

# Il Sole 24 ORE nòva

*Il compromesso tra dati e libertà*

01/03/2015

***Tecnologie e privacy-by-design sono le due soluzioni per cercare un equilibrio e rispondere alle esigenze dell'utente. Che vuole conoscere ma anche salvaguardarsi***

Nel 2000 gli utenti della telefonia mobile erano il 12% della popolazione mondiale. A fine 2014 sono il 96%, ovvero 6,8 miliardi di persone. I cellulari in circolazione sono il 128% degli abitanti dei paesi sviluppati ed il 90% degli abitanti dei paesi in via di sviluppo. E gli usi che facciamo del cellulare si sono estesi a dismisura, insieme ai sensori incorporati, in grado di registrare posizione, accelerazione, acquisire immagini e video, interagire con altri dispositivi e ovviamente essere sempre connessi alla rete. L'effetto forse più sorprendente e dirompente dell'umanità connessa sono i dati che ci lasciamo dietro, come sette miliardi di pollicini digitali. Dati che permettono di osservare le attività umane su scala globale, e quindi di misurarle, quantificarle e, in definitiva, prevederle. Perché solo gli indovini e i consulenti cercano di prevedere il futuro senza dati, come dice Laszlo Barabasi. Se invece abbiamo dati dettagliati su un fenomeno, anche inaspettato o bizzarro – un cigno nero – possiamo prevederlo. Figuriamoci, quindi, se non si possono prevedere tanti aspetti del nostro comportamento quotidiano, come i nostri movimenti o i nostri acquisti, data la regolarità della nostra routine! E quando scopriamo che osservando i “like” che lasciamo su Facebook è possibile capire se siamo gay con una accuratezza oltre il 95%, o che dalle foto che pubblichiamo è possibile stimare il nostro reddito e che dai nostri spostamenti del venerdì è possibile capire se siamo musulmani, ecco che ci rendiamo conto di quanto la dimensione in cui siamo entrati è davvero un passaggio epocale. Niente sarà come prima.

## **L'altra faccia di big data**

Si sta materializzando un nuovo strumento potentissimo, in grado di farci vedere il mondo e la società con occhi che finora non abbiamo avuto. Che può farci prevedere epidemie, instabilità, crisi economiche. Che può aiutarci a prevedere le conseguenze delle nostre decisioni, a livello collettivo e individuale, e quindi a farci scegliere meglio, renderci più consapevoli, farci comprendere e magari gestire la complessità della società plurale e interconnessa che abitiamo. Migliorare il nostro benessere. Ma al tempo stesso, big data ci trasforma in topolini sotto la lente, micro-organismi sul vetrino del microscopio. E se dai dati è possibile risalire alla nostra identità e alla nostra sfera personale più intima, ecco che possiamo ritrovarci alla mercé degli apprendisti stregoni alla ricerca di presunti terroristi, o di occasioni di business privo di scrupoli. Non c'è dubbio che, fino ad ora, big data nel nostro immaginario sia soprattutto legato al venir meno della privacy, alla paura della sorveglianza e del controllo sociale, piuttosto che alle magnifiche sorti e progressive legate all'uso dei dati come bene comune. Eppure non possiamo permetterci di fare a meno di questa sorgente di conoscenza, ne abbiamo assolutamente per affrontare le sfide della povertà, dell'energia, della disoccupazione, della disuguaglianza, del cibo, dell'ambiente, della salute. Della democrazia. E al tempo stesso non possiamo rinunciare a gestire le nostre informazioni e le nostre comunicazioni in libertà, condividendo ciò che ci va di condividere con chi ci pare e piace – altrimenti va in crisi totale l'idea stessa di democrazia.

## **Uso dei dati e privacy**

Quindi urgono soluzioni che possano far coesistere capra e cavoli, conoscenza e libertà. Uso dei dati e privacy. Pia illusione? No. Il fatto che in questa fase iniziale della società misurabile esistano pochi "latifondisti" che concentrano i dati di moltitudini di persone in grandi raccolte secretate non significa che questo sia l'unico modello possibile, né il più efficiente – al contrario, è dimostrato da mille esperienze che l'accesso aperto a dati interessanti stimola creatività, nuove idee di business, nuovi lavori. Il punto è come farlo in sicurezza, in un contesto etico facilitato da una tecnologia responsabile e trasparente. C'è tanta discussione su questi temi in giro per il mondo, e qua in Europa stanno nascendo infrastrutture per la ricerca sui big data che mettono l'etica e la privacy al centro. Come? Le linee di attacco sono due: una tecnologia per i dati di oggi, la *privacy-by-design*, e un "*new deal*" per i dati di domani. *Privacy-by-design* significa tener conto delle salvaguardie per la protezione dei dati personali fin dall'inizio del progetto di nuovi servizi basati su questi dati, in modo da rendere basso il rischio legato alla privacy delle persone coinvolte. Quello che gli scettici del tipo "la privacy è morta" ignorano che qualità dei servizi e privacy possono convivere. Che la

stragrande quantità di servizi rivolti, per esempio, ai cittadini per ottimizzare i propri spostamenti o ai decisori per ottimizzare le reti di trasporto pubblico, possono essere messi a punto facendo leva su dati personali che escono dai sistemi protetti solo dopo essere stati trasformati in modo da rendere trascurabile il rischio che uno specifico utente venga riconosciuto e che le sue informazioni personali (ad esempio, dove va il venerdì sera) siano svelate. Questo approccio è utilizzabile, ad esempio, sui dati telefonici e sulle tracce Gps dei navigatori delle nostre auto, per sfruttare il giacimento di conoscenza nelle banche dati degli operatori telecom e dei servizi telematici.

### **Privacy-by-design**

Nei nostri laboratori di ricerca usiamo l'approccio *privacy-by-design* ormai da anni per sviluppare progetti e prototipi insieme a vari partner industriali, come Wind, OctoTelematics, Toyota. Col centro Toyota InfoTechnology di Tokyo, ad esempio, stiamo mettendo a punto un sistema per la valutazione empirica del rischio nell'outsourcing di dati di mobilità personale per lo sviluppo di servizi innovativi di terze parti. Ad oggi, però, *privacy-by-design* non è ancora pratica corrente nella big data analytics, e non solo in Italia. Presto sarà un punto specifico della nuova direttiva europea sulla privacy, speriamo che questo dia una spinta all'uso sicuro dei dati ed eviti la tragedia dei "data commons", una espressione di Jane Yakowitz della Brooklyn Law School per indicare lo spreco di una risorsa comune di grande valore per una concomitanza di interessi particolari, arroganza, ignoranza, paura e mancanza di regole e di fiducia. Questo per oggi, o al più per domani mattina. In prospettiva bisogna rompere lo schema della raccolta passiva dei dati e far partire la partecipazione, la consapevolezza del valore dei dati di ciascuno di noi. Rendere ciascuno padrone dei propri dati. Ogni anno ogni persona lascia dietro di sé circa 3 gigabytes di briciole digitali nei sistemi e nei servizi più disparati che usiamo per le nostre azioni quotidiane: muoverci, comunicare, pagare beni, bollette o cibo, cercare in rete, leggere, giocare, fare sport, messaggiare, scrivere, postare, twittare, fare operazioni in banca. Tre giga, senza contare foto e video, altrimenti il conto sale di parecchio. Una valanga di informazioni personali che in larga parte vanno disperse (come lacrime nella pioggia). Eppure solo ciascuno di noi potrebbe collegare tutte queste informazioni su di sé in uno spazio personale dei dati. Nessun Google o Facebook ha un simile potere oggi, e dovremmo guardarci molto attentamente dal darglielo in futuro.

## Alla ricerca di senso

Immaginiamo per un momento di disporre di questo meccanismo – un “*personal data store*” – che non si limiti a mettere insieme tutte queste nostre tracce, ma che ne estragga senso e ci proponga una immagine di noi stessi, l’immagine riflessa nello specchio digitale. Ci aiuti a capire i nostri pattern di comportamento, di alimentazione, di spesa, o almeno come questi sono desumibili dalle nostre tracce. E ci dia la possibilità di confrontarci con i pattern della collettività, con possibili alternative, per aumentare la consapevolezza di noi e la capacità di cambiare e migliorare. Questo è fattibile, si sta sperimentando nel vivo di comunità reali in esperienze in giro per il mondo, ci stiamo provando anche noi con qualche centinaio di volontari che stanno creando un ecosistema di *personal data store* focalizzati su mobilità, servizi telecom e acquisti al supermercato, in collaborazione con Telecom Italia e con Coop. Ci sono sfide etiche importanti davanti a noi, ma è chiaro che questa idea di rendere ciascuno padrone dei propri dati cambia il gioco. Se riusciamo a capire l’importanza dei dati personali nella vita quotidiana per semplificare, risparmiare, diversificare, ecco che allora si può pensare ad un ecosistema diverso, in cui i dati possono fluire senza bisogno di concentrarli in grandi depositi centralizzati. In cui ognuno, invece di cedere i propri dati firmando una oscura liberatoria, decide se rispondere o meno a domande da parte di altri soggetti, in base al proprio interesse a partecipare ed alla fiducia dell’interlocutore. Un ecosistema che dall’esterno sembra un grande database a cui si possono porre domande, ma che in realtà è una rete di persone con i loro *personal data store* che possono scegliere di collaborare per rispondere. È chiaro che mancano pezzi di tecnologia, sia informatica che etica, per rendere funzionante questo ecosistema, ma in tanti ci stiamo lavorando. Alex Pentland di Mit Media Lab, anche insieme al centro Skill di Trento, sta sviluppando architetture per il *personal data store* e studia modelli di business che lo rendano sostenibile e sicuro. Protocolli come Ubiquitous Commons stanno proponendosi come meccanismi sicuri di comunicazione *peer-to-peer* che possono supportare la crescita dal basso di reti di pari per la condivisione di conoscenza. Occorre sviluppare meccanismi decentralizzati di reputazione per persone e organizzazioni, perché reti di questo tipo si sostengono sulla fiducia e sulla capacità di auto-regolarsi. Occorre sviluppare meccanismi di remunerazione e di incentivazione. Occorre sviluppare visualizzazione e story-telling per rendere la conoscenza, individuale e collettiva, fruibile da parte di tutti: decisori, imprenditori, cittadini. Può sembrare un percorso tortuoso e improbabile, quello che porta all’armonia fra conoscenza individuale e conoscenza collettiva. Ma ci arriveremo. Perché il potenziale della conoscenza individuale, l’energia imprigionata in quei 3 giga l’anno di dati personali, prima o poi sarà percepita da molti, e l’onda sarà inarrestabile. Nell’attesa, una buona notizia: la Commissione Europea, nel programma

Horizon 2020, ha deciso di finanziare una nuova infrastruttura di ricerca dedicata alla big data analytics e al social mining. Una rete di centri europei che mettono a disposizione dati, strumenti, risorse e competenze di *data scientist* per condurre grandi e piccoli esperimenti di big data analytics da parte di ricercatori, innovatori, startupper, policy-makers, istituzioni pubbliche. L'infrastruttura europea si chiama SoBigData, ed ha una testa italiana: è coordinata da Fosca Giannotti del Cnr di Pisa e si poggia sul nostro laboratorio di Big Data Analytics & Social Mining in rete con analoghi centri di Sheffield, Hanover, Zurigo, Helsinki, Tallinn, Londra. Si propone come un luogo dove sperimentare e far crescere l'ecosistema dei dati e della conoscenza: un embrione di quel sistema nervoso digitale per la nostra società, aperto, etico e partecipato, di cui abbiamo bisogno come il pane.



**DINO PEDRESCHI**

Professore ordinario di Informatica all'Università di Pisa