

ARTICOLO DI PUNTOSICURO

Anno 23 - numero 5020 di Martedì 05 ottobre 2021

È possibile utilizzare i big data per migliorare la salute e sicurezza?

Una tesi di laurea si sofferma sull'uso dei big data nella sorveglianza in materia di salute e sicurezza. La sfida dei big data, i limiti attuali della sorveglianza, gli studi pilota, l'epidemiologia tradizionale e l'epidemiologia digitale.

Urbino, 5 Ott ? L'**evoluzione della tecnologia**, come ricordato in molti [articoli](#) e [interviste](#), anche precedenti all'emergenza [COVID-19](#), può avere notevoli conseguenze in materia di prevenzione e tutela della salute nei luoghi di lavoro, anche con riferimento alla possibilità dell'utilizzo di **sistemi di sorveglianza e monitoraggio intelligenti**.

Ne abbiamo parlato con riferimento alla presentazione di una tesi di laurea - Diritto della sicurezza del lavoro, Dipartimento di Scienze Giuridiche, [Università degli Studi di Bologna](#) - dal titolo "**Business intelligence: data-driven per la sorveglianza delle malattie correlate al lavoro**".

La tesi, presentata da Chiara Ciolli, oltre a soffermarsi su alcuni processi e tecniche di Business intelligence (BI), presenta alcune applicazioni e studi di **sorveglianza intelligente**. E si sofferma anche sull'**utilizzo dei big data** in materia di salute e sicurezza sul lavoro (SSL).

Ne parliamo oggi con riferimento ai seguenti argomenti:

- [L'uso dei big data nella sorveglianza in materia di salute e sicurezza](#)
- [Gli studi sullo sfruttamento dei big data per le malattie professionali](#)
- [Epidemiologia tradizionale ed epidemiologia digitale](#)

Pubblicità

<#? QUI-PUBBLICITA-SCORM1-[EL0779] ?#>

L'uso dei big data nella sorveglianza in materia di salute e sicurezza

La tesi indica che "sebbene i big data abbiano travolto ogni settore, dall'economico al sociale al produttivo, la lettura scientifica attribuisce diverse accezioni al termine ma poche si riferiscono all'ambito della salute". In particolare la **Commissione europea nel 2016** ("*Study on Big Data in Public Health, Telemedicine and Healthcare*") si è espressa affermando che '*i big data in sanità si riferiscono a grandi set di dati raccolti abitualmente o automaticamente, che vengono acquisiti e archiviati elettronicamente. Il termine è riutilizzabile nel senso di dati multiuso e comprende la fusione e il collegamento di banche dati esistenti allo scopo di migliorare il sistema sanitario e le prestazioni che esso offre*'.

Tuttavia la tesi segnala che tale definizione, come le altre fornite dalla comunità scientifica, può risultare incompleta "poiché non si è tenuto conto delle **cinque 'V'** 'Velocità, Varietà, Volume, Veridicità e Variabilità', gli attributi che distinguono i big data".

In particolare "la prima V si riferisce al processo di produzione e analisi delle informazioni in tempo reale (o quasi), grazie ad algoritmi che lavorano senza l'intervento umano. Con Volume si intende, invece, 'l'alta' quantità di dati prodotti, impossibile da gestire attraverso le tecnologie tradizionali. La Varietà e Variabilità, considerabili come un'unica caratteristica, riguardano l'eterogeneità delle fonti e il diverso formato in cui i Big data si presentano, ovvero, possono essere strutturati, non strutturati o semi strutturati. Infine, i dati devono essere Veritieri, ossia in grado di fornire segnali reali".



In ogni caso per quanto riguarda il campo prevenzionistico della SSL "fino ad ora, la definizione di Big data e il loro utilizzo è stato quasi del tutto assente". E la sorveglianza della SSL "si è **limitata ad utilizzare, elaborare, aggregare semplici dati** - non **big data** - in formato elettronico solo di tipo strutturato, spesso con un unico obbiettivo, derivanti da un numero limitato di fonti e non in tempo reale". E infatti, le prime implementazioni di monitoraggio intelligente, presentate nella tesi, utilizzano solamente "dati inseriti dagli operatori in database al fine di renderli digitali. Aggregano dati di natura differente, amministrativa e medica, ma integrano pur sempre informazioni di solo due o tre database".

Infatti la sorveglianza in materia di salute e sicurezza "**non è ancora in grado di gestire i dati non strutturati**, come le immagini, le informazioni raccolte dai social network, quelle derivanti dalle applicazioni mobili o dalle tecnologie indossabili, e non è in grado di analizzarli in tempo reale e di prevedere il futuro".

Dunque ? continua la tesi - al giorno d'oggi i big data "rappresentano sia una **sfida imminente per coloro che si occupano delle strategie prevenzionistiche**, sia la base di un nuovo metodo, insieme alle tecnologie d' apprendimento automatico, che indirizza gli ispettori del lavoro verso aziende aventi un più alto rischio professionale".

Gli studi sullo sfruttamento dei big data per le malattie professionali

Benchè la sorveglianza della SSL non sia ancora in grado d'utilizzare i Big data, "è stato realizzato uno **studio** definito «pilota» che ha utilizzato i dati derivanti dal web, per analizzare l'interesse correlato alla **silicosi** a livello di comunità scientifica, di copertura mediatica e delle nuove sorgenti di dati".

Si ricorda, a questo proposito, che, nel secolo scorso, "uno dei fattori principali che ha causato la morte di molti lavoratori è stato l'esposizione prolungata a biossido di silicio nelle fonderie, industrie siderurgiche, miniere o durante i lavori di costruzione di dighe e gallerie. La silicosi è una pneumoconiosi, ovvero un'affezione dei polmoni provocata dall'inalazione di polvere". E "grazie all'efficienza di strategie prevenzionistiche e di normative in tale ambito, sia la morbilità che mortalità sono progressivamente diminuite nel tempo".

Tuttavia se negli ultimi anni sono emersi "nuovi cluster di silicosi", grazie alle "nuove tecnologie d'informazioni e di comunicazione", è stato possibile operare un'analisi.

Il metodo adottato ? continua la tesi ? "si basa sullo studio dei dati generati da Google trends, il quale si occupa di attività web al fine di tracciare le impronte digitali in Internet. I dati provenienti da questa nuova sorgente sono stati confrontati con le pubblicazioni online scientifiche (PumMed e Google Scholar), con la copertura mediatica (Google news) con il traffico di Wikipedia (Wikitrends) e con l'utilizzo dei media (Youtube e Twitter) sul tema della silicosi. Queste fonti di dati raccolgono informazioni non strutturate, generate dall'uomo in modo spontaneo, non programmato e non intenzionale".

E i **risultati della ricerca** hanno confermato che "dal 2010 al 2011 è stato registrato un aumento delle ricerche del web sul tema della silicosi. Sia le pubblicazioni della comunità scientifica sia le news da parte dei giornalisti sono aumentate in modo considerevole, e allo stesso tempo anche l'interesse e il coinvolgimento della popolazione, individuato tramite una crescita del numero di mi piace, commenti, hashtag e ri-tweet".

Tuttavia è anche emerso che, "tra tutti i documenti caricati e i commenti postati, pochi contenevano valide informazioni scientifiche. Questo caso pilota dimostra com'è possibile, nella società dei dati, impiegare i big data per esaminare le correlazioni che esistono tra scienza, governi, copertura mediatica e popolazione".

Epidemiologia tradizionale ed epidemiologia digitale

La tesi ricorda, alla fine del secondo capitolo dedicato alla sorveglianza intelligente della salute e sicurezza sul lavoro, che tradizionalmente l'**epidemiologia** "ha informato, a livello scientifico, l'ambito della salute pubblica e anche quello della prevenzione delle malattie di natura professionale, attraverso la pubblicazione di considerevoli scoperte scientifiche".

Tuttavia oggi il "travolgente fenomeno della digitalizzazione dei dati, l'impiego sempre più diffuso dei dispositivi smart unito alle nuove potenze di calcolo e dei nuovi metodi di analisi dei dati", influenzano il mondo dell'epidemiologia e ci permettono di parlare di **epidemiologia digitale**.

L'epidemiologia digitale ha il medesimo obiettivo dell'epidemiologia tradizionale ma "cambiano le metodologie impiegate, al passo con le recenti tecnologie, al fine di utilizzare nuove fonti - i big data - le tracce digitali che lasciamo quotidianamente e inconsapevolmente quando interagiamo con internet".

Si segnala che gli studiosi "si sono già serviti d' informazioni digitali in modo da fornire un supporto alla sorveglianza della salute per il monitoraggio delle malattie infettive, il comportamento e le percezioni della popolazione" (il documento riporta alcune esperienze e studi relativi al terremoto ad Haiti e alla diffusione della malaria in Kenya).

In definitiva il prossimo scoglio che dovrà affrontare l'epidemiologia occupazionale "sarà quello di **ampliare le sorgenti dei dati** - derivanti dall'uso di internet - integrandoli con quelle tradizionali". In questo modo la sorveglianza della SSL, basata esclusivamente sui dati che vengono registrati da personale medico, "muterà in una **sorveglianza intelligente** fondata sull'integrazione dei tradizionali dati sanitari con quelli generati dal web".

Rimandiamo, in conclusione, alla lettura integrale della tesi che riporta anche precise indicazioni sulle fonti e sulla bibliografia utilizzata per la sua elaborazione e segnaliamo, riguardo ai problemi correlati all'uso dei big data, anche i contributi di Adalberto Biasiotti sul tema.

RTM

Scarica il documento da cui è tratto l'articolo:

"Business intelligence: data-driven per la sorveglianza delle malattie correlate al lavoro". tesi di Chiara Ciolli - Tesi di Laurea in Diritto della sicurezza del lavoro, Dipartimento di Scienze Giuridiche, Università degli Studi di Bologna ? Relatore Prof.ssa Patrizia Tullini, Correlatore Prof.ssa Raffaella Brighi, anno accademico 2019/2020.



Questo articolo è pubblicato sotto una Licenza Creative Commons.

www.puntosicuro.it