

LABORATORIO DAL BASSO

OPENDATA

IN ACTION!

Cosa sono gli Open Data e a cosa servono? Dal Business al giornalismo alla cittadinanza attiva, passando per la ricerca archeologica scopriremo che sono il motore di tante nuove professionalità. Scopri quali!

11 Ottobre . 27 Novembre 2013



17
docenti

Vittorio Alvino
Michele Barbera
Luca Corsato
Simone Cortesi
Salvino Fidacaro
Gabriele Gattiglia
Nicolas Kayser-Bril
Flavia Marzano
Jon McClure
Maurizio Napolitano
Andrea Nelson Mauro
Michele Osella
Francesco Passantino
Aline Pennisi
Morena Ragone
Duccio Schiavon
Andrea Zedda

60
ORE
DI
LEZIONE



■ Introduzione agli Open Data (38%) ■ Data Journalism (18%)
■ Open Geodati (22%) ■ Open Government (7%)
■ Open Data for Business (15%)



OPEN DATA IN ACTION!

**11 OTTOBRE
27 NOVEMBRE 2013**

RELAZIONE DI FINE PROGETTO



Open Data: un'opportunità ancora poco sfruttata

Gli **Open Data** sono definiti da *Open Knowledge Foundation* come «*dati che possono essere liberamente utilizzati, riutilizzati e ridistribuiti da ciascuno – con la sola condizione, al massimo, di attribuzione e condivisione allo stesso modo*» (definizione disponibile all'URL: <http://okfn.org/opendata/>)

Essi costituiscono un'opportunità economica ancora poco sfruttata, che può aiutare da un lato le aziende a fare impresa, utilizzando il patrimonio informativo pubblico, e dall'altro la Pubblica Amministrazione a migliorare gli interventi sul territorio. Nuovi servizi privati, basati sui dati pubblici, aprono anche nuovi scenari economici e nuovi mercati, per un valore complessivo stimabile in decine di miliardi di Euro all'anno nella sola Europa (VICKERY G. 2012, *Review of recent studies on PSI re-use and related market developments*, meta-study for the European Commission, disponibile on line all'URL: <http://ec.europa.eu/digital-agenda/en/news/economic-analysis-psi-impacts>).

Emergono nuovi modelli di collaborazione pubblico-privata basati sulla progressiva disponibilità di informazioni e dati pubblici, ai quali anche i singoli cittadini possono contribuire, sfruttando le opportunità del web di mettere a sistema la conoscenza dispersa tra milioni di persone, molte delle quali esperti in particolari settori. Le stesse modalità di collaborazione utilizzate per la redazione collettiva di Wikipedia, possono essere applicati a particolari servizi pubblici.

A queste motivazioni di carattere economico e di efficacia dell'azione pubblica, si somma la considerazione che gli open data consentono di realizzare le istanze di trasparenza e accountability delle politiche pubbliche.

Tuttavia, vi è un problema di accesso a tali banche dati aperti: essi potranno accrescere il proprio valore solo se si forma una massa critica di persone capace di analizzarli e interpretarli. Tale problema è legato al cd. *data divide*: gli Open Data rischiano di essere utilizzati solo dalla ristretta élite di cittadini dotati di quelle competenze adatte a elaborare dati e estrarne il significato.

Il Laboratorio dal Basso **Open Data in Action** ha inteso supportare un'ecosistema imprenditoriale pugliese che, utilizzando gli Open Data, sia in grado di creare valore per la collettività e di generare profitto.



A chi si è rivolto il Laboratorio

Il Laboratorio si è rivolto ad imprese, associazioni, liberi professionisti che utilizzano o vogliono utilizzare gli Open Data nel proprio business, ivi inclusi progetti no-profit.

Il laboratorio, inoltre, si è rivolto a sviluppatori di tecnologie web e mobile che permettono di utilizzare o acquisire e distribuire dati in formato aperto.

Un apposito modulo è stato dedicato anche al Data Journalism Driven, una tematica di interesse per giornalisti, blogger, reporter, ricercatori e studiosi dell'evoluzione della società.

Esso è stato rivolto anche a tutti gli studenti universitari, con lo scopo di sensibilizzarli rispetto a una tematica fortemente di impatto in termini di mercato.

Infine, il laboratorio ha rappresentato un'occasione per far emergere i soggetti che si occupano di Open Data e per facilitare la nascita di reti e cluster d'innovazione nel settore.

Associazione Proponente

L'Associazione **AkuBari** è vincitrice di Principi Attivi 2012 e nasce il 24 giugno 2013. Aku in lingua giapponese significa "aperto, aprire, varco": l'intento dell'associazione di promozione sociale è appunto quello di spingere la città di Bari ad aprirsi verso l'esterno attraverso la valorizzazione del suo patrimonio culturale per attirare e consolidare il flusso turistico attualmente sottodimensionato. Questo avverrà anche sotto forma di data storytelling sul modello di progetti come <http://civic.io/civic-hacking-by-example/>: partendo dai dati a disposizione sulla città e cogliendo l'occasione per richiederne, si potranno realizzare progetti di data journalism di approfondimento per sensibilizzare la cittadinanza sull'importanza degli Open Data, della trasparenza e stimolarne così la partecipazione.

In collaborazione con

L'Associazione **OpenBSK**, vincitrice di Principi Attivi 2012 e costituita il 27 giugno 2013 è focalizzata sugli Open Data. Nasce come un'iniziativa sperimentale per la realizzazione di una piattaforma fruibile tramite interfacce web e Application Programming Interface (API) che aggrega OD e che ne consenta l'analisi e il riuso da parte di servizi, anche terzi. La tecnologia di OpenBSK consiste in un Software as a Service (SaaS) e un Data as a Service (DaaS) che integra acquisizione, archiviazione, analisi, visualizzazione e redistribuzione di OD.

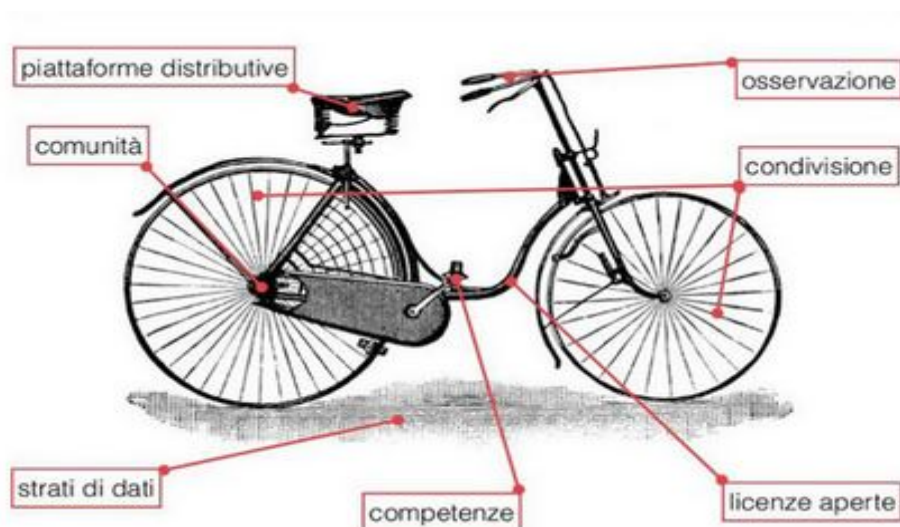
11/10/2013: Introduzione agli Open Data e teoria dei modelli di business sui dati aperti.

Il ciclo del dato.

Docente: LUCA CORSATO

h di lezione: 2

partecipanti: 12



Gli Open Data fanno parte di un ampio e complesso processo di condivisione della conoscenza e dei suoi strumenti. Tutti noi abbiamo il delizioso ruolo di fruitori (quando possibile) ma anche di inconsapevoli produttori. In questo modulo il docente ci ha presentato in modo completo ed esaustivo il ciclo di

produzione e fruizione dei dati aperti. Partendo dalla definizione etimologica della parola "dato" il discorso si è sviluppato intorno alla metafora della bicicletta, la cui parti sono andate a descrivere ogni singolo elemento cardine fondamentale per un virtuoso ciclo di produzione dei dati aperti.

La sella rappresenta le piattaforme distributive e collaborative (wikipedia, OKFN) e i portali di pubblicazione e ricerca di dataset aperti (come ad esempio i sitiweb opendata della pubblica amministrazione). La catena è invece costituita dalle comunità territoriali o dalle community (gruppi di interesse opendata, wikipediani, osmers) che richiedono ed utilizzano i dati aperti: un ruolo fondamentale di distribuzione dell'interesse anche economico che impedisce l'imposizione di soluzioni operative ai soli grandi gruppi economici. La strada è formata dagli strati di dati che sono a nostra disposizione da cui è possibile estrarre informazione e valore, racconto sulla realtà territoriali. Per fare questo tuttavia è necessario avere o raffinare le competenze (i pedali) digitali informatiche o di ricerca. Ciò che regge tutta questa intelaiatura di informazione e competenze sono necessariamente le licenze di utilizzo: senza licenze aperte l'utilizzo dei dati come fonte di



informazione e partecipazione e consapevolezza non ha fondamento.

L'ultima fase per la realizzazione di un ciclo virtuoso di produzione e diffusione dei dati è la condivisione: affinché qualunque processo virtuoso possa davvero partire è necessaria la più ampia e aperta condivisione che permetta un processo di prodigiosa e dirompente intelligenza collettiva come motore finale di innovazione, partecipazione e avanzamento della conoscenza.

Business model theory e policy innovation in Open Data.

docente: MICHELE OSELLA

h di lezione: 2

partecipanti:12

Riuso dei dati pubblici e loro potenziale economico: da un'analisi su alcuni casi studio, il gruppo di ricerca dell'Istituto Superiore Mario Boella di Torino ha dato una prima risposta sul rapporto tra qualità dei dati pubblici rilasciati, e in generale della PSI e la derivante competitività dei soggetti del territorio che si occupano di questo specifico settore economico. L'attività di ricerca condotta è di natura esplorativa e ha comportato la conduzione di 13 casi di studio internazionali i quali sono stati selezionati da una long list di oltre 130 aziende recensite. Da questa prima analisi sembra che dati pubblici aperti e di qualità producano più competitività e diversificazione, in parte smorzando rischi di monopoli e accentramento di capitali.

Tuttavia secondo lo studio effettuato, nonostante le rosee aspettative ed il fervente impulso proveniente dalla comunità scientifica, nel panorama italiano il processo di diffusione e riuso dell'informazione pubblica stenta ancora a decollare. La ricerca è partita da queste domande:

- Esistono organizzazioni orientate al profitto capaci di generare ricavi a partire dall'informazione prodotta dal settore pubblico?
- Quali sono i meccanismi sottesi alla creazione del valore in questo comparto? Come viene elaborata l'informazione grezza nell'ottica di renderla attrattiva ed utilizzabile?
- Alla luce della non rivalità e della non escludibilità al consumo che caratterizzano la PSI quando essa si configura come Open Data, quali sono i fattori di successo che giocano un ruolo nevralgico in chiave competitiva?

Tra le conclusioni di questa ricerca una ci è sembrata molto significativa riguardo l'obiettivo specifico del nostro laboratorio. "Con l'incalzante affermarsi del modello Open Data, sembra emergere all'orizzonte un cambio di paradigma concernente la fruizione dell'informazione. In un mondo contraddistinto da barriere che ostacolano l'accesso dell'informazione, la fruizione di quest'ultima è vincolata ad un esborso finanziario. Con l'avvento di un paradigma spiccatamente orientato all'apertura del dato (c.d. open by default), invece, l'accesso all'informazione avviene gratuitamente mentre forme di pagamento possono essere richieste allo scopo di mantenere riservato il risultato contenuto in opere derivate. La consapevolezza del ruolo cruciale che l'attore

pubblico gioca nella valorizzazione della risorsa PSI ha condotto all'estrapolazione di alcune considerazioni di policy; le suddette indicazioni, benché orientate principalmente alla committenza, potrebbero risultare di interesse anche per altri decisori politici che intendano comprendere più a fondo e sostenere attivamente il ruolo della PSI nella promozione dello sviluppo socioeconomico.”

1

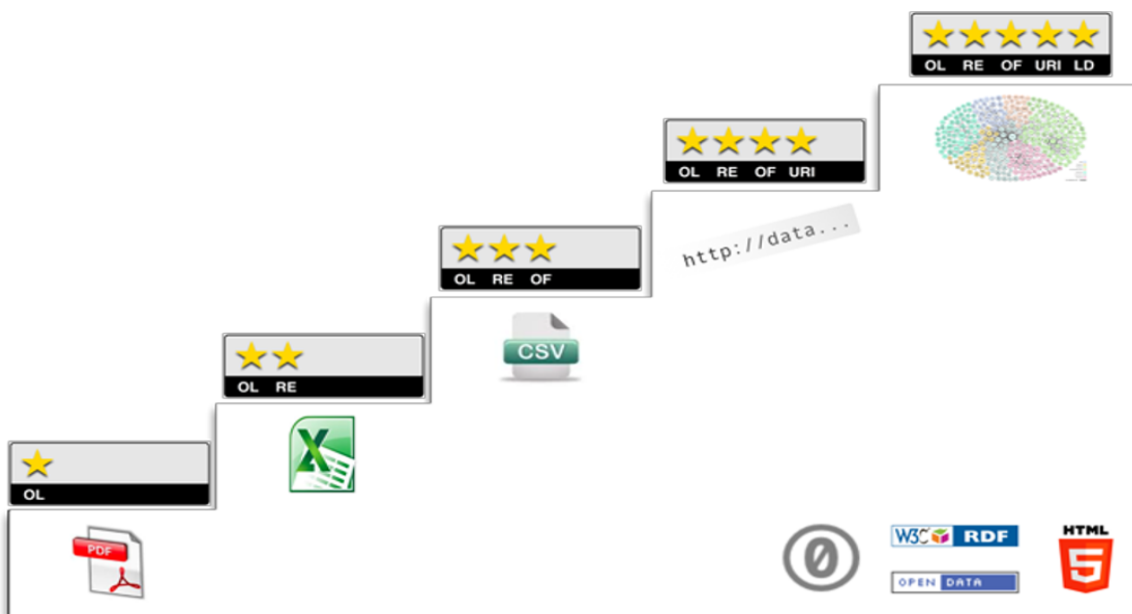
12/10/2013: Case History.

Modelli di business su dati aperti, il caso Spazio Dati, From source to LOD (con sessione pratica).

docente: MICHELE BARBERA

h di lezione: 6

partecipanti: 6



Caso studio di un modello di business sui dati aperti: il caso SpazioDati. Dal racconto del fondatore della giovane Start Up SpazioDati, Michele Barbera, si è approfondita la struttura dei formati di più alta qualità dei dati aperti: RDF e Linked Open Data, utili a connettere tra loro dati di dataset diversi che si trovano in server diversi. E' quindi la chiave per l'interoperabilità. E' l'ideale per relazionare

¹ Enrico Ferro, Michele Osella, *Modelli di Business nel Riutilizzo dell'informazione pubblica*
<http://www.osservatorioict.piemonte.it/it/images/phocadownload/modelli%20di%20business%20nel%20riuso%20dellinformazione%20pubblica.pdf>



grandi moli di dati provenienti da fonti diverse, come importanti istituzioni o reti di imponenti archivi. Ma il loro valore sta anche nel fatto che può collegare tanti dataset piccoli, come per esempio quelli di piccole biblioteche o associazioni. L'attività di SpazioDati si concentra appunto nella creazione di applicazioni web basate su queste tecnologie come per esempio. La prima parte della lezione è stata dedicata ad una ricognizione storica sulla nascita di Internet e sulle tecnologie che sono alla base del suo funzionamento. Nella seconda parte il docente si è soffermato più nello specifico nella descrizione di strumenti e applicazioni utili per la pubblicazione di set di dati in formato linked.

13/10/2013: GeoDati.

Cos'è Open Street Map e strumenti a disposizione per la mappatura geografica di dati aperti.

docente: SIMONE CORTESI

h di lezione: 4

partecipanti: 12

Visualizza Modifica Cronologia entra iscrivi

Gruppo di modifiche: 18330587

[Gruppo di modifiche XML](#) | [modificheOsm XML](#)

« 16330586 | 18330588 » « 18329934 | **simone** | 18331031 »

Creto il: dom 13 ott 2013, 11:21:44 +0000
Chiuso il: dom 13 ott 2013, 11:21:47 +0000
Appartiene a: [simone](#)

Etichette: [comment = bari - parcheggi](#)
[created_by = JOSM/1.5 \(6238 en\)](#)

Riquadro di selezione:

41.130761		
16.8896435	(riquadro)	16.8752061
		41.1214158

Le mappe con cui ci orientiamo ci dicono davvero come è fatto il territorio? Nell'immagine qui sotto potete osservare una parte della città di Bari in diverse versioni, dove vengono messe in evidenza diverse caratteristiche.



In alto a sinistra la foto satellitare mostra la foto aerea di Bari Vecchia.openstreetmap map compare - geofabrik.

In alto a destra la versione standard di OpenStreetMap. In basso a sinistra la versione standard di Google Maps, in basso a destra, invece, una seconda versione OpenStreetMap, costruita a partire dagli stessi dati, ma completamente diversa per aspetto.

OpenStreetMap è una mappa fatta dalle persone, per le persone, senza nessun controllo da parte dell'amministrazione pubblica. Chiunque può aggiungere modifiche migliorative alla mappa.

E' libera, nel senso che ognuno è libero di scaricare i dati sorgenti e di utilizzarli in maniera creativa, senza alcun limite alla fantasia. Molte mappe che normalmente consideriamo libere, in realtà non lo sono, non si ha ad esempio il diritto di scaricare i dati di Google Maps e di utilizzarli all'interno di un progetto aziendale, non si ha il diritto di aggiungere i dati mancanti o di correggere quelli errati.

Questo è un esempio di come si possono inserire dati non previsti dalle amministrazioni, come ad esempio i parcheggi abusivi della città di Bari.





14/11/2013: Open Government. Open Government: il caso Open Polis e best practice.

docenti: VITTORIO ALVINO, FLAVIA MARZANO

h di lezione: 4

partecipanti: 8

Se lasciati soli soletti sui portali governativi loro dedicati, gli Open Data non assicurano che l'istituzione di riferimento stia attuando un piano d'azione di Open Government. Il caso di Open Polis è un eccellente esempio di riutilizzo degli Open Data. Il suo aggiornamento quotidiano e costante rispetto alle attività parlamentari da idea della promozione della coscienza civica che gli Open Data, se ben utilizzati, possono stimolare.

Openpolis si occupa di progetti per l'accesso alle informazioni pubbliche, civic media che promuovono la trasparenza e la partecipazione democratica dei cittadini della rete. Ogni giorno migliaia di persone accedono alle loro piattaforme web per consultare e arricchirne le informazioni, che in origine sono dati pubblici estratti dai siti governativi. Un volume di informazioni che è a disposizione di tutti gratuitamente, e da aggregazione di dati diventa data-driven journalism, cioè un'indagine profonda e continua sui dati aperti che porta alla luce notizie. Questo lavoro ha trasformato in breve tempo Open Polis in un osservatorio civico della politica che analizza quotidianamente i meccanismi complessi che muovono l'Italia.

La prima piattaforma web da cui prendeva il nome l'associazione, Openpolis, oggi Open politici, è del 2007 e conta decine di migliaia di iscritti e una redazione distribuita che monitora e documenta il lavoro dei politici italiani. Voi siete qui, test di orientamento politico-elettorale, ha attirato centinaia di migliaia di persone in occasione delle ultime tornate elettorali. Open parlamento è il sito di maggiore successo con milioni di accessi e una grande notorietà. Open parlamento è un progetto del 2009 e da sito d'informazione è diventato uno strumento di lavoro per centinaia di persone, oltre che un luogo di relazione per i cittadini italiani che spesso ricercano il confronto coi politici proprio tramite il sito. Poi c'è Camere aperte, il rapporto annuale di Openpolis cui è collegato l'indice di produttività parlamentare, il sito dove i dati principali del rapporto vivono in costante aggiornamento. Insieme, questi due strumenti costituiscono una base di informazioni unica per il giornalismo dei dati.

Infine il progetto Open bilanci è finalizzato a "rendere trasparenti", comprensibili, tracciabili e condivisibili dai cittadini i bilanci dei comuni italiani. È un'operazione che attinge ad un enorme patrimonio di dati pubblici, lo storico degli ultimi dieci anni, finalmente accessibili anche ai non addetti ai lavori.



20/11/2013

GeoDati: i sistemi GIS con gli Open GeoDati, tipologie e struttura delle mappe digitali, formati e sistemi per la gestione dei dati cartografici.

Cosa sono i GEO dati: gli Open Data in ambito geografico.

Dopo una breve introduzione al modello dei dati geografici (cosa sono, dove trovarli) ed alle diverse estensioni in cui sono reperibili, si è svolta un'esercitazione sul sistema GIS, condotta importando in Quantum GIS i dati geografici del Comune di Bari.

Si è fatto un merge tra i dati relativi alle circoscrizioni e quelli relativi al numero degli abitanti per ciascuna circoscrizione stessa; sono stati poi importati e diversamente ricombinati i dati sulle postazioni dei siti raccolta farmaci/ pile esauste/ umido ed infine quelli del bike-sharing. Per ottenere la corretta sovrapposizione di due set di dati geografici realizzati con sistemi di coordinate differenti, si è eseguita una trasformazione geodetica per riportarli ad un unico sistema di riferimento. Quantum GIS permette dunque di far confluire dati provenienti da diverse fonti in un unico progetto di analisi territoriale. I dati, divisi in layers, possono essere analizzati e da essi viene creata l'immagine mappa con il graficismo che può essere personalizzato dall'utente e eventualmente rispondere alle analisi tipiche del GIS - gradazione di colori, sfumatura di colore, valore unico. La mappa può essere arricchita da icone e da etichette dipendenti dagli attributi degli elementi cartografici. In questa maniera abbiamo creato varie mappe personalizzate. Tra i plugin che sono stati utilizzati la Mappa di concentrazione, che crea una mappa raster di concentrazione a partire da un vettore di punti, e quello OpenStreetMap, che rende possibile scaricare e visualizzare nella vista i livelli di OpenStreetMap.

Docente: MAURIZIO NAPOLITANO

h di lezione: 4

partecipanti: 11

Mappa project: web-GIS e predittività. l'utilizzo di dati aperti in archeologia

Nel 2011 un gruppo di archeologi, matematici e geologi, paladini dell'Open Source e dell'Open Access, misero in condivisione i propri saperi specifici per studiare strumenti di calcolo predittivo applicabili al potenziale archeologico di un'area urbana, con l'intenzione di creare uno strumento utile alla tutela, alla ricerca ed alla governance della città e del suo patrimonio sepolto.

L'archeologia stratigrafica ha prodotto una enorme mole di dati archeologici, nell'ultimo decennio anche in formato digitale, ma essi giacciono per lo più inutilizzati negli archivi. Nel frattempo sono stati sviluppati anche strumenti di calcolo informatico in grado di processare una grande quantità di



dati digitali.

Dunque, date queste premesse, la sfida del gruppo di menti all'opera per il Progetto Mappa è stata questa: realizzare il primo open digital archaeological archive italiano, rendendo accessibili tutti i dati pubblici delle indagini archeologiche, nella convinzione che garantire l'accesso e la condivisione del patrimonio storico comune sia il mezzo migliore per sviluppare una società della conoscenza veramente libera e democratica. Nell'epoca dell'interattivo Web 2.0, la rete può diventare infatti un potentissimo mezzo di comunicazione in progress delle fasi di lavoro e dei risultati.

Senza annoiare il lettore con lunghissime dissertazioni accademiche, per capire bene cosa si intende con "tutti i dati pubblici delle ricerche archeologiche" è necessaria una piccola digressione su che tipo di documentazione una ricerca archeologica produce.

Lo scavo è un'azione distruttiva, al termine della quale tutto ciò che resta a testimonianza del procedimento interpretativo della stratificazione è la documentazione, unica base analitica attendibile su cui fondare la ricostruzione delle vicende storiche del sito indagato, composta di schede di unità stratigrafica, registri, relazioni, diagrammi stratigrafici, immagini fotografiche e rilievi.

Questa documentazione (apparato archeografico + relazioni) è il "dato grezzo" dell'archeologia e consentirebbe, qualora fosse resa disponibile e liberamente accessibile, di reiterare il processo di interpretazione alla luce delle nuove acquisizioni. A questo scopo le sole pubblicazioni scientifiche, se pure in Open Access, risultano inutili, perché non sono corredate di questi "dati grezzi". L'archeologia è sempre e comunque un'attività di ricerca (e non un ozio intellettuale!), poiché produce dati unici ed irripetibili, ma non vi è ricerca finché non vi è pubblicazione del dato. I "dati grezzi" possono già essere considerati una pubblicazione scientifica, proprio per questo nell'archivio del Progetto Mappa la loro proprietà intellettuale viene tutelata, attribuendo ad essi un codice DOI (Digital Object Identifier), un "ISBN" digitale, ed utilizzando una licenza CC BY/ CC BY SA.

Per il territorio della città di Pisa, la raccolta di questi dati ha consentito di calibrare un algoritmo che serve a stimare la presenza/consistenza dei depositi archeologici. Si è realizzato un webGis in cui sono stati incrociati dati paleoambientali (particolarmente interessanti a Pisa i dati relativi alle paleotracce dei letti fluviali dell'Auser/ Serchio, dell'Arno e dell'Osoli/ Ozzori) e dati storico-archeologici e tramite un modello matematico è stato possibile stimare il potenziale del deposito non ancora indagato, in base a relazioni di tipo spaziale o funzionale fra i ritrovamenti già effettuati. Un esempio di come funziona l'algoritmo Mappa è questo: se è noto il prospetto di una casa-torre di età medioevale, è probabile che nei suoi pressi si conservino tracce della corte interna, dei vicoli laterali e della strada principale su cui il prospetto stesso si affacciava. L'algoritmo agisce in maniera simile a quelli page-rank, utili per assegnare l'importanza delle pagine web nei motori di ricerca (ed applicabile a tutti gli insiemi di oggetti collegati da citazioni e riferimenti reciproci).

Docente: GABRIELE GATTIGLIA

h di lezione: 4



partecipanti: 11

26/11/2013 Introduzione al Data Journalism e modello di business sul Data Journalism.

Docente: NICOLAS KAYSER-BRIL

h di lezione: 4

partecipanti: 9

Attualmente l'innovazione della rete e dei nuovi media stanno creando cambiamenti sconvolgenti nel mondo dell'informazione e nel giornalismo, soprattutto per la carta stampata, in netto calo di vendite tendenzialmente dappertutto. Fra le varie conseguenze che ciò ha avuto, in Italia e non solo, c'è la chiusura di tantissime testate e la drammatica difficoltà per le giovani leve del giornalismo a trovare un impiego. Eppure, mai come oggi l'informazione appare viva e l'interesse del pubblico si conferma sempre ampio. Non sarà forse che il mestiere del giornalista necessita un ripensamento, una ridefinizione?

L'intervento di Nicolas Kayser-Bril, programmatore francese di stanza a Berlino e fra gli ideatori del collettivo europeo Journalism++, illustra con chiarezza e con tanti esempi concreti come gli Open Data su internet possano consentire al giornalista una miriade di approfondimenti e nuove rappresentazioni della realtà che il mezzo tradizionale della stampa non può per sua natura permettere. Non si tratta soltanto di dare vita a colorate infografiche e magari dimenticarsi l'importanza del tradizionale articolo di approfondimento, dell'inchiesta o dell'editoriale. Le moderne tecnologie e il cosiddetto Data-Driven Journalism possono al contrario aggiungere nuovi strati di conoscenza: grazie ai dati aperti presenti sul web, il giornalismo investigativo può raggiungere vette di incredibile ampiezza. Proprio perché la mole di dati è ingente e le possibilità sono potenzialmente infinite, non mancano però le difficoltà.

Come ricavare i dati? Come usarli per offrire rappresentazioni pertinenti e interessanti dei fenomeni in atto? Come saperli usare? Come evitare errori nell'utilizzo e nella loro interpretazione? Alla luce delle difficoltà economiche dei media tradizionali e dei giornalisti, come monetizzare questi dati per i fini giornalistici?

Leggiamo nel manifesto di Journalism++: *"We believe that the presentation of journalism can be vastly improved, that journalists have not been able to think outside of the old boxes that we used to call print, radio and television"*.

La professionalità del giornalista e le sue competenze non sono messe in discussione. Forse con questi presupposti i giornalisti possono pensare a nuove possibilità per reinterpretarsi alla luce di nuovi strumenti e competenze che possono definire questa professione al di là delle polverose regole di accesso all'Ordine.



27/10/2013: Introduzione al Data Journalism.

Case history su elaborazione dati di OpenCoesione.

Docenti: ALINE PENNISI, ANDREA ZEDDA

h di lezione: 4

partecipanti: 5

Il portale OpenCoesione traccia i fondi (europei e nazionali) con cui vengono finanziate le politiche di coesione, che finanziano programmi e singoli progetti con lo scopo di cambiare qualcosa in meglio sul territorio, laddove le politiche ordinarie non arrivano o non possono arrivare. Sulla home page del portale leggiamo che i progetti monitorati sono 709.718, per un finanziamento totale di 65,2 mld di euro.

La natura ed il tema dei finanziamenti sono molto differenti tra di loro: per questo si dice che le politiche di coesione finanziano “dall’ago al reattore”.

I

I dati rilasciati in formato aperto sul portale OpenCoesione possono essere utilizzati per condurre un’analisi a diversi livelli di approfondimento: dal “macro”, valutazione totale delle politiche, al “micro”, valutazione del singolo progetto attraverso la prassi del monitoraggio civico.

Per iniziare ad utilizzarli, è necessario sapere come essi sono strutturati: poiché la piattaforma OpenCoesione è incentrata sui progetti (diversamente da altre piattaforme europee che pure si occupano del monitoraggio dei fondi strutturali ma sono incentrate sul soggetto) diventa fondamentale il CUP, una sorta di codice fiscale nazionale ed il codice locale del progetto, che permette di incrociare le informazioni più facilmente. Su uno stesso progetto per esempio possono convergere più soggetti. Inoltre un progetto può insistere su diversi territori.

I dati scaricabili hanno differenti insiemi di variabili:

- 76 variabili scaricabili sul sito del DPS
- 32 variabili scaricabili sul sito OpenCoesione
- 22 variabili scaricabili sul sito CIPE

Il monitoraggio dei fondi strutturali è di fondamentale importanza, soprattutto in un periodo di forte riduzione di investimenti delle politiche nazionali: dobbiamo saper sfruttare la capacità delle politiche di coesione di finanziare dall’ago al reattore e tradurlo in opportunità, dal momento che, come ha dimostrato un’inchiesta del The Guardian, l’Italia da più soldi all’UE di quanti ne spende!

Alla luce di questa informazione, è risultata quasi indispensabile la successiva esercitazione sono stati esplorati i soggetti attuatori che stanno realizzando i progetti finanziati dalle politiche di coesione in Puglia.

Sono stati scaricati in csv i dataset (più semplici) dalla piattaforma OpenCoesione, filtrando attraverso la natura dell’investimento infrastrutture e territorio (ovviamente Puglia!), ed i dataset (più complessi) dal sito del DPS sui soggetti attuatori per regione.



Si sono pulite le variabili con Open Refine, lasciando solo quelle che interessavano ai fini dell'indagine sui soggetti attuatori e per fare merge tra i due differenti dataset si è sfruttato l'unico elemento comune ed univoco ovvero il codice locale progetto.

08/11/2013: contesto giuridico e licenze.

Docente: MORENA RAGONE

h di lezione: 2

partecipanti: 5

Il lungo cammino che ha condotto al Codice dell'Amministrazione Digitale (CAD) muove i primi passi dalla legge Bassanini (legge n° 59/97), in cui nell'Art. 15, comma 2, si dichiarano validi e rilevanti a tutti gli effetti di legge atti, dati e documenti formati con strumenti informatici e telematici. Da questo articolo deriva il regolamento per la formazione, archiviazione e trasmissione di documenti con strumenti informatici e telematici (DPR 513/1997), sostituito nel 2000 dal testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa (DPR 445/2000).

La conversione in legge dell'Art. 9 del DL 179/2012 (Ulteriori misure urgenti per la crescita del paese), che modifica l'art.52 del CAD, ha reso normativo il principio dell' Open Data by default.

A partire dal 18 marzo 2013 dati e documenti pubblicati online dalle amministrazioni titolari - senza una esplicita licenza d'uso che ne definisca le possibilità e i limiti di riutilizzo – sono da intendersi come dati aperti, quindi dati che possono essere liberamente acquisiti da chiunque e riutilizzabili anche per fini commerciali. A partire dall'analisi dell'Art. 68 del CDA in cui viene esplicitato cosa si intende con dati di tipo aperto, nella seconda parte dell'intervento ci si è soffermati su un momento specifico del ciclo del dato: la scelta della licenza.

Nelle licenze Creative Commons, le condizioni di utilizzo sono quattro e ad ognuna è associato un simbolo grafico per renderne più facile il riconoscimento: attribuzione; condivisione allo stesso modo; non commerciale; non opere derivate. Dalla combinazione delle quattro clausole nascono le licenze CC n uso: CC by; CC by-SA; CC by-ND; CC by-NC; CC by-NC-SA; CC by-NC-ND.

Ma in che misura le licenze sono interoperabili? Come potete vedere dall'immagine successiva, la CC0 non presenta alcun conflitto con altri tipi di licenze. Buona l'interoperabilità anche della CC by e della CC by NC.

Lo scorso 26 Novembre sono diventate operative le licenze CC 4.0, che quindi non sono presenti nelle precedenti immagini. Qui il comunicato ufficiale rilasciato da Creative Commons: queste licenze, che potranno essere utilizzate in tutto il mondo, sono adatte ad "aprire" database ed a pubblicare open government data, anche perché ora comprendono i diritti sui generis di database. Il nostro piccolo percorso attraverso il quadro normativo relativo ai dati aperti si è concluso con il recente decreto sul "Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e



diffusione di informazioni da parte delle pubbliche amministrazioni" (Dlgs 33/2013). Questo decreto, che si profila quasi come testo unico degli obblighi di trasparenza e pubblicità a carico degli enti pubblici, con l'Art. 5 ha riconosciuto ai cittadini un diritto di accesso più ampio sull'attività amministrativa, introducendo il diritto di "accesso civico" ed attribuendo ai cittadini stessi il ruolo di "guardiani" sull'effettiva pubblicazione dei dati. Ci si è cimentati nella compilazione di un'istanza di accesso civico, predisposta dal gruppo di Spaghetti Open Data in occasione del Festival Internazionale del Giornalismo ("Cara PA ti scrivo...", alla cui stesura la stessa Morena ha dato un contributo fondamentale).

Data Mining

Docente DUCCIO SCHIAVON

ore 4

partecipanti 14

Avere a disposizione delle informazioni non equivale automaticamente a possedere della conoscenza. L'overload di dati che vengono generati ogni giorno non ci consente di stare al passo con l'informazione che viene costantemente rilasciata in rete, a meno che non ci affidiamo a tool automatici o a tecniche di analisi statistica. Il Data Mining è l'insieme di questo tipo di tecniche e di 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 di tale sapere in ambito scientifico, industriale o operativo.

La risorsa migliore dove trovare riferimenti sul data mining è kdnuggets.com: qui è possibile reperire programmi software, news, dataset, opportunità di lavoro, aziende che si occupano del tema, corsi, incontri e webinar.

E' significativa la definizione di data mining fornita dal prof Adelchi Azzalini del dipartimento di scienze statistiche dell'Università di Padova: "Il data-mining rappresenta l'attività di elaborazione, in forma grafica o numerica, di grandi raccolte o di flussi continui di dati con lo scopo di estrarre informazioni utili a chi detiene i dati stessi". Una tipico esempio di attività di data mining (nella quale siamo inconsapevolmente coinvolti tutti i giorni) lo ritroviamo all'interno dei supermercati quando effettuano la "market basket analysis", cioè quel processo di analisi di affinità che analizza le abitudini di acquisto dei clienti nella vendita al dettaglio al fine di trovare associazioni su i diversi prodotti comprati. Tale processo è utile per l'adozione di strategie di marketing ad hoc.

Il più famoso risultato ottenuto applicando la tecnica del "market basket analysis" è quello di "birra e pannolini", che asserisce che per chi compra pannolini esiste statisticamente un'alta probabilità che compri anche birra. La spiegazione razionale di tale asserzione è data dal fatto che, avendo un bambino, con maggiore probabilità i genitori tenderanno a restare a casa con il proprio figlio. A partire da questo storico risultato i supermercati si sono evoluti e, soprattutto attraverso la grande distribuzione, ad oggi vengono fatte analisi basate sugli scontrini e sulla geo-localizzazione grazie alle carte fedeltà.

Al pari dei supermercati, anche le compagnie telefoniche effettuano raccolte dati in base all'utilizzo



registrato dai propri utenti, offrendo nuovi piani tariffari ed individuando nuovi segmenti di clienti target. L'uso delle tecniche di data mining non si limita alle sole indagini di mercato, ma spazia anche nel campo della ricerca scientifica in microbiologia dove vengono utilizzate per sequenziare il DNA. Inoltre il data mining viene utilizzato anche per le previsioni meteo.

Tuttavia, la vera sfida che bisogna affrontare è quella di applicare le tecniche di data mining ai dati presenti su web, (attività probabilmente molto impegnativa): ci sono così tanti dati che non sono facilmente numerabili, la gran parte dei quali è organizzata in forma non-strutturata (testi, valori senza metadati che li descrivano, audio, video, ecc..).

Inoltre all'interno di un singolo documento, gli elementi oggetto di ricerca non sempre sono pertinenti, come ad esempio nel contenuto di una qualsiasi pagina web, dove è possibile avere diversi tipi di contenuti che si riferiscono anche ad ambiti diversi.

Come espresso in precedenza in altri termini, il problema è la "dimensionalità" dei dati: l'abbondanza di informazioni non consente l'estrazione di conoscenza utile (termine volutamente generico perchè quasi mai nel Data mining è specificato a priori l'oggetto di interesse, che si individua proprio mentre si scava nei dati). Ne risulta che, tutto sommato, il data mining è una disciplina giovane, collocata nel punto di intersezione di varie aree scientifiche, specialmente quelle relative alla statistica e all'informatica.

Dato che l'obiettivo spesso non è dichiarato a priori la situazione classica nella quale ci si ritrova è che, andando alla ricerca di nuove informazioni, si finisce col trovare comunque qualcosa, anche se questo qualcosa in realtà non esiste.

A tal proposito, non potevano mancare alcune citazioni di personaggi celebri in questo ambito:

"If you torture the data long enough, nature will always confess"

"Se torturi i dati abbastanza a lungo, la natura confesserà sempre"

(George E. P. Box)

"All models are wrong but some are useful"

"Tutti i modelli sono sbagliati, ma alcuni sono utili"

(Ronald Coase)

In conclusione, il data mining è un processo complesso di identificazione di tendenze nei dati, con lo scopo di cercare modelli o trend validi che siano potenzialmente utili e comprensibili e che consentano all'utente di prendere decisioni cruciali.

Per chi si volesse cimentare con il data mining, un tool molto utile ma semplice da imparare a utilizzare è il [Weka](#), sviluppato dal "Machine Learning Group" dell'università di *Waikato*.

Strumenti Google per la raccolta di dati aperti.



Grazie al GDG Bari è stato possibile portare ad Open Data in Action Salvino Fidacaro e Francesco Passantino. Salvino Fidacaro è il manager del GDG Nebrodi, esperto di web marketing e SEO che con la sua azienda si occupa soprattutto di web, digital e social marketing.

Il credo della sua azienda è “Non importa quale canale viene utilizzato ma è importante la strategia, l’esecuzione e come misurare l’impatto del nostro lavoro. Consideriamo i mezzi interattivi come un’opportunità unica di gestire e misurare ogni singolo aspetto della relazione con il consumatore.”

Francesco Passantino invece è il manager del GDG Palermo, ricercatore, consulente e formatore, nell’ICT. Molto attivo nel panorama startup di Palermo, è inoltre l’organizzatore dello Startup Weekend Palermo e si prodiga nella diffusione della cultura degli Open Data sul territorio palermitano.

Il Knowledge Graph di Google

Docente Salvino Fidacaro

Ore 2

Presenze

Il Knowledge Graph (in italiano "grafico della conoscenza") è una funzione di ricerca che è stata introdotta il 16 maggio 2012 in google.com, mentre nella versione italiana è stata attivata il 4 dicembre 2012.

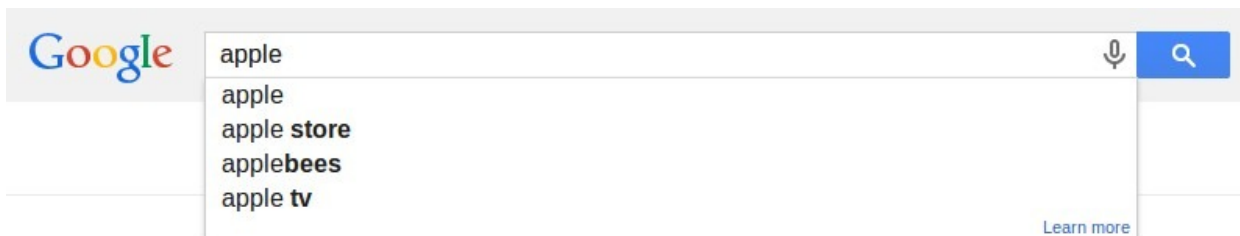
Il Knowledge Graph (KG) è il primo passo verso una ricerca semantica: grazie a questa funzione, il motore di ricerca di Google associa alle parole cercate un oggetto e metterà in relazioni oggetti in modo da avere una ricerca più veloce e accurata. Come è cambiato l’approccio ai dati da parte di Google e degli utenti con l’introduzione di questa tecnologia?

Attualmente il KG contiene più di 700 milioni di oggetti, questo non si attiva su ogni ricerca, ma solo su quelle per cui riconosce di possedere del contenuto.

In Italia la percentuale di ricerche che coinvolge il KG è attorno al 4% contro il 9% di quelle negli Stati Uniti. Il KG inoltre viene sfruttato dall’applicazione Android Google Now.

Attualmente l’algoritmo di ricerca di Google valuta più di 200 parametri, quello che oggi fa il SEO (acronimo di search engine optimization) è cercare nel contesto semantico delle parole.

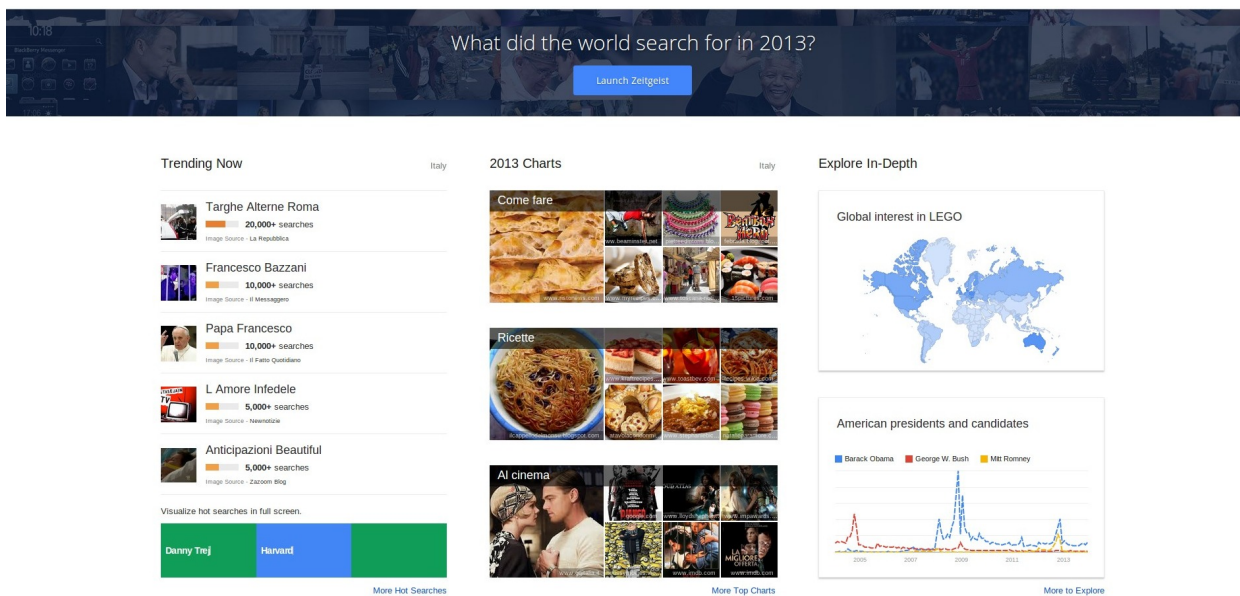
Cosa vuol dire? Vuol dire trovare quelle parole associate alla parola su cui vogliamo essere indicizzati il più in alto possibile nelle ricerche effettuate dagli utenti. Per trovare queste parole si può utilizzare Google Instant, la funzione di autocompletamento delle ricerche di Google.



Posso così creare una mappa delle relazioni tra le parole, ad esempio se voglio essere indicizzato sulla parola “Ipad”, la mia rete di parole comprenderà: Apple, viewer, prezzo.

Un'altra caratteristica importante nell'approccio alle ricerche è la stagionalità di certe parole chiave, pensate ad esempio alla parole “regalo” o simili nel periodo natalizio.

Per fare un'analisi di questo tipo uno strumento molto utile è [Google insight for search](#) (Google aka Trends)



Analizzando nello specifico il KG, come si costruisce e quali sono i risultati?

Google ovviamente non fa trapelare molte informazioni al riguardo, ma alcune di queste si possono desumere da best practice o tramite reverse engineering.

Ad esempio è fortemente consigliato utilizzare all'interno del tag a di un link l'attributo “author” per identificare l'autore della risorsa linkata.

Un forte contributo a ciò che viene valutato dal KG deriva dalle attività degli utenti su Google+:

- le condivisioni hanno un peso maggiore dei +1

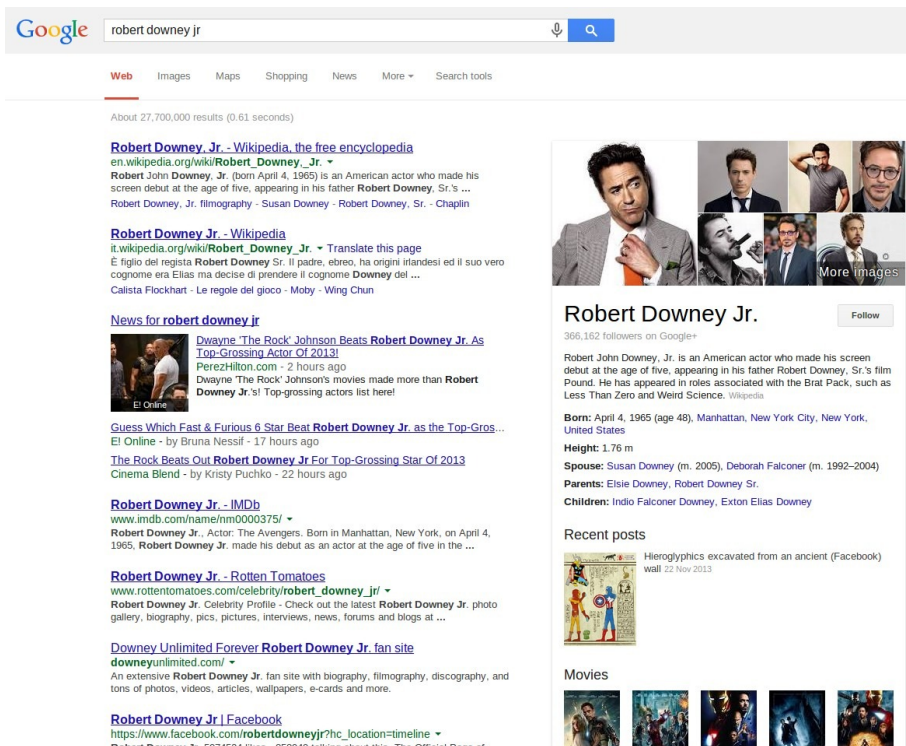
- i post sui profili personali sono più efficaci dei post come pagina

Qual'è il risultato del KG?

Analizziamo alcuni esempi tipici, cerchiamo una persona, un film ed un luogo.

Cerchiamo ad esempio un personaggio pubblico come Robert Downey Jr, Google, oltre alla classica lista di siti web, ci fornirà sulla destra un box come è possibile vedere nello screenshot seguente.

Il contenuto del riquadro è un'entità del KG a cui ne sono associate altre (ad esempio i film in cui ha recitato)



Se invece cercassimo il film Iron Man 3, il box sulla destra ci restituirebbe l'entità del film con descrizione, data di uscita e le entità correlate come i membri del cast, con ancora il solito Robert Downey Jr.



Google ironman 3

Web Images Maps Shopping Videos More Search tools

About 71,100,000 results (0.57 seconds)

Iron Man 3 - Marvel.com
marvel.com/ironman3
IRON MAN 3: THE OFFICIAL GAME In this intense endless runner game, you'll have access to an amazing arsenal of Iron Man suits, including iconic MK42 ...

Iron Man 3 (2013) - IMDb
www.imdb.com/title/tt1300854/
★ ★ ★ ★ Rating: 7.4/10 - 305,676 votes
Directed by Shane Black. With Robert Downey Jr., Guy Pearce, Gwyneth Paltrow, Don Cheadle. When Tony Stark's world is torn apart by a formidable terrorist ...
Full Cast & Crew - (2013) Poster - Cast & Crew - Trivia

Iron Man 3 - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Iron_Man_3
Iron Man 3 (stylized onscreen as Iron Man Three) is a 2013 American superhero film featuring the Marvel Comics character Iron Man, produced by Kevin Feige ...
Ty Simpkins - Krakoa - Extremis - Mandarin


Iron Man 3 Teaser Trailer UK - Official Marvel UHD - YouTube
www.youtube.com/watch?v=5EjG-1U3wqA
Oct 23, 2012 - Uploaded by MarvelUK
Watch the first official Iron Man 3 trailer. Marvel's Iron Man 3 - coming to UK cinemas April 24th 2013 ...

Iron Man 3 - Movie Trailers - iTunes
trailers.apple.com/trailers/marvel/ironman3/
Marvel's Iron Man 3 pits brash-but-brilliant industrialist Tony Stark/Iron Man against an enemy whose reach knows no bounds. When Stark finds his personal ...

Iron Man - Iron Man 3: Fan Power | Facebook
https://www.facebook.com/ironman/app_47025606352542
Iron Man. 16896940 likes · 61850 talking about this. Marvel's Iron Man 3 in theaters May 3, 2013.

Iron Man 3 - Rotten Tomatoes
www.rottentomatoes.com/m/iron_man_3/
Review: With the help of its charismatic lead, some impressive action sequences, and even a few surprises, Iron Man 3 is a witty, entertaining...

Iron Man 3 Reviews - Metacritic
www.metacritic.com/movie/iron-man-3
★ ★ ★ ★ Rating: 62% - 44 reviews








Iron Man 3
2013 Film
★ ★ ★ ★ 7.4/10 - IMDb
★ ★ ★ ★ 62% - Metacritic


Iron Man 3 is a 2013 American superhero film featuring the Marvel Comics character Iron Man, produced by Kevin Feige of Marvel Studios and distributed by Walt Disney Studios Motion Pictures. [Wikipedia](#)

Initial release: April 14, 2013 (Paris)
Director: Shane Black
Running time: 130 minutes
Initial DVD release: September 24, 2013 (USA)
Featured song: Blue (Da Ba Dee)

Cast

 Robert Downey Jr. Iron Man	 Gwyneth Paltrow Pepper Potts	 Ben Kingsley Mandarin	 Don Cheadle War Machine	 Jon Favreau Happy Hogan
--	--	---	--	---

People also search for



In fine cerchiamo la città di San Francisco, il box questa volta ci restituirà informazioni geografiche fornendoci già una preview della mappa, informazioni demografiche e meteo, oltre alla solita lista di entità correlate quali prossimi eventi e foto dei punti di interesse.


Google San Francisco

Web Images Maps Shopping News More Search tools

About 1,280,000,000 results (0.82 seconds)

San Francisco - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/San_Francisco
San Francisco (Spanish for "Saint Francis") was founded on June 29, 1776, when colonists from Spain established a fort at the Golden Gate and a mission ...
Consolidated city-county - History of San Francisco - San Francisco Bay Area

News for san francisco


House under construction collapses in San Francisco
San Jose Mercury News - 29 minutes ago
SAN FRANCISCO - Authorities say a house under construction has collapsed in the hilly San Francisco neighborhood of Twin Peaks.


Keeping key vets vital for 49ers
ESPN (blog) - 3 hours ago

San Francisco's rock and roll landmarks
San Francisco Chronicle (blog) - 1 hour ago


San Francisco Travel Guide: Things to Do, Hotels, Events ...
www.sanfrancisco.travel/
Share your unique experience in our amazing city. Tag your SF Instagram photos with the hashtag #MagicHourSF and your photo could appear on our SF ...
Things to do in San Francisco - Visitor Information Center - Contact Us - About

Welcome to SFGOV City and County of San Francisco Official site
www.sfgov.org/
SFGOV is the official website of the government of the City and County of San Francisco, providing information about departments, meetings, legislation, ...

Images for san francisco - Report images



San Francisco City Guide | Hotels, Restaurants, Nightlife, Real ...
www.sanfrancisco.com/
A complete San Francisco, CA Travel & Tourism Guide specializing in hotels, attractions, restaurants, real estate, nightlife and local business information.








San Francisco
City in California






San Francisco, officially the City and County of San Francisco, is the leading financial and cultural center of Northern California and the San Francisco Bay Area. [Wikipedia](#)

Weather: 12°C, Wind NW at 0 km/h, 64% Humidity
Local time: Tuesday 8:44 AM
Population: 825,863 (2012)
Colleges and Universities: San Francisco State University, More

Upcoming events

 San Francisco 49ers vs... Dec 23	 Cirque du Soleil Amaluna Dec 20	 Chance The Rapper's... Dec 18	 The Book of Mormon Dec 21	 Holograms - TV Ghost - G. Green Dec 18
--	---	---	---	--

Points of interest

 Golden Gate Bridge	 Fisherman's Wharf	 Alcatraz Island	 Golden Gate Park	 San Francisco cable car...
---	--	--	---	---

Come è possibile desumere da questo breve racconto, anche Google sta adottando un approccio



sempre più strutturato ai dati, costruendo dataset e mettendoli in relazioni con quelli già disponibili online sia sotto forma di open data che come dataset proprietari e quelli che vediamo ora sono solo i primi risultati di quello che è possibile fare nell'ambito delle ricerche dei dati.

2. Strumenti Google per trattare gli open data & best practice from Palermo

Docente Francesco Passantino

Ore 3

Presenze

Sul web sono disponibili una moltitudine di strumenti che si possono utilizzare per manipolare i dati indipendentemente che vogliamo renderli disponibili in formati aperti o meno.

[Open Refine](#) è sicuramente il più importante, prodotto inizialmente dalla società Metaweb (acquisita nel 2010 da Google) è stato in seguito rilasciato in modalità open source ed è uno strumento che consente di pulire, manipolare e collegare i dati tra di loro.

Uno strumento molto utile è [Table Capture](#), un'estensione di Chrome che consente di estrarre tabelle da pagine web, per consentire di elaborarne i dati.

Anche Google Spreadsheet della suite di Google Docs fornisce una serie di funzione per lavorare sui dati, in particolare ricordiamo:

- la possibilità di determinare la lingua di un testo
- la possibilità di tradurre un testo
- la possibilità di importare un feed rss

SeoTools invece è un'estensione di Microsoft Excel che consente di fare analisi sui sui like di facebook, sugli share di twitter e sui +1 di Google+ per quanto riguarda un link che gli viene fornito.

Avviando un discorso di standardizzazione dei contenuti sul web, Google ed altri grandi player del settore hanno avviato iniziative come [sitemaps.org](#) che definisce un protocollo per descrivere la mappa di un sito web e [schema.org](#) che fornisce una serie di schemi, tag html e markup che i webmaster possono utilizzare per migliorare la SEO delle pagine.

[Fusion Table](#) è un'applicazione web e un'estensione per Chrome fornito da Google per la gestione dei dati, fusion table fornisce strumenti per la visualizzazione di grafici, mappe e timeline in maniera intuitiva.

In particolare ad esempio georeferenzia in automatico gli indirizzi senza limiti di utilizzo.

Un altro strumento a disposizione per trattare dati è [Public Data Explorer](#) che permette di inviare



grosse moli di dati ed interrogarli tramite interfaccia. E' necessario però accreditarsi come data provider.

Altre risorse ugualmente importanti sono:

- freebase (dataset di Google)
- OpenDataKit (un google forms per app Android)
- Schoolofdata.org
- geojournalism handbook
- datajournalism handbook

A conclusione del suo intervento Francesco ci ha parlato inoltre dell'esperienza palermitana sugli open data, ed è notizia di questi giorni l'approvazione da parte del Comune delle [linee guida](#) riguardanti gli Open Data.

Data Visualization

Alessio Cimarelli, Data Visualization in d3.js

Ore 4

Presenze 33

La Data Visualization è la creazione e lo studio della rappresentazione visiva dei dati.

Secondo Friedman, "l'obiettivo principale della Data visualization è quello di comunicare in modo chiaro ed efficace le informazioni attraverso mezzi grafici. Ciò non significa che la visualizzazione dei dati ha bisogno di essere noioso per essere funzionale o estremamente sofisticati per apparire bella. Per comunicare efficacemente idee, forma estetica e funzionalità devono andare di pari passo, fornendo approfondimenti in una piuttosto radi e complessi set di dati, comunicando i suoi aspetti chiave in un modo più intuitivo. Eppure i progettisti spesso non riescono a raggiungere un equilibrio tra forma e funzione, creando stupende visualizzazioni di dati che non riescono a servire il loro scopo principale, comunicare informazioni.

Infatti , Fernanda Viegas e Martin M. Wattenberg hanno suggerito che una visualizzazione ideale non solo dovrebbe comunicare in modo chiaro , ma stimolare l'impegno e l'attenzione spettatore.

La visualizzazione dei dati è strettamente legato alla grafica di informazioni , la visualizzazione delle informazioni, visualizzazione scientifica e grafica statistici . Nel nuovo millennio , la visualizzazione dei dati è diventata un'area attiva di ricerca, insegnamento e sviluppo.

KPI Biblioteca ha sviluppato la " Tavola periodica dei metodi di visualizzazione ", un grafico interattivo , visualizzando le varie modalità di visualizzazione dei dati . Esso comprende sei tipi di metodi di visualizzazione dei dati. Dati, le informazioni , il concetto , la strategia, la metafora e composto.

D3.js è una libreria JavaScript per la manipolazione di documenti basati su dati . D3 aiuta a portare i



dati a vita usando HTML , SVG e CSS . L'enfasi di D3 su standard web vi offre tutte le funzionalità dei browser moderni, senza legarsi a un quadro di proprietà , che unisce potenti componenti di visualizzazione e di un approccio data-driven per la manipolazione del DOM.

Data Visualization

Andrea Nelson Mauro, Data Visualization per il Data Journalism

Ore 4

Presenze 8

In questo incontro il docente ha analizzato il processo di gestazione di un lavoro giornalistico data driven simulando un workflow di ricerca dati, cleaning, visualizzazione e ricerca della storia. La visualizzazione dei dati è la trasformazione di dati e conoscenze in forma visuale. IL suo obiettivo è in primo luogo esaminarne i contenuti, capirne le meccaniche ed ottenere prodotti finali che permettano di percepire in maniera immediata le informazioni in essi presentati. Oltre che al lavoro finale di presentazione è importante visualizzare i dati anche nella fase di costruzione della storia. Con delle prove pratiche abbiamo

provato ad associare ogni tipo di grafico ad una fase di ricerca di un ipotetico articolo.

- **Grafico a linee:** Il più usato per mostrare i cambiamenti o gli andamenti nel tempo.
- **Istogramma:** Quando i valori di ciascun anno, o di ciascuna serie storica in genere, non hanno una relazione reciproca gli uni con gli altri, o quando ad esempio è importante mostrare la crescita di alcuni valori, è preferibile utilizzare un istogramma piuttosto che un grafico a linee.
- **Grafico a barre:** Questa variante è molto utile se hai molte etichette da mostrare. Qui le informazioni sono nella parte sinistra della visualizzazione, in modo da lasciare più spazio e rendere più facile e piacevole la lettura.
- **Grafico a torta:** è utile se vuoi mostrare dati che sommati insieme rappresentano un totale (e presi singolarmente rappresentano le parti del totale) – come ad esempio tutti gli smartphone venduti in un paese.

L'ultima parte della lezione è stata invece dedicato alla spiegazione dello strumento di visualizzazione per mappa CARTODB.