

Big data e fake news ai tempi della COVID

A colloquio con **Walter Quattrociocchi**

Direttore del Laboratory of Data Science and Complexity, Università Ca' Foscari, Venezia e Coordinatore del gruppo di lavoro Web data e impatto socioeconomico per la Task Force Dati per l'emergenza COVID-19 voluta dal Ministero dell'Innovazione

Mai come nel periodo che stiamo vivendo il tema dei big data è diventato essenziale per gestire al meglio la pandemia e in particolare la difficile fase di convivenza con il virus. La loro analisi si può infatti trasformare in leva delle strategie che i diversi governi stanno mettendo a punto per uscire dalla situazione di emergenza. "Ecco quindi che grandi corporazioni come Facebook e Google cominciano a collaborare con le istituzioni e con le università, mettendo a disposizione il loro petrolio per risolvere problemi relativi al bene comune", ci spiega Walter Quattrociocchi in questa intervista. La sfida è quella di riuscire a far interagire il mondo degli scienziati con quello della politica, perché i risultati della scienza si possano trasformare in suggerimenti politici efficaci.



Facebook ha aperto la porta dei suoi dati ai ricercatori impegnati nel contrasto all'epidemia attraverso il programma "Data for Good". Vuole spiegarci di che cosa si tratta e perché è utile poter utilizzare queste informazioni?

In questo momento i dati hanno lo stesso valore dell'oro o del petrolio. Non era mai accaduto prima. E il fatto che grandi corporazioni come Facebook e Google cominciano a collaborare con le istituzioni e con le università, mettendo a disposizione "il loro petrolio" per risolvere problemi relativi al bene comune è secondo me un passaggio molto importante.

Il mio gruppo di ricerca è entrato in contatto con il progetto "Data for Good", promosso da Facebook, che ha messo a disposizione dei ricercatori i dati di mobilità dei propri utenti. Volendo entrare più nel dettaglio di questo progetto per quello che riguarda l'Italia, dobbiamo immaginare il nostro territorio coperto da una griglia costituita da celle larghe più o meno 500 x 500 metri. Ogni 8 ore viene rilevato il numero di persone che si trova all'interno della cella e il numero di quanti si sono spostati da una cella all'altra. La privacy è completamente preservata perché nessuno sa chi si è spostato, ma si conosce solo il numero di persone che si spostano da una cella all'altra.

La possibilità di accedere a queste informazioni ci ha permesso di compiere delle analisi strutturate. Nel momento in cui si conosce il dato, si possono infatti cominciare a costruire delle matrici di indagine che permettono di superare l'approccio puramente speculativo e di essere più solidi nelle argomentazioni.

Abbiamo quindi potuto costruire una matrice della mobilità italiana pre e post lockdown, costituita da una rete che connette tutti i Comuni, rappresentati da dei nodi. L'arco tra questi nodi identifica le persone che si spostano. In questo modo sappiamo quante persone si sono spostate da un Comune all'altro. Costruendo delle metriche di efficienza, dove per efficienza intendiamo "quanto è facile raggiungere ogni altro posto a partire da un qualunque posto sulla rete", abbiamo notato che con il lockdown diventa ovviamente più difficile raggiungere i posti. Le persone si muovono di meno e sono facilitati la comunicazione e il movimento nel vicinato a discapito delle distanze lunghe. Questo è un effetto endogeno di tutti gli choc e di tutte le catastrofi naturali.

A quali risultati economico-sociali siete giunti con questo vostro studio?

L'analisi dei dati sulla mobilità per il mio gruppo di ricerca è stata utile su due fronti. Il primo è quello dello studio delle dinamiche di diffusione del virus. Una delle assunzioni che abbiamo fatto in partenza è che la mobilità sia in qualche modo un buon indicatore del contatto e di conseguenza del contagio. Bisogna però stare molto attenti all'interpretazione di questo dato, perché se c'è tanta

mobilità, ma tutti si muovono in maniera separata e ben divisa, la possibilità di contagio è pressoché nulla.

Abbiamo costruito dei modelli simulativi matematici per disegnare eventuali scenari di uscita dal lockdown e abbiamo riscontrato che mantenere la distanza sociale tra la categoria degli anziani e quella dei giovani minimizzerebbe di molto la possibilità di ripartenza del virus.

Il secondo fronte è stato quello delle ripercussioni economiche del lockdown. Preso un determinato Comune, ne abbiamo studiato la mobilità prima e dopo il lockdown, e abbiamo visto che i Comuni più colpiti dal delta, ossia quelli che hanno un delta di differenza maggiore nella mobilità, sono i Comuni con indice di deprivazione più alto. Quindi quelli fondamentalmente più poveri e collocati nelle aree più periferiche del Paese.

Come coordinatore del team Web data e impatto socio-economico della Task Force voluta dal Ministero dell'innovazione ha avuto davanti a lei la sfida di trasformare gli strumenti scientifici, e in particolare l'analisi dei big data, in leve della strategia del governo nella gestione della pandemia. Il nostro Paese è preparato a tutto ciò?

Uno degli effetti 'positivi' generato da questa grande tragedia della pandemia è di essersi resi conto che non possiamo affrontare questa emergenza come si è fatto in passato, ma che abbiamo anche altre possibilità rappresentate dall'utilizzo dei dati. Finora abbiamo avuto difficoltà a impiegarli principalmente per due motivi. Il primo è che in Italia chi si occupa oggi di data science spesso proviene da altri settori e quindi riutilizza in questo contesto gli strumenti che impiegava nelle di aree competenza originarie in modo alcune volte positivo altre completamente anacronistico. Dall'altra è in corso una forte diatribe, di carattere accademico, su chi debbano essere i giusti interpreti della data science: gli informatici, i fisici o gli statistici? Questo perché in Italia è ancora molto difficile mettere in atto l'idea di lavorare insieme per il bene comune. Si lavora tutti per piccoli potentati che cercano di prevaricare sugli altri a discapito poi una cooperazione e questo allontana dall'obiettivo principale che sarebbe quello di risolvere il problema.

In questo contesto l'iniziativa del Ministero dell'Innovazione di creare una task force che, riunendo competenze diverse, si occupi di analisi dei dati per supportare il governo nella messa in atto di politiche per contrastare la COVID mi è sembrata un'idea eccellente, quasi necessaria. Ovviamente questa iniziativa si è portata dietro anche i problemi che ho descritto prima. Non è così facile infatti creare un cluster di persone provenienti da background differenti che sia effettivamente in grado di operare in maniera attiva, proficua ed efficiente. C'è poi un problema di linguaggio perché quello scientifico è diverso da quello politico e i risultati della ricerca scientifica vanno masticati per essere poi trasformati in suggerimenti politici. Far sì che questa task force sia in grado di interagire direttamente con l'interlocutore politico è un altro passaggio molto delicato.

La sua formazione professionale la rende esperto anche di comunicazione, con una particolare attenzione al monitoraggio delle

fake news. A questo proposito lei stesso ha definito la pandemia come la tempesta perfetta. Vuole spiegarci perché e se si sta riuscendo, secondo lei, a contrastare efficacemente la diffusione delle false notizie?

No, non ci stiamo riuscendo per una serie di motivi.

Il punto è che abbiamo tre livelli di comunicazione tra loro interconnessi e in azione in questo momento: la comunicazione istituzionale, che per sua stessa natura deve essere la più lucida e sobria possibile; il filone giornalistico, che elabora a volte in parte anche l'informazione istituzionale; e infine i flussi di comunicazione sui social, che fungono da amplificatori.

La comunicazione istituzionale ha più o meno cercato di offrire delle indicazioni rispetto a una situazione nuova e molto complessa e articolata. Ha quindi cercato di fare il meglio che poteva con quello che aveva disposizione.

La comunicazione giornalistica ha conosciuto momenti di vittoria, ma anche di sconfitta, dovendo fronteggiare una tematica che ha praticamente monopolizzato l'informazione. Il problema è che la pandemia si è rivelata mediaticamente come la tempesta perfetta, perché rompe il velo di Maya. Ci siamo cullati per anni con la falsa credenza, o meglio la fake news, che la scienza fosse qualcosa di monolitico e in grado di dare risposte certe. Questa convinzione è stata in qualche modo avallata da personalità autorevoli, che discutono nei salotti e nei talk show, facendo passare un messaggio molto sbagliato di cosa sia la dinamica scientifica. La dinamica scientifica è semplicemente un tentativo di "capirci qualcosa", ma questo sfugge ai più. Nel momento in cui è stata creata un'aspettativa quasi dogmatica rispetto alle risposte che la scienza è in grado di dare, scoprire che invece gli scienziati possono avere posizioni diverse e che ricorrono a regole terze per cercare di mettersi d'accordo attraverso il metodo scientifico (alla faccia de "la scienza non è democratica") ha alimentato in realtà un meccanismo di sfiducia verso la scienza.

Questa scienza, che era stata venduta come religione al punto che molti attendevano con trepidazione il rituale liturgico delle 18.00 in attesa dei risultati della Protezione Civile, si è dimostrata invece solo come un metodo onesto per cercare di trovare delle risposte. Intorno all'incertezza rispetto al tema COVID, perché di fatto lo stiamo conoscendo adesso, si è creato un circo mediatico nel quale hanno trovato posto anche tutte le informazioni fasulle. È possibile contrastare questo meccanismo? Assolutamente no, perché più è alta l'incertezza maggiore è la probabilità che l'argomento diventi in qualche modo materia di diffusione di informazioni false o fuorvianti o strumentali.

A proposito di questa pandemia si è creato un marasma in cui persone sconosciute al mondo accademico sono diventate esperte di epidemiologia o gente che faceva altri mestieri si è proposta come opinionista senza avere alcuna competenza. Tutto questo a scapito del sistema scientifico in sé, perché la scienza, poi, quando deve parlare non parla, facendo però parlare chiunque. Questo è l'altro problema. ■ ML