

ASPETTI ETICI DELLA RICERCA E
DELL'INNOVAZIONE TECNOLOGICA:
IL CASO DELL'INFORMATICA
MONDI VIRTUALI, RETI SOCIALI, SISTEMI AUTONOMI
“INTELLIGENTI”

CARLO GHEZZI (*)

SUNTO. – Gli sviluppi della ricerca e dell'innovazione tecnologica nel settore dell'Informatica stanno rivoluzionando la società contemporanea. Viviamo in un mondo “ciber-fisico” in cui le persone interagiscono e collaborano tra di loro in maniera continua nello spazio fisico e nello spazio “virtuale”, e agenti “virtuali” automatizzano ogni tipo di attività, incluse quelle che tradizionalmente venivano riservate all'intelligenza umana. Ogni individuo può accedere a dati e informazioni ovunque si trovi e in ogni momento; ogni barriera è rimossa, ma al tempo stesso sono possibili utilizzi distorti. Qualunque azione degli individui risulta mediata dalla tecnologia, che potenzialmente invade ogni angolo della sfera privata. Questi cambiamenti sono avvenuti e avvengono così rapidamente e con effetti così spettacolari che le loro implicazioni etiche sono state a lungo ignorate, per poi esplodere troppo tardi in maniera drammatica. Questo contributo ha l'obiettivo di introdurre brevemente i cambiamenti profondi generati dai progressi dell'Informatica e le conseguenti implicazioni etiche. Inoltre suggerisce che solo attraverso un approccio interdisciplinare si possono affrontare i problemi etici generati dalla ricerca e dall'innovazione e che ciò debba avvenire durante gli sviluppi della ricerca e della tecnologia, e non mediante tardive e inefficaci riflessioni a posteriori.

ABSTRACT. – Advances in research and technological innovations in Informatics are revolutionizing our society. We already live in a cyber-physical world: interaction and collaboration among humans occurs seamlessly both in physical and in virtual spaces and virtual agents automate tasks at all levels, including those that traditionally relied on human intelligence. Access to data and information is available to everybody, anywhere and at any time, eliminating barriers but also generating potential distorted usage.

(*) Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Milano, Italia. E-mail: carlo.ghezzi@polimi.it

Whatever humans do is inevitably mediated by technology, which potentially invades the private sphere of individuals. Changes are happening so fast and their effects are so spectacular that their ethical implications are often understood too late or underestimated. This contribution briefly introduces the disruptive changes produced by advances in Informatics and their severe ethical implications. It argues that an interdisciplinary effort is needed to focus on ethical issues while research and innovation are developed rather than as an after-the-fact reflection.

L'informatica sta alla base della rivoluzione digitale che ha investito la società contemporanea. Le tecnologie digitali, il software e gli algoritmi sempre più sofisticati hanno creato un mondo globale interconnesso. In questo mondo, entità artificiali che posseggono livelli di autonomia decisionale e capacità operative sofisticate sempre più interagiscono e cooperano con l'uomo e l'ambiente fisico, realizzando un mondo ciber-fisico che solo qualche anno fa era inimmaginabile. Sempre più viene utilizzato il termine suggestivo **intelligenza artificiale**, anziché il più prosaico termine *informatica*, per indicare la radice scientifica, metodologica e tecnologica che ha portato alla costruzione di sistemi che sono in grado di sostituire e addirittura migliorare abilità che fino ad ora hanno richiesto l'intelligenza umana per essere attuate.¹

In queste note vorrei offrire una riflessione sulla molteplicità dei risvolti etici che sono introdotti dai recenti sviluppi della ricerca e innovazione informatica, in particolare soffermandomi su ciò che mi pare fondamentalmente nuovo rispetto alle riflessioni che storicamente sono state fatte sulle relazioni tra scienza/tecnologia ed etica. L'identificazione del "nuovo" dovrebbe dunque portare a nuove riflessioni teoriche, da cui far seguire nuovi modelli comportamentali e nuove "regole".

Al fine di meglio comprendere la varietà e la specificità dei problemi etici posti dalla scienza e tecnologia informatica ritengo opportuno partire da elementi di riflessione su alcuni fatti reali di grande importanza osservati nella società contemporanea, correlati ai tratti salienti

¹ Non mi soffermo qui sulla discussione circa il termine "Intelligenza Artificiale", coniato nel 1955 da John McCarthy, e la sua adeguatezza a definire esattamente i confini di una disciplina. Si veda ad esempio l'articolo scientifico " D.L. Parnas, The Real Risks of Artificial Intelligence", Comm ACM, Ottobre 2017 (<https://cacm.acm.org/magazines/2017/10/221330-the-real-risks-of-artificial-intelligence/fulltext>), o anche la posizione espressa in numerosi interventi pubblici dal filosofo Luciano Floridi (per esempio quello apparso sul Corriere della Sera – Innovazione, 1-12-2017).

dell'evoluzione scientifico-tecnologica avvenuta nell'ultimo ventennio.

La pubblica opinione è stata recentemente posta di fronte al grave problema delle “fake news”, che si è dimostrato aver avuto un ruolo primario nell'influenzare la pubblica opinione riguardo a importanti decisioni politiche, quali la Brexit o le elezioni presidenziali negli Stati Uniti, o più recentemente le elezioni presidenziali in Brasile. Il fenomeno consiste nel fare propaganda diffondendo notizie mirate false, attraverso mezzi di comunicazione che sono in grado di raggiungere i singoli individui. Tutto ciò è reso possibile dal fatto che oggi esiste un unico, gigantesco e pervasivo mezzo di comunicazione (la rete Internet), che sostituisce ormai la pleora di mezzi trasmissivi del passato (dal vecchio telefono al telegrafo, alla posta, la radio, i giornali, ...). Tutto oggi passa su un'unica infrastruttura ed arriva sullo smartphone di ogni singolo individuo. Questa gigantesca infrastruttura che tutto connette, e che ormai connette tutti gli oggetti che ci circondano (la cosiddetta “Internet of Things”) consente ad osservatori interessati di raccogliere informazioni su tutto quanto transita sulla rete. In particolare, consente di usare queste informazioni per costruire un “profilo” delle persone attive nella rete, vale a dire una caratterizzazione di ciascuno di noi in termini di interessi personali, preferenze, capacità di spesa, ecc. Da qui, la possibilità di propaganda personalizzata e di azioni dirette verso individui specifici con determinate caratteristiche al fine di orientarne il comportamento in una certa direzione per un ben preciso fine, senza che l'individuo coinvolto se ne renda conto. Negli esempi citati, il profilo dell'utente è stato utilizzato per una campagna volta ad orientare gli elettori in direzioni politiche ben precise.

Ovviamente il profilo si basa su dati che spesso noi stessi esplicitamente rilasciamo e talvolta (ma non sempre) non su dati fraudolentemente acquisiti. Va peraltro osservato che non sempre questo rilascio esplicito deriva da una scelta razionale ben informata e consapevolmente presa.² Altrettanto ovviamente, l'esistenza di questo profilo personale a volte ci fa comodo; ad esempio, quando Amazon ci suggerisce l'acquisto di certi libri dopo che abbiamo evidenziato i nostri interessi in base agli acquisti passati o come quando un filtro è in grado di eliminare

² L'UE ha recentemente affrontato con grande lungimiranza e in modo molto innovativo il tema della *privatizza* dei dati personali con il GDPR 2018 (Regolamento Generale sulla Protezione dei Dati).

automaticamente dalla nostra casella di posta i messaggi indesiderati, in base al nostro profilo personale relativo a quali siano le sorgenti che consideriamo indesiderate. Ci disturba invece il suo utilizzo a scopo di propaganda politica, o da parte delle compagnie assicurative, che possono utilizzare informazioni in merito alla vita privata al fine di decidere il costo di una polizza da proporre, o da parte di una società a cui un giovane fa domanda di assunzione, che può utilizzare informazioni sulle preferenze politiche o sull'orientamento sessuale per decidere se fare un'offerta di lavoro.

Questi scenari sono ormai tipici della realtà di oggi (non di un ipotetico futuro) e sono possibili grazie a quattro fondamentali componenti tecnologiche:

- l'esistenza di un'infrastruttura globale attraverso cui passano tutte le informazioni (Internet),
- la possibilità di archiviare come dati permanenti tutti gli accadimenti e i fenomeni osservati,
- l'esistenza di dispositivi (gli smartphone, i personal computer, ...) che possono essere utilizzati da ogni persona per accedere ai dati di Internet e per creare nuovi dati,
- la possibilità di elaborare tutti questi dati, attraverso applicazioni software disseminate sulla rete, al fine di scoprire (imparare) correlazioni significative tra i dati e quindi scoprire come interagire con gli individui e con gli elementi fisici che sono presenti nella rete per fini specifici.

I termini *Big Data* e *Machine Learning* vengono comunemente utilizzati per far riferimento alla enorme quantità di dati oggi disponibili, e inimmaginabile in passato, e alla possibilità di apprendere automaticamente interessanti correlazioni tra questi dati. Il termine *Machine Learning* merita una riflessione un poco più approfondita. La capacità di apprendere contraddistingue l'uomo da (quasi) tutti gli esseri viventi: costituisce uno dei tratti essenziali di ciò che consideriamo un comportamento intelligente. Non a caso, l'*Automatic* (o *Machine Learning*) è fin dall'inizio stato uno dei settori costitutivi dell'Intelligenza Artificiale. Bisogna però anche osservare che oggi ciò che va sotto questo nome non ha nulla a che fare con quanto i ricercatori intendevano con questo termine fino a pochi anni fa.³ Oggi con esso si intende l'uso

³ Questo concetto è stato espresso in modo molto convincente e in molti

di sofisticati algoritmi, basati su metodi statistici e in particolare su reti neurali (i cosiddetti metodi di *Deep Learning*) che consentono di scoprire correlazioni partendo da grandi masse di dati, mentre originariamente si pensava di riprodurre il funzionamento della mente umana razioncinante, realizzando i meccanismi deduttivi della logica matematica con metodi di calcolo.

Questa digressione fa capire dunque che il termine *learning* viene utilizzato per indicare la capacità di correlare e generalizzare leggi di dipendenza tra fenomeni osservati, tipicamente al fine di classificare i comportamenti osservati all'interno di classi specifiche. Per apprendere, l'algoritmo non deve *capire* un fenomeno osservato, ma può riprodurlo se ha potuto osservare un sufficientemente grande insieme di esempi. E' così che un algoritmo è in grado di scoprire errori di ortografia, o di suggerire cose da comprare. Queste tecniche di *machine learning* hanno avuto in alcuni campi successi clamorosi. Per esempio, si è riusciti a progettare un agente artificiale che è stato in grado di sconfiggere il campione mondiale di Go. Go è un gioco da tavolo strategico molto popolare in Asia ed estremamente complesso strategicamente. Analogamente, si sono avuti enormi successi nell'ambito del riconoscimento automatico di immagini con tecniche di *Deep Learning*. Questi sistemi stanno rendendo possibile la realizzazione di veicoli in grado di "vedere" e distinguere automaticamente altri veicoli, i pedoni, e in generale i diversi agenti che possono essere presenti in uno scenario in cui i veicoli possono muoversi autonomamente, senza l'intervento di un guidatore umano.

E' possibile dunque affermare che la grande evoluzione scientifico-tecnologica degli ultimi decenni nasce da un incontro tra gli algoritmi avanzati che consentono elaborazioni complesse, la capacità di realizzare in software affidabile questi algoritmi, la disponibilità di enormi quantità di dati accessibili mediante una rete globale ed elaborati dal software e la disponibilità di infrastrutture di elaborazione sempre più potenti e distribuite.

Oggi viviamo dunque in mondo a due dimensioni (mondo ciberfisico): fisico e virtuale. La dimensione virtuale diventa sempre più pre-

interventi pubblici da un importante ricercatore, Nello Cristianini (si veda, ad esempio, il suo intervento su "Living in a Data Obsessed Society" nell'ambito del convegno *Ethics for the Technological Innovation*, Politecnico di Milano, 19-10-2017).

valente. E il mondo virtuale non è solo il luogo in cui noi operiamo come soggetti, ma anche il luogo in cui noi siamo oggetti.

Un aspetto fondamentale della rivoluzione tecnologica portata da Internet è che a ogni individuo viene data una possibilità di operare come protagonista in un mondo *virtuale* che fa a meno di infrastrutture di intermediazione. Chiunque, usando ad esempio Facebook sul proprio telefono, può pubblicare un filmato, può comunicare o collaborare con chiunque. In linea di principio, chiunque può acquistare e vendere in un mercato planetario. In questo mondo virtuale nascono nuove forme di lavoro, nuovi modelli di fare *business*, nuovi modi di comunicare, di pubblicare, di produrre e fruire contenuti (musica, testi, immagini, ...).

Apparentemente, il mondo virtuale si presenta come un mondo di compiuta ed estrema *democrazia*. Nulla di più illusorio, se si osserva il fenomeno in maniera critica. In realtà, non è affatto vero che sono scomparsi gli intermediari. E' vero che sono scomparsi gli intermediari tradizionali, quelli che esistevano nella "vecchia" società nella quale abbiamo vissuto fino a ieri, che avevamo imparato a conoscere, e rispetto alla quali avevamo definito precise regole di comportamento (in molti casi, regolamentate attraverso la norma giuridica). A questi se ne sono però sostituiti altri, che ancora non abbiamo ben capito e che già stanno provocando i loro buoni e cattivi effetti.

Un altro aspetto cruciale del mondo virtuale e dell'apparente scomparsa di intermediari è la progressiva scomparsa dei meccanismi tradizionali di selezione e certificazione dell'informazione, di produzione, verifica e custodia della conoscenza. Si vengono così a creare movimenti di opinione in cui l'inesistente nesso vaccini-autismo sostenuto in maniera fraudolenta dall'ex-medico Andrew Wakefield viene spiegato come congiura delle case farmaceutiche che "lucrano" sui vaccini. Risposte a quesiti medici vengono cercate consultando Internet invece del medico specialista. Più in generale, l'accesso diretto ai dati viene confuso con la possibilità di far a meno dell'*esperto*, visto come intermediario-casta, generando quindi il rigetto della scienza che sa spiegare e dello scienziato che sa scoprire che cosa sta dietro i dati.

Si potrebbe qui fare una riflessione sull'influenza che questa visione ha anche sul modo di fare scienza. Qualcuno pensa che i successi del *machine learning* dimostrino la fine di un modo di fare ricerca scientifica che cerca di capire il *perchè* di certi fenomeni e costruisce *modelli* capaci di spiegarli. L'obiettivo diventa scoprire certe correlazioni e farne uso, non tanto cercare di spiegarle.

In realtà la necessità di spiegare il perché, cercando di passare da semplici correlazioni a relazioni di causalità, è una componente fondamentale della ricerca scientifica: l'approccio *bottom-up* dell'apprendimento automatico e quello *top-down* di tipo modellistico possono e devono invece essere visti come complementari.

Dopo questo excursus è bene ritornare a cercare di sistematizzare gli aspetti etici della ricerca e innovazione informatica che ho tratteggiato.

Gli aspetti di natura etica nascono da motivi di diversa natura. Innanzitutto, in quanto si tratta di tecnologia *pervasiva* che oggi tocca ogni individuo, nelle sue relazioni con altri individui, con il mondo fisico e l'ambiente nel quale l'individuo è immerso, con altri individui ai quali è interconnesso, con le organizzazioni (di ogni tipo) con le quali interagisce. Ovviamente sono toccate anche le organizzazioni, le interazioni tra le organizzazioni e tra organizzazioni e individui. Un tipico esempio di problema di natura etica riguarda la tensione tra la protezione della sfera individuale (per esempio, la riservatezza di dati personali) e la propria presenza nel mondo virtuale, che di fatto rende tracciabili i comportamenti individuali. Tipicamente, la tensione tra privacy e sicurezza personale. Il cosiddetto *consenso informato* alla cessione di dati personali è sicuramente un importante riconoscimento della necessità di protezione di dei dati personali, ma fornisce solo un pallido e talvolta inefficace esempio di norma di salvaguardia del principio etico di riservatezza. Analogamente, si pensi al problema del diritto all'oblio per dati personali resi disponibili in passato e alla reale possibilità di renderlo totalmente efficace.

La tecnologia informatica non solo automatizza funzionalità esistenti nel mondo fisico reale, ma crea nuovi *mondi virtuali*, mai visti in precedenza. Si prendano ad esempio i *social network* come Facebook, i nuovi concetti e ruoli di "amico", "follower", etc. Si prenda un altro caso oggi molto discusso: quello del denaro e del suo utilizzo. Da qualche anno è nato e si sta sviluppando in maniera inarrestabile il concetto di denaro digitale. Bitcoin, la prima moneta digitale introdotta nel 2008, sta già avendo un enorme successo: il suo valore rispetto al dollaro si è decuplicato dall'inizio dell'anno e va moltiplicato per 30000 rispetto al valore nel 2011! Il bitcoin si basa sullo scambio diretto e certificato tra due contraenti e non richiede l'intermediazione delle banche. Può così nascere un nuovo mondo finanziario, diverso da quello che abbiamo imparato a conoscere da secoli. A questi nuovi mondi virtuali creati dalla ricerca e dall'innovazione tecnologica è difficile se non

impossibile estrapolare in maniera diretta i concetti e principi etici che si sono consolidati nel mondo fin qui conosciuto, nonché le norme del diritto che erano applicabili in passato.

Alla difficoltà di replicare concetti e principi etici del passato, nonché ovviamente le norme e le leggi che le regolano nella società, va aggiunto che nuovi principi e nuove norme devono essere continuamente studiati e attuati, a una velocità mai vista in passato. Il progresso scientifico e tecnologico sta avvenendo a una velocità che rende del tutto inefficace l'approccio metodologico tradizionale, di intervento ex-post, quando l'innovazione è diventata tecnologia di uso quotidiano. Va tenuto infatti presente che la diffusione della tecnologia può avvenire in tempi estremamente rapidi. Per fare un esempio, Facebook, il primo esempio di social network lanciato nel 2006, a 10 anni di distanza ha più di 2 miliardi di utenti.⁴ A me pare che sia necessario che fin dal momento della ricerca scientifica debbano essere prese in considerazione le potenziali implicazioni etiche, e che ciò debba continuare con la progettazione delle soluzioni tecnologiche, prima che queste entrino a valanga in un mondo impreparato ad accoglierle, e potenzialmente vulnerabile da un punto di vista etico.

Ciò comporta un nuovo approccio, necessariamente interdisciplinare, che vede gli scienziati degli algoritmi e i tecnologi informatici collaborare con filosofi, sociologi, antropologi, giuristi nello sviluppo di ricerca e innovazione che veda al proprio centro l'uomo.

⁴ Si noti che i cinesi non possono sottoscrivere Facebook, e che un analogo di Facebook (WeChat) è stato creato per l'enorme mercato cinese e per consentire il livello di controllo dei dati ai fini politici.