Immuni non efficace? Meglio dati aperti, geo-referenziati, anonimi e piattaforme “data driven”

Pandemia e qualità dei dati

La gestione della pandemia, inclusa la complessa campagna di vaccinazione di massa, richiede “cruscotti” informativi basati su dati coerenti, organizzati e continuamente aggiornati per agevolare la comprensione del quadro complessivo e informarne la popolazione.

Come più volte rilevato nelle varie fasi dell’emergenza pandemica Covid-19 (*contact tracing*, colori dei territori per aree di rischio, vaccinazioni, …), si assiste spesso a incertezze, situazioni di difficoltà e lentezze nel fronteggiare le emergenze (ad es. scarsità di vaccini), talvolta favoritismi (ad es. vaccinazioni di categorie a minor rischio rispetto ad altre), difficoltà di controllo dei dati territoriali, ecc. sostanzialmente a causa della **mancanza di dati aperti ed affidabili** su cui poter ragionare e su cui le Autorità competenti possano prendere decisioni consapevoli.

I dati mancanti, strutturati o non strutturati, sono i dati mai raccolti perché difficili da raccogliere o mai ritenuti utili in passato; i dati chiusi sono esistenti ma non disponibili; i dati aperti hanno un formato che li può rendere accessibili a tutti, umani e macchine, ma di cui a volte non si conosce la qualità, la semantica e la ridondanza.

Tali distinzioni vanno realisticamente oltre la differenza tra dati amministrativi o non, dati pubblici o privati, e così via. E' una distinzione funzionale che mira al raggiungimento della **qualità dei dati,** che non riguarda solo la riservatezza e protezione dei dati personali (*privacy*), ma anche la loro accuratezza, attualità, completezza, coerenza, credibilità, comprensibilità, … per citare solo alcune delle quindici caratteristiche dell'ISO / IEC 25012 sul modello di qualità dei dati, sintetizzabile nel sentire comune con: **occorrono dati certi.**

L'attenzione alla qualità sposta la prospettiva dal "controllo" *su* tutti al "contributo" *di* tutti, dal **timore** – più o meno consapevole ed esplicito – di essere oggetto di analisi e possibili profilature e "schedature", alla **fiducia** in un sistema in crescita, sviluppo e maggiore occupazione. Fiducia nelle Istituzioni come elemento centrale per la condivisione e il rispetto da parte della popolazione delle misure adottate dalle Autorità, incluse le limitazioni alle libertà personali per cause di forza maggiore, accettazione delle indicazioni sui vaccini, ecc.

Dati di qualità migliorano i rapporti tra Governanti e Governati, supportano le decisioni con aderenza alla realtà dei fatti e delle esigenze, e consentono a ricercatori, Enti di studio, gente comune, … di adeguare i propri comportamenti in funzione di auto-convincimenti, più che imposizioni dall'alto, e di aumentare le possibilità di analisi di correlazioni spesso nascoste (V. ad es. recente tentativo di correlazione tra numero di vaccini e riduzione della morbosità).

Piattaforme tecnologiche "data driven"

In precedenti articoli su questa testata (ad es. "Dati Bene Comune. Covid-19 e aree territoriali: i dati contano, se rac-

contano bene") era stata intrapresa una strada di avvicinamento a questi concetti, spinti dalle esigenze informative della pandemia Covid-19 ma traguardando una possibile migliore organizzazione di una auspicabile Sanità del Futuro, sempre più basata su processi e piattaforme ICT (*Information & Communication Technology*) "data driven", per meglio organizzare e complementare le procedure di assistenza medica (**personalizzata** e quanto più **tele-monitorata**) alla popolazione.

Trattata con sufficienza, sinora, nei processi sanitari, la *tecnologia* basata sui *dati* dovrebbe essere inserita con piena dignità nei sistemi sanitari, anche come prezioso supporto alle "3T" (tracciamento, test e terapia), cui andrebbe aggiunta una quarta T (*trust*), la *fiducia* di cui si diceva sopra.

Per passare dai dati mancanti a quelli aperti, riteniamo condivisibile una cultura quantitativa per la raccolta e l'utilizzo dei dati. Esiste peraltro un possibile e semplice approccio, poco costoso se programmato a priori con piattaforme adatte, che aiuti a questo scopo, sulla base dei dati rilevati ogni giorno da vari attori nell'ambito dei propri interessi e competenze istituzionali (Amministrazioni pubbliche e private, Enti ed Istituzioni, gli stessi cittadini proprietari dei dati, …), ma che vengono spesso ignorati, dispersi e non valorizzati dagli stessi attori.

Un tale approccio consisterebbe nel pianificare la continuità e coerenza tra dati nazionali, regionali, provinciali e comunali, definendo un flusso di **dati aperti tra Amministrazioni**, non solo tra Amministrazioni e cittadini. Ogni dato nazionale dovrebbe essere derivato da dati regionali, ogni dato regionale derivato da dati provinciali, ogni dato provinciale derivato da dati comunali. Per non parlare dei dati delle città metropolitane, che dovrebbero derivare da sotto-insiemi ancora più specifici come i quartieri delle città o municipi o quant'altro, con

“granularità” definibile secondo le necessità.

Come suggerito nel citato articolo, si ribadisce pertanto l’estrema rilevanza di una visione “telescopica” dei dati, una visione che – come il telescopio – fa apparire vicine cose lontane, e più chiare cose apparentemente complesse e confuse. Il territorio diviene così una miniera di informazioni comparate, un campo di suggerimenti di miglioramento oculato e foriero di soluzioni delle cause dei problemi esistenti. Gli stessi divari Nord-Sud o Est-Ovest, problemi globali, ecc. possono beneficiare di nuovi elementi di sviluppo resi possibili dall’analisi dei dati, lasciando alla Politica il ruolo di interpretazione, priorità, armonizzazione e decisione con spirito democratico su bene comune e impostazione etica.

Rafforzare il territorio è del resto in linea con il Titolo V della Costituzione, con l’approccio di allargamento delle *Smart City* alle *Smart Land* (fino alla *Smart Nation*), con le politiche dei trasporti, del turismo, della mobilità, del pendolarismo, della relativizzazione dei dati alla consistenza della popolazione, …

> ***Con gli stessi obiettivi, il concetto di “contact tracing” dovrebbe progredire verso un approccio di “(geo)location tracing” degli eventi, per la gestione di vicende non solo epidemico-sanitarie ma anche e soprattutto economico-sociali e ambientali, generando valori rappresentativi sia assoluti che percentuali nei vari territori.***

Piattaforme di analisi *big data* possono essere di grande aiuto per il contrasto alle epidemie, e riuscire a supportare il “contact tracing” con strumenti automatici rappresenta un elemento (*building block*) di notevole rilevanza. Oggettivamente il termine “tracing” non piace, ci fa sentire in qualche modo spiati e vulnerabili – ed infatti nelle documentazioni ufficiali si parla di “exposure notification”. Allo stesso tempo, sarebbe importante poter sapere se, quando

e *dove* sono avvenuti contatti tra soggetti infetti (o a rischio) e soggetti sani (almeno sulla carta), o incontri tra persone sane (asintomatiche), anche per poter gestire con consapevolezza e dovizia di informazioni sia le politiche di gestione pandemica che le regole di ritorno alla normalità.

> ***La gestione del luogo è essenziale per la comprensione delle cause di contaminazione, e ciò è possibile anche nel rispetto della privacy individuale. In altre parole, allo stato attuale sarebbe estremamente utile conoscere statisticamente i luoghi e le circostanze dei contagi attraverso gli spostamenti nel tempo delle persone – che possono e devono rimanere anonime e i loro dati riservati.***

Contact tracing e geolocalizzazione

Al riguardo, mentre una soluzione ICT / app di tracciamento dei contatti non avrebbe molto senso nel pieno dell'epidemia, con decine di migliaia di nuovi contagi al giorno, essa (ri)diventerà cruciale con il progressivo incremento delle vaccinazioni. Esse (auspicabilmente) ridurranno di molto i casi positivi, che dovrebbero a questo punto essere **tracciati con precisione geo-localizzata** per identificare e delimitare focolai e ridurre sempre più le occasioni di contagio.

E' quindi fondamentale sapere non solo **se** ci siano stati contagi, ma anche **dove** (e quando) essi siano avvenuti.

La quasi totalità delle app di *contact tracing* sono progettate per permettere ad un soggetto che si scopre positivo di segnalare (anonimamente) il proprio *status* ad un Centro di Controllo Sanitario, il quale potrebbe così comunicarlo a tutti quei soggetti con cui il positivo è entrato in contatto (tipicamente per almeno 10 minuti ed entro 2 metri circa, ove si utilizzi la tecnologia *Bluetooth* degli *smartphone*) nei 14

giorni precedenti alla scoperta dell'infezione.

Entrambe le tecnologie GPS e *Bluetooth* presenti nei nostri *smartphone* potrebbero essere utilizzate per tracciare i contatti a rischio e avvisare i potenziali contagiati. Il GPS identifica i percorsi degli individui che si sono incrociati, mentre il *Bluetooth* rileva connessioni tra *device* a distanza ravvicinata (fino a pochi metri), **senza sapere dove**. Per tali motivi, la tecnologia *Bluetooth* è stata considerata la meno invasiva per la *privacy*, permettendo agli *smartphone* di accorgersi se e quando siano vicini, ma senza informazioni geo-referenziate che possano far risalire in qualche modo all'identità degli utenti. Quindi, niente GPS e nessuna geo-localizzazione su dove siano avvenuti i vari incontri.

> ***D'altro canto, le mutate necessità di tracciamento "mirato" a seguito della disponibilità del vaccino anti-Covid suggerirebbero di provare a riesaminare l'intera vicenda, approfondendo i pro e contro dell'utilizzo della geo-localizzazione dei contagi. A detta di molti esperti, utilizzando un tracking GPS opportunamente anonimizzato ad integrazione del Bluetooth – e.g. tramite un protocollo PPSI (Privacy Preserving Set Intersection) -, i problemi di tutela della privacy sarebbero analoghi, sarebbe possibile tracciare geograficamente contagi e focolai, e basterebbe il 20% di utilizzo della app da parte della popolazione rispetto ad oltre il 60% con il solo Bluetooth (Low Energy, BLE) per una ipotizzata efficacia dei risultati.***

Il rischio – per quanto detto, probabilmente sovrastimato – di possibili violazioni della *privacy* individuale ha sinora escluso qualsiasi iniziativa in tale direzione. Rimane aperta la possibilità di gestire una qualche **geo-localizzazione dei dati anonimi ed aggregati** – al momento non presente nelle app di *contact tracing* della maggioranza dei Paesi nel mondo -, in modo da poter verificare i pericoli di contagio nelle varie situazioni operative – trasporti (V. rilevazioni di virus

pandemico sui bus / metro delle città), scuole, ristoranti, sport, ecc.

Dati “telescopici” e cruscotti informativi

Nei convegni di settore, come quello *online* sui #datiBeneComune tenutosi l’8 aprile scorso, è emersa una vibrata richiesta di coinvolgimento della Società Civile, un diritto ad essere informati, una esigenza di essere trattati da adulti da parte di un popolo di età media giovanile più di quella che si è abituati a vedere in genere negli eventi istituzionali o convegni scientifici di settore. Un popolo di giovani competenti e preparati che ha interiorizzato i metodi statistici come soluzione per uscire dalla complessità dei sistemi moderni, dalle insidie delle mancanze di supervisione, non solo a livello medico regionale e mondiale, ma di tutti i settori che la digitalizzazione, più che risolvere nel suo complesso, ha per ora solo fatto emergere, generando una forte richiesta di approfondimento e partecipazione.

Non tutto è gratuito, non tutti i dati possono essere disponibili, ci sono criteri e pianificazioni da definire. Non ci si deve meravigliare se alcuni dati sono regolamentati dalla *privacy* o se altri dati pubblici hanno un costo di vendita – i dati sono, ormai lo sappiamo, una miniera a cielo aperto. Ma ciò che occorre migliorare è la consapevolezza che **senza dati non si va più da nessuna parte,** e che la “telescopicità” verso il basso va portata anche ai livelli più alti, come a quelli europei e mondiali, viste le inevitabili implicazioni della globalizzazione.

I dati aperti non sono solo un problema “tecnico” di formato, di un “cloud” avvolgente, di un “digital twin” automatico o di una strategia democratica o meno. I dati aperti vogliono scuotere la scienza e la coscienza, andare oltre gli elementi

fondamentali della natura analizzati da Talete e dalla cultura ellenica di millenni fa dell'acqua, terra, aria, fuoco, per completare la conoscenza dei prodotti della Persona e delle Organizzazioni. La seconda era informatica che si sta vivendo, dopo il primo periodo dagli anni '60 ad oggi, può impostare un nuovo approccio di analisi basato sulla ri-organizzazione dei dati, in cui, prima ancora che l'Intelligenza Artificiale possa dare il suo contributo, una "Intelligenza Naturale aperta" possa stabilirne le direzioni fondamentali.

L'attenzione ai dati non è una moda del momento. E' **parte integrante dello sviluppo digitale ed economico**, una delle vie necessarie per la qualità che la Società moderna esige.

Conclusioni – sanità del futuro "data driven"

La gestione della pandemia Covid-19, inclusa la presente complessa campagna di vaccinazione di massa, richiede "cruscotti"informativi (*dashboard*) immediatamente **comprensibili e rappresentativi** della totalità delle informazioni necessarie. **Cruscotti** basati su dati coerenti e continuamente aggiornati per poter prendere decisioni consapevoli e informare opportunamente la popolazione. **Dati** aperti, affidabili e utilizzabili per il contrasto alla pandemia (morbosità, vaccinazioni, ecc.), **geo-referenziati** e organizzati "**telescopicamente**" dal livello comunale a quelli provinciale, regionale e nazionale, possibilmente europeo, con granularità a piacere.

A seguire l'auspicabile, rapida fuoriuscita dalla crisi pandemica, processi organizzativi e strumenti tecnologici basati su piattaforme "data driven" sono fondamentali per un Piano Nazionale di *e-health* che realizzi una efficace Trasformazione Digitale del nostro Servizio Sanitario Nazionale.

Una Sanità “personalizzata” sui bisogni dei singoli individui e delle collettività, per meglio organizzare le procedure di assistenza medica alla popolazione, con enfasi su tele-monitoraggio / -medicina, medicina territoriale di base come 1° livello di assistenza e *screening* della popolazione, accessi agli Ospedali e visite di persona solo nei casi necessari, ecc.

“Occorre *conoscere* per poter *deliberare*” (Luigi Einaudi)

Fulvio Ananasso, Presidente di Stati Generali Innovazione, Socio onorario UNINFO e Consigliere CDTI

Domenico Natale, Presidente Commissione UNI/CT 504 “Ingegneria del software”, Socio onorario UNINFO, Consigliere CDTI e Socio SGI

https://www.key4biz.it/immuni-non-efficace-meglio-dati-aperti-geo-referenziati-anonimi-e-piattaforme-data-driven-per-la-sa/355920/