

Torino, 09 Aprile 2015



Open data tra diritto e tecnologia

Una lezione per capire cosa sono i dati aperti, a
cosa servono e come si fa Open data

Andrea Violato

violato.a@gmail.com
andrea.violato@nomotila.it

Chi sono?

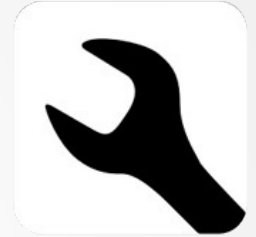
- Mi occupo di Open data, ontologie (giuridiche), compliance
- Lavoro in Nomotika (spin-off di Informatica Giuridica)
- Sono un giurista... come voi
- Le mie e-mail:
 - violato.a@gmail.com
 - andrea.violato@nomotika.it

Indice degli argomenti

- Open data, una introduzione
 - Cosa sono i dati aperti?
 - A chi/cosa servono?
 - Un po' di storia...
 - Inquadramento normativo
- Come fare Open data (#1)? Questioni di 'diritto'
 - Diritto d'autore: copyright, copyleft e licenze
 - Licenze aperte applicabili in Italia
- Come fare Open data (#2)? Questioni di 'tecnologia'
 - Formati aperti: CSV, XML, RDF, ..., Open Document Standard
 - Come pubblicare i dati?
 - Linked Open Data (LOD)
- Cenni sulle 'Linee Guida per l'interoperabilità semantica attraverso i Linked Open Data' di AgID
- Quando fare Open data?
 - Storie di successo ed #EpicFail
 - Un 'calcolo' per gli Open data

Open data, una introduzione

- **Cosa sono i dati aperti?**



Corrente di pensiero - movimento sociale - stile di vita che perora la causa dell'apertura, del (ri)uso (anche commerciale) e della diffusione libera ed incondizionata dei dati di 'dominio collettivo' detenuti dal settore pubblico

I dati aperti possiedono una dimensione economica, etico-sociale, culturale

Open data, una introduzione

- **A chi/cosa servono?**

- 1) Cittadini → Trasparenza, partecipazione
- 2) Imprese → 'New Economy', meritocrazia
- 3) Professionisti (web e non solo) → Creazione di nuovi servizi 2.0
- 4) Giornalisti, studiosi, analisti → 'Approfondimento', indagine
- 5) ... lo stesso settore pubblico → 'Check-up' completo, contrasto a fenomeni 'tumoriali'

Open data, una introduzione

- **Un po' di storia**

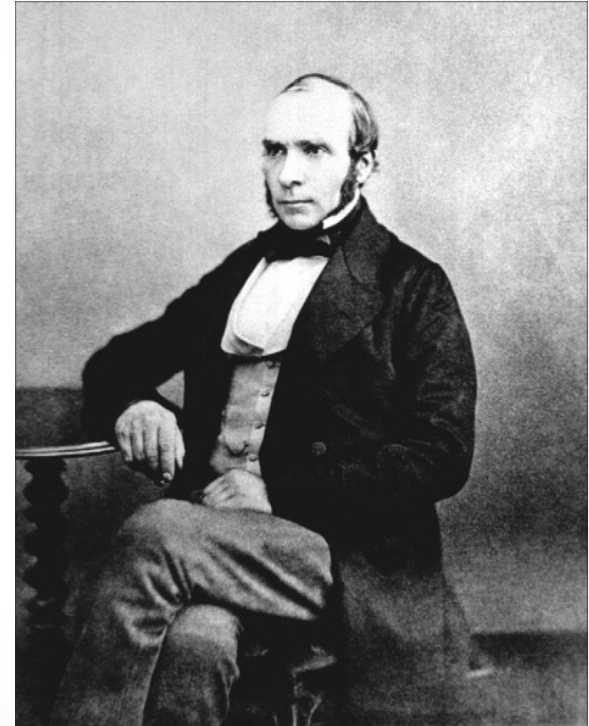
1) 'Doc' John Snow (1850)

2) Barak Obama (2009)

3)

4) Dati Piemonte (2010)

5) Dati.gov (2011)



Open data, una introduzione

- **Inquadramento normativo**

Europa

- Direttiva 2003/98/CE del Parlamento europeo e del Consiglio, approvata il 17 novembre 2003 e pubblicata nella GUCE n. L 345 del 31 dicembre 2003; «accesso e riutilizzo»
- Proposta di modifica alla direttiva europea 2003/98/CE sull'opendata (26 giugno 2013); proposta estensiva; entrerà in vigore nel 2015; «rimedio al diniego»

Italia

- Decreto legge "Misure urgenti per l'innovazione e la crescita: agenda digitale e startup" (Decreto Crescita 2.0)
 - Modifiche al CAD (art. 52 su Accesso e riutilizzo; art. 68 su Dati pubblici aperti) e previsione di standard tecnici
 - Linee guida di AgID
- Quadro normativo definito dal Vademecum Open Data
 - Beta dal 2011
 - Prevede come applicabili normative «analogiche»

Come fare Open data (#1)? Questioni di 'diritto'

- **Diritto d'autore: copyright, copyleft e licenze**

“Tutela del dato”, “titolarità del dato”, “licenze per dati” non sono la stessa cosa

Oggetto del diritto d'autore non è l'informazione in sé, ma la modalità 'ingegnosa' attraverso la quale si estrinseca il dato

Il copyright riguarda la struttura organizzativa con la quale è data una forma all'informazione, non l'informazione stessa

Diritto d'autore è, dunque, l'insieme di diritti (moralì, di disposizione economica, 'connessi') su opere d'ingegno

Come fare Open data (#1)? Questioni di 'diritto'

- **Diritto d'autore: copyright, copyleft e licenze**

Cenni su L. 633/1941 (legge sul Diritto d'autore)

- Artt. 1 + 2 → estensione della disciplina 'analogica' anche a programmi per elaboratore e basi di dati
- Art. 5 + 11 → riconoscimento di una 'proprietà intellettuale' all'ente pubblico. I 'testi degli atti ufficiali' rappresentano un doppio problema definitorio:
 - Atti ufficiali → art. 11 lascia poche speranze, in apparenza
 - Testi → un XML?
- Art. 6 → applicazione della definizione di 'titolare' anche nella creazione di un software (è colui che finanzia lo sviluppo)
- Diritto '*sui generis*' (EU) in capo al costituente di un database (diverso da creatore). Le tempistiche variano se il prodotto è:
 - 'creativo' (70 anni)
 - 'compilativo' (15 anni) → 'atti di rinuncia' in licenze di Public Domain

Come fare Open data (#1)? Questioni di 'diritto'

- **Diritto d'autore: copyright, copyleft e licenze**

Licenze: modalità attraverso le quali si decide 'come' condividere il dato o le strutture di dati con le quali sono raccolti:

- Copyright ©
- Open licensing
 - no Copyleft (no opere derivate)
 - Copyleft debole (GPL – Mozilla License)
 - Copyleft forte (opere derivate devono avere la *stessa* licenza dell'opera originaria)
- Public Domain (70-15 + 'atto di rinuncia')

Come fare Open data (#1)? Questioni di 'diritto'

- **Licenze aperte applicabili in Italia**

- 1) Italian Open Data License - <http://www.dati.gov.it/iodl/2.0/>
- 2) Creative commons - <http://creativecommons.org/> - <http://www.creativecommons.it/>
- 3) Open Data Common Database License - <http://opendatacommons.org/licenses/odbl/> (OpenStreetMap)



	CC0	CC-BY	CC-BY-SA	CC-BY-SA-NC	CC-BY-ND	IODL 1.0	IODL 2.0	ODbL	ISA Open Meta data 1.0
Portabilità interlingua	X	X	X	X	X			X	
Riconoscibilità internazionale del logo	X	X	X	X	X			X	
Uso per fini commerciali	X	X	X				X	X	X
Uso gratuito del lavoro soggetto alla licenza	X	X	X	X	X	X	X	X	X
Possibilità di modificare i dati	X	X	X	X		X	X	X	X
Possibilità di preservare la paternità		X	X	X	X	X	X	X	X
Compatibilità interlicenza per lavori derivati	X	X					X		

Come fare Open data (#2)? Questioni di 'tecnologia'

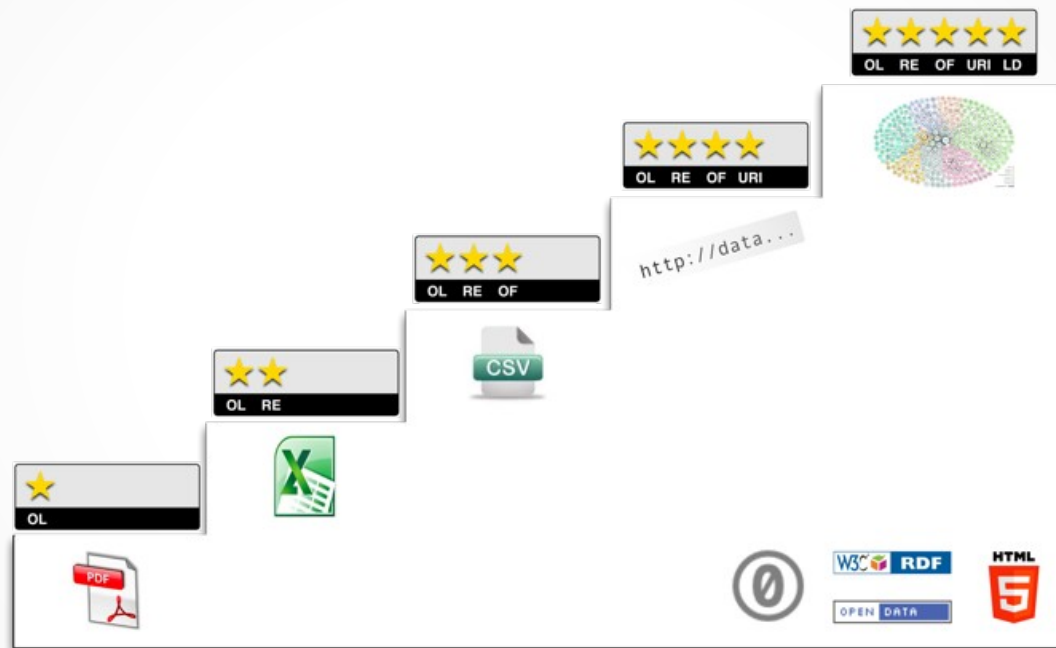
- **Formati aperti: CSV, XML, RDF, ..., Open Document Standard**

Il dato aperto è la base sulla quale costruire una infrastruttura tecnologica (non necessariamente IT) mediante la quale far circolare l'informazione

Il formato nel quale il dato è rilasciato deve essere anch'esso aperto, nel senso che:

- Non deve obbligare chi vuole accedere e (ri)usare il dato ad acquistare tecnologie proprietarie (es: al posto del pacchetto MSOffice, meglio usare LibreOffice)
- Deve essere 'interoperabile'
- Deve seguire le 'best practices' per agevolare la pubblicazione, l'accesso e il riuso del dato (es: predisposizione 'naturale' di API)

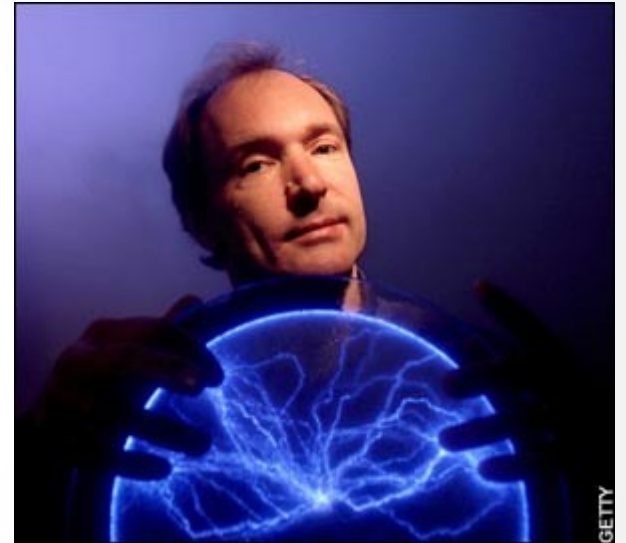
Come fare Open data (#2)? Questioni di 'tecnologia'



Come fare Open data (#2)? Questioni di 'tecnologia'

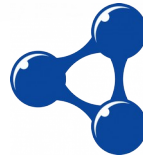
- **Come pubblicare i dati aperti?**

- 1) URI per identificare la risorsa
- 2) HTTP come protocollo di comunicazione
- 3) RDF (+ SPARQL) per la presentazione/pubblicazione dei dati
- 4) Creazione dei link tra le varie URI → Linked Open Data

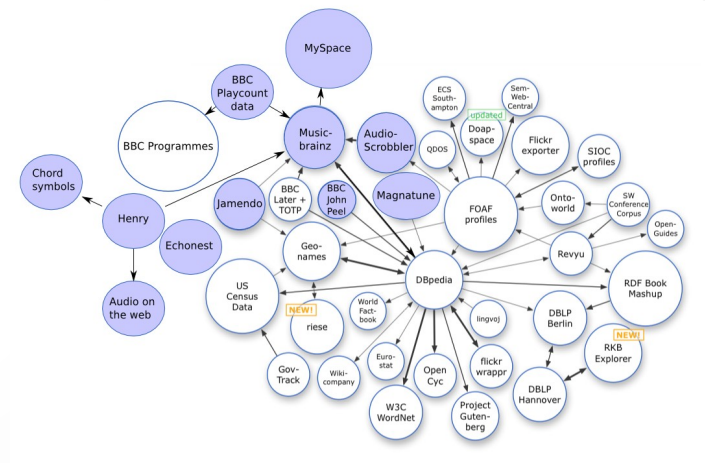


Come fare Open data (#2)? Questioni di 'tecnologia'

- **Linked Open Data (LoD)**



Letteralmente 'dati aperti collegati', è il massimo valore aggiunto attribuibile a singoli Open data che, una volta raccolti e congiunti secondo criteri di coerenza e sistematicità, sono capaci di trasmettere 'nuove' informazioni



Cenni sulle "Linee guida per l'interoperabilità semantica attraverso i linked open data" di AgID

- Prevedono standard tecnici, modelli di riferimento e indicazioni operative
- Definiscono, anche implicitamente, nuovi processi interni o definiscono quali modifiche devono essere apportate a quelli esistenti per:
 - Creare dati aperti
 - Raccogliere dati aperti
 - Mettere a disposizione dati aperti
- Sanciscono ufficialmente che un dato è aperto se è «interoperabile»
 - XML
 - Ontologie
 - Linked Open Data

Quando fare Opendata?

- **Storie di successo e clamorosi flop**

1) Ebola Tracking: <http://ebolatracking.org/> → UP

2) Soldi Pubblici: <http://soldipubblici.gov.it/it/home> → MAH!

3) Corte di Cassazione (ITA):

<http://www.italgiure.giustizia.it/sncass/> → LOL

4) Taxi and Limo Commission (2014, NYC) → #EpicFail

Quando fare Opendata?

- Un "calcolo" per gli Opendata

$$P \times B + D > C$$

[Arnaud Sahuguet, David Sangokoya, 2015]

- 1) P = probabilità che l'apertura del dato provochi un qualche effetto (positivo)
- 2) B = il beneficio che il dataset riceve dalla sua apertura
- 3) D = impatto globale sull'ecosistema
- 4) C = il costo complessivo dell'apertura

Se i Costi sono inferiori alla somma tra il prodotto di Probabilità e Beneficio con l'Impatto globale, allora SI DEVE fare Open data

Lecture

- Open Data Handbook → <http://opendatahandbook.org/>
- Linked Open Data → <http://linkeddata.org/>
- A 'calculus' for Open data → <https://medium.com/@sahuguet/p-b-d-c-1218ee894400>
- Linee Guida AgID → http://www.agid.gov.it/sites/default/files/documentazione_trasparenza/cdc-spc-gdl6-interoperabilitasemopendata_v2.0_0.pdf
- ... altro (su richiesta)