

Generazioni connesse3 - 14/11/2018
Miur - Sala della comunicazione

Algoritmi, big data e costruzione identitaria

Teresa Numerico

Dip. Filosofia, comunicazione e spettacolo

Università di Roma III

teresa.numerico@uniroma3.it

La stampa internazionale sul gatekeeping dei contenuti

- Mike Allen, Axios, 9 nov. 2017 Sean Parker unloads on Facebook exploiting human psychology
<https://www.axios.com/sean-parker-unloads-on-facebook-2508036343.html>
- Cathy O'Neal, Bloomberg, 7 nov. 2017, maybe facebook is broken
<https://www.bloomberg.com/view/articles/2017-11-07/maybe-social-media-is-broken>
- Roger McNamee, Guardian, 11 nov. 2017 How Facebook and Google threat public health - and democracy
<https://www.theguardian.com/commentisfree/2017/nov/11/facebook-google-public-health-democracy>

Adolescenti e relazioni amoroze: tecnologia come controllo?

- Durante o dopo una relazione il 31% (26% durante) sperimenta il controllo online del partner
- Intorno al 10% degli adolescenti che hanno avuto una relazione hanno avuto accesso a un account del partner, hanno danneggiato il profilo del/della ex online
- Il 21% riporta che un ex o un partner attuale ha letto i loro messaggi senza permesso

Teens, Technology and Romantic Relationships Pew research center, 1 oct. 2015 <http://www.pewinternet.org/2015/10/01/teens-technology-and-romantic-relationships/>

Desiderio dello sguardo di approvazione

- Un'erosione dell'anonimato è un prodotto dei servizi di social media pervasivi [...] e forse un cambio del modo in cui le erosioni vedono quello che deve essere pubblico o privato
- Il vecchio stratagemma del *panopticon* è stato implementato[...] in maniera universale.
- il vecchio incubo del *panopticon* (non sono mai da solo) è stato trasformato nella speranza di "non essere più da soli" (abbandonati, ignorati, negletti)
- La paura di aprirsi è stata soffocata dalla gioia di essere notati

Def. NSF & NIH

- L'espressione "big data" in esame fa riferimento ad ampi, diversi, complessi, longitudinali e/o distribuiti set di dati generati da strumenti, sensori, transazioni Internet, Email, video, numero di click e/o altre fonti digitali disponibili oggi e nel futuro

NSF-12-499 citato in Floridi 2017

Cosa sta succedendo?

- La digitalizzazione consente con grande facilità la cosiddetta datificazione, che non è esclusivamente il frutto del digitale, ma il digitale offre un acceleratore a questo processo
- La costruzione dei dati non riguarda solo fenomeni scientifici riguarda la nostra stessa vita
- Siamo noi che accettando di usare i telefoni cellulari e le loro smart app datifichiamo la nostra vita

Quantified self e selftracking

- Activity monitor (Apple Watch, Fitbit tracker Jawbone)
- Sleep specific monitors (Pillow - Smart Sleep Cycle Alarm Clock Drowzy, Lark, SleepBot, Sleep as Android)
- Reproductive health (Clue, Glow, kindara)
- Diet and weight (Fitbit aria scale, my fitnessPal, smalt)
- Mood (lume personal tracker, moodTracker,
- Altro (23andMe, WorkMeter ecc.)

SECONDO EVGENY MOROZOV QUESTI DISPOSITIVI SONO UN MODO DI INTROIETTARE LA SILICON VALLEY NELLE NOSTRE ABITUDINI

La purezza dei dati

- Come se questi dati discendessero direttamente dai fatti e avessero un rapporto diretto, incontaminato con la 'realtà'
- La politica dei big data segnala e suggerisce che non c'è niente da decidere e che tutto discende direttamente dal funzionamento dei meccanismi di costruzione delle correlazioni (algoritmi) di cui pochi conoscono la logica e nessuno sa perché le correlazioni trovate dovrebbero avere senso

Il potere dell'algoritmo

- È basato sull'uso di probabilità per anticipare le azioni degli individui o le loro preferenze a livello di profilazione che non è in controllo del singolo, ma che è gestito anche in riferimento al singolo
- Fronteggiamo una normatività che, pur sembrando democratica, in realtà crea con le sue cieche categorizzazioni un forte impatto sulla rappresentazione della realtà a proposito di schemi che riguardano la politica, la società il genere, l'origine etnica ecc. Ripristinando esattamente gli stereotipi e le istanze di esclusione eliminate da anni di lotte politiche

Che cos'è l'algoritmo

- Gli algoritmi non sono necessariamente software
- Sono procedure codificate per trasformare certi input formalizzati in entrata in output di uscita sempre formalizzati, basandosi su precise istruzioni di calcolo che devono essere finite e produrre risultati in tempo finito
- Gli algoritmi vengono molto prima dei computer, un esempio informale è l'algoritmo per il massimo comun divisore che si trova negli elementi di Euclide (circa 323-283 a.c.)
- Il termine algoritmo prende il nome dal matematico, geografo persiano Al-khwārizmī (780-850 d.c.)

Computer come algoritmi

- La prima definizione formale di algoritmo avviene con la Macchina astratta di Turing, l'antecedente del computer
- I computer fanno esclusivamente calcoli che sono eseguiti come procedure codificate di regole da applicare ai dati in input per produrre degli output
- Possiamo pensare quindi ai programmi di computer come a degli algoritmi
- Tuttavia un algoritmo può anche essere eseguito a mano (e prima dei computer spesso lo era)

Qual è il problema?

- Il problema è che gli algoritmi sono sempre più centrali nella presa di decisione in moltissime questioni anche vitali per le persone, come ottenere un lavoro, avere un mutuo, quantificare il premio dell'assicurazione, accedere all'università (per adesso in US) ma noi non abbiamo nessun dettaglio su come agiscano e quale logica seguano nella maggior parte dei casi
- È importante comprendere in che modo l'introduzione degli algoritmi nelle pratiche della conoscenza può avere delle conseguenze politiche, come ingiustizie, mancanza di equanimità nella presa di decisioni che riguardano l'essere cittadini

Vedi Gillespie 2014

Rilevanza politica degli algoritmi

- Modelli di inclusione: cosa c'è dietro la costruzione di un indice, cosa è escluso come i dati sono preparati per l'algoritmo
- Cicli di anticipazione: le implicazioni dei tentativi dei fornitori di algoritmi per comprendere e predire le abitudini degli utenti e come possono incidere
- La valutazione della rilevanza: in che modo gli algoritmi determinano cosa è rilevante, la loro oscurità e come agiscono sulle scelte politiche
- La promessa di obiettività algoritmica: e come reagisce di fronte alle controversie
- L'influenza sulle pratiche: cioè come gli utenti riorganizzano le proprie pratiche per essere efficaci per gli algoritmi da cui dipendono
- La produzione di pubblici calcolati: come la presentazione algoritmica del pubblico ha effetto sulla autopercezione del pubblico e chi è nella posizione di trarne beneficio (Gillespie 2014, 168)

I problemi epistemologici di alcuni metodi sui big data

- La costruzione dei dati relativi all'oggetto di studio
- Le tecniche algoritmiche usate per creare modelli interpretativi
- La tendenza a concentrarsi sull'archiviazione dei dati senza le dovute cautele sulle motivazione e l'esterno dell'archiviazione
- I meccanismi di categorizzazione che tendono a privilegiare la valutazione delle variabili datificabili e quantificabili a discapito delle altre caratteristiche
- La mancanza di una discussione approfondita sulle infrastrutture nascoste che impongono i loro criteri interpretativi

Frenologia e big data

- La frenologia era un modello che si basava su una stupidaggine pseudoscientifica allo scopo di formulare asserzioni autorevoli e per decenni nessuno l'ha verificata. I Big Data possono cadere nella stessa trappola.
- I modelli[...]possono sbarrare la strada a determinate persone, anche se la "scienza" sulla quale si basano non è niente di più che un'accozzaglia di ipotesi non verificate.

O'Neil 2017, 178-179

Macchine e umani rispetto ai dati

- [I dati] hanno urgente bisogno del contesto, del buon senso e dell'equità che solo l'uomo è in grado di fornire. Ma se lasciamo la questione in mano al mercato, che premia l'efficienza, la crescita e il cash flow (tollerando nel contempo un determinato livello di errore), qualsiasi intromissione da parte dell'uomo sarà scoraggiata e gli operatori riceveranno istruzioni di tenersi alla larga dalle macchine (O'Neil 2017,227)
- Ma il processo decisionale umano, benché spesso viziato, ha una grande virtù: di evolvere. [...]I sistemi automatici, invece, rimangono fermi nel tempo fino a quando gli ingegneri non decidono di modificarli (O'Neil 2017, 294)

Di chi è la responsabilità degli algoritmi?

- Chi si chiede come programmare gli algoritmi per garantire equità, conciliazione vita lavoro, non discriminazione di minoranze o delle differenze (fragilità emotiva dei lavoratori)?
- Chi invece si chiede solo come massimizzare i profitti?
- Dietro le macchine e i sistemi informativi ci sono progetti industriali e obiettivi precisi

L'essere umano ha bisogno di [...] seguirle [le macchine], di proteggerle, e di mettere a punto gli algoritmi come nostri aiutanti digitali, poiché gli algoritmi, nonostante l'apparenza, non sanno aiutare se stessi, ma seguono semplicemente una regola meccanica (Bunz 2014, 112)



La costruzione delle abitudini

- Le abitudini non sono virus che si propagano
- Per quanto difficile sia cambiare abitudini, le abitudini sono acquisite non 'incorporate' come gli atti involontari, come respirare
- È scorretto pensare le abitudini come dipendenze
- Gli studi sull'omofilia delle abitudini non sono fatti per sostenere la giustizia nell'analisi sociale

Chun 2016, pp. 8-15

Conoscenza e segreto

- Conoscenza è potere. Controllare gli altri mentre si evita di essere oggetto di controllo è la più importante forma di potere
- Le aziende cercano i nostri più intimi dettagli di vita come potenziali clienti e dipendenti, ma danno ai regolatori il minimo possibile di informazioni circa le loro statistiche e procedure.
- Le aziende internet raccolgono sempre più dati sui loro utenti ma lottano contro i regolatori che vorrebbero che gli utenti esercitassero un qualche controllo sui risultanti dossier digitali.

(Pasquale 2015 pp.3-4)

Di chi sono questi dati?

- Secondo Rodotà (2014) seguendo l'idea di *Habeas Data* i dati sono nostri.
- Nessuno dovrebbe poter vendere e comprare le nostre email per giunta avendo attaccato a ciascuna di esse una categorizzazione da marketing del tipo "cliente non affidabile", "elevate spese mediche", "Guida pericolosa", "reddito in declino" (Cardon 2016, 73)

La categorizzazione e i suoi demoni

- La categorizzazione è un sistema di potere per governare la realtà presuntamente costituita di fatti oggettivi (Bowker, Leigh Star 2000)
- Si stabilisce cosa è sano e cosa è malato, chi ha più merito e chi ne ha meno per trovare lavoro, o come valutare la prestazione, quanto vale un premio di assicurazione, chi ha diritto al prestito bancario

Big data e algoritmi di machine learning aggiungono a tutte queste variabili aleatorie dei precisi vincoli sulla datificazione e sulla misurabilità di ogni aspetto da valutare secondo i propri criteri. Quello che non è rappresentabile in forma quantitativa non è analizzabile dalle tecniche di big data, perché esse sono intessute dei programmi di computer che devono analizzare i dati (gli algoritmi di machine learning) Non si dà algoritmo senza base dati organizzata per lui

Aziende broker di dati

- Acxiom <http://www.acxiom.com/>
- Exelate/Nielsen <http://exelate.com/>
- TowerData formerly RapLeaf
<http://intelligence.towerdata.com/> e
il suo listino prezzi per le email
[http://intelligence.towerdata.com/
pricing-append](http://intelligence.towerdata.com/pricing-append)
- BlueKai comprata da Oracle nel 2014
[http://www.oracle.com/us/corporate/
acquisitions/bluekai/index.html](http://www.oracle.com/us/corporate/acquisitions/bluekai/index.html)

IL PROBLEMA DEI FALSI POSITIVI

- L'obiettivo dei big data non è solo il presente ma piuttosto il futuro e la capacità di prevederlo
- Nella prevenzione di frode e crimine o terrorismo per esempio non sarà ottenere falsi positivi, quanto essere in grado di cogliere tutti i veri positivi.
- Non ci saranno mai delle mancanze perché la logica del sistema è quella del depistaggio piuttosto che di una vera diagnostica

(Rouvroy, berns 2013, 7)

Che fare?

- Educare gli adolescenti alla consapevolezza che quello che si fa online è pubblico anche se sembra avvenire in privato
- Costruire una consapevolezza sui sistemi di categorizzazione e sul loro funzionamento ad excludendum
- Secondo Chun costruire una nuova appartenenza alla rete che consenta il diritto all'oblio non attraverso la protezione legale, ma per la consapevolezza della nostra fragilità
- Rivendicare il diritto a non considerare un'abitudine come una dipendenza
- Lavorare per una regolamentazione degli strumenti di profilazione, impedendo di usarli senza controlli e soprattutto impedendo la segretezza dei metodi quando le loro conseguenze hanno effetto sulla cittadinanza e sulla democrazia

Fonti citate

- Bowker G.C. Leigh Star S. (2000) *Sorting things out*, Mit Press, Cambridge, MA.
- Bunz M. (2014) "Algoritmi della conoscenza e trasformazione del lavoro", in Pasquinelli M. (2014) (a cura di), *Gli algoritmi del capitale*, Ombre Corte, Padova.
- Cardon D. (2016) *Cosa sognano gli algoritmi*, Mondadori, Milano.
- Chun W. H.K. (2016) *Updating to remain the same*, MIT Press, Cambridge, MA.
- Deleuze G. (1990/2015) "Postscritto sulla società del controllo", *Pourparler*, Quodlibet, Macerata, pp. 234-241.
- Floridi L. (2017) *La quarta rivoluzione*, Raffaello cortina, Milano.
- O'Neal C. (2016) *Weapons of math destruction*, Allen Lane, Penguin Books, St. Ives.
- Pasquale F. (2015) *The black box society*, Harvard Univ. Press, Cambridge, MA.
- Pew and Internet Research center (2016) Privacy and information sharing
<http://www.pewinternet.org/2016/01/14/privacy-and-information-sharing/>
- Rodotà S. (2014) *Il mondo nella rete*, Laterza, Roma.