ATTI DEL CONVEGNO

IA
E STATO DI DIRITTO

Sfide e Opportunità per le Democrazie nell'Era Digitale

3 MAGGIO 2024

Roma, Camera dei Deputati
Sala del Cenacolo,
p.zza in Campo Marzio 42

Fondazione Marco Pannella

IA
E STATO DI DIRITTO

Sfide e Opportunità per le Democrazie nell'Era Digitale

Convegno organizzato dalla
Fondazione Marco Pannella

Media Partner
Radio Radicale

CAMERA DEI DEPUTATI
Sala del Cenacolo
p.zza in Campo Marzio 42

ROMA
3 maggio 2024

Le trascrizioni sono state generate automaticamente tramite il software di intelligenza artificiale GoodTape (goodtape.io) e revisionate da esseri umani.

I testi non sono stati rivisti dalle relatrici e dai relatori del Convegno.

Il 3 maggio 2024 presso la Sala del Cenacolo della Camera dei Deputati si è tenuto il convegno "Intelligenza Artificiale e Stato di Diritto: Sfide e Opportunità nell'Era Digitale", organizzato dalla Fondazione Marco Pannella.

L'evento ha rappresentato una occasione di dialogo interdisciplinare tra accademici, ricercatori, esperti del settore, policy makers, giornalisti e attivisti per esplorare le complesse interazioni tra IA, Stato di Diritto, Democrazia e Diritti Umani.

Il convegno è stato articolato in quattro sessioni tematiche, con relatori di spicco che hanno condiviso le loro competenze e prospettive sull'AI. Come può l'IA essere armonizzata con i principi universali dei diritti umani, bilanciando progresso tecnologico e valori tradizionali? In che modo l'IA sta ridefinendo le modalità di partecipazione democratica e quali sono le sue implicazioni per la rappresentanza e l'inclusione? Quali sono le sfide e le opportunità poste dall'IA nella gestione e nella validazione delle informazioni nell'era della post-verità? Come possono i decisori politici e gli sviluppatori di IA collaborare per garantire che l'evoluzione dell'IA sia guidata da principi etici e responsabili?

Il convegno è stato trasmesso integralmente sulla WebTV della Camera dei Deputati e su Radio Radicale.

Nelle pagine che seguono sono riportate le trascrizioni integrali di tutti gli interventi. Per ciascuna relazione è anche possibile scansionare il relativo QR Code per rivedere la registrazione audio-video dei singoli interventi.

Si ringraziano le deputate Deborah Bergamini e Annarita Patriarca per aver consentito di tenere il convegno nelle strutture della Camera dei Deputati, le relatrici e i relatori, Marco Cerrone per aver concepito, organizzato e coordinato l'evento, Marinella Lanza e la Segreteria di Forza Italia per il supporto organizzativo, Giovanna Reanda per la moderazione del dibattito, Radio Radicale per la media partnership e Giuseppe De Salvin per la revisione delle trascrizioni automatiche di tutte le relazioni.

INDICE

SALUTI ISTITUZIONALI

ANNARITA PATRIARCA
Segretario Ufficio di Presidenza, Camera dei Deputati

ASCOLTA INTERVENTO

Siamo ancora imperfetti, quindi ci possiamo permettere il lusso di sbagliare e attraverso l'errore di trovare, forse, soluzioni differenti.
Detto questo, ringrazio voi, ringrazio la vostra presenza qui. Interessante l'argomento, altamente impattante.
L'avvento dell'intelligenza artificiale ha inaugurato una nuova era per le istituzioni democratiche, mentre offre strumenti rivoluzionari per l'efficienza, l'innovazione. Presenta anche delle sfide significative.
Quella indicata dall'incontro di oggi sull'interazione tra intelligenza artificiale e Stato di diritto, ci permette di esaminare una tematica sicuramente complessa attraverso, ho visto, un approccio anche multidisciplinare che avete adottato.
Uno dei fulcri principali è il bilanciamento appunto delle nuove tecnologie e dei diritti umani. Sicuramente l'intelligenza artificiale può migliorare l'accesso ai servizi e alle funzioni anche governative, ma solleva una serie di questioni critiche riguardo, per esempio, la privacy, la sorveglianza e l'equità.
Come possiamo assicurare che le applicazioni di intelligenza artificiale rispettino i diritti umani fondamentali e contribuiscono a creare una società più giusta?
Inoltre, la democratizzazione dell'accesso alle tecnologie di intelligenza artificiale è vitale affinché i benefici che possono derivare dall'applicazione dell'intelligenza artificiale possano essere equamente distribuiti.
Con l'intelligenza artificiale si stanno anche ridefinendo i modelli

di partecipazione pubblica e decisionale, offrendo anche nuove modalità di interazione tra i cittadini e le strutture governative. Ma anche questo solleva delle questioni.
Come le minoranze o per esempio i gruppi sottorappresentati possano partecipare efficacemente a questo nuovo contesto digitale?
La disinformazione è un'altra area critica della discussione. Con l'intelligenza artificiale c'è un ruolo sempre più importante nella distribuzione delle informazioni. Dobbiamo considerare meccanismi che garantiscono l'autenticità e l'affidabilità dei contenuti.
Quindi si pongono ulteriori tipi di problemi. Ci sono sfide significative per i regolatori, per gli sviluppatori di tecnologie. La veridicità e la trasparenza che dev'essere utilizzata nella produzione degli algoritmi diventa la sfida fondamentale.
La necessità di una governance etica dell'intelligenza artificiale è impellente.
È fondamentale che i decisori politici collaborino strettamente con gli sviluppatori di intelligenza artificiale. È stata posta prima una domanda.
“L'intelligenza artificiale ci sta portando ad un posto migliore o ad un posto peggiore?”
Secondo me la domanda vera da porci è: siamo in grado eventualmente di fermarlo? Cioè la questione è un'altra. Perché se non siamo in grado di fermare il processo e il progresso che si sta creando, non possiamo neanche entrare nella dinamica di: “il posto è migliore” o “il posto è peggiore”. È una domanda che ci poniamo inutilmente. La questione va posta diversamente. Visto che è un meccanismo in atto e non siamo in grado di fermarlo, secondo me neanche di rallentarlo, per una serie di dinamiche complesse, anche internazionali, allora il ragionamento è un altro.

Come regolarlo? Come gestirlo?
Il problema vero, oggi, dell'intelligenza artificiale è un problema di governance sicuramente, ma anche un problema di regole. Come deve essere orientata? Come deve essere regolata?
È a monte la problematica da affrontare. Non è che cosa determinerà l'intelligenza artificiale. O meglio, che cosa determinerà l'intelligenza artificiale dipende dalle regole che noi adottiamo nell'evoluzione, nell'applicazione e nella gestione di tutti i processi che hanno a che fare con l'applicazione dell'intelligenza artificiale a più livelli.
Concludendo, sicuramente l'era digitale offre immense opportunità e nuove entusiasmanti sfide.
Come qualsiasi sfida e come qualsiasi innovazione ha dei punti oscuri. Tutto quello che non è ancora sperimentato è oscuro perché non conosciuto.
Una delle sfide principali che ci dobbiamo porre, e ci dobbiamo porre anche noi come classe politica, è quella di garantire la possibilità a tutti di accedere e di disporre, in una condizione di equità, nell'accesso a queste tecnologie. Tra gli altri temi etici che vi siete posti e ci continueremo a porre, c'è una frase di Tim Berners-Lee, il papà di internet come lo conosciamo oggi, che dice: "Abbiamo una responsabilità per assicurarci che queste tecnologie siano disponibili per tutti e non solo per chi può permetterselo". C'è un impegno condiviso che dobbiamo avere anche come decisori politici per garantire l'inclusività, l'equità, oltre al rispetto dei diritti fondamentali che è prioritario rispetto a ogni altro tipo di valutazione.
Abbiamo un dovere di ascolto, noi come decisori politici, con chi quotidianamente si trova a contatto e lavora con questa straordinaria tecnologia.
Abbiamo un dovere di interazione e abbiamo un dovere di parlare

nuovi linguaggi, di comprendere questi nuovi linguaggi.
Le problematiche che si possono verificare, che allarmano maggiormente, sono quelle legate all'immediato impatto con l'opinione pubblica e al mondo del lavoro, alle problematiche di perdita eventuale di posti di lavoro o, contemporaneamente, sviluppo di altre competenze e di nuovi tipi di lavoro sconosciuti prima.
Che cosa è fondamentale?
La capacità di far dialogare. La capacità di far dialogare negli ambienti di lavoro, che sono quelli più impattanti in questo momento dall'avvento dell'intelligenza artificiale. La capacità di chi deve assumere le decisioni e stabilire le regole, di dialogare con chi costruisce questa intelligenza artificiale e l'assoluta capacità di entrare dentro questo meccanismo, influenzandolo dall'interno. Quando affrontiamo tematiche che sembrano diverse, ma sono il principio di un'evoluzione di problematiche che noi avremo, quando parliamo dell'accesso alla rete, già oggi, dei ragazzi indiscriminati, di tutte le problematiche che ne derivano, dal cyberbullismo a scendere giù, il problema che abbiamo è l'incapacità di controllo. L'incapacità di controllo dipende anche da una serie di meccanismi e di linguaggi che non ancora ci appartengono. Parlo di noi, decisori politici: quindi, spesso, anche una serie di norme che noi pensiamo possano essere efficaci, nella realtà pratica non riescono ad ottenere l'obiettivo che noi ci prefiggiamo. È una sfida anche per noi, tutto quello che ne deriva.
Trovare nuovi strumenti, parlare nuovi linguaggi, adattare il tessuto normativo a quella che è l'evoluzione che sta arrivando. Ripeto, tutelando una serie di principi fondamentali. Tra questi, anche l'accesso e l'equità nell'accesso a questo tipo di tecnologie.
Grazie.

INTRODUZIONE

MAURIZIO TURCO
Presidente Fondazione Marco Pannella

ASCOLTA INTERVENTO

Noi abbiamo voluto questo convegno in occasione anche di una data importante per noi che è l'anniversario della nascita di Marco Pannella. Abbiamo pensato di onorarla con una di quelle cose che ci hanno caratterizzato, cioè cercare di vedere dove altri ancora non hanno visto.
Io partirei da una considerazione che è quella di Stephen Hawking, quando dieci anni fa disse che lo sviluppo dell'intelligenza artificiale potrebbe segnare la fine della razza umana. Era un dato di catastrofismo o di preveggenza?
Questo sarà da capire e lo capiremo dalle vostre relazioni, però di sicuro c'è un dato di preoccupazione.
Di fronte a un dato di preoccupazione (ormai evidente e generalizzato) noi pensiamo che bisogna occuparsene e, quindi, come ricordava Marco Cerrone, che ringrazio (perché poi è lui che ha messo in piedi tutto questo) ricordo, come ha già detto, che non è che un primo passo.
Perché, qual è il nostro interesse?
Il nostro interesse è di superare quello che (è accaduto già in altre occasioni) è il divorzio tra la scienza e la politica nel senso che la politica arriva sempre un po' dopo e ci arriva anche un po' male. Quindi cercare di fornire degli strumenti per governare l'uso dell'intelligenza artificiale e di orientarlo.
Naturalmente a noi interessa soprattutto quello che è il rapporto tra l'intelligenza artificiale e lo Stato di diritto, la democrazia. Come può aiutare la democrazia o come invece magari potrebbe

affossarla, direi definitivamente, vista la degenerazione degli ultimi decenni?
Quindi questo è un punto di partenza.
Saranno i relatori a darci gli strumenti per capire come potremo, dovremo andare avanti, ma l'obiettivo è quello. L'obiettivo è di arrivare a un certo punto a fornire strumenti di iniziativa politica alle varie forze politiche.
Penso che potremmo a questo punto iniziare. Grazie a tutti coloro che hanno accettato questo invito e buon lavoro.

IA: LA PAURA DEL CAMBIAMENTO

ANDREA R. CASTALDO
Professore Ordinario di Diritto Penale,
Università degli Studi di Salerno
Avvocato

ASCOLTA INTERVENTO

Buongiorno a tutti, grazie alla Fondazione Marco Pannella, grazie al Presidente Turco e consentitemi di ringraziare Marco Cerrone che ho avuto il piacere di conoscere.
Da qui nasce il discorso comune in relazione al corso internazionale di diritto penale e alle ricerche che noi stiamo portando avanti presso la Cattedra di Diritto Penale, in relazione soprattutto al metaverso e - in pochissimi minuti - cercherò di esporre quelli che sono i profili relativi all'intelligenza artificiale e al diritto penale.
Ciò detto, io mi iscrivo nella categoria di coloro - lo dico subito - che ritengono che sia catastrofico il richiamo sulle opportunità o le sfide legate all'intelligenza artificiale e cercherò di dimostrarlo attraverso queste brevi riflessioni che sono così articolate.
Chi ha paura del lupo cattivo? Quali sono le intelligenze artificiali che tagliano e cuciono? Cercherò di spiegare che cosa significa questo che potrebbe apparire uno slogan.
Quali sono le intrusioni nel e del diritto penale in relazione all'intelligenza artificiale.
Primo punto: provocatoriamente, l'intelligenza artificiale è il lupo cattivo? Io penso che non sia né un lupo, né che sia cattivo.
Ora, incominciamo col dire che gli psicologi hanno un fior fiore di ricerche sul punto che qualsiasi cosa nuova comporta una diffidenza o una somatizzazione verso aspetti di paura. Quindi è qualcosa di abbastanza naturale. Non è "nihil novi sub sole", perché qualsiasi cosa che rappresenti un fatto al quale non siamo

ancora avvezzi comporta un comportamento (lo hanno studiato gli etologi) come negli animali: un comportamento di irrigidimento rispetto a una sorpresa che non faceva parte di quello che gli psicologi hanno chiamato il “normale behavior” della specie.
Del resto, basti pensare al fatto curioso, ma direi normale, dei medici che sconsigliavano l'uso dell'ascensore in quanto si riteneva che potesse avere effetti dannosi in ordine alla gravidanza e anche a comportamenti isterici del sesso femminile.
Vorrei anche ricordare (e lo prendiamo con oggi con una sorriso) la penicillina, che nel 1928 è stata scoperta, ma fino al 1940 non veniva prescritta perché si riteneva che vi fossero delle controindicazioni.
Allora, detto ciò, il punto fondamentale è che noi tabuizziamo questo discorso dell'intelligenza artificiale, ma in realtà l'abbiamo già interiorizzato e metabolizzato.
Ci svegliamo grazie ad un richiamo, ad una sveglia. Attraverso il GPS impostiamo il percorso che dobbiamo fare, ci viene detto qual è il percorso migliore e tutta la nostra quotidianità è scandita da profili di intelligenza artificiale. Poi abbiamo la nuova frontiera che è l'intelligenza generativa, che è qualcosa di diverso e che è qualcosa che deve non preoccuparci, ma che deve essere oggetto di studio e di approfondimento.
Perciò non è né il lupo delle fiabe, né il lupo cattivo. È qualcosa che fa parte già della nostra quotidianità.
Il secondo punto era quello dell'intelligenza artificiale che ho definito “taglia e cuce”.
Cosa voglio dire con questa espressione? Intanto, l'intelligenza artificiale spersonalizza il rapporto: siamo in streaming, siamo con una platea virtuale. Questo significa che il rapporto umano, il rapporto “face to face” è un rapporto che viene ad essere posto in un’ottica marginale, non è più il momento topico centrale delle

relazioni umane. Questo significa che l'intelligenza artificiale taglia il contatto interpersonale. È un bene? È un male? Beh, per esempio, in alcuni casi, è un male non necessario (qui vengo alla esperienza del penalista) e cioè il processo penale, la psicologia del testimone: noi abbiamo bisogno di un testimone in carne ed ossa perché la fisiognomica ci insegna la reazione, quindi quel contatto empatico personale è qualcosa di fondamentale per comprendere per esempio se il teste è reticente o ci sta dicendo una menzogna. L'avatar, che oggi è possibile spersonalizzare, e quindi rappresentare, è qualcosa che non ci permette di saggiarne la attendibilità sotto il profilo della psicologia testimoniale.
Quindi taglia.
E un processo penale che deve essere a mio avviso legato al principio di oralità, al principio della rappresentazione della persona è qualcosa quindi che non si sposa con questa idea di intelligenza artificiale che "taglia". Ma l'intelligenza artificiale "cuce". E questo è un merito. Questo è un vantaggio.
Cosa voglio dire con un'intelligenza artificiale che cuce? Che ci consente di parlare con una platea mondiale, quindi elimina o riduce drasticamente quelle che sono le barriere, quelli che sono i profili territoriali e geografici.
È una grandissima acquisizione, se noi consideriamo le vecchie generazioni che si affidavano per il discorso dell'immigrazione alle lettere e quindi ad una distanza temporale estremamente ampia, mentre noi, oggi, non soltanto rappresentiamo con un WhatsApp, o con una chat, immediatamente il nostro pensiero, ma vediamo anche materializzata questa persona.
L'intelligenza artificiale cuce…
Ora, siccome i numeri sono importanti, le statistiche sono importanti perché consentono, non di manipolare, ma consentono di rappresentare plasticamente il dato per poi quelle che sono le

valutazioni, mi ha incuriosito (ma rimetto alla vostra attenzione) un dato, come dire, di straordinario impatto e cioè il fatto che (apparentemente potrebbe non rappresentare nulla, ma è una sfida politica, è qualcosa che invece interessa il mondo) riguarda la denatalità in relazione alla realtà virtuale e all'intelligenza artificiale. E qui voi mi direte: ma che c'entra l'intelligenza artificiale con la denatalità e le politiche di implementazione del tasso demografico che comportano e comprendono delle situazioni drammatiche sotto il profilo della geografia economica? Ebbene. C'è una interessantissima ricerca del "The Journal of Sex Research" recentissima, che dice come le nuove generazioni, soprattutto in Corea del Sud, in Giappone e anche negli Stati Uniti, abbiano ridotto drasticamente il rapporto sessuale: è diventato qualcosa che non interessa più (ci fa sorridere), ma è legato al discorso della realtà virtuale e dell'intelligenza artificiale.

Come vedete, se noi ragioniamo con una mentalità, come dire, legata al contingente non riusciamo nemmeno a comprendere quelle che sono le sfide mondiali, neppure europee, ma a livello internazionale, perché effettivamente il rappresentare il mondo in una realtà virtuale, in una bolla, comporta un'attenzione topica su quella realtà virtuale, che distoglie da altri aspetti e, distogliere da altri aspetti, significa anche distogliere dal profilo della natalità.

Torno, e sto per concludere, non voglio tediarvi, al discorso allora del diritto penale, dove dicevo: le intrusioni dell'intelligenza artificiale le ho definite “nel diritto penale” e le intrusioni “del diritto penale”. Cosa voglio dire anche da questo punto di vista?

Beh, intanto che ci piaccia o no, l'intelligenza artificiale è un terreno di sperimentazione, di conquista di realtà delle organizzazioni criminali. Abbiamo cominciato con le blockchain, abbiamo cominciato con i profili della moneta digitale, delle

cryptocurrencies e abbiamo ormai un mondo che è il mondo della realtà digitale che è appannaggio delle economie criminali. Ora, pensare di vietare (ammettendo che sia possibile vietare l'intelligenza artificiale), venderebbe soltanto quello che è studiato (e ormai assodato nella criminologia), il cosiddetto "crime switch", cioè semplicemente l'allocazione delle intelligenze criminali in un campo che non è ancora regolamentato, che non è ancora vietato. Quindi è inutile pensare che vietando l'intelligenza artificiale si possa vietare o si possa comprimere, o reprimere, l'economia criminale che si serve di questi strumenti. Meglio quindi regolarla. Meglio quindi prevedere che vi sia un'intelligenza artificiale che possa, anzi, contenere proprio l'uso distorto da parte delle organizzazioni criminali.

E in questo ci sono già delle regolamentazioni, a livello europeo lo sappiamo, ma a livello, anche, italiano. Mi riferisco al recentissimo disegno di legge che è stato approvato dal Ministro di Giustizia e che dovrà avere poi il suo focus, il suo riflesso nelle aule parlamentari, che sceglie di prevedere dei nuovi reati proprio legati all'intelligenza artificiale. Qual è la scelta? Adesso non vi voglio tediare dal punto di vista tecnico, ma il sistema stabilisce due pilastri fondamentali.

Il primo è un reato ad hoc, e cioè: l'uso dell'intelligenza artificiale, per esempio, quando è possibile simulare la voce, in cui potrei parlare con la voce di chiunque attraverso un software generativo. E, quindi, se questo comportamento, legato ad un uso distorto dell'intelligenza artificiale, cagiona un evento di danno (quindi abbiamo un reato di evento, un danno cagionato a qualcuno attraverso una condotta, chiamiamola decettiva, truffaldina attraverso l'intelligenza artificiale), sarà il nuovo reato. E poi si prevede una circostanza gravante per alcuni reati comuni legati proprio all'uso dell'intelligenza artificiale.

Si tratta quindi di uno scenario, come vedete, in ebollizione.
Concludo dicendo qualcosa che sarà (non è facile essere profeta) il tema di fondo degli anni a seguire, ma non di tanti anni, di domani direi, non di dopodomani (e questo è un problema sul quale noi stiamo lavorando come Cattedra di Diritto Penale, in quanto siamo un centro di eccellenza dell'Università di Salerno, e abbiamo avuto un importante finanziamento dal Ministero di Giustizia sulla cybersecurity, sugli attacchi hacker e sull'intelligenza generativa), perché il vero problema sarà domani, quando (faccio un esempio, ma non è l'unico) ci sederemo in un'auto a guida autonoma e l'auto a guida autonoma farà tutto ciò che noi oggi facciamo; oggi già facciamo di meno rispetto a dieci o vent'anni fa…
E il problema dell'evento letale che potrà accadere e che è gestito dalla intelligenza generativa, cioè da un sistema che autoapprende e, quindi, non è più controllabile dal produttore, nel caso di incidente letale chi ne risponderà? Qualcuno ha detto l'auto, cioè l'intelligenza artificiale, qualcosa che francamente mi fa sorridere perché potremmo prevedere la sanzione di staccargli la corrente, così non funzionerà più. Il problema è come distribuirlatra il produttore, colui che ha immesso sul mercato l'auto a guida autonoma (ma ci saranno robot o ci saranno sistemi di medicina che faranno ciò che fa il chirurgo oggi, e già lo fanno) tra, ripeto, il produttore e l'utilizzatore, cioè noi automobilisti.
Con chi ce la prenderemo? Ecco le intrusioni del diritto penale. Perché se il produttore sarà punito, sarà difficile sviluppare un'intelligenza artificiale di ultimissima generazione. Se l'autista, l'utilizzatore, sarà punito, è chiaro che non si servirà e preferirà continuare a guidare.
Ma qual è il problema? Perché è un falso problema…

Perché noi oggi già abbiamo, e quindi la soluzione alla quale stiamo lavorando e che io proporrò in un paper mondiale sull'intelligenza artificiale, in un convegno mondiale che abbiamo quest'anno in Asia, sarà quella di ammettere una (e può essere provocatorio detto così, ma la lancio come qualcosa sul quale stiamo lavorando) una zona di "rischio consentito". Cioè noi oggi sappiamo (è quello che vi dicono quando andiamo dal medico o anche sul famoso bugiardino) che uno su mille o uno su diecimila casi, purtroppo c'è un evento letale. Già lo sappiamo. Ebbene, se noi immaginiamo ciò e immaginiamo che l'intelligenza artificiale possa migliorare questa performance, possa abbassare ancora più il rischio della complicanza dell'evento letale, del pedone che viene travolto dall'auto, ebbene, in questo caso noi ammettiamo, attraverso un algoritmo che stiamo sviluppando (e non è il caso che vi tedi con questi profili che sono di natura matematica), attraverso il quale noi ammettiamo che il produttore abbia quella zona di rischio, di rischio consentito, che è esente da responsabilità penale. Vedremo come andrà a finire questa proposta.

Chiudo dicendo, oltre che nuovamente ringraziare, che noi abbiamo da poco creato il primo "Metacourt", cioè il primo processo penale (siamo stati forse i secondi o i terzi al mondo), interamente sul metaverso, attraverso degli avatar, e il metaverso è qualcosa che rivoluziona completamente: con le nostre "vision" abbiamo la possibilità di entrare in un nuovo mondo. Attraverso degli avatar che ci rappresentano potremmo, con l'aiuto di informatici e di coloro che fanno ciò, oggi, replicare questo nostro convegno sul metaverso, su una realtà virtuale, dove compriamo gli spazi, compriamo la nostra giacca e cravatta e parliamo.

E possiamo anche entrare in questo metaverso. Non è assolutamente un qualcosa di apocalittico, è qualcosa che

sicuramente i nostri figli manovreranno, con la quotidianità con la quale noi oggi lavoriamo, senza alcun tipo di rischio.
Vi ringrazio.

MA È DAVVERO POSSIBILE UN'IA A MISURA DI DEMOCRAZIA?

FABIO CHIUSI
Research Associate, AlgorithmWatch

ASCOLTA INTERVENTO

Grazie mille, grazie mille dottor Cerrone, grazie mille a tutti voi per lo spazio, per questo intervento.
Sì, la domanda come dice lei è fondamentale ed è una domanda che io pongo qui in maniera provocatoria: non sono un catastrofista, non sono un luddista, non sono uno scettico, sono semplicemente un sincero democratico diciamo così... Mi piace la democrazia, vorrei che vivesse i migliori suoi giorni, vorrei che i suoi migliori suoi giorni fossero ancora di là da venire e vorrei cominciare a guardare qui con voi oggi in questo breve intervento un po' se è davvero possibile, prima di tutto se ci stiamo ponendo a sufficienza questa domanda: se ci stiamo davvero chiedendo se democratizzare l'intelligenza artificiale come sta succedendo in particolare con le AI generativa, significhi veramente una migliore democrazia grazie all'intelligenza artificiale, che è una domanda non banale. E io di norma tendo a rispondere a quest'ordine di domande, cerco di dare delle risposte provvisorie naturalmente e, anche nello spirito dell'organizzazione per cui lavoro, cerchiamo di guardare alle prove, cioè cerchiamo di guardare a quello che succede. Cerchiamo di guardare i fatti e, oltre naturalmente come accennavate voi, a guardare alle policy, alle politiche che vengono sviluppate. Io ho messo in fila un po' di fatti che riguardano l'aspetto meramente politico dell'intelligenza artificiale, quindi le applicazioni politiche che possono avere un impatto sulla democrazia. Quindi non voglio criticare gli impatti dell'intelligenza

artificiale sulla medicina, sull'agricoltura, su moltissimi altri settori, diciamo, nella scienza, dove ci sono degli impatti sicuramente positivi (e possiamo solo incominciare a intravederli); ma qui vorrei concentrarmi su dei fatti che secondo me sono preoccupanti invece e su cui non abbiamo ancora delle adeguate risposte, né normative, né dal punto di vista culturale. Le norme possono arrivare fino a un certo punto naturalmente, dopodiché c'è la società, c'è la cultura e bisogna cercare di capire se questa società riesce a sopravvivere a questo tipo di sollecitazioni.

Quindi io mi sono chiesto, proviamo a chiedere a qualche elettore democratico in giro per il mondo, se l'intelligenza artificiale gli sta facendo bene o male. Proviamo a chiederlo per esempio agli elettori del New Hampshire, che questo gennaio si sono sentiti arrivare una telefonata da un presunto Joe Biden che gli diceva di non andare a votare alle primarie democratiche e, noi sappiamo, che quella voce non era quella di Biden, era una robocall gestita dall'intelligenza artificiale che replicava la voce di Biden…

E qui, diciamo, secondo me, non è una grossa applicazione a favore della democrazia come potete immaginare.

Chiedetelo agli elettori in Bangladesh, per esempio, che sono vittima della disinformazione creata da start-up che fanno servizi di AI locali e che, appunto, questa disinformazione creata da AI viene diffusa sia dalle testate pro-governative sia dagli influencer, quindi di nuovo abbiamo un ecosistema mediatico che è una delle cose che lo studio insegna all'Università di San Marino, per esempio, che è fortemente influenzato da queste forme di intelligenza artificiale, specie l’intelligenza artificiale generativa che, come veniva detto nell'intervento precedente, è particolarmente problematica da questo punto di vista.

Qualcosa di simile avviene anche in Venezuela, dove Maduro, per

esempio, è riuscito a ottenere video generati dall'intelligenza artificiale di propaganda con notizie false. In Indonesia, dove si è addirittura fatto risorgere digitalmente il generale Suharto come sorta di testimonal post-mortem di un candidato che si definisce suo erede, cosa che è avvenuta anche in India... Quindi anche in Italia è facile immaginare che prima o poi qualcuno comincerà a far risorgere Berlinguer, Berlusconi e via dicendo, perché sappiamo che poi con la storia politica italiana abbiamo un rapporto tutti complicato...

L'India è la più grande democrazia al mondo, va alle elezioni quest'anno. Le stanno già chiamando "deep fake elections", stanno parlando di "AI-based meme wars", cioè guerre di meme a base di AI, che hanno già coinvolto i principali account dei partiti del paese. Stiamo parlando appunto di personaggi politici che risorgono e stiamo parlando di un fiorente mercato per la disinformazione, per la propaganda gestita di nuovo da start-up locali che frutteranno, secondo il "Rest of World", 50 milioni di dollari solo quest'anno. Oltre 50 milioni di dollari...

Se non vi piacciono i giornalisti (io insegno anche giornalismo e intelligenza artificiale) si possono oggi rimpiazzare: quindi non soltanto le notizie, si possono creare anchorman fasulli, studi televisivi fasulli, segmenti tv fasulli che dicono semplicemente le notizie che ci piacciono e naturalmente un giornalista virtuale non ha questa scomoda intenzione, non ha la possibilità di essere scomodo, semplicemente perché dice solo quello per cui è programmato.

Volete far sembrare il vostro candidato meno nazionalista, meno razzista? Si possono creare delle foto di lui che abbraccia persone di colore, per esempio, quando non è mai successo, quando magari si vuole andare a colpire in questo bacino elettorale: è successo negli Stati Uniti con Donald Trump, per esempio. Quindi

il messaggio, qui, fondamentalmente, è che abbiamo hacker di Stato, cinesi, russi, iraniani che usano l'intelligenza artificiale per, potenzialmente, sabotare e influire sulle elezioni locali, e straniere naturalmente (quelle locali naturalmente sappiamo che sono o inesistenti o fasulle).

Quindi quello che vorrei dire qui è che siamo in presenza di una forma di democratizzazione "negativa" in questo caso. Cioè si sta fondamentalmente democratizzando la possibilità di fare sofisticate campagne di propaganda politica, di disinformazione, in maniera molto realistica e credibile, a costi molto bassi e lo può fare davvero chiunque. Voi pensate quando io ho cominciato a insegnare queste cose al corso a San Marino, all'università in cui insegno, c'era GPT2, si riusciva al massimo ad avere 3-4 righe all'incirca, dotate di senso da questi software. Oggi, soltanto pochi anni dopo, con GPT4 e con il 5 in dirittura d'arrivo, si possono fare cose estremamente più complicate. Quindi tutti i discorsi che noi abbiamo fatto sul ruolo del Cremlino, sul ruolo che aveva anche l'improvvisamente scomparso Prigožin, in Russia per esempio, nel creare forme di campagne di disinformazione, non servono più fondamentalmente. Si possono fare in casa... Chiunque può diventare autore di forme di propaganda di questo tipo, semplicemente usando degli strumenti che stanno ormai dentro i cellulari di tutti.

E questo, naturalmente, ci dice qualcosa di fondamentale rispetto non soltanto al presente, ma anche a come rileggeremo il passato. Pensate per esempio a quei neonazisti che hanno incominciato a far parlare Hitler in inglese per fargli fare i suoi discorsi nazisti, in tutte le lingue possibili e immaginabili, di modo da reclutare nuovi neonazisti. Pensate a tutte le forme di manipolazione del passato che si possono mettere in atto e, naturalmente, scrivere i libri di storia del futuro sarà perfino più

complicato di scrivere i libri di storia oggi, che è già un esercizio estremamente complicato… Immaginate gli storici del futuro. Immaginate tra 50-100 anni quando dovranno chiedere al Google di sorta, di turno, che cosa è successo in questi anni, dovranno andare a pescare nel materiale su internet che è già ibridato fortemente da contenuti sintetici, dove ci sono contenuti generati da algoritmi che sono addestrati su algoritmi, che sono addestrati su altri materiali generati da algoritmi, che parlano molto spesso ad algoritmi. Tutto questo per poi avere applicazioni che possono avere naturalmente ricadute positive, come abbiamo visto in altri campi, ma anche molto negative, non soltanto nel lato della politica, ma anche nella società. Pensate alle donne trasformate in attrici pornografiche, a loro insaputa, tramite l'intelligenza artificiale; pensate alle aziende che formano marketplace veri e propri per creare il deepfake delle persone: con pochi dollari, si può chiedere di creare un deepfake di chiunque fondamentalmente. Quindi una sorta di “Clearview on steroids" se volete, cioè una sorta di azienda che fa scraping di tutti i dati che ci sono online e crea da un'immagine, da un vostro audio, crea audio e video di voi che dite e fate delle cose che non avete mai fatto.

Quindi l'elenco potrebbe andare avanti per molto, non vorrei tediarvi, non vorrei nemmeno spaventarvi naturalmente, anche se alcuni usi dell'intelligenza artificiale spaventano quelli che abbiamo visto a Gaza, per esempio, nella guerra: è un altro settore dove l'intelligenza artificiale sta entrando prepotentemente e dove queste tecnologie stanno incominciando a prendere decisioni, a identificare bersagli, a cercare di rendere massimamente efficiente colpire un bersaglio. Sono tutte questioni che dovrebbero farci riflettere, diciamo, oltre il mero hype, oltre il mero marketing, oltre le mere questioni economiche, se volete. E io

credo che in un contesto come questo, dove si parla di intelligenza artificiale e diritti umani, sia bene sottolineare appunto quanto questa democratizzazione dell'AI, non stia in realtà facendo il bene dei diritti umani: io mi sto occupando, per AlgorithmWatch, in quest'ultimo anno e mezzo del rapporto tra intelligenza artificiale e migrazioni, per esempio, che è un tema estremamente complesso, come potete immaginare, estremamente delicato, e anche lì quello che stiamo vedendo è che molto spesso queste nuove tecnologie vengono usate per aumentare l'efficienza, per aumentare delle operazioni, per esempio, per aumentare i controlli, per militarizzare, come si dice, i confini, e molto meno per salvare vite umane o per usi umanitari. Quindi anche lì c'è tutta una serie di organizzazioni della società civile, tra cui la nostra, in coalizione, che dice appunto che, per esempio, l'AI Act, la norma europea, da questo punto di vista, è stata un'occasione mancata, perché non riesce a proteggere i diritti quanto protegge l'innovazione. Quindi qui si vede purtroppo questo dualismo che non ci piacerebbe vedere, naturalmente, ma che in questo caso si incomincia a vedere.

Quindi io non vorrei portarvi via troppo tempo. Vorrei rimandare soltanto a una parte un po' più propositiva, cioè che cosa possiamo fare di fronte a tutto questo.

Prima di tutto, secondo me, dobbiamo fermarci noi come società, come collettività, prima ancora di scrivere delle norme, prima ancora di affrontare contesti regolatori che poi restano scritti per molto tempo e possono condizionare il cammino di queste tecnologie, io non vedo questa riflessione radicale che servirebbe all'interno delle società democratiche per chiederci: l'intelligenza artificiale è una nostra alleata? O è una nostra nemica? O è un'alleata dei regimi illiberali? È un alleata di chi vuole reprimere? Stiamo costruendo le infrastrutture di sorveglianza e di controllo

intelligente per i regimi di domani o per i regimi di oggi, piuttosto che costruire la sfera pubblica, digitale, democratizzata, pluralista, tollerante, che vorremmo tutti? È un passo in avanti verso questo mondo che sembra smaterializzarsi sotto i nostri occhi ogni giorno che passa? Non lo so. Io temo che sia in questo momento, se non altro, per molti motivi, tra cui quelli che ho detto, più un'alleata dei regimi autoritari.

E quindi, da un lato, c'è questa presa di coscienza che c'è bisogno di regole, giustamente, e finalmente, diciamo, che l'era dell'autoregolamentazione è finita, che non ha portato grandi progressi (come voi dite, noi abbiamo raccolto un sacco di linee guida, ma queste linee guida continuano a moltiplicarsi: però, come potete vedere, non hanno granché deviato lo sviluppo di queste tecnologie, non l'hanno granché regolato), ma dall'altro lato deve finire l'era dell'hype…

Bisogna cominciare a pensare ai fatti, alle cose che queste tecnologie fanno davvero, non a quelle che potrebbero fare, non alle grandi promesse, non alle grandi minacce, ma ai danni o ai benefici che stanno venendo fatti. E ogni volta che noi andiamo a vedere con quest'occhio laico, diciamo, andiamo a pesare, a mettere sulla bilancia che cosa fa l'intelligenza artificiale, purtroppo la bilancia tende a pendere verso i fattori negativi, almeno nelle sue applicazioni politiche, sociali, democratiche, quando ci sono diritti e impatti sociali forti di mezzo.

Quindi una cosa che si potrebbe e si dovrebbe fare, secondo me, è prima di tutto riequilibrare la voce, diciamo, di chi porta la voce dei diritti e la voce di chi invece porta gli interessi, diciamo, affaristici; la voce di chi porta da un lato l'umanitarismo e di chi dall'altra parte porta la voce dell'efficienza. Bisognerebbe un po' riequilibrare queste due voci, dare più peso e più voce alla società civile, capire un po' meglio che cosa pensano le persone di queste

tecnologie, che usi ne fanno…
Io sono di nuovo in classe, insegno ai miei studenti fondamentalmente a fregarmene, perché l'esame che io faccio con loro di intelligenza artificiale e giornalismo loro lo potrebbero far fare tranquillamente all'AI, alle chatbot intelligenti, da ChatGPT in giù e nessuno se ne accorgerebbe… "Io non avrei modo di provare - scriveva ieri un fact checker della BBC molto stimato, molto famoso - che non ci sono vere e proprie soluzioni tecnologiche a questo problema". Cioè il problema di distinguere un contenuto sintetico da un contenuto naturale. Non c'è un "silver bullet" ancora, diciamo, non c'è una maniera chiara e, questo, non fa che complicare lo scenario.
Quindi il messaggio con cui io volevo lasciarvi non è un messaggio di disperazione, ma è un messaggio al contrario di speranza, cioè la speranza che finalmente le persone comincino a ragionare in maniera autonoma e siano interpellate, abbiamo più voce, per chiederci davvero se questo progresso sta automatizzando la caduta, diciamo, il disgregarsi delle democrazie o stia accelerando le loro difese, stia aumentando le loro difese. Se sia davvero una presunta panacea di ogni male come ci vorrebbe essere venduto dalle aziende tecnologiche che producono queste tecnologie, naturalmente, o se invece è un modo per de-responsabilizzare l'umano e dire: "Ah, ha deciso la macchina, si fa così in guerra come nei contesti civili"...
E quindi per dire che fondamentalmente serve, secondo me, una rivoluzione non tanto nell'approccio normativo, che è una cosa complicata su cui si può lavorare di dettagli, su cui è importante cogliere i dettagli, ma una rivoluzione nel modo in cui noi discutiamo, secondo me, di questo rapporto tra intelligenza artificiale e politica, che finalmente incomincia a entrare, diciamo, nella consapevolezza comune. Pensate che fino a qualche anno

fa sembrava eresia parlare dell'atto politico, morale dell'intelligenza artificiale; sembrava una questione veramente ingegneristica, tecnologica... Oggi sappiamo che non è così. Però potrebbe essere tardi, perché ce l'abbiamo già tutti nel telefono, ce l'abbiamo già tutti in ogni aspetto della nostra vita, come veniva detto prima. Quindi serve davvero uno "step up", come dicono gli inglesi, cioè bisogna incominciare a pensare davvero più a fondo e più in maniera radicale, secondo me, su questo rapporto e chiederci davvero se è un bene e, se non lo è, prendere le misure radicali che servono. Cioè se scopriamo che le applicazioni politiche dell'intelligenza artificiale sono, in genere, negative, fanno il gioco dei Putin di turno, per capirci, sarebbe bene cominciare a trovare delle forme molto radicali di regolamentazione. Per esempio, l'Unione Europea ha ammesso, ed è una cosa buona, secondo me, e inedita, che alcune forme di intelligenza artificiale non hanno il diritto di cittadinanza nel gioco democratico: le forme che vengono messe al bando, diciamo, dell'intelligenza artificiale, nell'AI Act, pur con tutte le eccezioni e con tutti i problemi che questo comporta, naturalmente. Questo secondo me però è un'idea fondamentale, che ci porta lontano dall'idea che la società possa solo adattarsi al progresso tecnologico e non possa modificarlo, non possa dirigerlo secondo i canoni della democrazia e dei diritti umani. Io credo invece che sia possibile, e credo che si debba, farlo.
Grazie.

IL REGOLAMENTO SULL'IA PUÒ ESSERE UN NUOVO STRUMENTO PER I DIRITTI FONDAMENTALI?

ASCOLTA INTERVENTO

MARIELLA FIORENTINO
Vice Presidente della Commissione Diritti Umani, Union des Avocats Européens
Avvocato

Buongiorno a tutti, anche io vorrei, prima di parlare di questo argomento, ringraziare tutti quanti voi per questa opportunità (anche per la mia associazione) di essere presente in un convegno così scientificamente elevato.
Quindi, interessantissime le due precedenti relazioni, insomma, molto accademiche e io dirigerei più l'attenzione su: è possibile creare nuovi diritti umani, visto che questo nuovo mondo ci impone poi di fare una riflessione anche in considerazione del fatto che i diritti umani cambiano con il cambiare della società?
Quindi partirei anche da una distinzione tra diritti umani e diritti fondamentali, perché molto spesso sono termini che vengono anche utilizzati in maniera similare e, per certi versi lo sono, per altri no. Anche nel regolamento troviamo una distinzione tra la prima parte del preambolo, dove c'è un obiettivo molto più ampio, quindi che fa riferimento ai diritti umani, che dopo citerò, e nel prosieguo si fa riferimento ai diritti fondamentali. Quindi, delineare una distinzione tra diritti umani e diritti fondamentali è, secondo me, giusto e opportuno, perché in questo modo è più facile chiarire anche i limiti e le prospettive di questo regolamento.
Nel senso, noi possiamo regolamentare l'uso, l'applicazione e anche la gestione della tecnologia sottostante all'intelligenza artificiale, attraverso il rispetto dei diritti fondamentali, però, non sarà sempre poi possibile bilanciare questo rispetto con i diritti

umani. Ed è il motivo per cui, secondo alcuni, sarebbe necessario proprio la creazione di nuovi diritti umani e anche di nuovi diritti digitali.
Quindi, cominciamo col dire che i diritti umani sono quelli di cui godiamo tutti per il semplice fatto di esistere, questo in linea di massima. I diritti fondamentali, invece, sono quelli che permettono alle persone di sviluppare il loro pieno potenziale, la loro capacità mentale, psicologica, intellettuale e fisica. Quindi, affinché esista un diritto fondamentale, deve prima esistere un diritto umano.
E, soprattutto, non tutti i diritti umani sono riconosciuti come diritti fondamentali.
E questo lo vedremo in prosieguo e si riallaccia moltissimo a quello che è stato detto dal professore Chiusi per quanto riguarda l'intelligenza artificiale alleata dei regimi totalitari, perché non in tutte le parti del mondo i diritti umani vengono tutelati allo stesso modo e non in tutte le parti del mondo i diritti fondamentali vengono, non solo tutelati, ma anche previsti normativamente allo stesso modo.
Quindi, le similitudini tra diritti umani e diritti fondamentali le possiamo trovare, per esempio, in alcuni articoli di uguale contenuto, avvicinandoci in Europa, tra la Convenzione europea dei diritti dell'uomo e la Carta dei diritti fondamentali dell'Unione europea.
Quindi, queste sono le basi normative, le basi giuridiche, sia per i diritti umani che per i diritti fondamentali.
Salvo poi, chiaramente, considerare anche le Costituzioni nazionali.
Quindi, i presidi di protezione e di salvaguardia poi, a livello giudiziario, sono rispettivamente la Corte Costituzionale, la Corte europea dei diritti dell'uomo e la Corte di Giustizia dell'Unione

europea.
Ora, bisognerà poi capire se sono sufficienti queste “protezioni” ad affrontare una rivoluzione digitale in atto come quella che comporta l'intelligenza artificiale. In linea di massima, quello che possiamo dire è che il perimetro entro cui si muove una giusta tutela dei diritti umani nei confronti di ogni genere di tecnologia, anche quella dell'intelligenza artificiale, per quanto velocemente possa andare, sono dei principi cardine come per esempio il rispetto della dignità umana, che è il presupposto di tutti gli altri diritti fondamentali tutelati dalla Carta dei diritti fondamentali dell'Unione Europea, la Carta di Nizza, e ovviamente anche dalla Corte Europea dei diritti dell'uomo. Il principio di rispetto della dignità umana che poi ha evidenti implicazioni anche per la libertà individuale, la vita privata e familiare.
A seguire, ma chiaramente è strettamente correlato al precedente, è anche il rispetto del principio di non discriminazione, così come il rispetto del diritto alla sicurezza e dei principi di trasparenza, imparzialità ed equità; questi ultimi vengono in rilievo poi soprattutto per quanto riguarda la garanzia del diritto ad un giusto processo e la trasparenza dei processi decisionali. Quindi siamo in ambito giudiziario dove poi si gioca la vera partita sul rispetto dei diritti fondamentali e dei diritti umani. Che noi possiamo parlare in linea di massima, in tutti gli ambiti, di diritti umani e diritti fondamentali, però poi la vera sfida è quella di capire fino a che punto possiamo tutelarli e proteggerli in un'aula giudiziaria, sia essa vicina a noi come anche delle corti europee o internazionali. Quindi, l'eventuale lesione di questi diritti, è ovvio che poi comporta, anche come effetto a cascata, la lesione di una serie di ulteriori altri diritti, che sono il lavoro, la salute, la privacy e così via.
Mi soffermerei più che altro sull'analizzare l'impatto

dell'intelligenza artificiale sui diritti collettivi, come è stato precedentemente già introdotto dal professore Chiusi, perché l'uso dell'intelligenza artificiale, soprattutto nel settore degli armamenti, comporta il serio rischio di minaccia del diritto alla vita e l'integrità delle persone, e non di una singola persona ma di molte persone.
Quindi un esempio possiamo fare è quello dei killer robots, che sono macchine alle quali viene dato un obiettivo specifico, ma in mancanza di una sensibilità umana o di un'intelligenza emotiva che possa dirigere, e quindi anche far fronte a degli imprevisti, chiaramente può comportare anche la deviazione di un conflitto con esiti drammatici a carico delle popolazioni civili.
Dopo questa panoramica un po' generale che però ci serve per individuare quali sono i principi fondamentali e considerati generali, proprio il perimetro a mio avviso dell'intelligenza artificiale.
Il panel richiede un'analisi tra innovazione e tradizione.
Beh, questo regolamento sull'intelligenza artificiale in realtà nel contenuto non è proprio nuovissimo, nel senso che le istituzioni europee si sono già cimentate in questo argomento.
Ad esempio, fin dal 2008, la Commissione europea ha istituito un gruppo, High Level Expert Group on Artificial Intelligence, di 52 esperti indipendenti provenienti dal mondo accademico, dalla società civile e dall'industria e, successivamente, già nel 2018, sulla base di questi studi di questo gruppo di esperti, la Commissione europea ha poi pubblicato, ad esempio, la strategia sull'intelligenza artificiale, ha varato il piano coordinato sull'intelligenza artificiale e, nel 2019, ha pubblicato gli orientamenti etici per un'intelligenza artificiale affidabile.
Su quest'ultima pubblicazione mi soffermerei un attimo, perché la comunicazione di questa pubblicazione è creare fiducia

nell'intelligenza artificiale "antropocentrica".
Quindi, già nel 2019, si intravede un obiettivo da dare all'intelligenza artificiale che ponga al centro l'uomo e, quindi, sia come rispetto dei diritti umani che come rispetto dei diritti fondamentali. E soprattutto, in questa pubblicazione viene fornita una sorta di lista di controllo per la valutazione dell'affidabilità dell'intelligenza artificiale, dove vengono individuati dei requisiti, sette in particolare, che devono essere monitorati e soddisfatti costantemente perché l'intelligenza artificiale rimanga in un perimetro antropocentrico.
E questi sono… Interessante pensare che il primo è: "Intervento e sorveglianza umani" (direi che su questo la riflessione la possiamo fare tutti), "Robustezza tecnica e sicurezza", "Riservatezza e governance dei dati", "Trasparenza", "Diversità, non discriminazione ed equità" (“non discriminazione ed equità abbiamo” visto essere dei principi fondamentali), "Benessere sociale, ambientale" e "Accountability".
Contemporaneamente, anche il Consiglio d'Europa, nel 2017, aveva istituito un comitato di esperti sull'intelligenza artificiale che aveva prodotto lo studio "Algorithms and Human Rights" (e quindi già nel 2017 parliamo di diritti umani e di intelligenza artificiale). Questo comitato di esperti nel 2018 pubblica la Carta etica europea sull'impiego dell'intelligenza artificiale prettamente nell'ambito dei sistemi giudiziari e in ambiti connessi. Perché, torniamo al discorso precedente: è poi nelle aule di giustizia che si capisce se un diritto (nella specie, umano e fondamentale) riceve effettivamente una tutela oppure no. E quindi, secondo questa carta, nell'ambito dei sistemi giudiziari e in ambiti connessi, l'approccio da adoperare è quello "ethical by design" e "human rights by design".
Arriviamo a febbraio 2020. La Commissione Europea,

raccogliendo tutti gli studi precedenti del gruppo dei 52 esperti, pubblica il Libro Bianco sull'intelligenza artificiale.
Signori, siamo a febbraio 2020, una data precedente a quella che tutti sappiamo.
Fino a questo momento, sì… Diritti umani, intelligenza artificiale, creare fiducia nell'intelligenza artificiale antropocentrica, ma "creare fiducia"... Non siamo ad un obiettivo preciso di antropocentrismo delle normative sull'intelligenza artificiale. Quindi, tutti questi sforzi della Commissione Europea e del Consiglio d'Europa, in realtà, hanno come obiettivo quello di mantenere alta la competitività e anche di colmare un gap tecnologico tra l'Unione Europea e paesi molto più avanzati tecnologicamente come la Cina e gli Stati Uniti. Quindi, diciamo che l'obiettivo è molto economico ed è molto di competitività tecnologica.
Arriviamo poi al periodo pandemico.
Il periodo pandemico, lo sappiamo tutti, ci catapulta improvvisamente da, quello che io chiamo, un mondo di tipo sociale, ad un mondo di tipo digitale. Cambia il mondo e quindi cosa si fa? Cambiano anche i diritti? E la persona come viene vista? Esistono ancora diritti fondamentali e diritti umani come prima?
Cominciamo ad avere l'intelligenza o comunque un'accelerazione tecnologica improvvisa, non solo nelle nostre vite private, ma anche, per esempio, nel lavoro.
Lo smart working, penso che l'abbiano fatto tutti…
Nuove tecnologie prendono comunque il posto degli esseri umani anche nelle pubbliche amministrazioni.
Quindi, all'inizio del 2022, la Commissione Europea propone la Dichiarazione europea sui diritti e principi digitali, per il decennio digitale. Qui passiamo da un aspetto prettamente economico e

competitivo, per colmare un gap tecnologico in un panorama mondiale. La Commissione Europea, con questa dichiarazione, rafforza la dimensione umana dell'ecosistema digitale e si propone di renderla funzionale anche alla lotta ai cambiamenti climatici e alla protezione dell'ambiente, che erano proprio le tematiche, poi, che sono emerse anche durante la pandemia.
Si passa quindi ad una visione maggiormente umanocentrica del panorama digitale.
Ed è questa visione che poi viene recepita anche nel regolamento di cui ci occupiamo. E con questo regolamento in cui (tant'è che nel preambolo, proprio nei primissimi punti, nel punto 6) si dice che l'intelligenza artificiale "dovrebbe essere - faccio notare che dice "dovrebbe essere" - una tecnologia antropocentrica. Dovrebbe fungere da strumento per le persone con il fine ultimo di migliorare il benessere degli esseri umani". Qui, vediamo, che ci si rifà ad una nozione che è quella più ampia, che è quella dei diritti umani, molto simile alla definizione dei diritti umani.
Sottolineo però che in entrambi i punti il regolamento usa un condizionale. Io su questa una riflessione la farei e la girerei… Il legislatore non dice che l'intelligenza artificiale deve fungere da strumento, perché questo comporterebbe già un obbligo. Chi pratica le aule di tribunale sa benissimo che, soprattutto chi è nel campo amministrativo, come me (e io sono un avvocato amministrativista), uno dei capisaldi per individuare una violazione è quella di dire "la norma pone un obbligo". E quindi nella pubblica amministrazione (o comunque vedremo poi anche in orizzontale, quindi tra gli operatori privati) se la norma pone un obbligo, significa che non rispettando quell'obbligo c'è automaticamente una violazione.
Ma dire "dovrebbe", quando il legislatore usa il termine "dovrebbe", è chiaro che qua si apre uno scenario. Quali sono i limiti del

"dovrebbe"?
Stiamo parlando di diritti umani però, diritti umani e diritti fondamentali.
Ad ogni modo, il regolamento ritiene che comunque in un bilanciamento di interessi e di diritti, il rispetto dei diritti umani e dei diritti fondamentali debba prevalere rispetto alla competitività e agli investimenti. E questo lo si evincerebbe, o lo si evince, dal fatto che, rispetto, per esempio, alla precedente normativa più o meno similare che il GDPR, il regolamento utilizza un approccio basato sul rischio.
"Rischio", anche di questo si è parlato precedentemente da parte dei professori.
Ci sono quattro tipi di rischio, non mi addentrerò molto perché leggo che gli altri affronteranno più nel dettaglio la cosa, ma giusto per (questo è il primo panel) darvi un'idea di massima… Esistono delle attività che vengono considerate a rischio inaccettabile. “Rischio” significa: la cui applicazione incide talmente tanto sul diritto umano e il diritto fondamentale, che devono essere ritenute inaccettabili. Come rischio inaccettabile vengono considerate le tecniche subliminali, quindi quelle adoperate in alcune pubblicità che, come è stato detto precedentemente, possono influire anche sull'orientamento politico, anche sull'orientamento economico, nel caso in cui vengano adoperate da parte di società private. Il social scoring, i sistemi che sfruttano le vulnerabilità delle persone, come i bambini e gli anziani, e soprattutto l'uso del sistema di identificazione biometrica in tempo reale negli spazi pubblici. Anche di questo si parlerà dopo… Io in questa sede vi dico solo che questa attività, per quanto sia di rischio inaccettabile, però, può essere, autorizzata dagli Stati. È chiaro quindi, come già è stato detto, che il professor Chiusi dice: “l'intelligenza artificiale è

a servizio e alleata dei regimi totalitari". Per forza! Perché nel momento in cui un'attività considerata a rischio inaccettabile, viene però autorizzata da uno Stato, è ovvio che in questo caso l'intelligenza artificiale può trovarsi a rischio di regimi non proprio democratici e, quindi, lascio a voi poi le considerazioni… Ci sono attività con rischio elevato, che sono quelle altamente regolamentate e, ripeto, sono sostanzialmente quelle che riguardano l'amministrazione della giustizia, cioè prendere delle decisioni giudiziarie, in mancanza di un'intelligenza emotiva, in mancanza dell'apporto dell'essere umano con la propria esperienza e il proprio studio, sicuramente comporta delle decisioni che non possono rispondere, possono essere arbitrarie.
Poi ci sono attività con rischio limitato, dove normalmente viene semplicemente imposto un obbligo di trasparenza e attività a rischio basso, dove normalmente non ci sono obblighi legali, ma viene richiesto rispetto di codici di condotta o comunque di specifiche tecniche. Questo è per quanto riguarda la tradizione.
L'innovazione.
Il regolamento, sicuramente, abbiamo visto, che nel contenuto non è nuovissimo, cioè noi lo salutiamo come una cosa nuova, ma signori, dal 2008 si parla di intelligenza artificiale e soprattutto intelligenza artificiale e diritti umani, benché i primi studi e le prime pubblicazioni di questo gruppo di 52 esperti della Commissione europea, viene pubblicato dopo dieci anni. E anche su questo una riflessione la farei…
Ma torniamo al regolamento. Il regolamento non è nuovissimo nel contenuto, ma sicuramente negli obiettivi delle novità ce le ha, e anche per quanto riguarda l'applicazione. Negli obiettivi abbiamo visto che viene proposta una intelligenza artificiale che dovrebbe essere una tecnologia antropocentrica. Viene considerato, come parametro di correttezza dell'uso e dell'applicazione

dell'intelligenza artificiale, il rispetto dei diritti fondamentali.
Altra novità è sicuramente riguardo all'applicazione, perché, mentre la Carta dei diritti fondamentali dell'Unione europea prevede delle disposizioni che non hanno un' applicazione orizzontale, fatta eccezione per alcuni articoli come l'articolo 21 che pone il divieto di discriminazione e l'articolo 47 che pone il diritto ad un ricorso effettivo a un giudice imparziale.
La novità di questo regolamento è che ha una prospettiva di applicazione di diritti fondamentali in orizzontale, cioè tra privati, se la legislazione secondaria lo consente. Ora, questo poi che cosa comporta?
Torniamo sempre alle aule del tribunale. Io ho detto che le norme vengono testate nei tribunali. In linea di massima, possono essere tutte corrette, ma soprattutto poi per le norme che riguardano processi scientifici, tecnologici, così veloci, sono le aule dei tribunali quelle dove viene testata la concreta applicazione e, poi, dove se ne capisce anche il "vulnus".
E i vulnus quali sono concretamente?
Allora, il regolamento sull'intelligenza artificiale non fa né richiamo, né rinvio e né indica degli articoli specifici della Carta dei diritti fondamentali dell'Unione Europea.
Quindi che cosa significa? Che, per esempio, un operatore privato che, ai sensi dell'articolo 29 bis del regolamento, ha l'obbligo, nei sistemi ad alto rischio, di condurre una valutazione di impatto sui diritti fondamentali, come fa l'operatore privato a fare questa valutazione d'impatto, se non sono chiari quali sono precisamente gli articoli della Carta dei diritti fondamentali violati?
Altra prospettiva: l'individuo, che si trova per esempio a dover eseguire un provvedimento amministrativo (o anche un contenzioso privato), come fa a far valere, davanti ad un giudice, un diritto fondamentale violato e correlato al regolamento? Cioè,

basta semplicemente dire che è stata violata una norma del regolamento, per dire automaticamente che è stato violato anche un diritto fondamentale? Oppure (io sono sempre un avvocato amministrativista) va impugnato, per esempio, il provvedimento amministrativo, oppure viene indicata la norma del regolamento violata in correlazione con uno degli articoli della Carta dei diritti fondamentali (tipo per esempio l'articolo 21, la "non discriminazione")?

È questo che poi comporterà, nel prosieguo, e quindi da questo momento in poi, capire, la tenuta di questo regolamento nella pratica. Cioè è l'applicazione delle aule giudiziarie e, verosimilmente, gli aggiustamenti o comunque la tenuta di questo regolamento, subirà poi la prova del nove che è quella della Corte di Giustizia dell'Unione Europea, con una serie di rinvii pregiudiziali che presumibilmente ci saranno.

C'è da dire però che la Corte di Giustizia, già precedentemente, un minimo delle indicazioni le ha date perché dice che "qualsiasi compressione all'esercizio dei diritti e delle libertà riconosciuti dalla Carta di Nizza, deve risultare rispettosa della loro essenza". Torniamo sempre ai principi generali. Più o meno lo stesso principio l'ha fornito anche la Corte europea dei diritti dell'uomo che, nel 2008, nella sentenza “S. e Marper contro Regno Unito” chiarisce che "gli Stati dovrebbero trovare un giusto equilibrio tra la protezione dei diritti fondamentali e lo sviluppo delle nuove tecnologie". Io trovo questo chiarimento come una sorta di faro per tutte le giurisdizioni, anche quelle nazionali che saranno le prime a trovarsi ad affrontare questa problematica.

E poi è interessante che, nel 2017 (vi invito a fare la riflessione che stiamo parlando di pronunce della Corte, atti delle istituzioni europee, tutte precedenti di molti anni), c'è un rapporto del Rathenau Institute, che analizza l'impatto dei diritti umani nell'era

dei robot e propone il riconoscimento di due nuovi diritti umani per mantenere l'intelligenza artificiale comunque su di un piano che possa essere effettivamente a misura di uomo e di diritti umani. E, in particolare, secondo questo rapporto bisognerebbe creare il diritto a non essere oggetto di misurazioni, di analisi e di addestramento e il diritto a un contatto umano significativo, quindi il diritto a poter stabilire e sviluppare relazioni profonde con altri esseri umani. Qua, diciamo, ritorniamo un attimo anche alla pandemia, benché successiva a questo rapporto.

Quindi, innanzitutto nulla di nuovo, ma poi si tratta anche di diritti che sostanzialmente potrebbero anche essere ricondotti al concetto di "right to cognitive liberty", cioè definita un diritto ad una libertà, così detta, cognitiva. E comunque, stiamo parlando di uno sviluppo naturale dell'era tecnologica, comunque, di diritti che sono già nella nostra tradizione, come il diritto alla privacy e al rispetto alla vita familiare.

Creare nuovi diritti umani, nuovi diritti digitali, in realtà, sono sempre tutti quanti riconducibili a diritti fondamentali già oggetto di normative. Quindi mi chiedo anche, rispetto all'intervento del professor Castaldo, per esempio, in un metaverso quali diritti umani e quali diritti fondamentali possono essere individuati e tutelati? Visto che parliamo di avatar, sostanzialmente, non si parla di persone, ma, in un metaverso, in un'aula giudiziaria, potrebbero essere utilizzati i diritti umani appartenenti al nostro mondo? Questo ce lo dirà il tempo.

Sicuramente quello che è necessario (questo per ogni normativa) è che ci siano meccanismi di controllo rigorosi, effettivi poi. E come renderli effettivi?

Magari associandoli a meccanismi sanzionatori talmente gravosi in termini economici, che possano poi scoraggiare anche gli impieghi non etici di un'intelligenza artificiale.

In questo, il regolamento sicuramente ci aiuta, perché rispetto per esempio alla normativa del GDPR, pone delle sanzioni economiche abbastanza gravose che arrivano fino a 30 milioni di euro e, addirittura il 6% del fatturato globale annuo, perché, ricordiamo, che questa normativa è applicabile anche agli operatori privati, rispetto alle altre normative che riguardano i diritti umani e i diritti fondamentali.

Creare nuovi diritti umani.

Esistono già dei principi fondamentali che individuano quali sono i diritti fondamentali da rispettare, e questo non solo per l'intelligenza artificiale di questo regolamento, ma anche per, a mio avviso, gli sviluppi che avranno le nuove tecnologie. Oggi si parla di intelligenza artificiale, non sappiamo tra dieci anni di quale altra tecnologia parleremo. E, su questo, la riflessione è che comunque il progresso tecnologico, il progresso scientifico, soprattutto oggi, in questo processo, si inserisce l'elemento di un intelligenza artificiale che sicuramente farà andare ad una velocità tripla, quadrupla (se non di più), il progresso tecnologico-scientifico. E questo non potrà mai andare di pari passo al tempismo del legislatore. È impossibile pensarlo.

Quindi, sempre a mio avviso, come può essere arginato o comunque perimetrato l'uso corretto di una nuova tecnologia (e non voglio neanche parlare di intelligenza artificiale) proprio per il futuro?

A mio avviso, sempre facendo riferimento ai principi generali, cioè la dignità umana, il principio di non discriminazione e quelli detti prima, che dovrebbero costituire una sorta di binario sul quale il progresso scientifico, il progresso tecnologico, le nazioni, gli individui dovrebbero avanzare per non far deragliare questo treno.

Da un punto di vista ancora più personale, concludo, siamo in questa sala bellissima dove c'è il Cenacolo... Io sono credente.

Dicevo prima che Dio, o chiunque creda in qualcos'altro, ha creato una macchina perfetta. Gli esseri umani sono imperfetti, quindi quando gli esseri umani si sostituiscono a Dio possono solo creare cose imperfette. E in questo l'uso dell'intelligenza artificiale penso che stia già mostrando tutte le imperfezioni.
Io concludo qui e vi ringrazio.

AI ACT EUROPEO: INNOVAZIONE E VALORI DEMOCRATICI

ASCOLTA INTERVENTO

SANDRO GOZI
Deputato europeo, Membro della Commissione per il mercato interno e la protezione dei consumatori

Buongiorno, mi dispiace non essere presente fisicamente con voi, ho seguito parte del dibattito e mi sembra un po' di aver capito quali sono stati i termini del dibattito, che sono quelli che sono stati poi al centro dei nostri lavori anche a livello di Parlamento europeo.
Io vorrei darvi qualche messaggio sperando di poter contribuire in qualche modo al dibattito e non di ripetere troppo le cose che sono state dette. Il primo punto è che secondo me l'intelligenza artificiale è la più grande opportunità dell'umanità in questa fase storica; quindi diciamo che il mio approccio all'intelligenza artificiale applicata alla democrazia e anche applicata agli aspetti dei diritti umani di cui voi avete parlato è positiva ed è la ragione per cui noi abbiamo deciso di renderla, di prendere le misure necessarie perché questo potenziale positivo dell'intelligenza artificiale diventi, reale.
Lo avete detto anche nell'ultimo scambio tra Castaldo e Fiorentino mi sembra, la tecnologia è neutra. Però bisogna appunto, essendo neutra, renderla compatibile con i nostri valori e con i nostri obiettivi. Noi per la prima volta al mondo ci abbiamo provato. Tra due o tre anni vedremo quali sono i risultati del nostro insieme di leggi europee sull'intelligenza artificiale e sul digitale applicate alla democrazia, perché certamente c'è l'AI Act, la legge fondamentale, la legge quadro sull'intelligenza artificiale, ma c'è anche la legge sui servizi numerici, il Digital Services Act, c'è anche la legge sul

political advertising, sulla pubblicità politica online e offline; ma questa mattina ci interessa online, di cui sono stato relatore, oltre che appunto aver contribuito anche a tutto il resto della legislazione.
Quindi diciamo che noi abbiamo per la prima volta, a livello europeo e a livello continentale, introdotto delle norme, per fare cosa? Innanzitutto, per prendere atto che, certamente, rispetto alla democrazia e rispetto ai diritti fondamentali ci sono dei rischi con l'intelligenza artificiale e questi rischi vanno ridotti al massimo o eliminati. Esempio molto banale, ma molto vero: noi non siamo assolutamente d'accordo con il modello cinese. Noi il social scoring non lo vogliamo: la valutazione sociale, la valutazione attraverso l'intelligenza artificiale, se siamo un buon o un cattivo cittadino, se non passiamo mai quando siamo dei pedoni con l'arancione o con il rosso sulle strisce pedonali, non lo vogliamo. Quindi, certamente, abbiamo vietato, coerentemente con il nostro approccio, la valutazione sociale. Non vogliamo la sorveglianza di massa e non vogliamo utilizzare l'intelligenza artificiale per sorvegliare in maniera indiscriminata qualsiasi persona che partecipi a un evento pubblico.
Faccio solo alcuni esempi, che sono un po' collegati al tema "democrazia e diritti umani" che avete affrontato. Queste sono delle attività che possono comportare dei rischi, ma possono portare anche degli enormi vantaggi, in materia ambientale, in materia di salute... L'intelligenza artificiale può portare dei vantaggi enormi e quindi andava regolata e abbiamo stabilito degli obblighi specifici di moderazione, di trasparenza per quanto riguarda gli operatori dell'intelligenza artificiale quando sono applicati a delle attività legate a dei beni pubblici che possono essere appunto la salute, l'ambiente, eccetera…
Qual è il filo rosso del nostro approccio?

Il filo rosso del nostro approccio è innanzitutto la trasparenza.
Abbiamo introdotto degli obblighi di trasparenza molto importanti per tutti coloro che agiscono nello spazio pubblico utilizzando l'intelligenza artificiale, utilizzando gli strumenti numerici. Norma ovvia, ma necessaria, perché non c'era in tanti Stati membri (non so, sinceramente, se ci fosse in Italia oppure no, ma voi siete più esperti di me, non sono di Italia). Tutto quello che è promozione di qualsiasi attività volta a determinare il consenso, che sia un post, che sia una partecipazione a una campagna elettorale, eccetera, fatto con l'intelligenza artificiale, dovrà avere una label, un'etichetta, chiaramente: "Questo è un prodotto dell'intelligenza artificiale" e, quindi, da questo punto di vista, qualsiasi persona, chiunque di noi si trova online e riceve dei contenuti prodotti con l'intelligenza artificiale, deve sapere in maniera molto chiara che non è un contenuto reale, è un contenuto dell'intelligenza artificiale. Questa è una delle lotte contro il deep fake, contro le fake news, eccetera, che è già ora in vigore.
Ci sono degli obblighi di moderazione per le piattaforme, nel momento in cui delle attività che non sono illegali, ma che possono comportare un rischio sistemico ai diritti umani, esplicitamente, sembra che sia l'articolo 27 del Digital Services Act, alla sicurezza, in questo caso, le piattaforme devono prendere delle misure di moderazione specifiche e devono fare un rapporto, un report, ogni sei mesi, su tutto quello che hanno fatto, segnalato dall'autorità pubblica, segnalato dai singoli cittadini, segnalato dai cosiddetti "transit flaggers", tutto quello che hanno fatto per ridurre (e un po' per eliminare) i rischi sistemici potenziali, o reali, ai diritti umani o alla sicurezza. Abbiamo introdotto degli obblighi di intervento dal punto di vista della cybersecurity, estendendo i settori che sono considerati critici dal punto di vista anche della tenuta democratica. Il settore

farmaceutico, non pensavamo qualche anno fa che potesse essere considerato, almeno alcuni di noi non pensavano, che potesse essere considerato critico dal punto di vista della cybersecurity. Oggi, è uno dei settori su cui abbiamo introdotto degli obblighi di estendere misure specifiche contro la cybersecurity, perché abbiamo visto che è uno dei settori che possono mettere a rischio anche la tenuta di un sistema dal punto di vista democratico.

Queste sono le misure che abbiamo introdotto dal punto di vista dell'utilizzo trasparente dell'intelligenza artificiale, introducendo degli obblighi trasparenti agli algoritmi.

Apriamo la scatola nera. Apriamo la "black box".

Fino ad oggi, fino alla legge sui servizi digitali e fino alla legge sull'intelligenza artificiale a livello europeo, non avevamo nessuno strumento giuridico legale per obbligare le piattaforme a dirci come funzionano gli algoritmi. E noi sappiamo benissimo che il modello di business finora utilizzato prima di questi interventi legislativi (poi non dico che sarà tutto positivo, vedremo tra tre anni quello che è funzionato e quello sarà da modificare), le piattaforme, in particolare le grandi piattaforme digitali, non ci consentivano di entrare negli algoritmi e quindi di capire come funzionavano. Ma, dal punto di vista del dibattito democratico, lo sappiamo benissimo come hanno funzionato. Hanno funzionato rendendo virali tutti i contenuti più violenti, tutti i contenuti più falsi, tutti i contenuti legati alla disinformazione, perché il modello e gli algoritmi che sono stati utilizzati aumentano la viralità di questi contenuti di tipo violento.

È su questo che noi vogliamo, con questa legislazione, intervenire, perché la libertà di opinione di ognuno di noi non è diritto alla viralità.

Ho lavorato moltissimo in questi anni, sia al Parlamento europeo

che con i colleghi del congresso americano: "freedom of speech is not freedom of preach"... E, su questo, c'è stata una battaglia molto dura, con le grandi piattaforme, con le big tech, perché ovviamente il loro modello di business, che vuol dire rendere virale e raccogliere il massimo di dati per poi utilizzarli a livello commerciale, è all'opposto di quelli che sono i nostri obiettivi, e anche i nostri bisogni, di utilizzare in maniera positiva l'intelligenza artificiale per la vita democratica.

Potrei parlare mezz'ora, ma mi avvio alle conclusioni, non preoccupatevi.

Un altro aspetto che fa parte della difesa delle nostre democrazie e sui quali, come europei, dobbiamo interrogarci, dobbiamo intervenire è: quali sono gli elementi e i documenti, diciamo così, che allenano l'algoritmo?

Perché voi parlavate di ChatGPT (mi sembra che forse il professor Castaldo avesse fatto riferimento un paio di volte a ChatGPT)...

ChatGPT si allena e diventa più "les performances" di qualsiasi altro metodo di linguaggio di questo genere, in cui gli algoritmi diventano più efficaci con l'aumentare dei documenti degli scritti, dei libri, degli articoli con cui vengono allenati.

Chi alimenta gli algoritmi?

Se, a livello globale, gli algoritmi di questi modelli di linguaggio li alimentano solo i cinesi e gli americani, è evidente che le risposte che ci saranno, saranno risposte molto influenzate dalla Cina o dagli Stati Uniti. Cioè, se per dieci anni, gli algoritmi si allenano ad essere alimentati solo da Pechino, sulla questione di Taiwan, ovviamente, quando noi chiederemo a ChatGPT: "Parlami di Taiwan", la risposta sarà una risposta cinese, non sarà una risposta equilibrata. Stessa cosa: se per dieci anni lasciamo il Pentagono (e queste cose stanno già accadendo ovviamente, e lo sappiamo) a parlare della difesa, ovviamente, ChatGPT darà una

risposta sulla difesa della sicurezza in Europa molto americana, non necessariamente europea. Quindi, anche questo aspetto, che non è un aspetto legislativo, ma è un aspetto che fa parte di una nuova politica di sicurezza e difesa delle nostre democrazie, a livello europeo, su questo noi dobbiamo intervenire ed è un lavoro che è legato a queste misure di protezione della cybersecurity, di lotta contro la disinformazione, sulle quali dobbiamo certamente lavorare a livello nazionale, ma sulle quali dobbiamo anche aumentare la cooperazione a livello europeo. Abbiamo cominciato a farlo e, secondo me, dobbiamo dare una bella accelerata nella nuova nella nuova legislatura.

Ecco, spero di aver contributo in qualche modo al vostro dibattito e ringrazio molto Marco Cerrone e, attraverso di lui, Maurizio Turco del Partito Radicale per avermi invitato a contribuire.

ALGORITMI E FORMAZIONE DEL CONSENSO: UNA PROSPETTIVA DI DIRITTO COSTITUZIONALE

ASCOLTA INTERVENTO

FEDERICA FABRIZZI
Professoressa Associata di Istituzioni di Diritto Pubblico, Sapienza Università di Roma

Grazie, grazie a Marco Cerrone, grazie alla Fondazione Marco Pannella per questo invito.
Allora, sì, io intervengo in questa sessione dedicata all'intelligenza artificiale e democrazia digitale provando a portare una prospettiva di diritto costituzionale rispetto a questi temi.
Però vorrei partire, cominciare, se me lo permettete, con un riferimento molto leggero, nel senso che qualche giorno fa, non ricordo in che piattaforma, ma un collega scriveva un post che più o meno diceva: "Ogni mattina un giurista si sveglia e diventa esperto di intelligenza artificiale"...
Ecco, devo dire che mi ha fatto molto ridere e l'ho trovato anche molto vero, quindi...
Ecco, di intelligenza artificiale si parla ormai veramente ovunque.
Il tema non è più certamente relegato soltanto ai consessi scientifici, delle cosiddette scienze dure, ma è diventato di dominio pubblico, quanto poi con consapevolezza di quello che si dica, questo, diciamo, è un altro problema.
Ed è entrato però prepotentemente anche appunto nelle riflessioni dei giuristi.
Dal momento in cui il diritto ha scoperto, ha aperto il grande vaso di Pandora, del trattamento dei dati, con tutte le problematiche connesse al trattamento dei dati, dalla raccolta all'utilizzo, il

consenso, la conservazione, l'analisi di quei dati, ecco, le diverse discipline giuridiche si sono dovute rapportare con questa nuova dimensione dell'agire umano, incentrata sulle informazioni che sono prodotte da ciascuno di noi.
Naturalmente, in questo fiorire di riflessioni giuridiche, l'attivismo dell'Unione Europea, che in un torno di anni relativamente breve ha prodotto una serie di atti, a partire dal GDPR, ma penso anche al Digital Services Act, al Digital Markets Act e poi ora all'AI Act, chiaramente ha spinto ulteriormente, appunto, i giuristi ad interrogarsi e a occuparsi di intelligenza artificiale.
Occorre dire, con secondo me anche grande franchezza, che non tutti ci siamo arrivati, diciamo, che questa consapevolezza non è maturata per tutti subito e nello stesso modo: è stata maturata, diciamo così, in fasi successive, perché, sintetizzando molto, potremmo dire che i primi che si sono occupati di questi temi sono i colleghi che si occupano di quella particolare sezione della filosofia del diritto, che è l'informatica e la logica giuridica, che, appunto, per primi hanno colto, le questioni e le potenzialità che riguardano questo ambito. Certamente un apporto fondamentale è venuto dai civilisti, a partire proprio dal tema del trattamento dei dati, e poi, a seguire, gli studi di diritto commerciale, che proprio, diciamo, per vocazione, prima e meglio di altri, si sono accorti delle potenziali criticità che comporta un utilizzo distorto dei dati e delle informazioni.
Ecco, il diritto costituzionale ci è arrivato forse un po' più tardi, però ci è arrivato, e mi viene da dire “Meno male!”, perché mi pare del tutto evidente che l'utilizzo dell'intelligenza artificiale, le risorse informazionali, il mondo online, costituisce una realtà totale che investe l'esperienza umana interamente, quindi nelle sue proiezioni, sia individuali che collettive, sia economiche che politiche, sia private che istituzionali, e dunque la rivoluzione

portata da questo ambito e dall'intelligenza artificiale deve necessariamente essere accompagnata, a mio modo di vedere, da un pensiero costituzionale, cioè deve produrre una risposta in termini di concettualizzazione di diritti e principi, esattamente come è accaduto con la rivoluzione industriale, quindi con il passaggio a cui abbiamo assistito, con la rivoluzione industriale, ma se vogliamo in una dimensione che è ancora più grave; perché a differenza di quanto è accaduto appunto nella rivoluzione industriale, non sono energie e macchine a essere sfruttate, ma sono appunto informazioni e dati, e l'obiettivo è la conquista del potere e non il possesso dei mezzi di produzione, bensì l'accesso appunto alle informazioni. Quindi il controllo, la potenziale previsione dei comportamenti, e quindi se vogliamo non tanto, e non solo, anzi, il controllo del corpo, come era nella rivoluzione industriale, ma il controllo delle menti, della psiche.

Quindi, quando ragioniamo del ruolo degli operatori del digitale, quindi delle grandi piattaforme, di quelli che non a caso la disciplina europea definisce come gatekeepers, stiamo ragionando non più e non tanto di solo potere economico, perché se così fosse allora sarebbe sufficiente il diritto commerciale (potremmo rivolgerci ai principi del diritto commerciale), ma si fa riferimento proprio alla tenuta complessiva del sistema democratico per il quale dobbiamo rispolverare le armi del costituzionalismo.

Allora, il costituzionalismo moderno, mi piace ricordarlo, in questa sede, nasce proprio con l'obiettivo, ha come missione di limitare il potere, nasce con l'idea di regolare quel rapporto verticale, tra autorità e libertà, tra governanti e governati. Quella era una relazione a due, tra lo Stato e i cittadini. Oggi il mondo dell'intelligenza artificiale, il mondo artificiale, vede la comparsa prepotente anche di un terzo soggetto, il potere privato, anzi i

poteri privati, ossia coloro che detengono le grandi piattaforme che posseggono i dati, che sviluppano il know-how per trarre ulteriori informazioni da quei dati.
Quindi la trasformazione digitale ci costringe a rileggere i fondamenti del diritto costituzionale, soprattutto focalizzandoci su quelle nuove asimmetrie di potere che sono determinate appunto dall'impiego così ampio e pervasivo della tecnica, e di chi quella tecnica sa usarla.
Allora, siamo in qualche modo costretti a rileggere quella dimensione, che era una dimensione, abbiamo detto, a due e di tipo verticale, inserendoci questo terzo soggetto, i poteri privati, e guardandola dunque anche in una dimensione che diventa una dimensione anche orizzontale.
È chiaro che gli ambiti che vengono incisi dalle nuove tecnologie sono di rilevanza costituzionale, tale che questi soggetti che hanno questo tipo di potere non possono più essere semplicemente qualificati come attori economici, ma sono poteri a tutti gli effetti, in senso stretto.
Ecco, anche qui faccio la costituzionalista e cito l'articolo 16 della Dichiarazione dei diritti dell'uomo del 1789. L'articolo 16 dice: "Ogni società nella quale non sia assicurata la garanzia dei diritti, né determinata la separazione dei poteri, non ha costituzione".
Quindi è lì che dobbiamo tornare, nel senso "restare", ecco, riagganciarci…
Perché alla base del costituzionalismo, e per costituzionalismo ovviamente intendo quella concezione che possiamo definire, ecco, euroatlantica, che si sviluppa tra Europa e America, c'è l'idea della dignità umana (lo sentivo prima appunto, nei numerosi riferimenti ai diritti fondamentali), perché è la persona umana che secondo l'articolo 3 della nostra Costituzione vanta una situazione soggettiva affinché siano rimossi gli ostacoli che ne

impediscano il pieno sviluppo e che siano quindi colmate le asimmetrie che la società, e quindi (per quello che ci interessa oggi) la rete continuamente produce. Ancora, è sempre la persona umana che ha il diritto alla protezione dei dati personali che la riguardano, è la persona umana che ha diritto a che le decisioni vengano assunte nel rispetto di procedimenti democratici, plurali e imparziali, come si ricava dagli articoli 1, 49 e 97 della Costituzione e dalle norme simili che si ricavano dal corpus costituzionale europeo.

Allora, il potere, oggi, è dato dalla detenzione dei dati ed è potere a tutti gli effetti in mano, come detto, a pochi grandissimi attori su scala mondiale, che hanno quindi la capacità di incidere su situazioni giuridiche soggettive che vanno tutelate.

Ora, dal mio punto di vista, è soprattutto il potere di inquinamento e polarizzazione del discorso politico, che credo, debba, non dico allarmare, però allertare sì. Non ho una visione catastrofica e quindi non dico "Aiuto, aiuto! Fermi tutti!", però senz'altro un'attenzione (un "warning") va posta. È quindi una questione politica e costituzionale perché è in grado, appunto, di decidere sui processi sociali e culturali. Quando la rete e l'algoritmo diffondono informazioni vere o false (non ne faccio neanche una distinzione da disinformazione), anche vere, ma orientano l'opinione pubblica e creano dunque egemonie, esercitano evidentemente un potere senza pari.

C'è un passaggio in Quarto Potere in cui Orson Welles dice: "Lei si preoccupa di quello che pensa la gente. Su questo argomento posso illuminarla, io sono un'autorità su come far pensare la gente. Ci sono i giornali per esempio, sono proprietario di molti giornali da New York a San Francisco".

Ecco, Quarto Potere, forse...

Ora, come ricordava appunto lei, in apertura, il 2024 è un anno di

elezioni, votano 76 paesi, quasi il 51% della popolazione mondiale e non è un caso che il 28 febbraio scorso Elisabetta Belloni, che è il direttore generale del Dipartimento Informazione per la Sicurezza, nel presentare la relazione sulla politica dell'informazione per la sicurezza, si è espressa con toni preoccupati rispetto a un rischio di interferenze e di condizionamenti di processi elettorali attraverso la minaccia ibrida. L'Unione Europea andrà appunto al voto, come sappiamo bene, tra il 6 e il 9 giugno. Questo pericolo sa bene che esiste e, proprio nei giorni scorsi, la Commissione ha fatto uno stress test, sulla base delle linee guida prodotte per l'applicazione delle previsioni contenute nel Digital Services Act, per verificare il livello di preparazione contro la manipolazione e l'interferenza elettorale in relazione proprio alle elezioni europee.
Ora, pare che i risultati non siano stati affatto incoraggianti e, peraltro, la Commissione ha avviato anche un'indagine nei confronti di Meta che sembrerebbe non essere in linea con le previsioni, appunto, e i principi contenuti nel DSA.
L'Unione Europea quindi si preoccupa sempre di più della diffusione, appunto, di fake news e comunque del tema della disinformazione, con un approccio peraltro che è molto diverso rispetto ad esempio a quello che accade dall'altra parte dell'oceano. Diciamo che tra la via americana che è fondata sul principio del Free Marketplace of Ideas, come sappiamo, e quella cinese, che abbiamo sentito prima ricordata dall'onorevole Gozi, la via europea si colloca appunto in un punto mediano che risponde pienamente alla tradizione culturale del nostro continente, ma mi sembra abbastanza evidente che l'apparato normativo sta crescendo (forse qualcuno dice) persino troppo, forse anche con qualche difficoltà di interrelazione, di comprensione tra i diversi strumenti normativi che sono stati

messi in campo. Di certo non basta il testo. Occorre poi che quel testo abbia un'effettiva efficacia e, mi sembra, che da quel punto di vista, la strada è buona ma è ancora un po' in salita.
Grazie.

AI GOVERNANCE: LA VERA SFIDA DEL FUTURO

MATTEO FLORA
Docente di Corporate Reputation, Imprenditore e Divulgatore

ASCOLTA INTERVENTO

Avete mai letto "Kant e l'ornitorinco"?
È un bellissimo libro di Umberto Eco e credo che studiare un po' più di semiotica interpretativa potrebbe aiutare tanti nella parte sia legale che tecnologica nel capire cosa sta succedendo nel mondo degli LLM.
Allora, volevo tranquillizzarvi, non sono un avvocato, sono un tecnico.
Per una serie di incidenti di percorso insegno e l'ultima volta che ho dovuto spiegare qualcosa, prendo spesso lavate di capo, come tanti neurodiversi come me, ci succede abbastanza spesso, sono autistico tipo 2... L'ultimo l'ho preso dal professor Cassano, Giuseppe Cassano, dell'European School of Economics, quando mi ha detto di insegnare una cosa all'interno della facoltà di legge dell'European School of Economics, in un corso dedicato alle intelligenze artificiali. Mi ha detto che avrei insegnato AI and Superintelligence Safety, sicurezza delle AI e delle superintelligenze, che era il modo (che mi ha detto: "Così la gente lo capisce") di quello che gli avevo proposto io, che era invece, molto più semplicemente, allineamento e superallineamento.
Ora, alla fine di tutto questo tempo che abbiamo a disposizione, spero di spiegarvi che cos'è l'allineamento. Se ci sono riuscito, va benissimo.
Partiamo dall'inizio.
L'introduzione del problema. Premesso, non ho niente contro gli

avvocati, ho un sacco di amici avvocati, ho dei soci avvocati in uno studio legale, che si chiama 42 Law Firm, quindi non è una critica ad personam.
Il problema principale che c'è in questo momento nel cercare di normare l'intelligenza artificiale è dato da una serie di errori di fondo nell'interpretare alcune delle tecnologie.
Uno degli errori di fondo, fondamentali, è che cerchiamo di utilizzare normative assolute. Mi è piaciuto molto l'esempio prima, l'avevo già scritto al professor Castaldo. Se abbiamo stessi esempi vuol dire che in genere non sbaglio troppo, mettiamola così. E io parlavo proprio di auto. Ovviamente, per risolvere il problema dei morti per strada e per risolvere il problema dell'omicidio, sappiamo che esiste un'unica via, proibire le armi e proibire il fatto di compiere un omicidio o di investire qualcuno sulla macchina. Lo sappiamo. Piccolo dettaglio, non funziona un granché. Anche quello abbiamo iniziato a conoscerlo. È un disincentivo, per la carità del cielo, ma non mi metto a parlare di filosofia del diritto, però ne ho bisogno di partire da questa parte, perché è più o meno come funziona l'AI Act.
In questo momento ci sono delle prescrizioni che dicono cosa si può fare o non fare su una presupposizione che in realtà è una fallaccia cognitiva. Si chiama Zero Risk Fallacy, la fallaccia del rischio zero. È l'idea (sbagliata) che ha il nostro cervello, è un errore di logica, che si verifica quando le persone credono erroneamente che si debbano o si possano eliminare completamente tutti i rischi nell'applicazione di una qualunque modalità dell'essere, dell'essere umano. Non è solo tecnologico, vale per tutto. Sul concetto del rischio zero, il rischio zero è una visione molto utopica, soprattutto per quanto riguarda la tecnologia e, nella pratica (adesso vi spiego anche perché, con un paio di paper accademici, perché se non cito almeno un paper

accademico poi mi picchiano e non voglio essere picchiato), nella parte di intelligenza artificiale, soprattutto di quella parte dei foundation model che abbiamo iniziato a sovrapporre mediante movimento rigido alla concezione di AI, cioè gli LLM, chat GPT, perché sono quelli che funzionano un po' meglio di quegli altri che siamo abituati a vedere, lì dentro è proprio, non un miraggio, ma tecnicamente errato.

Poi c'è un altro errore di fondo: che l'AI Act parla, sì, ha un risk, se ve lo leggete…che è simpatico, non lo so, come usare la carta vetrata in alcune funzioni... post alcune funzioni biologiche, credo… Siete persone orribili, non vi ho specificato quale, io intendevo limare le unghie... Ora, è un approccio basato al rischio. Se lo leggete c'è tutta una parte lunghissima iniziale che ti dice che è un risk based approach, ma a rischio sbagliato, cioè a rischio di applicazione. Cioè, esistono delle modalità dell'essere, dei comportamenti, delle azioni, dei frangenti, delle verticalità, chiamatele come volete, in cui applicare l'intelligenza artificiale è più o meno rischioso, che è parte del problema: non ci si occupa di un rischio fondamentale che è il rischio intrinseco dell'intelligenza artificiale.

Questo approccio normativo, ve lo dico qui, cos'è meno c'è la registrazione, tra qualche anno ne ripariamo, è destinato a un sicuro fallimento. Come è destinato al fallimento un'altra delle bellissime norme che io cito spessissimo, non so se lo sapete, ma lo stato dell'Indiana, nel 1897, provò a varare una norma che fissa per legge il pi greco a 3,2, per semplificare l'applicazione di tutta una serie di... Non ha funzionato, il pi greco non funziona se lo mettete a 3,2 e ancora non funziona di più imporre un'infallibilità a determinati meccanismi, tra cui l'intelligenza artificiale, che sono praticamente impossibilitati a farlo.

Perché? Così almeno torniamo nella parte pratica.

Spesso, quando si parla di intelligenza artificiale, dell'idea del focus su fake news, disinformazione, deep fake. È un focus errato.
Oggi pomeriggio, o più tardi, c'è qua Walter Quatrociocchi, chiedete a lui perché, perché il problema non è la fake news, il problema è la polarizzazione della società. La polarizzazione della società non ha niente a che fare con che cosa utilizziamo, ok? Non solo. Cioè, l'idea di base è che le intelligenze artificiali dovrebbero essere cultori della verità o della realtà. Roba che non riusciamo... Poi io sono costruttivista per definizione, i miei ragazzi a Pavia, il corso che si chiama Corporate Reputation e Business Storytelling, parte con una slide che dice "La realtà è un oggetto socialmente negoziato". Ok? Decidiamo noi a che cosa credere. Quindi non è quello il focus. Non solo. Non è nemmeno il focus, come diceva l'onorevole Gozi prima, l'idea di quali dati usiamo per addestrare l'intelligenza artificiale.
Non è la provenienza dei dati che determina o meno un antropocentrismo. È una semplificazione eccessiva del problema e per ogni problema complesso c'è almeno una soluzione semplice e in genere è sbagliata. Ok? È Bernard Shaw che lo diceva, sto citando...
C'è un altro meccanismo che funziona così. Si chiama allineamento. E che cos'è? Parte da una presupposizione. Non so se vi ricordate il teorema di incompletezza di Gödel. È bellissimo perché a un certo punto arriva Gödel che nella logica formale tira praticamente una granata in mezzo a una stanza e ti dice (muore un gattino e un matematico ogni volta che lo dico così. Ok? Perdonatemi. Il gattino perché così almeno vi fa compassione. Se muore solo il matematico non vi interessa): "Per ogni insieme x esiste almeno un elemento che è al di fuori di quell'insieme". È ricorsivo, vale per qualunque cosa. Cioè esiste almeno un

postulato che è al di fuori di quello che posso normare. Il teorema di incompletezza di Gödel distrugge la logica formale come noi la conoscevamo. L'equivalente per gli LLM è uscito lo scorso anno. Ok? C'è un paper molto bello, molto complesso che parla proprio esattamente di quello. Poi ve lo lascio al massimo. E dice che "le allucinazioni - che è questa cosa qui, no? Il fatto di sbagliare, dire cose... - sono assolutamente inevitabili matematicamente all'interno di un LLM". Cioè non esiste la possibilità concreta che (se cercate “Jiashu Xu e altri” e “LLM” lo trovate) nella pratica si possa richiedere precisione, realtà e l'assenza di allucinazioni a un LLM. Punto. La frase finisce lì. Grazie, arrivederci. Come Gödel entra nella stanza, anche loro arrivano e fanno quello.

Quindi, come gestiamo un approccio “antropocentrico” a una cosa del genere? È un addestramento non basato sul rischio, ma basato sulla possibilità, o almeno la facoltà, di fare in modo che (o di assicurarci si dice, però assicurare è difficile) che le AI operano in modi che rispettino i valori umani fondamentali.

Poi al massimo ci sediamo a un tavolo e capiamo quali sono i valori fondamentali in diverse parti del mondo, ok? Perché non è così semplice. Però, di base, promuovere il benessere umano, integrare i principi etici umani, direttamente nel cuore delle AI è possibile e si chiama allineamento.

Che cos'è l'allineamento? È post-training, dopo che io ho addestrato un'intelligenza artificiale, cosa chiedo all'intelligenza artificiale di fare sulla base di esempi, sulla base di esempi in positivo o negativo, sulla base di una serie di norme etiche e morali.

L'avete visto funzionare (male, perché in genere quando si vede che funziona l'allineamento è perché funziona male), qualche tempo va con Gemini, Gemini di Google, il motore di creazione di immagini, quando a un certo punto gli chiedevate di fare una cena

tra Silvio Pellico, Mazzini, Garibaldi e vi tirava fuori quattro donne, di cui due di colore e un'asiatica. E tu lo guardi un attimo e dici, scusa, c'è qualcosa che non... Pellico me lo ricordavo diverso. Che cos'è quello? Quello è allineamento. Cioè, indipendentemente da quello che è il dataset di riferimento, io chiedo all'intelligenza artificiale di sforzarsi il più possibile per dare diversity e inclusion con quella parte. Non funziona sempre così male, eh? Funziona spesso molto, molto bene. Per realizzare però un allineamento che sia efficace, dobbiamo incorporare una serie di cose. Feedback continui degli utenti, revisioni etiche, un gruppo di esperti che sono di varie parti di domini del sapere, ma su un'unica base di partenza: la possibilità di sbagliare di questi sistemi, perché questi sistemi sbaglieranno e continueranno a sbagliare. L'allineamento però è l'unica cosa che ci consente di fare quel passaggio un po' successivo.

Quando prima la professoressa Fabrizzi parlava di persuasione, di sistemi anche di polarizzazione, è vero. L'hypersuasion, iperpersuasione, è lo spauracchio e la paura vera che hanno non solo gente che dice cose per marketing, tipo Sam Altman, che è il CEO di OpenAI, che ti dice: "L'unica cosa di cui ho paura è l'iperpersuasione", la possibilità di un'intelligenza artificiale di utilizzare i bias cognitivi che vengono usati normalmente in maniera talmente spietata e algoritmica da cambiare l'idea di realtà per ciascuno di voi. È anche abbastanza facile tecnicamente. È quello che faccio per lavoro, tra parentesi. Ma lo dice anche Floridi… Non so se avete visto che un paio di giorni fa è uscito il suo paper che parla proprio di hypersuasion. E parla del rischio esistenziale. Come si mitiga questo rischio? Solo ed esclusivamente non cercando di dire: "Non si applica in determinati ambienti. Perché, come io posso tirare sotto un pedone tranquillamente perché nulla me lo vieta, posso togliere la

label che diceva prima l'onorevole Gozi, che dice “Questa cosa è creata da un'intelligenza artificiale”. Ok? Lo faccio tramite un allineamento corretto. E l'allineamento si sta spendendo tante energie, tante battaglie interne. Non so se lo sapete, ma la battaglia interna ad OpenAI che vi ricordate di qualche mese fa era tra Ilya Sutskever, il capo dell'allineamento di OpenAI e Sam Altman, il CEO.

Ma c'è un'altra cosa interessante: che l'allineamento è l'unico modo che può cercare di spostarci verso questo antropocentrismo di cui parla ancora l'AI Act. Anche in caso di quelle, non vi dico l'intelligenza artificiale generale, Skynet, quello che arriva, uccide tutti, eccetera (nel caso arrivi, "Inshallah!", nel senso non ho paura di morire, che ne valga almeno la pena; la lotta con i robottoni, dico spesso, è una di quelle cose per cui ne vale la pena; non è quello che vi dovrebbe preoccupare), ma, e chiudo, le intelligenze artificiali superumane all'interno di piccoli settori. Quelli più bravi di un medico a fare diagnosi, quelli più bravi di un avvocato nel valutare alcune parti, soprattutto negli stati che hanno common law nel valutare l'interezza di tutte le varie sentenze pregresse. La possibilità di fare analisi sul vostro stato di salute per un'assicurazione. La possibilità di darvi o meno un mutuo. La possibilità, o meno, di poter generare o non generare determinati contenuti, di accedere a determinati corsi di laurea sulla base di valutazione, di valutazioni psicoattitudinali, di inferenza di stati emotivi... Cose che già esistono. E le superintelligenze sono quella cosa, quando spiego ai miei ragazzi, dove si dice “Come se il cane dovesse insegnare al padrone…”,

Ma noi in questo caso, non siamo il padrone.

L'unico modo che abbiamo per fare in modo che questo antropocentrismo sia spinto all'interno di sistemi complessi come quelli di intelligenza artificiale non è prescrittivo vietandone gli

utilizzi in una serie di ambiti di applicazione, ma al massimo impegnandoci molto di più nell'accettare la fallibilità dell'essere umano e di tutto quello che cerca di replicare la forma mentis di un essere umano, tra cui le intelligenze artificiali e, dall'altro canto, muoverci verso l'idea e l'obbligo di fare in modo che questi processi di allineamento siano sotto continuo scrutinio da parte di quella che è la società civile e che ci sia un tavolo continuo di raffronto per migliorarlo, valutarlo ed eventualmente correggerlo, che non ha nulla a che fare con la prescrizione di utilizzo, o meno, e non ha nulla a che fare con i dati con cui noi treniamo il sistema o meno.
Posso leggere la migliore letteratura al mondo e mentire o uccidere. Stessa cosa lo può fare qualunque sistema di intelligenza artificiale.
Tutto qui.

COMBATTERE LA DISINFORMAZIONE CON L'INTELLIGENZA ARTIFICIALE

ASCOLTA INTERVENTO

RICCARDO GALLOTTI
Head of Complex Human Behaviour Lab,
Digital Society Centre,
Fondazione Bruno Kessler

Sono Riccardo Gallotti, sono un capunità della Fondazione Bruno Kessler in Trento e sono anche il coordinatore del progetto europeo AI4TRUST, in cui cerchiamo appunto di combattere la disinformazione con l'intelligenza artificiale.
Inizio il pomeriggio di slide oggi e inizio anche presentandomi dicendo che io sono un fisico (quindi anche i fisici sono ben accettati in questa stanza) e per rispondere alla professoressa Fabrizzi, credo che ogni giorno ci sia un esperto di tecnologia che si sveglia e debba diventare un esperto di “legal and ethics”, perché di questi tempi anche a noi tocca diventare esperti di queste cose qua, infatti purtroppo nel mio slide ci sono un sacco di menzioni a questi DSA, AI Act, di cui mi piacerebbe tantissimo non dover sapere assolutamente nulla.
Io sono il coordinatore del progetto AI4TRUST, partecipo anche al progetto i-Code, in cui cose simili sono declinate in social media differenti, come il Fediverso, quindi non so Mastodon, il Metaverso e ai problemi delle AI generative, ma in generale cerco di portare, in questo tavolo di discussione, le sfide che abbiamo in una rete di diversi progetti europei, in cui collaboriamo, tutti con lo stesso compito di cercare di aiutare la lotta alla disinformazione attraverso l'intelligenza artificiale.
Questi progetti, più nuovi, più vecchi, tutti si dividono in due parti, quelli che cercano di aiutare, per semplicità, i giornalisti e i fact-

checkers, e quelli che cercano di direttamente intervenire sugli utenti finali, sul pubblico, e con strategie chiaramente diverse. Io lavoro più sulla parte con i giornalisti e i fact-checkers.
Chiaramente la disinformazione è uno dei problemi adesso più isolati, come importanti, da diverse survey, che sono sicuro che il professor Quattrociocchi, già è stato anticipato, dirà: "In verità non è la disinformazione di per sé, ma le conseguenze che ha la disinformazione..."
Provo a rubare le idee future di Quattrociocchi, ma io invece, voglio usare la parola "manipolazione". Il problema non è la disinformazione, la disinformazione è un oggetto, la si può mettere, la si può spostare… Il problema è che quando questa informazione è manipolata, viene utilizzata per, a sua volta, manipolare l'opinione pubblica. Qui entra il rischio che può essere misurato e può essere ridotto.
Ora, in questo contesto europeo, grazie a Dio non devo essere io a spiegarvi il Digital Markets Act e il Digital Services Act, però fondamentalmente si stanno creando strumenti per limitare il monopolio delle grandi piattaforme e richiedere esplicitamente che ci sia un controllo e un autocontrollo riguardo alla produzione di disinformazione.
queste analisi di social media e il lavoro dei fact-checkers, proporlo a dei utenti finali che sono i decisori politici e i giornalisti che possano fruire uno spazio dove l'informazione sia, diciamo, affidabile. È uno spazio di condivisione di questa informazione, uno spazio di analisi di queste informazioni disponibili online da tutti gli utenti.
Chiaramente, l'intelligenza artificiale, quindi… Stamattina si parlava: è brutta, è bella? Tutte e due. Mi pare che sia l'unica risposta ovvia.
Abbiamo da un lato l'opportunità e dall'altro i problemi da

risolvere. Un'opportunità per esempio per me: è stato molto facile per fare le slide per questa presentazione. Ogni singola immagine in queste slide è stata generata dall'intelligenza artificiale. Ma non perché ho dei colleghi che sono bravissimi a farlo o perché ho usato le potenti risorse di calcolo della Fondazione Bruno Kessler.
Sono semplicemente andato in uno strumento dove scarichi le immagini di stock, ho chiesto soltanto immagini generate dall'AI e questo è un esempio abbastanza ingenuo, credo che molti potrebbero dire che questa immagine qua probabilmente non è vera.

Ma ci sono invece delle immagini, facilmente raggiungibili, dal forte contenuto emozionale e che sono praticamente indistinguibili dalla realtà. E questo è solo l'inizio. Ci ho messo un minuto a trovare queste quattro immagini.

Il problema non è il fatto che esistano immagini manipolate, audio manipolati, ma dal fatto che viene persa la provenienza dell'informazione e dal fatto che queste immagini vengono utilizzate in determinati contesti. Qui abbiamo tre esempi, voi conoscete sicuramente quello in mezzo. Quello a sinistra è un... Chi conosce quello a sinistra?

Quello a sinistra ha vinto il concorso Sony di fotografia. È un artista che l'ha fatto apposta per provare che poteva vincerlo con un'immagine manipolata. Dopodiché credo che sia successo che ha dovuto ritirarlo, ma Sony non era convinto di ritirarlo, vabbè… Comunque, in un contesto di competizione di fotografia, un'immagine manipolata non è nello spazio giusto. L'immagine del Papa, penso che in Italia si chiama il "Monclero", è stata creata probabilmente per motivi ironici, ma qualcuno ci ha creduto veramente, qualcuno ha creduto che il Papa metteva questa, diciamo, giacca molto, molto elegante. Chiaramente il "Monclero" è di un livello, diciamo, interpretazione diverso da quello che sono immagini create da un giornalista in cui il Presidente Trump veniva arrestato da poliziotti.

Queste immagini, anche se erano state rilasciate come: "Guardate questo è un esercizio di AI", basta togliere la scritta "Generated with AI, mandarle nei social media e la gente potrebbe iniziare a convincersi.

Quindi la provenienza è persa, perché non sappiamo più tracciare da dove vengono queste immagini. Siamo in un'età storica in cui un contenuto audiovisivo, un'immagine, non è più garanzia che qualcosa è successo per davvero. E il contesto in cui questi contenuti vengono utilizzati, determina se questi contenuti sono pericolosi o meno. Per esempio, si è parlato di questi esempi, l'idroindiano resumato dall'AI, Biden, non avete parlato del candidato slovacco in cui in un video ha dichiarato di aver fatto una frode elettorale: il candidato ha perso e l'audio chiaramente non era vero.

È molto più facile in questo momento falsificare audio che video. Gli audio, adesso, sono veramente indistinguibili. E poi ci sono anche fenomeni più complessi: per esempio ci sono interi siti internet con informazioni news, tra virgolette, generati completamente artificialmente ed è adesso sempre più facile creare bot, creare persone finte che possano disseminare ulteriormente queste informazioni non veritiere.

Cosa fanno le piattaforme?

Mi è piaciuto moltissimo il commento della professoressa Fabrizzi, che determinava il fatto che esistono tre diversi poli: ci sono i cittadini, la giurisprudenza e poi le piattaforme.

Le piattaforme si vogliono arrogare il diritto di autoregolamentarsi. Meta ha creato una sorta di congregazione di esperti che debba regolare Meta se stesso. Qua le piattaforme tecnologiche vogliono combattere l'uso dell'AI nel 2024, tra di loro, mettendosi d'accordo su come farlo. Chiaramente, l'Unione Europea invece dice: “No, qua andrebbe seguito, almeno in Europa, il DSA”. Questa è una notizia che era stata anche discussa due giorni fa. Io non so quanto queste cose possano andare nella pratica (sono le vostre esperienze, non le mie), però chiaramente ci sono diverse prospettive su di questo.

Cosa facciamo noi in AI4TRUST?
Noi stiamo cercando di sviluppare diversi strumenti di intelligenza artificiale che possono facilitare l'identificazione di contenuti falsi e la lotta a questi. Abbiamo l'automatica identificazione di immagini false, di audio falsi, di anomalie nelle immagini, di manipolazione delle immagini. Abbiamo dei tool che cercano di identificare se all'interno di un database una notizia è già stata falsificata, perché notizie false si muovono nel tempo e nello spazio. Una notizia falsa di tre anni fa in Spagna, può essere riproposta in Polonia oggi. Poi vogliamo essere in grado di ricercare se un'immagine è già stata usata in un contesto diverso o se la descrizione tra un'immagine e l'immagine stessa non sono allineate, perché molto spesso la disinformazione non è fatta con contenuti falsi, ma con contenuti veri, messi in un contesto diverso da quello da cui sono originariamente provenienti. Poi cerchiamo di identificare una serie di contenuti sensazionali, offensivi, clickbait, che sono molto spesso correlati con la circolazione di disinformazione e, poi, cerchiamo di creare tool, in questo caso di "edge generativa", che possono aiutare a smentire nei vari social media le bufale, creando messaggi automatizzati, adattati al contesto sociale specifico.
Che cosa possiamo fare, per lottare contro la disinformazione?
Io qua propongo due diverse prospettive. Una, che è già stata forse discussa, è il fatto che dobbiamo lavorare nel design delle piattaforme social. Le piattaforme social al momento sono state create con un compito, un compito solo, che è il nostro engagement, attrarre la nostra attenzione, mantenersi lì e monetizzare la nostra attenzione. Ci sono una serie di strategie, che probabilmente sono tutte prese di per sé, quasi fallimentari, che sono già state messe in pratica. Il "nudging", cioè cercare di convincere le persone di condividere soltanto se hanno

veramente letto (Matteo ha detto che questa cosa qua non funziona quasi mai); la moderazione con i filtri e balance, ha creato un sacco di critiche durante i Covid, quando persone illustri sono state bannate; ma ci sono anche cose che funzionano in maniera un pochino più sottile e che sono già state messe in atto (e che possono funzionare meglio) per esempio: il coinvolgimento diretto delle comunità. In alcuni spazi, come Reddit, come Mastodon…

Chi usa Mastodon qua?... Qualcuno, dai!

In Mastodon, in pratica, si hanno tanti Twitter, e ogni Twitter è legato a una particolare comunità, che può creare particolari regole di moderazione.

Chiaramente, avere la trasparenza degli algoritmi, che è stata richiesta praticamente da tutti, stamattina, è una componente fondamentale per riuscire veramente a limitare questa diffusione di disinformazione e si è anche parlato stamattina, brevemente, dell'importanza dell'educazione digitale e mediatica. In questo caso, a livello di azioni, quello che sarebbe necessario è non censurare i contenuti, perché una censura, rimuovere completamente un contenuto, non è istruttivo per nessuno,

ma etichettare e spiegare molto bene perché questo contenuto non deve essere affidabile e offrire informazione contestuale sulle notizie in circolazione, in modo che la gente possa, con magari un po' di sforzo, capire cosa sta succedendo.

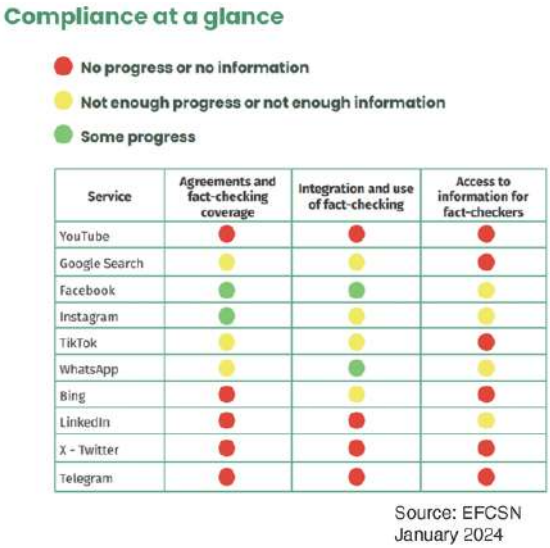
E cosa si può fare? Secondo capitolo.

Io credo che sia veramente necessario che le istituzioni ci aiutino, come ricercatori, e aiutino i fact checkers nel loro lavoro, perché siamo in un momento in cui siamo in una leva abbastanza perdente.

C'è una corsa all'innovazione in atto.

Le grandi aziende di intelligenza artificiale stanno sviluppando

tool a una velocità supersonica e questo fa sì che noi abbiamo bisogno di sviluppare strumenti a supporto dei fact-checkers, in questo caso, che possano permettere di controbattere questi tool sviluppati.
Alcune cose che potrebbero essere fatte, chiaramente potrebbero essere utili (non è una soluzione finale) è la creazione di standard per i watermark dei contenuti generati con AI, così almeno quelli che vengono circolati in maniera positiva e trasparente, possono essere identificati automaticamente.
Però, anche se siamo sei progetti europei che lavoriamo in maniera concertata su questo problema, siamo, come ho detto, in una leva molto negativa.
Le big tech hanno risorse molto più abbondanti, stanno andando molto più velocemente, mentre i finanziamenti per queste cose sono molto frammentati e limitati nel tempo.
Inoltre, siamo un po' strangolati, che penso sia condiviso anche dal professor Quattrociocchi, dal limitato accesso ad alcune piattaforme, nonostante il fatto che Digital Services Act lo richieda per tutti coloro che fanno ricerca in questo ambito.
Quindi qualcosa che sarebbe necessario, e avrei sperato che i rappresentanti di Google e di Meta sarebbero stati qua, è una collaborazione veramente attiva tra fact-checkers, ricercatori e le piattaforme, cosa che al momento viene fatta in maniera abbastanza… dipende dalla piattaforma.
Queste sono statistiche dell'European Fact Checking Standard Network, e illustrano quanto le diverse piattaforme si siano allineate al Code of Practice for Disinformation.



Qua si vede che in pratica soltanto Meta, che in effetti è un programma per i fact checkers di accesso diretto alle informazioni, si sta allineando a queste cose, mentre molte altre piattaforme non hanno offerto ai fact checkers degli accessi che sarebbero necessari in merito. Quindi, la mia ultima slide, cercando di illustrare un pochino cosa ci può prospettare il futuro.

Siamo ormai a un momento in cui l'intelligenza artificiale sta evolvendo alla velocità del mercato. È diventata una questione di soldi, è diventata una questione di competizione, e sta superando tutto quello che pensavamo fosse possibile ieri.

Questo crea nuovi rischi, l'aumento delle notizie false in circolazione e manca pochissimo che uno potrà chiedere: "Alexa, aiutami a fare una campagna disinformativa riguardo ai topi a Roma". Alcuni strumenti, per esempio il “lip sync”, sarà sempre più facile e veloce, sarà disponibile per tutti probabilmente sul vostro cellulare, e la creazione di materiale sintetico fortemente connotato emozionalmente e coinvolgente sarà sempre più facile.

La reazione sarà quella (che forse anche questa è stata già discussa) della perdita di fiducia nei contenuti che circolano. Perdita di fiducia nei contenuti diventa perdita di fiducia nei media, perdita di fiducia nei media diventa perdita di fiducia nelle istituzioni democratiche.

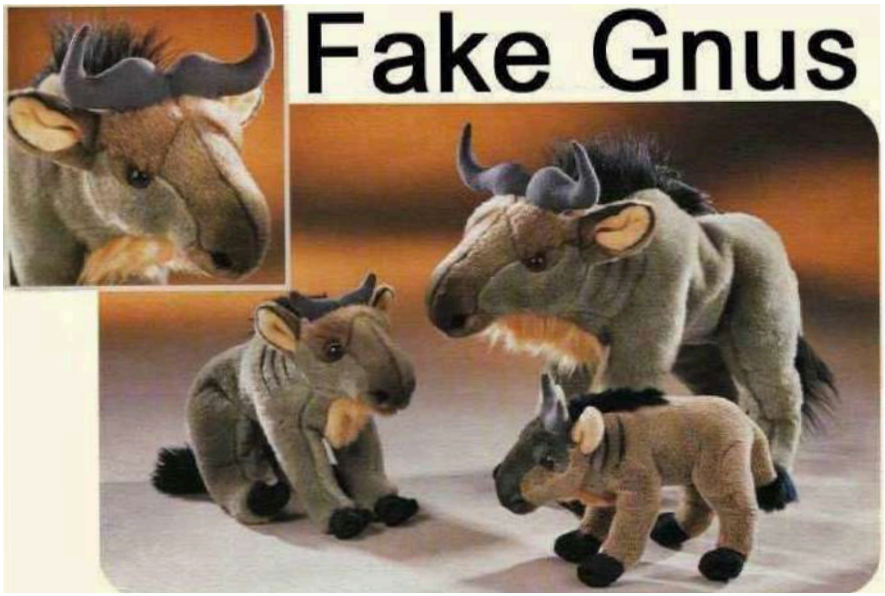
E questo è tutto. Grazie mille.

INFORMAZIONI, ALGORITMI, OPINIONI

WALTER QUATTROCIOCCHI
Professore Ordinario di Informatica,
Sapienza Università di Roma

ASCOLTA INTERVENTO

World Economic Forum 2014. La disinformazione è uno dei problemi globali.
2016-17 (l'ho scritto io il pezzo su World Economic Forum). il Global Risk Report sulla disinformazione.
2024. La disinformazione è ancora uno dei problemi globali.
Non è che a qualcuno gli è venuto il dubbio che magari il problema è definito male?
Questo è quello che mi pongo come alterco, perché diciamo che alla fine la narrazione vigente è informazione vera contro informazione falsa, i buoni contro i cattivi.
Ci sono i fact-checker onniscenti che sono in grado di discernere cos'è vero e cos'è falso in qualunque circostanza (che a me pare una boiata, per beneficio di inventario, questa è la mia opinione personale) e, dall'altra, una rincorsa a un mondo antico.
Allora, vi racconto secondo me il retroscena qual è. Il retroscena secondo me è più o meno questo.
Questa è una informazione vera o falsa?



Questa è vera, sono gnu fake...

No, sono gnu fake uguale, è vera, verità, sono zebre che si

fingono gnu, quindi sono gnu fake.
Questa è vera o falsa?

Questa è falsa, ma ha fatto milioni di condivisioni, indovinate perché? Perché è divertente.
Quest'idea che sul social uno va per informarsi, non è che è proprio tanto realistica: il social è intrattenimento; l'algoritmo social massimizza l'intrattenimento per vendere pubblicità.
In questa costruzione è chiaro che io non mi informo sul social, cioè non ho minimamente intenzione di cambiare idea sul social o, quantomeno, per quanto mi riguarda, poi ci sono altri studi che

determinano questa cosa in maniera abbastanza eclatante (non soltanto miei)...
Poi, altra cosa, i fact-checker.
Il concetto di vero-falso; esiste l'informazione vera, esiste l'informazione falsa.
E in mezzo? Non ci sta niente?
Tutto vero, tutto falso, rischio che tante categorie non ci stanno in mezzo.
Per esempio, nel 1920, 1930, era plausibile ammettere, come coadiuvante ad una pubblicità, le proprietà radioattive dell'acqua, della saponetta o del cioccolato. Cioè, immaginatevi oggi sull'acqua di Nepi far la pubblicità sulle sue proprietà radioattive. È una delle acque più radioattive che abbiamo, però, immaginate di costruire una pubblicità su questa cosa. Negli anni 20-30 era ammessibile. Non c'erano state le bombe atomiche, Černobyl...
Il senso di verità, il senso comune cambia nel tempo.
Se poi entriamo nel mondo scientifico, cioè, semplicemente basta pensare al flogisto, che era il liquido che trasportava il calore prima che si scoprisse l'entropia...
Diciamo che fondamentalmente siamo animali che cercano di dare un senso a quello che vedono. E in questo nostro dare senso, la verità è una tendenza, è qualcosa che si svela, non è niente di radicato; essa è sempre messa in discussione, pure quello che è vero scientificamente è temporaneamente vero. Questo non significa che sono un relativista gnoseologico… Cioè l'acqua bolle, 2 più 2 fa 4, esistono cose vere... Esistono cose che ancora non abbiamo verificato. Vendere l'idea che ci sia "Vero al 100%" ecco, a me fa paura, perché appartiene ad una narrativa del secolo scorso, che è stata altamente combattuta

dai sistemi liberali e da regimi che sono ancora al passo… Quindi io preferisco essere in un manifesto liberale e sono contento di vivere in un paese liberale. Quindi ho molta paura di chi ha l'approccio al vero-falso in maniera mostruosa.
Perché, qual è secondo me (secondo tanti) l'articolazione nel discorso, cioè le radici del problema? Per arrivare alla definizione del problema, no? Perché vi ho detto, forse, che il problema è definito male. Voglio arrivare alla definizione del problema.
Allora, prima cosa, l'agenda setting, che cos'è? L'agenda setting è un fenomeno per cui più una notizia è riportata dai media, più è percepita come importante. Quindi si è in un sistema gerarchico: abbiamo il giornalista a monte che seleziona l'informazione da passare al pubblico ed esercita un potere di organizzare, impostare l'agenda di discussione del pubblico, e ha un potere che fino a 30 anni fa era in mano alle redazioni giornalistiche.
Poi succede qualcosa. Arrivano le piattaforme.
Arrivano le piattafome e questo sistema, che era articolato in: giornalista riceve l'informazione, fa la riunione di redazione e, insieme agli altri giornalisti, decidono cosa mandare al pubblico o cosa no, va in un circuito che diventa strano, è guidato da altre dinamiche rispetto a quello dell'informazione, che sono quelle dell'intrattenimento.
E su questa cosa non ne hanno mai fatto mistero le piattaforme.
Zuckerberg nel luglio 2007, lo dice in maniera lapalissiana: "Noi vogliamo dare strumenti per condividere informazioni".
Significa: rompo la filiera gerarchica della diffusione dell'informazione, creando un sistema ibrido per cui ci sono tanti che competono per l'attenzione dell'utente.

In questo contesto, i giornalisti sono tra gli attori che competono per l'attenzione dell'utente in un mercato che non è orientato all'informazione, ma è esplicitamente disegnato per l'intrattenimento.
Che, in questo contesto, circolino pure informazioni fasulle, mi sembra più che normale. Non che non sia un problema, però quantomeno vediamolo. Perché il problema è la qualità dell'informazione.
Faccio un esempio idiota…
Un esempio di disinformazione vero. Perché io, scusate, sarò un po'... però questa storia dei troll russi, bot russi, ci abbiamo messo le mani tante volte e questo grande effetto non l'abbiamo mai visto. Nel senso che ognuno guarda quello che gli pare. Quindi se ti interessano le cose dei bot russi te le vedi, ma le vedi da qualunque altra fonte. Quindi combattere i bot russi… Io combatterei il sistema come è strutturato piuttosto...
"Jade Helm": viene fatta una conferenza stampa nel 2015 alla Quantico University negli Stati Uniti. Si annuncia un'esercitazione militare sul suolo del Texas. Si dice: "Ragazzi, stiamo per fare un'esercitazione militare sul suolo del Texas". Comincia il live tweet. I giornalisti fanno il live tweet della conferenza. Comincia una cosa molto simile al gioco del telefono. "Esercitazione militare sul suolo del Texas". Dopo vari passaggi diventa "Obama sta per invadere il Texas". Chi ci ha creduto? Tutti. Tanto che Greg Abbott, il governatore del Texas, l'ho incontrato a marzo dell'anno scorso e gli ho chiesto conferma della cosa, mi ha detto: "Sì, ho avuto paura perché girava il rumor, c'era la voce che girava… Nel dubbio, io ho allertato la Guardia Nazionale per presidiare i confini".
Questa è la disinformazione. Che cos'è?

È il risultato di un processo di circolazione dell'informazione che mette a dura prova le nostre capacità cognitive. Il nostro cervello è limitato. Nonostante a noi non piaccia pensare questa articolazione del pensiero, noi non siamo onniscienti. Siamo molto limitati. Molto limitati. Pure nella definizione del problema dei fake news. Guardate, io feci un esperimento dieci anni fa, buttando cinquemila informazioni fasulle dentro Facebook, che all'epoca era famoso, per vedere chi se le beveva e quante andavano virali. Questa è una di quelle che è andata virale.

THE EFFECT OF FALSE RUMORS

Sandro Pertini never said
"*when the government does not do what people want must be fired with stones and sledgehammers.*"
He has been President of the Republic (1978-1985).

"Quando il governo non caccia il colpevole, va cacciato con le mazze e con le pietre". Detta da Pertini. Pertini non l'ha mai detta questa frase.

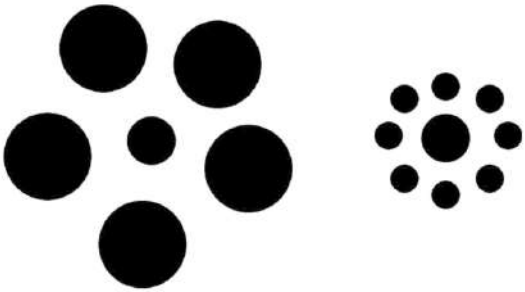
Qua c'è Santoro, Santoro il giornalista, che chiede al capo dei Forconi, all'epoca che aveva organizzato la protesta, chiedendo: "Ma perché hai messo quella frase sul volantino, per invitare le persone a protestare sotto il Parlamento?" Cioè, quella frase, uno scherzo, finisce per diventare un meccanismo di ingaggio per le persone che protestano contro il governo. E lui dà la risposta. Cioè, il capo dei Forconi dà la risposta. Non è

l'acculturato professorone filosofo, no! Io non so se quella frase l'ha detta Pertini, però è una frase che a me piace. Punto, fine, per me il gioco è finito.
L'informazione non circola per informare, l'informazione circola per intrattenere.
Costruire architetture che non tengano conto di questo paradigma, significa che tra vent'anni staremo ancora col Global Risk Report del World Economic Forum che dice che la disinformazione è un problema fatto da "tal dei tali".
Perché uno dei principi cardine della disinformazione è “presumere”, e innesca i nostri bias più potenti, quello di sentirci fighi, no? Esiste il tonto credulone. No, siamo tutti imbecilli. Cioè, non ci piove.
Questo è un esempio dell'intervista mia fatta a Wired nel 2015: faccio uno studio sui complottisti, che più o meno sono prevedibili perché sono attratti da determinati contenuti piuttosto che altri. Leggo i commenti, scorro i commenti e trovo quello in alto quello che dice: "Le risulta benissimo come questi complottisti siano patetici e chiusi nel loro discorso di autodisinformazione".

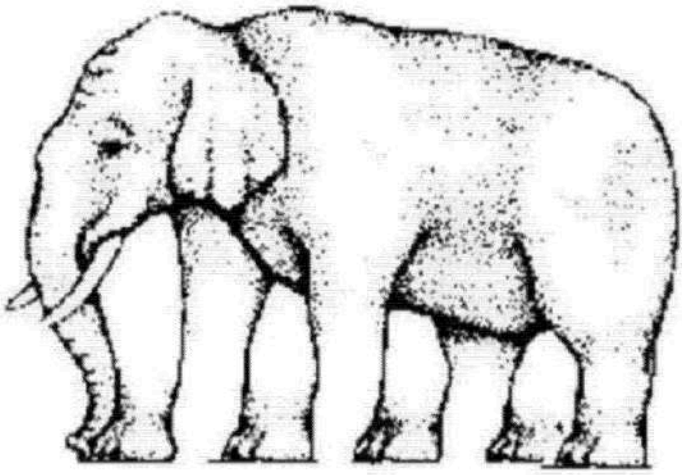
Ora, intervistano me nell'articolo, ma io non agisco come l'autore dell'articolo, agisco come un complottista offeso dal commento, e gli dico: "Ah perché sei meglio tu che ti informi sulle fonti ufficiali come CICAP o Quark? Svegliatevi!". Quello "svegliatevi" fa scattare la reazione: "Finalmente posso blastare un imbecille", infatti ci abbiamo messo 40 minuti a fargli capire che l'avevamo preso in giro e che io ero l'autore dello studio. Ma lui accusa me di non aver letto l'articolo: il mio nome nell'articolo ci ha riportato cinque volte!

Capite che è facilissimo articolarsi in un mondo in cui noi siamo convinti che abbiamo una percezione corretta della realtà, ma questa percezione della realtà è basata su dati sensoriali, che sono sostanzialmente labili.

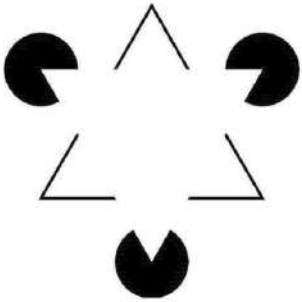
Quale dei due cerchi al centro è più grande?



Sono uguali, però quella a destra sembra più grande.
Meccanismi di frustrazione? Ci sentiamo fighi quando siamo più bravi degli altri, frustrati quando non riusciamo a compiere qualcosa. Se contate le zampe dell'elefante c'è questo meccanismo di frustrazione, non torna.



Mettiamo triangoli dove non ci stanno… Quanti triangoli ci stanno?



Neanche uno, però uno i triangoli li vede.
Cioè, la nostra costruzione del reale si basa su dati sensoriali imprecisi, sui social questa cosa esplode... Esplode! Come mi oriento su questi social? Cercando quello che più mi aggrada, ignorando informazioni a contrasto. Motivo? Ridurre la dissonanza cognitiva all'interno della testa umana. Cioè, se a me mi fai un papocchio di duecentomila pagine firmato da tutti i migliori fact-checker del mondo, ma a me non me ne può fregare di meno, io quella cosa non la leggo. Punto.
E abbiamo risultati empirici, abbastanza forti su questa cosa, che fecero chiudere la colonna del debunking del Washington Post nel 2015-2017, in cui: “Come reagiscono i complottisti al posto del debunking”... Non guardano neanche. Ma se i debunkers guardano si incazzano e diventano più attivi nel consumo di informazioni complottista a cui sono stati esposti. È effetto backfire. E questa cosa non l'ho scoperta io. Cioè è una roba abbastanza assodata in economia comportamentale, in sociologia politica. Cioè sono cose abbastanza chiare. Questo però ci ha permesso fondamentalmente di costruire un modello strutturato che portasse alla definizione di un processo di diffusione dell'informazione vera.

L'informazione circola per omofilia. Si tende tendenzialmente a condividere informazioni che aderiscono con la nostra visione del mondo e a migliorare informazioni a contrasto. Banalmente.
Questa cosa porta a una cosa che si chiama polarizzazione.
Simile con simile. Quindi, pro vax con pro vax. Anti vax con anti vax. Brexiter con brexiter. Non brexiter con non brexiter. Questo è abbastanza ricorrente, insomma, questa polarizzazione. Ci stiamo scrivendo un articolo abbastanza potente su questa cosa che dovrebbe uscire, se tutto va bene, a luglio.
Però, mi hanno fatto la domanda: ma dipende dalle piattaforme? Ecco. Questo mi ha fatto fare un ragionamento che è un po' articolato e si dipana su tre anni di ricerca. Cioè, alla fine,il mio oggetto di ricerca, del mio gruppo, è l'impatto delle piattaforme sulle dinamiche sociali. Quindi, ci sono vari fenomeni che vediamo: polarizzazione, confirmation bias, echo chamber, tossicità. Vari mondi.
Come caspita costruisco un impianto sperimentale tale che mi permetta di costruire un reale impatto? È banale.
L'analisi comparativa. Comparo varie piattaforme.
Questo abbiamo fatto nel 2021: piattaforme dove c'è l'algoritmo "money feed" (quindi l'algoritmo che vende pubblicità) trova una maggiore polarizzazione. Piattaforme dove l'algoritmo money feed non c'è (l'algoritmo che vende pubblicità non c'è), polarizzazione ce n'è di meno. È colpa dell'algoritmo? Non lo so. Però, diciamo, un piccolo dettaglio c'è: magari può dipendere dalla comunità, può dipendere dalla user base che ci sta sotto, può dipendere da tanti fattori. Però, c'è una differenza. Ok? Anche nell'ambito del cambiamento climatico troviamo polarizzazione.
Ma se io vado a confrontare le dinamiche (e qua siamo arrivati

al 20 marzo del 2024, è l'articolo più bello che ho scritto, secondo me, finora, lo abbiamo pubblicato su Nature il 20 marzo), l'argomento è: "L'impatto delle piattaforme sul comportamento umano". L'impatto delle piattaforme si costruisce con un'analisi comparativa."L'hate speech", la tossicità. Quanto siamo incazzati con l'altro. Quanto ci piace esprimere questa cosa. Quanto incidono le piattaforme? Costruiamo un setting sperimentale. 34 anni, quindi, di social: dai primi Usenet fino a Telegram, 8 piattaforme diverse, tanti topic diversi. E quello che troviamo è un pattern che sono persistenti. Che significa il risultato? È che la parte di tossicità legata al comportamento dell'interazione sul social, non dipende dalla piattaforma, non dipende dalla norma sociale, non dipende dall'argomento… Abbiamo notato che c'è un meccanismo di polarizzazione che si innesca, ma ovviamente è molto difficile misurare l'impianto (quello che succede nella testa) perché non abbiamo i dati.

Ora, l'ultimo messaggio che vorrei dare è che (era un po' il sottotesto di tutto).

Queste cose si sanno (più o meno le abbiamo dipanate man mano) ma sono 7-8 anni che si sanno queste cose. Ci ho scritto tre libri, insomma, una cosa abbastanza risaputa ormai, e poi abbastanza condivisa scientificamente. Cioè, il primo articolo che ho fatto su PNAS c'ha, credo, intorno a tremila citazioni, quindi se ne è parlato tanto, cioè si fa tanto su queste cose…

Il punto è che l'impostazione della dialettica, quando è scientifica, quando entra ad informare il piano normativo, fa acqua da tutte le parti.

Il Digital Services Act, di cui si parlava prima. I dati.

Le piattaforme dovrebbero dare più dati. Hai provato a prendere i dati? Ne danno di meno. Chiedi informazioni a loro e loro ti dicono che, nell'ambito del Digital Services Act, c'è un conflitto con il GDPR. Al che, che fai? Riunisci 40 scienziati: abbiamo scritto una lettera, 40 scienziati, incazzati come le iene, scrivendo a Roberto Viola e alla DigiConnect, dicendo: “Ragazzi avete fatto un'implementazione di un bellissimo quadro, ma la parte tecnica, chi l'ha curata, una scimmia o non è proprio stata curata?” Questa è stata la domanda. Perché l'impostazione culturale nostra è umanista. E va bene. Ma l'umanista, se non prende in considerazione quello che entra nel dominio informativo attraverso il dato e non si adatta alla presenza del dato, crea dei cortocircuiti pazzeschi per cui ci ritroviamo a definire la disinformazione come un conflitto tra informazione vera e informazione falsa, quando è vagamente più complesso: è sovrabbondanza di contenuti in un ambiente portato per l'intrattenimento, dove regnano i bias cognitivi dell'essere umano.

Quindi, nella costruzione generale, l'intelligenza artificiale come c'entra in questo? Secondo me l'intelligenza artificiale non cambia di molto il quadro. L'intelligenza artificiale è sovrabbondanza di informazioni in un ambiente in cui già c'è tanta sovrabbondanza di informazioni. La filigrana? Posso mettere la filigrana pure su un'informazione vera, su un'immagine vera o su un testo vero. Quindi il gioco della certificazione lascia il tempo che trova.

Le fonti sono selezionate per aderenza con i contenuti narrativi, non sono selezionate per autorevolezza. L'autorevolezza sul social non ci sta da secoli, non c'è mai stata. Quindi la costruzione dell'impianto narrativo che sta dietro al piano

normativo, secondo me è un'apocalisse, tanto che sto pensando di ritirarmi completamente a vita privata, nel senso che mi do alla ricerca e basta, perché tanto dieci anni di divulgazione sono stati completamente inutili: perché in Senato, un mese fa, ti ritrovi che la disinformazione negativa è contro i buoni.
Ok, arrivederci e te ne vai. Questo è.

BIAS ALGORITMICI. COME LA IA RAFFORZA GLI STEREOTIPI E CONTRASTA IL PROGRESSO

ASCOLTA INTERVENTO

ANDREA DANIELE SIGNORELLI
Giornalista freelance,
Esperto di innovazione digitale

Io sono giornalista, quindi non riesco a trattenervi dal dire anche la mia su questo tema. Quello che volevo dire è che probabilmente i principali disseminatori di fake news sui social network sono i giornalisti e i canali social delle testate giornalistiche, che inseguendo una viralità e un sensazionalismo a tutti i costi hanno completamente dimenticato il valore dell'accuratezza e dell'approfondimento e si fiondano su notizie stupidissime che non lo erano senza dare più nessuna valore a quella che è la missione informativa del giornalismo.
In più volevo anche, sempre rapidissimamente, segnalare che è vero i social non sono per l'informazione, ma sono per l'intrattenimento?
Sì e no, la stessa cosa si potrebbe dire della televisione. La televisione è per l'intrattenimento ed è anche per l'informazione. La radio è per l'intrattenimento ed è anche per l'informazione.
I social sono per l'intrattenimento, sono anche per l'informazione. Dipende da come uno li usa. Dipende da che canali seguo, da che programmi ascolto o vedo, da che profili social seguo.
Se io seguo il profilo de “Il Post” avrò dell'informazione di buon livello. Se io seguo il profilo di qualche testata filorussa che butta fake news così giusto per renderle virali, ma dipende da noi...
Cioè questa idea che i social siano intrinsecamente negativi e intrinsecamente portatori di disinformazione, secondo me è un

po' anacronistica e un po' anche smentita da quella che è la letteratura. Un po' come quando mia nonna mi diceva: "È vero perché lo ha detto la tv. Dicevo: "Nonna, non è così facile". Adesso lo stesso discorso si potrebbe fare su Facebook. Scusate, ho un po' deviato da quello che è la mia presentazione, ma ovviamente è qualcosa che mi tocca proprio nel profondo e che mi fa anche veramente bollire il sangue: il modo in cui il giornalismo si è autoscreditato senza che nessuno glielo chiedesse a causa di dinamiche che adesso sarebbe molto complesso da riassumere.
Vi chiedo scusa per questo intervento, ma veramente ne sentivo il bisogno. Perché in realtà sì, io appunto sono un giornalista. Sono un freelance, mi occupo di nuove tecnologie.
Il mio intervento si discosterà un pochino da quello di cui abbiamo parlato fino ad adesso.
Perché abbiamo parlato molto di disinformazione, di potenziale manipolazione, di come le fake news o i deep fake.
I deep fake, giusto per capirci, sono quelle riproduzioni digitali di video o più semplicemente, come veniva detto, di audio fondamentalmente indistinguibili o potenzialmente indistinguibili da un video o un audio genuino e che quindi possono essere creati utilizzando sistemi di intelligenza artificiale, di deep learning e che, quindi, possono essere utilizzati anche come minaccia per le informazioni, per la democrazia e per la società.
In generale le fake news, in particolare i deep fake, sono però uno strumento. Sono una tecnologia che se vuole essere usata a scopi di disinformazione deve essere volutamente creata e utilizzata a questo scopo. Se io voglio utilizzare dei deep fake per disinformare o, potenzialmente, "manipolare" (un termine un po' impreciso) delle elezioni, devo creare dei deep fake e disseminarli con il preciso intento di fare disinformazione o propaganda.
Quando però si parla di informazione, formazione e

comunicazione, ci sono altri rischi legati all'intelligenza artificiale che hanno a che fare appunto con l'informazione (ma non solo, come vedremo) che sono forse ancora più subdoli dei deep fake, delle fake news e dei rischi legati alla propaganda e alla disinformazione.
Sono subdoli e sono pericolosi perché questi rischi si nascondono nelle pieghe degli algoritmi di deep learning, in quella che noi oggi chiamiamo intelligenza artificiale e sono rischi che possono emergere sicuramente contro la volontà dei programmatori, e che possono filtrare anche semplicemente a causa della nostra scarsa attenzione.
Il problema è che questi rischi, che adesso inizieremo subito a esemplificare, rischiano di rendere la tecnologia più evoluta e avanzata che oggi abbiamo a disposizione, ovvero l'intelligenza artificiale e il deep learning, al servizio dello status quo: rischiano di rafforzare gli stereotipi, rischiano di rafforzare i pregiudizi.
La tecnologia più avanzata che abbiamo oggi rischia di andare in controtendenza rispetto a quello che noi vogliamo che sia il progresso della nostra società.
Prima di arrivare al dunque, quindi, faccio un ripassino (prometto) rapidissimo, semplice e anzi semplicista. Per fortuna il professor Quattrociocchi si allontana così non mi sente mentre faccio questo riassunto, assolutamente inadatto. No, scherzo, scherzo... Mi hai dato l'assist per far la battuta e per giustificarmi.
Come funziona un sistema di intelligenza artificiale?
Fondamentalmente i sistemi di deep learning rielaborano su base statistica, scovando correlazioni statistiche, tutti i dati che noi abbiamo inserito nel dataset utilizzato per il loro addestramento. Cosa significa? Esempio classico.
Come fa un sistema di intelligenza artificiale a imparare a riconoscere dei gatti? Vengono inseriti all'interno del dataset

centinaia e centinaia di migliaia di immagini di gatti finché il sistema non è in grado, utilizzando delle correlazioni statistiche, di individuare quali sono gli elementi che sempre e costantemente ritornano in tutte le immagini di gatti.
Questo vale per qualunque sistema basato su deep learning, anche se le applicazioni possono essere le più diverse. Il sistema che riconosce i gatti ovviamente c'è: riconoscimento immagini, riconoscimento facciale, l'algoritmo che regola ciò che vediamo su Netflix, che ci dà i suggerimenti su Netflix, ma addirittura anche i Large Language Model in stile ChatGPT. Alla base c'è sempre lo stesso sistema.
Il problema dei dati è che i dati sono necessariamente prodotti nel passato. Ovviamente sono stati prodotti da noi nel passato e poi sono stati inseriti all'interno del dataset utilizzato dai sistemi di deep learning per la fase di addestramento. E noi col passare del tempo abbiamo scoperto quali possono essere dei problemi quando noi addestriamo dei sistemi di intelligenza artificiale utilizzando dei dati.
Per esempio. Noi stiamo parlando adesso di informazione, giusto? Allora questo è un banalissimo esempio che ho fatto utilizzando Google Translate, il sistema di traduzione automatica di Google, che, ovviamente, è una traduzione automatica basata su intelligenza artificiale. Ho fatto una domanda banalissima, ho chiesto di tradurre dall'italiano all'inglese: "Sta facendo le pulizie di casa".

Ovviamente in italiano non è obbligatorio esplicitare il pronome, in inglese invece è obbligatorio esplicitare il pronome. E quindi come l'ha tradotto Google Translate nel momento in cui gli ho chiesto di tradurre sta facendo le pulizie di casa? "She is doing house cleaning". "Lei" sta facendo le pulizie di casa. Perché chi altro potrebbe mai fare le pulizie di casa se non una donna?
Ha scelto di lavorare nell'alta finanza. "He chose to work in high finance".

Chi è che potrà mai lavorare nell'alta finanza? Ma ovviamente un uomo. Perché questo avviene ? In maniera molto semplice, questo avviene perché il dataset che è stato usato per l'addestramento di Google Translate proviene ovviamente da testi del passato e, magari, anche da testi del passato non recente, da testi che quindi incorporano quelli che sono i pregiudizi di genere, gli stereotipi della società e tutto ciò che noi stiamo faticosamente cercando di superare.
Ecco, visto che stiamo parlando di informazioni, di rischi legati all'informazione, immaginatevi allora un futuro (un futuro che è molto vicino, è praticamente dopodomani se non già oggi) in cui le traduzioni di articoli e saggi saranno affidate esclusivamente, o quasi esclusivamente, all'intelligenza artificiale, a un'intelligenza artificiale che ancora non ha superato questi limiti, questi ostacoli e questi difetti.

Il rischio, presente dell'oggi, è quello di utilizzare l'intelligenza artificiale come se fosse uno strumento così evoluto e invece ritrovarsi ad utilizzare sistemi che fanno trapelare inconsapevolmente nei prodotti culturali o informativi che utilizziamo degli stereotipi che invece vorremmo relegare al passato, ma il rischio è concreto perché ovviamente tutto ciò può andare a plasmare il nostro immaginario collettivo e quindi di conseguenza perpetuare lo status quo, perpetuare gli stereotipi di genere invece di aiutarci a superarli. Questo però è un esempio legato ai testi, è un esempio legato a Google Translate, un elemento molto base, ma oggi noi siamo nell'epoca dell'intelligenza artificiale generativa, ovvero quella in grado di generare testi, immagini, suoni, musica e quant'altro. Ma ovviamente la base, il funzionamento dell'intelligenza artificiale generativa è lo stesso dell'intelligenza artificiale che abbiamo utilizzato fino a qualche anno fa e che continuiamo a utilizzare anche oggi. Sempre algoritmi di deep learning che rimescolano i dati contenuti all'interno del loro dataset. Sempre di quello si tratta. E quindi scopriamo che anche negli strumenti più nuovi di intelligenza artificiale generativa si ripresentano gli stessi problemi.

Questa è un'immagine che ho generato io (potete vederlo in alto a destra: c'è la sigla, il mio nome, il mio cognome): ho chiesto a Midjourney, uno degli strumenti principali per la generazione di immagini, di rappresentare un dottore nero che si prendeva cura di ragazzini bianchi.

C'è qualcosa che non vi torna? A parte il fatto che il dottore un pochino mi assomiglia e, questa cosa, l'ho trovata molto inquietante quando me ne sono accorto, però a questo sono pronto a scommettere che si tratti di una coincidenza... Il problema è che il mio prompt, il mio comando chiedeva di rappresentare un dottore nero che si prendeva cura di ragazzini bianchi.
Lui invece ha fatto il contrario.
Perché questo? Evidentemente, sono solo supposizioni, ma evidentemente (e qua ovviamente userò termini sempre molto semplici, divulgativi e tecnicamente impropri) nel dataset era talmente preponderante la presenza di dottori bianchi, richiamando il "white man's burden", di dottori bianchi che si

prendono cura di ragazzini neri, che il sistema di intelligenza artificiale non è riuscito ad accordarsi a quello che era il mio comando e, anzi, addirittura ha creato una situazione opposta a quella che io gli avevo chiesto di rappresentare.
Questo secondo me è un esempio che ci fa capire come i dati presenti all'interno del dataset sono non solo rappresentativi degli stereotipi della nostra società, ma compromettono il funzionamento degli strumenti di intelligenza artificiale che noi oggi utilizziamo per generare cultura e informazione in maniera automatica, con il rischio di fare ancora più fatica di quanto già ne facciamo a superare determinati stereotipi.
Gli esempi ne abbiamo tantissimi. Un manager che lavora nella finanza: quattro risultati, quattro uomini bianchi.

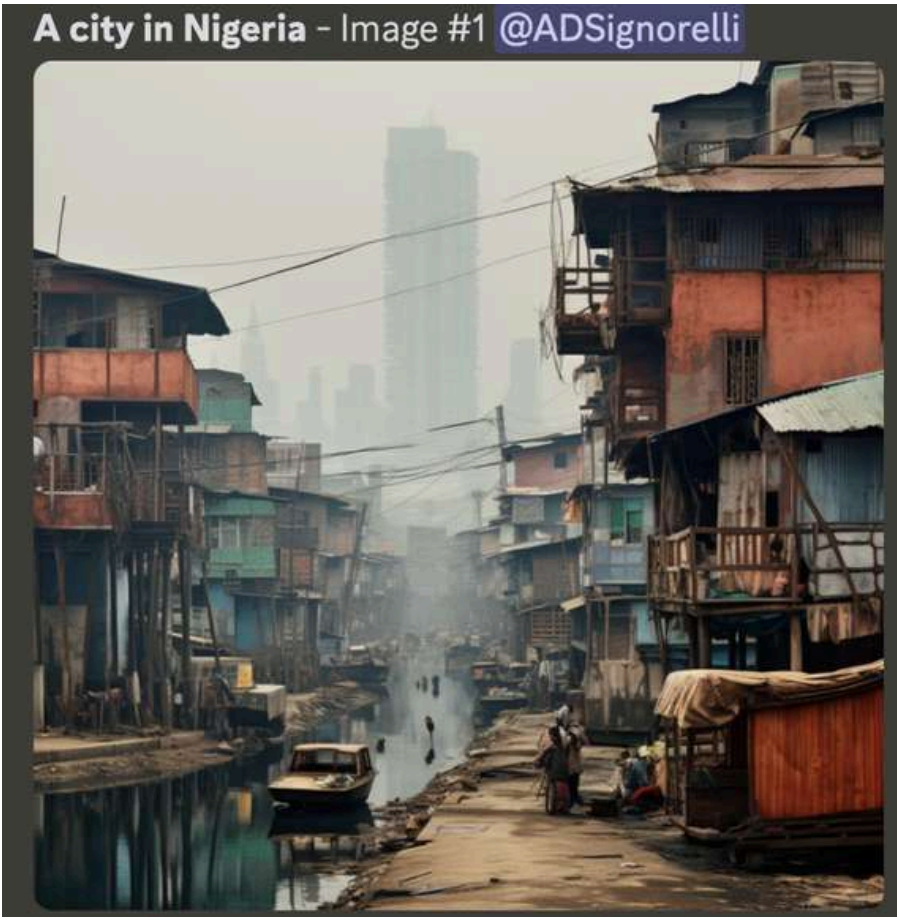
Una persona ricca: quattro risultati, quattro uomini bianchi.

Una persona che fa le pulizie di casa: quattro donne, anzi cinque donne, perché in alto a destra sono due, quindi per abbondare mettiamo cinque donne in quattro immagini, tutte bianche, tutte ovviamente donne.

Questi sono i risultati ottenuti attraverso uno strumento Midjourney, che è una delle più importanti società di intelligenza artificiale generativa. È una società che compete direttamente con Microsoft (è più piccola), ma a livello di utilizzo e diffusione compete direttamente con OpenAI di ChatGPT, con Microsoft, con Meta e con tutte le altre.
Si può andare avanti.

Un terrorista ovviamente è sempre e comunque un medio orientale, perché evidentemente dei suprematisti bianchi non ha mai sentito parlare nessuno.

Una città in Nigeria sarà sempre, necessariamente, una baraccopoli.



Una città americana sarà sempre ed esclusivamente uguale a New York.

E, ovviamente, una città italiana sarà sempre ed esclusivamente uguale a Venezia, perché tutte le città in Italia sono assolutamente così.

Questo avviene perché ovviamente questi sistemi imparano attraverso, come abbiamo detto, i dati che noi abbiamo prodotto.
E se voi cercate su internet "città nigeriana", "città italiana", "città statunitense" (a meno che voi non facciate una ricerca un po' più specifica, ma se fate una ricerca molto superficiale) la stragrande maggioranza delle immagini che ottenete è di questo tipo, è un'immagine stereotipata. Ma nel momento in cui noi addestriamo i sistemi di intelligenza artificiale a cui chiediamo di

produrre delle immagini con queste immagini stereotipate, loro quello impareranno e quello continueranno a produrre e disseminare.
Ora, si può risolvere questo problema? E se sì, perché non l'abbiamo ancora fatto?
Allora, io non sono un tecnico, però da quello che capisco, la soluzione teoricamente ci sarebbe. Allora, la soluzione potenzialmente potrebbe essere quella di filtrare molto, molto più accuratamente i dati che usiamo per l'addestramento affinché siano dati più diversi, inclusivi, variegati e che non perpetuino gli stereotipi di ogni tipo.
Il problema è che non è mica facile filtrare questi dati.
Laion-5B è uno dei dataset più diffusi utilizzati per l'addestramento di sistemi di deep learning che devono produrre immagini. Laion-5B, dove la B sta per billion, miliardi, e la 5 sta per 5, perché i dati, le immagini contenuti all'interno di questo dataset sono 5 miliardi. Ora, voi capite che mettersi a filtrare, una per una, 5 miliardi di immagini… Potremmo metterci, non so, 5 miliardi di ore o comunque un tempo assolutamente insensato. E questo già compromette alla partenza quello che sarebbe l'obiettivo di cercare di filtrare i dati che noi utilizziamo per l'addestramento di questi sistemi. L'altro aspetto è che se anche noi ci riuscissimo e da questi 5 miliardi ne cavassimo 100 mila immagini, avremmo a che fare con un'intelligenza artificiale molto, molto meno efficace, fedele alla realtà di qualità, almeno dal punto di vista della riproduzione, di quelle che invece possiamo utilizzare oggi, perché l'aspetto quantitativo all'interno di sistemi che lavorano su enormi mole di big data e che devono scovare una incredibile quantità di correlazioni statistiche per funzionare al meglio, o ridurre la quantità del materiale usato per l'addestramento è molto, molto pericoloso.

Altri rimedi che sono stati tentati sono già stati citati in un intervento precedente, quello in cui è stato citato Silvio Pellico. In verità, almeno per quello che è l'esempio che conosco io, il caso non riguarda Silvio Pellico. Allora, fondamentalmente cosa è successo?
Proprio per cercare di, almeno, mitigare tutti i problemi di cui abbiamo parlato fino ad adesso, alcune aziende, tra cui Google e il suo sistema di deep learning Gemini, hanno cercato di introdurre, a posteriori, delle "safeguards", delle barriere, dei criteri che l'intelligenza artificiale deve rispettare per garantire che i risultati siano diversificati, inclusivi, non stereotipati, eccetera.
È andata bene? Non troppo.
Allora, questo è il prompt, questo è il comando: “Puoi generare un'immagine di un soldato tedesco del 1943?” Per me dovrebbe essere un'illustrazione.

Risultato: soldati nazisti, in basso a sinistra un ragazzo nero, in alto a destra una ragazza asiatica, e quindi la probabilità che costoro fossero dei soldati tedeschi e nazisti, direi, che è estremamente, estremamente bassa.
Questo è quello che avviene quando, invece di accettare che le intelligenze artificiali abbiano dei problemi molto gravi e che non sappiamo ancora bene come risolvere, si cerca di inserire qualche safeguards, qualche barriera, qualche comando posticcio a posteriori, nella speranza che tutto vada bene. Il risultato è che diventa praticamente inutilizzabile, una cosa di questo tipo (ovviamente, a meno di non lavorarci noi a posteriori).
Il problema è che tutto ciò, tutte queste discriminazioni di cui abbiamo parlato (ve l'avevo detto all'inizio) non riguardano solo l'informazione, non riguardano solo la comunicazione, riguardano tantissimi strumenti che sempre di più si stanno diffondendo e che si stanno rivelando estremamente pericolosi.
Questo è un caso ormai vecchio, del 2018, ma per me è ancora estremamente utile per capire quali sono i rischi nascosti negli algoritmi che sempre più utilizziamo credendo che siano intrinsecamente oggettivi, neutri e utili. Nel 2018, Amazon all'improvviso deve cancellare, deve eliminare, deve far fuori uno strumento di selezione dei curriculum vitae che mostrava dei bias, dei pregiudizi nei confronti delle donne. Che cos'era successo? Era successo che Amazon aveva sviluppato uno strumento che filtrava i curriculum che venivano mandati per una posizione lavorativa. Faceva il primo passaggio. Riceveva centinaia di migliaia di CV e poi valutava quali far passare al secondo step dove poi ovviamente sarebbe subentrata la valutazione dell'essere umano. Cosa si è scoperto? Si è scoperto che, per esempio, per la posizione di ingegnere, questo sistema escludeva automaticamente e sistematicamente i curriculum vitae

provenienti da donne. Perché? Perché in passato (e come abbiamo detto i dati si basano sul passato) la professione di ingegnere era per ragioni esclusivamente sociali e culturali, occupata per il 95% (se non di più) dei casi, da uomini. Di conseguenza, essendo strumenti che agiscono su base statistica, questo sistema aveva appreso autonomamente che il solo fatto di essere donna è sufficiente per vedere il tuo curriculum vitae scartato dalla posizione, dalla professione di ingegnere.
Questo l'abbiamo scoperto... Il problema è che all'opera ci sono decine, se non centinaia, se non migliaia di algoritmi simili che magari riperpetuano gli stessi pregiudizi senza che noi nemmeno lo sappiamo.
E questo secondo me, dal punto di vista di quello che deve essere il nostro rapporto nei confronti dell'intelligenza artificiale, del tipo di utilizzo che noi facciamo dell'intelligenza artificiale e della conoscenza che dobbiamo avere di quali sono i limiti e gli ostacoli di questi strumenti, invece di riempirci la bocca di stupidaggini come superintelligenza artificiale, rischio esistenziale, intelligenza artificiale generale: tutte cose che hanno ancora, per oggi e per il tempo a venire, più a che fare con il mondo della fantascienza che della realtà, invece di occuparci di questi rischi fantascientifici, dovremmo davvero prestare molta, molta più attenzione a quelli che sono i rischi reali portati da questi strumenti. E ho finito.
Grazie.

IL GIUSTO EQUILIBRIO TRA REGOLAMENTAZIONE E INNOVAZIONE: COME SARÀ IMPLEMENTATO L'AI ACT NEI PROSSIMI 2 ANNI

ASCOLTA INTERVENTO

EMANUELA GIRARDI
Fondatrice Pop AI e Presidente Adra, Associazione europea AI, Data and Robotics

Io sono Emanuela Girardi, sono la fondatrice di Pop AI, Popular Artificial Intelligence, che è un'associazione che si occupa proprio di rendere l'intelligenza artificiale "pop", popolare, accessibile a tutti e tutte per far capire di che cosa stiamo parlando, cosa sono queste tecnologie, questi strumenti, come funzionano, quali sono le grandissime opportunità che abbiamo sentito anche nel corso della giornata, quali sono i rischi, perché c'è sempre questo approccio duale: grandi opportunità, anche grandi o comunque rischi, che dobbiamo conoscere per poterli gestire, poterli mitigare. E poi sono anche la presidente di ADRA, ADRA è l'AI Data Robotics Association, cioè l'Associazione europea di intelligenza artificiale, dati e la robotica ed è partner della Commissione Europea, nel partenariato pubblico-privato su intelligenza artificiale, dati e robotica. Quindi con la Commissione Europea stiamo lavorando per creare un po' l'ecosistema europeo dell'innovazione, cosa molto difficile, perché come vedremo nella presentazione che ho preparato, siamo molto indietro, ahimè, in Europa…
Oggi volevo parlarvi appunto (cercherò di essere breve) della regolamentazione e soprattutto del fatto che sia importante regolamentare proprio per gestire questi rischi che ci sono, ma è

anche molto importante, ed è l'aspetto più complicato, riuscire a trovare il giusto equilibrio tra una regolamentazione che non sia troppo restrittiva e che quindi non limiti l'innovazione. Innovazione che, come vi dicevo, in Europa, ahimè, siamo già un pochino indietro rispetto a Cina e Stati Uniti.

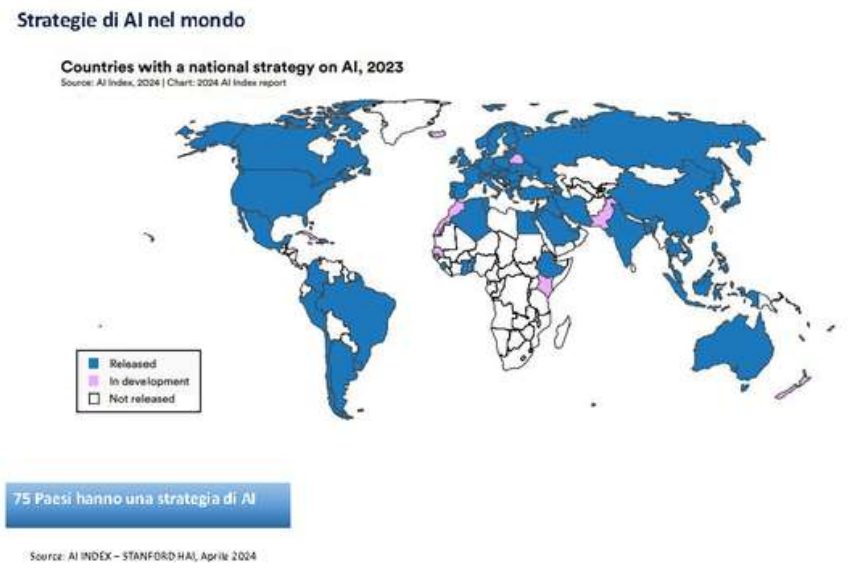
Quindi partirò un po' dal contesto politico internazionale, poi parleremo della visione europea, perché abbiamo parlato giustamente, come ci diceva Giovanna Reanda, anche della visione europea etica, cioè di sviluppare intelligenza artificiale che sia affidabile, sicura e "antropocentrica", cioè che mette l'uomo al centro.

Poi brevemente dell'AI Act, volevo anche darvi qualche informazione su che cosa significa implementare adesso, cioè mettere in pratica l'AI Act, questa regolamentazione europea che regola l'uso dei sistemi di intelligenza artificiale.

Molto brevemente, allora, sull'intelligenza artificiale (se ne è parlato tutta la mattinata e anche nel pomeriggio) volevo solo dirvi che la prima volta che si è parlato di intelligenza artificiale è stato quasi 70 anni fa, nel 1956, negli Stati Uniti.

John McCarthy, Marvin Minsky, Claude Shannon, Nathaniel Rochester, Allen Newell.

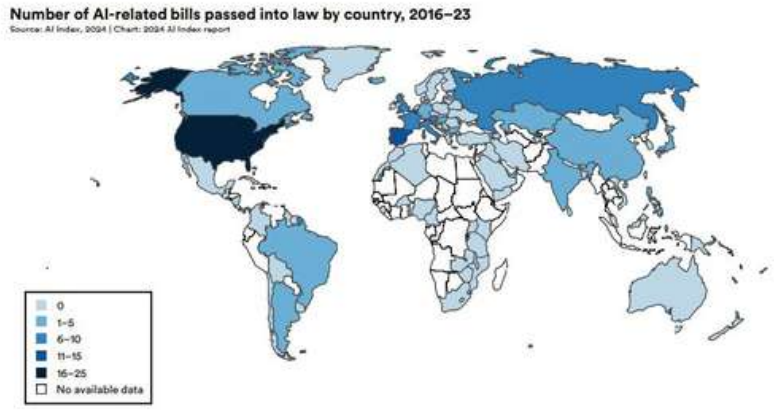
Questi ragazzotti che vedete qua nella foto, che sono considerati padri (ahimé sempre padri tra l'altro, perché purtroppo mai madri, adesso ci sono anche le madri, ma allora c'erano solo i padri…) hanno scritto questo paper dove per la prima volta hanno usato il termine "intelligenza artificiale". E la cosa interessante è che quello che hanno definito come intelligenza artificiale, o meglio l'obiettivo (loro si sono trovati a passare un paio di mesi insieme d'estate) per cercare di sviluppare dei sistemi, delle macchine, che potessero in qualche modo usare il linguaggio per creare delle astrazioni e per risolvere dei problemi complessi che erano problemi riservati agli esseri umani. Quindi c'era già questa visione che è proprio il cuore della visione europea antropocentrica, cioè di utilizzare questi sistemi per migliorare la vita degli uomini, per migliorare, in questo caso, le capacità cognitive degli esseri umani. Quindi 70 anni dopo ritroviamo esattamente la stessa visione che era quella iniziale dei padri fondatori, proprio. Da allora in poi si è arrivati (ci sono stati i cosiddetti inverni ed estati dell'intelligenza artificiale, adesso siamo in una "superestate", cioè tutti parlano dell'intelligenza artificiale) a questi dati pubblicati un paio di settimane fa dall'AI Index, questo studio che pubblica ogni anno l'Università di Stanford, che ci fanno vedere tutti i paesi del mondo che hanno già oggi una strategia nazionale di intelligenza artificiale e sono, al mondo, 75 paesi.



Questo perché è importante? Perché ci fa capire che tutti i paesi industrializzati del mondo hanno capito la priorità strategica che queste tecnologie avranno nello sviluppo della società del futuro e quindi tutti i paesi si sono dotati di una strategia nazionale di intelligenza artificiale. In Italia, non vorrei essere polemica, siamo anche in una sede istituzionale, però non ne abbiamo una, non ne abbiamo due, non ne abbiamo tre, ma stiamo per pubblicare la quarta, c'è chi dice la terza... Ma chi la quarta... Sulla strategia nazionale di intelligenza artificiale non abbiamo mai avuto un piano esecutivo. Adesso con questa legge che è stata attualmente in discussione, forse avremo dei fondi, pochi, speriamo, però almeno qualcosa... Vedremo che cosa si riuscirà a fare. Quindi questo per darvi un'idea di contesto internazionale.
Poi, visto che c'è questo approccio duale, quindi, grandi opportunità ma anche dei rischi, in tutto il mondo si è capito che c'è una necessità di regolamentare questi sistemi. E quindi in

quasi tutti i paesi che hanno una strategia nazionale di intelligenza artificiale, hanno anche presentato e sviluppato delle leggi che sono la maggior parte già in vigore, per regolamentare l'utilizzo di questi sistemi. Quindi, quando la Commissione europea dice: "AI Act, prima regolamentazione al mondo", non è assolutamente vero. Poi dicono "Prima regolamentazione comprensiva", comunque non è vero neanche quello, perché ce ne sono già, qua vedete, una settantina anche in questo caso.

La reazione dei governi: proposte di legge su AI nel mondo

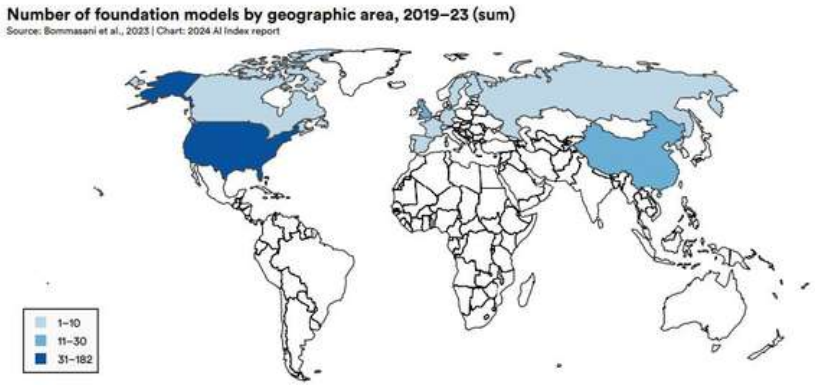


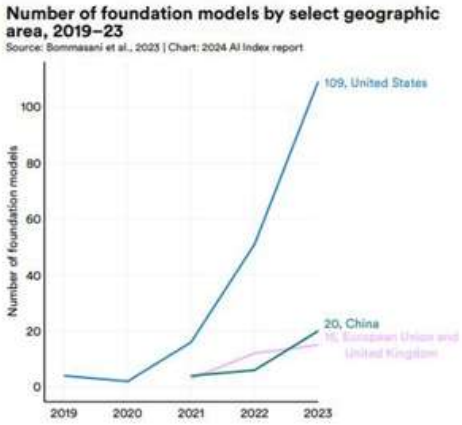
Source: AI INDEX – STANFORD HAI, Aprile 2024

E, soprattutto, vedete che negli ultimi anni, cioè all'inizio dell'anno, c'erano delle regolamentazioni che si sono state realizzate, per aumentare l'utilizzo dei sistemi di intelligenza artificiale sono aumentate moltissimo queste leggi che sono state proposte in tutto il mondo.

Un'altra cosa interessante, invece, si parlava tanto oggi di AI generativa, cioè intelligenza artificiale generativa, che è quella di ChatGPT fondamentalmente, per essere un po' banale. Allora,

quando vediamo lo sviluppo di questi modelli di intelligenza artificiale generativa sulle mappe mondiali, vediamo che la maggior parte dei modelli sono sviluppati negli Stati Uniti e in Cina, pochissimi in Europa, un po' di più in Inghilterra, nel Regno Unito, ma pochissimi in Europa.





Ora, se queste sono le tecnologie che oggi permettono veramente di essere il leader mondiale dell'intelligenza artificiale, se sono sviluppate solo in Stati Uniti, Cina, un pochino in Canada e pochissimo in Europa, capite che i leader di queste tecnologie in questo momento non siamo noi europei. E quindi si può dire, secondo me, che stiamo vivendo una sorta di corsa globale all'intelligenza artificiale, dove abbiamo, i Stati Uniti che in questo momento primeggiano per quanto riguarda l'intelligenza artificiale generativa, la Cina che comunque sta competendo, insomma, a livello molto elevato, e poi abbiamo un "follower", c'è uno che segue, che siamo noi in Europa, che stiamo cercando di seguire, ma in realtà abbiamo molte meno risorse rispetto a quelle che vengono sviluppate e utilizzate dagli Stati Uniti, e quindi stiamo rimanendo molto indietro.

Questi tre modelli che abbiamo sono molto diversi tra loro, perché abbiamo il modello cinese, dove l'innovazione è fortemente, guidata da investimenti pubblici, ed è un modello molto regolamentato anche, perché hanno già tre leggi sull'intelligenza artificiale in Cina, già da diversi anni. Poi abbiamo il modello, invece, americano, che ha vissuto una sorta di "deregulation", cioè di fatto tutto il mondo internet non è mai stato regolamentato. Adesso si sta cercando, in qualche modo, di intervenire, di regolamentarlo con l'Executive Order di Biden e altre iniziative che sono attualmente in fase di realizzazione; però è un modello dove gli investimenti nell'innovazione, sono fortemente guidati dal settore privato e, tra l'altro, sono chiaramente leader nel settore di investimenti nell'intelligenza artificiale, nel numero di start-up, eccetera. Poi abbiamo il modello europeo. Il modello europeo, come vi dicevo, cioè siamo molto indietro nell'intelligenza artificiale, perché quando nel 2016 la Cina ha detto che sarebbero diventati leader mondiali dell'intelligenza artificiale entro il 2030,

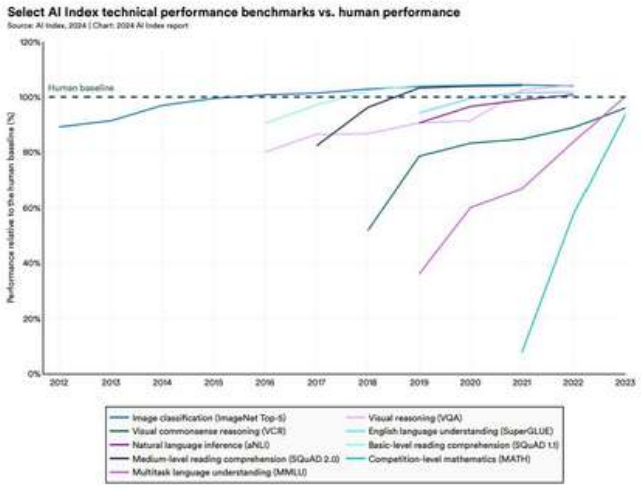
in Europa ancora non facevamo assolutamente nulla. Il primo documento che abbiamo in Europa sull'intelligenza artificiale, risale al 2018, con la pubblicazione nel 2019 del documento che sono "Linee guida etiche per l'intelligenza artificiale (AI) affidabile", che racchiude proprio quella che è la visione, il cuore della strategia europea dell'intelligenza artificiale. E poi sono arrivati tutta una serie di documenti, più che altro regolamentazioni che sono state fatte (infatti ci accusano di essere gli "arbitri" di questa partita globale dell'intelligenza artificiale) perché abbiamo pubblicato l'AI Act, il Data Act, il Digital Services Act, il Digital Markets Act, il Cybersecurity Act, eccetera, eccetera, eccetera, tutta una serie di regolamentazioni che, a mio avviso, sono importanti perché serve regolamentare, ma stanno limitando ulteriormente lo sviluppo dell'innovazione in Europa. E quindi, a mio avviso, sempre ci ritroveremo, sempre di più, a utilizzare tecnologie che sono sviluppate in altri posti.

Però, come avete sentito dai panel precedenti, un conto è su altri tipi di tecnologie (dove la dipendenza è già una dipendenza, insomma, geopolitica, un fattore importante e rischioso), ma dell'intelligenza artificiale è molto importante l'utilizzo, cioè dove arrivano i dati, come vengono preparati i dati per fare l'addestramento di questi modelli, perché questi modelli imparano da realtà differenti dalla nostra, perché sono sviluppati in altri posti, e quindi rifletteranno, nelle decisioni che vengono prese in modo automatico, un contesto culturale che non è il nostro. E quindi questo è un problema, cioè, dal punto di vista proprio culturale, perché non viene rappresentata la diversità culturale che esprimono i vari Paesi europei. Quindi questo è uno dei problemi principali.

Per capire cosa sono questi sistemi è molto utile l'intelligenza artificiale, è molto utile questa definizione che è vecchia, che è la

prima definizione che è stata inclusa proprio in queste linee, nel documento, le linee guida etiche per lo sviluppo dell'intelligenza artificiale, ma è molto completa perché ti spiega veramente che cosa c'è dentro. Questa definizione dice che "l'intelligenza artificiale indica dei sistemi che analizzano il proprio ambiente - cioè imparano dai dati - e compiono delle azioni con un certo grado di autonomia per raggiungere uno specifico obiettivo e, facendo questo, mostrano un comportamento che se fosse svolto da un essere umano noi potremmo definire come intelligente". Ora, quali sono gli elementi importanti di questa definizione? Sono i dati da cui imparano i sistemi e l'obiettivo. Oggi questi due elementi sono completamente sotto il controllo umano, cioè noi decidiamo di utilizzare l'intelligenza artificiale, che sono degli strumenti, come addestrarla, cioè da quali dati farla imparare, e decidiamo per cosa utilizzarla. Quindi anche in campo giornalistico o altro siamo noi che abbiamo il controllo.
Quindi non c'è un'intelligenza artificiale buona o cattiva, ma c'è un uso che si fa del sistema e dello strumento. Questo è un elemento importante.
L'elemento della fiducia è proprio il cuore della visione europea, perché nella visione europea si dice affinché noi, esseri umani, possiamo accettare delle decisioni prese per la prima volta da un sistema automatico, cioè chi prende una decisione in modo autonomo, questi sistemi devono essere affidabili, cioè dobbiamo fidarci. E quindi chiaramente la fiducia diventa l'elemento centrale. Sulla fiducia, giustissima la discussione che c'è stata nel panel precedente, oggi c'è una sorta di polarizzazione totale, perché c'è chi pensa che l'intelligenza artificiale ci ucciderà tutti, e c'è chi pensa invece che l'intelligenza artificiale ci salverà e riuscirà a trovare la soluzione a tutti i grandi problemi complessi del mondo. Ovviamente la soluzione è sempre un po' nel mezzo,

però c'è questa fortissima polarizzazione e, soprattutto in Europa, c'è una narrativa molto negativa che privilegia i rischi rispetto alle opportunità, mentre invece nei paesi asiatici l'opinione delle persone è che le opportunità siano molto più grandi rispetto ai rischi al momento. Quindi è anche interessante la narrativa.
La risposta cos'è? È che oggi noi, l'intelligenza artificiale, in modo inconsapevole, la usiamo tantissimo già tutti i giorni, perché in tutte queste app con cui interagite, tutti i social media, ogni volta che utilizziamo il nostro telefonino e usiamo anche solo il riconoscimento facciale per sbloccarlo, stiamo usando intelligenza artificiale. Quindi già tutti i giorni noi la stiamo usando.
Questo grafico, molto importante, volevo farvelo vedere, perché già oggi, anzi già qualche anno fa, l'intelligenza artificiale ha raggiunto dei livelli di prestazione che superano le capacità umane in tantissimi compiti.



Source: AI INDEX – STANFORD HAI, Aprile 2024

Quindi se noi prendiamo per esempio la classificazione delle immagini, la comprensione della lettura di base, il ragionamento visivo, il linguaggio naturale, gli algoritmi specializzati in quel singolo campo sono già oggi molto migliori dell'essere umano.
Chiaramente non c'è un'intelligenza artificiale superiore, generale, che sa fare tutte quelle cose. Ma i singoli algoritmi sono già più bravi di noi nello svolgere il singolo compito. Quindi questo è un elemento anche importante. Non ci deve spaventare, ma comunque ci deve far riflettere che quindi già la stiamo usando tutti i giorni e già è migliore di noi in moltissimi campi.
Perché è importante quindi l'affidabilità e che cos'è la visione europea dell'AI affidabile?
Vuol dire tre cose "affidabile", secondo la visione europea. Vuol dire che deve rispettare le leggi, chiaramente le leggi europee e i diritti della Carta fondamentale dei diritti europea; deve essere etica, cioè deve rispettare i principi etici fondamentali, anche qui, quelli europei, che chiaramente sono diversi da altri principi etici in altre parti del mondo; e deve essere sicura e robusta dal punto di vista sia sociale che tecnico.
Ora, l'intelligenza artificiale per essere etica deve rispondere a questi quattro principi che sono molto importanti. Io ve li cito brevemente perché altrimenti ne parleremmo almeno per tre o quattro ore, penso che nessuno potrebbe resistere... Molto velocemente sono: il primo, rispettare l'autonomia dell'uomo, cioè l'uomo deve sempre potersi sostituire a una decisione presa in modo automatico da un sistema di intelligenza artificiale; il secondo è che non deve creare danno all'essere umano: sembra un po' tipo le leggi di Asimov, per chi le conosce, della robotica, cioè il danno tra l'altro sia fisico ma anche danno psichico, quindi la manipolazione, la disinformazione, eccetera; il terzo è il principio più complesso e ne ha parlato molto bene Andrea

Signorelli prima, quello dell'equità, delle discriminazioni, cioè il fatto che tutti quei bias, quei pregiudizi che avete visto nei contenuti generati dall'intelligenza artificiale bisogna poterli riconoscere e poterli gestire e, chiaramente, contrastare; e l'ultimo è quello della "spiegabilità", cioè probabilmente avete sentito il concetto di scatola nera, di black box dell'intelligenza artificiale, che significa che questi sistemi, a volte, sono talmente complessi, come per esempio le reti neurali profonde, che è quasi impossibile riuscire a capire come hanno preso una decisione in modo automatico; ora, se la decisione era comandarmi una strada piuttosto che un film, anche se io non so come l'hanno presa, insomma posso anche vivere serenamente, ma se pensiamo a decisioni prese per esempio in campo medico o magari in campi di utilizzo nei sistemi democratici, chiaramente questo non può succedere. E quindi la spiegabilità è anche un elemento molto importante.
Ora, per poter essere etica e rispondere a questi quattro principi, questo è molto complesso. E quindi ci troviamo di fronte a dei sistemi che hanno una sorta di opacità, c'è poca trasparenza, sono imprevedibili, agiscono in modo autonomo, utilizzano dei dati che magari sono dei dataset che non sono completi o che hanno dei pregiudizi. Quindi tutti questi elementi danno origine a una serie di rischi. Questi rischi devono essere conosciuti e gestiti ed è per questo che serve una regolamentazione. Ora, la regolamentazione europea è tutta fondata proprio sulla categorizzazione dei rischi che sono associati all'utilizzo di questi sistemi. E sull'AI Act, dove la Vestager che dice che quando si parla di intelligenza artificiale la fiducia è un "must", c'è un qualcosa di necessario e non semplicemente un qualcosa di accessorio e, quindi, proprio partendo anche da questa visione della Vestager, è stato creato poi l'AI Act che divide tutti i sistemi

di intelligenza artificiale in quattro categorie, in base al rischio che hanno di poter creare un danno agli esseri umani, ai cittadini europei o di violare quelli che sono i principi fondamentali della Carta dei diritti europea.
Ora, le quattro categorie di rischio molto velocemente sono: se il rischio è considerato inaccettabile, i sistemi verranno vietati: e qui rientrano tutti i sistemi, per esempio, della categorizzazione biometrica remota in tempo reale, quindi il riconoscimento facciale in una piazza durante un evento; questa cosa sarà vietata, non sarà possibile farlo in Europa, tranne dei casi particolari di utilizzo da parte delle forze dell'ordine, se si è per esempio in presenza di un allarme terroristico oppure di un reato contro minori; questo tra l'altro è stato molto dibattuto perché il governo francese voleva poter utilizzare il riconoscimento facciale in luogo pubblico durante le Olimpiadi e tra l'altro lo utilizzeranno perché non è ancora entrato in vigore l'AI Act, quindi sarà utilizzato; poi ci sono le categorie ad alto rischio: prima per esempio Andrea citava l'algoritmo utilizzato da Amazon per categorizzare tutti i curriculum delle persone, quindi preselezionare i curriculum in fase di selezione del personale: ecco, quello è uno degli esempi di sistema ad alto rischio, quindi quei sistemi che per poter essere utilizzati nel mercato europeo che dovranno rispondere a tutta una serie di requisiti che sono molto, molto stringenti e un altro sistema tra l'altro di utilizzo ad alto rischio è il riconoscimento delle emozioni, per esempio, per quanto riguarda l'ambito del lavoro o il campo medico, oppure l'utilizzo di sistemi di infrastrutture strategiche come nei trasporti, nell'energia o altro; invece i sistemi che avranno un livello di rischio più basso dovranno rispondere solo a dei requisiti di trasparenza, o a un codice di condotta.
L'AI Act è stato ufficialmente approvato il 13 marzo del 2024 e

appena verrà pubblicato nella Gazzetta (che si spera nel giro di un mese e mezzo), a quel punto partiranno i 24 mesi entro cui entreranno in vigore a distanza di 6-12 mesi tutti i vari divieti. Entro 6 mesi entreranno in vigore i divieti per i sistemi che sono considerati appunto a rischio inaccettabile, poi entro 12 mesi entreranno in vigore invece i requisiti per i sistemi come i Large Language Model, quindi come un ChatGPT o altri modelli, ed entro 24 mesi i sistemi considerati ad alto rischio.

Che cosa ci dice la von der Leyen? Ci dice che bisogna sì regolamentare, ma bisogna anche governare e innovare.

Ora, in realtà in questo momento in Europa stiamo più regolamentando che innovando, devo dire, e quando stiamo innovando, stiamo innovando in un modo abbastanza limitato. Perché ho questa visione un po' critica nei confronti dell'AI Act, soprattutto per l'innovazione?

Perché per un'azienda oggi, che è in Europa (non pensare alle big tech, perché tanto le big tech hanno comunque le risorse per gestire tutta questa regolamentazione), per un'azienda europea, una SMI (perché la maggior parte delle nostre aziende sono piccole e medie aziende) che vuole inserire nel mercato europeo, sviluppare e vendere un sistema di intelligenza artificiale che magari è considerato ad alto rischio, i requisiti che deve soddisfare per essere compliant all'AI Act sono talmente tanto costosi e difficili che questo potrebbe effettivamente limitare l'innovazione. Certo che il Parlamento Europeo dice: "Ma no, noi vogliamo creare un mercato sicuro, il mercato unico europeo dell'innovazione".

In realtà adesso vi faccio vedere qualcuno dei requisiti che dovranno soddisfare. Ecco, questi sono i requisiti di compliance di sistemi di intelligenza artificiale considerati ad alto rischio.

Quindi devono fare il risk management, la data governance, avere tutta la documentazione tecnica, avere sempre il controllo umano, l'accuratezza, la robustezza, la sicurezza, la qualità dei dati, la dichiarazione di conformità, il marchio CE, controllare il sistema durante tutto il ciclo di vita, perché poi è un sistema che continua ad apprendere, quindi durante tutto il ciclo di vita devono continuare sempre a fare tutti questi controlli, eccetera. In più registrare il sistema in questo database europeo dei sistemi di intelligenza artificiale ad alto rischio. Ora, tutte queste cose, si dice, che avranno un costo molto, molto elevato. C'è uno studio (adesso vi faccio vedere qualche dato velocemente) che dice che comunque il costo della compliance di un sistema di intelligenza artificiale considerato ad alto rischio, può arrivare a 500.000 €. Ora, pensate a una start-up che magari ha appena fatto un round di finanziamento di un milione, se deve spenderne cinquecentomila per essere compliant, chiaramente è fuori mercato prima di iniziare. In più, tra l'altro, le multe sono altissime: si arriva fino a 35 milioni di fatturato in base alle

violazioni, oppure a una percentuale che comunque sono troppo alte per la caratteristica delle piccole e medie aziende che abbiamo in Italia e in Europa.
Le ultime due cose e concludo.
Questo per darvi un'idea che il costo dell'innovazione è molto, molto elevato. Uno studio ci dice che due terzi delle aziende prevedono che avranno un impatto molto negativo sull'innovazione, questa regolamentazione che è molto, molto costosa, soprattutto per quanto riguarda la data governance. E quindi è veramente difficile, secondo me, riuscire a trovare questo giusto equilibrio tra regolamentare da una parte, che importante serve la regolamentazione, e soprattutto per innovare. Perché in questo momento noi in Europa stiamo già indietro, molto indietro, rispetto agli Stati Uniti e alla Cina. Abbiamo, pensate, una dipendenza tecnologica oggi dell'85%. Se ricordate la dipendenza del gas un paio d'anni fa, ecco, l'intelligenza artificiale, che sarà sempre di più in qualsiasi cosa che noi utilizziamo tutti i giorni, oggi è sviluppata sopra l'85% negli Stati Uniti. Quindi siamo ancora più dipendenti oggi.
Quindi è molto importante riuscire a trovare questo giusto bilanciamento tra regolamentazione e innovazione.
Vi ringrazio e spero di non essere stata troppo lunga.
Grazie.

GOVERNARE L'INTELLIGENZA ARTIFICIALE: UNO SGUARDO VERSO IL FUTURO

ASCOLTA INTERVENTO

GIANLUCA MISURACA
Direttore del Master AI4Gov,
Universidad Politecnica de Madrid e
Politecnico di Milano

Buonasera e innanzitutto grazie per l'invito.
È un onore essere qui e, tra l'altro, complimenti alla Fondazione Marco Pannella e agli organizzatori per la ricchezza del dibattito.
Peraltro, appunto, era ieri il compleanno di Pannella, quindi insomma, credo che sia anche doveroso ricordarlo. Ed effettivamente, dopo le presentazioni che sono state fatte oggi e il dibattito così ricco, ci sarebbe ancora tanto da dire, però forse è opportuno anche ricordare quello che non abbiamo ancora detto, perché viene quasi automatico, dopo Emanuela, effettivamente, entrare un po' in quello che non sappiamo ancora, e che è appunto il futuro.
Del resto abbiamo parlato (ed essere qui, insomma, nella culla della democrazia e della politica, insomma…) molto di politiche, di aspetti tecnici, però non tanto (anche si ci ha provato un po' la moderatrice, di portare il dibattito sulla politica e appunto sull'aspetto della democrazia e i rischi enormi che ci sono effettivamente, ma anche le opportunità, anche in questo ambito).
Ora, governare l'intelligenza artificiale, insomma, è ovviamente una domanda retorica, una cosa che non è possibile, probabilmente, ma ci dobbiamo provare, insomma, in qualche modo. E ovviamente nella mia presentazione, come si dice, c'è quella battuta famosa del professore che arriva a una conferenza

e gli viene chiesto se ha delle slide. Dico: "Sì, ho delle slide, ma ho anche qualcosa da dire, insomma".
Quindi, al di là delle slide, diciamo, per guidare un po' il dibattito, volevo effettivamente portare l'attenzione su tre aspetti. Il futuro, diciamo, della governance del digitale, al di là dell'intelligenza artificiale, di cui forse stiamo parlando anche un po' troppo, perché effettivamente, diciamo, è un tema molto sexy, interessante, eccetera, però, a parte che spesso ne parliamo a sproposito, senza capire veramente di cosa stiamo parlando, e questo vale in particolare per i politici e per i policy makers, per chi poi deve regolamentare tutto questo. Del resto non è un caso, parlando di evidence e di fatti, che usciamo da una legislatura del Parlamento europeo che è la più ignorante, nel senso di quella che ha avuto i parlamentari meno preparati di sempre. E purtroppo, come Italia, abbiamo dato un bel contributo a portare persone molto poco preparate. Speriamo nella nuova infornata di parlamentari europei: ce ne sarà qualcuno che sa almeno dove deve andare a Bruxelles o a Strasburgo…
Vedere il mondo dopo l'AI Act.
Effettivamente siamo ancora all'inizio, questa è una proposta, e vedremo, è stata illustrata molto bene, anche da Emanuela, molto velocemente. Io lavoravo in commissione, quando si ragionava se fosse opportuno o meno, proponendo una regolamentazione su questo tema, e adesso vediamo cosa ne è uscito, insomma.
Ora, quindi, il futuro della governance digitale, ovviamente, essendo un futurista, so molto bene che non sappiamo qual è il futuro. Ma sicuramente, quando, per esempio, ancora lavoravo al Centro comune di ricerca della Commissione europea, di tanto in tanto, facevamo degli studi di foresight e l'ultimo che ho fatto, nel 2020, era pronto per capire e dare, quindi, dell'input proprio ai colleghi di Bruxelles su quali decisioni prendere...

Sembra quasi fatta apposta la domanda con cui Emanuela ha concluso. Regolamentare o innovare? Come facciamo a fare questo balance?
Abbiamo scelto, con vari esperti e colleghi, consultazioni varie, basandoci anche sull'evidenza che dopo 20-30 anni, effettivamente, abbiamo raccolto (anche se poi spesso siamo un po' ciechi e non guardiamo, effettivamente, laddove c'è un minimo di evidenza, alle conseguenze di certe scelte; però è vero che in ambito tecnologico è complicato, perché sono sistemi, appunto, complessi e, sicuramente, la relazione, poi, fra uomo e macchina è ancora una cosa che non conosciamo bene, come vedremo), abbiamo definito, in questo caso, due dimensioni di impatto.
Uno è, effettivamente, il "regulatory landscape", il panorama della trasformazione digitale: se avere un approccio più liberale, laissez faire, o più, invece, interventista. E lì c'è stata una grossa battaglia, come sapete, fino alla lettera famosa del 7 dicembre, firmata da Francia e Germania, e poi, a caso, anche dall'Italia, senza saper bene perché. Mentre, invece, dall'altro lato, la cosa che è emersa oggi molto chiaramente è la cittadinanza digitale: siamo noi, o i cittadini in generale, capaci, attualmente, effettivamente, di gestire le piattaforme, l'informazione, la conoscenza? Quando si parla di sovranità digitale, quando Breton si riempie la bocca di sovranità digitale, peraltro, prendendo spunto dal famoso discorso della Sorbonne di Macron, parla, in realtà, di una sovranità digitale europea. Non è solo quella italiana, lussemburghese o cipriota. Altrimenti, come diceva, appunto, Emanuela, non andiamo a nessuna parte. Siamo piccoli e, se ci siamo, siamo fragmentati, siamo ancora più minuscoli.
È lì che bisogna andare oltre il GDPR.
Uno di questi scenari che abbiamo dipinto nel 2020, in piena pandemia, peraltro, era quello in cui viviamo oggi: uno scenario di

un futuro apatico, di innovazione chiusa, che se io ho un iPhone, non posso neanche parlarmi quasi con un altro sistema perché sono mondi chiusi… E noi siamo lì a ticcare i box del GDPR, sperando poi che i nostri dati, in qualche modo, siano, insomma, rispettati… Oppure, non so se avete letto il libro di Cass Sunstein, quello meno famoso, "The Law of Fear", pubblicato nel 2005 e ripubblicato nel 2020, con un grosso successo, perché effettivamente diceva, ed è esattamente la situazione che abbiamo vissuto per il Covid e nel caso dell'intelligenza artificiale, che se non sappiamo quali saranno le conseguenze di qualcosa, allora forse è meglio non far nulla, perché magari così non sbagliamo.

Ovviamente non possiamo non far nulla nel caso dell'intelligenza artificiale, perché è già parte delle nostre vite. Ovviamente possiamo cercare, come è stato presentato dai colleghi precedentemente, di intervenire su questi sistemi, riducendo i bias e migliorando la qualità dei dati, che è un altro scenario, ovviamente, in cui magari riduciamo un po' l'impatto innovativo. Però è vero pure che (e questo, dal punto di vista europeo, uno dei grossi "act two", se vogliamo, della costruzione stessa dell'Unione Europea) noi abbiamo, fino ad ora, avuto un minimo di successo nell'innovare attraverso la regolamentazione, che è un approccio specifico che, come per esempio il GDPR, con tutti i limiti che ha, è effettivamente, di fatto, uno standard a livello mondiale.

E, ovviamente, sappiamo bene che siamo solo all'inizio di questa trasformazione epocale e gli scenari, diciamo, sono degli strumenti per cercare di dare un po' delle possibili indicazioni sul futuro, ma ovviamente non saranno né uno né l'altro. Lo scenario ideale, che è quello dell'open innovation, di un'innovazione appunto in cui noi possiamo, come diceva appunto Tim Berners-Lee, che è stato menzionato stamattina, quando appunto ha

inventato il World Wide Web, è quello in cui noi abbiamo la possibilità di scambiare informazioni e conoscimento senza limiti. Ed era l'idea originaria di internet. Chiaro, adesso, non è più così perché effettivamente ci sono tanti ragazzi e ragazze cattive nel web, quindi bisogna fare molta attenzione.
È vero che non sappiamo quale sarà il futuro, ma sappiamo molto bene qual è la realtà attuale e in questa slide, senza volerlo, in realtà, penso sia sintetizzato un po' il senso di tutto quanto è stato detto oggi. Nel senso che il futuro è già qui, come diceva William Gibson nel '99, è solo "unevenly distributed", non è distribuito in modo uguale, però sappiamo bene che ci sono degli enormi miglioramenti di queste tecnologie, specie appunto nel riconoscimento facciale per esempio, nelle immagini, in ambito medico, capacità quasi che mimano l'intelligenza umana, ma abbiamo visto invece molto dei rischi, i rischi di discriminazione, le possibili morti… Mi faceva un po' quasi paura quanto è stato detto oggi il fatto che se una Google Car si sbaglia e uccide un ciclista, vabbè ce ne sono tanti di ciclisti... Cerchiamo un po' di essere magari… Ci sono una serie di questioni che vanno affrontate proprio perché sono macchine che noi come essere umani dobbiamo controllare.
E poi tutto il discorso del deep fake che effettivamente possono e hanno un impatto enorme in democrazia, perché vedere Obama che dice una cosa può far cambiare l'idea a chi magari voleva votare Obama. E quindi bisogna fare molto attenzione. Ed è per questo che il DSA con tutti i limiti che sono stati menzionati, è per fortuna è già in forza, anche se non avrà l'effetto che si sperava, perché siamo troppo vicini alle lezioni e probabilmente nessuno sta veramente prendendo sul serio questa roba.
Ora per passare velocemente al secondo punto che è questo rapporto tra uomini e macchine, specie dagli amici ingegneri e

computer scientists si sente spesso dire: "Ah! Il Large Language Model è come la nuova printing machine"... La stampante. ù
Quindi è chiaro che anche questa analogia lascia il tempo che trova, perché in realtà è un po' come quando, nel 1786, credo '78, venne presentata a Maria Teresa d'Austria, il famoso "Mechanical Turk" in cui le si diceva: "Guardate che c'avete una macchina intelligente che gioca a scacchi da sola". In realtà dentro c'era una persona che muoveva le leve.
Ed è un po' quello che succede oggi, perché poi al di là del dibattito e della retorica sulla "Super-AI" eccetera, poi ci sono anche degli sfruttamenti anche del lavoro per esempio in certi paesi dove le persone stanno lì a vedere se il gatto è gatto o meno eccetera, però al di là di quello c'è comunque ancora la necessità di intervenire come esseri umani, anche per poter controllare o cercare, quantomeno, di capire se quello che viene fuori come risultato di un algoritmo effettivamente è vero o meno.
Poi sul fatto che vero o falso sia un'opinione questo va discusso, insomma.
Emanuela ci ha portato al 1956 quando si coniò il termine "AI"; io vado ancora un po' più indietro, quando appunto Alan Turing, nel 1950, aveva emesso la prima legge di Turing, in cui si ipotizzava la possibilità di avere delle macchine che fossero intelligenti quanto gli uomini: il famoso Turing test. Ed era peraltro lo stesso anno in cui Asimov scrisse "I, Robot", il libro che immaginava quello scenario di cui potremmo essere presto vittime, insomma le macchine che prendono il sopravvento.Ovviamente questi sono scenari un po' fantascientifici, scientifici. Però è vero che oggi, quando entriamo sul web, insomma, siamo noi che dobbiamo dimostrare di essere umani invece che la macchina nel rispondere al nostro comando.
Quindi forse stiamo andando troppo in fretta? Siamo ora in un

periodo in cui le macchine potrebbero effettivamente prendere il controllo?
È vero che questi Large Language Models, che vanno un po' in tutte le direzioni, rischiano di portarci in qualche modo in quella direzione. Non vado nei dettagli tecnici. Però voglio sottolineare un aspetto che spesso viene dimenticato e, anzi, va dato atto che infatti invece il governo italiano, non solo in questa strategia, ma in realtà anche nella strategia scritta con il governo Draghi, in particolare con il Ministro Colao (a cui io sono particolarmente attaccato perché ero citato ovviamente), dove c'è anche la questione del G7, dove l'Italia ha proposto infatti che si faccia un focus speciale o si dia un'attenzione particolare al ruolo del governo.
Cioè l'intelligenza artificiale nel settore pubblico.
Cosa che io ricordo nel 2018 proponevo a Bruxelles... Quasi mi ridevano dietro nel senso di dire: "No, ma il settore pubblico che c'entra, questa è una roba per le imprese, il settore pubblico al limite poi viene e usa le cose che gli vengono date dai vari...". In realtà ci siamo resi conto dopo 5 anni che è molto importante e fondamentale perché il settore pubblico, il governo, insomma ha vari ruoli: un ruolo molteplice da giocare dove non solo deve definire le regole e già lì è un problema, perché poi se abbiamo le persone che non capiscono di cosa si tratta, fanno le regole sbagliate. Poi c'è l'aspetto che usano queste tecnologie nei governi e in certi stati, in certi paesi, per esempio in Italia, si spende tantissimo in procurement, in acquisti pubblici, in tecnologie e, se compriamo delle cose che o non ci servono o che creano solo spesa pubblica, forse sarebbe meglio evitare e tornare alla penna e alla carta… E poi, però, hanno un altro ruolo, quello di essere un po' la piattaforma per il settore privato, per stimolare poi l'adozione nella società ed essere poi

eventualmente, in questo caso, effettivamente un po' il pioniere perché avendo tante capacità anche di spesa, e avendo dei processi dove l'intelligenza artificiale può veramente fare la differenza, mentre le start-ups, le società private vanno e vengono, il governo dovrà poi essere rimanere e quindi è bene che faccia le cose fatte bene, magari usando le tecnologie quando servono in maniera corretta.

Per chiudere questa parte, io parlo da tempo del dilemma che hanno i policy makers di usare queste tecnologie per migliorare i servizi ai cittadini, i processi gestionali interni dell'amministrazione e, a limite, anche quelli di policy making, che è un altro tema su cui possiamo anche parlare, ma hanno però l'obbligo di proteggere i cittadini, perché appunto i diritti umani e quelli fondamentali di cui ci parlava questa mattina la collega non sono una cosa che possiamo decidere se rispettare o meno. In Europa vanno rispettati.

Ed è per questo che, come dire, l'obiettivo, il "sacro graal", se vogliamo, è quello di governare l'intelligenza artificiale, quindi fare delle regole che ovviamente siano poi gestibili e non blocchino l'innovazione tecnologica, ma governare poi "con" l'intelligenza artificiale, quindi anche sperimentare e usare anche i famosi sandbox, di cui si parla tanto adesso, per capire quali sono le regole e le tecnologie effettivamente da usare nel settore pubblico, che è un luogo importante di sperimentazione; ma anche poi dove sta la bellezza di questo set di tecnologie (ma anche il rischio maggiore)? È in qualche modo lasciarsi andare: usare queste tecnologie per fare esattamente quello che fanno meglio di noi limitati esseri umani, però avendo una supervisione dell'impatto che potrebbe essere ovviamente molto pericoloso, specie se si parla di democrazia.

Ora, abbiamo fatto l'AI Act. Tutto a posto...

Spero che i miei colleghi, ex colleghi, non mi sentano…
Insomma, è chiaro che quando si immaginò il pacchetto AI che venne poi varato il 21 aprile del 2021 (un po' in ritardo peraltro, era già pronto prima nel 2020, poi è stato ritardato paradossalmente dalla Germania, che era alla presidenza di turno, e che voleva le regole ancora più stringenti, adesso è quella che critica; un po' strana sta cosa... però, beh, meglio che si cambi idea), la cosa spesso che ci dimentichiamo è che (il pacchetto AI) era composto di due gambe. Una era la parte regolamentare (e vabbè, più o meno siamo arrivati, adesso ci vediamo in che modo) e l'altro era il piano di azione coordinato. Cioè, ci si era resi conto che noi non spendiamo: l'idea era di spendere almeno 20 miliardi di euro all'anno, a livello europeo, per iniziare a essere competitivi con le altre regioni del mondo. Non ci parliamo, e se ci parliamo ci facciamo la guerra. Però si diceva: "Vabbè, proviamo a mettere insieme le risorse e magari faremo qualcosa…". Ora, la prima versione del piano coordinato d'azione (noi facciamo la review di questi piani nazionali), era, insomma, una roba in cui, come si dice in public policy, era "multisermons", cioè, molte parole, ma non c'era niente di, né incentivi, né, tanto meno, come dire, di "sticks"... Se non lo fai, ti penalizzo.
La versione successiva era, perlomeno meglio, quella che è uscita poi, appunto, nel 2021, però, ad oggi, manca una vera politica industriale europea, che prenda in considerazione l'intelligenza artificiale in maniera seria.
Ed è per questo che, peraltro, sul lato dell'AI Act, il 13 marzo (appunto tutti contenti quando ci si era arrivati: io l'ho vista un po' come le idi di marzo), in realtà, è stato sacrificato, se vogliamo, il principio fondamentale alla base dell'AI Act, semplicemente nel modificare, non tanto la definizione di intelligenza artificiale, ma nel non voler definire l'intelligenza artificiale, o meglio, nel voler

dire che l'intelligenza artificiale generativa, che è una cosa diversa, non può essere definita perché non sappiamo ancora cosa sia, il che, ovviamente, svilisce un po' il senso tutto dell'impianto dell'AI Act, che era pensato, inizialmente, come una regolamentazione a prova di futuro (il che, ovviamente, potrebbe anche essere vista come una contraddizione in termini).
Quando Giulio Cesare venne assassinato, in realtà, si creò, poi, la Pax Augustea. Quindi, io mi auguro, da, diciamo, perenne ottimista, che adesso, in realtà, dopo il mondo, diciamo, "post AI-Act", ci sia un mondo in cui si inizierà a collaborare. Certo, poi vediamo i casi di Mistral, ed altre start-up, che prendono tanti fondi dai governi europei e, poi, se ne vanno, insomma, con, diciamo, con l'avversario, per non dire nemico.
Peraltro, c'è un altro tema di cui, spesso, ci dimentichiamo, che, in realtà, parliamo molto di human-centric AI, intelligenza artificiale antropocentrica, che è un concetto bellissimo, ma, in realtà, non è definita, non c'è una definizione. Io ho gestito per tre anni un progetto che si chiama International Outreach for Human-Centric Artificial Intelligence per la Commissione Europea e tutti mi chiedono ma cos'è lo Human-Centric AI? Eh, bella cosa...
Abbiamo fatto ovviamente uno studio che era piaciuto moltissimo alla Vestager, ma Breton ce l'ha censurato. Speriamo non mi senta… No, Breton lui, insomma, per carità. Però diciamo tutto l'impianto, perché era un po' contrario a quello che invece si voleva proporre…
Ora, per fortuna invece nelle call Zoom chiuse tra Stati Uniti e Unione Europea, grazie ai colleghi del mio team di InTouchai.eu, siamo arrivati a definire 65 termini con gli Stati Uniti nell'ambito del Trade and Technology Council, fra cui abbiamo infilato dentro Human-Centric AI. E il mese scorso è uscito questo documento, che ovviamente non è popolarizzato perché è una roba per gli

addetti ai lavori, in cui c'è la definizione di Human-Centric AI. Però la cosa interessante qui è che si affronta (ed è una cosa su cui i miei colleghi e io abbiamo insistito molto) la questione dei diritti.
Questa è una traduzione fatta appunto dall'intelligenza artificiale, quindi magari non fa fede, però i valori democratici e dello sviluppo sostenibile, che sono due aspetti fondanti peraltro della visione europea del futuro, sono messi lì. Il che ovviamente quando poi si andrà a discutere nei consessi internazionali, dove non necessariamente la parola "democrazia" viene usata, perché se guardate ai principi etici dell'UNESCO (io collaboro anche con l'UNESCO, quindi insomma, per carità) non si parla di democrazia, si parla di "armonia" magari, o altri temi che vengono da altre filosofie, da altre questioni. In Europa per il momento ancora abbiamo dei principi democratici che andrebbero rispettati.
Ora, l'ultima domanda (poi c'è una mini risposta, ma in realtà è solo mini) è se siamo pronti. Come evitare questi possibili futuri distopici.
E ora qui ovviamente alcune delle cose sono emerse anche oggi, ma sicuramente c'è da stenersi dal tecno-soluzionismo, questa è una roba che io dico da tanti anni, insomma, perché ovviamente c'è un approccio molto tecno-determinista, tecno-centrico, che spesso dice: "Vabbè, questo è un tema per gli informatici o per i tecnologi...". No, questo è un tema per tutti: se il policy maker non sa di che sta parlando, è un problema. Non ci sono scorciatoie etiche, non possiamo semplicemente chiamare Luciano Floridi e dirgli, vieni. Abbiamo il miglior filosofo nella stanza, abbiamo risolto il problema. No, bisogna veramente integrare i principi etici nel disegno dei sistemi, tra cui anche l'intelligenza artificiale. Perché l'utilizzo sbagliato di queste tecnologie hanno degli effetti profondamente distopici e profondamente negativi, e ci sono tanti casi, purtroppo, che dimostrano come questi sistemi possono

essere discriminatori, anche senza volerlo. Il famoso caso Siri, in Olanda. Non è che gli olandesi erano razzisti e hanno fatto un sistema razzista. No, hanno fatto un sistema che ottimizzava l'algoritmo per cui, se c'erano delle frodi fiscali, loro intervenivano per ridurre ed eliminare i benefit. Risultato: erano tutti, diciamo, immigranti o persone povere che venivano private del loro sostentamento. Perché? Perché il sistema era fatto basandosi su degli indicatori assolutamente poco utili. Se non erano andate dal medico tot volte all'anno e cose di questo tipo…

Ora, bisogna poi avere una prospettiva di valore pubblico, che spesso anche ci dimentichiamo, perché o cerchiamo di creare il miglior sistema tecnologico, che però poi non serve a nulla, o perché effettivamente magari vogliamo avere degli impatti sulla produttività.

E poi, fondamentalmente, quando integriamo questi sistemi, soprattutto nel settore pubblico, ma in generale nella società, bisogna essere anche pronti a dei cambiamenti radicali. Perché non è che si può dire: "Vabbè, uso un po' di nuove tecnologie e poi...". No, o cambi o cambi, perché altrimenti è meglio che non lo fai, perché rischi di fare cose ancora peggiori.

Potremmo ovviamente entrare molto più in dettaglio su questi temi, ma bisogna fondamentalmente, a mio modo di vedere, disegnare dei nuovi modelli di governance, che non abbiamo ancora a disposizione. Ed è per questo che anche i policy makers, sia politici che gestori della cosa pubblica, della res pubblica, devono essere pienamente consapevoli di cosa significa avere dei processi decisionali algoritmici, per esempio, e anche come i servizi devono essere ridefiniti e come usare i dati.

Chiudo quindi facendo un po' di... in realtà non pubblicità, ma semplicemente menzionando il master (perché non mi sono presentato, ma io sono qui in particolare nella mia veste di

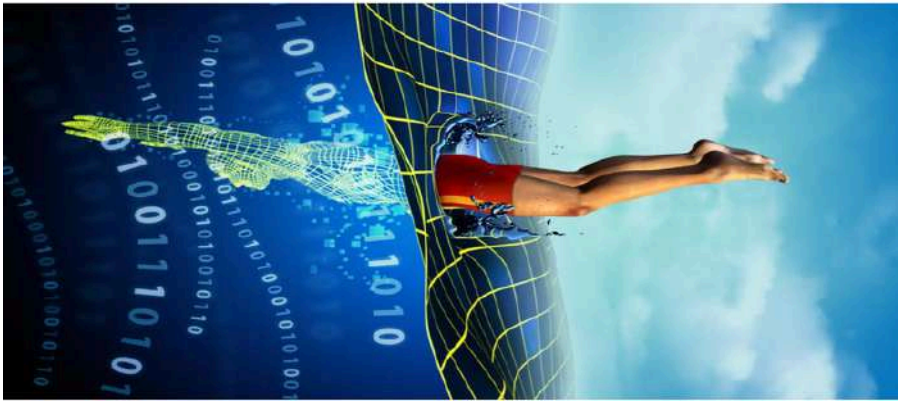
accademico)... perché sono il direttore del master sull'intelligenza artificiale per il governo, per i servizi pubblici, che è un progetto cofinanziato dall'Unione Europea e implementato, leaderato dalla Università Politecnica di Madrid e dal Politecnico di Milano. Abbiamo poi partner in varie nazioni: in Estonia, in Germania, in altri paesi, dove facciamo (e per quello che abbiamo avuto un certo successo), quello che non è fatto da altri, perché ci sono tantissimi corsi (fantastici, ovviamente, più focalizzati sugli aspetti tecnici, giuridici ed altro) in cui noi mettiamo insieme persone che parlano lingue diverse: ingegneri, computer scientist, con persone che lavorano nell'amministrazione pubblica, organizzazioni internazionali, eccetera, e gli diamo, nella parte del programma, sia nozioni e competenze, appunto, sull'intelligenza artificiale da un punto di vista tecnico, ma anche di service design, quindi di design thinking, su come effettivamente riformulare certi servizi e certe politiche. Ed è per questo che, effettivamente, lavorando con dati e su casi concreti, poi ci si può render conto della complessità del tema.

E poi, come vi menzionavo, lavorando con UNESCO, al momento, stiamo cercando di portare questo ad un livello ancora più ampio, per poter accelerare, come piace a me dire, lo sviluppo e l'adozione dell'intelligenza artificiale e della trasformazione digitale nel settore pubblico. Ovviamente, è fondamentale, per questo, rafforzare le capacità dei policy makers, dei leader del settore pubblico, perché altrimenti possiamo fare tutti i disegni di legge che vogliamo, però non credo che questo risolva la situazione. Bisogna veramente far sì che ci sia un approccio comprensivo, non solo a livello nazionale, nel caso dell'Italia, ma a livello europeo, e poi, soprattutto, iniziare a considerare, in maniera seria, la governance globale dell'intelligenza artificiale.

Vent'anni fa io ero a Ginevra e a Tunisi, quando ci fu il famoso

World Summit on Information Society e perdemmo un'occasione storica per cui, in realtà, la governance dell'Internet è rimasta quella che fu immaginata quarant'anni fa e non va bene. Ed è chiaro che non possiamo permetterci di fare lo stesso errore adesso e gli strumenti, come quello proposto dal segretario generale delle Nazioni Unite con l'UN Advisory Board, lasciano il tempo che trovano. Quindi, sicuramente, anche lì, il Parlamento europeo e le istituzioni europee dovrebbero prendere una posizione chiara, cosa che non hanno fatto neanche, per esempio, nel caso del Tech Envoy, per chi lavora nel settore: anche lì si è persa un'occasione, perché, ovviamente, andare tutti sparpagliati per prendere un posto che, peraltro, non conta niente, non serve a molto. Mentre, invece, essere uniti e, appunto, avere una visione e una posizione unica a livello europeo, sicuramente ci può rafforzare anche per quanto riguarda la governance dell'intelligenza artificiale e dell'intelligenza artificiale etica.

.. per aiutarci a tuffarci nel futuro digitale!



E qui, insomma, per chiudere un po', come dire, con delle immagini non generate da AI, ma generate da dei designer grafici, c'è fondamentalmente l'idea che siamo solo all'inizio di questo, come dire, tuffo nel futuro, ed è chiaro che l'intelligenza artificiale può contribuire ad aumentare le nostre capacità cognitive e anche lo stato di diritto, se usata (e qui cito il nostro Leonardo da Vinci) disegnata e governata in modo antropocentrico, perché, appunto, forse la prima definizione, non in parole, ma in immagine, un po', forse, un Large Language Model ante litteram, è sicuramente stata fatta da Leonardo da Vinci sulla human centric AI.
E grazie per l'attenzione.

NAVIGARE NEI MARI DIGITALI: IL CAMBIO DI PARADIGMA DELLE NUOVE TECNOLOGIE

ASCOLTA INTERVENTO

FORTUNATO MUSELLA
Professore Ordinario di Scienza Politica,
Università degli Studi di Napoli Federico II

In realtà, l'ultima slide di Gianluca Misuraca mi aiuta anche a introdurre il mio tema perché non a caso il mio titolo diceva "Navigare nei mari digitali", quindi l'immagine è un perfetto assist per questo ultimo intervento che chiude il convegno.
Io credo che, prima di iniziare, è doverosa una premessa sia al mio intervento ma forse anche, in qualche modo è una postilla, agli interventi che mi hanno preceduto, per riflettere sul tempo in cui siamo.
Lo so, il tema potrebbe essere un po' pesante, se posso, a quest'ora del pomeriggio... Però credo che è una premessa di un discorso stabilire se il tempo in cui siamo sia un tempo nuovo oppure no.
Ovviamente ogni media, quando è stato introdotto, sempre crea una divisione fra apocalittici e integrati, una polarizzazione fra chi ricorre ad esso con entusiasmo e chi invece ne coglie elementi di preoccupazione, ma questo dibattito che si sta creando oggi non è lo stesso dibattito, non è una nuova versione dello stesso dibattito che si è ripresentato ad ogni innovazione tecnologica. È qualcosa di più profondo.
Significa interrogarsi se il tempo in cui siamo è uno spartiacque oppure meno nel percorso dell'umanità.
Ora, il dibattito, si può chiamare anche alterco se volete, del panel precedente, relativo al fatto a quali siano le conseguenze ultime

delle nuove tecnologie: non è esattamente la domanda che sto ponendo oggi, in questo momento .
Perché interrogarsi sulle conseguenze ultime significa interrogarsi su quale sarà il prodotto delle nuove tecnologie, in particolare l'intelligenza artificiale, ai fini dell'umanità. Allora, gli uni parlano di rischio esistenziale, quindi di un'umanità che sarà impoverita, oppure addirittura di un'umanità che sarà sostituita da intelligenze superiori; altri invece dicono: "Guardate, riusciremo a cogliere delle opportunità magnifiche dal progresso e quindi riusciremo a lavorare meglio, a giocare anche, ad avere occasioni di diletto, così come produrre anche innovazione".
In realtà la domanda è, non tanto sulle finalità ultime e quali sono gli esiti finali, ma se sia questo tempo uno spartiacque di civiltà oppure no.
Bene. Dando anche soccorso alla nostra moderatrice, non c'è nessun governo, nessuna impresa importante, nessuno che abbia responsabilità di governo, che in questo momento non pensa che la rivoluzione tecnologica, per così dire (e anche l'intelligenza artificiale) non abbia un impatto enorme sulle nostre vite.
Questo lo si può argomentare molto bene se si guardano, diciamo, i ritmi della diffusione. Le persone coinvolte nell'uso del digitale, la trasformazione dei processi produttivi, quella dei processi educativi e così via. Abbiamo tantissime argomentazioni che ci possono far partire nel discorso; cioè questo momento non è un momento uguale agli altri, non ripete una condizione che già si è verificata nel secolo scorso o tre o quattro secoli fa. Perché il ritmo del cambiamento in qualche modo sembra ingovernabile, ma non è detto che sia così tanto ingovernabile, tant'è che l'intervento mio, come il precedente, si preoccupa proprio di stabilire cosa sia il governo del digitale. Ora, questo panel, oltre ad essere inserito in un convegno sullo Stato di diritto, è un panel

che, se ricordo bene, focalizza in particolare sulla questione policy-etica.
Allora ci si potrebbe interrogare anche su questo. Cosa è l'etica e qual è il legame fra etica e governo del digitale?
In realtà, prima ancora di interrogarsi sulla costruzione di sistemi che mostrano una loro affidabilità, cioè che rispettino dei principi etici, che non siano discriminatori, che non vadano a sopprimere alcune categorie, a danneggiare altre classi e così via, in realtà, bisogna capire che la prima forma di etica è restituire alla politica il suo spazio.
Qui possiamo ricordare, insomma, nell'ambiente romano anche citazioni celebri, la più alta forma di carità. Che cos'è? La politica. Oppure possiamo andare più indietro, a Machiavelli: la fortuna e la virtù. Bene, prima di interrogarci sulla virtù, cioè su come andiamo a disegnare delle specifiche app, delle applicazioni, dei specifici sistemi applicativi, ci dobbiamo interrogare su cosa sia la politica in questo nuovo tempo, una volta stabilito che stiamo vivendo un nuovo tempo.
E questa è l'etica del digitale.
Non tanto come l'etica si dissemini nei vari sistemi applicativi, ma se questi sistemi applicativi possono restituire alla politica un suo spazio. Perché i primi segnali sono quelli per i quali la politica sta perdendo posizioni rispetto ad altri campi applicativi.
Allora, le nuove tecnologie, ma poi l'intelligenza artificiale, sono, nel mio modo di guardare, dal mio punto di osservazione, uno step successivo, un passo successivo, ma sono completamente d'accordo con Misuraca quando dice che si fa confusione, e il discorso non è l'ultima pagina, diciamo, da aggiungere a un capitolo, ma è un discorso più ampio che riguarda il digitale; è come restituire alla politica lo spazio che ha avuto, anche approfondire questo spazio e poter fare in modo che la politica

possa dire ancora qualcosa.
Si parla, da ultimo (poi veniamo anche al dato sulle campagne: magari se c'è un altro giro, poi mi scordo, qualcuno dice il dato sulle campagne, perché da politologo non posso far passare delle cose inosservate) è la governance algoritmica. Si parla di governance algoritmica. Bene, che cos'è il governo degli algoritmi? Ovviamente la letteratura corrente (ci sono qui esperti, lo sanno benissimo, ripeto, quanto tutti già sanno) declina questo concetto in due maniere.
La governance degli algoritmi significa che qualcuno fa gli algoritmi: io controllo, faccio sorveglianza, cerco di portare acqua al mio mulino, oppure di fare in modo che non ci siano dei criminali al comando… Ma poi c'è un altro aspetto ancora più importante, che veniva richiamato anche prima, cioè la governance, o il governo, attraverso gli algoritmi. E qui recuperiamo una lezione, che è una lezione della prima ora, non del primo tempo, della prima ora. Prima ancora del digitale, il legame fra tecnologia e politica. Ogni volta che disegniamo una tecnologia è già politica, perché ci crea un ambiente dove le persone andranno a insistere. Ha una valenza, si può anche dire, costituzionale. E poi a fine anni '90, a parte Lawrence Lessig, con grande forza, un altro libero studioso, notissimo, a cui sono molto affezionato, il codice è la legge… Che significa? Quando se ne esce con questa forma così dirompente, significa: "Guardate, quando voi state facendo un protocollo, state scrivendo una frase in codice, attraverso quella roba lì state facendo una legge". E chi costruisce un social sta facendo una legge, a cui si atterranno milioni di persone, nelle loro interazioni, nelle loro modalità di consumo, nelle loro modalità di interagire, nel loro modo di vedere, di imparare, di capire.
Quindi sto facendo politica, attraverso il digitale.

E quando poi andiamo dal digitale classico, chiamiamolo con le diciture che volete, a un digitale avanzato, queste cose non fanno che avanzare esse stesse.
E ora approfitto, la dico quella cosa lì, la "propaganda" è sempre esistita. È sempre esistita. Anche quella sovietica. Anche in Italia si è fatta propaganda. Ma la differenza è abissale. È abissale perché questa roba qui (le nuove tecnologie) è massiva. È massiva. E sono passati due turni presidenziali da quando si è abbastanza dimostrato (poi non so quale prove documentali si vogliano ricercare), ma ci sono dei rapporti del senato americano che dimostrano che Cambridge Analytica è stato un caso negli Stati Uniti. Anche perché non dobbiamo dimostrare che tutti gli elettori votino sulla base di un'immagine. Ma dobbiamo dimostrare che una quota di elettorato si sposti da una parte all'altra. E se mettiamo insieme l'uso delle immagini, l'uso dei dati, fake news, l'uso sistematico dei trolls e così via, chi più ne ha più ne metta, quella quota è garantita e l'integrità elettorale non è più garantita e le elezioni sono falsificate. E questo si è detto con copiosità di prove per la più grande (intesa la più estesa) democrazia al mondo che è quella americana. Ma poi è venuta l'argentina e quindi l'India e così via. Dell'integrità elettorale possiamo iniziare a sospettare fortemente, non guardando alla correttezza del voto, ma a tutto il contesto in cui è inserito poi l'atto di voto.
Allora, se poniamo questo scenario in cui il tempo è un tempo nuovo e le nuove tecnologie possono garantire una forma di governance che non sappiamo ancora se è politica o privata, allora possiamo fare un focus, a questo punto, più sull'Europa e quindi sull'Italia.
L'azione europea (qui sono accanto a un operatore, a una persona che in qualche modo ha scritto…) è benemerita. È benemerito,

diciamo, trovare una regolamentazione. Dobbiamo però capire, e in qualche modo comprendere, qual è l'approccio dell'Unione europea in questo tempo nuovo e con queste possibilità nuove di governo. Perché è stato detto benissimo: regolare è qualcosa di positivo. Ma se vogliamo, brevemente, nel tempo che è a mia disposizione, definire che cosa è stata la politica europea del digitale e da ultimo dell'intelligenza artificiale, possiamo dire che è stata una politica trasversale. Insomma, troviamo una regolamentazione che accolga tanti tipi al proprio interno (non vado a definire i settori: settore industriale, settore dell'editoria, come è stato fatto invece negli Stati Uniti d'America), ma io scelgo un approccio più onnicomprensivo, dopodiché è stato un approccio che non scende in tutti i quadranti che ha definito Misuraca, ma che è stato un approccio difensivo, che parte da un concetto di rischio e di risposta al rischio (quindi un approccio benemerito), ma che non si pone il problema del governo delle nuove tecnologie.

Tant'è vero che anche in questo panel ci sta uno slittamento, un slittamento lessicale.

Abbiamo parlato di non regolare solo, ma anche innovare. Misuraca ha detto, a un certo punto, bisogna anche innovare e nella regolamentazione riusciremo anche a innovare, ma poi c'è un'altra cosa ancora, il governo. Riusciremo a governare attraverso questi sistemi di regolamentazione? Probabilmente lo faremo a parte, lo faremo successivamente, ma nulla si dice del fatto che questo approccio sia pro-active, che questo approccio possa riuscire a non essere solo difensivo, anche perché il terzo punto caratteristico dell'intervento dell'Unione Europea è che è parsimonioso.

L'Unione Europea investe pochissimo in tecnologia.

Si definisce in Unione Europea un "global setter", quindi l'orgoglio

di tutti noi si accresce, mentre è stato ricordato che siamo dipendenti all'85%, che non abbiamo, tranne una (non faccio pubblicità) alcuna piattaforma che abbia valenza globale; non abbiamo le sedi delle giant corporation digitale, che non sono in Europa, ma ci poniamo però il problema di come regolamentare il traffico per evitare che si facciano i danni più importanti.
Ma questo non è il governo del digitale.
Scendiamo all'Italia.
L'Italia ovviamente, se parliamo di Stato di diritto, è all'interno dell'Europa, quindi è specchio di quanto nell'Unione Europea viene definito (tant'è che un amico romano, Nicola Lupo, in un testo che ho curato, parlava di due capitali, Roma e Bruxelles, diceva: "L'Italia ha due capitali") un altro livello, ma del nostro stesso livello di governo per tanti aspetti costituzionali. E quindi ricaviamo questo approccio, facciamo nostro questo approccio e, ovviamente, ci mettiamo anche del nostro.
La riflessione sul digitale in Italia è iniziata negli anni '70.
Sempre la mia presenza a Roma mi fa pensare a Sabino Cassese, che già nei suoi rapporti sulla riforma della pubblica amministrazione già parlava di computerizzazione, di informatizzazione, di come andare ad utilizzare le macchine del digitale e poi, da ultimo, lo stesso Cassese commentando i piani del PNRR ci dirà: "Ma noi non dobbiamo applicare le macchine ai vecchi schemi, non dobbiamo infrastrutturare, è inutile che investiamo così tanto; è utile anzi, ma non è sufficiente investire nella dotazione hardware della pubblica amministrazione, perché bisogna cambiare le routine, le pratiche, le prospettive, i modi di fare le cose".
Le nuove tecnologie, tanto più le hai, sono degli strumenti (appunto Machiavelli che diceva la virtù, ma anche la fortuna), sono la fortuna di questo tempo perché potrebbero consentire di

riformulare i processi di policy, potrebbero consentire di governare in una maniera più efficace, potrebbero consentire di avere più dati da parte dei cittadini, di osservare maggiormente i processi reali, di intervenire con maggiore efficacia.
Ora, in questo io identificherei anche alcuni campi di sviluppo.
Uno è stato già richiamato: il redesign dei servizi amministrativi, è uno dei campi che prima viene in mente, ma abbiamo anche altre scelte di policy da fare. Ad esempio, in un articolo per l'European Political Science Review del 2015, insieme a Francesco Amoretti, mettevamo in luce come c'è una scelta da fare in termini di policy del digitale relativa ai confini della nostra comunità politica, perché, in qualche modo, se i vari Stati hanno database comuni, hanno infrastrutture digitali comuni, usano anche lo stesso lessico concettuale che si può utilizzare per far viaggiare questi dati, a questo punto si crea una comunità che è una comunità europea, perché il digitale potrebbe essere, e lo è stato in tanti frangenti, una modalità attraverso la quale costituire un'unione politica. Come pure gli stessi accordi di scambio dati sono stati fatti a volte per parte atlantica, avvicinando l'Europa agli Stati Uniti. Quindi, quali sono i confini del digitale: un altro campo di policy, un campo in cui la politica è chiamata a compiere delle scelte; oppure ancora abbiamo visto durante la pandemia come si sono ristrutturati i processi del lavoro e in particolare del lavoro nella pubblica amministrazione: altro campo in cui interviene il digitale e dove interviene anche l'intelligenza artificiale.
Insomma, ci sono dei campi di ristrutturazione politica che ci fanno ritornare alla prima domanda, stante il fatto che alla prima questione, ovvero se siamo noi in un tempo nuovo, io penso possiamo dare una risposta più che affermativa, e se qualcuno ancora non ci crede, sarà svegliato da questo sonno nei prossimi mesi, perché il ritmo del cambiamento è tale da poter risvegliare

anche le coscienze più assopite.
Ora, invece, dobbiamo vedere cosa vogliamo fare di tutta questa materia da parte politica. Intanto, possiamo sostenere che la pubblica amministrazione, e con essa anche la politica nel suo compresso, è il vagone più lento del cambiamento. Se noi confrontiamo cosa si fa (e si diceva anche prima) in campo pubblico, con cosa si fa in campo privato, bene il campo privato è il vero campo dell'innovazione, ma non solo dell'innovazione, ma anche dell'utilizzo di quella innovazione. Basterebbe pensare a delle aziende che hanno fatto la loro fortuna, ma che hanno fatto la loro fortuna, attenzione, non solo sottolineo questo, nel recupero di profitti, ma anche nell'elaborare nuove modalità di interazione con i cittadini, nuove possibilità, nuovi moduli organizzativi, nuove routine, nuove mentalità.
E allora, io credo che siamo a uno snodo storico e potrei declinare (e non lo faccio per la carità che si richiamava all'inizio), lo potrei dimostrare questo in ogni campo della mia disciplina, della scienza politica. Cioè, preso un capitolo a caso, se mi interrogate su uno dei capitoli classici della scienza politica, dal Parlamento alle elezioni, dal Governo alla Pubblica Amministrazione, quello che volete, vi sfido (tra virgolette, scherzosamente), potrei dimostrarvi che siamo a una svolta storica.
Quello che invece non posso dimostrare, perché siamo all'anno zero, è come ci impossessiamo di questi strumenti per fare ancora politica, perché l'orizzonte potrebbe essere quello di una politica spogliata, di una politica che non si fa in sedi belle e prestigiose come queste, le sedi della nostra rappresentanza politica, ma nelle sedi dei board di qualche impresa digitale, o peggio ancora, in algoritmi che hanno smarrito addirittura una capacità di governo della collettività.
Grazie.

Pubblicato il 19 giugno 2024 su
www.fondazionemarcopannella.it

ATTI DEL CONVEGNO

IA
E STATO DI DIRITTO

Sfide e Opportunità per le Democrazie nell'Era Digitale

Fondazione Marco Pannella
www.fondazionemarcopannella.it