

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

DIPARTIMENTO DI INFORMATICA - SCIENZA E INGEGNERIA
Corso di Laurea in Ingegneria e Scienze Informatiche

LOREM IPSUM
DOLOR SIT AMET
EXPECTO TITOLORUM
(TODO: alla fine)

Relatore:
Dott. Giovanni Delnevo

Presentata da:
Francesco Valentini

Sessione IV
Anno Accademico 2023/2024

Introduzione

TODO

Indice

Introduzione	i
1 Contesto	1
1.1 Open Data	2
1.1.1 Concetto di Open Data	2
1.1.2 Caratteristiche principali degli Open Data	3
1.1.3 Formati principali	4
1.1.4 Confronto con dati privati	9
1.1.5 Contesto italiano e internazionale	9
1.2 Bologna e progetto BolognaWiFi	12
1.2.1 Bologna Smart City	12
1.2.2 Progetto BolognaWiFi	13
1.3 Visualizzazione dei dati	13
1.3.1 Importanza della visualizzazione	13
1.3.2 Caratteristiche di una buona visualizzazione	14
1.3.3 Strumenti di visualizzazione	14
1.4 Sfide e opportunità	15
1.4.1 Sfide tecniche	15
1.4.2 Aspetti etici e legali	15
1.4.3 Opportunità	17
2 Chapter 3	19
2.1 Il caso studio: dati del BolognaWiFi	19

Bibliografia

19

Elenco delle figure

1.1	Open Data su Google Trends	2
-----	--------------------------------------	---

Elenco delle tabelle

Contesto

TODO: descrizione di una pagina sugli argomenti che andremo a trattare in questo capitolo

1.1 Open Data

Le nuove tecnologie permettono di creare servizi in grado di migliorare la vita dei cittadini e di far funzionare più efficientemente governi e società. Molti dei dati necessari a raggiungere questi obiettivi sono prodotti da organismi pubblici, tuttavia spesso tali dati non sono disponibili in formati che li rendano facili da manipolare. Per questo motivo, da diversi anni si utilizza la nozione di *open data*, i cosiddetti dati aperti, e più specificamente gli *open government data*, ovvero quelli riguardanti il settore pubblico del governo, intesi come informazione, pubblica o privata, accessibile e riutilizzabile da chiunque e per qualsiasi fine. Come possiamo vedere in Figura 1.1, l'uso comune del termine inizia nel 2009, quando diversi governi hanno annunciato nuove iniziative per l'apertura della loro informazione pubblica, tra cui quelli di Stati Uniti, Regno Unito, Canada e Nuova Zelanda [1].

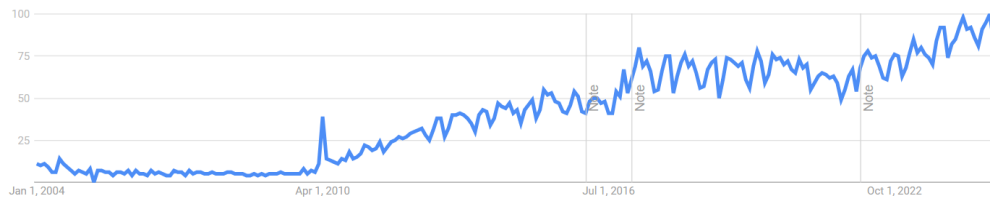


Figura 1.1: Ricerche su Google relative al termine Open Data effettuate dal 2004 al 2025. Notare l'aumento successivo al 2009 e il trend in crescita [2].

1.1.1 Concetto di Open Data

Gli *open data*, specialmente quelli della pubblica amministrazione, costituiscono una formidabile risorsa che ancora oggi rimane in larga parte inutilizzata. Molte persone e organizzazioni raccolgono una vasta gamma di svariati tipi di dati per svolgere le proprie attività. In questo ambito, il governo riveste particolare importanza, sia per la quantità e la centralità dei dati che raccoglie, sia perché la maggior parte dei dati del governo sono

pubblici per legge, quindi potrebbero diventare aperti e resi disponibili per l'utilizzo da parte di chiunque. Questo riveste particolare interesse in quanto si ritiene che gli *open data* possano apportare benefici in molti settori, ed esistono altrettanti esempi di come siano già stati usati con successo. Ci sono inoltre svariati gruppi di persone e organizzazioni che possono trarre vantaggi dalla disponibilità di *open data*, incluso il governo stesso. Allo stesso tempo, è impossibile effettuare previsioni esatte su come e dove verranno apportati questi benefici in futuro, in quanto è nella stessa natura dell'innovazione che gli sviluppi spesso partano dai posti più inaspettati [3].

1.1.2 Caratteristiche principali degli Open Data

In base a quanto stabilito dalla *Open Definition*, gli *open data* sono dati che possono essere utilizzati liberamente e ridistribuiti da chiunque e sono sottoposti, al massimo, al riconoscimento obbligatorio di una potenziale attribuzione o crediti in generale [4]. Nella *full Open Definition* vengono specificate in dettaglio le proprietà più importanti, ovvero:

- **Disponibilità e Accesso:** i dati devono essere disponibili per intero, senza pagare più di un costo ragionevole di riproduzione, preferibilmente scaricandoli dalla rete Internet. I dati devono, inoltre, essere disponibili in un formato conveniente e modificabile.
- **Riutilizzo e Ridistribuzione:** i dati devono essere forniti sotto licenza che ne permetta liberamente il riutilizzo e la ridistribuzione, incluso specialmente il mescolamento con altri *dataset*.
- **Partecipazione Universale:** tutti devono essere in grado di utilizzare, riutilizzare e ridistribuire i dati, senza discriminazioni contro campi di applicazione o contro persone o gruppi particolari. Ad esempio, non sono permesse le licenze ristrette ai soli scopi non-commerciali, che impedirebbero l'utilizzo commerciale, e nemmeno qualsiasi altra restrizione all'uso per certe finalità, come quella esclusivamente educativa.

Queste regole sono necessarie per garantire chiarezza sul significato di apertura e, soprattutto, per assicurarne l'*interoperabilità*. Quest'ultima denota la capacità di lavorare insieme, ovvero interoperare, tra diversi sistemi e organizzazioni, che in questo caso si riferisce all'interoperabilità, detta anche mescolanza, tra dataset differenti. È una capacità molto importante perché permette a diversi componenti di collegarsi per lavorare insieme e questo è essenziale per realizzare sistemi grandi e complessi, la cui costruzione altrimenti sarebbe quasi impossibile senza interoperabilità [5].

Il principio cardine di un *commons* di dati o codice, ovvero una raccolta di essi, è che un qualunque pezzo di materiale aperto lì contenuto possa essere mescolato liberamente con altro materiale aperto. Questa interoperabilità è assolutamente necessaria a livello pratico per realizzare i principali benefici ottenibili dall'apertura dei dati, ovvero l'abilità notevolmente migliorata di combinare insieme diversi dataset e di conseguenza la capacità di sviluppare un numero maggiore di prodotti e servizi aventi qualità migliore. Questa definizione data riguardo all'apertura assicura che, dati due dataset aperti ottenuti da due fonti diverse, possiamo essere sempre in grado di combinarli insieme, evitando di trovarci di fronte a un numero enorme di dataset aventi scarsa o nulla capacità di essere mescolati insieme in sistemi di più larga scala, senza quindi avere la capacità di apportare valore aggiunto. Il punto chiave di aprire i dati deve essere il *focus* su dati non-personali, ovvero i dati non devono contenere informazioni riguardanti persone specifiche. Inoltre, per alcuni tipi di dati ottenuti dal governo, potrebbero applicarsi restrizioni nazionali per motivi di sicurezza [3].

1.1.3 Formati principali

Gli open data possono essere utilizzati in svariati formati, di seguito vediamo quali.

JSON

Quello più utilizzato in ambito software è JSON, un semplice formato file che è molto facile da leggere per qualsiasi linguaggio di programmazione. La sua semplicità lo rende più veloce da processare per un computer rispetto ad altri formati come XML.

XML

XML è un formato molto utilizzato per lo scambio di dati, in quanto fornisce buone opportunità di mantenere la struttura dei dati e il modo in cui i file sono costruiti, permettendo agli sviluppatori di scrivere parti della documentazione insieme ai dati senza interferire con la loro lettura.

RDF

RDF è un formato raccomandato dal W3C e rende possibile rappresentare i dati in una forma in cui sono più facili da combinare con altri dati provenienti da più fonti. Dati RDF possono essere salvati in varie serializzazioni, tra cui XML e JSON. RDF incoraggia l'utilizzo di URL come identificatori, fornendo un modo conveniente di interconnettere direttamente iniziative open data già esistenti sul Web. Non è ancora molto diffuso, ma è diventato un trend tra le iniziative di Open Government, tra cui annoveriamo i progetti di Open Data collegati tra i governi di Regno Unito e Spagna. L'inventore di Internet, Tim Berners-Lee, ha recentemente proposto uno schema chiamato *fivesstar* che include i dati RDF collegati, ponendolo come obiettivo da raggiungere per le iniziative di open data.

Spreadsheet

Sono utilizzati da molte autorità, le quali hanno ancora delle informazioni rimaste in documenti come quelli di Microsoft Excel. Questi dati possono essere utilizzati immediatamente con la corretta descrizione del significato delle varie colonne. Tuttavia, in alcuni casi ci possono essere delle macro e delle

formule nei fogli di calcolo, che possono diventare potenzialmente più complicate da gestire. Di conseguenza, è consigliabile documentare questo genere di calcoli di fianco allo *spreadsheet*, dato che è generalmente più accessibile alla lettura da parte degli utenti.

Comma-Separated Values

I file CSV possono essere un formato molto utile data la loro compattezza e possono quindi essere adatti al trasferimento di grandi quantità di dati all'interno della stessa struttura. Tuttavia, questo formato è talmente spartano che i dati sono spesso inutili senza appropriata documentazione, dato che può essere quasi impossibile indovinare il significato delle diverse colonne. Diventa quindi particolarmente importante che la documentazione dei singoli campi sia accurata. Inoltre, è essenziale rispettare la struttura del file, in quanto la singola omissione di un campo è in grado di disturbare la lettura di tutti i rimanenti dati del file senza avere una reale opportunità di aggiustamento, perché non si può determinare come vanno interpretati i dati rimanenti.

Documento di testo

I documenti classici in formati come Word, ODF, OOXML o PDF potrebbero bastare per mostrare alcuni tipi di dati, come delle *mailing lists* relativamente stabili o equivalenti, senza richiedere un grande sforzo. Questo formato non offre alcun supporto per il mantenimento di una struttura coerente, il che spesso significa che diventa difficile inserire dati mediante strumenti automatizzati. Bisogna assicurarsi di utilizzare dei *templates* come base dei documenti che mostreranno dati per il riuso, così da rendere possibile ottenere informazioni da tali documenti. Inoltre, per supportare ulteriormente l'utilizzo dei dati si potrebbero utilizzare dei *markers* tipografici il più possibile, così da rendere più facile per una macchina la distinzione degli *headings*, ovvero i titoli di qualsiasi tipo, dal contenuto, e così via. Ge-

neralmente è raccomandabile non esibirli in un formato di *word processing*, se i dati esistono in formati differenti.

Testo non formattato

I documenti di testo non formattato (.txt) sono molto semplici da leggere per un computer, ma generalmente escludono metadati strutturali dall'interno del documento, quindi gli sviluppatori devono creare un *parser* che possa interpretare ogni documento nell'ordine in cui appare. Possono verificarsi alcuni problemi dallo scambio di file *plain text* tra diversi sistemi operativi, in quanto Windows, macOS e altre varianti Unix hanno ciascuno il loro modo di indicare al computer che si è raggiunta la fine della riga, quindi 3 diversi caratteri di *newline*.

Immagini scannerizzate

Probabilmente questo è il metodo meno indicabile per la maggioranza dei dati, ma sia TIFF che JPEG-2000 sono in grado di dargli una documentazione di ciò che si trova nell'immagine, e possono addirittura indicare testualmente l'intero contenuto del documento all'interno di un'immagine dello stesso documento. Potrebbe essere rilevante mostrare dati come immagini i cui dati non sono stati creati elettronicamente, un esempio sono i materiali di archivio, per i quali avere un'immagine è sicuramente meglio di nulla.

Formati proprietari

Alcuni sistemi dedicati hanno i propri formati di dati con cui possono essere salvati o esportati. Talvolta può bastare esporre i dati in tale formato, specialmente se ci si aspetta che un ulteriore utilizzo avvenga in un sistema simile a quello da cui provengono. Va sempre indicato dove si possono trovare ulteriori informazioni su questi formati proprietari, per esempio fornendo un link al sito web del fornitore. Generalmente, se fattibile, è raccomandabile mostrare i dati in formati non-proprietari.

HTML

Al giorno d'oggi, molti dati sono disponibili in formato HTML su vari siti e questo potrebbe tranquillamente essere sufficiente se i dati sono molto stabili e con ambito limitato. In alcuni casi, però, potrebbe essere preferibile avere i dati in una forma più facile da scaricare e manipolare, ma dato che il riferimento a una pagina su un sito web è una modalità molto semplice ed economica, potrebbe costituire un buon punto di partenza nella visualizzazione di dati. Tipicamente, sarebbe più appropriato utilizzare delle tabelle in documenti HTML per contenere i dati, ed è importante che i vari campi di dati siano visualizzati e assegnati degli ID che li rendano facili da trovare e manipolare i loro dati. Per effettuare lo scraping, Yahoo ha sviluppato un *tool* in grado di estrarre informazioni strutturate da un sito, e tali strumenti sono in grado di effettuare molte altre operazioni purché i dati siano accuratamente etichettati [6].

Formati Open File

Anche se le informazioni sono fornite in formati elettronici, dettagliati e comprensibili da un computer (*machine-readable*), potrebbero ugualmente verificarsi problemi relativi ai formati del file stesso. I formati in cui le informazioni sono pubblicate, ossia la base digitale in cui le informazioni sono salvate, possono essere sia aperti che chiusi. Un formato aperto è uno in cui le specifiche per il software sono disponibili a tutti e in maniera gratuita, così che chiunque possa utilizzare queste specifiche nel loro software senza limitazioni al riutilizzo imposte dai diritti sulla proprietà intellettuale. D'altro canto, se un formato file è chiuso, potrebbe essere sia perché è proprietario e le specifiche non sono disponibili al pubblico, o perché il formato file è proprietario ma il riutilizzo è limitato, anche se le specifiche sono state rese pubbliche. Se le informazioni sono rilasciate in un formato chiuso, questo può causare ostacoli significativi al riutilizzo delle informazioni in esso racchiuse, costringendo chi sia interessato a utilizzarle a comprare il software necessario. Il beneficio dei formati file aperti è che permettono agli sviluppatori di produrre più

pacchetti software e servizi che utilizzano questi formati, minimizzando così gli ostacoli al riuso delle informazioni che contengono. Utilizzare formati file proprietari per i quali le specifiche non sono disponibili pubblicamente può generare una dipendenza su software di terze parti o sui detentori delle licenze del formato file, il che potrebbe significare, nello scenario peggiore, che sia possibile leggere le informazioni esclusivamente utilizzando certi pacchetti software, i quali potrebbero avere un costo proibitivo, o anche diventare obsoleti. Di conseguenza, per i dati aperti forniti dal governo vige la forte preferenza per cui le informazioni dovrebbero essere rilasciate in formati file aperti e *machine-readable*, leggibili da un computer [7].

1.1.4 Confronto con dati privati

Differenze principali

Vantaggi e svantaggi rispetto agli Open Data

1.1.5 Contesto italiano e internazionale

Ad oggi, è già possibile indicare un cospicuo numero di settori nei quali i dati aperti del governo stanno già apportando un certo valore, tra cui:

- Controllo di trasparenza e democrazia.
- Partecipazione.
- Autopotenziamento.
- Miglioramento o creazione di nuovi prodotti e servizi da parte di privati.
- Innovazione.
- Miglioramento dell'efficienza di servizi del governo.
- Miglioramento dell'efficacia di servizi governativi.
- Misurazione dell'impatto di certe politiche.

- Nuova conoscenza ottenuta dalla combinazione di fonti di dati e dalla ricerca di *pattern* all'interno di grandi volumi di dati.

Riguardo alla trasparenza, esistono progetti come il finlandese *tax tree*, l'albero delle tasse, e il britannico *where does my money go*, dove vanno i miei soldi, per visualizzare in che modo i ricavi derivanti dalle tasse vengono spesi dal governo. Inoltre, possiamo ricordare come gli open data abbiano risparmiato al Canada \$3.2 miliardi in una frode fiscale relativa a un ente di beneficenza. Altri siti web includono il danese *folketsting.dk* per tracciare l'attività in parlamento e i processi di creazione di leggi, così da poter vedere esattamente gli eventi che stanno accadendo, e quali parlamentari sono coinvolti. Gli open data del governo possono anche aiutarci a prendere decisioni migliori nella nostra vita, o a diventare più attivi nella società. Una donna in Danimarca ha sviluppato *findtoilet.dk*, che mostra tutti i bagni pubblici danesi così che le persone che conosceva con problemi alla vescica potessero diventare più sicuri di se stessi quando andavano nuovamente fuori. In Olanda è disponibile un servizio, *vervuilingsalarm.nl*, che ci avvisa con un messaggio se domani la qualità dell'aria nelle vicinanze raggiungerà una soglia che abbiamo definito in precedenza. A New York possiamo facilmente capire dove possiamo portare il cane a passeggio, insieme a trovare altre persone che frequentano gli stessi parchi. Servizi come *mapumental* nel Regno Unito e *mapnificent* in Germania permettono di trovare dei posti in cui vivere, prendendo in considerazione la durata del tragitto per arrivare a lavoro, il prezzo delle case e la qualità di una certa zona. Tutti questi servizi, per funzionare, si appoggiano a dati forniti apertamente dal governo. Economicamente, gli open data sono altrettanto importanti: diversi studi hanno stimato il valore economico degli open data in alcune decine di miliardi di Euro su base annua, e questo solamente nell'Unione Europea. Nuovi prodotti e aziende stanno riutilizzando dati aperti, come ad esempio il danese *husesweb.dk* che aiuta a trovare dei modi per migliorare l'efficienza energetica della propria casa, insieme alla pianificazione finanziaria e alla ricerca di costruttori che possano svolgere il lavoro. Tutto è basato sul riutilizzo

delle informazioni catastali e informazioni sui sussidi da parte del governo, insieme al registro commerciale locale. Google Traduttore utilizza un enorme quantità di documenti dell'UE che compaiono in tutte le lingue europee per addestrare gli algoritmi di traduzione e quindi migliorare la sua qualità del servizio. Gli open data offrono del valore anche al governo stesso, ad esempio per aumentarne l'efficienza. Il Ministero Olandese dell'Istruzione ha pubblicato online tutti i suoi dati relativi all'istruzione, permettendone il riuso. Da allora, il numero di domande che hanno ricevuto è crollato, riducendo il carico di lavoro e i relativi costi, mentre per i dipendenti pubblici è più facile rispondere alle rimanenti domande, in quanto diventa maggiormente chiaro dove possono essere trovate le informazioni rilevanti. Gli open data hanno anche reso più efficiente la pubblica amministrazione, il che porta a un'ulteriore riduzione dei costi. Il dipartimento olandese dei beni culturali sta rilasciando attivamente i loro dati e collabora con società amatoriali di storia e gruppi come la *Wikimedia Foundation* per eseguire i loro compiti in maniera più efficiente. Questo non solo porta a miglioramenti nella qualità dei loro dati, ma contribuisce anche a rimpicciolire le dimensioni del dipartimento stesso. Anche se ci sono già numerose istanze relative ai modi in cui gli open data stanno creando del valore aggiunto sia in ambito sociale che in quello economico, non sappiamo ancora nulla riguardo alle nuove cose che diventeranno possibili. Nuove combinazioni di dati possono creare ulteriore conoscenza, il che può potenzialmente portare a scoprire campi di applicazione completamente nuovi. Lo abbiamo già visto in passato, per esempio quando nel XIX secolo a Londra il Dott. Snow ha scoperto il legame tra inquinamento dell'acqua potabile e colera, combinando i dati delle morti attribuibili al colera con la posizione dei pozzi d'acqua. Questo portò alla costruzione dei sistemi fognari di Londra e migliorò notevolmente la salute pubblica della popolazione. Probabilmente potremmo vedere accadere nuovamente degli sviluppi simili, scaturiti dalla combinazione di diversi dataset aperti. Questo potenziale, ancora non sfruttato, può essere rilasciato se trasformiamo i dati pubblici del governo in open data, ma solo se sono davvero aperti, ovvero se

non sottoposti ad alcun tipo di restrizione legale, finanziaria o tecnologica relativa al suo riutilizzo da parte di terzi. Ogni restrizione impedisce alle persone di riutilizzare i dati pubblici, o quantomeno lo rende più difficile. Per poter realizzare appieno il loro potenziale, i dati pubblici devono essere aperti, ovvero degli open data [3].

Iniziative globali

Open Government Partnership

Normativa italiana

CAD (Codice dell'Amministrazione Digitale)

Portali di Open Data

dati.gov.it

Open Data Bologna (<https://opendata.comune.bologna.it/>)

Casi studio e iniziative significative

INSPIRE

Copernicus

1.2 Bologna e progetto BolognaWiFi

1.2.1 Bologna Smart City

Digitalizzazione e innovazione urbana

Progetti principali

Politiche locali per gli Open Data

Strategie e risultati

Infrastrutture digitali recenti

Implementazioni chiave

1.2.2 Progetto BolognaWiFi

Obiettivi principali

Scopi e finalità del progetto

Tipologie di dati raccolti

Flussi utenti

Accessi WiFi

Affluenza in aree urbane

Statistiche recenti

Dati sull'uso del WiFi pubblico a Bologna

Risorse utili

Documentazione ufficiale del progetto

Report e analisi pubblicati dal Comune

1.3 Visualizzazione dei dati

1.3.1 Importanza della visualizzazione

Analisi dei dati urbani

Esempi pratici e benefici

Supporto decisionale

Impatti sulle amministrazioni pubbliche

1.3.2 Caratteristiche di una buona visualizzazione

Chiarezza e semplicità

Interattività e Accessibilità

1.3.3 Strumenti di visualizzazione

Panoramica degli strumenti comuni

Tableau

D3.js

Vantaggi e svantaggi degli strumenti

Breve descrizione comparativa

Applicazioni pratiche: dati sul traffico in Regno Unito

Andrew Nicolson è uno sviluppatore software che è stato coinvolto in una campagna di successo contro la costruzione di una nuova strada, la *Westbury Eastern bypass*, nel Regno Unito. Nell'aprile 2010, Andrew era interessato a ottenere accesso e a utilizzare i dati sul traffico della strada che stavano venendo utilizzati per giustificare le proposte e così è riuscito a ottenere alcuni dei dati rilevanti mediante la libertà di richieste di informazioni, ma il governo locale gli ha fornito i dati in un formato proprietario che può essere letto solamente utilizzando software prodotto da un'azienda chiamata Saturn, che è specializzata in modellazione di traffico e sulle relative previsioni. Non esiste una versione del software di sola lettura, quindi il gruppo di Andrew non aveva altra scelta se non acquistare una licenza software, finendo per pagare £500 (€600) utilizzando uno sconto per studenti. Il pacchetto software principale veniva offerto da Saturn a partire da £13 000, oltre €15 000, un prezzo fuori misura per la maggioranza dei normali cittadini. Anche se nessuna legge nel diritto dell'informazione garantisce un diritto ad avere l'accesso ai dati in formati aperti, le iniziative open data del governo stanno

iniziando ad essere accompagnate da documenti di *policies* che stabiliscono che le informazioni ufficiali debbano essere rese disponibili in formati di file che siano aperti. L'amministrazione Obama ha definito il *gold standard*, con la *Open Government Directive* emessa nel dicembre 2009, la quale decreta che, nella misura praticabile e soggetta a restrizioni valide, le agenzie debbano pubblicare informazioni online in un formato aperto che possa essere recuperato, scaricato, indicizzato e ricercato da applicazioni di ricerca web di utilizzo comune. Un formato aperto è inteso come uno che sia indipendente dalla piattaforma, leggibile dalla macchina e reso disponibile al pubblico senza restrizioni che impediscano il riuso di tali informazioni [7].

1.4 Sfide e opportunità

1.4.1 Sfide tecniche

Gestione di dataset eterogenei

Problemi e soluzioni

Scalabilità e performance

Tecnologie per ottimizzare

1.4.2 Aspetti etici e legali

Quando parliamo di database, dobbiamo innanzitutto effettuare una distinzione tra la struttura e il contenuto: utilizzando il termine *dati* ci riferiamo al contenuto del database, mentre gli elementi strutturali fanno riferimento ad attributi come i nomi dei campi e il modello della struttura dati, ovvero l'organizzazione dei campi sopra citati e le loro interrelazioni. In molte giurisdizioni è probabile che gli elementi strutturali di un database siano coperti da copyright, anche se ciò dipende in una certa misura dal livello di creatività che è stato coinvolto nella creazione della struttura. Tuttavia, in

questo caso siamo particolarmente interessati ai dati. Quando parliamo di dati dobbiamo prestare attenzione, in quanto tale parola non è particolarmente precisa e può riferirsi sia a un insieme che a un singolo elemento, ma può anche significare una grande raccolta di dati, come tutto il contenuto del database. Per evitare confusione, riserveremo il termine *contenuto* per riferirci ai singoli elementi, mentre la parola *dati* riguarderà l'intera raccolta. Diversamente da materiali multimediali come testi, musica o film, la situazione legale per i dati varia notevolmente tra diversi Paesi. Tuttavia, la maggior parte delle giurisdizioni gli garantisce alcuni diritti in qualità di raccolta dati. La distinzione tra il contenuto del database e l'intera raccolta diviene particolarmente cruciale per i database fattuali, dato che nessuna giurisdizione garantisce il diritto di monopolio sui fatti individuali, ovvero il contenuto, anche se potrebbe garantire loro dei diritti in qualità di raccolta dati. Per fare un esempio, consideriamo un semplice caso di un database che elenca i punti di fusione di varie sostanze. Mentre l'intero database potrebbe essere protetto dalla legge, così da non permettere a nessuno di accedere, riusare o ridistribuire i suoi dati senza averne il permesso, questo non impedirebbe mai a nessuno di affermare il fatto che la sostanza Y si scioglie alla temperatura Z. Le forme di protezione per le raccolte dati ricadono largamente in due casi: diritto d'autore e diritto *sui generis*. Come abbiamo già sottolineato, non abbiamo nessuna regola generale e la situazione varia in base alla giurisdizione. Di conseguenza, bisogna procedere considerando singolarmente ciascun Paese per delineare, se applicabile, quale dei due approcci sia utilizzato in una particolare giurisdizione. Dobbiamo inoltre notare come, in assenza di protezione legale, molti fornitori di database chiusi, ovvero non aperti, possono utilizzare un semplice contratto combinato con disposizioni legali per proibire la violazione dei meccanismi di controllo dell'accesso e ottenere così dei risultati simili a quelli dati da un diritto formale sull'indirizzo IP. Ad esempio, il provider di un database di citazioni può raggiungere qualsiasi insieme di termini e condizioni che voglia implementare, semplicemente richiedendo agli utenti di autenticarsi tramite una password, o fornendo agli

utenti un account e una password alla condizione che l'utente accetti i termini e le condizioni del servizio [8].

Privacy e anonimizzazione dei dati

Strumenti e tecniche utilizzate

Licenze e diritti d'uso

Regolamenti e pratiche comuni

Casi noti di violazioni

Esempi significativi e lezioni apprese

1.4.3 Opportunità**Trasparenza e coinvolgimento**

Benefici per cittadini e istituzioni

Pianificazione urbana

Utilizzo dei dati per migliorare i servizi

Collaborazioni pubblico-privato

Partnership per sfruttare al meglio gli Open Data

Chapter 3

2.1 Il caso studio: dati del BolognaWiFi

Bibliografia

- [1] Open data handbook - introduzione, 2025. URL <https://opendatahandbook.org/guide/en/introduction/>.
- [2] Google trends, 2025. URL <https://trends.google.com/trends/explore?date=all&q=%2Fm%2F0119t0j6&hl=en>.
- [3] Open data handbook - perché scegliere gli open data, 2025. URL <https://opendatahandbook.org/guide/en/why-open-data/>.
- [4] The open definition, 2025. URL <https://opendefinition.org/>.
- [5] The full open definition, 2025. URL <https://opendefinition.org/od/2.1/en/>.
- [6] Yahoo website scraping tool, 2025. URL <http://developer.yahoo.com/yql/>.
- [7] Open data handbook - formati file, 2025. URL <https://opendatahandbook.org/guide/en/appendices/file-formats/>.
- [8] Open data handbook - diritti legali, 2025. URL <https://opendatahandbook.org/guide/en/appendices/what-legal-ip-rights-are-there-in-databases/>.
- [9] Leaflet doc, 2024. URL <https://leafletjs.com/reference.html>.
- [10] Vue.js doc, 2025. URL <https://vuejs.org/guide/introduction.html>.

- [11] Open data bolognawifi - matrice spostamenti, 2025. URL <https://opendata.comune.bologna.it/explore/dataset/bolognawifi-matrice-spostamenti/information/?disjunctive=hour>.
- [12] Open data bolognawifi - affollamento, 2025. URL <https://opendata.comune.bologna.it/explore/dataset/iperbole-wifi-affollamento/information/?disjunctive=ora>.
- [13] Open data bolognawifi - affluenza, 2025. URL <https://opendata.comune.bologna.it/explore/dataset/iperbole-wifi-affluenza/information/>.
- [14] Open data handbook - cosa sono gli open data, 2025. URL <https://opendatahandbook.org/guide/en/what-is-open-data/>.