

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Triennale in Informatica per il Management

**Open Data:
I nuovi modelli di
business**

Tesi di Laurea in Base di Dati

**Relatore:
Chiar.mo Prof.
Danilo Montesi**

**Presentata da:
Valentina Dallolio**

**Sessione II
Anno Accademico 2011/2012**

Indice

1	Introduzione	5
2	Open Data: definizioni e concetti base	7
2.1	Public Sector Information e Open Government	7
2.2	Open Knowledge Foundation e la definizione di Open	9
2.3	Cosa sono i “Data”? Quali “Data?”	10
2.4	Open Data Definition	11
2.5	Linked Open Data: Raw Data e Linked Data	11
2.6	Closed Data vs. Open Data	13
3	Nascita dell’Open Data: come e dove	15
3.1	Dall’Open Source e Free Software all’Open Data	15
3.2	Open Data nel mondo e in Italia	16
3.3	Open Data Institute	19
3.4	Le Pubbliche Amministrazioni e le Aziende	22
4	Aspetti tecnici e legali: creare e diffondere i dati	25
4.1	Trasparenza e Privacy	25
4.2	Copyright e Licenze Aperte	27
4.3	Il formato dei dati	29
4.4	Pubblicazione	30
5	Benefici, ostacoli e valore economico dell’Open Data	31
5.1	Benefici	31
5.1.1	Alcuni Esempi	34
5.1.2	Caso di studio: Come l’Open Data ha salvato al Canada 3,2 miliardi di dollari	35
5.2	Costi, svantaggi e Ostacoli	37
5.3	Perchè fare Open Data?	38
5.4	Open Data come strategia economica e nuove opportunità	41
5.4.1	Caso di studio: I dati sul trasporto pubblico	45

6	I Nuovi Modelli di business degli Open Data	51
6.1	Apps e nuovi servizi di informazione per i cittadini	53
6.1.1	Caso di studio: Pocket Ranger - Nuova Zelanda	56
6.2	Pulizia e organizzazione dei dati	59
6.3	Data mining	61
6.4	Data broking e sviluppo di APIs	62
6.5	Efficienza operativa o ottimizzazione	65
6.6	Prodotti finanziari	66
6.7	Customer Attraction e Customer Retention	68
6.8	Supporto delle attività principali	68
6.9	Servizi di supporto all'apertura dei dati	71
6.10	Freemium	71
6.11	Dual Licensing	72
6.12	Sponsorizzazione	72
6.13	Charging for Changes	73
6.14	L'aumento della qualità attraverso la partecipazione	73
7	Conclusioni	75
8	Bibliografia ed elenco delle figure	77

Capitolo 1

Introduzione

L'Open Data e l'Open Source Software sono due delle principali aree di sviluppo per i governi, le organizzazioni e gli individui. Gli obiettivi principali sono: condividere le informazioni liberamente, risparmiare denaro e dar vita a nuove opportunità per creare nuovi prodotti o servizi a beneficio delle imprese o delle società in generale.

L' Open Data è un movimento internazionale ed un insieme di pratiche secondo cui l'accesso ai dati pubblici (vale a dire di interesse pubblico e raccolti con soldi pubblici) deve essere libero e gratuito per tutti, senza restrizioni che ne limitino la riproduzione e l'accesso.

Vediamo quali sono i concetti base che fanno da contorno all'Open Data, le definizioni principali e le caratteristiche dei dati aperti.

Descriviamo la nascita e la diffusione dell'Open Data, dove vengono utilizzati e pubblicati i dati aperti e gli aspetti tecnici e legali con cui vengono distribuiti. Evidenziamo i benefici, gli svantaggi e gli ostacoli della pratica dell'Open Data.

Infine ci dedichiamo all'analisi degli aspetti economici dell'Open Data, analizzando tutti i possibili modelli di business attuali e futuri.

La pratica dell'Open Data è iniziata solamente da pochi anni e si sta diffondendo così velocemente da suscitare interesse in tutto il mondo. Il tema è assolutamente nuovo e tecnologico, tanto che, attualmente, le maggiori informazioni sono reperibili solo nel Web con pochissimi riferimenti cartacei.

Capitolo 2

Open Data: definizioni e concetti base

Al fine di capire cosa sono, come e dove vengono utilizzati i dati aperti è necessario spiegare alcuni concetti base e definizioni fondamentali che fanno parte del mondo dell'Open Data.

2.1 Public Sector Information e Open Government

Il **Public Sector Information (PSI)** genera, gestisce e usa i dati che vengono raccolti dalle Pubbliche Amministrazioni. I dati che le PA producono, raccolgono, organizzano e utilizzano nello svolgimento dei loro compiti istituzionali hanno un valore intrinseco che va oltre l'assolvimento dei compiti istituzionali dell'ente che le detiene/gestisce. L'informazione detenuta dal settore pubblico rappresenta un patrimonio immenso in ogni ambito - dalla cartografia, alla normativa, alla statistica - fatto di informazioni, dati e contenuti.

L'**Open Government** rappresenta un modello di amministrazione aperto e trasparente da parte degli enti e delle istituzioni pubbliche verso i cittadini. Ha come tratti distintivi:

- la centralità del cittadino
- un'amministrazione partecipata e collaborativa
- la trasparenza intesa come apertura dei dati e delle informazioni

- la condivisione dei dati attraverso le nuove tecnologie digitali (Internet e Web)

L'Open Data, per l'accesso e l'utilizzo in rete di dati e risorse pubblici, è necessario per lo sviluppo dell'Open Government all'interno delle amministrazioni centrali e locali. A Novembre 2010 è stato presentato il "Manifesto per l'Open Government" che elenca gli 11 modi per fare Open Government:

1. Governare con le persone.
2. Governare con la rete.
3. Creare un nuovo modello di trasparenza.
4. Trattare l'informazione come infrastruttura.
5. Liberare i dati pubblici per lo sviluppo economico del terzo millennio.
6. Informare, coinvolgere, partecipare per valorizzare l'intelligenza collettiva.
7. La rete moltiplica il potenziale delle intelligenze coinvolte e aumenta l'efficacia dell'azione amministrativa.
8. Educare alla partecipazione.
9. Promuovere l'accesso alla Rete.
10. Costruire la fiducia e aumentare la credibilità della pubblica amministrazione.
11. Promuovere l'innovazione permanente nella pubblica amministrazione.

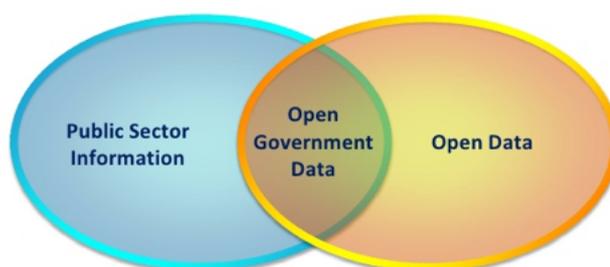


Figura 2.1: Relazione tra PSI, Open Government e Open Data

2.2 Open Knowledge Foundation e la definizione di Open

OKFn è un gruppo di lavoro che si prefigge come scopo quello di promuovere la creazione, la condivisione e l'applicazione di Open Knowledge (*informazioni aperte*) che ognuno può usare, condividere e modificare, nell'era del Web. OKFn ha individuato le caratteristiche affinché un lavoro possa definirsi "Open" e quindi rientrare nella più ampia definizione di Open Data. Sono stati individuati 11 punti/condizioni di cui gli ultimi 5 estrapolati dalla Open Source Definition (che contiene a sua volta le 10 caratteristiche che un software deve avere affinché possa essere considerato Open Source):

1. **Accesso:** l'opera deve essere disponibile in una forma conveniente e modificabile.
2. **Ridistribuzione:** la licenza deve permettere di ridistribuire l'opera modificata o arricchita.
3. **Riuso:** la licenza deve permettere di modificare o utilizzare l'opera.
4. **Assenza di restrizioni tecnologiche:** l'opera deve essere accessibile senza nessuna forma tecnologica che ne ostacoli la disponibilità.
5. **Attribuzione:** la licenza può imporre di specificare i collaboratori e i creatori dell'opera.
6. **Integrità:** la licenza potrebbe prevedere che l'opera modificata venga distribuita con un nome diverso o un altro numero di versione rispetto all'originale.
7. **Nessuna discriminazione verso singoli o gruppi di persone.**
8. **Nessuna discriminazione verso i settori di applicazione** (es. in a business or for genetic research).
9. **La licenza deve essere distribuibile.**
10. **La licenza non deve essere specifica a un pacchetto:** i diritti connessi ad un'opera non devono dipendere dal fatto che essa è parte di un determinato pacchetto/prodotto.
11. **La licenza non deve impedire la distribuzione di altre opere.**

2.3 Cosa sono i “Data”? Quali “Data?”

Con “Data” si intende ogni singolo pezzo di informazione di qualsiasi tipo (dalle immagini ai numeri, definizioni testuali, mappe, audio, ecc.) che:

- sono descrizioni dirette di fatti (es. le temperature medie di una città, le fasce di reddito in alcuni paesi, il percorso di un fiume tracciabile su una mappa, ...) o sono strettamente legate a fatti, quindi non sono copiabili;
- sono riproducibili senza ambiguità quando il metodo usato per generarli è conosciuto in tutti i suoi dettagli;
- sono parti, o possono essere immediatamente usati come parte, di informazioni più grandi;
- hanno (quasi sempre) molto più significato e valore quando sono legati tra loro e completati da “metadata” (dati riguardanti altri dati, per esempio il nome dell’autore o la data di creazione);
- possono, grazie alle caratteristiche sopra descritte, essere espressi e archiviati in formati digitali, anche se non originariamente generati sotto questa forma.

Secondo la filosofia Open Data vedremo che i dati più diffusi sono: dati scientifici (es. dati geografici, medici, genoma, composti chimici, ...) o dati governativi che sono raccolti dalle Pubbliche Amministrazioni (es. censimenti) o che riguardano le Pubbliche Amministrazioni (es. bilanci).



Figura 2.2: Tipi di dati

2.4 Open Data Definition

Per Open Data si intende il movimento che promuove la libera distribuzione e il libero riutilizzo dei dati pubblici senza alcuna restrizione.

La definizione tecnica che OpenDefinition.org dá di Open Data è la seguente: *“Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and share-alike.”*

I dati aperti sono dati che devono essere:

- accessibili (preferibilmente via Internet) a non meno del costo di riproduzione, senza limitazioni basate sull'identità o sull'intenzione dell'utente;
- in formato digitale e leggibile dalla macchina per l'interoperabilità con altri dati;
- liberi da restrizioni sull'uso o sulla redistribuzione nelle loro condizioni di licenza.

2.5 Linked Open Data: Raw Data e Linked Data

I **Raw Data** sono dati grezzi, atomici, non elaborati, disorganizzati e disaggregati. È difficile analizzarli, identificare tendenze e fare considerazioni complessive. I dati organizzati ed interpretati sono più facili da comprendere e da ricordare, ma nascondono i dati di partenza, rendendo difficile la verifica o la creazione di analisi e raggruppamenti diversi da quelli forniti. Una buona pratica dell'Open Data è quella di rendere pubblici i “raw data” e non i dati già analizzati ed elaborati.

I **Linked Data** sono un metodo di esporre, condividere e connettere dati attraverso URI deferenziabili sul Web.

L'idea dietro ai Linked Data è quella di “hyperdata links” che permetteranno alle persone o alle macchine di trovare dati correlati sul Web che non erano precedentemente “linked”.

Tim Berners-Lee ha descritto quattro principi sui dati collegati:

1. Usare gli URI per identificare gli oggetti.
2. Usare HTTP URI in modo che questi oggetti possano essere cercati e reperiti da persone ed applicazioni web.

3. Fornire informazioni utili sull'oggetto quando il suo URI è dereferenziato (trovato), usando formati standard come RDF/XML.
4. Includere collegamenti ad altri oggetti nei dataset esposti (con lo stesso meccanismo) per aumentare e migliorare il reperimento di altre informazioni correlate nel web.



Figura 2.3: Linked Open Data

Il modello **Linked Open Data** propone un approccio tecnologico e metodologico per collegare tra loro gli Open Data, per confrontarli con altre informazioni da risorse differenti e renderli parte di un unico spazio informativo globale e condiviso.

L'interoperabilità è uno dei vantaggi piú importanti del modello Open Data. I dati, se isolati, hanno poco valore; viceversa, il loro valore aumenta sensibilmente quando dataset differenti, prodotti e pubblicati in modo indipendente da diversi soggetti, possono essere incrociati liberamente fra loro (mashup) da terze parti. Sono importanti anche per una strategia di Open Government.

L'obiettivo del progetto *Linking Open Data* del W3C è di estendere il Web pubblicando diversi open dataset, come RDF, sul Web e impostando link RDF tra tutti i dati provenienti da risorse differenti.

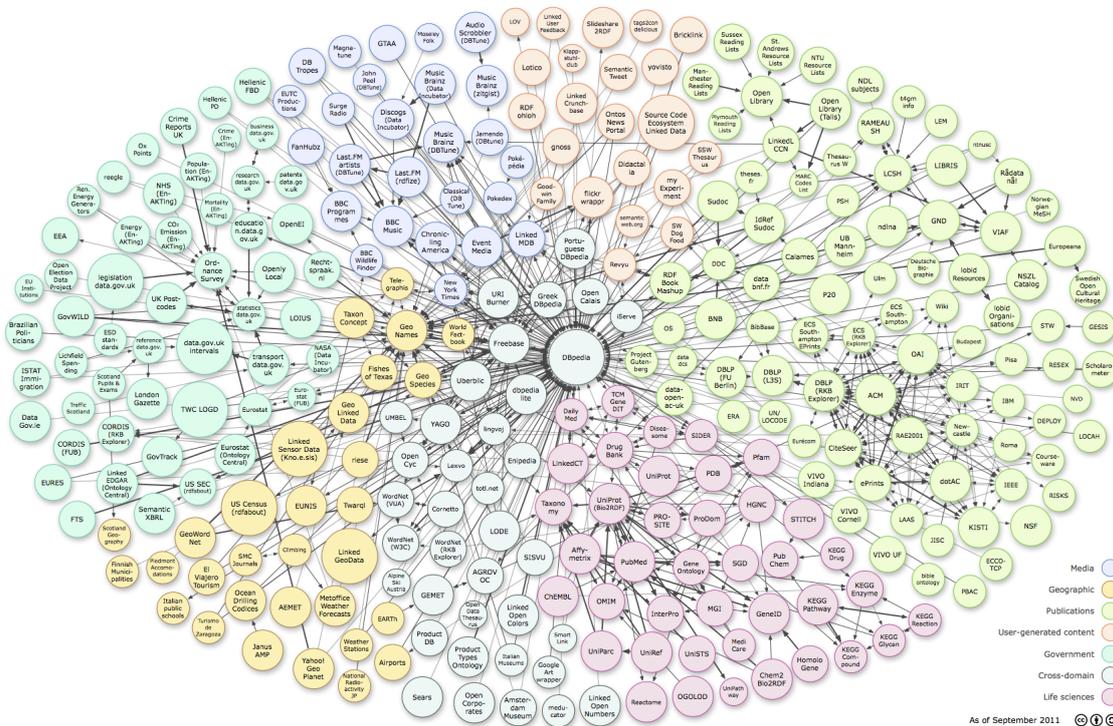


Figura 2.4: Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>

2.6 Closed Data vs. Open Data

Fanno parte dei **Closed Data** tutti i dati che non sono liberamente accessibili o riutilizzabili. Diversi meccanismi impediscono l'accesso e il riutilizzo dei dati, ad esempio: licenze, copyright, brevetti, barriere tecnologiche, costi, ecc. Praticando una politica di Closed Data si crea:

- una situazione di privilegio per chi riesce ad accedere ai dati e un rischio maggiore di corruzione politica e commerciale;
- un profitto economico per chi li distribuisce e una chiusura verso la società;
- valore commerciale;
- tempi più lunghi e difficoltà nella raccolta e l'analisi dei dati.

Gli **Open Data** si contrappongono ai Closed Data in quanto sono accessibili a tutti e possono essere riutilizzati e ridistribuiti liberamente. In questo modo si crea:

- fiducia verso le PA e lo Stato, collaborazione e partecipazione di tutti;
- apertura e zero costi;
- valore sociale;
- un'immediata disponibilità di informazioni.

Capitolo 3

Nascita dell'Open Data: come e dove

Gli scopi del movimento Open Data sono simili a quelli dei movimenti “Open” come Open Source, Open Content, e Open Access. La filosofia dietro gli Open Data esiste da tempo, ma il termine “Open Data” è recente. Questo termine ha guadagnato popolarità con la nascita di Internet e del World Wide Web e, specialmente, con il lancio delle iniziative di Open Data Government come Data.gov.

3.1 Dall'Open Source e Free Software all'Open Data

Open Source, Free Software e Open Data condividono la necessità di rendere libero l'accesso alle informazioni:

Open Source e Free Software: libero accesso al codice sorgente dei software e quindi anche alle loro informazioni.

Open Data: libero accesso ai dati pubblici nel loro formato grezzo.

Il Free Software si propone di garantire 4 libertà:

- Libertà di eseguire il programma per qualsiasi scopo.
- Libertà di studiare il programma e modificarlo.
- Libertà di ridistribuire copie del programma in modo da aiutare il prossimo.
- Libertà di migliorare il programma e di distribuire pubblicamente i miglioramenti in modo tale che tutta la comunità ne tragga beneficio.

Queste sono le stesse libertà e gli stessi principi per cui nasce l'Open Data.

Richard Stallman, fondatore del movimento Free Software, ha sottolineato che *Open Source è un metodo di sviluppo del software; Free Software è un movimento sociale.*

L'Open Data può essere visto a sua volta come un metodo per gestire i dati pubblici, per garantirne l'accesso e la loro libera utilizzazione. Allo stesso tempo ha un risvolto etico in quanto si prefigge di ottenere, così come il software libero, una libera circolazione delle idee, scambio di informazioni, conoscenza, libertà di pensiero e di espressione e benefici per la comunità.

In base a tutto quello detto finora possiamo azzardare anche la definizione di Free Data diversa da quella di Open Data, ricordando che il movimento Free Data nasce in contrapposizione alla creazione di database proprietari, così come lo sviluppo di software proprietario aveva portato alla nascita del movimento Free Software.

Free Data Movimento che presta maggiore attenzione agli aspetti filosofici ed etici dell'accesso libero ai dati, così come è per il Free Software. Si fonda maggiormente sul concetto di “gratuito” (ad esempio, nell'editoria, il NYT online dal 2007 è gratuito).

Open Data Movimento che presta maggiore attenzione agli aspetti pratici, così come è per l'Open Source. I dati possono essere riutilizzati anche per scopi commerciali o per ottenere vantaggi economici.

3.2 Open Data nel mondo e in Italia

Una serie di data store governativi sono stati aperti negli ultimi anni in tutto il mondo. Il primo e più celebre, che ha sostanzialmente fatto da modello ai successivi, è stato Data.gov americano, lanciato dal governo Obama a seguito della Direttiva sull'Open government nel dicembre 2009 come misura anti-crisi, per permettere di produrre beni e servizi attraverso i dati aperti intesi come materia prima. Di lì a pochi mesi anche il Regno Unito ha aperto il suo Data.gov.uk fortemente voluto e sponsorizzato da Tim Berners-Lee, l'inventore del World Wide Web, seguito dalla Nuova Zelanda con data.govt.nz. In pochi anni la pratica degli open data e dei data store governativi si è estesa fino in Norvegia con data.norge.no (Aprile 2010), Australia con Data.gov.au (Marzo 2001), Canada con Data.gc.ca (Marzo 2011), Marocco con Data.gov.ma (Aprile 2011), Kenya con opendata.go.ke (Luglio 2011), Cile con datos.gob.cl (Settembre 2011), Olanda con data.overheid.nl (Ottobre

2011), Spagna con datos.gob.es (Ottobre 2011), Uruguay con datos.gub.uy (Novembre 2011), Francia con Data.gouv.fr (Dicembre 2011), Brasile con dados.gov.br (Aprile 2012), India con data.gov.in (2012), Estonia con open-data.ee, Portogallo con dados.gov.pt, Moldavia con date.gov.md, Belgio con data.belgium.be.

Data.gov americano ha attualmente disponibili 378529 dataset, Data.gov.uk ha invece 8741 dataset. Forniscono quindi una grande quantità di dati aperti in formato grezzo e contengono anche articoli, librerie, casi di studio, materiale di formazione, blog, elenchi di applicazioni sviluppate, ecc. Su questo modelli si basano anche tutti i gli altri siti governativi sopra elencati.



Figura 3.1: Data.gov

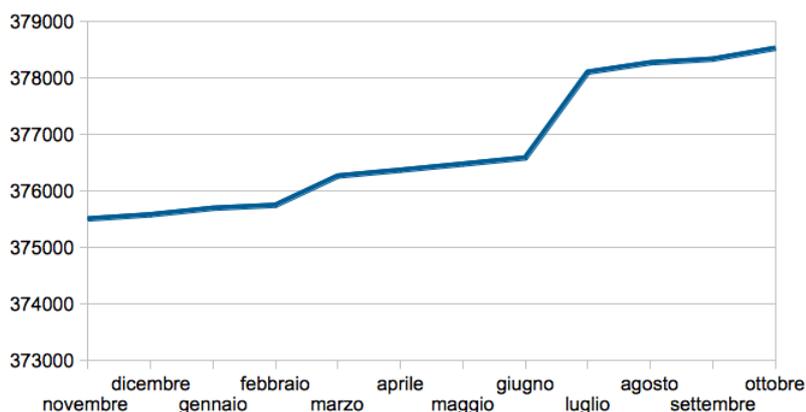


Figura 3.2: 3034 dataset liberati negli Stati Uniti, a partire dal mese di Novembre 2011 fino a Ottobre 2012 compresi

Sono molteplici anche le iniziative d'apertura del patrimonio informativo avviate in **Italia** da parte di pubbliche amministrazioni centrali e locali. Il primo data store italiano è stato quello della Regione Piemonte (2010), dati.piemonte.it, all'interno del quale sono catalogati dati aperti riconducibili ai vari enti regionali (comuni, province, ...). Dopo circa un anno dalla nascita del data store piemontese anche la Regione Emilia-Romagna, ad ottobre 2011, pubblica online il suo catalogo di dataset (dati.emilia-romagna.it). Con il lancio del portale dati.gov.it, avvenuto il 18 Ottobre 2011, si è aperta una nuova stagione per l'innovazione e la trasparenza nella PA, una strada verso l'Open Data italiano. Il processo di gestione delle informazioni prodotte sta progressivamente evolvendo, verso modelli più aperti.

Nell'ordinamento Italiano però non esiste una norma che impone di rendere pubblici i dati in possesso delle PA, come la Direttiva americana lanciata da Obama. La maggior parte degli archivi sono ancora "scatole chiuse".

In Italia, attualmente, sono disponibili oltre 3600 dataset rilasciati in formato aperto.

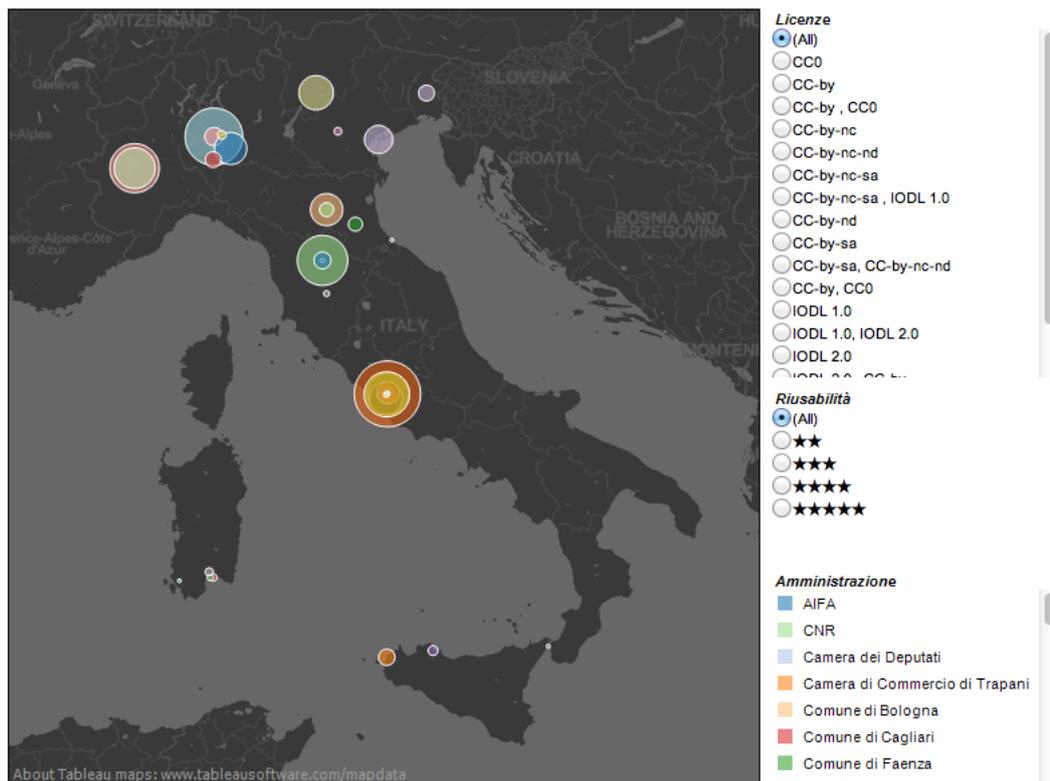


Figura 3.3: Distribuzione geografica delle amministrazioni italiane che rilasciano open data

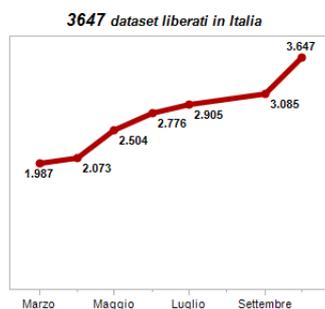


Figura 3.4: Numero totale di dataset liberati in Italia, a partire dal mese di Marzo 2012 fino a Ottobre 2012

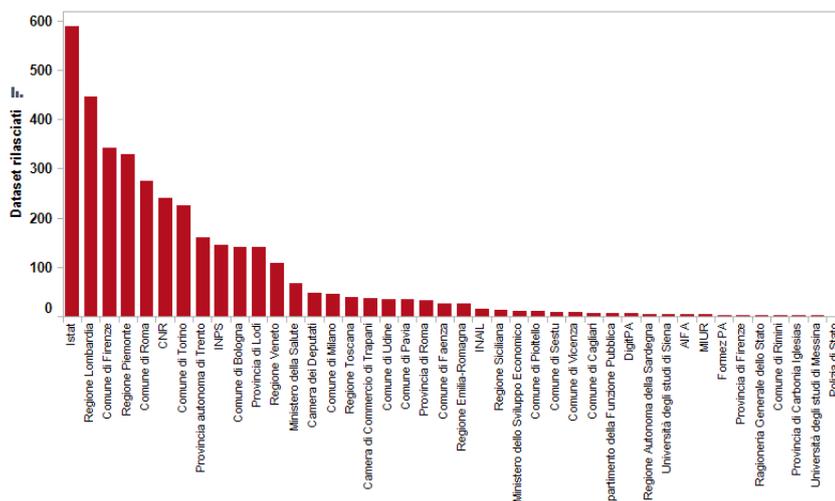


Figura 3.5: Confronto per istogrammi tra le pubbliche amministrazioni italiane che pubblicano i propri dati in formato aperto

3.3 Open Data Institute

A Settembre 2012 è nato, a Londra (Shoreditch), l'Open Data Institute (ODI) che vuole essere il primo centro Open Data nel mondo, sotto la guida di Tim Berners-Lee e del professore Nigel Shadbolt.

L'ODI ha come scopo quello di formare imprenditori, programmatori, tecnici e creativi che possano sviluppare idee innovative con l'utilizzo degli open data. La sua missione può essere riassunta in 6 punti di cui i primi tre

riguardano il core business, incubazione e innovazione, mentre gli altri tre sono volti alla creazione di un ecosistema di supporto:

1. Incubare nuove opportunità di business dove la crescita economica sarà data dall'utilizzo dell'Open Data.
2. Supportare ed espandere l'utilizzo dell'Open Data nel commercio e fornire le competenze necessarie per l'utilizzo dei dati aperti a un numero sempre più grande di persone.
3. Sviluppare nuovi modelli di business.
4. Migliorare la comprensione delle opportunità legate all'utilizzo degli open data nel settore pubblico.
5. Supportare la ricerca e lo sviluppo dei presupposti necessari al fine di sfruttare al meglio i dati aperti.
6. Consolidare la propria presenza e la propria reputazione a livello internazionale attraverso la collaborazione con altri enti governativi e non-governativi in tutto il mondo.

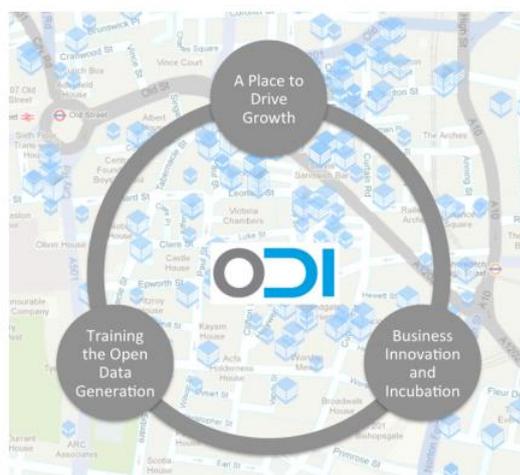


Figura 3.6: ODI

Molte altre informazioni possono essere ricavate dall'ODI Implementation Plan disponibile sul sito dell'istituto.

Inoltre è disponibile l'ODI Business Plan 2012-2017 nel quale vengono identificati 4 clusters: business sector, public sector, mondo accademico (University

sector), interessi all'estero (International Dimension). Questi clusters rappresentano settori complementari e lo scopo dell'ODI è quello di supportare la cooperazione, l'integrazione e la costruzione di un ponte tra loro.

La prima base fisica dell'ODI è in Shoreditch (Londra) intorno all'iniziativa Tech City. Questa posizione permette l'accesso immediato alle varie reti che compongono il fenomeno Silicon Roundabout. Con dalle 600 alle 800 start-ups attive ovunque - delle quali molte cercano di sfruttare e utilizzare l'Open Data - questa rete sarà una parte essenziale per garantire questo talento innovativo che circonda l'ODI. È un insieme di talento che collaborerà con l'ODI e userà gli outputs e le risorse dell'ODI. Queste start-ups, composte da molti giovani laureati, hanno una naturale affinità per la ricerca e l'etica dello sviluppo tipica delle università. Come tale Tech City offre un ambiente adatto per stare a cavallo sia delle imprese sia del mondo accademico. L'ODI provvederà a fornire anche spazi di incontro, formazione e innovazione. L'obiettivo dell'ODI sarà realizzato attraverso quattro principali linee di attività concentrate sui 4 cluster precedentemente elencati (imprese, settore pubblico, università e collaborazione internazionale). Esso comprenderà anche una forte linea di sostegno alla formazione della prossima generazione di Open Data Technologists.



Figura 3.7: I 4 clusters dell'ODI

3.4 Le Pubbliche Amministrazioni e le Aziende

L'Open Data è principalmente rivolto alle Pubbliche Amministrazioni, in quanto sono i più grandi produttori di informazioni e di dati relativi al lavoro, alle imprese, alla cartografia del territorio, alla salute, ai livelli di inquinamento, alla criminalità, ecc. Perciò le PA devono utilizzare le nuove tecnologie per essere aperte e trasparenti verso i cittadini attraverso la distribuzione dei dati, evidenziando così l'efficienza e l'utilità del loro lavoro. Gli obiettivi delle PA sono:

- Aumentare la trasparenza e il coinvolgimento dei cittadini, istituzioni e tutti gli altri stakeholders
- Migliorare il mercato
- Favorire l'innovazione tecnologica, lo sviluppo e la diffusione di nuove tecnologie
- Creare valore sui dati e benefici per i cittadini



Figura 3.8: Open Data, Pubbliche Amministrazioni e Aziende

Per quanto riguarda le Aziende, soprattutto in Italia, sono più restie al rilascio e alla condivisione dei loro dati, a causa disinteressi personali e vantaggi nel non rendere pubblici i propri dati. Nel resto del mondo però alcune aziende hanno iniziato a rendere pubbliche le loro informazioni.

Nel 2010 è stato rilasciato il sito opencorporates.com che ha come scopo idealistico quello di possedere un URL per ogni azienda nel mondo, più materialmente quello di condividere informazioni riguardanti le aziende. Permette di effettuare ricerche su tutte le giurisdizioni, filtrando il tipo di compagnia, lo

stato, il paese e molti altri attributi. Il sito è semplice da usare e facilita il filtraggio e l'accesso dei dati. Inoltre, OpenCorporates, utilizzando il mashup dei dati, ha aiutato molti enti statali ad individuare errori nei loro dati. Inoltre sta ottenendo sempre più informazioni direttamente dai registri delle aziende e ha aiutato molte Companies House a creare il loro nuovo sistema URI per identificare le compagnie. OpenCorporates permette l'accesso diretto ai suoi dati attraverso un' API (Application Programming Interface) aperta, e consente l'utilizzo di strumenti per la pulizia dei dati, come Google Refine, per unire altri dataset alle compagnie. Molti hanno già usufruito di questo sito: giornalisti, organizzazioni come OpenSpending, e anche dipartimenti governativi nel mondo.

Il co-fondatore di OpenCorporates, Chris Taggart, ha affermato, *“In a global world, where companies are actually complex networks of corporate legal entities, it critical that people have not just free access to company data, but the right to freely reuse it and combine it with other data. When financial crises strike, or disasters such as the Gulf Oil spill happen, it takes months or years to unpick the various corporate entities that form the large groups involved. OpenCorporates massively simplifies that job. At the other end of the scale, OpenCorporates makes it trivial to search for cross-border information on small and medium-sized companies information that is often difficult to get from any other source.”*.

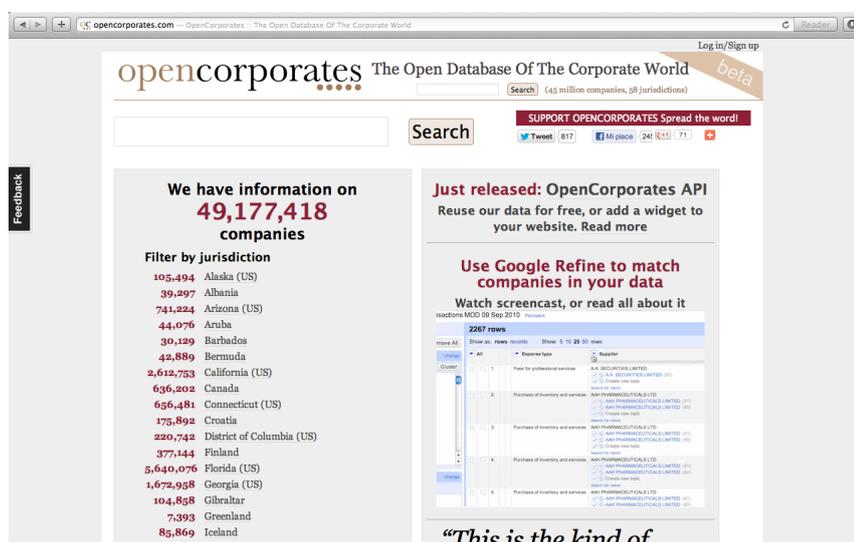


Figura 3.9: OpenCorporates

Un altro esempio simile a OpenCorporates è OpenSpending dell'Open Knowledge Foundation, che ha fatto grandi progressi nel mappare i pagamenti

da parte dei governi alle compagnie, costruendo una piattaforma che estende l'originale Where Does My Money Go, e ora può includere spese governative di ogni tipo.



Figura 3.10: OpenSpending

Capitolo 4

Aspetti tecnici e legali: creare e diffondere i dati

Ora vediamo ed analizziamo alcuni aspetti legali e le principali caratteristiche tecniche dei dati aperti, come vengo creati i datasets e pubblicati i dati.

4.1 Trasparenza e Privacy

La **trasparenza**, o più in generale la disciplina del rapporto dialogico tra Pubblica Amministrazione e cittadino, è stata prevista dalla Legge n. 241/1990 (“Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi”).

La trasparenza è nata come il diritto di accesso agli atti e ai documenti per coloro che avessero specifico e concreto interesse. Con l'introduzione dell'informatica nell'attività amministrativa e il ripensamento sul ruolo del settore pubblico la definizione sopra data non è più sufficiente. Per esempio vediamo i seguenti principi introdotti nel 2005 dal Codice dell'Amministrazione digitale (CAD): *“le pubbliche amministrazioni nell'organizzare autonomamente la propria attività utilizzano le tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di efficienza, efficacia, economicità, imparzialità, trasparenza, semplificazione e partecipazione”* (D. Lgs. n. 82/2005, art. 12, CAD); inoltre le amministrazioni pubbliche sono obbligate per legge ad assicurare *“la disponibilità, la gestione, l'accesso, la trasmissione, la conservazione e la fruibilità dell'informazione in modalità digitale”* (D. Lgs. n. 82/2005, art. 2, CAD); a ciò si aggiunga che *“i dati delle pubbliche amministrazioni sono formati, raccolti, conservati, resi disponibili e accessibili con l'uso delle tecnologie dell'informazione e della comunicazione che ne consentano la fruizione e riutilizzo da parte delle altre pubbliche*

amministrazioni e dai privati” (D. Lgs. n. 82/2005, art. 50, CAD).

Il Legislatore nel 2009 ha così introdotto una nuova concezione di trasparenza: *“La trasparenza è intesa come accessibilità totale, anche attraverso lo strumento della pubblicazione sui siti istituzionali delle amministrazioni pubbliche, delle informazioni concernenti ogni aspetto dell’organizzazione, degli indicatori relativi agli andamenti gestionali e all’utilizzo delle risorse per il perseguimento delle funzioni istituzionali, dei risultati dell’attività di misurazione e valutazione svolta dagli organi competenti, allo scopo di favorire forme diffuse di controllo del rispetto dei principi di buon andamento e imparzialità. Essa costituisce livello essenziale delle prestazioni erogate dalle amministrazioni pubbliche ai sensi dell’articolo 117, secondo comma, lettera m), della Costituzione.”* (articolo 11, comma 1, D.Lgs. n. 150/2009).

Infine il 25 gennaio 2011 è entrato in vigore il terzo decreto correttivo al CAD, il D. Lgs. n. 235/2010, che ha completato l’art. 11 D. Lgs. n. 150/2009, facendolo diventare un vero e proprio diritto soggettivo: il nuovo comma 1-bis dell’art. 52, infatti, afferma che le pubbliche amministrazioni devono promuovere *“progetti di elaborazione e di diffusione dei dati pubblici di cui sono titolari”*, nonché assicurarne la pubblicazione *“in formati aperti”*, al fine di *“valorizzare e rendere fruibili”* i dati stessi.

Con tutte le leggi introdotte la trasparenza è diventata un obiettivo principale delle Pubbliche Amministrazioni. Non c’è trasparenza se l’Amministrazione non rende pubbliche on-line le informazioni formate e raccolte nell’esercizio della propria attività.

La **Privacy** rappresenta il più grande ostacolo alla trasparenza.

Fare Open Data non vuol dire che l’Ente deve rendere pubblici tutti i dati formati nell’esercizio delle proprie attività istituzionali. I limiti alla conoscibilità dei dati rimangono quelli previsti dalle leggi e dai regolamenti vigenti (ad esempio in materia di segreto di Stato) e con riferimento alla riservatezza dei soggetti a cui i dati si riferiscono (che andrà garantita ai sensi del D. Lgs. n. 196/2003). Il rispetto della riservatezza degli individui come presupposto per la pubblicazione dei dati e del mantenimento della fiducia è fondamentale.

La Privacy può essere tutelata pubblicando i dati in forma anonima o comunque adottando tutte le cautele idonee ad evitare che gli individui cui i dati si riferiscono possano essere identificati. La deliberazione n. 88/2011 ha lo scopo di definire un primo quadro unitario di misure e accorgimenti finalizzati a individuare opportune cautele che i soggetti pubblici sono tenuti ad applicare in relazione alle ipotesi di pubblicazione dei propri dati sul Web. In ogni modo, gran parte dei dati pubblici non possono essere classificati come personali, in quanto non riconducibili ad un soggetto. Sempre

nel provvedimento 88/2011 il Garante della Privacy ha sostenuto che il perseguimento della finalità di trasparenza dell'attività delle PA può avvenire anche senza l'utilizzo di dati personali. Secondo il Garante, infatti, *“non si ravvisa la necessità di adottare alcuna specifica cautela qualora le pubbliche amministrazioni ritengano di pubblicare sul sito web informazioni non riconducibili a persone identificate o identificabili (ad esempio dati quantitativi aggregati per uffici riguardanti i livelli retributivi ed accessori risultanti dai contratti collettivi o da atti interni di organizzazione; tassi di assenza e di maggiore presenza del personale; informazioni relative alla performance dell'amministrazione; obiettivi assegnati agli uffici insieme ai relativi indicatori e ai risultati complessivi raggiunti; l'ammontare complessivo dei premi collegati alla performance stanziati e di quelli effettivamente distribuiti; dati relativi al grado di differenziazione nell'utilizzo della premialità, informazioni concernenti la dimensione della qualità dei servizi erogati, notizie circa la gestione dei pagamenti e le buone prassi)”*. Quindi, prima di intraprendere un'attività che comporta la pubblicazione di dati personali, è necessario valutare se la finalità di trasparenza e di comunicazione possa essere perseguita senza divulgare tale dati o rendendo pubblici atti e documenti senza indicare dati identificativi, usufruendo di modalità per identificare l'interessato solo quando necessario.

4.2 Copyright e Licenze Aperte

Esistono due tipi di licenze con cui è possibile pubblicare i dati: “closed” e “open”.

Le **licenze chiuse** consentono di riservare “tutti i diritti”, utilizzando il simbolo © che viene usato per indicare il titolare del *copyright* sull'opera. Un dato pubblico reso disponibile con una licenza chiusa non rispetterebbe il diritto di riusabilità tipico dell'Open Data, in quanto senza il consenso di chi detiene i diritti non può essere copiato, ripubblicato o modificato; il dato potrebbe essere solo consultabile.

Le **licenze aperte** tutelano l'autore del dato, attribuendogli la paternità del dato stesso ed evitando che quanto pubblicato possa subire alterazione senza un controllo e senza l'imposizione di regole ben precise. Allo stesso tempo garantiscono una serie di diritti a chi entra in possesso del dato, permettendo di ridistribuirlo liberamente e, in alcuni casi, anche manipolarlo al fine di creare opere derivate.

Le singole Amministrazioni decidono quale licenza aperta utilizzare per pubblicare i dati.

A partire dal 2010 sono state sviluppate licenze “aperte” create dai Governi di alcuni Paesi allo scopo di dare a tutte le Amministrazioni uno strumento chiaro e “certificato” in grado di facilitare la diffusione e il riutilizzo dell’informazione del settore pubblico. I primi Paesi a dotarsi di tali strumenti sono stati il Regno Unito con la Open Government Licence e la Francia con la Licence Information Publique, più recentemente l’Italia con la IODL (Italian Open Data Licence).

Più in generale le licenze che possono essere adottate per fare Open Data sono:

Open Knowledge Foundation Licence (Open Data Commons): le licenze Open Data Commons rappresentano un progetto di Open Knowledge Foundation rivolto ad affermare l’uso di licenze aperte anche per i database. Le più usate per i dati aperti sono:

- PDDL (Public Domain Dedication and Licence): Dominio Pubblico per data/database.
- ODC-By (Open Data Commons-Attribution Licence): Attribuzione per data/database.
- ODC-ODbL (Open Data Commons-Open Database Licence): Attribuzione (occorre sempre rendere possibile e garantire l’uso del database o delle opere da esso derivate secondo i termini della licenza ODbL), Share-Alike (l’uso di versioni adattate del database, nonché la creazione e distribuzione di database derivati o adattati, deve essere effettuata sempre nel rispetto dei termini della licenza ODbL), Keep open (se si redistribuisce il database o una sua versione adattata, è necessario non utilizzare sistemi che ne limitino l’uso).

Licenze Creative Commons: le licenze Creative Commons nascono dall’esigenza di permettere la libera circolazione del materiale creativo protetto dal diritto d’autore dal momento che sono ispirate al modello “alcuni diritti riservati”. Due delle clausole principali di queste licenze sono: Attribuzione e Condividi allo stesso modo. La prima impone che si debba riconoscere la paternità dell’opera all’autore originario e con essa viene imposto di segnalare sempre la fonte. La seconda, conosciuta anche come clausola virale della licenza, se applicata stabilisce che l’alterazione, trasformazione o sviluppo dell’opera, obbliga a redistribuire l’opera risultante soltanto per mezzo di

una licenza identica a quella attribuita all'opera originaria, garantendo che le libertà concesse dall'autore siano attribuite anche alle opere derivate. Le licenze CC per l'Open Data sono:

- CC Zero
- CC by: Attribuzione
- CC SA: Condividi allo stesso modo
- CC-BY-SA: Attribuzione e Condividi allo stesso modo

Italian Open Data License (IODL): questa licenza è rivolta principalmente alle Pubbliche Amministrazioni ed è una licenza Share-Alike, cioè che impone di mantenere le opere derivate sotto la stessa licenza o una delle licenze compatibili. IODL è compatibile (a senso unico) sia con ODbL sia con CC-BY-SA 3.0.

4.3 Il formato dei dati

I dati, secondo gli standard Open, devono essere:

- pubblicati sul Web utilizzando una licenza aperta, per facilitarne l'accesso e il riuso;
- strutturati e leggibili dai computer (per esempio Excel invece di un'immagine scannerizzata di una tabella);
- primari, cioè disponibili attraverso formati non proprietari (es. CSV - Comma Separated Value - invece di Excel), vale a dire in formato grezzo, libero da vincoli di presentazione, lettura o interpretazione;
- collegati ad un URI (Uniform Resource Identifier): Linked Open Data. Quindi devono rispettare le specifiche W3C (RDF e SPARQL) per identificare le cose di cui si parla, in modo che gli utenti possano puntare (linkare) a ciò che viene pubblicato. Tutto questo permette di collegare tra loro dati pubblicati da entità differenti per produrre contenuti più ampi, interessanti e utili.

Affinché i dati siano riutilizzabili nel migliore dei modi devono essere in formato non proprietario o grezzo, ad esempio CSV. Quasi tutti i formati ODBC (Open DataBase Connectivity) possono essere convertiti in CSV. Più difficile di una semplice conversione è fare del "data scraping" se ci si trova di fronte

a un dato strutturato, ad esempio, in formato PDF. In questo caso bisogna usare strumenti intelligenti come PDF Excel online per riportarli in ODBC, controllarli e, solo allora, convertirli in CSV.

Un insieme di dati pronti per il riuso (insieme definito dalla tipologia, piuttosto che dalla dimensione, dal contesto, dal tempo o dal luogo, ecc.) viene detto dataset. Un dataset, accompagnato dalla sua licenza aperta e reso disponibile sul Web è Open Data.

4.4 Pubblicazione

Per pubblicare dei dataset Open Data è sufficiente renderli in qualunque modo disponibili e scaricabili dal Web, attraverso un server dedicato o in hosting presso qualsiasi provider Web.

È possibile inoltre utilizzare:

- software open source di gestione dei contenuti Web (siti Internet Open Data) per presentare i dataset al pubblico;
- servizi online come Google Spreadsheet e Google Sites per una rapida diffusione dei dataset a costo zero.

È buona pratica aggiungere una descrizione del dataset su CKAN Italia o CKAN, un catalogo di dataset Open Data e Linked Open Data gestito dalla Open Knowledge Foundation.

Per pubblicare Linked Open Data servono alcuni accorgimenti tecnologici aggiuntivi che è possibile trovare sul sito dell'Associazione Linked Open Data Italia.

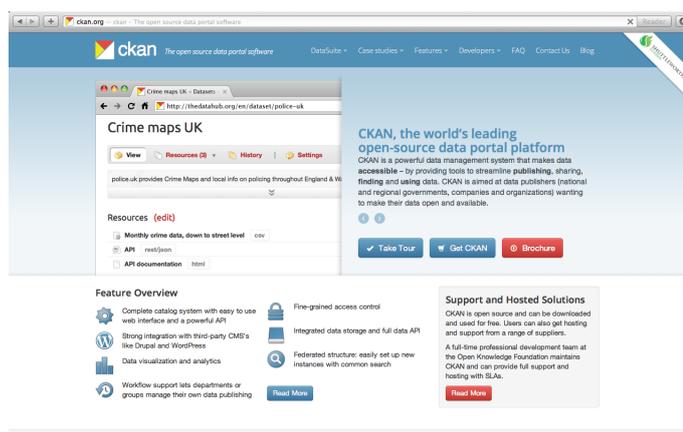


Figura 4.1: CKAN

Capitolo 5

Benefici, ostacoli e valore economico dell'Open Data

Analizziamo i benefici e gli svantaggi di una pratica Open Data, tutti i riscontri economici, gli ostacoli e i principali aspetti favorevoli nel fare Open Data.

5.1 Benefici

Una buona pratica dell'Open Data può apportare i seguenti benefici sociali:

- Creazione di trasparenza e aumento della fiducia.
- Coinvolgimento dei cittadini.
- Possibilità per tutti di creare nuovi servizi informativi ed effettuare analisi di qualsiasi tipo attraverso i dati disponibili.
- Partecipazione di più menti alla rielaborazione e allo studio dei dati.
- Miglioramento dell'allocazione di risorse pubbliche.
- La maggiore quantità di informazioni può incidere nei comportamenti delle persone/aziende (per esempio, minor uso dell'automobile, oppure minor consumo di acqua).
- Efficienza e miglioramento dei servizi pubblici.
- Creazione di valore economico e sociale.

L'apertura dei dati può avere anche vantaggi economici:

- Risparmi dal mancato sviluppo di applicazioni e servizi
- Risparmi nella gestione dei database
- Risparmi per minori costi di analisi e report
- Maggiori entrate derivanti dalla “vendita” parziale/totale dei dati
- Risparmi derivanti dall’acquisto di dataset di dati pubblici
- Vendita di applicazioni Web e/o di servizi informativi realizzati con i dati liberati

Andando piú in dettaglio possiamo vedere e analizzare i benefici interni ed esterni alle Pubbliche Amministrazioni dell’Open Data.

Benefici interni alle PA:

- Miglioramenti di produttività ed efficienza nella gestione dei processi amministrativi interni grazie alla tempestiva disponibilità del dato.
- Riduzione dei costi legati allo scambio delle informazioni (interoperabilità) delle basi dati internazionali.
- Monitoraggio della spesa pubblica sostenuta dalle diverse amministrazioni nell’esercizio della propria missione istituzionale, del suo funzionamento e per le diverse finalità di sviluppo territoriale sia di origine nazionale che comunitaria.
- Risparmi di spesa derivanti dalla realizzazione e gestione di nuovi servizi digitali erogati da terze parti grazie all’utilizzo solo ed esclusivamente di dati (tecniche di mashup) secondo l’approccio “The State brings its data and they do the rest”.
- Risparmi derivanti dalla riduzione dei costi di transazione per il reperimento dei dati di e da altre amministrazioni.

Benefici esterni alle PA:

- Ridurre e, al limite, azzerare le “asimmetrie informative” tra i fornitori di prodotti e servizi della PA (sia le aziende che i cittadini possono usufruire dei dati).
- Ampliare i mercati concorrenziali, riducendo le rendite di posizione di mercati monopolistici e/o oligopolistici. Una maggiore quantità ed un’elevata qualità delle informazioni territoriali e delle amministrazioni consente l’abbattimento di barriere all’ingresso nei mercati (piú fornitori).

- Aumentare la diversificazione dei prodotti e servizi data da più imprese che tendono a specializzarsi su molte filiere produttive.
- Ottenere un aumento del gettito fiscale a seguito degli investimenti necessari a creare nuovo business intorno alle informazioni pubbliche (in seguito alla nascita di nuove imprese).
- Aumentare la quota degli Investimenti Diretti Esteri rivolti al Paese che mette a disposizione una rilevante quantità di dati e informazioni. Credibilità e fiducia acquistata dalle PA del Paese considerato.
- Informare adeguatamente le ONG su dove investire e quanto, sia a livello di Paese che di città considerata.
- Agevolare un benchmark sulla trasparenza dei bilanci e della attività amministrativa di diversi Paesi che aderiscono a progetti comuni promossi da organismi e/o accordi internazionali per analizzare, ad esempio, i fenomeni di corruzione.

Impatti interni alla PA		Impatti esterni alla PA	
BENEFICI			
<i>Monetari (benefici che possono essere quantificabili economicamente)</i>	<i>Non Monetari (non quantificabili)</i>	<i>Monetari (benefici che possono essere quantificabili economicamente)</i>	<i>Non Monetari (non quantificabili)</i>
<ul style="list-style-type: none"> • Risparmi dal mancato sviluppo di applicazioni e servizi • Risparmi nella gestione dei database • Risparmi per minori costi di analisi e report • Maggiori entrate derivanti dalla "vendita" parziale/ totale dei dati 	<ul style="list-style-type: none"> • Maggiore trasparenza • Maggiore partecipazione dei cittadini 	<ul style="list-style-type: none"> • Risparmi derivanti dall'acquisto di dataset di dati pubblici • Vendita di applicazioni Web e/o di servizi informativi realizzati con i dati liberati 	<ul style="list-style-type: none"> • Maggiore trasparenza e quindi maggiore fiducia nel sistema • Migliore allocazione di risorse pubbliche • La maggiore quantità di informazioni può incidere nei comportamenti delle persone/ aziende (per es. minor uso dell'automobile, oppure minor consumo di acqua) • Sviluppo di nuovi servizi informativi
COSTI			
<i>Monetari (costi che possono essere quantificabili economicamente)</i>	<i>Non Monetari (non quantificabili)</i>	<i>Monetari (costi che possono essere quantificabili economicamente)</i>	<i>Non Monetari (non quantificabili)</i>
<ul style="list-style-type: none"> • Costi per "liberare" i dati • Premi in denaro per i contest 	<ul style="list-style-type: none"> • Superamento delle resistenze interne 	<ul style="list-style-type: none"> • Tempo di lavoro per sviluppare le applicazioni Web 	

Figura 5.1: Impatti interni ed esterni alle PA in termini di benefici e costi nel praticare l'Open Data

5.1.1 Alcuni Esempi

Migliorare i servizi pubblici Nei paesi dove la corruzione è diffusa, i servizi e i fondi spesso non raggiungono i cittadini e coloro che provvedono a fornire i servizi. E anche se i servizi raggiungono le persone, la loro qualità è spesso incredibilmente scarsa: i risultati di alcune indagini in Bangladesh, Ecuador, India, Peru e Uganda hanno trovato tassi di assenza fino al 20% e il 35% per gli insegnanti delle scuole e per gli operatori sanitari. In molti casi, il personale è malamente formato. Il rilascio di dati sulla fornitura di servizi può contribuire a ridurre la corruzione e a migliorare i servizi pubblici. In Uganda, i ricercatori hanno fornito informazioni ai genitori attraverso la pubblicazione, nei giornali locali, di dati sui finanziamenti per un sottoinsieme casuale di scuole. Di conseguenza, la corruzione è notevolmente diminuita, mentre i risultati scolastici sono notevolmente migliorati.

Smascherare la corruzione politica Le informazioni possono anche esporre politici corrotti. Il governo federale del Brasile, per esempio, ha iniziato a selezionare e controllare dei comuni a caso, rilasciando le relazioni sui controlli effettuati ai media. I ricercatori hanno trovato che i risultati dei controlli hanno avuto un impatto significativo sulla probabilità di rielezione dei politici: quelli esposti per corruzione sono stati puniti alle elezioni, e l'impatto è stato più forte nelle zone in cui la diffusione di informazioni è stata favorita da una radio locale.

Aumentare l'efficienza del mercato Una storia di pescatori nel sud dell'India fornisce un altro esempio di come le informazioni siano in grado di migliorare l'efficienza del mercato. Studiando l'adozione di telefoni cellulari in Kerala, i ricercatori hanno trovato prove convincenti del fatto che l'accesso alle informazioni attraverso i telefoni cellulari hanno aiutato i pescatori a vendere il pescato al mercato dove il prezzo era più alto (e il pesce più richiesto).

Benefici finanziari Il *United Nations e-Government Summary* del 2010 ha evidenziato due esempi dagli Stati Uniti: in California costa 21.000 dollari implementare il sito web sulla trasparenza delle spese dello Stato e i suoi costi operazionali annuali sono stimati al di sotto di 50.000 dollari. I visitatori del sito riportano le spese non necessarie al governo e, dopo solo pochi mesi di operatività, il sito ha già risparmiato allo Stato più di 20 milioni di dollari. In modo simile il sito sulla trasparenza in Texas, appena pochi mesi dopo il suo avvio, ha già aiutato a raggiungere risparmi di oltre 5 milioni di dollari.

5.1.2 Caso di studio: Come l'Open Data ha salvato al Canada 3,2 miliardi di dollari

Questo caso di studio è stato realizzato da David Eaves nel 2010 e mostra come l'Open Data ha aiutato ad esporre una delle più grandi frodi fiscali nella storia del Canada.

Tutto ha inizio nel 2007 quando un cliente ha richiesto di analizzare il settore della beneficenza a Toronto. Considerandolo un semplice progetto di consulenza, è stata contattata la Canada Revenue Agency (CRA) e le è stato chiesto di tutti i T3010s (gli Annual Information Returns dove gli enti di beneficenza rilasciano al CRA le loro ricevute di beneficenza e altre informazioni) del 2005 a Toronto. Dopo aver aspettato diverse settimane e risposto ad alcune domande, la CRA ha fornito le informazioni richieste.

In seguito sono stati riordinati i dati ed è iniziata l'analisi del settore di beneficenza nella Greater Toronto Area. Un pomeriggio, per gioco, sono stati ordinati gli enti di beneficenza in base alla dimensione delle donazioni fiscali quietanzate.

A questo punto è importante fare delle osservazioni. La United Way of Greater Toronto è uno degli enti di beneficenza più grandi nel Nord America, infatti la sua più recente donazione di beneficenza è stata la più grande nel continente. Nel 2008 - l'anno in cui è iniziata la crisi finanziaria - la United Way of Greater Toronto ha raccolto 107,5 milioni di dollari.

Con una certa sorpresa, dopo aver ordinato gli enti di beneficenza in base agli importi delle donazioni del 2005, è stato scoperto che la United Way of Greater Toronto non era prima sulla lista, ma era terza.

Questa fu un'enorme sorpresa. Da qualche parte a Toronto, senza che nessuno lo sapesse, due enti di beneficenza avevano raccolto più soldi della United Way of Greater Toronto (che nel 2005 aveva raccolto \$96,1M). La più grande, l'International Charity Association Network (ICAN) ha raccolto \$248M nel 2005. L'altra, la Choson Kallah Fund of Toronto ha raccolto \$120M.

Infatti, 4 delle top 15 associazioni di carità sulla lista, includendo Millennium Charitable Foundation e Banyan Tree, non erano conosciute da coloro che sono attivi nel settore della beneficenza da molto tempo.

In tutto, è stato stimato che queste associazioni di carità operando illegalmente hanno protetto approssimativamente mezzo miliardo di dollari nel 2005. Infatti, i giornali hanno poi confermato che nel 2007, le donazioni fraudolente erano circa un miliardo di dollari l'anno, con circa 3,2 miliardi di dollari illegalmente al riparo, una somma che rappresenta il 12% di tutte le donazioni in Canada.

Pensiamo: un miliardo di dollari l'anno è circa il 6% del budget annuale del Governo Federale.

CRA non aveva ancora preso provvedimenti verso queste organizzazioni. Le frodi fiscali erano ancora identificate dalla CRA come beneficenze qualificate e stavano ancora sollecitando i donatori con l'approvazione del governo.

Allora, furono date le informazioni alla leadership del settore di beneficenza e queste organizzazioni come un gruppo le hanno portate alla leadership al CRA. Dalla fine del 2007 fino al 2009 la divisione di carità CRA - ora sotto una nuova leadership - ha sistematicamente chiuso i rifugi fiscali di beneficenza e sta continuando a farlo. Uno dopo l'altro, International Charity Association Network, Banyan Tree Foundation, Choson Kallah Fund, the Millennium Charitable Foundation e altri che erano stati identificati hanno perso il loro stato di associazioni di beneficenza. I segnalati 3,2 miliardi di dollari in gettiti fiscali sostenuti da 100000 filers fiscali canadesi non sono stati finora riconosciuti o sono in corso di discussione. Una "class action" lanciata da centinaia di donatori contro le organizzazioni e uno studio legale di Banyan Tree Foundation è stata recentemente certificata. È la prima volta. Forse la CRA stava già investigando questi casi.

Questo significa che, al meglio, i dati governativi - informazioni che dovrebbero essere rese più accessibili e aperte in un formato leggibile dalla macchina e non strutturato - hanno aiutato a rivelare uno dei più grandi scandali di evasione fiscale nella storia del paese. Ma se il CRA stava già investigando, l'esame minuzioso di questi dati da parte del pubblico è servito ad uno scopo differente - aiutando a portare queste questioni fuori, all'aperto, ha forzato il CRA ad intraprendere un'azione pubblica (sospendendo il diritto di queste organizzazioni di sollecitare donazioni) al più presto. Essenzialmente da prima del 2005-2007 dozzine di associazioni di beneficenza stavano operando illegalmente. Se i dati riguardo le loro ricevute di beneficenza fossero stati disponibili al pubblico, qualcuno avrebbe potuto notare queste frodi e portarle alla luce molto prima. Forse poteva essere avviato anche un sito web di monitoraggio delle donazioni. Questo avrebbe esposto le associazioni di beneficenza che avevano grandi donazioni anormali con pochi programmi per giustificarle. Inoltre, avrebbe potuto dare ad alcuni dei 100000 canadesi ora soggetti di revisione uno strumento per valutare le associazioni di beneficenza a cui stavano dando i soldi.

Nel mondo del computer c'è qualcosa chiamata Legge di Linus, che afferma: "dato un numero sufficiente di occhi, tutti i bug vengono a galla". Lo stesso può essere detto riguardo diverse politiche pubbliche o questioni di corruzione. Per diversi dataset i cittadini non dovrebbero fare una richiesta. Non dovrebbero neppure avere da rispondere a domande sul perchè vogliono i dati. I dati dovrebbero essere scaricabili nella loro interezza. Non intrappolati dietro a qualche inutile motore di ricerca. Quando il dato è reso disponibile in un formato leggibile dalla macchina, più occhi possono veder-

lo. Questo significa che qualcuno che conosce il settore, è più probabile che individui qualcosa che un dipendente pubblico in un'altra città non potrebbe vedere perché non ha il giusto contesto. E se a quel dipendente pubblico non è consentito di esporre il problema, può poi condividere queste informazioni con i cittadini.

Questo è il potere dell'Open Data: il potere di trovare i problemi in situazioni/ambienti complicati e possibilmente anche di prevenire la loro creazione.

5.2 Costi, svantaggi e Ostacoli

Gli Open Data non creano solo benefici economici e sociali, hanno anche alcuni costi e svantaggi:

- Costi per liberare i dati;
- Premi in denaro per i contest;
- Tempo di lavoro per sviluppare le applicazioni Web;
- Non sempre i dati aperti possono favorire chi li pubblica, interessi personali prevalgono su interessi sociali;
- I dati devono essere continuamente aggiornati per avere valore;

I principali ostacoli alla pratica dell'Open Data possono essere riassunti in:

- Ignoranza sull'uso e sulla pubblicazione dei dati aperti
- Privacy
- Incompletezza dei dati raccolti
- Non tutta la popolazione possiede ancora o sa usare Internet e le nuove tecnologie

Per superare queste difficoltà è necessario istruire le persone, utilizzando diverse piattaforme informative, e far conoscere a tutti come diffondere i dati nel migliore dei modi e nel rispetto della legge.

5.3 Perchè fare Open Data?

Dopo aver visto tutti gli aspetti positivi e negati del fare Open Data è lecito chiedersi nuovamente perchè fare Open Data. Questi possono essere alcuni dei motivi principali:

- Trasparenza verso i cittadini
- Efficienza e miglioramento dei servizi pubblici
- Rilascio di valore sociale e commerciale
- Partecipazione e coinvolgimento
- Innovazione e competizione
- Creare nuovi modelli di business e lavoro

Il miglioramento dei servizi pubblici ha come conseguenza diretta benefici economici. Ad esempio, il rilascio dei tassi di infezione MSRA negli ospedali del Regno Unito è un esempio calzante. Nel 2008, MRSA costava al NHS 45 milioni di sterline per il trattamento di circa 5000 pazienti infetti - 9000 sterline per paziente. Da Luglio 2010 ogni mese i tassi di infezione MSRA in tutti gli NHS ospedali sono stati pubblicati su data.gov.uk. Da Agosto 2011 i tassi di infezione sono calati sotto i 100 per mese attraverso la NHS Trust. Certamente la pubblicazione di open data non può essere stabilita come l'unica causa di miglioramento e i concomitanti 34 milioni di sterline di risparmi. Tuttavia, lo scambio riportato di buone pratiche tra ospedali e trust combinate con l'effetto di classifiche che mostrano i peggiori ospedali e trust avrà avuto un effetto significativo.

Vediamo qualche esempio di servizi no-profit realizzati con i dati aperti.

EveryBlock EveryBlock è il modo migliore per seguire le notizie sul vicinato e mettersi in contatto con i propri vicini in 19 città degli Stati Uniti. L'obiettivo di EveryBlock è quello di aiutarti a essere un vicino migliore, fornendoti notizie sul quartiere del vicinato frequentemente aggiornate, oltre a strumenti per avere significative conversazioni con i vicini. Il sito è semplice da usare: basta creare un account e scegliere quali luoghi desideri seguire - per esempio, la vostra casa e il vostro posto di lavoro. La tua home page personalizzata su EveryBlock, creata e aggiornata durante la giornata, vi mostrerà ciò che sta accadendo vicino ai luoghi che hai scelto, oltre a quello di cui la gente parla in quei luoghi. Partecipa alla conversazione del vicinato

quando hai qualcosa con cui contribuire - una domanda per i vostri vicini, una notizia, un elenco di eventi, o solo un promemoria in merito a qualcosa che la gente nel tuo quartiere dovrebbe conoscere.

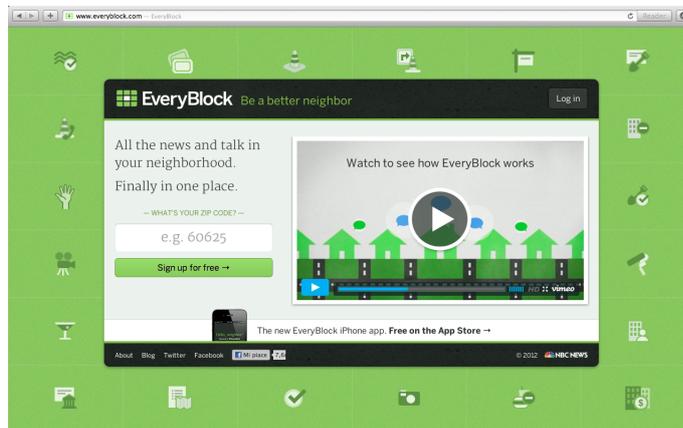


Figura 5.2: EveryBlock

WhereDoesMyMoneyGo È un sito web che ha come scopo quello di promuovere la trasparenza e il coinvolgimento dei cittadini attraverso l'analisi e la visualizzazione di informazioni riguardanti le spese pubbliche del Regno Unito. È nato nel 2007 ed è gestito dalla Open Knowledge Foundation.



Figura 5.3: WhereDoesMyMoneyGo

DataMasher È un sito web che consente di incrociare dati statistici a due a due allo scopo di evidenziare eventuali correlazioni. Ad esempio, confrontare

40CAPITOLO 5. BENEFICI, OSTACOLI E VALORE ECONOMICO DELL'OPEN DATA

il tasso di obesità degli adulti sovrappeso con il numero di fast-food per 100.000 abitanti.



Figura 5.4: DataMasher

Gasofa È un sito web spagnolo, nato nel 2011, che mette a disposizione informazioni sulla maggior parte delle stazioni di servizio in Spagna, con i prezzi che ciascuna offre. Vengono presentate le informazioni geografiche con una mappa, il prezzo corrente delle stazioni di servizio e i dati statistici sull'evoluzione storica dei prezzi. Contiene anche un elenco degli ITV più vicini, dei punti di concentrazione di incidenti e dei tratti pericolosi delle reti stradali della Spagna.



Figura 5.5: Gasofa

5.4. OPEN DATA COME STRATEGIA ECONOMICA E NUOVE OPPORTUNITÀ⁴¹

RM School Finder RM Education è una azienda degli Stati Uniti che fornisce risorse educative e ITC per scuole e genitori. In Giugno 2012 ha lanciato RM School Finder, un sito web che usa dati aperti per aiutare i genitori a trovare la giusta scuola per i loro bambini. Il sito è facile da usare e permette ai genitori di scegliere una locazione, selezionare un gruppo di scuole e confrontarle. Gli obiettivi di RM School Finder sono: presentare le informazioni che saranno più utili ai genitori per aiutarli a scegliere una scuola, aiutare i genitori a rispondere alle domande che solitamente si pongono quando devono scegliere una scuola, come cosa conseguiranno gli alunni, cosa dicono gli ispettori, e se saranno in grado di ottenere un posto presso la scuola. Per rispondere a queste domande, RM School Finder usa dati sulle scuole pubblicati dal Department for Education and Ofsted includendo School Performance Tables, GCSE Subject Results e Ofsted inspection outcomes. Le scuole hanno la possibilità di aggiungere informazioni riguardanti la scuola, per esempio come i parenti saranno coinvolti, le attività extra-curricolari e come la scuola si descrive.



Figura 5.6: RM School Finder

5.4 Open Data come strategia economica e nuove opportunità

La vera opportunità commerciale nel praticare Open Data è l'aiutare le persone a scoprire tutti questi dati, a vedere il loro potenziale e a realizzare come possono essere utilizzati per avviare attività di mercato, prendere decisioni migliori e identificare nuove opportunità. Inoltre la manutenzione dei datasets e l'aggiornamento continuo dei dati sono importantissimi per creare valore

e sviluppare efficienti servizi pubblici o buone opportunità commerciali. Affinché creino valore, in particolar modo se entrano a far parte di un modello di business, i dati devono possedere un livello di aggiornamento garantito (in termini di qualità del dato e tempi certi di rilascio). Il valore dell'Open Data sta nel suo utilizzo, non nella sua grezza esistenza.

In America la pratica dell'Open Data è stata utilizzata anche come misura anti-crisi ed è considerata l'infrastruttura della nuova economia dell'immateriale. Tali dati, infatti, una volta liberati dalla segretezza, diventano la materia prima per la produzione di beni e servizi del mondo digitale e quindi un'opportunità per l'imprenditore e per lo startupper che, raccogliendo le informazioni su un dato settore, possono sviluppare un'applicazione web, creando un modello di business.

L'Open Data può essere visto anche come una strategia economica. Per esempio, Washington DC government ha organizzato l'"Apps for Democracy" (appsfordemocracy.org), un concorso a premi per le migliori applicazioni realizzate con dati forniti dall'amministrazione, dando 30 giorni di tempo e 50.000 dollari per premi. Ha ottenuto come risultato lo sviluppo di 47 applicazioni (Web, iPhone, Facebook, ecc.) dal valore di mercato di 2.600.000 dollari. Il risparmio della collettività ammonta a 2.550.000 dollari.

Figura 5.7: Apps for Democracy

Similmente, New York City, con BigApp3.0, ha lanciato una sfida agli sviluppatori software che consiste nel creare applicazioni che sfruttano i dati aperti della città per rendere NY migliore, aiutando i residenti, i visitatori

5.4. OPEN DATA COME STRATEGIA ECONOMICA E NUOVE OPPORTUNITÀ43

e le imprese. NYC BigApps3.0 offre 50.000 dollari e altri premi come ricompensa alle migliori applicazioni: *“NYC BigApps Competition 3.0 is a joint initiative of New York City Economic Development Corporation and the New York City Department of Information Technology & Telecommunications. The Competition is intended to provide recognition to individuals, teams of individuals, corporations and nonprofit organizations for developing innovative software applications that utilize City of New York data available on NYC Open Data at www.nyc.gov/data. The goals of the Competition are twofold: (1) to stimulate the development of applications that improve access to information and government transparency, making it easier and more fun to visit, live and work in the City; and (2) to encourage innovation and the creation of new intellectual property with commercial potential by individuals, startups and small organizations.”* Questo è sicuramente un esempio di investimento nella creazione di applicazioni che possono creare valore per la città, ad esempio attraverso la promozione di servizi o la descrizione di attività nella zona.



Figura 5.8: BigApp3.0

Anche altre città hanno promosso dei contests per lo sviluppo di applicazione, simili agli esempi già visti, tra queste vi è Ottawa (Canada) con www.apps4ottawa.ca, Apps for Italy (www.apps4italy.org), Londra (UK) con London Open Data Budget App Contest (www.budgetappcontest.com), San Diego (California) con San Diego Apps Challenge (ideas.sdappschallenge.com), Amsterdam (Olanda) con Apps for Amsterdam (www.appsforamsterdam.nl/en), ecc.

Vediamo ora altri esempi sul valore economico dell'Open Data in diverse parti del mondo.

Australia: uno studio condotto dal governo Australiano (<http://www.finance.gov.au/publications/gov20taskforcereport/chapter5.htm#a3>) riguardo gli impatti economici dei dati territoriali sull'economia nazionale ha suggerito che i dati territoriali aperti e i sistemi di posizionamento ad alta precisione possono aumentare la produttività di miliardi di dollari australiani in una vasta gamma di settori industriali. Per esempio le entrate date dall'industria delle informazioni territoriali nel 2006-2007 avrebbero potuto aggiungere un guadagno di AUD 6.43-12.57 miliardi (USD 6.7-13 miliardi di dollari) , equivalenti al 0.6-1.2% del GDP.

Danimarca: uno studio in danese (<http://digitaliser.dk/resource/1021067>) sulla quantificazione del valore dell'Open Government Data ha mostrato che banche, assicurazioni ed energia hanno indicato che un migliore accesso alle informazioni del settore pubblico potrebbe avere un valore significativo, con il settore energia che stima, in collaborazione con il settore delle costruzioni, che il mercato nazionale potenziale per miglioramenti energetici, attingendo da varie fonti di dati del governo, è di 0.54-2.7 miliardi di euro (USD 0.55-3.3 miliardi di dollari).

Commissione Europea: una recente Comunicazione della Commissione Europea sull'Open Data (http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/open_dat_communication/en.pdf) prevede che, nel complesso, i vantaggi economici nell'aprire i dati pubblici potrebbero ammontare a 40 miliardi di euro (USD 48,7 miliardi di dollari) l'anno nell'Unione Europea.

Irlanda: uno studio effettuato dal TASC (<http://www.tascnet.ie/upload/file/An%20Economic%20Argument.pdf>) ha mostrato che la tariffazione per i dati non porta al recupero dei costi, che la legge FOI del paese non è così "costosa", e che il valore stimato del potenziale business nel riutilizzo dei dati pubblici è tra 83-399 milioni di euro (USD 101-486 milioni di dollari) l'anno.

Kenya: il Kenya è il primo paese africano che ha digitalizzato le sue sue informazioni (www.opendata.go.ke) come parte di un'iniziativa Open Data su vasta scala. L'iniziativa sta producendo un'economia secondaria con una gamma di applicazioni che vengono create per aiutare le persone ad accedere alle informazioni governative, gestirle e commentarle.

Spagna: uno studio (<http://www.ontsi.red.es/ontsi/en/estudios-informes/characterization-study-infomediary-sector-june-2011>) commissionato dal governo ha scoperto che il volume di business nel paese direttamente associato con il rilascio dei dati aperti da parte del governo nazionale era 550-650 milioni di euro (USD 669-791 milioni di dollari) e che tra 5000 e 5500 impiegati erano direttamente assegnati alle attività legate al riutilizzo delle informazioni.

USA: l'apertura dei dati meteorologici americani (http://www.nap.edu/openbook.php?record_id=10610&page=23) ha portato l'industria commerciale del meteo ad ottenere ricavi lordi di USD 400-700 milioni di dollari l'anno con 400 imprese che danno lavoro a 4000 persone.

In confronto, l'Europa aveva un'economia di dimensioni analoghe, ma con dati sul meteo in gran parte chiusi, e aveva solo 30 imprese con 300 dipendenti e dei ricavi pari a USD 30-50 milioni di dollari l'anno.

Negli Stati Uniti, i dati meteorologici sono sempre stati aperti e resi liberamente disponibili dal governo federale come dati o "real time streams" scaricabili. Secondo l'American Meteorological Society, la dimensione totale del mercato privato del settore delle condizioni climatiche è più di 1,5 miliardi di dollari l'anno. La ricerca ha stimato il valore economico diretto di accesso ai dati meteorologici del governo USA in 500 milioni di dollari l'anno. Una revisione della letteratura di "Arzberger et al" (2004) ha identificato anche il ruolo di questi dati pubblici nel sostenere una rapida crescita dell'industria che gestisce i rischi meteorologici sottoscrivendo strumenti per la gestione del rischio finanziario, per un valore di circa 8 miliardi di dollari. A partire da Novembre 2011, i dati del servizio meteorologico pubblico del Regno Unito sono ora disponibili come Open Data e vi è motivo di credere che un valore sostanziale sarà creato, in seguito, come risultato.

5.4.1 Caso di studio: I dati sul trasporto pubblico

L'articolo relativo a questo caso di studio del 2011 è di David Eaves ed è stato tradotto, con il permesso dell'autore, da Michele Barbera per il blog dell'associazione Linked Open Data Italia.

TransLink la società che gestisce il trasporto pubblico nella regione di Vancouver / Lower Mainland, ha lanciato un'applicazione di tracciamento degli autobus in tempo reale, che utilizza dati GPS per comunicare quanto sia distante l'autobus che stiamo aspettando.

Naturalmente per coloro che sono interessati all'innovazione nei governi e

nelle politiche pubbliche, sorge una domanda: “i dati GPS sono o non sono aperti (Open Data)?”

TransLink rende disponibile l'orario dei trasporti con licenza aperta non-commerciale. Alcuni dirigenti di TransLink potrebbero chiedersi “perchè, in questo momento di crisi economica, dovremmo rendere i nostri dati liberamente disponibili?”.

La risposta è che TransLink dovrebbe rendere aperti i propri dati, inclusi quelli GPS in tempo reale, con una licenza che ne permetta il riuso sia commerciale, sia non commerciale. Non perchè questa sia la cosa socialmente giusta da fare, ma perchè è la scelta più sensata per TransLink dal punto di vista economico.

In primo luogo, non è facile identificare modelli di business con cui Translink possa generare ricavi direttamente dai dati. Consideriamo due possibili modelli:

- vendita diretta di un'applicazione trasporti (o di pubblicità all'interno della stessa);
- vendita di un servizio di “prossimo autobus” ad aziende che ritengono che questo tipo di informazione possa essere utile per i propri clienti o dipendenti (per esempio bar e esercizi commerciali).

TransLink ha abbandonato l'idea di sviluppare applicazioni sui trasporti in favore dello sviluppo di un sito web per dispositivi mobili - ma anche se vendesse applicazioni per 1\$ a download i ricavi sarebbero ridicoli. Stimando, per eccesso, una base di 100.000 utenti, TransLink potrebbe ricavare circa 85.000 dollari (tenendo conto delle royalty trattenute da Apple per l'iPhone e assumendo che non vi sia alcuna royalty per Android). Si noti che non si tratta di un flusso di ricavi annuali, ma di un ricavo una tantum. Forse si può stimare che 10-20.000 persone scarichino la app ogni anno perchè hanno aggiornato il proprio telefono o perchè arrivano a Vancouver per la prima volta. In questo caso i ricavi annuali ammonterebbero a 15.000 dollari. Nell'arco di 5 anni, TransLink genererebbe un ricavo di circa 145.000 dollari. Non male, ma niente di eccezionale.

Al contrario, la gratuità incoraggerebbe l'utilizzo dell'applicazione. Dobbiamo considerare anche un costo opportunità: potrebbe accadere che una migliore disponibilità di informazioni sui trasporti induca alcune persone a scegliere il trasporto pubblico anzichè andare a piedi, in taxi o usare l'auto. Lo scorso anno TransLink ha gestito 211,3 milioni di viaggi. Assumiamo che la maggiore accessibilità dei dati generi un incremento dei viaggi dello 0,1%. Si tratta di un incremento infinitesimale, che però si traduce in 211.300 viaggi in più. Assumendo che ogni viaggiatore paghi un biglietto di \$2.50, si

5.4. OPEN DATA COME STRATEGIA ECONOMICA E NUOVE OPPORTUNITÀ47

genererebbe un ricavo addizionale di 528.250 dollari.

Se estendiamo l'analisi su 5 anni, come nel caso precedente, si ottiene un maggior ricavo totale di 2,46 milioni di dollari. Molto meglio dei 145.000 dollari dello scenario precedente! Questo scenario si limita ai benefici economici per l'azienda e non tiene conto della minor congestione delle strade, della riduzione dello smog e di una migliore impronta energetica.

Se applichiamo il modello alla distribuzione dei dati grezzi, si ottengono gli stessi risultati. La UBC sarà disposta a pagare per avere dati TransLink in tempo reale sui terminali degli edifici dei collettivi studenteschi? Improbabile.

I bar e caffè posizionati in zone strategiche della città sarebbero disposti a pagare? Forse.

Ovviamente i consumatori di dati dovrebbero comunque pagare per avere insegne elettroniche, ma aggiungervi ulteriori costi annuali di licenza potrebbe scoraggiare molti di essi ad offrire un simile servizio. Tenendo conto di tutti i costi connessi alla gestione delle insegne, i costi legali, la gestione dei contratti e i costi di vendita, è difficile immaginare che TransLink possa guadagnare di più di quanto non possa fare incoraggiando altri a installare insegne che generino più clienti per il proprio core business: spostare persone dal punto A al punto B.

Tanto per avere un'idea dei numeri, supponiamo che la gratuità dei dati spinga gli esercizi commerciali che non sarebbero stati disposti a pagare una licenza a installare delle insegne "il prossimo autobus" e che questo induca anche solo 1000 persone a fare 40 viaggi in più ogni anno. Questo genererebbe per TransLink un incremento di ricavi di 100.000 dollari all'anno senza dover sostenere alcun costo. Altri potrebbero installare e mantenere le insegne e TransLink non sarebbe costretta a gestire alcun contratto, licenza o struttura vendita.

Da un punto di vista economico è impossibile immaginare uno scenario in cui a TransLink non convenga permettere il riuso anche commerciale dei propri dati. TransLink non dovrebbe concentrarsi nel generare qualche dollaro in più offrendo licenze d'uso a pagamento dei propri dati, ma piuttosto concentrarsi nello spostare il vantaggio competitivo dall'accesso all'accessibilità.

Essere il monopolista dei dati sul trasporto non genera benefici per TransLink e fa sì che meno persone vedano e utilizzino i dati. Al contrario, liberare i dati fa sì che l'accesso agli stessi non sia più un vantaggio competitivo. Nel momento in cui chiunque (TransLink, Google, sviluppatori indipendenti, ecc.) è in grado di accedere ai dati, la competizione si sposta dall'accesso all'accessibilità. I consumatori smettono di rivolgersi a chi detiene i dati e si rivolgono a coloro i quali li rendono più facilmente utilizzabili.

Per esempio TransLink ha previsto che nel 2011 effettuerà un numero record

di viaggi. Una parte di me si domanda quanta parte di questi siano derivati dall'aver reso accessibili i dati su Google Maps. La verità è che per pianificare viaggi con i trasporti pubblici, Google Maps è molto più facile da usare del sito di TransLink. E questa è un'ottima cosa per TransLink! Proviamo a immaginarci cosa succederebbe se una moltitudine di aziende condividessero i dati di TransLink: dai caffè Starbucks e Blenz, ai college e le università, fino ai grandi edifici pubblici cittadini. Il vero crimine è che oggi TransLink concede a Google un monopolio de facto. Gli sviluppatori indipendenti o le piccole aziende del territorio che pagano le tasse possono utilizzare i dati per usi commerciali? Stando alla licenza, sono tagliati fuori. TransLink, dovrebbe desiderare un ecosistema in cui ognuno possa competere sull'accessibilità. Da cui trarrebbe comunque vantaggio incrementando vendite e ricavi. Ma facciamo un ulteriore passo avanti. Se TransLink aprisse i suoi dati anche per uso commerciale, ci sarebbero altri vantaggi.

Approvvigionamento. Alcuni lettori obietteranno che esistono già delle fermate di Vancouver che mostrano dati sui “prossimi autobus” (per esempio quanti minuti mancano al prossimo autobus). Se TransLink rendesse disponibili i propri dati attraverso un API (Application Public Interface) potrebbe alterare i processi di approvvigionamento per l'acquisto e la manutenzione delle insegne. Qualsiasi attore sul mercato potrebbe sapere come sono strutturati i dati e quindi gestire la manutenzione delle insegne oppure sperimentare processi innovativi e più economici per produrle. Un effetto simile si genererebbe sui bandi di gara per il sito di TransLink. Lasciando i dati liberamente disponibili TransLink potrebbe semplicemente chiedere agli sviluppatori quale ritengono sia il modo migliore di visualizzare i dati. Un numero maggiore di aziende potrebbe decidere di partecipare ai bandi, incrementando la probabilità di generare innovazione e di diminuendo i costi.

Analisi. Rendere disponibili i dati GPS potrebbe avere un altro effetto positivo. Aziende locali potrebbero utilizzarli per calcolare i flussi di traffico e predire gli ingorghi. Potrebbero queste aziende essere disponibili a pagare per ottenere i dati? Forse, ma probabilmente non abbastanza per giustificare i maggiori costi di vendita e le spese legali. Anche TransLink beneficerebbe di tali analisi, potendole utilizzare per aggiustare i propri orari e per avvertire in anticipo i guidatori di autobus di eventuali problemi. Naturalmente l'intera cittadinanza ne beneficerebbe, poiché automobilisti meglio informati potrebbero modificare i propri comportamenti (scegliendo magari anche di non utilizzare l'auto), riducendo la congestione, lo smog, l'impronta energetica, ecc.

Le opportunità di analisi di dati GPS sono potenzialmente illimitate. Blogger

5.4. OPEN DATA COME STRATEGIA ECONOMICA E NUOVE OPPORTUNITÀ⁴⁹

e studenti universitari possono fare molto. È lecito immaginare che sia possibile correlare dati sul trasporto pubblico con molti altri tipi di dati (crimine, meteo, orari degli spostamenti dei pendolari) ottenendo informazioni utili a TransLink per la pianificazione dei trasporti. Non c'è alcuna possibilità che TransLink disponga delle risorse necessarie per fare tutte le analisi possibili, quindi lasciarle fare ad altri può soltanto essere una buona idea.

Conclusioni. Per un'azienda di trasporti simile a TransLink (o un'autorità dei trasporti) è quindi consigliabile:

- Aggiungere i dati GPS al proprio portale Open Data.
- Modificare le licenze. Abbandonare le restrizioni all'uso commerciale. Fanno male al proprio business e sono anti-competitive (perché Google può usare i dati per un'applicazione commerciale mentre gli sviluppatori locali non possono?).
- Aggiungere un feed RSS ai dati GTFS. Come Google, anche noi saremmo contenti di sapere quando aggiornare i dati. Dato che noi viviamo qui e siamo anche noi utenti, sarebbe opportuno estendere lo stesso servizio che date a Google anche a noi.
- Organizzare un "Transit Data Camp" per invitare sviluppatori e imprenditori locali a incontrare il proprio staff e incoraggiarli ad utilizzare i propri dati.

Capitolo 6

I Nuovi Modelli di business degli Open Data

Un Modello di Business è la logica con cui un'organizzazione crea valore. Consiste in una serie di soluzioni organizzative e strategiche per acquisire vantaggio competitivo.

Vediamo come gli Open Data possono favorire la nascita di nuovi modelli di business.

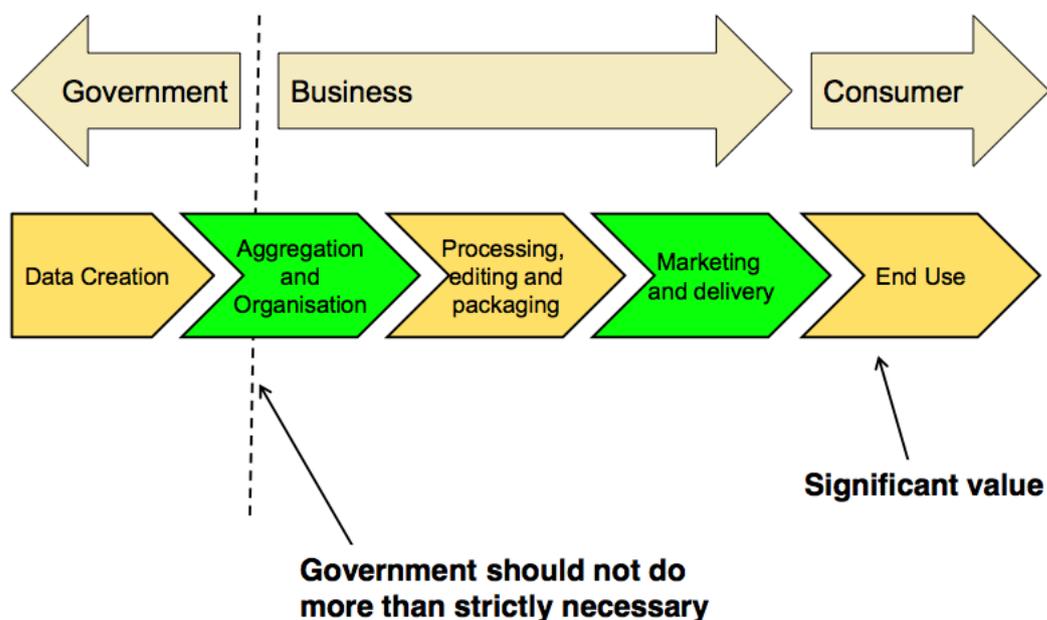


Figura 6.1: Catena del valore del riuso dei dati pubblici

Le amministrazioni pubbliche generano basi informative, ne regolano il rilascio e, a loro volta, usufruiscono dei dati prodotti da altri organismi pubblici nell'ambito delle loro funzioni. Altri soggetti, imprese o singoli, acquisiscono i dati per praticarne il riutilizzo, anche a fini commerciali, creando, a partire dall'informazione, servizi a valore aggiunto rivolti a un particolare bacino di utenza.

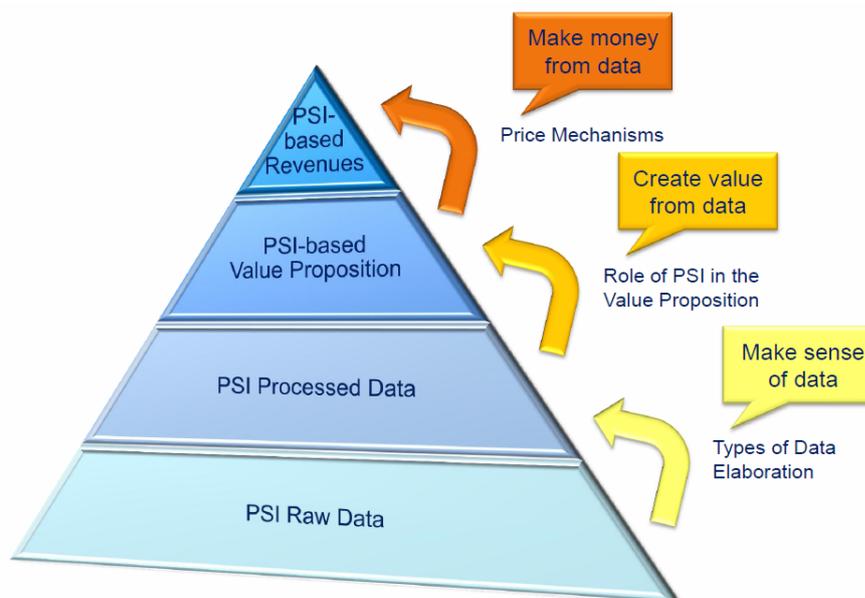


Figura 6.2: Framework per l'analisi del modello di business del PSI (piramide del valore dell'Open Data) presentata dall'Istituto Superiore Mario Boella

Il valore sta nel trasformare i dati grezzi disponibili gratuitamente in qualcosa che è realmente più utile e che i consumatori saranno disposti a pagare. Come mostra la piramide ci sono diversi steps tra i dati grezzi e il prodotto finale. È comune trovare aziende di successo costruite sulla base del software open source. Aziende come Open Corporates e Yucat si stanno sviluppando sull'Open Data usando alcune delle stesse idee:

- tariffare abbonamenti o royalties;
- il modello freemium dove un servizio base è offerto gratuitamente ma con tariffe per servizi premium;
- pubblicizzare terze parti all'interno di un'applicazione;
- contributi trasversali così che la propria applicazione sia parte della propria pubblicità; offrendo servizi a cui i consumatori potrebbero prov-

vedere loro stessi ma che sono più convenienti ed efficienti se vengono pagati per essere forniti.

Ci sono anche interessanti idee con cui un'azienda come Yucat o Listpoint fa risparmiare denaro alle amministrazioni pubbliche curando e processando i loro dati prima di renderli disponibili. La differenza tra quello che costerebbe ad una pubblica amministrazione curare i propri dati e quello che le costerebbe pagare una terza parte è dove l'azienda può realizzare un profitto, mentre la PA risparmia sulle proprie spese. Nel caso di Open Corporate, ogni arricchimento dei dati deve essere fatto sotto una licenza share-alike. Se un'azienda vuole evitare la condivisione della sua versione arricchita di dati su OC allora deve pagare una certa quota.

Per creare business è importante avere la sicurezza che i dati governativi siano disponibili, accessibili e in un formato usabile. La mancanza di tale sicurezza può essere la barriera più grande alla vera innovazione e ai benefici economici.

Nelle sezioni successive vediamo più in dettaglio alcuni modelli di business basati sul riuso dei dati aperti.

6.1 Apps e nuovi servizi di informazione per i cittadini

Lo sviluppo di applicazioni per dispositivi mobili (tablet e telefoni cellulari) è stata la prima idea di business che si è sviluppata con l'apertura dei dati pubblici. Sono tantissime le Apps disponibili create con i dati rilasciati dalle PA e la maggior parte dei siti governativi contiene una sezione in cui vengono descritte le Apps sviluppate. Sicuramente questo settore di business non è quello che porta maggiori entrate economiche ma sicuramente è il più sfruttato. Meteo, trasporti e altre utilità pubbliche sono tra le Apps più scaricate per gli Smartphone. Spesso vengono anche organizzati dei concorsi per lo sviluppo di applicazioni con i dati forniti dalle PA, che creano valore economico per lo Stato. Molte applicazioni sono basate sulla geolocalizzazione e permettono, ad esempio, di individuare tutti i servizi presenti in una determinata zona, anche facendo uso di più dataset in modo da offrire un'informazione più esaustiva. Infatti anche molte zone turistiche utilizzano applicazioni e siti web per mostrare i principali luoghi di interessi e svago, al fine di pubblicizzare le attività della zona. Più persone conoscono un servizio, più aumenta la possibilità che questo servizio venga utilizzato (es. trasporti, ristorazione, negozi, luoghi di svago come cinema, bowling, discoteche, ecc.).

Goolzoom Questo è un esempio di un modello di business legato all'informazione geografica. Goolzoom è un portale sviluppato principalmente dal programmatore spagnolo Jesús Barrio, che appartiene all'azienda (spagnola) GoolInvent S.L.U. Goolzoom collega e mostra in un'unica interfaccia dati geografici e amministrativi presi in tempo reale da vari server via Internet, facendo risparmiare moltissimo tempo a geometri, ingegneri, architetti, urbanisti, avvocati, agenti immobiliari... oppure a privati cittadini che cercano casa. Sono disponibili account a pagamento con servizi di analisi aggiuntivi. A giugno 2010 Goolzoom integra circa 200 mappe differenti e conta intorno a 250000 visite al mese, con 120000 utenti distinti e 80000 paganti nel corso dell'anno'.

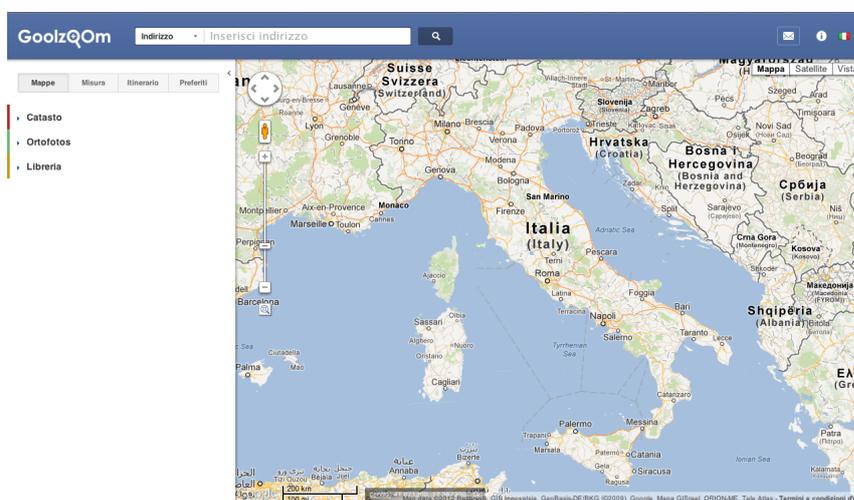


Figura 6.3: Goolzoom

Parkopedia Parkopedia è una compagnia innovativa, basata sull'Open Data, che mescola locazioni geografiche con altri dati locali. Una piccola azienda situata nel Regno Unito che usa dati in tempo reale da autorità locali per aiutare i guidatori a trovare spazi di parcheggio per le macchine gratuiti. Parkopedia è cresciuta fino a diventare la principale risorsa mondiale di informazioni sui parcheggi, fornendo informazioni su oltre 20 milioni di spazi in 25 paesi. Usato da milioni di guidatori, il servizio Parkopedia include uno strumento di prenotazione che permette ai guidatori di prenotare online un parcheggio e possiede informazioni in tempo reale sulla disponibilità di parcheggi. Parkopedia lavora anche con altre organizzazioni (es. AA, the UK's Automotive Association) per integrare i suoi dati in pianificazioni di viaggio per applicazioni e navigatori satellitari.

6.1. APPS E NUOVI SERVIZI DI INFORMAZIONE PER I CITTADINI⁵⁵



Figura 6.4: Parkopedia

iCommute SF iCommute SF è un'applicazione per iPhone che aiuta a localizzare, organizzare e accedere ad informazioni stradali e previsioni di arrivo in tempo reale per il sistema Muni di San Francisco. L'applicazione costa 2,99\$ e sicuramente aiuta i cittadini di San Francisco a muoversi meglio con i trasporti e i servizi pubblici della città.

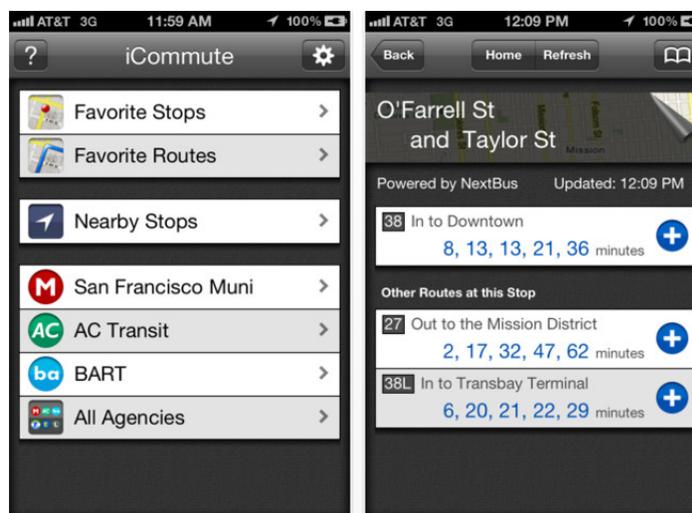


Figura 6.5: i Commute SF

Food Sprout Food Sprout combina differenti dataset per rendere trasparente come viene prodotto il cibo, su tutta la catena di fornitura. Le risorse

dei dati di Food Sprout sono: dati che raccoglie il team interno di Food Sprout, dati che gli utenti inseriscono nel sistema che vengono verificati da Food Sprout, terze parti come agricoltori non-profits che hanno dei dati, agenzie governative e database di cibo, reports investigative dove difficilmente il team di Food Sprout trova i dati. Il servizio permette agli utenti di visualizzare informazioni sul loro cibo, verificare quale cibo è realmente sostenibile e quale non lo è ed effettuare ricerche. Il modello di business dietro questo servizio è costituito da: strumenti SaaS per le aziende, consulenza usando i dati, sponsorizzazione / pubblicità.



Figura 6.6: Food Sprout

6.1.1 Caso di studio: Pocket Ranger - Nuova Zelanda

Pocket Ranger è un'applicazione per smartphone (iPhones o telefoni Android) che è stata sviluppata da una collaborazione tra il Department of Conservation (DOC) e Project Tongariro, una comunità con base a Turangi e registrata come ente di beneficenza. L'applicazione si basa principalmente sui dati aperti resi disponibili dal DOC. Project Tongariro ha gestito lo sviluppo dell'applicazione, coinvolgendo uno sviluppatore software di Auckland. L'applicazione dà agli utenti tutto quello di cui hanno bisogno per conoscere il Tongariro Alpine Crossing: descrizione dei percorsi e del loro ambiente circostante, punti di interesse, storie, foto, mappe, informazioni sulla sicurezza e servizi locali disponibili.

In Pocket Ranger è incluso anche un lettore di codice Quick Response

(QR). Lungo i sentieri sono installate delle postazioni con un codice QR su esse; quando viene letto, da un telefono con Pocket Ranger installato, compaiono informazioni rilevanti in base alla posizione dell'utente. La versione successiva dell'applicazione farà uso della funzionalità GPS nei cellulari per provvedere ad una caratterizzazione "tu sei qui".

Impatti:

- La polizia e la Search and Rescue locali sono contente che buone informazioni di sicurezza e consigli riguardo quale equipaggiamento gli escursionisti dovrebbero portare con loro sono incluse nell'applicazione. Rendere disponibili tutte queste informazioni prima che gli escursionisti raggiungano il parco dovrebbe garantire la loro sicurezza.
- Project Tongariro ha reso l'applicazione disponibile gratuitamente includendo informazioni riguardanti i servizi locali come alloggi, trasporti, punti di ristorazione, e altre attività nell'area. Le imprese locali pagano per avere i loro servizi inclusi, in modo da finanziare l'applicazione, e aumentare la loro visibilità ai viaggiatori che utilizzano l'applicazione.
- Una caratteristica fondamentale è la capacità di fornire, attraverso l'uso di codici QR, informazioni basate sulla posizione per gli escursionisti effettivamente nel parco, senza la necessità dell'installazione di una segnaletica di informazione di grandi dimensioni (che ha un impatto negativo sul paesaggio). Ciò riduce anche la necessità di produrre libri e metodi di visualizzazione che possono essere costosi.
- Project Tongariro possiede la proprietà intellettuale per il modello (template) dell'applicazione, che può essere acquistato ad una frazione del prezzo di sviluppo e applicato ad eventuali altri sentieri per passeggiate, parchi nazionali e piste ciclabili.
- L'applicazione è stata scaricata da oltre 2000 persone. La seconda versione rilasciata, in Dicembre 2011, ha incorporato feedback dagli utenti precedenti e nuove caratteristiche che includono il lettore di codice QR. La fiducia nell'applicazione è cresciuta e in primavera 2012, quando è iniziata la nuova stagione per gli escursionisti, è stata condotta una campagna.

Impatti economici e sociali:

- una comunità ha sviluppato l'applicazione beneficiando il pubblico (viaggiatori nazionali e internazionali)

58CAPITOLO 6. I NUOVI MODELLI DI BUSINESS DEGLI OPEN DATA

- i guadagni delle entrate sono stati ricavati attraverso la vendita del modello dell'applicazione
- i guadagni delle entrate sono stati ricavati attraverso la pubblicità dei servizi locali
- l'applicazione provvede un'innovativo nuovo canale pubblicitario per le aziende locali
- miglioramento della sicurezza pubblica dato dal fatto che i viaggiatori sono meglio informati riguardo le condizioni locali e le attrezzature necessarie.
- vi è un minore impatto sul paesaggio grazie all'uso di codici QR sui posti invece di grandi cartelli informativi.

Impatti di efficienza:

- i costi per pubblicare le informazioni riguardo il Tongariro Alpine Crossing sono ridotti;
- i costi di ricerca e salvataggio sono potenzialmente ridotti;
- il modello può essere applicato ad altri percorsi e piste ciclabili.

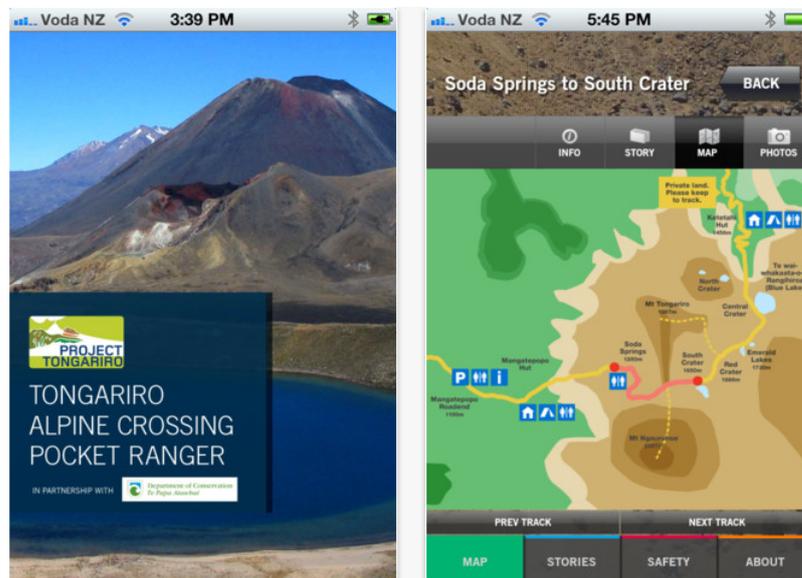


Figura 6.7: Pocket Ranger

6.2 Pulizia e organizzazione dei dati

Un'altra opportunità di business riguarda la pulizia e l'organizzazione dei dati aperti. I dati sono molto utili se organizzati e raccolti in modo efficiente. Alcune aziende traggono reddito dall'analisi delle informazioni messe a disposizione dalla PA, come DataMarket, Factual e Infochimps. Un aspetto importante da valutare è la Business Intelligence (BI), un insieme di processi aziendali per raccogliere ed analizzare informazioni strategiche, nonché la tecnologia utilizzata per realizzare questi processi e le informazioni ottenute come loro risultato. Applicare la BI ai dati aperti crea valore aggiunto e apre le porte a un nuovo modello di business.

Spikes Cavell Questo è un esempio della Business Intelligence a livello di acquisto dei dati. Spikes Cavell fornisce, ai responsabili delle decisioni nel settore pubblico, la Business Intelligence, gli strumenti on-line e le capacità di analisi di cui hanno bisogno per trasformare il loro modo di procurarsi beni e servizi.

Al centro della proposta Spikes Cavell è l'Osservatorio - una piattaforma online che facilita la distribuzione della spesa e restringe la visibilità rapidamente, in modo conveniente e con poco sforzo, sul ruolo dell'agenzia o dell'istituzione.

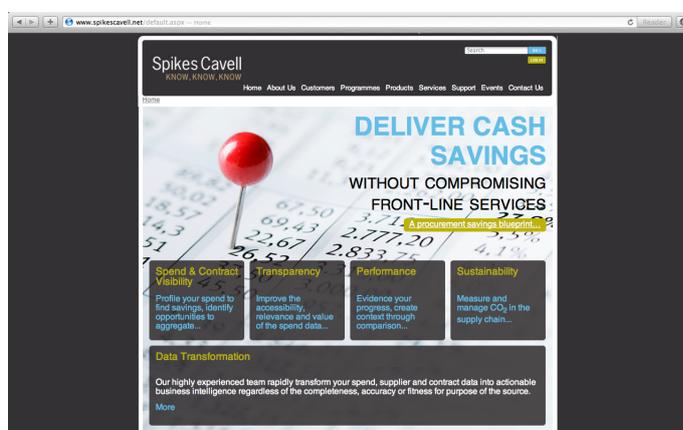


Figura 6.8: Spikes Cavell

DataMarket DataMarket è una compagnia che ha integrato molte risorse di dati pubblici aperti al suo portale e usa il seguente approccio: i dati che fin dall'inizio sono aperti e gratuiti resteranno aperti e gratuiti anche su DataMarket, solo più semplici da trovare e usare.

60CAPITOLO 6. I NUOVI MODELLI DI BUSINESS DEGLI OPEN DATA

È stata avviata, in Islanda, nel 2010 e il suo modello di business si riassume in due parti:

- Alcune caratteristiche del sistema, come personal dashboards, scheduled email reports e API access sono disponibili solo per utenti abbonati che hanno pagato.
- L'accesso ai dati premium, quali ricerche di mercato, dati di mercato finanziari e le previsioni degli analisti, è venduto a pagamento. Come implica il nome della compagnia, questo è un mercato per i dati, anche se molta "merchandise" è gratuita.

Tenere i dati aperti e gratuiti è una parte importante di questo modello:

- Porta una base di utenti al sito che potrebbero a un certo punto diventare abbonati, sia per l'account pro o alcuni dei dati premium.
- Significa che gli utenti sono pieni di contenuti gratuiti per realizzare il valore dei loro strumenti e conoscere cosa aspettarsi se comprano qualcosa nel sito; e ultimo, ma non meno importante ...
- ...Rende i fornitori dei dati del settore pubblico di gran lunga più positivi verso l'approccio del servizio, piuttosto che se la compagnia nascondesse i loro dati dietro un "paywall".

DataMarket aiuta gli utenti aziendali a trovare e capire i dati, e i fornitori di dati a pubblicare in modo efficiente i loro dati e a raggiungere un nuovo pubblico.

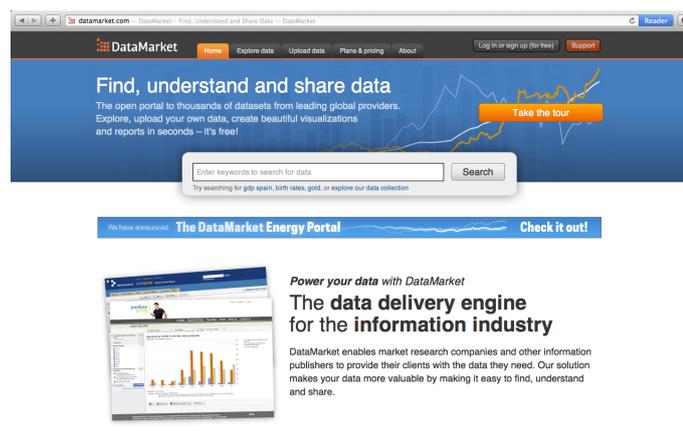


Figura 6.9: DataMarket

6.3 Data mining

Il data mining è l'insieme di tecniche e metodologie che hanno per oggetto l'estrazione di un sapere o di una conoscenza a partire da grandi quantità di dati (attraverso metodi automatici o semi-automatici) e l'utilizzo scientifico, industriale o operativo di questo sapere.

La grande quantità di dati aperti rilasciati dalle PA permette quindi di creare servizi che, mostrando i dati aperti, danno la possibilità di ricavare informazioni specifiche di un determinato settore.

PharmaTimes PharmaTimes fornisce una piattaforma, per il settore farmaceutico e per i sistemi sanitari, che rende disponibili le ultime notizie, i pensieri, e le iniziative in questa importante area. Il sito, infatti, contiene dati sulle prescrizioni, sugli esiti dei pazienti, registri, cartelle cliniche e registri sul livello di educazione degli allievi.

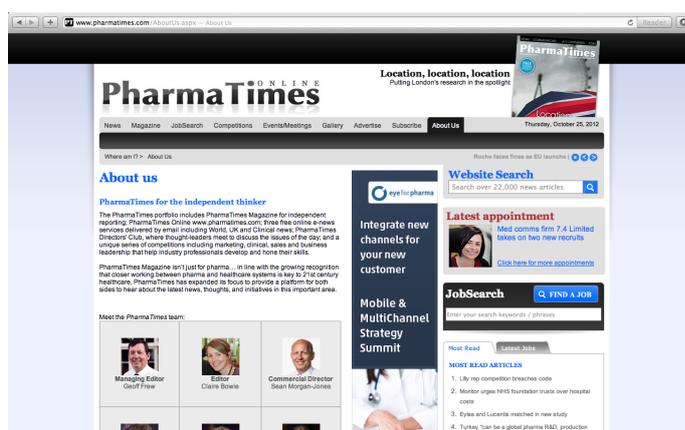


Figura 6.10: PharmaTimes Online

Graydon Graydon Belgium è una compagnia belga che aiuta a valutare correttamente e a controllare la salute dei rapporti commerciali. Mette a disposizione dei clienti tutte le informazioni rilevanti e aggiornate che permettono di valutare la situazione finanziaria dei rispettivi interlocutori in un batter d'occhio. Un'analisi finanziaria approfondita permetterà di prendere le giuste decisioni creditizie.

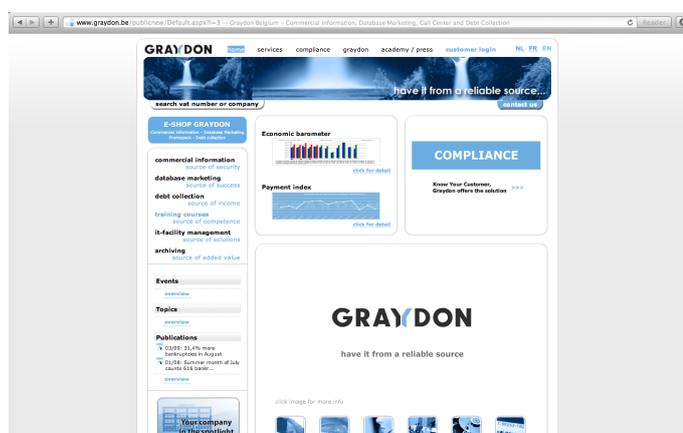


Figura 6.11: Graydon

6.4 Data broking e sviluppo di APIs

Con Data broking si intendono i dati come dati di intermediazione.

Le Application Programming Interface sono l'insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per l'espletamento di un determinato compito all'interno di un certo programma. Quindi la creazione di API che permettono di sviluppare programmi e applicazioni con l'utilizzo dei dati aperti o che consentono la visualizzazione dei dati in modo efficace, può essere sicuramente un modello di business.

Windows Azure Marketplace Windows Azure Marketplace è un mercato online per l'acquisto e la vendita di applicazioni SaaS (Software as a Service) finite e set di dati di livello premium. Con Windows Azure Marketplace è possibile collegare aziende in cerca di soluzioni innovative basate sul cloud con partner che hanno sviluppato sistemi pronti all'utilizzo. Nel sito è presente una sezione dati dove è disponibile un'ampia gamma di dati, demografici, ambientali, finanziari, relativi alle vendite al dettaglio e allo sport, che possono essere utilizzati nelle applicazioni software di Microsoft Office, in strumenti di business intelligence e in applicazioni proprietarie personalizzate. Inoltre c'è una sezione Applicazioni dove è possibile scoprire e provare applicazioni basate sulla piattaforma Windows Azure e acquistate attraverso un'unica fonte attendibile.

Nel 2012 ha avviato il servizio Bing Search API con cui è possibile incorporare web, immagini, notizie e risultati di ricerca video, così come ricerche correlate e suggerimenti di ortografia, per creare applicazioni e servizi unici.

Permette anche di integrare informazioni da dataset nella propria applicazione attraverso servizi standardizzati sui dati, e analizzare i dati online usando lo strumento Service Explorer, PowerPivot in Microsoft Excel e altri numerosi strumenti. In altre parole questo servizio nuovo consente agli sviluppatori di integrare e personalizzare i risultati della ricerca in applicazioni o siti web utilizzando XML o JSON; aggiungere funzionalità di ricerca a un sito web, creare applicazioni unicamente per il consumatore o aziendali, o sviluppare nuovi mash-up. Questo è un chiaro esempio di orientamento al cliente fidato e che fa uso di grandi quantità di dati e APIs.



Figura 6.12: Windows Azure Marketplace

Swirrl Swirrl è una micro SME situata a Stirling e a Manchester, specializzata in tecnologie per la pubblicazione di “5-star” Linked Open Data. Nell’ultimo paio d’anni l’azienda ha rilevato un incremento lento, ma costante nella domanda per lo sviluppo di software su misura per l’hosting e la consultazione di linked data. La risposta a questa domanda è stata la creazione della piattaforma PublishMyData e la vendita di servizi correlati. PublishMyData è il principale prodotto di Swirrl e fornisce la tecnologia e le infrastrutture necessarie per pubblicare i linked data, insieme ad un framework flessibile per la presentazione dei dati agli utenti, sia attraverso pagine web sia attraverso API. Questo è attualmente offerto come un servizio completamente hosted e gestito, con una *open source Community Edition* uscita nei primi mesi del 2012. Poche persone vogliono lavorare direttamente con i dati grezzi: la priorità, con PublishMyData, è quella di rendere il lavoro dei “data intermediaries” il più semplice possibile, permettendo la creazione di applicazioni, visualizzazioni e servizi sui dati con valore aggiunto. Utilizzando PublishMyData come base, Swirrl aiuta i proprietari di dati

a pubblicare i loro dati in modo utile. Swirrl stessa è coinvolta nello sviluppo di applicazioni per utenti finali oltre che nel sostenere gli altri sviluppatori a farlo. Il prossimo passo per Swirrl è di dimostrare il valore dei dati aperti, non solo per la trasparenza, ma anche per la creazione di servizi a valore aggiunto per l'utilizzo da parte del pubblico, delle imprese e del governo stesso. Gli editori di dati devono fornire dati di elevata qualità tramite API affidabili che si adattano alle esigenze degli sviluppatori. Gli sviluppatori hanno bisogno di concentrarsi sul valore concreto per gli utenti finali.

Swirrl, pur essendo orgogliosa sostenitrice dei Linked Open Data, crede fermamente nel dare alla gente quello di cui ha realmente bisogno e vuole. Ciò significa non forzare alcuno standard particolare o tecnologia sugli utenti o clienti, o assumere che ogni problema sia “un chiodo per i suoi martelli RDF”. Per i gestori dell'azienda, questo significa un uso giudizioso di standard del caso, pur riconoscendo che non tutti sono “nerd” SPARQL e RDF come loro. Ma non riguarda tutto la tecnologia. Uno dei motivi per cui i linked data non entrano a far parte della popolare conoscenza del Web è che non esistono ancora pattern per interfacce utenti così evoluti per la descrizione e la navigazione dei tipi di dati multi-dimensionali che la modellazione RDF produce. La tecnologia è relativamente semplice, le parti difficili sono la comprensione del problema principale del cliente, e la creazione di una “user experience” per i vari tipi di utenti efficace e piacevole per la visualizzazione e la gestione dei dati. Questo non è affatto un problema risolto, ma l'azienda sta già avviando un buon lavoro per rendere i dati accessibili, e sta costantemente cercando di migliorare in questo settore. Swirrl ha un modesto ma sano profitto dalle sue attività sui linked data e questo le dà fiducia per continuare a investire.

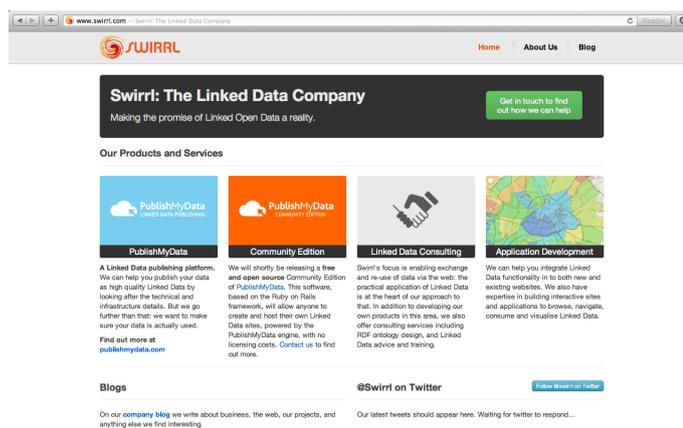


Figura 6.13: Swirrl

Geowise Con base ad Edinburgo, è una piccola medio impresa che permette alle organizzazioni di presentare grandi volumi di dati pubblici spesso complessi, in una formato visivo più semplice da capire usando il suo principale prodotto software InstantAtlas. Gli utenti possono pubblicare online reports dinamici, con collegamenti tra grafici, tabelle, mappe che mettono insieme i dati riguardanti aree e popolazioni locali. I reports di InstantAtlas possono essere piazzati su un sito web, in un “content management system”, resi disponibili per il download, inviati per email o distribuiti su CD. Questi reports possono anche essere esportati come PDFs per essere inseriti in presentazioni e documenti scritti. Moltissime nazionali e internazionali organizzazioni commerciali, pubbliche e del settore del volontariato usano InstantAtlas per visualizzare e mostrare i dati.



Figura 6.14: Geowise - InstantAtlas

6.5 Efficienza operativa o ottimizzazione

Gli Open Data possono generare valore se utilizzati per migliorare l'efficienza operativa di molti servizi, oppure se utilizzati per ottimizzare servizi pubblici o situazioni quotidiane, come la gestione del traffico, dei parcheggi, ecc.

Ito! ITO World Ltd. è una compagnia del Regno Unito nata nel 2006 con lo scopo di fornire servizi sul web per i professionisti e gli utenti dei trasporti. È specializzata nel mappare e visualizzare i dati sul trasporto. I servizi sono basati su un modello multimediale del sistema di trasporto del Regno Unito, che include strade e trasporti pubblici, mostrando i risultati con l'uso tecniche di effetti visivi. ITO ha lavorato su un numero di standards di

chiavi europee e dati inglesi ed è il maggiore partner industriale in 5 anni di progetti di ricerche governative per capire alcuni usi innovativi delle tecnologie informative da parte degli utenti dei trasporti. ITO ha da sempre lavorato con il Department for Transport, National Rail Enquiries, Guardian Media Group e Vodafone. ITO World ha lavorato anche con Google, supportando la fornitura di un piano di viaggio con il trasporto pubblico per Londra utilizzando Google Maps. Questo servizio era basato sui dati ufficiali del *Transport for London*, rilasciati come Open Data. In Aprile 2012, ITO World ha lavorato con Google sull'interazione di informazioni in tempo reale, riguardo le interruzioni lungo la Metropolitana di Londra, nel loro servizio. Quando c'è una interruzione del servizio di trasporto, ai viaggiatori viene presentata una strada alternativa e stimato il tempo di viaggio basato su dati in tempi reali. ITO World è stato coinvolto anche in OpenStreetMap (www.openstreetmap.org) per un certo numero di anni e possiede un insieme di strumenti di editing che permettono la collezione di enormi quantità di dati "crowdsourced".

Con questo esempio abbiamo visto che le informazioni in tempo reale sui blocchi o sui lavori stradali permettono una maggiore efficienza logistica.

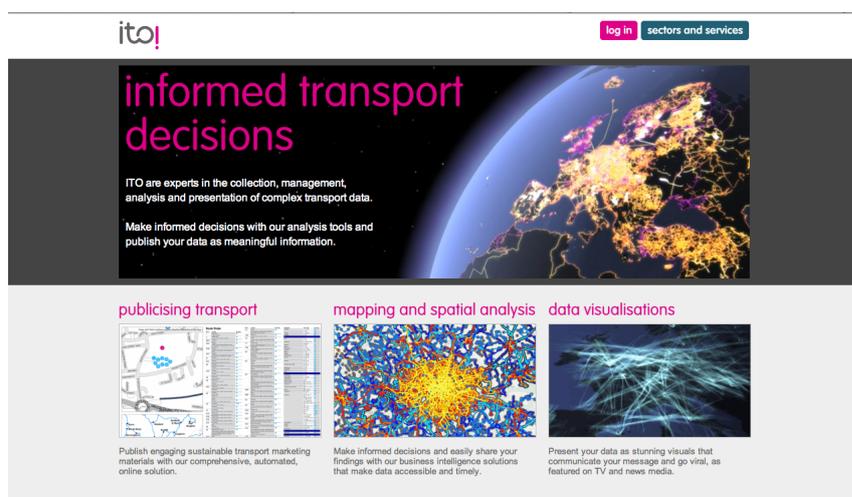


Figura 6.15: Ito!

6.6 Prodotti finanziari

Un altro modello di business che può essere adottato è quello di produrre prodotti finanziari usufruendo di analisi sui dati aperti.

Ad esempio, le assicurazioni sui rischi ambientali, metereologici.

Lloyd's Anche noti come Lloyd's of London, sono un mercato di assicurazioni del Regno Unito che tratta grandi rischi dividendoli fra un grande numero di operatori. Utilizzano l'analisi dei dati aperti per vendere assicurazioni. Ad esempio, negli Stati Uniti la gestione del rischio sul tempo (meteo) ha un valore contrattuale annuo di 4 miliardi di dollari. Vengono fatte anche assicurazioni sulle inondazioni in base alle topografie dettagliate e ai registri dei fiumi.

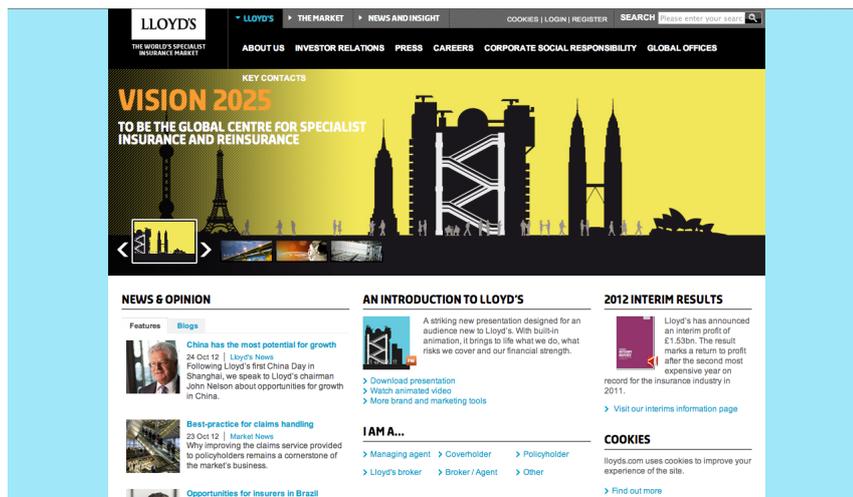


Figura 6.16: Lloyd's

6.7 Customer Attraction e Customer Retention

Per Customer Attraction si intende la capacità di attrarre nuovi consumatori, mentre per Customer Retention si intende la capacità di mantenere i consumatori, fidelizzarli. In questo caso i dati aperti possono essere utilizzati all'interno di un servizio o prodotto per attrarre potenziali clienti o mantenerli, fornendo maggiori informazioni. Vediamo di seguito un esempio di un sito web in cui i prezzi delle case locali attraggono potenziali clienti e in cui vengono mostrate liste di agenti immobiliari, servizi finanziari, edifici e altri servizi locali.

Zoopla Nel Regno Unito, Zoopla.co.uk è il sito web immobiliare più completo, che mira ad attrarre nuovi utenti con le risorse di cui hanno bisogno per prendere decisioni immobiliari disponendo di una quantità maggiore di

68CAPITOLO 6. I NUOVI MODELLI DI BUSINESS DEGLI OPEN DATA

informazioni migliori. Aiuta i consumatori sia a trovare la loro casa futura sia a ricercare il mercato combinando centinaia di migliaia di annunci immobiliari con dati di mercato, informazioni locali e strumenti comunitari. Zoopla è sostenitore della trasparenza e ogni cosa che fa è finalizzata a rendere il mercato più efficiente allo stesso modo sia per i consumatori immobiliari sia per gli inserzionisti. Il sito è diventato rapidamente la principale destinazione online, nel Regno Unito, per i consumatori di immobili che intendono cercare casa e fare le loro ricerche di mercato, e nello stesso tempo è diventato anche il partner di marketing online preferito dalle agenzie immobiliari del Regno Unito, lasciando conoscere le agenzie e gli sviluppatori di immobili. Lanciato nel 2008, Zoopla.co.uk è stato finora uno dei siti web che è cresciuto più velocemente nel Regno Unito, attraendo ora più di 20 milioni di visitatori al mese e collezionando numerosi premi e riconoscimenti, includendo l'essere nominato uno dei Top 10 UK Tech Companies (Guardian) e uno dei Top 10 Most Innovative UK Companies (Smarta).

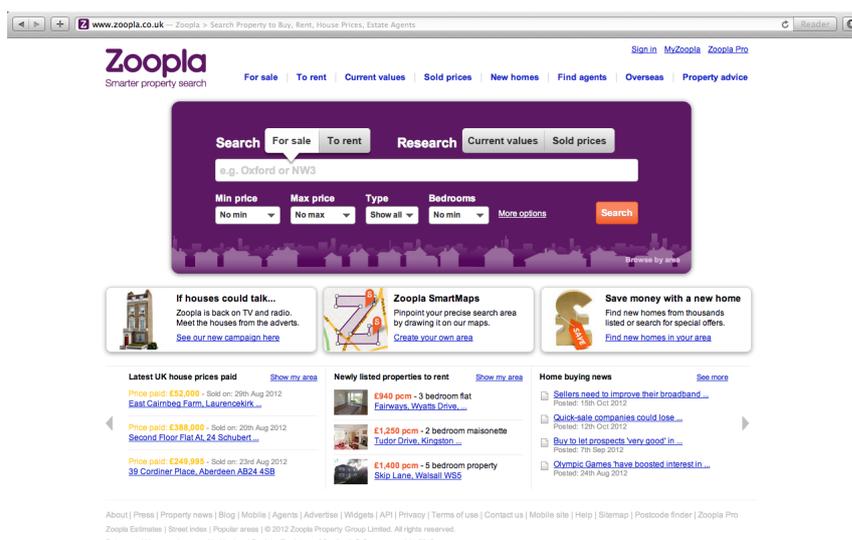


Figura 6.17: Zoopla

6.8 Supporto delle attività principali

Un'altro modello di business può essere quello in cui il rilascio dei dati aperti supporta l'obiettivo primario di business dell'organizzazione. Un esempio è la Barclays Cycle Hire a Londra, in cui il rilascio dei dati aperti sulle biciclette ha portato allo sviluppo di applicazioni che rendono più facile per i potenziali clienti di avvalersi del servizio, portando così delle entrate nel core

business.

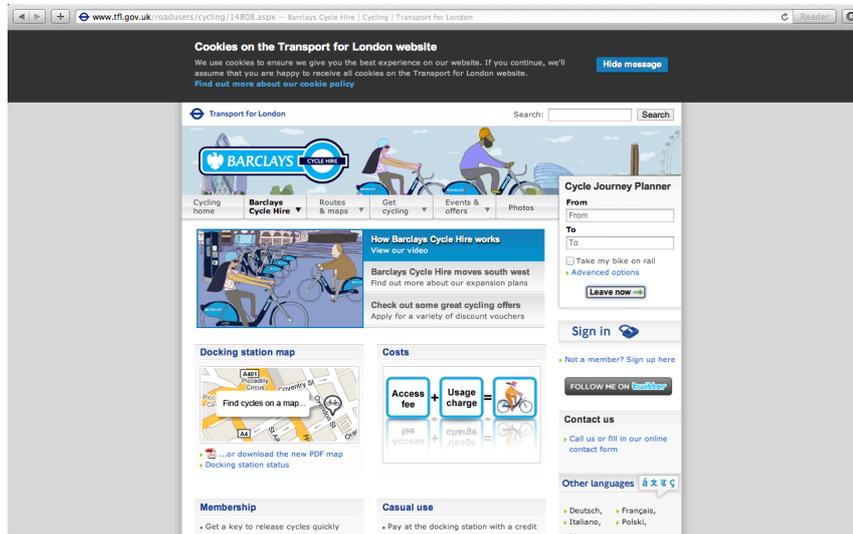


Figura 6.18: Barclays Cycle Hire

Un altro esempio è il recente rilascio dei dati relativi ai giocatori di calcio del Manchester City, che, si spera, porterà le persone a creare un modo migliore per misurare la performance dei giocatori, da cui essi potranno trarre dei vantaggi.

Sotto questa categoria possono essere inserite anche le situazioni in cui l'organizzazione che pubblica i dati finisce per migliorare il proprio uso dei propri dati utilizzando gli strumenti che vengono creati da terze parti grazie all'apertura e alla disponibilità dei dati.

Un esempio è l'utilizzo di TheyWorkForYou.com da parte dei deputati. Questo sito aiuta le persone ad informarsi sul nome dei loro deputati, il loro collegio elettorale e le loro idee.

Una grande opportunità per il settore pubblico è quella di creare un mercato per gli strumenti che gli permettono di lavorare in modo più efficiente, aprendo i propri dati.

Nei casi in cui le organizzazioni rilasciano i dati per sostenere la loro attività principali, potrebbero anche trovare, nello stesso tempo, un supporto utile a tali emissioni con "Hack Days", competizioni e così via, al fine di guidare la creazione iniziale di alcuni prodotti.

Un'altra chiave di lettura di questo modello può essere quella dell'utilizzo di dati aperti per pubblicizzare e lanciare un prodotto di un'azienda in

70CAPITOLO 6. I NUOVI MODELLI DI BUSINESS DEGLI OPEN DATA

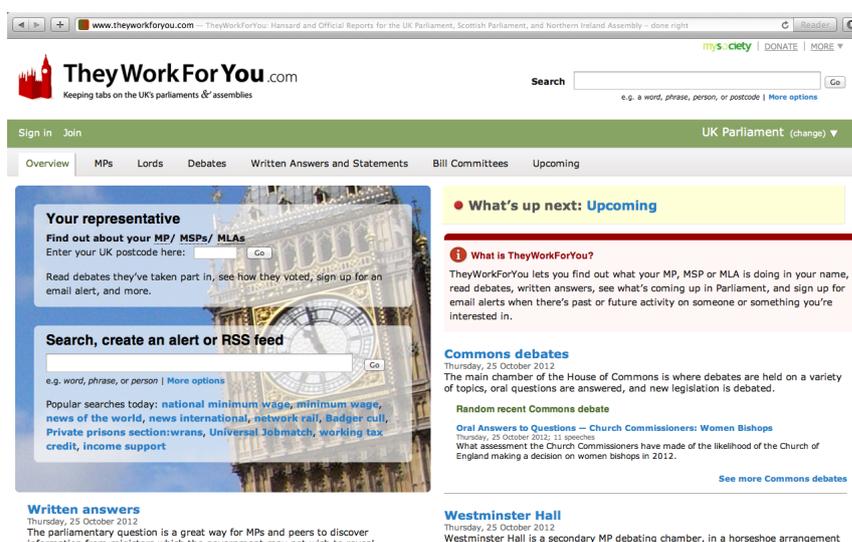


Figura 6.19: TheyWorkForYou

una determinata area geografica o settore, oppure per una certa fascia della popolazione. Ad esempio, analizzando e facendo il mashup di dati aperti riguardanti alcune aree geografiche e le principali malattie in quelle aree, se scopro che la maggior parte delle persone sono diabetiche, in certe zone, posso lanciare prodotti senza zucchero con un alta probabilità di guadagno.

6.9 Servizi di supporto all'apertura dei dati

L'offerta di supporto e servizi è un modello di business che sembra funzionare bene per compagnie costruite intorno all'open source. Nel mondo dell'Open Data, i fornitori di dati potrebbero offrire pacchetti a pagamento con:

- garanzie sulla disponibilità di dati
- priorità su correzioni di bug (sia nei dati sia nella loro fornitura) per pagare i clienti
- aiuto tempestivo per clienti che utilizzano i dati
- servizi di visualizzazione dei dati, analisi e confronto con altri dati

Questi tipi di servizi tendono ancora ad essere accoppiati con licenze nel modo dei dati, mentre nell'open source sono stati liberati con successo.

6.10 Freemium

Il modello Freemium (o Free Premium) è stato usato con qualche successo per i servizi web-based; potrebbe anche funzionare per l'Open Data. In base a questo modello, un'organizzazione dovrebbe pubblicare i dati aperti in una forma base - forse con qualche limitazione sui formati e riducendo le chiamate alle API - e offrire un accesso avanzato a quelli che sono intenzionati a pagare.

Ci sono molti modi in cui i dati aperti possono essere più utili rispetto a pubblicazioni statistiche su fogli elettronici o a un API base; secondo un modello freemium alcuni di questi punti dovrebbero essere offerti solo a quelli che pagano per essi:

- disponibilità di differenti formati leggibili dalla macchina
- numero libero di chiamate alle API
- accesso a scariche di dati piuttosto che attraverso un API (o viceversa)
- fornitura di feeds di cambiamenti ai dati
- valorizzazione dei dati attraverso informazioni aggiuntive
- accesso veloce ai dati
- fornitura di dati su DVD o dischi rigidi oltre che sulla rete

6.11 Dual Licensing

Coloro che pubblicano i dati potrebbero rendere i dati disponibili con una licenza aperta per certi propositi, e con una licenza chiusa per altri. Questa tecnica è stata adottata per certi prodotti Open Source. I “certi propositi” potrebbero non essere semplicemente “non commerciali”: gli editori potrebbero ancora incoraggiare le start-up a usare i dati da caricare in base alle dimensioni o alle entrate dell'organizzazione. O la licenza potrebbe stabilire che i dati possono essere usati in prodotti ma non possono essere usati in ulteriori feeds di “valore aggiunto” dei dati senza essere autorizzati (questo è grossomodo equivalente al dual licensing con una licenza share-alike). Questo è il modello usato da OpenCorporates.

6.12 Sponsorizzazione

Se ci sono persone o organizzazioni che credono fortemente che un particolare dataset dovrebbe essere aperto e disponibile per tutti, queste potrebbero essere favorevoli a sponsorizzare la pubblicazione di tali dataset (che non è lo stesso di concedere una licenza; la conseguenza è che il dato è aperto a tutti, non solo a quelli che pagano). Alcuni dati che sono stati aperti dal governo, così come dalla Ordnance Survey, sono stati essenzialmente aperti attraverso la sponsorizzazione della Tesoreria: essi hanno pagato affinché i dati fossero aperti, per ragioni loro (facendolo con la speranza che questo sarà in grado di far crescere l'economia).

Come si può persuadere altri a sponsorizzare l'apertura dei dati?

- Se è qualcosa a cui essi avrebbero voluto comunque applicare una licenza, forse è la situazione migliore per affrontare ogni rottura, data dal fatto che i dati sono liberamente disponibili, con i loro competitori.
- Se il tipo di dataset è ancora difficile da chiudere dopo che è stato aperto, essi potrebbero rischiare che tale dataset abbassi i loro costi a lungo termine.
- Se essi vendono analisi o prodotti di visualizzazione che chi usa i dati trova molto utili, aiuterebbero molto il loro business se rendessero i dati disponibili.

6.13 Charging for Changes

In alcuni casi, gli individui o le organizzazioni sono obbligati a fornire informazioni agli enti pubblici (e essi hanno un dovere legale di collezionarle), così che siano disponibili nel governo e più generalmente nella società. Questi enti pubblici possono (e ogni tanto fanno) addebitare a coloro che forniscono queste informazioni i “costi di amministrazione”. Esempi di questo sono le informazioni delle Companies House, le Gazette, Land Registrations, VAT Registrations e così via.

In questi casi, coloro che inseriscono le informazioni nel registro sono obbligati dalla legge, perciò sarebbe possibile far pagare loro tutto il necessario per supportare la fornitura dei dati come dati aperti. Infatti, il fornire i dati come dati aperti è come aumentare il loro utilizzo (sia nel governo sia più largamente), e quindi la pressione politica per mantenere il registro e, quindi mantenere la sua longevità.

6.14 L'aumento della qualità attraverso la partecipazione

Il modello utilizzato al legislation.gov.uk si basa sulla crescita della qualità dei dati che devono essere pubblicati - mantenendo l'aggiornamento del libro dello statuto - usufruendo dell'aiuto di altri soggetti che potrebbero beneficiare dall'aver un libro dello statuto aperto e aggiornato. Dato che, altrimenti, sarebbe molto costoso entrare in possesso di queste informazioni, sono presenti alcuni potenziali contributori, tra cui editori, avvocati, accademici, e il governo stesso.

Questo modello non copre integralmente i costi di apertura dei dati: i contributori generalmente non pagano denaro per essere coinvolti, ma donano aiuto per il mantenimento dei dati pubblicati. Così questo modello di business non copre completamente i costi, ma è molto utile per le organizzazioni che hanno l'obbligo di pubblicare le informazioni, ma non hanno le risorse per farlo bene.

Capitolo 7

Conclusioni

L'impatto dell'Open Data sul mercato può essere riassunto in due termini: innovazione e competizione. Sul lato della competizione esiste una semplice legge che collega l'efficienza dei mercati alla quantità delle informazioni disponibili agli agenti che vi operano. Questo vuol dire che più informazione c'è meglio è - si tratta chiaramente di una semplificazione ma rende l'idea. Pensiamo ad esempio ad un mercato immobiliare in cui chiunque possa facilmente avere (ad esempio su internet) le informazioni sulle compravendite. Il livello di consapevolezza del consumatore aumenterebbe e il prezzo degli immobili tendenzialmente calerebbe, perchè diminuirebbero le asimmetrie informative e le condotte scorrette. Perciò l'Open Data ha le potenzialità di aumentare la competizione e migliorare il funzionamento del mercato. Di conseguenza si aprono nuovi spazi per servizi innovativi. Nel caso del mercato immobiliare si potrebbe pensare a un app su Internet e smartphone che colleghi i dati sulle compravendite con informazioni su ambiente, trasporti, criminalità per offrire all'utente maggiori opportunità di scelta.

L'Open Data ha quindi tre principali attori: *data providers*, *decision makers* e *active citizens* (o pubblico). Un *data provider* è quello che pubblica i dati in formato grezzo e li rende accessibili per tutti; può essere pubblico o privato anche se le agenzie pubbliche e governative sono i più comuni. I *decision makers* promuovono l'Open Data per aumentare la trasparenza e per legittimare le loro proposte e decisioni. I cittadini attivi vengono incoraggiati a partecipare nei processi decisionali, affinché la presentazione dei dati sia il più possibile semplice e comprensibile.

In conclusione ritengo che l'Open Data sia una grande risorsa, ancora da scoprire pienamente, ma che può creare grandi benefici e valore per i cittadini, le aziende e i governi. Il maggiore problema è che attualmente non tutti riescono a capire come usufruire di una grande quantità di dati aperti, né sanno come renderli disponibili, oppure non si ha abbastanza fiducia nella risorsa

da creare un rilascio di dati aggiornato, affidabile e completo. Ritengo che adottare misure che mostrino alle persone come diffondere e utilizzare i dati sia indispensabile, così come fornire esempi di qualità sull'utilizzo dei dati aperti possa incoraggiare ad utilizzarli. Detto tutto questo sottolineo il fatto che se fosse così semplice praticare l'Open Data le situazioni attuali sarebbero differenti. In Italia ci si sta provando, ma in confronto ad altri paesi come gli Stati Uniti o il Regno Unito siamo ancora molto indietro e la nostra mentalità di certo non aiuta. Ritengo che una buona pratica dell'Open Data possa aiutare molto i governi, le aziende e i cittadini. Tutti devono avere le stesse possibilità di sviluppo di nuovi business, non solo chi ha già le competenze, lo stato e le astuzie necessarie per crearsi un monopolio sui dati, come Google. Per sfruttare al meglio la grande quantità di dati disponibili bisogna essere attivi nel creare piattaforme, strumenti, posti, in cui l'esistenza e il significato dei dati aperti diventi visibile a tutti. Bisogna creare una cultura del dato tra tutti i tipi di soggetti: agenzie pubbliche, partners nella catena del valore, piccole medie aziende, gruppi di consumatori, ecc. C'è bisogno di mediatori dei dati e "dataactivists". Se i dati, nonostante aperti, rimangono inaccessibili e misteriosi alla maggior parte delle persone non ci saranno mai benefici dati dall'innovazione, contribuzione e discussioni di qualità. Il punto centrale nell'apertura dei dati è quello di generare più interpretazioni produttive e usi possibili, creare terreno fertile per discussioni su problemi/questioni comuni, permettere alle persone di contribuire e relazionarsi con i dati delle aziende.

Capitolo 8

Bibliografia ed elenco delle figure

Bibliografia

- [1] Marco Fioretti. *Open Data, Open Society*. DIME e Scuola Superiore Sant'Anna di Pisa. 2010. <http://www.dime-eu.org/files/active/0>
- [2] Marco Fioretti. *Open Data: Emerging trends, issues and best practices*. DIME e Scuola Superiore Sant'Anna di Pisa. 2010. <http://www.lem.sssup.it/WPLem/odos>
- [3] Associazione Italiana per l'Open Government. *Come si fa Open Data?*. Versione 2.0. Maggioli Editore. 2011. <http://www.scribd.com/doc/55159307/Come-Si-Fa-Opendata-Ver-2>
- [4] Stefano Languardia. *Dall'Open Source all'Open Data*. Associazione Italiana per l'Open Government. 2010. <http://www.datagov.it/2010/11/25/dalopen-source-allopen-data/>
- [5] Open Knowledge Foundation. *Open Data - An introduction*. <http://okfn.org/opendata/>
- [6] *Open Data Institute*. <http://theodi.org>
- [7] Open Data Institute. *Business Plan 2012-2017*. 2012. <http://www.theodi.org/sites/default/files/ODI%20Business%20Plan%20-%20May%20Release.pdf>
- [8] Open Data Institute. *Implementation Plan 2012 and beyond*. 2012. <http://www.theodi.org/sites/default/files/ODI%20Implementation%20Plan%20-%20May%20Release.pdf>
- [9] Tim Berners-Lee. *Linked Data*. W3C. 2006. <http://www.w3.org/DesignIssues/LinkedData>
- [10] Associazione italiana per l'Open Government. *Manifesto per l'Open Government*. 2011. <http://www.datagov.it/il-manifesto/>

- [11] Guo. *The Benefits of Open Data - Evidence from Economic Research*. Open Knowledge Foundation Blog. 2012. <http://blog.okfn.org/2012/10/05/the-benefits-of-open-data-evidence-from-economic-research/>
- [12] David Eaves. *The Economics of Open Data - Mini-Case, Transit Data & TransLink*. 2011. <http://www.linkedopendata.it/leconomia-degli-open-data-un-mini-caso-di-studio-i-dati-sul-trasporto-pubblico>
- [13] David Eaves. *Case Study: How Open data saved Canada \$3.2 Billion*. 2010. <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>
- [14] Jonathan Gray. *The Business of Open Data*. Open Knowledge Foundation Blog. 2010. <http://blog.okfn.org/2010/07/06/the-business-of-open-data/>
- [15] Marco Fioretti. *Open Data for a new management of communities and the environment*. IUAV. Venezia. 2011.
- [16] Governo Italiano. *dati.gov.it - I dati aperti della PA*. <http://www.dati.gov.it>
- [17] United States Government. *data.gov - Empowering People*. <http://www.data.gov>
- [18] HM Government. *data.gov.uk - Opening up Government*. <http://data.gov.uk>
- [19] Ric Roberts. *Linked Open Data Business Model*. Swirrl. 2012. <http://blog.swirrl.com/articles/linked-data-business-models>
- [20] Hjalmar Gislason. *The Business of Open Data*. Web 2.0 Expo. 2011. <http://www.web2expo.com/webexsf2011/public/schedule/detail/16559>
- [21] Chris Taggart. *Introducing a new list just for open data on companies*. Open Knowledge Foundation Blog. 2011. <http://blog.okfn.org/2011/08/01/introducing-a-new-list-just-for-open-data-on-companies/>
- [22] Jeni Tennison, member of the Cabinet Office's Open Data User Group. *Open Data Business Models*. 2012. <http://www.jenitennison.com/blog/node/172>

- [23] Paul Malyon, Senior Product Manager, Experian QAS. Open Data - Building the Business Case. AGI GeoCommunity '12. <http://www.agi.org.uk/storage/GeoCommunity/AGI12/Papers/PaulMalyon.pdf>
- [24] HM Treasury. *Public Data Group Business Case*. 2012. http://www.hm-treasury.gov.uk/d/foi_130212.pdf
- [25] Daniel Kaplan, CEO, Fing. *Towards corporate and user-generated open data: benefits and issues around the extension of open data*. Open Data Week Keynote Speech - Nantes, May 25, 2012. http://www.opendataweek.org/wp-content/uploads/2012/04/DK_ODW_Keynote-2505.pdf
- [26] Charlotte Alldritt. *Open Data Case Studies*. Data.gov.uk Blog. 2012. <http://data.gov.uk/blog/open-data-case-studies>
- [27] HM Government. *Open Data White Paper*. Presented to Parliament by the Minister of State for the Cabinet Office and Paymaster General by Command of Her Majesty. June 2012. http://data.gov.uk/sites/default/files/Open_data_White_Paper.pdf
- [28] Jonathan Gray. *Examples of EU companies using open data?*. Open Knowledge Foundation, LOD2. 2011. <http://lod2.okfn.org/2011/12/05/examples-of-eu-companies-using-open-data/>
- [29] Richard Best. *Open Data Stories*. 2011. <http://opendatastories.org>
- [30] Government ICT Directions and Priorities. *Open Data Case Studies*. New Zealand. <http://ict.govt.nz/programme/opening-government-data-and-information/open-data-case-studies>
- [31] Andrew Stott. *Experience, Business Models and Case Study of Open Data in UK*. UK Transparency Board formerly Director, data.gov.uk & UK Deputy GCIO. 2012. http://robertlu.med.ncku.edu.tw/data/????3_0??Open%20Data????_??.pdf
- [32] Open Government Partnership Blog. *Open Data and economic growth: which link, if any?*. 2012. <http://blog.opengovpartnership.org/2012/07/open-data-economic-growth/>

- [33] Olaf Gorlitz, Gerd Groner, Christian Hachenberg, Stefan Scheglmann, Matthias Thimm. *The Open Data Workflow: Towards an Open Data Middleware*. Institute for Web Science and Technologies (WeST), University of Koblenz-Landau, Germany. 2012.

Elenco delle figure

2.1	Relazione tra PSI, Open Government e Open Data	8
2.2	Tipi di dati	10
2.3	Linked Open Data	12
2.4	Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/	13
3.1	Data.gov	17
3.2	3034 dataset liberati negli Stati Uniti, a partire dal mese di Novembre 2011 fino a Ottobre 2012 compresi	17
3.3	Distribuzione geografica delle amministrazioni italiane che rilasciano open data	18
3.4	Numero totale di dataset liberati in Italia, a partire dal mese di Marzo 2012 fino a Ottobre 2012	19
3.5	Confronto per istogrammi tra le pubbliche amministrazioni italiane che pubblicano i propri dati in formato aperto	19
3.6	ODI	20
3.7	I 4 clusters dell'ODI	21
3.8	Open Data, Pubbliche Amministrazioni e Aziende	22
3.9	OpenCorporates	23
3.10	OpenSpending	24
4.1	CKAN	30
5.1	Impatti interni ed esterni alle PA in termini di benefici e costi nel praticare l'Open Data	33
5.2	EveryBlock	39
5.3	WhereDoesMyMoneyGo	39
5.4	DataMasher	40
5.5	Gasofa	40
5.6	RM School Finder	41
5.7	Apps for Democrecy	42

5.8	BigApp3.0	43
6.1	Catena del valore del riuso dei dati pubblici	51
6.2	Framework per l'analisi del modello di business del PSI (piramide del valore dell'Open Data) presentata dall'Istituto Superiore Mario Boella	52
6.3	Goolzoom	54
6.4	Parkopedia	55
6.5	i Commute SF	55
6.6	Food Sprout	56
6.7	Pocket Ranger	58
6.8	Spikes Cavell	59
6.9	DataMarket	60
6.10	PharmaTimes Online	61
6.11	Graydon	62
6.12	Windows Azure Marketplace	63
6.13	Swirrl	65
6.14	Geowise - InstantAtlas	66
6.15	Ito!	67
6.16	Lloyd's	67
6.17	Zoopla	69
6.18	Barclays Cycle Hire	69
6.19	TheyWorkForYou	70