Libro bianco sull'intelligenza artificiale
Un approccio europeo all'eccellenza e alla fiducia
(ESTRATTO)

1) La Commissione si impegna a favorire i progressi scientifici, a preservare la leadership tecnologica dell'UE e a garantire che le nuove tecnologie siano al servizio di tutti gli europei e ne migliorino la vita rispettandone i diritti.

2) La Commissione sostiene pertanto un approccio normativo e orientato agli investimenti con il duplice obiettivo di promuovere l'adozione dell'IA e di affrontare i rischi associati a determinati utilizzi di questa nuova tecnologia.

3) Oggi la maggior parte dei dati riguarda i consumatori ed è conservata e elaborata in infrastrutture centrali basate su cloud. Un'ampia fetta dei dati di domani, ben più abbondanti di quelli di oggi, proverrà invece dall'industria, dalle imprese e dal settore pubblico e sarà conservata in vari sistemi, in particolare dispositivi di calcolo funzionanti ai margini della rete. Ciò apre nuove opportunità per l'Europa, che ha una posizione forte nel settore del digitale e delle applicazioni tra imprese (business-to-business), ma ha una posizione relativamente debole per quanto riguarda le piattaforme destinate ai consumatori.

4)Semplificando possiamo dire che l'IA è un insieme di tecnologie che combina dati, algoritmi e potenza di calcolo. I progressi compiuti nell'ambito del calcolo e la crescente disponibilità di dati sono pertanto fattori determinanti per l'attuale crescita dell'IA. L'Europa può combinare i suoi punti di forza industriali e tecnologici con un'infrastruttura digitale di elevata qualità e un quadro normativo basato sui suoi valori fondamentali per diventare un leader mondiale nell'innovazione nell'economia dei

dati e nelle sue applicazioni, come indicato nella strategia europea per i dati3. Su questa base l'Europa
può sviluppare un ecosistema di IA che consenta alla sua società e alla sua economia nel loro
complesso di godere dei benefici apportati dalla tecnologia: ……..

5) Dato l'impatto significativo che l'intelligenza artificiale può avere sulla nostra società e la necessità di
creare maggiore fiducia, è essenziale che l'IA europea sia fondata sui nostri valori e diritti
fondamentali quali la dignità umana e la tutela della privacy.

6)· il quadro strategico che stabilisce misure per allineare gli sforzi a livello europeo, nazionale e
regionale. Tramite un partenariato tra il settore pubblico e privato, l'obiettivo di tale quadro è
mobilitare risorse per conseguire un "ecosistema di eccellenza" lungo l'intera catena del
valore, a cominciare dalla ricerca e dall'innovazione, e creare i giusti incentivi per accelerare
l'adozione di soluzioni basate sull'IA, anche da parte delle piccole e medie imprese (PMI);
· gli elementi chiave di un futuro quadro normativo per l'IA in Europa, che creerà un
"ecosistema di fiducia" unico. A tal fine, deve garantire il rispetto delle norme dell'UE,
comprese le norme a tutela dei diritti fondamentali e dei diritti dei consumatori, in particolare
per i sistemi di IA ad alto rischio gestiti nell'UE7. La costruzione di un ecosistema di fiducia è
un obiettivo strategico in sé e dovrebbe dare ai cittadini la fiducia di adottare applicazioni di
IA e alle imprese e alle organizzazioni pubbliche la certezza del diritto necessaria per innovare
utilizzando l'IA. La Commissione sostiene con forza un approccio antropocentrico basato sulla
comunicazione "Creare fiducia nell'intelligenza artificiale antropocentrica" 8 e terrà conto anche dei contributi ottenuti durante la fase pilota degli

orientamenti etici elaborati dal gruppo di esperti ad alto livello sull'IA…….

7) · Azione 4: la Commissione collaborerà con gli Stati membri per garantire che almeno un polo dell'innovazione digitale per Stato membro sia altamente specializzato in IA. I poli dell'innovazione digitale possono ricevere sostegno nell'ambito del programma Europa digitale.

8)· Azione 5: nel contesto di Orizzonte Europa, la Commissione istituirà un nuovo partenariato pubblico-privato per l'IA, i dati e la robotica, al fine di unire gli sforzi, garantire il coordinamento della ricerca e dell'innovazione nell'IA, collaborare con altri partenariati pubblico-privati di Orizzonte Europa e lavorare insieme alle strutture di prova e ai poli dell'innovazione digitale precedentemente menzionati.

9) H. ASPETTI INTERNAZIONALI
L'Europa si trova nella posizione ideale per esercitare una leadership mondiale costruendo alleanze a ttorno ai valori condivisi e promuovendo l'uso etico dell'IA. Il lavoro svolto dall'UE nel campo dell'IA ha già influenzato le discussioni a livello internazionale. Nell'elaborare i propri orientamenti etici, il gruppo di esperti ad alto livello ha coinvolto una serie di organizzazioni extra UE e vari osservatori governativi. Parallelamente l'UE è stata coinvolta da vicino nello sviluppo dei principi etici per l'IA dell'OCSE 25 . Il G20 ha successivamente approvato tali principi nella dichiarazione ministeriale sul commercio e l'economia digitale del giugno 2019. Allo stesso tempo l'UE riconosce che sono in corso importanti lavori sull'IA in altri consessi multilaterali, tra cui il Consiglio d'Europa, l'Organizzazione delle Nazioni Unite per l'educazione, lascienza e la cultura (UNESCO), l'Organizzazione per la cooperazione e lo sviluppo economici (OCSE), l'Organizzazione mondiale del commercio e l'Unione internazionale delle telecomunicazioni (UIT). ell'ambito delle Nazioni Unite, l'UE sta lavorando per dare seguito alla relazione del gruppo ad alto livello sulla cooperazione digitale, anche per quanto riguarda la raccomandazione sull'IA.
L'UE continuerà a cooperare in materia di IA con i paesi che condividono gli stessi principi, ma anche con gli attori globali, seguendo un approccio che promuova le regole e i valori dell'UE (ad esempio sostenendo la convergenza normativa verso l'alto, garantendo l'accesso alle risorse fondamentali, compresi i dati, e creando condizioni di parità). La

Commissione monitorerà da vicino le politiche dei paesi terzi che limitano i flussi di dati e affronterà le restrizioni indebite nei negoziati commerciali bilaterali e mediante iniziative nell'ambito dell'Organizzazione mondiale del commercio. La Commissione è convinta che la cooperazione internazionale sulle questioni riguardanti l'IA debba basarsi su un approccio che promuova il rispetto dei diritti fondamentali, tra cui la dignità umana, il pluralismo, l'inclusione, la non discriminazione e la protezione della privacy e dei dati personali26, e si adopererà per esportare i suoi valori nel mondo27. È inoltre chiaro che lo sviluppo e l'uso responsabili dell'IA possono essere una foza trainante per conseguire gli obiettivi di sviluppo sostenibile e portare avanti l'Agenda 2030.
La Commissione ha pubblicato una comunicazione31 in cui accoglie con favore i sette requisiti
fondamentali individuati negli orientamenti del gruppo di esperti ad alto livello:

· intervento e sorveglianza umani,
· robustezza tecnica e sicurezza,
· riservatezza e governance dei dati,
· trasparenza,
· diversità, non discriminazione ed equità,
· benessere sociale e ambientale, e

L'IA può essere estremamente utile, ad esempio in quanto può rendere prodotti e processi più sicuri, ma può anche risultare dannosa. Tale danno può essere di natura sia materiale (quando incide sulla salute e sulla sicurezza delle persone, provocando perdite di vite umane o danni patrimoniali) sia immateriale (quando causa una perdita della privacy, restrizioni alla libertà di espressione, pregiudizi alla dignità umana o discriminazioni, ad esempio nell'accesso all'occupazione), e può riguardareun'ampia gamma di rischi. Il quadro normativo dovrebbe concentrarsi su come ridurre al minimo i
diversi rischi di danno potenziale, soprattutto quelli più significativi.
I rischi principali connessi all'uso dell'IA riguardano l'applicazione di norme intese a tutelare i diritti fondamentali (comprese la protezione dei dati personali e della privacy e la non discriminazione), nonché le questioni legate alla sicurezza32 e alla responsabilità.

Rischi per i diritti fondamentali, comprese la protezione dei dati personali e della privacy e la non discriminazione L'uso dell'IA può pregiudicare i valori su cui si fonda l'Unione e causare violazioni dei diritti fondamentali33, compresi i diritti alle libertà di espressione e di riunione, la dignità umana, la non discriminazione fondata sul sesso, sulla razza, sull'origine etnica, sulla religione sulle convinzioni personali, sulla disabilità, sull'età o sull'orientamento sessuale (ove applicabili in determinati settori), la protezione dei dati personali e della vita privata34 o il diritto a un ricorso giurisdizionale effettivo e a un giudice imparziale, nonché la tutela dei consumatori. Tali rischi potrebbero derivare da difetti nella progettazione complessiva dei sistemi di IA (anche per quanto riguarda la sorveglianza umana) o dall'uso di dati senza che ne siano state corrette le eventuali distorsioni (ad esempio se un sistema è addestrato utilizzando solo o principalmente dati riguardanti gli uomini, il che comporta risultati non ottimali per quanto concerne le donne).

L'IA può svolgere molte funzioni che in precedenza potevano essere esercitate solo dagli esseri umani.
I cittadini e le persone giuridiche saranno pertanto sempre più soggette ad azioni realizzate e a
decisioni prese da sistemi di IA o con la loro assistenza; si tratta di un risultato non sempre facile da
comprendere e cui è difficile opporsi in modo efficace ove necessario. L'IA aumenta inoltre le
possibilità di seguire e analizzare le abitudini quotidiane delle persone. Vi è ad esempio il rischio
potenziale che l'IA venga utilizzata, in violazione della normativa dell'UE sulla protezione dei dati e di
altre norme, dalle autorità statali o da altri soggetti a fini di sorveglianza di massa, oppure dai datori di
lavoro per osservare il comportamento dei loro dipendenti. Poiché permette di analizzare grandi
quantità di dati e di individuare collegamenti tra di essi, l'IA può essere usata anche per ricostruire e
deanonimizzare dati riguardanti le persone, il che implica nuovi rischi per la protezione dei dati
personali anche in relazione a set di dati che di per sé non contengono dati personali. L'IA è utilizzata

anche dagli intermediari online per stabilire priorità riguardo alle informazioni da fornire ai loro utenti oppure a fini di moderazione dei contenuti. I dati trattati, le modalità con cui vengono progettate le applicazioni e il margine di intervento umano possono incidere sui diritti alla libertà di espressione alla protezione dei dati personali e alla privacy nonché sulle libertà politiche.Le distorsioni e le discriminazioni rappresentano un rischio intrinseco di qualunque attività sociale od economica. Il processo decisionale umano non è immune da errori e distorsioni. Queste stesse distorsioni, se presenti nell'IA, potrebbero tuttavia avere effetti molto maggiori e colpire o discriminare numerose persone in assenza dei meccanismi di controllo sociale che disciplinano il comportamento umano35. Ciò può accadere anche quando il sistema di IA "apprende" nel corso del suo funzionamento. In tali casi, in cui i risultati non potevano essere evitati o anticipati in fase di progettazione, i rischi deriveranno non da difetti nella progettazione originale del sistema, bensì dagli effetti pratici delle correlazioni o dei modelli che il sistema individua all'interno di un ampio set di dati. Le caratteristiche specifiche di molte tecnologie di IA, tra cui l'opacità (effetto "scatola nera"), la complessità, l'imprevedibilità e un comportamento parzialmente autonomo, possono rendere difficile verificare il rispetto delle normative dell'UE in vigore volte a proteggere i diritti fondamentali e possono ostacolarne l'applicazione effettiva. Le autorità preposte all'applicazione della legge e le persone interessate potrebbero non disporre dei mezzi per verificare come sia stata presa una determinata decisione con il coinvolgimento di sistemi di IA e, di conseguenza, se sia stata rispettata la normativa pertinente. Le persone fisiche e giuridiche possono incontrare difficoltà nell'accesso effettivo alla giustizia in situazioni in cui tali decisioni possono avere ripercussioni negative su di loro …..

A norma della direttiva sulla responsabilità per danno da prodotti difettosi, il produttore è responsabile dei danni causati da un prodotto difettoso. Tuttavia, nel caso di sistemi basati sull'IA, come quelli delle auto a guida autonoma, può rivelarsi difficile provare che il prodotto è difettoso e dimostrare il danno cagionato e il nesso di causalità tra difetto e danno. In aggiunta non è chiaro come e in che misura si applichi la direttiva sulla responsabilità per danno da prodotti difettosi nel caso di alcuni tipi di difetti, ad esempio per quelli risultanti da carenze della cibersicurezza del prodotto.

Per quanto riguarda la protezione dei diritti fondamentali e dei diritti dei consumatori, il quadro legislativo dell'UE comprende atti normativi come la direttiva sull'uguaglianza razziale39, la direttiva sulla parità di trattamento in materia di occupazione e di condizioni di lavoro40, le direttive sulla parità di trattamento tra uomini e donne per quanto riguarda l'accesso a beni e servizi e in materia di occupazione41, una serie di norme sulla tutela dei consumatori42, sulla protezione dei dati personali e sulla privacy, in particolare il regolamento generale sulla protezione dei dati, nonché altre normative settoriali riguardanti la protezione dei dati personali, come la direttiva sulla protezione dei dati nelle attività di polizia e giudiziarie43. Inoltre, a partire dal 2025 si applicheranno le norme sui requisiti di accessibilità dei beni e dei servizi previste dall'atto europeo sull'accessibilità44. I diritti fondamentali devono inoltre essere rispettati in sede di attuazione di altre normative dell'UE, anche nel settore dei servizi finanziari, della migrazione o della responsabilità degli intermediari online. Se da un lato la legislazione dell'UE rimane in linea di principio pienamente applicabile, indipendentemente dal coinvolgimento dell'IA, dall'altro è importante valutare se tale normativa possa essere adeguatamente applicata per far fronte ai rischi derivanti dai sistemi di IA o se sia necessario apportare adeguamenti a determinati strumenti giuridici. Ad esempio, gli operatori economici rimangono pienamente responsabili della conformità dell'IA alle norme vigenti a tutela dei consumatori: è vietato utilizzare algoritmi che sfruttano il comportamento dei consumatori in violazione delle norme vigenti, e qualunque infrazione a tale divieto è punita di conseguenza.

Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità

La relazione che accompagna il presente libro bianco analizza il quadro giuridico pertinente, individuando le incertezze riguardanti l'applicazione di tale quadro giuridico in relazione ai rischi specifici derivanti dai sistemi di IA e da altre tecnologie digitali. La conclusione è che la legislazione vigente in materia di sicurezza dei prodotti sostiene già un concetto ampio di sicurezza, con l'obiettivo di proteggere da tutti i tipi di rischi derivanti dal prodotto in funzione dell'uso dello stesso. Per garantire una maggiore certezza del diritto si potrebbero tuttavia introdurre disposizioni che contemplino esplicitamente i nuovi rischi derivanti dalle tecnologie digitali emergenti.

· Il comportamento autonomo che mostrano alcuni sistemi di IA durante il loro ciclo di vita può comportare modifiche significative dei prodotti, che a loro volta hanno ripercussioni sulla sicurezza e possono rendere necessaria una nuova valutazione dei rischi. Potrebbe inoltre rendersi necessaria, come misura di salvaguardia, la sorveglianza umana dalla fase di progettazione e durante tutto il ciclo di vita dei prodotti e dei sistemi di IA. Ove opportuno potrebbero essere presi in considerazione obblighi espliciti per i produttori anche in relazione ai rischi per la sicurezza mentale degli utenti (ad esempio dovuti alla collaborazione con robot umanoidi). La legislazione dell'Unione in materia di sicurezza dei prodotti potrebbe prevedere prescrizioni specifiche che affrontino i rischi per la sicurezza derivanti dall'uso di dati errati in fase di progettazione, nonché meccanismi per garantire che sia mantenuta la qualità dei dati durante l'intero periodo di utilizzo dei prodotti e dei sistemi di IA.
· Il problema dell'opacità dei sistemi basati su algoritmi potrebbe essere affrontato mediante prescrizioni in materia di trasparenza.
· Potrebbe essere necessario adeguare e chiarire le norme vigenti in relazione ai casi di software indipendente immesso sul mercato senza altri componenti, oppure integrato in un prodotto dopo che quest'ultimo è stato immesso sul mercato, qualora ciò abbia un impatto sulla sicurezza. · Data la crescente complessità delle catene di approvvigionamento in relazione alle nuove tecnologie, disposizioni che richiedano specificamente una collaborazione tra gli utenti e gli operatori economici attivi lungo la catena di approvvigionamento potrebbero garantire la certezza del diritto. Le caratteristiche delle tecnologie digitali emergenti, come l'IA, l'Internet delle cose e la robotica, possono mettere in discussione alcuni aspetti dei quadri giuridici relativi alla responsabilità e renderli meno efficaci. Alcune di queste caratteristiche potrebbero rendere difficile far risalire il danno alla condotta di una persona, condizione indispensabile per invocare la responsabilità per colpa conformemente alla maggior parte delle norme nazionali in materia. Ne deriva la possibilità che i costi per le vittime aumentino in maniera significativa, e la potenziale difficoltà di avviare azioni per responsabilità nei confronti di soggetti diversi dal produttore e di ottenere elementi di prova a sostegno di tali azioni. Le persone che hanno subito un danno provocato con il coinvolgimento di sistemi di IA devono godere dello stesso livello di protezione delle persone che hanno subito danni causati da altre tecnologie; nel contempo occorre che l'innovazione tecnologica possa continuare a svilupparsi. · È opportuno valutare attentamente tutte le

opzioni che consentano di raggiungere tale obiettivo, compresa la possibilità di modificare la direttiva sulla responsabilità per danno da prodotti difettosi e ulteriori possibili interventi mirati di armonizzazione delle norme nazionali in materia di responsabilità. Ad esempio, la Commissione sta valutando se e in quale misura sia necessario mitigare le conseguenze della complessità adeguando l'onere della prova richiesto dalle norme nazionali in materia di responsabilità in relazione ai danni provocati dal funzionamento delle applicazioni di IA.

IA dovrebbe essere considerata ad alto rischio se soddisfa i due criteri cumulativi descritti di seguito. · In primo luogo, l'applicazione di IA è utilizzata in un settore in cui, date le caratteristiche delle attività abitualmente svolte, si possono prevedere rischi significativi. garantisce che l'intervento normativo sia mirato i settori in cui i rischi sono generalmente ritenuti più probabili. I settori interessati dovrebbero essere elencati in maniera specifica ed esaustiva nel nuovo quadro normativo. Ad esempio, settori dell'assistenza sanitaria, dei trasporti; dell'energia e parti del settore pubblico50. L'elenco dovrebbe essere periodicamente rivisto e modificato, ove necessario, in funzione dei pertinenti sviluppi nella pratica. In secondo luogo, l'applicazione dell'IA nel settore in questione è inoltre utilizzata in modo tale da poter generare rischi significativi. Questo secondo criterio riconosce il fatto che non tutti gli usi dell'IA nei settori selezionati comportano necessariamente rischi significativi. Ad esempio, per quanto l'assistenza sanitaria in genere possa essere certamente un settore rilevante, un eventuale difetto del sistema di prenotazione degli appuntamenti in un ospedale non presenta, in linea di massima, rischi tali da giustificare un intervento legislativo. La valutazione del livello di rischio derivante da un determinato uso potrebbe basarsi sull'impatto per i soggetti interessati. Ad esempio, usi delle applicazioni di IA che producono effetti giuridici, o effetti altrettanto rilevanti, sui diritti di una persona o di una società; usi che presentano il rischio di lesioni, morte o danni materiali o immateriali significativi; usi che producono effetti non ragionevolmente evitabili dalle persone fisiche o giuridiche……

Tenendo conto degli orientamenti del gruppo di esperti ad alto livello e di quanto esposto in precedenza, le prescrizioni per le applicazioni di IA ad alto rischio potrebbero riguardare i seguenti elementi essenziali, ulteriormente discussi nelle sottosezioni che seguono:

· --i dati di addestramento;
· --la tenuta dei dati e dei registri;
· --le informazioni da fornire;
· --la robustezza e la precisione;
· --la sorveglianza umana;
· --prescrizioni specifiche per determinate applicazioni particolari dell'IA, come quelle utilizzate a fini di identificazione biometrica remota.

· prescrizioni volte a fornire ragionevoli garanzie circa la sicurezza dell'uso successivo dei prodotti e dei servizi basati sull'IA; tale sicurezza è intesa come conformità agli standard stabiliti dalle norme dell'UE applicabili in materia di sicurezza (sia quelle vigenti, sia eventuali norme complementari). Si tratta ad esempio di prescrizioni volte a garantire che i sistemi di IA siano addestrati utilizzando set di dati sufficientemente ampi e che contemplino tutti gli scenari pertinenti, in modo da evitare situazioni pericolose;

· prescrizioni che impongano di adottare misure ragionevoli per garantire che tale uso successivo dei sistemi di IA non porti a risultati che implichino discriminazioni vietate. Tali prescrizioni potrebbero comportare, in particolare, l'obbligo di utilizzare set di dati sufficientemente rappresentativi, in particolare per garantire che rispecchino adeguatamente tutte le pertinenti dimensioni di genere, etnia e altri possibili motivi di discriminazione vietata;

· prescrizioni volte a garantire un'adeguata protezione della privacy e dei dati personali durante l'uso dei prodotti e dei servizi basati sull'IA. Tali questioni sono disciplinate dal regolamento generale sulla protezione dei dati e dalla direttiva sulla protezione dei dati nelle attività di polizia e giudiziarie per la parte che rientra nel loro rispettivo ambito di applicazione.

· registri accurati dei set di dati utilizzati per addestrare e sottoporre a prova i sistemi di IA, compresa una descrizione delle principali caratteristiche e delle modalità di selezione di tali insiemi;
· in taluni casi giustificati, gli stessi set di dati;

· la documentazione relativa alle metodologie, alle procedure e alle tecniche di programmazione53 e di addestramento utilizzate per costruire,

sottoporre a prova e convalidare i sistemi di IA, ove opportuno anche in materia di sicurezza, evitando distorsioni che potrebbero dar luogo a discriminazioni vietate.

· garantire che vengano fornite informazioni chiare sulle capacità e sulle limitazioni dei sistemi di IA, in particolare per quanto riguarda l'obiettivo per il quale i sistemi sono stati concepiti, le condizioni alle quali ci si può attendere che funzionino come previsto e il livello di precisione previsto nel raggiungimento dell'obiettivo specificato. Tali informazioni sono importanti soprattutto per coloro che applicano i sistemi, ma possono essere rilevanti anche per le autorità competenti e le parti interessate;
· informare i cittadini separatamente e con chiarezza quando questi interagiscono con un sistema di IA e non con un essere umano. Anche se la normativa dell'UE in materia di protezione dei dati contiene già alcune disposizioni analoghe54, potrebbero essere necessarie ulteriori prescrizioni per raggiungere gli obiettivi sopra indicati. In tal caso occorre evitare oneri inutili. Non è quindi necessario fornire tali informazioni, ad esempio, quando ai cittadini sia immediatamente evidente che interagiscono con sistemi di IA. È inoltre importante che le informazioni fornite siano obiettive, concise e di facile comprensione. Le modalità di presentazione delle informazioni dovrebbero essere adeguate al contesto specifico.

· prescrizioni che garantiscano la robustezza e la precisione dei sistemi di IA, o che almeno riflettano correttamente il loro livello di precisione, durante tutte le fasi del ciclo di vita;
· prescrizioni atte a garantire la riproducibilità dei risultati;
· prescrizioni atte a garantire che i sistemi di IA possano gestire adeguatamente gli errori o le incongruenze in tutte le fasi del ciclo di vita;
· prescrizioni volte a assicurare la resilienza dei sistemi di IA sia agli attacchi palesi sia ai tentativi meno evidenti di manipolare i dati o gli stessi algoritmi, e a garantire che in tali casi siano adottate misure di attenuazione

· il risultato del sistema di IA non diviene effettivo prima di essere stato rivisto e convalidato da

un essere umano (ad esempio, la decisione di respingere una richiesta di prestazioni di sicurezza sociale può essere presa unicamente da un essere umano);

· il risultato del sistema di IA diviene immediatamente effettivo, ma successivamente è garantito l'intervento di un essere umano (ad esempio, la decisione di respingere la richiesta di una carta di credito può essere adottata da un sistema di IA, ma successivamente deve essere possibile il riesame da parte di un essere umano);

· il sistema di IA può essere monitorato durante il suo funzionamento da un essere umano che può intervenire in tempo reale e disattivarlo (ad esempio, prevedendo su un'auto a guida autonoma un pulsante o una procedura di arresto che un essere umano può attivare qualora decida che il funzionamento dell'auto non è sicuro)

· in fase di progettazione, imponendo vincoli operativi al sistema di IA (ad esempio, in determinate condizioni di scarsa visibilità in cui i sensori divengono meno affidabili, un'auto a guida autonoma deve cessare di funzionare, oppure in qualunque condizione tale veicolo deve mantenere una determinata distanza rispetto al veicolo che lo precede).

La raccolta e l'uso di dati biometrici55 a fini di identificazione56 remota, ad esempio mediante la diffusione di sistemi di riconoscimento facciale in luoghi pubblici, comporta rischi specifici per i diritti fondamentali57. L'uso di sistemi di IA per l'identificazione biometrica remota ha implicazioni sui diritti fondamentali ce variano notevolmente a seconda dello scopo, del contesto e della portata del loro uso.
consegue che, conformemente alle vigenti norme dell'UE in materia di protezione dei dati e alla
Carta dei diritti fondamentali, l'IA può essere utilizzata a fini di identificazione biometrica remota unicamente ove tale uso sia debitamente giustificato, proporzionato e soggetto a garanzie adeguate.

La Commissione ritiene che in un futuro quadro normativo ciascun obbligo debba essere stabilito a carico dell'operatore o degli operatori che si trovano nella posizione migliore per affrontare eventuali rischi potenziali. Ad esempio, se da un lato gli sviluppatori dell'IA sono i più qualificati per affrontare i rischi derivanti dalla fase di sviluppo, dall'altro

la loro capacità di controllare i rischi durante la fase di utilizzo può essere più limitata. In tal caso il pertinente obbligo dovrebbe essere stabilito a carico del soggetto che applica l'IA.

In fase di progettazione e realizzazione di un sistema sulla base di valutazioni preliminari della conformità, occorre considerare in particolare:
·
che non tutti i requisiti sopra indicati possono essere verificati mediante una valutazione preliminare della conformità. Ad esempio, in generale l'adempimento dell'obbligo di fornire informazioni non si presta facilmente ad essere verificato mediante una simile valutazione;

la possibilità che taluni sistemi di IA si evolvano e apprendano dall'esperienza: tale possibilità può rendere necessarie valutazioni ripetute per tutta la durata dei sistemi di IA in questione;

· la necessità di verificare i dati utilizzati per l'addestramento nonché le pertinenti metodologie, procedure e tecniche di programmazione e di addestramento utilizzate per costruire, sottoporre
a prova e convalidare i sistemi di IA;

· che, qualora dalla valutazione della conformità emerga che un sistema di IA non rispetta alcune prescrizioni, ad esempio quelle relative ai dati utilizzati per l'addestramento, sarà necessario porre rimedio alle carenze individuate, ad esempio addestrando nuovamente il sistema nell'UE in modo tale da garantire il rispetto di tutte le prescrizioni applicabilITA’.

ETICHETTATURA SU BASE VOLONTARIA PER LE APPLICAZIONI DI IA NON AD ALTO RISCHIO

Per le applicazioni di IA che non sono considerate "ad alto rischio" (cfr. sezione C) e che pertanto non sono soggette alle prescrizioni obbligatorie di cui sopra (cfr. le sezioni D, E e F), oltre alla legislazione applicabile vi sarebbe la possibilità di istituire un sistema di etichettatura su base volontaria. Il marchio volontario consentirebbe agli operatori economici interessati di mettere in evidenza l'affidabilità dei loro prodotti e servizi basati sull'IA, e consentirebbe agli utenti di riconoscere facilmente che i prodotti e i servizi in questione sono conformi a determinati parametri di

riferimento obiettivi e standardizzati a livello dell'UE, che vanno oltre gli obblighi giuridici normalmente applicabili. Ciò contribuirebbe a rafforzare la fiducia degli utenti nei sistemi di IA e a promuovere l'adozione generale della tecnologia.
Tale opzione richiederebbe la creazione di un nuovo strumento giuridico che definisca il quadro relativo a un sistema di etichettatura su base volontaria per gli sviluppatori e/o i soggetti che applicano i sistemi di IA non considerati ad alto rischio. Sebbene la partecipazione al sistema di etichettatura sia facoltativo, una volta che gli sviluppatori o i soggetti che applicano l'IA hanno deciso di aderirvi le relative prescrizioni sarebbero vincolanti. La combinazione di attività di applicazione della legge ex ante ed ex post dovrebbe garantire il rispetto di tutte le prescrizioni.

GOVERNANCE

Occorre una struttura di governance europea in materia di IA, sotto forma di un quadro di cooperazione delle autorità nazionali competenti, per evitare la frammentazione delle responsabilità, aumentare la capacità degli Stati membri e garantire che l'Europa si doti progressivamente delle capacità necessarie per sottoporre a prova e certificare prodotti e servizi basati sull'IA. In tale contesto sarebbe utile sostenere le autorità nazionali competenti affinché possano adempiere il loro mandato nei casi di utilizzo dell'IA.
La struttura di governance dovrebbe garantire la massima partecipazione dei portatori di interessi, fra
cui organizzazioni dei consumatori e parti sociali, imprese, ricercatori e organizzazioni della società
civile, che dovrebbero essere consultati in merito all'attuazione e all'ulteriore sviluppo del quadro di RIFERIMENTO.

La struttura di governance proposta non dovrebbe duplicare le funzioni esistenti, bensì stabilire stretti legami con altre autorità competenti a livello nazionale e dell'UE nei vari settori, al fine di integrare le competenze esistenti e aiutare le autorità nel monitoraggio e nella sorveglianza delle attività svolte dagli operatori economici in cui intervengono sistemi di IA nonché prodotti e servizi basati sull'IA.

CONCLUSIONI

L'IA è una tecnologia strategica che offre molti benefici ai cittadini, alle imprese e alla società nel suo insieme, a condizione che segua un approccio antropocentrico, etico, sostenibile e rispettoso dei valori e dei diritti fondamentali. L'IA offre importanti vantaggi in termini di efficienza e produttività, che possono rafforzare la competitività dell'industria europea e migliorare il benessere dei cittadini. Può inoltre contribuire a individuare soluzioni ad alcune delle sfide sociali più urgenti, tra cui la lotta ai cambiamenti climatici e al degrado ambientale, le sfide legate alla sostenibilità e ai cambiamenti demografici, la protezione delle nostre democrazie e, ove tale uso sia necessario e proporzionato, la lotta alla criminalità.L'approccio europeo all'IA mira a promuovere la capacità di innovazione dell'Europa nel settore dell'IA, sostenendo nel contempo lo sviluppo e la diffusione di un'IA etica e affidabile in tutta l'economia dell'UE. L'IA dovrebbe apportare vantaggi alle persone ed essere un fattore positivo per la società.

La Commissione invita a formulare osservazioni sulle proposte presentate nel libro bianco mediante una consultazione pubblica aperta consultabile all'indirizzo: https://ec.europa.eu/info/consultations_en. Sarà possibile formulare osservazioni nel quadro della consultazione fino al 19 maggio 2020.

È prassi abituale della Commissione pubblicare le osservazioni pervenute in risposta a una consultazione pubblica. È tuttavia possibile chiedere che tali osservazioni rimangano in tutto o in parte riservate. In tal caso si prega di indicare tale richiesta sulla prima pagina della risposta e trasmettere alla Commissione anche una versione non riservata della risposta, che sarà pubblicata.