibilità di ampliare la collaborazione e accelerare il ritmo e la profondità della scoperte... siamo occupati e bloccare i dati e a prevenire l'uso di tecnologie avanzate che avrebbero un forte impatto sulla diffusione della conoscenza".

Sussistono ancora aree di forte resistenza laddove "un sacco di dati non possono essere 'aperti' per definizione, o per 'scopi commerciali', o per motivi di 'privacy', o a causa di restrizioni ereditate in precedenza, dati raccolti prima che il movimento Open Data si affermasse o che la gente ne prendesse consapevolezza". O ancora per motivi di pubblica sicurezza, come nel caso degli studi condotti sul virus H5N1 modificato, volti a studiarne alcune caratteristiche collegate alla sua trasmissibilità.<sup>10</sup> A seguito del suggerimento dato dalla commissione scientifica nazionale americana di biosicurezza<sup>11</sup> (NSABB) di non pubblicare i dati aperti delle ricerche H5N1, si è scatenato un acceso dibattito attorno alle questioni correlate alla diffusione di informazioni scientifiche che potrebbero essere usate a fini bioterroristici. La comunità scientifica ha percepito tali raccomandazioni della commissione NSABB come una censura inutile e il problema è stato attentamente valutato sia dagli scienziati sia dai direttori delle riviste che a malincuore hanno seguito le indicazioni della NSABB.

Quello che però si mina non pubblicando i dati chiave – hanno affermato gli studiosi che hanno comunque chiesto al NSABB di fornire un sistema interno controllato dove mettere comunque a disposizione i dati per le comunità di ricerca – è proprio l'effettiva efficacia di queste ricerche, che hanno bisogno del contributo di altri gruppi, nonché dell'esigenza di ripetere gli esperimenti per migliorare la salute e la sicurezza pubblica, studiando quindi sistemi di monitoraggio e prevenzione di eventuali epidemie.<sup>12</sup>

La Dichiarazione di Ghent,<sup>13</sup> tracciata dai revisori del progetto OpenAIRE,<sup>14</sup> intende favorire iniziative che, muovendosi dall'accesso aperto alla ricerca e alla conoscenza, vadano in direzione della creazione e dell'utilizzo di dati, software e risorse educative aperti, quattro iniziative che non solo promuovono uno sviluppo nell'ambito della ricerca e in quello educativo, ma ne riducono anche i costi. "Con simili strumenti, la potenziale rete sociale europea della conoscenza potrà fare i primi passi, creando innanzitutto i collegamenti tra i documenti, per ampliarli in seguito a nuovi legami di collaborazione tra individui, laboratori e istituzioni".

## Open Data al centro delle comunità Open

Di Open Data si sta parlando ovunque. *The Open Data Handbook*<sup>15</sup> è un semplicissimo e utilissimo manuale messo a punto dalla *Open Knowledge Foundation* rivolto a quanti sono interessati a conoscere meglio il fenomeno Open Data e a quelli che si considerano professionisti dei dati, a chi si occupa di erogare servizi pubblici, agli attivisti, giornalisti e ricercatori, politici e sviluppatori di politiche a vario titolo, *data geek* e a tutti quelli non hanno mai sentito parlare di API (Application Programming Interfaces).

Al pari di altre filosofie “open”, l’Open Data si propone l’obiettivo di rendere immediatamente accessibili a chiunque i dati primari della ricerca, senza limitazione di copyright, brevetti o altri meccanismi di controllo. Su Twitter esiste un *thread* molto vivace e attento alle notizie più innovative.

Luca De Biase,<sup>16</sup> a proposito delle comunità di rete, individua un settore in fermento, quello delle *smart communities*, spazi online dove i cittadini generano un filone di innovazioni possibili che vanno dalla riorganizzazione delle città e dei territori al miglioramento dell’efficienza nella gestione dei beni comuni della conoscenza, dall’inclusione sociale all’alfabetizzazione funzionale e tecnologica. In queste dimensioni sorgono progetti di “consumo collaborativo” come “Bicincittà”,<sup>17</sup> una rete per il *bikesharing*, o orientati alla sanità sul modello di “Pazienti.org”,<sup>18</sup> spazio online dove i cittadini possono valutare le strutture sanitarie e scambiarsi consigli in una community orizzontale, o punti di riferimento per la formazione come “Wikibooks”<sup>19</sup> che oltre ai manuali riunisce libri in pubblico dominio, come i classici della letteratura, o ancora “Openpolis”<sup>20</sup> che raccoglie in modo collaborativo le informazioni sulla politica e sulla pubblica amministrazione.

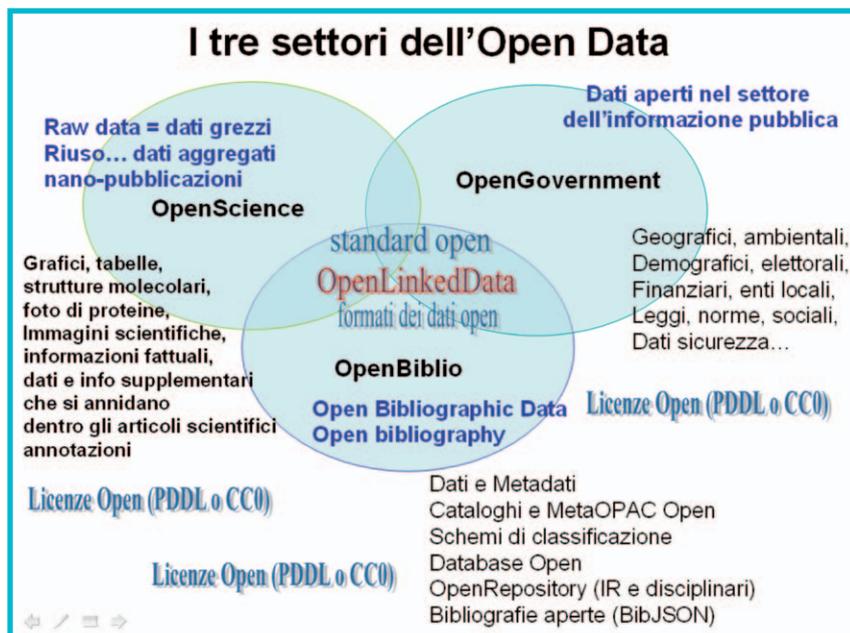
Tuttavia è indubbio che il movimento Open Data Commons, essendo il più giovane tra i movimenti fratelli, non è stato ancora in grado di dare una definizione precisa di Open Data. Del resto gli stessi strumenti legali collaudati per altri mondi open, nel mondo globalizzato degli Open Data hanno appena

iniziato il confronto tra differenti legislazioni e differenti assetti normativi.

Una buona definizione di partenza di cosa siano l’Open Data, l’Open Content, e l’Open Information è stata formulata dall’Open Knowledge Foundation, un’organizzazione no-profit, fondata nel 2004, con lo scopo di promuovere la conoscenza aperta, nel tentativo di rendere sempre più ampia e organizzata la condivisione di dati per renderli utilizzabili, riutilizzabili, ridistribuiti per scopi sociali benefici.<sup>21</sup> La suddivisione dei dati secondo OKF si differenzia in tre ambiti:

- a) scientifico, storico, geografico o in altro modo;
  - b) contenuti: musica, film, libri;
  - c) governativo e delle altre informazioni amministrative.
- Nell’Open Data ritroviamo la stessa matrice presente in altri movimenti “open” tra i quali:

- *Open Source*: riguarda le licenze di distribuzione di programmi e applicazioni;
- *Open Content*: facilmente confondibile con l’Open Data, ma mentre l’Open Data concentra l’attenzione sulla ricerca scientifica e i dati che le competono, l’Open Content si rivolge alla creatività delle opere;
- *Open Knowledge*: si ispira all’Open Content e mira alla “diffusione della conoscenza con le seguenti caratteristiche: libera (di proprietà di ogni essere umano), accessibile (raggiungibile da ogni essere umano), comprensibile (presentata nel modo più semplice possibile, e attraverso la definizione esaustiva di ogni significato coinvolto), universale (accessibile in ogni lingua)”;



Tre settori dell'Open data con al centro LinkedOpenData (LOD) [di Antonella De Robbio]

- *Open Access*: riguarda la comunicazione scientifica e si concretizza nella disponibilità dei contenuti digitali online, articoli di riviste, ricerche svolte in campo accademico, consentendone la pubblicazione e l'accesso/utilizzo gratuito;
- *Open Learning*: nuovo movimento per una didattica a distanza Open (a vari livelli): piattaforme (Moodle), fino ad arrivare a esperienze come la Khan Academy,<sup>22</sup> un'organizzazione educativa non a scopo di lucro creata nel 2006 con l'obiettivo dichiarato di "fornire un'educazione di alta qualità a chiunque, dovunque";
- *OpenCourseWare*:<sup>23</sup> l'iniziativa da cui ha tratto origine il movimento delle OER;
- *Open Educational Resources* (OER): una risorsa didattica aperta deve avere caratteristiche di duplice apertura: dal punto di vista dei diritti deve essere licenziata in modo da poter essere usata e ri-utilizzata, dal punto di vista tecnico deve avere il codice sorgente aperto per sviluppi ulteriori;
- *Massive Open Online Courses* (MOOC):<sup>24</sup> corsi di grandi dimensioni ad accesso massivo dove i partecipanti sono distribuiti a livello globale e i materiali del cor-

so sono anch'essi aperti e collocati in punto differenti del web;

- *Open Library*: consistente in progetti che mirano a creare pagine web per ogni libro pubblicato, permettendo l'accesso libero all'opera, la condivisione, la diffusione della stessa. Nello specifico l'*Open Library Project* nasce da Internet Archive, in collaborazione con la California State Library e la Kahle Austin Foundation;
- *OpenBibliographic Data*: sotto-insieme di OpenData;
- *Open Bibliography* in science, technology and medicine (STM) è una combinazione di strumenti Open Source, specifiche tecniche Open e Open bibliographic data. Usa BibJSON, un formato di dati a struttura testuale semplice (BibTex, Dublin Core, PRISM e JSON);
- *Open Linked Data* (LOD): affinché i dati siano realmente *open*, devono essere anche tecnicamente *connettabili*. In sostanza quando parliamo di dati aperti, non ci riferiamo solo a una "filosofia" generale o alle questioni del copyright, ma anche all'aspetto strettamente tecnico della *interoperabilità*.



Il thread #opendata di Twitter alla data del 3 marzo 2012 ore 14.35

## OpenData & LinkedOpenData

Un aspetto strettamente correlato al movimento Open Data è il Linked OpenData (LOD) del W3C il quale ha lo scopo di mettere in web insiemi di dati aperti e connessi in modo interoperabile. Per questa ragione potremmo dire che i dati LOD sono un aspetto stesso del web semantico, in quanto il termine “dati collegati” è usato per descrivere un metodo di esporre, condividere e connettere dati attraverso URI differenziabili, entro un modello RDF impostando link RDF tra i dati provenienti da differenti risorse, al di là dei “profile” o di standard specifici. Il *W3C Incubator Group*, attivo da maggio 2010 ad agosto 2011, ha pubblicato il *Library Linked Data Incubator Group Final Report*. Obiettivo del gruppo è stato aumentare l’interoperabilità dei dati delle biblioteche sul web, mettendo assieme persone interessate alle attività sul web semantico nella comunità bibliotecaria e delle scienze dell’informazione, basandosi su iniziative esistenti e identificando possibilità future di collaborazione.

Le raccomandazioni contenute nel report sono le seguenti:

- i responsabili dei dati dei sistemi bibliotecari devono identificare i set di dati idonei ad essere esposti come Linked Data e promuovere la discussione su Open Data e diritti connessi;
- chi si occupa di standard per le biblioteche deve promuovere la partecipazione delle biblioteche alle discussioni sul web semantico e la sua standardizzazione, sviluppare standard compatibili con i linked data, e diffondere buone pratiche per incentivare le biblioteche a muoversi in direzione linked data;
- bibliotecari e archivisti conservano i set di linked data e applicano l’esperienza delle biblioteche nella cura e nella conservazione a lungo termine dei set di dati.

In quest’ottica, di grande interesse è il Metadata Handbook,<sup>25</sup> un’iniziativa lanciata dal Public Domain e dall’Open Bibliographic Working Group dell’OKF, dove si stanno mettendo a punto via Wiki delle Raccomandazioni.<sup>26</sup>

Anche dai governi e organizzazioni internazionali, fino alle città e istituzioni locali, vi è un impegno diffuso all’apertura e al collegamento dei dati secondo i quattro principi che conducono al Linked Open Data (LOD):<sup>27</sup>

- quando si pubblicano dati bibliografici si deve usare una licenza esplicita e robusta;
- è necessario usare una licenza o una deroga appropriata al tipo di dati;

- se davvero si vuole che i dati prodotti siano realmente usati e arricchiti da altri, devono essere aperti nelle modalità definite dalla Open Definition; in particolare, non devono essere usate la clausola non-commerciale e altre clausole restrittive;
- ove possibile, è preferibile porre esplicitamente i dati bibliografici nel Pubblico Dominio tramite licenza PDDL (Public Domain Data Licence) o CC0 (Creative Commons Zero).

Un fenomeno di grande potenzialità (seppure ancora limitato) riguarda quelle riviste – di ambito prevalentemente biologico, ambientale e sociologico – che si sono dotate di un mandato di tipo *data-sharing* e che richiedono agli autori degli articoli di fornire dati aperti per un possibile ri-uso ai fini della replicabilità dei risultati in modo indipendente. A riguardo il Digital Curation Centre<sup>28</sup> ha pubblicato il documento *Cite Datasets and Link to Publications*,<sup>29</sup> guida che illustra la creazione di collegamenti tra le pubblicazioni di ricerca e i dati su cui si basano, così da rendere possibile la localizzazione dei dataset per chi legge un articolo, e viceversa.

Le potenzialità insite nell’esposizione dei set dei metadati ai fini della valutazione di un lavoro di ricerca fondato su set di dati, sono oggetto di un certo interesse da parte di alcune iniziative correlate agli archivi aperti.<sup>30</sup> In un recente lavoro focalizzato sul progetto pilota EASY si è descritto il workflow dei dataset ai fini di un loro utilizzo, di modo che anche i dataset allegati al paper di ricerca rientrino nella peer-review e siano valutati con approccio sia qualitativo sia quantitativo.<sup>31</sup>

David Shotton, in un recente articolo pubblicato su “D-Lib Magazine”,<sup>32</sup> analogamente a quanto proposto da Tim Berners-Lee con *Five Stars of Linked Open Data*, suggerisce di valutare gli articoli online tramite una costellazione di cinque criteri indipendenti all’interno di un universo multidimensionale. Per ciascuno dei cinque fattori (peer-review, accesso aperto, contenuto arricchito, set di dati disponibili e metadati leggibili dalla macchina) viene fornita una scala di valori di cinque punti, che vanno da 0 a 4. Il modello a cinque stelle può fornire un quadro concettuale con cui giudicare il grado di apertura di un articolo (e anche di una rivista) utile alla valutazione in termini di compatibilità a gradazioni più o meno open.

Uno studio depositato lo scorso novembre 2011<sup>33</sup> sull’archivio internazionale dei fisici ArXiv.org, propone di verificare il livello citazionale di articoli che contengono un collegamento ai dati a cui fanno riferimento,

rispetto ad articoli che invece non citano direttamente i dati primari della ricerca. Per gli autori la questione, oltre ad essere interessante da un punto di vista puramente accademico, è rilevante per il miglioramento e diffusione della scienza: la condivisione dei dati, infatti, consente di verificare gli esperimenti e di fare nuove scoperte. Inoltre il fatto che gli articoli siano corredati da dati verificabili, consente un impatto maggiore in quanto vengono più citati creando un ciclo virtuoso. Nonostante gli evidenti benefici della condivisione dei dati della ricerca, la pratica non è così diffusa: spesso mancano non solo la volontà, ma anche gli strumenti per collegare i dati alle pubblicazioni.

### I nuovi alfabeti dell'Open Biblio

L'Open Biblio<sup>34</sup> è il luogo dove le conversazioni sui dati aperti si aprono perché le strutture di dati sono descritte in modo adeguato. Tramite un modello di dati applicabile e usabile in modo interoperabile l'Open Biblio rende possibile funzionalità di *parsing* tra contenitori informativi differenti di modo che anche i motori di ricerca localizzino e recuperino le informazioni opportunamente connesse.

Risale a non più di un paio di anni fa il memorabile *talk* di David McCandless alla TED Conference del giugno 2010 dove la bellezza della visualizzazione dei dati costituiva l'architettura stessa dell'informazione come strumento di navigazione. In quella sede McCandless illustrava la trasformazione di insiemi di informazioni complesse come le spese militari nel mondo, i *buzz* dei media web 2.0, gli stati degli utenti di Facebook, l'impatto dei twitters entro le comunità ecc. in semplici ed eleganti diagrammi. Una mappa che aiuta a non perdersi tra i meandri dell'informazione, trovando percorsi unici e connessioni che potrebbero cambiare il modo in cui percepiamo il mondo reale. Sebbene come ci ricorda la metafora di Bateson,<sup>35</sup> la mappa non può coincidere col territorio, così come i simboli non coincidono con la realtà

Attraverso la visualizzazione dell'informazione, riusciamo ora a entrare dentro un paesaggio che possiamo esplorare con gli occhi, una sorta di mappa dell'informazione. E quando ti perdi dentro l'informazione, ecco che una mappa dell'informazione è un modello che ti aiuta. (...) Nel corso degli anni, in rete, abbiamo riversato grandi quantità di informazioni e di dati: li abbiamo irrigati attraverso le reti e la connettività per farli crescere. Sono stati coltivati con il lavoro volontario di persone e di Governi.<sup>36</sup>

Anche per il mondo Open Biblio possiamo parlare di nuovi alfabeti se ci riferiamo a modelli per registrare dati bibliotecari che si connettano a contesti di ricerca, istituzionali o governativi, emulando in parte quanto la tradizione MARC ci ha insegnato nella registrazione dei dati in diversi alfabeti. Se un alfabeto è associato ad una lingua o ad un linguaggio che costruisce le proprie parole sull'alfabeto, nel linguaggio "formale" della biblioteca i dati (metadati) – qualora aperti – possono costituire i simboli ideali per un sistema di relazioni, attraverso la connessione tra scienza e ambiente sociale. Il *social environment* inteso come ambiente sviluppato dall'uomo in contrasto con l'ambiente naturale è costituito dall'insieme delle condizioni individuali di vita e di lavoro, dal livello di reddito, dal grado d'istruzione e dalla comunità di cui ciascuno fa parte. Tutti questi elementi esercitano per esempio un forte impatto sulla salute e di conseguenza sulla qualità della vita. La grande diversità di ambienti sociali esistenti nei vari Paesi del Mondo contribuisce ad aumentare le disparità sul piano della qualità della vita, in quanto fenomeni come la diffusione di malattie, l'aspettativa di vita, variano notevolmente non solo tra ricchi e poveri, ma tra persone più o meno istruite. Per migliorare l'ambiente sociale sono state avviate diverse iniziative comunitarie e internazionali che mettono al centro i dati che devono essere aperti, condivisibili e riusabili. La biblioteca tramite l'Open Biblio può creare un ponte d'oro tra le istituzioni e il cittadino se biblioteca pubblica, tra scienza e società se biblioteca accademica o di ricerca.

Il JISC ha sviluppato per conto della Resource Discovery Taskforce un utilissimo documento presente sul sito "Guide to Open Bibliographic Data",<sup>37</sup> focalizzato su alcuni casi o esempi di applicazione possibili di Open Bibliographic Data come oggetto di studio.

Dal settore OpenBiblio si origina per esempio tutto il filone delle bibliografie aperte, fondamentali per gli aspetti correlati all'impatto citazionale tra le comunità di ricerca che operano in rete. Il concetto di Open Bibliography nella scienza, tecnologia e medicina è introdotto – da autori come Peter Murray Rust<sup>38</sup> – come una combinazione di strumenti OpenSource, di specifiche aperte e di dati bibliografici aperti.

Nell'universo bibliografico c'è un chiaro cambiamento di paradigma da formati record fissi alle ricombinabili dichiarazioni sui metadati. Le istituzioni e i sistemi bibliotecari che implementano archivi di dati bibliografici e relativi servizi dovranno necessariamente prevedere

processi e strategie di fornitura dati compatibili con il *framework* Linked Data e di licenze adeguate.<sup>39</sup>

In merito al rilascio dei metadati in modalità più o meno aperte, e/o più o meno LOD-compatibili, i cataloghi bibliotecari presentano tuttavia comportamenti fortemente disomogenei. Questa difformità è probabilmente dovuta ad una mancanza di consapevolezza sulle potenzialità dell'OpenData da una parte, ma anche da una inesperienza pressoché generalizzata sull'uso delle licenze Open Data Common e sulle clausole di utilizzo. In questo panorama si registrano esperienze diverse. Le grandi biblioteche nazionali europee hanno cominciato a pubblicare i propri metadati con una licenza aperta. Esempi sono la Bibliothèque nationale de France,<sup>40</sup> le biblioteche nazionali del Regno Unito, di Germania e di Spagna oltre alla biblioteca digitale d'Europa "Europeana" che espone i propri metadati in modalità Linked Open Data (LOD). I metadati aperti di Europeana a disposizione del settore pubblico e di quello privato sono utilizzabili per il web o per lo sviluppo di servizi innovativi tramite applicazioni su dispositivi mobili (smartphone, tablet...) con licenza *CC0 1.0 Universal Public Domain Dedication*.<sup>41</sup>

Grandi cataloghi come quello dei servizi informativi della biblioteca del CERN hanno fin da subito messo a disposizione funzionalità per lo scarico libero dei dati. Gli archivi del CERN, che comprendono anche i metadati del catalogo bibliografico, sono muniti di licenze del tipo Pubblico Dominio (*Public Domain Data License* o *PDDL*, e *Creative Commons 0 CC0*). I dati della CERN Library confluiscono anche entro l'Open Library Project<sup>42</sup> e gli utenti possono aggiungere commenti, immagini e informazioni in modo collaborativo. Anche l'intero dataset può essere scaricato<sup>43</sup> da soggetti terzi, mentre per il riuso massiccio i dati sono forniti di uscite Z39.50, SRU, e muniti di interfaccia OAI via *biblios.net*, un repository di open bibliographic data.<sup>44</sup>

Jens Vigen, direttore della biblioteca del CERN, dice:

I libri dovrebbero essere catalogati una volta sola. Attualmente il portafoglio pubblico paga più e più volte per avere lo stesso libro catalogato da più punti. I bibliotecari dovrebbero agire come predicatori: insiemi di dati creati grazie a finanziamenti pubblici devono essere resi disponibili gratuitamente a chiunque sia interessato. C'è un potenziale tremendo nel creare un ambiente naturale di condivisione con il riutilizzo dei dati bibliografici in un paesaggio ricco di servizi mash-up dove la maggior parte degli attori coinvolti, utenti o fornitori, potrebbero non essere semplicemente bibliotecari o utenti o fornitori di biblio-

teche. Per noi, qui al CERN l'Open Access è cosa naturale e pertanto crediamo nell'apertura ai dati e nel conseguente riutilizzo. Le nostre azioni sono condotte nello spirito della Dichiarazione di Berlino sull'accesso aperto alla conoscenza e i dati appartengono al patrimonio culturale. Tutti i firmatari della Dichiarazione dovrebbero allineare le loro politiche di conseguenza.<sup>45</sup>

Di potenziale economico parla anche Ilaria Fava a proposito dell'accordo sullo scambio dei metadati<sup>46</sup> alla base delle attività svolte da Europeana sui dati aperti: "Questo sostegno al settore commerciale nell'ambito del digitale è al centro della strategia di business di Europeana – scrive – e viene sottoscritto dalle biblioteche, dai musei, e dagli archivi che decidono di inserire i propri metadati in Europeana".<sup>47</sup>

Esempi concreti di cataloghi e/o servizi bibliotecari che hanno aderito pienamente alla filosofia OpenData sono quelli forniti dal registro CKAN (Comprehensive Knowledge Archive Network).<sup>48</sup>

Molto diversa è la situazione dei cataloghi universitari, che per loro natura sono decentrati, e hanno una situazione generalmente non chiara sulla titolarità dei diritti. 34 università italiane utilizzano il software proprietario Aleph. Se da un lato questo può essere visto come un problema di interoperabilità, dall'altro è opportuno ricordare che l'apertura dei dati non dipende da nessun software specifico (esistono standard internazionali per i dati bibliografici che ogni software è in grado di gestire ed esportare). Inoltre, Ex Libris, la casa produttrice di Aleph, ha reiterato la totale separazione del proprio software dai dati che esso gestisce: sta alle biblioteche gestirli in autonomia, e decidere sul loro rilascio. Esistono purtroppo situazioni documentate di cattiva gestione, in cui per una serie di fattori le banche dati sono "prigioniere" del sistema di gestione, e crediamo che il movimento #openbiblio abbia un valore importante anche per migliorare queste situazioni e garantire la conservazione dei dati bibliografici per la collettività. È comunque indubbio che l'adozione di software libero renderebbe più agevole l'apertura dei dati e la sua sostenibilità a lungo termine.<sup>49</sup>

Tuttavia anche in ambienti che apparentemente sembrerebbero più aperti, licenze che ricadono nel pubblico dominio come la CC0 e la PDDL non vengono adottate come dovrebbero. Queste licenze – estremamente aperte – non richiedono attribuzione e non pongono limiti a successivi riutilizzi, di conseguenza con la loro adozione si dichiara esplicitamente di rinunciare a tutti i diritti e ai crediti relativi alla titola-

rità o menzione d'autore su tutti i dati bibliografici del catalogo creati dall'istituzione. Per le ragioni su esposte l'Online Computer Library Center (OCLC)<sup>50</sup> ha preferito orientarsi su una licenza *Open Data Commons Attribution (ODC-BY)*,<sup>51</sup> una via di mezzo tra le licenze di pubblico dominio – sulle quali OCLC nutre parecchie riserve – e la *Open Database License (ODC-ODbL)*<sup>52</sup> più restrittiva in termini di interoperabilità, ma diventata un punto di riferimento anche in ambito pubblico.<sup>53</sup>

Degno di riflessione è il caso del progetto OpenStreetMap (OSM),<sup>54</sup> un progetto collaborativo finalizzato a creare mappe a contenuto libero del mondo. Il progetto punta ad una raccolta mondiale di dati geografici, il cui scopo principale è la creazione di mappe e cartografie. OpenStreetMap è infatti un servizio open source basato sul *crowdsourcing*, una specie di Wikipedia per la geografia. Gli utenti creano mappe *raw* (grezze) e le pubblicano entro il progetto OpenStreetMap, concedendo i dati con licenza Open Database License (ODbL), in una sorta di atlante globale in *crowdsourcing*.

Da questo contenitore open i servizi di geolocalizzazione – anche commerciali – prendono le mappe e le personalizzano per i dispositivi mobili o per servizi web. In cambio queste piattaforme di servizi creano nuove mappe e le mettono in Open Data dentro OpenStreetMap.

Recente il caso dell'uscita di Foursquare<sup>55</sup> – e di altre compagnie che forniscono servizi di geolocalizzazione – da Google Map che forniva loro i dati per le mappe. Da ora, per la costruzione di mappe nuove e proprietarie, Foursquare si affiderà alla startup MapBox, servizio per creare e personalizzare mappe i cui dati verranno poi integrati in modo open entro OpenStreetMap. In altri termini un crescente numero di compagnie sta optando per soluzioni alternative all'uso delle API a pagamento di Google Maps che chiede fino a 300mila \$ annui per il servizio di rilascio delle mappe. Anche StreetEasy, il servizio di geolocalizzazione circoscritto alla città di New York ha deciso di abbandonare le mappe di Google, e di iniziare a lanciare le proprie API mobili. Siamo alla guerra delle mappe, laddove l'utilizzo della geolocalizzazione si diffonde massicciamente tra gli utenti, e servizi commerciali hanno iniziato ad utilizzare i dati di OpenStreetMap, creando modelli di business alternativi basati sull'uso di librerie javascript open-source per potenziare le mappe in un lavoro condiviso e collaborativo, per servizi di geolocalizzazione web e per dispositivi mobili.

## Dati aperti per società aperte

I dati pubblici sono realmente utili laddove essi sono grezzi, davvero aperti e connessi. Esempi di dati aperti nel settore dell'informazione pubblica sono:

- dati geografici (mappe geografiche, stradali, geologiche, dati sui movimenti sismici, utilizzo del territorio...);
- dati ambientali (atmosferici, climatici, sull'inquinamento...);
- dati sul trasporto locale;
- dati demografici (relativi a censimenti, popolazione, flussi migratori...);
- dati elettorali;
- dati sulla produzione e sul consumo energetici;
- dati finanziari (di bilancio) e tributari (redditi e tasse);
- dati economici e delle attività degli enti locali (primo Comune italiano: Firenze – ottobre 2011 – ha pubblicato numerosi set di dati);
- dati sui beni immobili (catastali);
- dati sulla salute pubblica (dati aggregati);
- dati sullo stato dell'educazione;
- dati per la gestione e smaltimento;
- dati sulla gestione delle acque;
- dati sulla sicurezza pubblica;
- dati di fonte e documentazione pubblica (legislazione, giurisprudenza).

La gestione dei dati è inoltre essenziale per affrontare le sfide in campo ambientale. A titolo di esempio, si può citare l'analisi dei modelli di consumo energetico per migliorare l'efficienza energetica o dei dati sull'inquinamento per la gestione del traffico. L'adozione di decisioni strategiche consapevoli nei settori dei trasporti, dell'uso del territorio e dei cambiamenti climatici dipendono in misura crescente dall'analisi dei dati disponibili. Francesca Di Donato nel suo ottimo saggio su trasparenza di governo e cittadinanza attiva<sup>56</sup> sottolinea come siano proprio i momenti di grave crisi economica e finanziaria a fungere da volano per la *e-Governance*: “L'Open Government si configura così come una nuova modalità di amministrazione che nasce per riavvicinare governanti e governati e per restituire agli ultimi la fiducia nei primi. Per farlo, si riconosce la necessità di implementare i tre principi enunciati nel preambolo, vale a dire la trasparenza, la partecipazione e la collaborazione”.

Open Data come misura per il rilancio dell'economia è il focus di una ricerca realizzata dalla Scuola sant'Anna di Pisa<sup>57</sup> e coordinata da Marco Fioretti che propo-

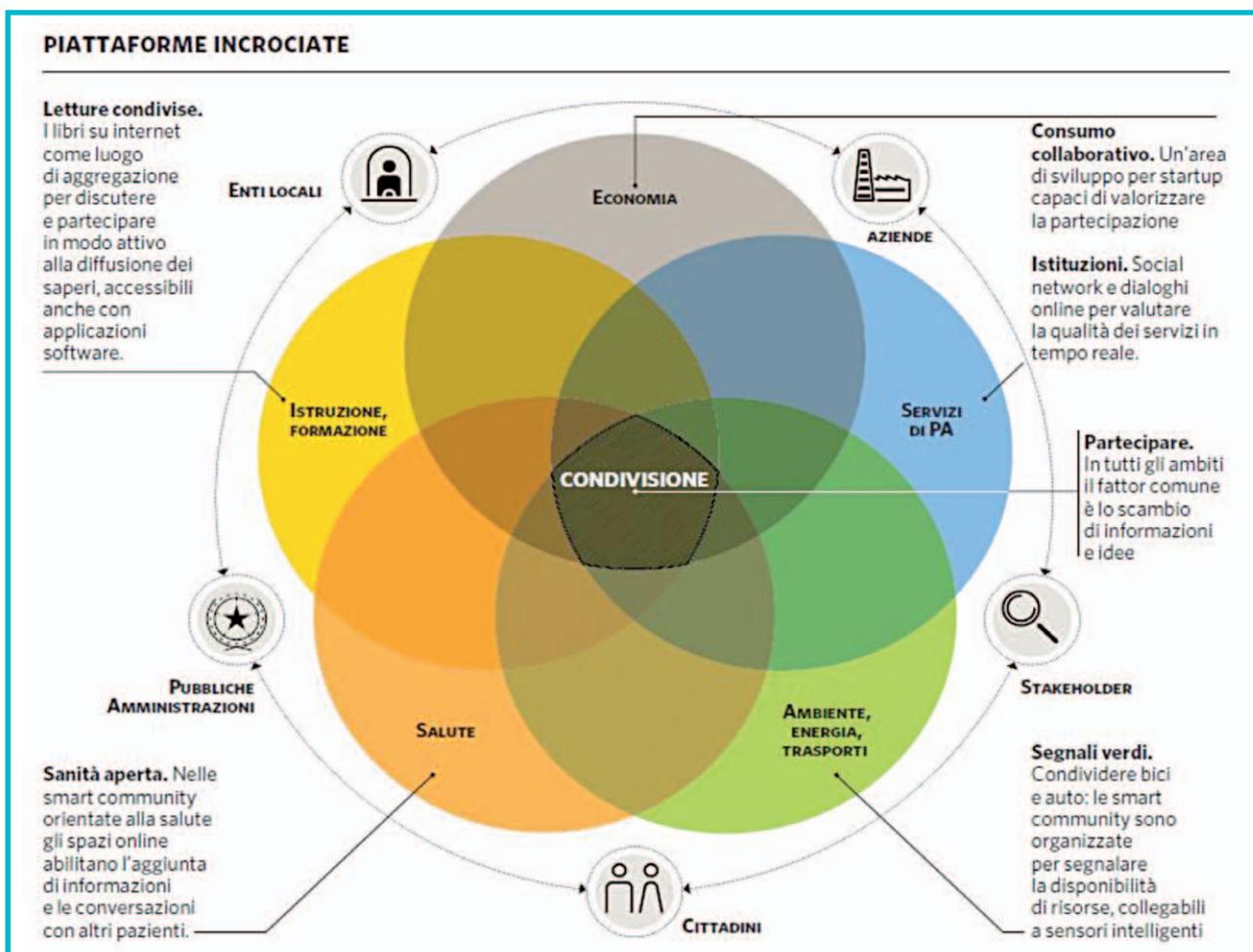


Immagine tratta dal post *Obiettivo smart city* [OpenDataBlog del Il Sole 24 ORE, lunedì 27 febbraio 2012 - 16:57]

ne sette linee di azione per impostare correttamente un “governo” dei dati aperti:

- definire chiaramente e spiegare sia il concetto di Open Data sia di public data;
- mantenere separate le questioni politiche da quelle economiche;
- distinguere tra dati esistenti e dati futuri;
- imporre licenze appropriate;
- educare la cittadinanza alla comprensione ed all'uso dei dati;
- focalizzarsi su questioni locali e specifiche al fine di suscitare l'interesse per l'Open Data;
- coinvolgendo ONG, enti di beneficenza ed associazioni imprenditoriali.

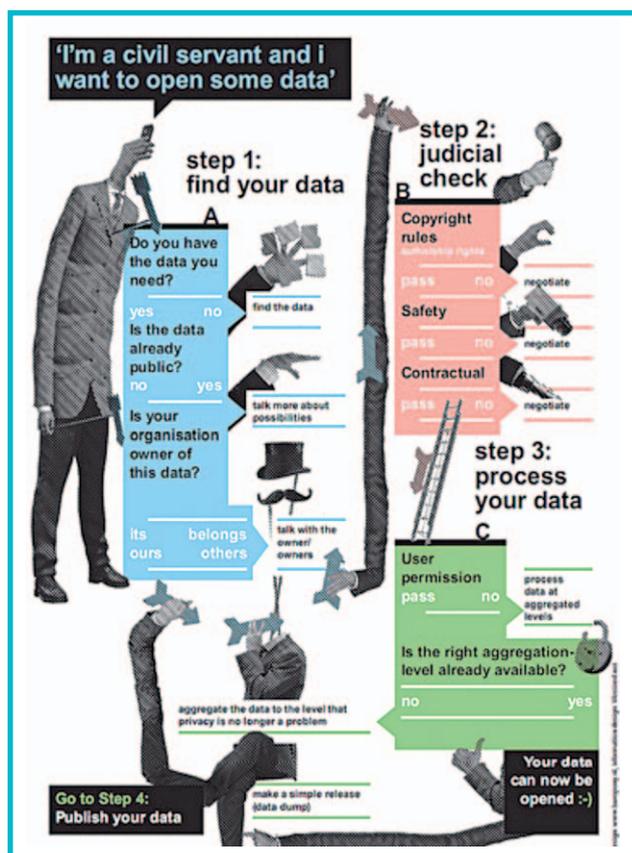
Il “Sole 24 Ore Online” ha aperto un blog dedicato all'argomento, “Open Data Blog”,<sup>58</sup> osservatorio ed aggregatore di novità sul tema, al cui interno è stata lanciata l'iniziativa “Porte Aperte”,<sup>59</sup> con l'intento di ospi-

tare segnalazioni, studi e conseguenti elaborazioni con tecniche di infografica.

Oltre al repository di CKAN citato prima, si segnala anche il repository Spaghetti Open Data.<sup>60</sup> Sono sorti anche corsi strutturati come quello organizzato sul sito dell'associazione di giornalismo investigativo e Radio Radicale “Open Data Journalism”.<sup>61</sup>

In altri termini il motto potrebbe essere: Open Data per ristrutturare i sistemi di governo rilanciando al contempo le economie e sconfiggendo la povertà,<sup>62</sup> come racconta Luca Conte a proposito dell'iniziativa delle Nazioni Unite “Global Pulse”,<sup>63</sup> progetto che punta a sviluppare – grazie anche alla collaborazione di aziende private esperte in Open Data – una piattaforma tecnologica open source a disposizione dei governi di tutto il mondo:

La rete di Pulse Labs a regime godrà almeno di dieci nodi nel mondo, seguendo il modello del primo laboratorio



L'Open Government Data Poster di LifeSized

aperto a New York. Le ricerche in corso spaziano dal monitoraggio delle questioni sulla sicurezza del cibo attraverso l'analisi degli articoli pubblicati dai media tradizionali, all'andamento della disoccupazione e del mercato del lavoro rilevato attraverso i social media, all'andamento del prezzo del pane in sei paesi dell'America Latina estraendo le informazioni dal web. Twitter è uno dei campi di indagine, con una ricerca tra Asia e Nord America sulla percezione dello stress legato alla crisi.<sup>64</sup>

In tale direzione alcune iniziative interessanti come il programma della Banca Mondiale accessibile entro un portale specializzato in dati economici mette a disposizione di ricercatori e semplici curiosi una mole straordinaria di dati,<sup>65</sup> accessibili in diversi formati, scaricabili attraverso API.

“Kenya Open Data” è un ambizioso progetto che vede il Kenya quale prima nazione, tra i paesi in via di sviluppo nel sud del mondo, ad aprire una piattaforma di Open Data suddivisi in sei tematiche: istruzione, energia, salute, popolazione, povertà, risorse idriche e servizi igienici. Le tabelle più scaricate riguardano la povertà per provincia e la spesa pubblica. Nella home è presente

una mappa che mostra le 47 contee della nazione, mentre l'area “community” segnala già le prime applicazioni sviluppate a partire dagli Open Data, tra le quali:

- *Huduma*: uno spazio dove i cittadini possono dare voce alle loro richieste per avere accesso alle infrastrutture di base, aiutati da una visualizzazione delle risorse disponibili nelle aree vicine;
- *Msema Kweli*: un'applicazione per cellulari che permette di vedere i piccoli progetti in fase di realizzazione sul campo che altrimenti sarebbe difficile raggiungere da una singola finestra di accesso.

Nelle reti sociali alcuni commenti sugli Open Data propongono di trasferire anche nel settore dell'informazione pubblica – in modo analogo a quanto avviene nella scienza – un meccanismo di revisione peer-review da parte dei cittadini rispetto a quanto promulgato dai governi relativamente a leggi, regolamenti, ordinanze. La necessità di conoscere i dati pubblici fa parte di un sistema democratico che aumenti la consapevolezza e la responsabilità dei cittadini. Allo scopo sono da segnalare alcuni prototipi per nuovi servizi che utilizzano dati governativi aperti come il progetto Roomware,<sup>66</sup> o il framework VURB,<sup>67</sup> quadro di politiche europee e di ricerca nel settore del design urbano attraverso sistemi computazionali, iniziative sfociate nell'*Open Government Data Poster*<sup>68</sup> sul sito di LifeSized.

Iniziative di un certo spessore sono promosse anche in Italia, in primis dall'Associazione italiana per l'Open Government<sup>69</sup> che riunisce un gruppo di esperti di diritto e di nuove tecnologie, funzionari pubblici e privati, docenti universitari ed altri componenti della società civile con l'obiettivo di sensibilizzare cittadini, imprese ed Amministrazioni e promuovere l'attuazione di strategie di Open Government nel nostro Paese. Numerose le istituzioni pubbliche italiane che stanno rispondendo in modo proattivo, tra le quali la Camera dei Deputati,<sup>70</sup> la Regione Piemonte<sup>71</sup> e la Regione Emilia Romagna,<sup>72</sup> l'Istat,<sup>73</sup> il CNR,<sup>74</sup> l'ENEL<sup>75</sup> e il Comune di Firenze.<sup>76</sup>

Dal 18 ottobre 2011 è attivo il *Portale dei dati aperti della PA*<sup>77</sup> italiano sviluppato dal Foromez PA e nato per promuovere il riuso delle informazioni pubbliche per cittadini, sviluppatori, imprese, associazioni di categoria e per le stesse pubbliche amministrazioni. La *Italian Open Data License (IODL)*, licenza italiana per banche dati pubbliche, nasce proprio grazie a queste sinergie.<sup>78</sup>

Il portale italiano del Foromez si ispira alla strategia europea per gli Open Data all'interno dell'Agenda digitale

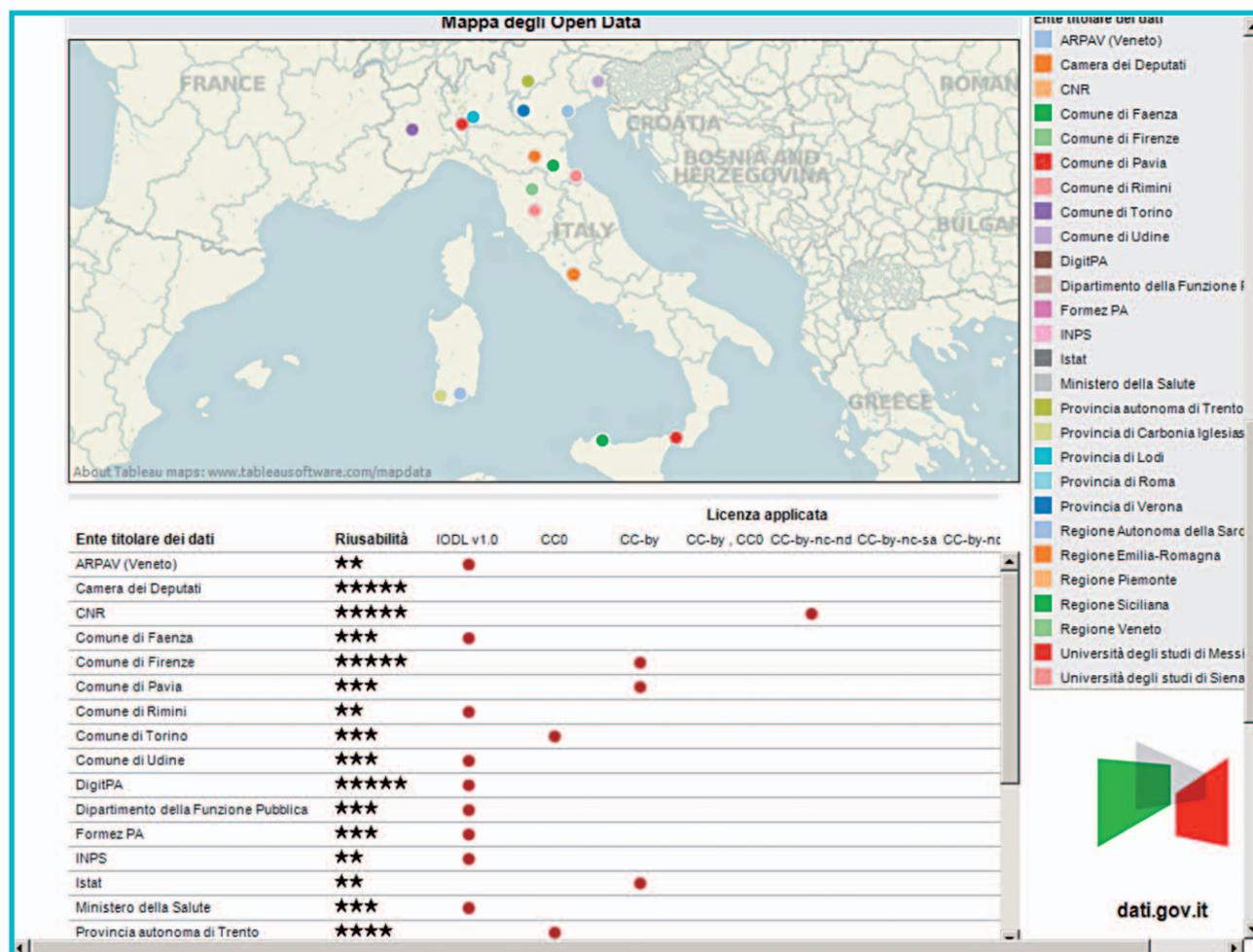
europea “Trasformare in oro i dati delle amministrazioni pubbliche”,<sup>79</sup> in quanto

le pubbliche amministrazioni europee sono sedute su una miniera d'oro dalle potenzialità economiche non valorizzate (...). Si tratta di interventi che collocano l'Unione europea all'avanguardia mondiale nel riutilizzo delle informazioni del settore pubblico e che daranno impulso al settore, già in forte crescita, che si occupa della trasformazione di dati grezzi in materiale da cui dipendono centinaia di utilizzatori delle tecnologie dell'informazione e della comunicazione: ad esempio applicazioni per gli smartphones, quali mappe, informazioni in tempo reale sul traffico e le condizioni meteo, strumenti di comparazione dei prezzi ecc. Tra gli altri principali beneficiari di queste iniziative figurano i giornalisti e gli accademici.<sup>80</sup>

Parallelamente in Horizon 2020 – il prossimo Ottavo Programma Quadro – le e-infrastructures sono viste

come la base fondamentale per l'e-science e l'open access come un requisito essenziale per i dati scientifici, consentendo ai ricercatori di spingersi oltre gli attuali confini della scienza.<sup>81</sup>

La strategia UE 2020<sup>82</sup> ha come obiettivo principale indirizzare le economie europee su un percorso di crescita elevata e sostenibile. A tal fine l'Europa dovrà rafforzare il proprio potenziale d'innovazione e ottimizzare l'uso delle proprie risorse, una delle quali è costituita dai dati pubblici – ovvero tutte le informazioni che gli organismi pubblici nell'Unione europea producono, raccolgono o acquisiscono a pagamento. Si tratta di informazioni dalle notevoli potenzialità – finora poco valorizzate – di riutilizzo in nuovi prodotti e servizi e che potrebbero permettere di migliorare l'efficienza delle amministrazioni. Entro la primavera del 2012 la Commissione UE renderà pubblici i dati in suo possesso tramite un nuovo “portale dei dati” che fungerà da punto di accesso unico ai dati ri-



Mappa degli Open Data in Italia

utilizzabili messi a disposizione da tutte le istituzioni, organismi e agenzie dell'Unione europea e dalle autorità nazionali.

L'applicazione di politiche in materia di dati aperti dovrebbe avvenire nel pieno rispetto dei diritti di proprietà intellettuale dei terzi e degli obblighi che incombono sull'Unione europea in virtù di trattati internazionali sui diritti di proprietà intellettuale.

Le competizioni attorno ai dati aperti sono laboratori di innovazione aperta, importanti non tanto perché favoriscono nuove applicazioni, ma perché possono ampliare i nostri orizzonti e cominciano fin d'ora a rivelare la profondità della nostra immaginazione in merito al potenziale delle opportunità che i dati aperti ci offrono.<sup>83</sup>

## NOTE

\* Questo articolo riprende i temi trattati dall'autrice in occasione del convegno "I nuovi alfabeti della biblioteca" che si è tenuto a Milano, presso il Palazzo delle Stelline, il 15 e 16 marzo 2012.

<sup>1</sup> Per approfondimenti vedi: ANTONELLA DE ROBBIO – SILVIA GIACOMAZZI, *Dati aperti con LODe*, "Bibliotime", 14 (2011), n. 2, <<http://didattica.spbo.unibo.it/bibliotime/num-xiv-2/derobio.htm>>.

<sup>2</sup> Rapporto della società McKinsey dal titolo: *Big Data, The next frontier for innovation, competition and productivity*, maggio 2011.

<sup>3</sup> RICCARDO LUNA, *Così gli archivi gratis sul Web cambieranno le nostre vite*, "la Repubblica", 18 ottobre 2011, <[http://www.repubblica.it/tecnologia/2011/10/18/news/open\\_data-23405552/?ref=HRERO-1](http://www.repubblica.it/tecnologia/2011/10/18/news/open_data-23405552/?ref=HRERO-1)>.

<sup>4</sup> John Wilbanks, direttore esecutivo di Science Commons.

<sup>5</sup> Gli autori dei *Panton Principles* sono: Peter Murray-Rust, Università di Cambridge (Regno Unito), Cameron Neylon, STFC (UK), Rufus Pollock, Open Knowledge Foundation, Università di Cambridge (Regno Unito), John Wilbanks, Science Commons (USA).

<sup>6</sup> <<http://www.arl.org/sparc/opendata/index.html>>.

<sup>7</sup> Biochimico all'Università di Cambridge (UK) - Unilever Centre for Molecular Sciences Informatics.

<sup>8</sup> <<http://opendatacommons.org/licenses/pddl/>>.

<sup>9</sup> <<http://creativecommons.org/choose/zero>>.

<sup>10</sup> I due studi paralleli sul virus H5N1 coordinati da Ron Fouchier dell'Erasmus Medical Center di Rotterdam e da Yoshihiro Kawaoka dell'Università del Wisconsin sono stati finanziati dagli istituti di sanità americani e sono stati sottomessi rispettivamente a "Science" e a "Nature".

<sup>11</sup> Il National Science Advisory Board for Biosecurity (NSABB) è una commissione che fa capo al governo americano.

<sup>12</sup> Cfr. ELEONORA VIGANÒ, *H5N1: censura inutile o preoccupazioni fondate?*, "Jeckyll. Comunicare la scienza", <<http://jeckyll.sissa.it/?p=6017>>.

<sup>13</sup> <<http://www.openaire.eu/it/component/content/article/223-seizing-the-opportunity-for-open-access-to-european-research-ghent-declaration-published>>.

<sup>14</sup> <<http://www.openaire.eu/>>.

<sup>15</sup> <<http://opendatahandbook.org/en/>>.

<sup>16</sup> LUCA DE BIASE, *Partecipazione dal basso per salvare i beni comuni*, "Nòva", domenica 26 febbraio 2012, <<http://www.ilsole24ore.com/art/tecnologie/2012-02-26/partecipazione-basso-salvare-beni-082142.shtml?uuid=AaeqcuxE>>.

<sup>17</sup> <<http://bicincitta.com/>>.

<sup>18</sup> <<http://progettokublai.ning.com/group/health20initaly>>.

<sup>19</sup> <<http://www.wikibooks.org/>>.

<sup>20</sup> <<http://www.openpolis.it/>>.

<sup>21</sup> <[http://it.wikipedia.org/wiki/Open\\_knowledge](http://it.wikipedia.org/wiki/Open_knowledge)>.

<sup>22</sup> <<http://www.khanacademy.org/>>.

<sup>23</sup> <<http://ocw.mit.edu/index.htm>>.

<sup>24</sup> Per comprendere cosa è un MOOC si veda il video su YouTube di Dave Cormier *What is a MOOC?*, <[http://www.youtube.com/watch?v=eW3gMGqcZQc&feature=player\\_embedded#!](http://www.youtube.com/watch?v=eW3gMGqcZQc&feature=player_embedded#!)>.

<sup>25</sup> <[http://en.wikibooks.org/wiki/Open\\_Metadata\\_Handbook](http://en.wikibooks.org/wiki/Open_Metadata_Handbook)>. Di grande interesse anche il Wiki che elenca i vari standard di metadati con la rispettiva descrizione: <[http://en.wikibooks.org/wiki/Open\\_Metadata\\_Handbook/Metadata\\_Standards#BibJSON](http://en.wikibooks.org/wiki/Open_Metadata_Handbook/Metadata_Standards#BibJSON)>.

<sup>26</sup> <[http://en.wikibooks.org/w/index.php?title=Open\\_Metadata\\_Handbook/Recommendations](http://en.wikibooks.org/w/index.php?title=Open_Metadata_Handbook/Recommendations)>.

<sup>27</sup> <<http://openbiblio.net/principles/it/>>.

<sup>28</sup> <<http://www.dcc.ac.uk/>>.

<sup>29</sup> <<http://www.dcc.ac.uk/resources/how-guides/cite-datasets>>. La guida fornisce esempi pratici che aiutano a risolvere le problematiche che si possono presentare, suggerendo le modalità per superare eventuali ostacoli.

<sup>30</sup> Vedi report "Peer Review in Scientific Publications", U.K. House of Commons Science and Technology Committee, 2011, <<http://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/peer-review/>>.

<sup>31</sup> MARJAN GROOTVELD – JEFF VAN EGMOND, *Peer-reviewed open research data: results of a pilot*, 2011. In 7th International Digital Curation Conference "Public? Private? Personal? navigating the open data landscape", Marriott Royal Hotel, Bristol, UK, 5 - 7 December 2011, <<http://eprints.rclis.org/bitstream/10760/16668/1/Grootveld-vEgmond.pdf>>.

<sup>32</sup> DAVID SHOTTON, *The Five Stars of Online Journal Articles: a Framework for Article Evaluation*, "D-Lib Magazine", January-February 2012, Volume 18, Number 1/2, <<http://www.dlib.org/dlib/january12/shotton/01shotton.html>>.

<sup>33</sup> EDWIN A. HENNEKEN, *Alberto Accomazzi Linking to Data - Ef*

fect on Citation Rates in Astronomy [arXiv:1111.3618v1]. Il case study si riferisce ad articoli pubblicati nel campo dell'astronomia, ma non è da dubitare che, se condotto su altre discipline, possa portare a risultati uguali o migliori.

<sup>34</sup> Per approfondimenti sull'Open Biblio si rimanda al paragrafo 5 di: ANTONELLA DE ROBBIO – SILVIA GIACOMAZZI, *Dati aperti con LODe*, cit.

<sup>35</sup> Per la la metafora della mappa e del territorio vedi Gregory Bateson, introduzione e cura di Marco Deriu, Milano, Bruno Mondadori, 2000, p. 139-140.

<sup>36</sup> Vedi nota 1 di DAVID MCCANDLESS, *The beauty of data visualization*, talk tenuto alla conferenza TED nel luglio 2010.

<sup>37</sup> <<http://obd.jisc.ac.uk/>>. La guida cerca di chiarire in termini generali, ma anche entro specifici contesti, 17 studi di casi concreti di utilizzo di dati che coprono aspetti concreti che ciascuna biblioteca potrebbe già aver incontrato o che potrebbe pianificare per uno sviluppo dei servizi bibliotecari verso l'utenza.

<sup>38</sup> Cfr. RICHARD JONES, MARK MACGILLIVRAY, PETER MURRAY-RUST, JIM PITMAN, PETER SEFTON, BEN O'STEEN and WILLIAM WAITES, *Open Bibliography for Science, Technology, and Medicine*, "Journal of cheminformatics", Volume 3, Number 1, 47, DOI: 10.1186/1758-2946-3-47, <<http://www.springerlink.com/content/g1630211473255n4/>>.

<sup>39</sup> Per un approfondimento sulle licenze open si veda ANTONELLA DE ROBBIO, *OLAOD: Open Licenses for Open Data*, "JLIS.it", 2011, Vol. 2, Issue 2, <<http://leo.cilea.it/index.php/jlis/article/view/4695>>, doi:10.4403/jlis.it-4695.

<sup>40</sup> <<http://data.bnf.fr/>>.

<sup>41</sup> Licenza che prevede la donazione al Pubblico Dominio Creative Commons Universal pubblicata all'indirizzo web: <<http://creativecommons.org/publicdomain/zero/1.0/>>.

<sup>42</sup> <<http://openlibrary.org>>.

<sup>43</sup> <<http://cern.ch/bookdata>>.

<sup>44</sup> <<http://biblios.net>>.

<sup>45</sup> Per ulteriori informazioni si veda il video su YouTube uscito in occasione del lancio del servizio <[www.youtube.com/watch?v=-CSmieTXbs](http://www.youtube.com/watch?v=-CSmieTXbs)>.

<sup>46</sup> Il Data Exchange Agreement, prevede che tutti i metadati presenti in Europeana siano ad accesso aperto e corredati della licenza per il pubblico dominio. Entrerà in vigore dal 1 luglio 2012: <<http://version1.europeana.eu/web/europeana-project/newagreement>>.

<sup>47</sup> Dal post del 1 marzo 2012 di Ilaria Fava su "CIBER Newsletter" di febbraio 2012, <<http://bib04.caspar.it/cibernewsletter/?p=15371>>.

<sup>48</sup> Repository creato da Open Knowledge Foundation Italia e il Centro NEXA su Internet & Società del Politecnico di Torino che offre data sets pubblici di ambito italiano, <<http://ckan.net/tag/library>>.

<sup>49</sup> Il 1. giugno si è svolta la seconda riunione del gruppo di lavoro sui dati bibliografici. Draft della riunione (via chat) del gruppo di lavoro sui dati bibliografici del 1 giugno 2011 sul sito di Open Knowledge Foundation Italia (OKF Italia nasce

nel 2010), <<http://it.okfn.org/2011/06/22/openbiblio-italia-seconda-puntata/>>.

<sup>50</sup> OCLC è un'organizzazione non-profit che agisce a livello mondiale e che fornisce alle biblioteche aderenti servizi per il recupero, la catalogazione e la conservazione dei libri. L'interessante dibattito in merito all'adozione della licenza più idonea ai membri del consorzio può essere seguito a questo link: <<http://blog.okfn.org/2011/06/06/oclc-worldcat-rights-and-responsibilities-and-open-data-licensing>>, mentre le relative policy si possono visionare qui <<http://www.oclc.org/worldcat/recorduse/policy/default.htm>>.

<sup>51</sup> <<http://www.opendatacommons.org/licenses/by/summary/>>. La licenza ODC-BY ha la caratteristica – determinata dall'etichetta BY che significa attribuzione – di richiedere espressamente che l'attribuzione della titolarità dei dati sia sempre mantenuta e rispettata, e quindi risulta perfettamente compatibile con gli obblighi delle clausole contenute nella sezione 3B delle nuove politiche sui dati.

<sup>52</sup> <<http://www.opendatacommons.org/licenses/odbl/>> *Attribution Share-Alike for data/databases*.

<sup>53</sup> Il Comune di Parigi nel 2010 ha votato una risoluzione che approvava l'uso di tale licenza per la diffusione di tutti i dati del Comune.

<sup>54</sup> <[http://wiki.openstreetmap.org/wiki/IT:Pagina\\_Principale](http://wiki.openstreetmap.org/wiki/IT:Pagina_Principale)>.

<sup>55</sup> <<https://it.foursquare.com/>>. Si tratta di un social network basato sulla geolocalizzazione disponibile tramite web e applicazioni per dispositivi mobili, La geolocalizzazione è l'identificazione della posizione geografica nel mondo reale di un dato oggetto, come ad esempio un telefono cellulare o un computer connesso a internet.

<sup>56</sup> Per approfondimenti su Open Government si rimanda all'ottimo saggio di FRANCESCA DI DONATO, *Lo stato trasparente: Linked open data e cittadinanza attiva*, <<http://www.linkedopendata.it/wp-content/uploads/statotrasparente.pdf>>.

<sup>57</sup> Progetto di ricerca realizzato da marco Fioretti del Laboratorio di economia e management della Scuola superiore Sant'Anna di Pisa. Report a cura di Marco Fioretti *Open Data: Emerging trends, issues and best practices*, <[http://www.lem.sssup.it/WPLem/odos/odos\\_2.html](http://www.lem.sssup.it/WPLem/odos/odos_2.html)>.

<sup>58</sup> OpenDataBlog - Il Sole 24 Ore, <<http://opendatablog.ilsole24ore.com/>>. Esiste anche una pagina Facebook correlata al Blog: <<http://www.facebook.com/OpenDataBLOG>>.

<sup>59</sup> "Porte Aperte: i vostri grafici, e studi su database aperti", <<http://opendatablog.ilsole24ore.com/2012/01/porte-aperte>>.

<sup>60</sup> <<http://www.spaghettiopendata.org/>>.

<sup>61</sup> <<http://www.opendatajournalism/>>.

<sup>62</sup> *Battere la povertà con gli open data* è l'articolo di Luca Conte pubblicato sul "Sole 24 Ore" del 4 marzo 2012: <<http://www.ilsole24ore.com/art/tecnologie/2012-03-04/battere-poverta-open-data-081935.shtml?uuid=AbjF821E&fromSearch>>.

<sup>63</sup> Global Pulse (<http://www.unglobalpulse.org/>) ha lo scopo di sostenere la comunità internazionale nel fornire nuove opportunità per stimare il benessere della popolazione, con anali-

si di indicatori in tempo reale, per monitorare come le comunità rispondono agli effetti delle crisi, contribuendo a migliorare l'efficienza delle politiche di protezione, ottimizzando la spesa pubblica. È focalizzata in tre ambiti di attività: la ricerca su dati e indicatori relativi al benessere sociale, lo sviluppo di tecnologie collaborative per l'analisi dei dati e una rete di laboratori in grado di catalizzare la ricerca in azioni di supporto e di diffondere buone pratiche.

<sup>64</sup> *Battere la povertà con gli open data* è l'articolo di Luca Conte pubblicato sul "Sole 24 Ore" del 4 marzo 2012: <<http://www.ilsole24ore.com/art/tecnologie/2012-03-04/battere-poverta-open-data-081935.shtml?uuid=AbjF821E&fromSearch>>.

<sup>65</sup> <<http://data.worldbank.org/data-catalog>>; Climate <<http://sdwebx.worldbank.org/climateportal/>>; Databank <<http://data.worldbank.org/indicator>>.

<sup>66</sup> Roomware è un framework open source fondato nel 2006 da James Burke, Robert Gaal e Tijs Teulings per realizzare spazi interni interattivi entro una rete globale di tipo peer-to-peer, usando tecnologie Bluetooth e RFID, <<http://roomware.org/>>.

<sup>67</sup> VURB ha sede ad Amsterdam, <<http://www.vurb.eu/>>.

<sup>68</sup> <<http://www.lifesized.net/2009/11/08/open-government-data-poster/>>.

<sup>69</sup> "Dati.gov.it" è progettato e sviluppato da Formez PA, su indicazione del Ministero per la pubblica amministrazione e l'innovazione: <<http://www.datagov.it/>>.

<sup>70</sup> <<http://dati.camera.it/>>.

<sup>71</sup> <<http://dati.piemonte.it/>>.

<sup>72</sup> <<http://dati.emilia-romagna.it/>>.

<sup>73</sup> <<http://dati.istat.it/>>.

<sup>74</sup> <<http://data.cnr.it/>>.

<sup>75</sup> <<http://data.enel.com/>>.

<sup>76</sup> <[http://www.comune.fi.it/opencms/opencms/amm/atti\\_e\\_documenti/open\\_data/index.html](http://www.comune.fi.it/opencms/opencms/amm/atti_e_documenti/open_data/index.html)>. Il Comune di Firenze ha avviato un percorso di condivisione dei propri progetti con la cittadinanza, mediante la metafora dei 100 luoghi e dal 2006 sta portando avanti un processo di razionalizzazione e bonifica del proprio patrimonio informativo. Come in altre istituzioni i dataset pubblicati sono pensati per essere utilizzati da professionisti del mondo GIS (sistemi informativi geografici), da sviluppatori software e da esperti in elaborazione dati, mentre i dati ad esclusivo scopo informativo sono consultabili negli altri canali della rete civica, le informazioni sulla mobilità o i dati statistici.

<sup>77</sup> <[www.dati.gov.it](http://www.dati.gov.it)>.

<sup>78</sup> <<http://www.formez.it/iodyl/>>.

<sup>79</sup> <<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1524&format=HTML&aged=0&language=IT&guiLanguage=en>>.

<sup>80</sup> <<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1524&format=HTML&aged=0&language=IT&guiLanguage=en>>.

<sup>81</sup> <[http://ec.europa.eu/research/horizon2020/pdf/proposals/com%282011%29\\_808\\_final.pdf](http://ec.europa.eu/research/horizon2020/pdf/proposals/com%282011%29_808_final.pdf)>.

<sup>82</sup> Strategia EU 2020: Open data An engine for innovation, growth and transparent governance, <<http://eur-lex.europa.eu>>.

<sup>83</sup> Le parole sono del sindaco di Vancouver *David Eaves* (*da*: <<http://opendatachallenge.org/>>).

## ABSTRACT

Data is not only the "new oil of knowledge" but also a new fertile soil which we can water with networks and connectivity. In other terms data is the new soil because they are created from one side by research communities and from the other by public information sector and if data are really open and interconnected, they become a fertile ground that creates not only intellectual wealth, but even the basis for economic recovery.

Any restriction on data and on its re-use limits any progress of research and communities development. Between Open Science and Open Government, we can place a third place that connects them both and that concerns the meta information. This middle land is called Open Bibliographic Data or OpenBiblio. This virtual place includes OPACs, databases and open archives. Speaking of Open Data means to place the defining elements within the framework of the Open Knowledge, in a dimension of knowledge that must analyse also the issue of Open Data licenses.