

 Contenuto archiviato il 2024-04-23

Articoli di approfondimento - Grandi dati al vostro servizio

L'aumento dell'uso delle TIC per gli affari, l'intrattenimento e i servizi pubblici sta portando all'accumulo di montagne di dati strutturati e in molti casi non strutturati. Questi cosiddetti "grandi dati" dovrebbero essere visti come un'opportunità e non un problema. La ricerca e le attività dell'UE per promuovere i dati aperti stanno aiutando a dare un senso e fare buon uso di questa risorsa.



"Dati aperti" è un concetto secondo il quale i dati dovrebbero essere liberamente accessibili a tutti, in modo che tutti possano usarli e ripubblicarli come desiderano, senza limiti di diritti di autore, brevetti o altri meccanismi di controllo. Ispirati al software open-source (non proprietario) e ai movimenti ad accesso aperto (editoria accademica), i dati aperti sono intesi come il libero movimento, uso, ri-utilizzo o distribuzione elettronica dei dati.

Una parte importante di questo movimento di "grandi dati" è l'uso a più ampio vantaggio della società delle informazioni non personali che i cittadini condividono con lo stato e i servizi pubblici. I dati statali aperti sono una risorsa importantissima che deve ancora essere sfruttata pienamente. "Lo stato raccoglie una vasta quantità di dati di alta qualità nell'ambito delle sue attività di lavoro ordinarie. Se questi dati vengono aperti, possono portare enormi vantaggi," osserva il sito Open Government Data (OGD), amministrato dal gruppo di lavoro per un governo aperto.

Secondo Rufus Pollock della Fondazione conoscenze aperte, rendere accessibili i dati permette alle aziende, ai singoli e al settore no-profit di costruire applicazioni e servizi utili e promuove la democrazia, la partecipazione al governo, la trasparenza e la responsabilità. "Perché non aprire i dati che sono già presenti e che si raccolgono

già?" dice.

Ci sono però non poche difficoltà - legali, tecniche, sociali e legate al mercato - che devono essere superate prima che i molti benefici dei dati governativi aperti possano essere efficacemente trasferiti ai cittadini.

Avanti in questa direzione...

Secondo le relazioni della recente Future Internet Assembly (FIA) tenutasi a Aalborg, Danimarca, "Tendenze come "grandi dati" e "internet delle cose" (internet of things o IoT), comprese "persone come sensori", stanno mostrando come cittadini/imprenditori/innovatori possono sviluppare nuovi servizi e applicazioni a beneficio delle città intelligenti." Il presentatore della FIA, Reinhard Scholl dell'Unione internazionale delle telecomunicazioni (UIT), dice che tra i buoni esempi ci sono l'iniziativa Dati aperti della Città di New York, il programma Smart City di Amsterdam, l'Open Data Gencat della Catalogna e la sfida Open Cities della Commissione.

Le migliori pratiche provenienti dagli USA, secondo Scholl, comprendono l'esperimento "Track Trash" del MIT che ha usato i sensori per monitorare dove va a finire la spazzatura. Mentre il servizio basato sui dati "scopri il crimine" di Oakland, ha detto, sta aiutando la città a migliorare la sicurezza.

Le informazioni del servizio pubblico (PSI) sono la più grande fonte di informazione in Europa, secondo il "DG Connect" della Commissione europea e comprendono cartine digitali e dati meteorologici, legali, sul traffico, finanziari, economici e altro. La maggior parte di questi dati grezzi può essere riutilizzata o integrata in nuovi prodotti e servizi per uso quotidiano, come sistemi di navigazione per automobile, previsioni del tempo, servizi finanziari e di assicurazione.

Il "riutilizzo delle informazioni del settore pubblico significa usare i dati in nuovi modi aggiungendovi valore, mettere insieme le informazioni provenienti da fonti diverse, fare mash-up e nuove applicazioni, a fini commerciali e non. Le informazioni del settore pubblico hanno un grande potenziale economico," spiega la Commissione nella sua [pagina web dedicata alle PSI](#) 

La ricerca UE si adatta al cambiamento

Anche il panorama della ricerca si è spostato per far posto ai rapidi cambiamenti che stanno avvenendo nella raccolta, l'elaborazione e la gestione dei dati. Per esempio, progetti finanziati nell'ambito delle attività del 7° PQ "Tecnologie per la gestione delle informazioni", come parte del tema "Contenuti e conoscenze", si sono occupati di una serie di campi di ricerca che comprendono i contenuti online, i media interattivi e sociali, lo sfruttamento del ragionamento e delle informazioni e la scoperta e la gestione della conoscenza.

Un'iniziativa, il progetto finanziato dall'UE Weknowit (["Emerging, collective intelligence for personal, organisational and social use"](#)) , ha sviluppato una piattaforma che converte ampi contenuti generati dall'utente da un problema di sovraccarico di informazioni a una nuova "intelligenza collettiva" con una serie di applicazioni, dalla gestione delle emergenze a un migliore turismo cittadino. Il progetto si è candidato per diversi brevetti e una manciata di prodotti e risultati sono destinati per il rilascio pubblico o commerciale.

"Usando una grande varietà di strumenti, la piattaforma Weknowit trasforma informazioni su larga scala e non ben strutturate in argomenti significativi, entità, punti di interesse, connessioni sociali ed eventi," dice il coordinatore del progetto Yiannis Kompatsiaris dell'Istituto di informatica e telematica (CERTH-ITI), Laboratorio di conoscenze multimediali in Grecia. Per fare ciò, i partner hanno sviluppato un'applicazione di middleware che si può usare su server per elaborare i dati che arrivano e instradarli efficacemente.

Hanno anche sviluppato diversi strumenti nell'ambito dei casi di studio del progetto, come uno scenario di reazione di emergenza e uno scenario di gruppo sociale del consumatore, mentre i partner hanno creato una dozzina di altri strumenti per compiti specifici. Nel frattempo, i partner CERTH-ITI, Yahoo! e l'Università di Koblenz stanno collaborando su aspetti in tempo reale per l'estrazione delle informazioni dei social media e stanno esaminando applicazioni nel settore delle informazioni e dei grandi eventi, come i festival del cinema.

Dati aperti anche per la scienza

Un uso migliore di dati strutturati è vantaggioso anche per la ricerca scientifica in modo più diretto, grazie ai progressi del cloud e grid computing o del supercalcolo. L'investimento europeo nelle e-infrastrutture, che usa la capacità "inutilizzata" dei computer distribuiti in tutto il mondo, significa che i ricercatori possono elaborare e analizzare insieme di dati più grandi rispetto a prima, svelando possibili risposte ad alcuni dei maggiori interrogativi della scienza, dalla fisica quantica ai modelli di cambiamenti climatici

Per esempio, i biologi che studiano uno specifico problema potrebbero creare un "ambiente virtuale di ricerca" (AVR) per collaborare nel grid, elaborare informazioni da una fonte in Estonia e analizzarle con strumenti software di data mining provenienti da un'altra fonte, per esempio, in Portogallo.

Muovendosi un passo più avanti, un progetto finanziato dall'UE, chiamato D4Science-II (["Data infrastructures ecosystem for science"](#)) , ha creato un quadro interoperabile per le e-infrastrutture che è come un ecosistema nel quale dati, risorse di calcolo e software appartenenti a diverse infrastrutture elettroniche possono

essere condivisi a prescindere dal luogo, dalla tecnologia, dal formato, dalla lingua, dal protocollo o dal flusso di lavoro. Il loro ecosistema ha sostenuto AVR in campi come la fisica ad alta energia, la biodiversità, la pesca e le risorse di acquacoltura. Ha aiutato ad aprire nuove aree di ricerca tra di loro e si sta espandendo verso nuovi settori.

Per esempio, D4Science-II ha sostenuto lo studio di mappatura delle specie marine Aquamaps. Aquamaps aiuta gli scienziati a incrociare i dati riguardanti la biodiversità marina con i registri del pescato per avere un quadro più chiaro di dove le riserve di pesce sono maggiormente a rischio. È un enorme esercizio di elaborazione di dati e numeri reso possibile grazie ai finanziamenti europei per le infrastrutture elettroniche e alla sua politica di dati aperti e di iniziative di ricerca.

"La cooperazione tra e-infrastrutture apre possibilità e campi di ricerca completamente nuovi. Possiamo analizzare dati scientifici rispetto a statistiche economiche, per esempio, per ottenere una prospettiva tutta nuova che prima non era disponibile," dice Donatella Castelli, un partner di D4Science-II presso l'Istituto di scienza e tecnologie dell'informazione (Alessandro Faedo) del Consiglio nazionale della ricerca italiano.

Editoria ad accesso aperto

Mentre le organizzazioni pubbliche stanno aprendo i loro dati ai ricercatori, potrebbe sembrare ironico che i risultati di tale ricerca finiscano per diventare inaccessibili in costose riviste. Nel tentativo di promuovere più pubblicazioni online ad accesso aperto - specialmente per la ricerca finanziata pubblicamente - la Commissione europea ha reso la pubblicazione ad accesso aperto obbligatoria per circa il 20% dei progetti del 7° PQ.

Inoltre, quando i progetti pubblicano una parte dei risultati in una serie di riviste tradizionali e una parte in pubblicazioni ad accesso aperto, le conoscenze sono frammentate ed è meno facile misurare il risultato prodotto da un progetto. Il progetto sostenuto dall'UE Openaire (["Open access infrastructure for research in Europe"](#))  si è proposto di cambiare le cose con l'idea di rendere tutto accessibile a tutti.

Il team di Openaire ha capito subito che la tecnologia non è che la metà della lotta per superare la frammentazione della ricerca e dei dati: "Una parte significativa del progetto si occupa di promuovere l'open access nella comunità del 7° PQ," dice Natalia Manola, manager del progetto, "sostenendo la pubblicazione ad accesso aperto in modo che i progetti possano contribuire appieno all'infrastruttura del sapere dell'Europa."

Con l'aiuto di progetti come Openaire e il suo follow-up, Openaireplus, la pubblicazione ad accesso aperto può dare una spinta all'economia e ai livelli di

innovazione dell'Europa, secondo Natalia Manola. Per gli impiegati di piccole imprese o per gli insegnanti, il costo dell'abbonamento a riviste scientifiche di alto livello può essere proibitivo e questo significa che della ricerca preziosa rimane inaccessibile. "Con l'open access, chiunque [può] usarla come desidera – è il modo migliore di usare al meglio la ricerca finanziata con fondi pubblici," conclude.

Riguardo a ciò, l'organizzazione per la ricerca nucleare, CERN, ha guidato un progetto finanziato dall'UE sullo studio dell'editoria ad accesso aperto chiamato SOAP ("[Study of open access publishing](#)"). in cerca di modelli di business sostenibili per promuovere l'editoria accademica. Il team ha documentato oltre 4.000 riviste e in seguito a qualche analisi, il team di SOAP ha scoperto che circa l'8% della produzione mondiale di articoli scientifici, o circa 120.000 articoli l'anno di una stima di 1,5 milioni, viene attualmente pubblicato come open access. Hanno concluso che un modello ibrido di open access (parzialmente basato sugli abbonamenti) è l'opzione più fattibile specialmente per l'editoria scientifica e di ricerca.

"Potenziando i modelli realizzabili di open access, i ricercatori europei - e in generale tutto il mondo – trarranno vantaggio dallo scambio di conoscenze e avranno accesso a un ampio materiale," secondo un rapporto CORDIS, "[Accesso a montagne di ricerca](#)", su SOAP.

I dati parlano

Anche se rendere accessibili i dati di proprietà pubblica, mettere insieme set di dati e pubblicare i risultati ad accesso aperto sono tutte cose vantaggiose per la scienza, monetizzare i dati strutturati commercialmente è una sfida più complessa. Alcuni progetti lanciati dall'UE stanno esaminando questa questione.

Il progetto finanziato dall'UE CODE ("[Commercially empowered linked open data ecosystems in research](#)"). è un'iniziativa guidata da PMI che si occupa del contenuto digitale e della parte linguistica dell'equazione dei grandi dati. I "dati aperti collegati" ("linked open data" o LOD) hanno un potenziale enorme come prossimo grande passo evolutivo di internet secondo il team di CODE. Questo potenziale rimane però largamente inutilizzato a causa della mancanza di strategie di utilizzo e di monetizzazione.

CODE, che ha cominciato a lavorare solo quest'anno, sta sviluppando un solido ecosistema per commercializzare i LOD sulla base di una catena di creazione di valore tra ruoli tradizionali (per es. fornitore e consumatore di dati) e non tradizionale (per es. analista di dati) sui mercati dei dati. I primi risultati sembrano promettenti.

Considerato il fatto che viviamo sempre più la nostra vita online, i partner del progetto finanziato dall'UE Limosine ("[Linguistically motivated semantic aggregation engines](#)"). stanno cercando di usare la lingua e la tecnologia di ricerca semantica per

migliorare questa esperienza online.

"Informazioni si accumulano su un'ampia gamma di attività umane, dalla scienza e i fatti ai contenuti personali, le opinioni e le tendenze," osserva il team del progetto. Il sistema multilingue di ricerca di opinioni di Limosine significa che internet può spostarsi dall'attuale ricerca incentrata sui documenti verso una maggiore aggregazione semantica. In altre parole, ottenere risultati di ricerca raffinati più velocemente attraverso strumenti più intelligenti che capiscono meglio e prevedono cosa si sta cercando.

Per esempio, se si cerca "dog's breakfast" (colazione per cani) usando gli standard di ricerca attuali si ottengono risultati su idiomi britannici o sul teatro canadese, quando invece una persona non madrelingua inglese probabilmente cercava più letteralmente un'alternativa sana da dare al proprio cane invece dei cereali! Gli strumenti di ricerca semantici possono essere in grado di contestualizzare la richiesta sulla base delle ricerche precedenti o di altre prove raccolte in precedenza.

Nel frattempo, progetti come [LIVE+GOV](#)  riunisce "Percezione della realtà, mining e aumento per un dialogo mobile cittadini-stato". Il progetto sta sviluppando una soluzione di "m-gouvernement" che permette ai cittadini di esprimere le loro esigenze allo stato attraverso tecnologie di percezione mobile già presenti negli smart phone, insieme a formati già accettati di e-partecipazione mobile.

Lubrificazione l'economia europea

Alla fine, i dati pubblici, generati da tutte le amministrazioni in Europa, dovrebbero diventare automaticamente ri-utilizzabili e stimoleranno l'innovazione e l'imprenditoria, che a loro volta alimentano nuove applicazioni e servizi, sia fissi che mobili.

"Così come il petrolio è stato assimilato all'oro nero, i dati prendono una nuova importanza e un nuovo valore nell'era digitale," ha commentato Neelie Kroes, vicepresidente della Commissione europea responsabile dell'Agenda digitale in occasione del lancio della Strategia per i dati aperti dell'UE a dicembre. Questo pacchetto per i dati aperti informa la nuova direttiva PSI che è adesso davanti al Consiglio e al Parlamento europei.

L'informazione del settore pubblico genera già circa 32 miliardi di euro di attività economiche ogni anno. Il nuovo pacchetto ha intenzione di far salire tale cifra fino a 70 miliardi di euro, il che secondo Neelie Kroes è "una spinta di cui la nostra economia ha estremo bisogno". Neelie Kroes ha mostrato il suo apprezzamento per il Regno Unito, la Danimarca e la Francia per le loro iniziative sui dati aperti e ha detto che la nuova strategia "cambierà radicalmente" il modo in cui le istituzioni dell'UE e la maggior parte degli enti pubblici in Europa condividono i loro dati.

Ha inoltre invitato gli stati a non aspettare che questo pacchetto diventi legge: "Potete condividere i vostri dati oggi e generare entrate e posti di lavoro e persino risparmiare grazie alla migliore informazione e le decisioni che ne deriveranno." Ha incoraggiato il settore privato ad aprire i loro dati per generare nuovi servizi. "I dati sono oro ... Cominciamo a estrarli!" ha esortato.

Tutti i progetti presentati in questo articolo sono stati sostenuti dal Settimo programma quadro (7° PQ) per la ricerca o da altre iniziative europee.

Link utili:

- [Europa 2020](#) 
- [FIA](#) 
- [Programma delle TIC del 7° PQ](#) 
- [7° PQ su CORDIS](#) 
- [Informazioni del settore pubblico - pagina web su Europa](#) 
- [Open Government Data](#) 
- [NTC Open Data](#) 
- [Amsterdam Smart City](#) 
- [Open Data Gencat](#) 
- [Open Cities](#) 
- [Science Commons](#) 
- [Weknowit](#) 
- [CODE](#) 
- [Limosine](#) 
- [Openaire](#) 
- [SOAP](#) 
- [D4Science-II](#) 
- [LIVE+Gov](#) 

Articoli correlati:

- [Data mining 3.0: dalle info all'"intelligenza collettiva"](#)
- [Accesso libero: i risultati dei progetti dell'UE diventano pubblici](#)
- [Accesso aperto a montagne di ricerca](#)
- [Salvare gli ecosistemi con ecosistemi di dati aperti e infrastrutture digitali](#)
- [I dati sono il nuovo oro, Neelie Kroes sul discorso Open Data Strategy, dicembre 2011](#) 

Progetti correlati



ARCHIVED

Live+Gov

Reality Sensing, Mining and Augmentation for Mobile Citizen-eGovernment Dialogue

5 Aprile 2023

PROGETTO

Questo articolo è contenuto in...

RIVISTA RESEARCH*EU

Catastrofi naturali e cambiamenti climatici: come la scienza si aspetta l'inaspettato

Ultimo aggiornamento: 8 Agosto 2012

Permalink: <https://cordis.europa.eu/article/id/89052-feature-stories-big-data-at-your-service/it>

European Union, 2025