

Indice

Introduzione	7
Lezione 1 Cosa sono i dati digitali e come se ne parla nel libro.....	17
Lezione 2 I modelli dei dati ci aiutano a rappresentare e comprendere il mondo	35
Lezione 3 L'utilizzo dei grafi semantici: il caso de <i>L'amica geniale</i>	51
Lezione 4 La qualità dei dati	73
Lezione 5 Il fenomeno delle fake news	99
Lezione 6 La visualizzazione dei dati.....	115
Lezione 7 Il valore dei dati	131
Lezione 8 Dati e machine learning.....	151
Lezione 9 L'etica dei dati digitali: l'accuratezza e l'equità.....	165
Lezione 10 Il Cosa e il Perché	191
Lezione 11 Il diritto europeo dei dati.....	207
Lezione 12 Previsione e prescrizione: la cattura del futuro da parte del passato.....	227
Autori.....	241

Introduzione

Giulio De Petra

1. Un nuovo universale antropologico

Da molti anni siamo immersi nel flusso crescente della trasformazione digitale in tutti i diversi ambiti della nostra vita. Dispositivi e applicazioni digitali hanno modificato, e stanno ulteriormente modificando, il nostro modo di lavorare, di relazionarci con gli altri, conoscere, curarci, produrre e commerciare, partecipare alla vita politica, fruire e produrre arte.

In ogni angolo del mondo, senza distinzioni di sviluppo o di ricchezza, in ogni momento della vita, dai momenti pubblici a quelli più intimi e privati, l'immagine di una donna o di un uomo, leggermente chini a sfiorare con un dito una piccola superficie luminosa raccolta nel palmo di una mano, concentrati e separati dal contesto che li circonda, è l'immagine più consueta che ci può capitare di osservare. Come ha osservato il filosofo Enrico Donaggio, questa immagine è come un nuovo “universale antropologico”.

Questa trasformazione ci è stata proposta e raccontata come benefica e ineluttabile.

Nel corso degli anni è stato costruito un senso comune che associa il diffondersi delle tecnologie digitali al progresso e al miglioramento della qualità della vita. A chi si assume il compito di promuovere questa diffusione è stato dato talvolta il nome di “digital evangelist”, come se si trattasse di una religione.

2. Critiche autorevoli e informate

È solo negli ultimi anni che qualche dubbio ha iniziato a diffondersi. Dapprima sono stati alcuni studiosi, come Evgeny Morozov che in “Inter-

net non salverà il mondo” (Mondadori, 2014) ha messo nel mirino della sua critica il “soluzionismo digitale”, cioè l’idea che per ogni problema esista una soluzione digitale. Più recentemente, dopo lo sviluppo esponenziale dell’utilizzo di piattaforme di comunicazione, (i cosiddetti “social”), ad essere oggetto di studi e pubblicazioni sempre più numerosi sono state le modalità di funzionamento delle grandi piattaforme di comunicazione, diventate rapidamente le imprese più ricche e potenti del mondo. Molti sono i testi, anche scritti con intento divulgativo, che raccontano il modo in cui le grandi imprese digitali estraggono valore dai dati che gli utenti consegnano loro inconsapevolmente con un utilizzo sempre più intenso e pervasivo. Tra quelli che hanno avuto più eco tra gli esperti, per la ricchezza dei casi analizzati e la completezza dell’analisi, *Il capitalismo della sorveglianza* di Shoshana Zuboff (Luiss, 2019). Mentre tra quelli scritti con esplicito intento divulgativo senza attenuare il rigore dell’analisi segnaliamo i molti libri scritti dal collettivo Ippolita, come ad esempio *Tecnologie del dominio* (Meltemi, 2017), che reca come sottotitolo “lessico minimo di autodifesa digitale”.

Anche l’università e la ricerca hanno iniziato negli ultimi anni ad analizzare e mettere in evidenza gli aspetti critici più evidenti della trasformazione digitale, non solo quelli relativi all’utilizzo dei sistemi digitali, ma anche quelli che caratterizzano la loro progettazione e produzione e, soprattutto, i nessi che legano le modalità di produzione e le caratteristiche architettoniche alle modalità di utilizzo. Tra i centri di ricerca più attivi in questa direzione e con reputazione internazionale si segnala il centro Nexa del Politecnico di Torino, che esplicitamente si occupa, con approccio multidisciplinare, di “Internet e società”.

3. La grande sproporzione

Malgrado libri, studi e ricerche sempre più numerosi, che ne mettono in evidenza aspetti critici sempre più significativi, nessuna perturbazione sembra scalfire la traiettoria della trasformazione digitale. Anzi, sempre più clamorosa appare la sproporzione tra la crescita della consapevolezza degli addetti ai lavori riguardo a rischi e criticità, e la possibilità effettiva

di porvi rimedio. Sproporzione che è determinata a sua volta dall'enorme differenza di potere tra le grandi aziende che determinano le caratteristiche della trasformazione digitale, e la debolezza di chi, istituzione scientifica o statale, si propone di porvi rimedio, o anche solo di diminuirne i rischi.

È in questo contesto che si è sviluppato lo sforzo di regolazione e di governo che caratterizza le più recenti iniziative dell'Unione Europea e di alcuni degli Stati che ne fanno parte. Il tentativo di individuare norme che possano orientare i comportamenti della grandi imprese digitali, impedendo alcuni sviluppi e favorendone altri, imponendo regole di comportamento o divieti, è stata la conseguenza "politica" della crescita di consapevolezza critica tra esperti e studiosi.

Va detto però che a stimolare l'attenzione delle istituzioni politiche europee non è stato soltanto il peso delle opinioni di studiosi e scienziati sociali e di qualche raro informatico. È stato soprattutto il fatto che le grandi imprese digitali, proprio in virtù del loro enorme potere di condizionamento delle opinioni e dei comportamenti sociali, hanno invaso anche il campo della politica, stravolgendone le procedure ed i comportamenti. Quando è emerso il modo in cui l'azienda Cambridge Analytica (con la complicità di Facebook) aveva condizionato le elezioni del 2018 negli Usa e in alcuni paesi europei, il sistema politico ha pensato a come difendersi, e ne è derivato un più forte impulso, in particolare in Europa, al tentativo di regolamentare la disponibilità e l'utilizzo dei dati personali.

4. Utilità e limiti della regolazione

La normativa europea, ed in particolare quella che si è costruita negli anni più recenti intorno alla protezione dei dati personali, mostra, con l'evidenza dell'esperienza quotidiana di ognuno di noi, la necessità ma anche la debolezza delle attività di regolazione. Non solo nella fase di controllo, ma proprio nella fase del processo di attuazione della normativa.

Come ricorda Giorgio Resta "Il meccanismo del consenso informato quale base normativa atta a legittimare il trattamento dei dati ha innescato un processo di burocratizzazione del consenso sia nei rapporti con i soggetti privati sia in quelli con i soggetti pubblici. Soprattutto nel contesto dei rapporti on-

line il consenso si è tradotto in quella famosa ‘foglia di fico’ a cui accennava già molti anni addietro Stefano Rodotà, atta a mascherare una realtà fortemente asimmetrica e in cui l’idea dell’autodeterminazione dell’interessato si è rivelata poco più che un’etichetta priva di riscontri operazionali”.

Il regolamento europeo per la protezione dei dati personali (GDPR) è il frutto più maturo del tentativo non solo di definire regole, ma anche di attuarle. L’attività dei garanti dei dati personali nei paesi europei, e il loro coordinamento a livello europeo, promosso e realizzato dal magistrato italiano Giovanni Buttarelli sono un importante punto di riferimento. La protezione dei Dati Personali non basta però da sola a realizzare un governo efficace, e altre e più ampie normative europee come il Data Governance Act stanno provando a definire un quadro di riferimento più completo.

Ma non è l’ampiezza della normativa il problema da superare e ogni tentativo di regolazione, pur necessario, rischia di essere inefficace e velleitario se non riesce a modificare diffusamente i comportamenti d’uso, se non riesce a far crescere nell’enorme platea degli utilizzatori un’adeguata consapevolezza dei meccanismi di funzionamento dei sistemi digitali.

5. La lezione dell’attualità

Quanto abbiamo finora scritto è stato dimostrato con grande evidenza da quanto è accaduto nei mesi più recenti, quando la possibilità di utilizzare liberamente sistemi di intelligenza artificiale (IA) capaci di generare linguaggio (ad esempio ChatGPT), ha prodotto un picco di attenzione su opportunità e rischi dei sistemi di IA in generale, e in particolare di quelli capaci di interagire mediante la comprensione e la produzione di linguaggio.

Imprese, ricercatori e stampa, hanno messo evidenza gli straordinari benefici che sarà possibile aspettarsi dall’utilizzo di sistemi sempre più potenti di intelligenza artificiale nei più diversi settori. Nello stesso tempo però alcuni degli stessi esperti direttamente responsabili delle tecnologie, in particolare quelle che consentono l’apprendimento automatico, hanno dichiarato pubblicamente le loro preoccupazioni, chiedendo moratorie nell’uso di questi sistemi o addirittura la messa al bando del loro utilizzo in

alcuni ambiti. Anche in questi caso il Parlamento Europeo ha votato, su proposta della Commissione, un impegnativo “AI act” che contiene divieti, prescrizioni e raccomandazioni, mentre i Garanti dei Dati Personali di alcuni paesi europei, a cominciare da quello italiano, hanno provato concretamente a negoziare, sulla base dei loro poteri, le modalità di funzionamento dei linguaggi generativi e di ChatGPT.

Suscitando l’interessato sarcasmo di chi, direttamente coinvolto nello sviluppo di questi sistemi, ha sostenuto che le leggi vigenti non possono essere applicate a prodotti basati sull’intelligenza artificiale, in virtù della loro novità e straordinarietà. Questi argomenti sono stati efficacemente definiti dalla filosofa Daniela Tafani “il mito dell’eccezionalismo tecnologico e del vuoto giuridico” (*Overlay journal of political philosophy*, Università di Pisa, maggio 2023).

Eppure, anche in questo caso, nonostante preoccupazioni autorevolmente espresse, proposte normative e concreti interventi sanzionatori, ChatGPT e altri prodotti della stessa natura sono diffusamente usati da ricercatori, giornalisti, studenti, avvocati o anche semplici curiosi sedotti dalla capacità affabulatoria dell’interlocutore digitale. Il che non è di per sé un male, ma rischia di diventarlo se l’utente non si pone il problema delle conseguenze della sua scelta, e dunque non diventa un soggetto sociale consapevole.

6. La Scuola Pop sui Dati Digitali

Cosa si può fare per intervenire positivamente nella complessità di questo contesto? Come è possibile contrastare le conseguenze negative della trasformazione digitale, senza perdere la possibilità di utilizzare la potenza delle tecnologie digitali per conseguire una maggiore giustizia sociale?

Il coinvolgimento largo e diffuso degli utilizzatori può essere una strada, forse la più promettente, da percorrere. A entrare in gioco devono essere infatti gli attori sociali coinvolti nella digitalizzazione e le loro forme organizzate. Ciò che serve, e ciò che attualmente manca, è un effettivo, ampio e consapevole coinvolgimento delle associazioni civiche di cittadinanza, delle organizzazioni che rappresentano il lavoro, di chi lavora con

e per le tecnologie digitali, di chi, anche non direttamente coinvolto, ne subisce le conseguenze, come ad esempio i cittadini utenti di servizi pubblici trasformati in servizi on-line o, in qualche caso, già erogati mediante l'uso di sistemi basati su intelligenza artificiale.

È come concreta conseguenza di queste convinzioni che si è sviluppata, nel corso del 2021, una originale esperienza formativa, la **“Scuola Pop sui Dati Digitali”**, da cui è nato questo libro.

Un'iniziativa di “formazione” come strumento per generare consapevolezza critica, responsabilizzare e rendere possibile un utilizzo attivo delle norme già disponibili, per consentire agli utilizzatori di partecipare alla progettazione e realizzazione dei sistemi digitali che li riguarderanno. Un'iniziativa formativa diversa, per contenuti e metodi, dalla formazione generalmente associata al tema della trasformazione digitale, che si propone prevalentemente di addestrare gli utenti all'utilizzo dei nuovi dispositivi, a superare le cosiddette “resistenze al cambiamento”, a formare forza lavoro specializzata nello sviluppo del software e nella gestione dei sistemi, a promuovere una “cultura dell'innovazione” che continua a ritenere la “digitalizzazione” un bene di per sé, indipendentemente dalle conseguenze del suo utilizzo.

Un'iniziativa formativa rivolta esplicitamente ai non “addetti ai lavori”.

Come si legge nella presentazione della scuola: “I soggetti cui il corso si rivolge sono tutti coloro che producono ed utilizzano dati digitali nella scuola, nel lavoro, nella politica, con particolare attenzione agli operatori delle associazioni di cittadinanza, e che avvertono la necessità e l'urgenza di dotarsi di strumenti critici per comprendere la realtà attraverso i dati digitali, e per poter operare in essa con autonomia e consapevolezza”.

Questi sono anche i lettori che ci auguriamo possano leggere queste pagine.

7. Perché i dati

Ma, identificati i destinatari, su quali aspetti della trasformazione digitale è utile concentrare lo sforzo formativo? In altri termini, tra i tanti contenuti possibili di un'iniziativa formativa dedicata a generare consa-

pevolezza critica sulla trasformazione digitale, quali selezionare e proporre all'attenzione dei destinatari?

La scelta è stata quella di assumere come perno di tutto il corso il tema dei “dati”.

Non solo per la centralità che i dati hanno nello sviluppo di tutti i sistemi digitali, ivi compresi i più recenti sviluppi dei sistemi di intelligenza artificiale basati sulle tecnologie di apprendimento automatico. Ma anche per la particolare caratteristica dei dati di poter essere compresi, analizzati e utilizzati anche indipendentemente dal loro utilizzo nei sistemi digitali, di essere cioè “prossimi” all'esperienza quotidiana di utenti non esperti, di possedere caratteristiche immediatamente descrivibili e riconoscibili anche senza l'utilizzo di un linguaggio specialistico.

Inoltre la Scuola è stata progettata e realizzata nel corso del 2021, cioè in piena pandemia Covid, e anche questo ha contribuito a definire le sue caratteristiche e la sua utilità.

In primo luogo perché, come sappiamo, la pandemia ha generato un gigantesco switch-off digitale ovvero l'impossibilità di proseguire in molti aspetti della nostra vita con modalità che non prevedessero il ricorso esclusivo ai dispositivi digitali. La sua rapidità ha costretto ognuna/o di noi ad un salto repentino nella propria capacità di uso degli strumenti digitali, pena l'essere esclusi, ad esempio, dal mantenere le proprie relazioni sociali e affettive. E questo ha grandemente contribuito ad ampliare la platea di chi era potenzialmente interessato ai contenuti della Scuola.

In secondo luogo perché proprio il confrontarsi con l'evolversi della pandemia ha portato a una grande attenzione ai dati, al loro significato, alla loro qualità, alla loro analisi, trattamento e comprensione.

Di tutto questo vi è un riflesso diretto nelle lezioni della scuola e nelle pagine del libro, in cui molti esempi sono proprio dedicati ad approfondire le caratteristiche dei dati epidemiologici relativi alla pandemia.

8. Un libro multidisciplinare

Alla realizzazione dell'esperienza della “Scuola Pop sui Dati Digitali” hanno contribuito principalmente tre soggetti: il Forum Disuguaglianze e

Diversità, che ha tra i suoi obiettivi quello di “Riorientare la trasformazione digitale verso obiettivi di giustizia sociale”, il Centro per la Riforma dello Stato, che ha promosso, a partire dal 2017, il progetto “Scuola critica del digitale”, con l’obiettivo di produrre e diffondere cultura politica critica sulla trasformazione digitale, e Carlo Batini, che proprio al tema dei dati ha dedicato la sua vita di docente, studioso e membro dall’Autorità per l’Informatica nella Pubblica Amministrazione e che aveva da poco pubblicato il libro *Scienza dei Dati* (<https://boa.unimib.it/handle/10281/295980>).

All’esperienza della Scuola hanno partecipato attivamente alcune decine di operatori delle associazioni di cittadinanza, di militanti sindacali, di insegnanti.

È anche dalla qualità della loro partecipazione e dal loro esplicito incoraggiamento che è nato il progetto di questo libro, i cui dodici capitoli traggono origine direttamente dalle dodici lezioni della Scuola. Al termine di ogni capitolo è indicato un link tramite il quale, chi è interessato, potrà seguire la registrazione della lezione corrispondente.

Nel terzo paragrafo della lezione 1 sono descritti tutti contenuti del libro, e ad esso rimandiamo per un’esposizione completa degli argomenti trattati.

Ci interessa qui mettere in evidenza che questo libro, come anche la Scuola, è frutto di una collaborazione tra docenti e studiosi di diverse discipline, sia scientifiche che umanistiche. A informatici come Carlo Batini, Federico Cabitza e Gaetano Santucci, si affiancano psicologi come Paolo Cherubini, giuristi come come Giorgio Resta e filosofe come Teresa Numerico. E non può essere che così quando si assume il tema dei dati come oggetto di studio e di formazione.

9. Si può fare

Da quanto detto finora sono chiari gli obiettivi di questo libro, ed esplicita l’ambizione a fare di queste pagine uno strumento che aiuti chi lo leggerà non solo ad orientarsi nel sempre più rapido svolgersi di quella che il PNRR chiama “transizione digitale”, ma anche a poter intervenire in essa non da utenti passivi, ma da protagonisti, con più forza e incisività.

A non delegare solo agli addetti ai lavori di determinare scopi e modalità dei processi di innovazione. A disturbare i manovratori dell'apparente ineluttabilità della direzione della trasformazione.

Ad esempio di quello che ci piacerebbe fosse l'uso di queste pagine vogliamo citare un episodio avvenuto a Roma nel maggio 2023.

Il 4 maggio il Consiglio di Istituto del Liceo Classico Pilo Albertelli di Roma ha respinto il progetto "Next Generation Lab" finanziato con fondi del PNRR Scuola (circa 125.000 euro).

Il progetto prevedeva lo sviluppo delle "professioni digitali del futuro". Le relative "competenze digitali specifiche" previste sono "saper girare video con uno smartphone, saper realizzare filmati e pillole per i social con attenzione crescente ai contenuti per le Instagram stories, saper analizzare i dati e i trend di ascolto streaming dei brani musicali".

Nel rispondere alle critiche della stampa e del Ministero i rappresentanti dei genitori hanno scritto: "Forse è troppo complesso da capire o troppo scomodo da dire che quanto sta avvenendo non è la contesa tra innovazione e opportunità da una parte e incompetenza e ideologia dall'altra? Ci dispiace deludere: tra noi ci sono ingegneri, informatici, fisici, matematici (ma anche insegnanti, operatori sociali, lavoratori autonomi, impiegati e operai); lavoriamo con le tecnologie e sulle tecnologie e sappiamo bene che il progresso tecnologico richiede una sempre maggiore complessità e profondità ed un pensiero critico che si nutre di conoscenza disinteressata. Solo con più cultura si può usare la tecnologia per il bene comune ed i mezzi tecnici possono restare tali e non trasformarsi in fini".

Ai genitori del Consiglio di Istituto del liceo Albertelli dedichiamo questo libro.