

Scienziato, prendi i miei dati

Viola Bachini

La condivisione dei dati rappresenta per molti un dilemma o perfino un enigma: il dilemma è rappresentato dalla capacità di disporre di una notevole quantità di dati acquisiti da strumentazioni sofisticate e la necessità di selezionarli, elaborarli e magari confrontarli con altri provenienti da fonti e/o settori disciplinari diversi al fine di acquisire nuove conoscenze. L'enigma deve invece essere sciolto a partire da un'incerta definizione di dato, analizzando scopi e processi di acquisizione, metodi scientifici e obiettivi da raggiungere ma anche modalità di elaborazione e conservazione proprie di specifici settori scientifici e sotto-gruppi disciplinari



Non solo scienziati, tutti possono dare un contributo significativo alla ricerca. Di *citizen science*, “cittadinanza scientifica”, si parla già da qualche anno: chiunque dovrebbe avere accesso alle informazioni e anche chi non è esperto in una certa disciplina dovrebbe essere coinvolto nei dibattiti e nei processi decisionali.

Ma si può andare anche oltre. Con la diffusione della rete, infatti, aumentano le possibilità di collaborazione tra scienziati e cittadini, in cui questi ultimi possono essere parte attiva.

I dati sull'accesso alla rete parlano chiaro: secondo il *The Guardian*, quest'anno circoleranno 1,03 miliardi di nuovi smartphone. In testa ci sono Cina e India, mentre l'Italia si piazza poco dopo i primi dieci posti. Un aspetto degno di nota per il nostro Paese riguarda il fatto che ben il 67% dei nuovi dispositivi finirà nelle tasche di connazionali che non hanno mai posseduto un cellulare dotato di una connessione internet.

Come può la tecnologia avvicinare i cittadini a un progetto scientifico fino a farli partecipare attivamente alle scoperte? Per esempio attraverso le app che raccolgono le segnalazioni degli utenti per progetti scientifici. Ce ne sono già diverse, da quella che chiede di fotografare e catalogare piante e animali incontrati durante le passeggiate nei boschi a quella che invita a scovare un nuovo corpo celeste.

I ricercatori si trovano così a disposizione grandi banche dati, alimentate da migliaia di persone che non lavorano con loro nei laboratori, ma che di fatto condizionano la buona riuscita della ricerca. Per alcune di queste, il gioco potrebbe valere una menzione su riviste del calibro di *Nature* come è accaduto allo scopritore di Orione e Sirio, le prime particelle di

polvere stellare identificate da una missione della Nasa.

Per la mole di dati da analizzare e il grande numero di appassionati, al momento la disciplina che ha avuto più successo con questo nuovo concetto di *citizen science* è l'astronomia. Che è stata anche la prima a lanciare un progetto partecipato.

CERCANDO TRA LE STELLE

Era il 1999 quando gli scienziati del programma *Seti at home* chiedevano l'aiuto dei cittadini. Lo scopo del progetto era quello di analizzare dati radioastronomici provenienti dallo spazio a caccia di segnali di intelligenze aliene. Per riuscire nell'impresa gli scienziati avevano bisogno di una grande potenza di calcolo, impossibile da realizzare con i soli calcolatori del centro di ricerca. Per superare il problema tecnico si pensò allora di coinvolgere volontari da tutto il mondo, che tramite internet avrebbero formato una rete di personal computer in grado di aiutare i ricercatori. Per aderire al progetto era sufficiente scaricare un software che, quando il computer non era dedicato ad altro, si collegava al server dell'iniziativa *Seti* e procedeva nell'elaborazione dei dati astronomici.

Fu un grande successo. Dopo un solo anno il programma era stato scaricato da più di due milioni di persone che, senza mai muovere i pc dalle loro scrivanie, avevano donato al progetto un totale di oltre 400.000 anni di elaborazione dati. Nel 2001 *Seti at home* entrò addirittura nel *Guinness World Record* come il più grande calcolo della storia.

Alla fine il progetto non raggiunse lo scopo di trovare intelligenze extra terrestri,

ma anche se le aspettative scientifiche non furono soddisfatte, il suo ruolo nella definizione dell'importanza della collaborazione tra scienziati e non esperti resta innegabile. Eppure ancora non si sfruttava tutto il potenziale di questa



La disciplina che ha maggiore successo sul fronte della citizen science è l'astronomia

collaborazione. Ciò che comunque mancava al progetto Seti era un coinvolgimento degli astrofili non solo per l'elaborazione ma anche per la produzione di dati. Sarà infatti dopo qualche anno che i progetti scientifici non chiederanno più i pc delle persone ma i loro occhi e il loro cervello. Il primo a compiere questo passo è stato *Galaxy Zoo*, nato con l'obiettivo di classificare più di un milione di galassie fotografate dal telescopio robotico della SDSS (*Sloan Digital Sky Survey*). La SDSS è la più estesa carta digitale del cielo, nata alla fine degli anni Novanta da una collaborazione americano-giapponese e che oggi copre circa un quarto del cielo, fino a una distanza di 1,5 miliardi di anni luce dalla Terra.

A differenza dell'elaborazione dei segnali radio, per il riconoscimento delle galassie non esiste nessun software in grado di fare il lavoro meglio dell'occhio umano. Da qui l'idea dei ricercatori dell'Università di Oxford: perché non affidare il censimento a un migliaio di volontari?

Gli scienziati misero così a disposizione un sito internet e un breve tutorial. Quando fu lanciato, nel 2007, nella sola prima giornata di operazioni sul portale arrivarono classificazioni di galassie al ritmo di circa 70.000 all'ora e in un anno si contarono quasi 150.000 partecipanti. Cittadini appassionati di astronomia che in solo dodici mesi riuscirono a catalogare circa un milione di immagini di galassie. Hanny van Arkel, giovane maestra olandese, riuscì addirittura a scoprire un oggetto astronomico di natura insolita che oggi porta il suo nome.

Dopo *Galaxy Zoo* si sono susseguiti diversi progetti di astronomia partecipata, molti promossi dalla Nasa, come *Planet Hunters*, che invita chiunque lo voglia a scrutare le curve di luce delle stelle per avvistare il transito di pianeti extrasolari. O *Be a Martian*, il sito in cui l'agenzia spazia-

le americana si serve del contributo dei cittadini per raffinare la mappa del Pianeta Rosso e dei suoi crateri. E in futuro? Il *Large Synoptic Survey Telescope* è un potente telescopio che si trova sul picco del Cerro Pachón, una montagna del Cile alta 2.682 metri. Lo strumento sarà in funzione dal 2018, e quando inizierà le acquisizioni produrrà 10 milioni di gigabyte di dati all'anno. Per poter elaborare tutte queste osservazioni sarà necessario superare le limitazioni dovute alla disponibilità e alla potenza dei calcolatori, ma saranno soprattutto necessari nuovi approcci e nuove soluzioni per riuscire a estrarre le informazioni a un ritmo paragonabile a quello con cui i nuovi dati vengono acquisiti. Che la soluzione sia, di nuovo, nel coinvolgimento degli astrofili?

SALUTE E SORVEGLIANZA PARTECIPATA

Anche l'ambito sanitario si sta aprendo alla raccolta diretta dei dati. Si chiama *participatory surveillance*, letteralmente "sorveglianza partecipata", e comprende tutte quelle pratiche in cui le persone decidono di condividere in rete e rendere pubbliche informazioni che le riguardano. Se nella sorveglianza partecipata i cittadini aderiscono su base volontaria e sono consapevoli dell'uso che viene fatto dei loro dati non si può parlare di invasione della privacy. Anzi, spesso queste raccolte dati sono volte al miglioramento della salute e della qualità dell'ambiente in cui le persone vivono e queste sono disposte a mettersi in gioco, anche quando non ne traggono direttamente un vantaggio, se c'è fiducia nell'istituzione che promuove il progetto.

Un esempio positivo di sorveglianza partecipata è rappresentato dalle *Health Map* dell'Organizzazione Mondiale della Sanità, che tutti i giorni raccolgono le segnalazioni delle patologie infettive da ogni angolo del pianeta. In Italia *Influweb* è un progetto scientifico per raccogliere informazioni sull'influenza grazie a volontari in tutto il paese che riportano il proprio stato di salute (l'aggiornamento viene fatto ogni settimana durante tutta la stagione influenza-



le). Registrandosi sul sito si entra nella rete di sorveglianza e si ricevono *reminder* settimanali, mentre una mappa interattiva mostra il numero dei casi di sindromi influenzali per ciascuna provincia aggiornato in tempo reale. «Non sembra essere troppo lontano il momento in cui, sul giornale, oltre alle previsioni del tempo, leggeremo in anticipo anche quanti casi di influenza ci saranno nelle varie regioni d'Italia», si legge sul sito del progetto. Questa visione potrebbe trasformarsi presto in realtà grazie a una collaborazione del team di Inluweb con l'Università di Boston. Il metodo matematico messo a punto dai ricercatori prevede l'utilizzo dei dati epidemiologici raccolti sulla piattaforma partecipativa e sarà in grado di fare previsioni accurate sull'andamento dell'influenza.

Ci sono casi in cui per raccogliere dati una piattaforma internet non basta. La ricerca sulle malattie genetiche, per esempio, necessita di campioni biologici, la cui raccolta e conservazione è decisamente più complessa. Tuttavia, con una buona organizzazione, anche per questo tipo di ricerca è possibile coinvolgere volontari.

Nel nostro Paese un caso virtuoso è quello dell'Associazione Ring 14, che si occupa di una sindrome molto rara che colpisce principalmente nel primo anno di vita. Stefania Azzali, presidente dell'associazione, racconta le difficoltà nell'incentivare la ricerca sulla malattia. «Ci siamo resi conto che dovevamo mettere a disposizione degli scienziati centinaia di campioni biologici, troppi per poter essere conservati da un solo istituto di ricerca». Così nel 2009 sono stati i primi in Italia ad attivare un accordo con la Biobanca dell'Ospedale Galliera di Genova per la raccolta di campioni di sangue e cute dei malati e dei loro familiari. Grazie a un paziente lavoro di reclutamento – «che non sarebbe stato possibile senza i social network», tiene a precisare Azzali – la Biobanca oggi conta 300 campioni, tutti raccolti con un kit spedito via corriere dall'associazione. Una collezione così ampia è molto appetibile per il mondo della ricerca e sta attirando scienziati di fama mondiale. «Bisogna sradicarsi dai legami con un unico centro

di ricerca perché più c'è apertura e più i ricercatori saranno interessati», conclude Azzali.

BANCHE DATI AMBIENTALI

Sono diverse, anche in Italia, le amministrazioni che hanno deciso di mettere in rete i database territoriali (che riguardano l'ambiente di una certa area), in modo che chiunque possa consultarli o elaborarli. Le potenzialità degli open data (dati aperti, scaricabili da internet) sono enormi, specialmente nell'ambito politico-amministrativo e di pianificazione degli spazi urbani e rurali. Perché allora non far partecipare alla raccolta dei dati territoriali chi in quelle zone ci vive? Uno dei parametri ambientali che preoccupa maggiormente i cittadini è l'inquinamento atmosferico. Gran parte delle città, infatti, fallisce nel rispettare le linee guida Oms sulla qualità dell'aria. Nel mondo, solo il 12% delle persone residenti in città in cui si effettua il monitoraggio, respirano aria pulita. Le città del nord Italia, secondo l'ultimo rapporto della Commissione ambiente europea, sono tra le peggiori in Europa.

A Firenze, dove la situazione non è tra le più rosee, c'è una centralina per il monitoraggio della qualità dell'aria che registra in continuo dati sulle concentrazioni delle polveri fini. Quello che la differenzia da tutte le altre sparse per la città è che non è stata l'Arpa a installarla, ma un gruppo di cittadini residenti nell'area, che volevano avere una fotografia accurata e il più possibile realistica dello stato di salute

 Il progetto "Firenze PM2.5" è inserito nei casi studio monitorati nei progetti del centro europeo *Joint Research Center*

ambientale del loro quartiere. «C'era una carenza di informazioni disponibili», commenta Annibale Biggeri, epidemiologo e statista dell'Università di Firenze. Biggeri fa parte della squadra di "Firenze PM2.5", l'attività di sorveglianza ambientale condotta dai cittadini e realizzata con il supporto della

Cooperativa Epidemiologia e Prevenzione G.A.Maccacaro. Lo scorso gennaio i dati raccolti per un anno dalla centralina e comparati con quelli istituzionali sono stati presentati durante un'assemblea pubblica. «Per salvaguardare l'ambiente i cittadini si sono fatti carico della produzione di conoscenza scientifica. E i dati che hanno raccolto sono di ottima qualità», commenta Biggeri. L'affidabilità di quei numeri è testimoniata dal fatto che anche l'Arpa Toscana ha deciso di utilizzarli, riconoscendo valore all'iniziativa. L'esperienza di "Firenze PM2.5" è attualmente oggetto di studio del *Joint Research Centre* della Commissione Europea come esempio positivo di sorveglianza diretta e di co-produzione di dati.

Un altro interessante esperimento di monitoraggio atmosferico partecipato è in corso in provincia di Napoli. Il centro di Ricerca Enea di Portici ha messo a punto un naso elettronico che i cittadini possono utilizzare per conoscere in tempo reale la situazione degli inquinanti atmosferici in città. Si tratta di un sistema portatile a basso costo e facilmente utilizzabile, che può essere montato su uno zaino o legato al manubrio della bicicletta e collegato a un'applicazione per smartphone. Lo scopo di Monica (Monitoraggio Cooperativo della Qualità dell'aria) è quello di favorire comportamenti virtuosi. Durante le passeggiate, infatti, gli utenti potranno identificare le aree maggiormente inquinate e condividere, tramite le piattaforme social, i percorsi alternativi per minimizzare l'esposizione. Così come per il caso di Firenze, anche i dati raccolti dal dispositivo portatile dell'Enea andranno ad affiancare le informazioni delle centraline di monitoraggio già presenti in città.

Quando si parla di ambiente, il contributo dei cittadini non è limitato al monitoraggio dell'aria. Esiste infatti anche una app pensata per trasformare il proprio telefono in un rilevatore mobile di rumori. Il vantaggio rispetto a una centralina fissa è che oltre a registrare il livello di rumore, è possibile descriverne la fonte e il grado di fastidio percepito. Tutte le misurazioni, poi, possono essere condivise sulla mappa dell'inquinamento acustico della città. Grazie a questo progetto i ricercatori avranno a disposizione una mappa dell'inquinamento acustico globale che potrebbe da utilizzare per capire come le persone percepiscono e reagiscono ai rumori. Una mappa che potrebbe essere utile anche agli urbanisti impegnati nell'abbattere questo fenomeno.