

IN COPERTINA SCIENZA E DEMOCRAZIA

Non solo Cambridge analytica l'altra faccia dei big data

Le tracce digitali che seminiamo in rete sono sterminate. I colossi del web le sfruttano, spesso a nostra insaputa, per fare profitti. Ma è possibile un uso alternativo con ricadute positive sulla qualità della vita di tutti, racconta Dino Pedreschi tra i massimi esperti nel campo

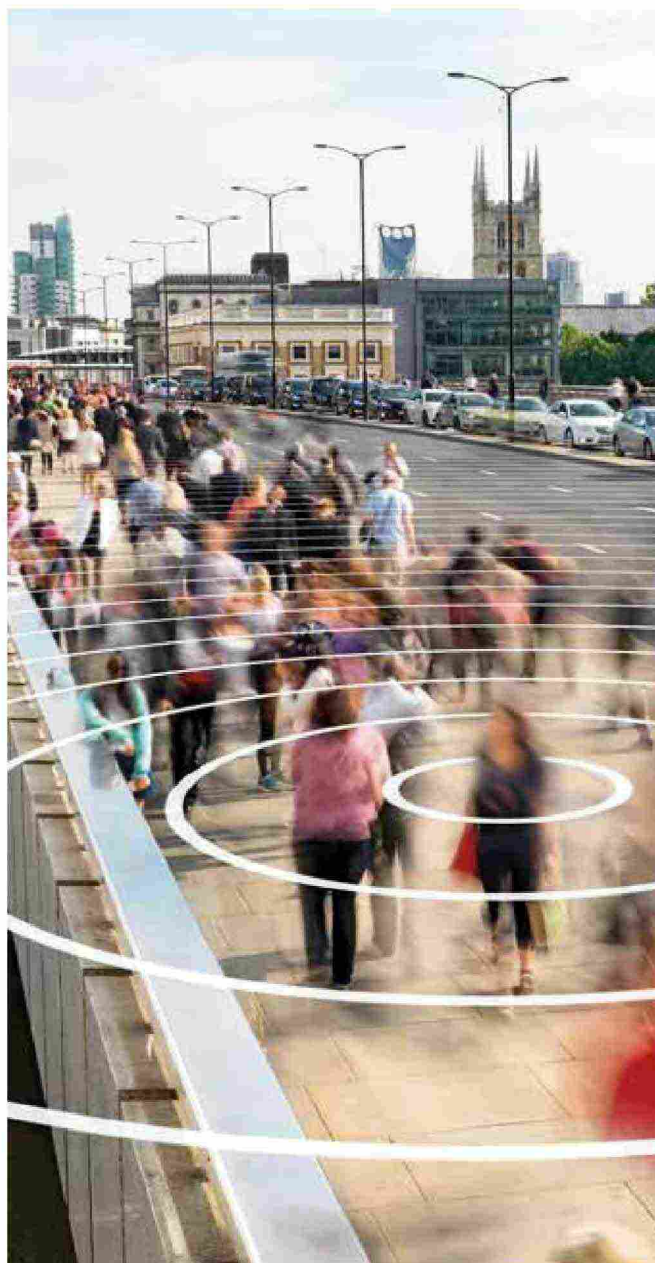
di **Leonardo Filippi**

Le informazioni che doniamo ai social network. Le abitudini di consumo che adottiamo nei grandi negozi online, come anche nei tradizionali supermercati in cui veniamo identificati con una tessera fedeltà.

Le prestazioni sanitarie erogate dall'ospedale, e quelle economiche di cui facciamo richiesta in banca. E molto, molto altro. Le tracce digitali che seminiamo ogni giorno, spesso senza nemmeno accorgercene, sono ormai le più svariate. Una mole di dati gigantesca - raccolta, stoccata ed elaborata - che offre, per la prima volta nella storia dell'umanità, la possibilità di osservare e analizzare con relativa semplicità il comportamento di una quantità enorme di persone.

Sono i cosiddetti big data. Secondo alcuni, si tratta solo del nuovo strumento di dominio delle grandi corporation del web. Il loro bene più prezioso. Informazioni, estratte grazie al "lavoro" di condivisione prestato gratuitamente sui social, utili alle aziende della Silicon valley per profilare milioni di utenti e incassare profitti col marketing mirato.

Vero. Ma è solo un lato, quello oscuro, della medaglia. «La scienza dei dati punta ad estrarre senso, a mettere a valore, le informazioni che cediamo quotidianamente. Ed è indubbio che oggi sia usata in gran parte come strumento per veicolare pubblicità» dice a *Left* è Dino Pedreschi. Ordinario di Informatica presso l'Università di Pisa Pedreschi è uno dei luminari italiani di questo nuovo ambito di ricerca. «Utilizzandoli per scopi di marketing - prosegue l'esperto che il 2 settembre terrà una lectio magistralis al [Festival della mente](#) a Sarzana -, i big data provocano indirettamente un altro problema: finiscono cioè per polarizzare sempre di più la nostra società». Sui social network, infatti, le nostre preferenze politiche, in nostri gusti musicali, le nostre simpatie vengono utilizzati per fornirci ulteriori informazioni in linea



con le nostre aspettative. Si chiama *filter bubble*, o problema dell'omofilia. «Uno studio recente conferma che si tratta di un meccanismo devastante, che crea "bolle mediatiche" e polarizza i consensi. Anche in casi in cui, in assenza di questo meccanismo - che sfrutta il nostro *bias confirmatorio*, la tendenza che ci porta a cercare conferme delle nostre convinzioni - tali divergenze non si creerebbero». Ma gli esempi di "cattivo uso" dei big data non finiscono qui.

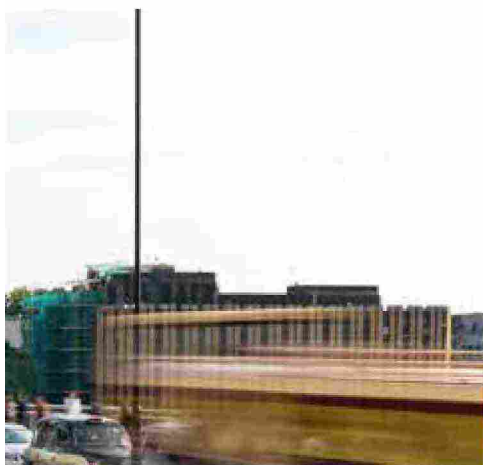
Basti pensare al celebre scandalo di Cambridge analytica, azienda che inferiva profili di personalità dai dati degli utenti Facebook, usandoli per tentare di influenzare l'esito delle presidenziali Usa. «La minaccia alla democrazia è veramente grossa, se si permette

di usare i dati in un modo illegittimo come questo - chiarisce Pedreschi - considerando anche che nelle elezioni moderne non è necessario spostare milioni di voti per far vincere o perdere una fazione». Ma a volte i guai possono accadere anche senza - apparentemente - violare alcuna legge. È il caso del megastore globale di Amazon. «Penso allo scandalo del *free same day delivery*, un'offerta di consegna dei prodotti nello stesso giorno dell'acquisto, non disponibile per alcuni

quartieri di città americane. Zone, guarda caso, con grande preponderanza di abitanti di colore. Ma i consumatori se ne accorgono, e Bloomberg fa esplodere il caso. La risposta del colosso di Bezos? È colpa dell'algoritmo». L'intervento umano nello studio dei dati, infatti, può essere ormai ridotto al minimo indispensabile. Grazie

Le regole Ue sulla privacy: diritto ad una spiegazione quando decide un algoritmo

all'intelligenza artificiale, a trattare i big data ci pensano supercalcolatori programmati per prendere decisioni, correggere errori statistici e addirittura elaborare veri e propri modelli decisionali. Si chiama *machine learning*, apprendimento automatico. «Stiamo passando da un'epoca in cui l'umano era responsabile di ogni singolo passaggio nella programmazione delle macchine - racconta Pedreschi - ad una situazione in cui l'algoritmo viene elaborato autonomamente, grazie allo studio di un numero enorme di esempi, in genere forniti da moltitudini di esperti umani. E quindi non sappiamo neanche bene come funziona, in che modo è stato in grado di generalizzare e fare scelte apparentemente in linea con quelli degli umani, certo, ma portandone con sé anche tutti i pregiudizi». Finendo magari con il negare un servizio in una parte di città - negli Usa si parla di *redlining*, tracciare linee rosse, ed è vietato per scopi creditizi - oppure ad uno specifico individuo. «Questi sistemi vengono usati, anche in Italia, per finalità assicurative o creditizie. Esistono sistemi di supporto alle deci-



IN COPERTINA

sioni che, quando vai a chiedere un mutuo, sulla base dei tuoi dati pregressi sono in grado di prevedere la tua affidabilità. Chiaro, in linea di massima può essere un sistema equo, ma è fondamentale che le logiche siano trasparenti», spiega il professore. L'importanza di sapere cosa c'è dentro il cofano degli algoritmi la si comprende bene se si pensa all'uso dei big data e del machine learning per scopi di polizia predittiva. Come nel caso del sistema di ausilio ai giudici Compas, usato negli States per esprimersi sulla scarcerazione di detenuti in base alla probabilità che commettano nuovamente reati: tacciato di razzismo in quanto stimava la possibilità di recidiva in media del doppio per i carcerati neri. Come denunciato da ProPublica. Il quotidiano britannico *Guardian*, a tal proposito, ha parlato di «nascita dei robot razzisti». Fortunatamente, l'Unione europea ha deciso di intervenire, col suo nuovo regolamento sulla privacy Gdpr. «La norma parla del diritto ad una *meaningful explanation*, il diritto a ricevere una spiegazione quando una decisione rilevante, non solo in senso legale, è stata presa da un sistema di intelligenza artificiale. Inoltre, per quel tipo di decisioni, è proibito l'uso di intelligenza artificiale senza intermediazione umana». Un bel passo in avanti, che fissa alcuni paletti. Ma, a prescindere dalle leggi, secondo Pedreschi i big data si possono prestare ad usi radicalmente opposti. Non per atomizzare, ma per rinsaldare i legami di una società. Non per speculare, ma per aumentare la conoscenza delle persone su se stesse e sull'essere umano. Fino ad intervenire sulle modalità positive di relazionarci con gli altri. «I big data possono essere valorizzati per la collettività. Mi spiego. Nei grandi sistemi complessi che abitiamo oggi, i problemi più grossi sono quelli "di allineamento". Pen-

siamo, all'economia, ai consumi, ma anche al traffico, eccetera. Quando riusciamo a tenere insieme le esigenze di tutti con quelle di ciascuno, siamo tutti più felici. E i big data possono aiutarci. Non solo. Se una persona diventa più consapevole delle abitudini di se stesso e degli altri, si può amplificare l'esperienza sociale, finché la nostra empatia». Alcuni esempi? «Noi - insiste il professore - all'interno del gruppo di ricerca internazionale Sobigdata, abbiamo curato alcuni studi. Grazie ai dati di posizionamento degli automobilisti, lavoriamo per monitorare e migliorare la mobilità delle persone. Ma questi dati possono servire anche, per esempio, a valutare scientificamente le tesi di un comitato che si oppone a un'infrastruttura ritenuta inutile. Un altro studio l'abbiamo compiuto sulla base dei dati di consumo dei soci Coop, più di due milioni di consumatori, relative agli ultimi 10 anni. Con lo scopo di valutare il tasso di integrazione dei "nuovi italiani". Risultato? Per la maggior parte dei migranti che si stabiliscono da noi, bastano solo 4-5 anni perché la distanza tra i loro carrelli e quella degli italiani si riduca sensibilmente. Un altro studio, infine, l'abbiamo condotto per capire la correlazione tra abitudini alimentari e malattie gastrointestinali, studiando cosa mettevano nel vassoio gli studenti della mensa dell'Università di Pisa. Conducendo rapidamente uno studio epidemiologico che, senza l'aiuto dei dati, avrebbe comportato un lungo periodo di analisi. Che spesso gli istituti pubblici non si possono permettere». Insomma, i big data usati per incentivare relazioni umane positive. «Esatto - aggiunge Pedreschi - li possiamo usare per sostenere una "globalizzazione gentile". Che non si dimentica delle persone e non lascia fuori **nessuno**».