

PREPARARE UN GIUSTO FUTURO

L'INTELLIGENZA ARTIFICIALE E I DIRITTI FONDAMENTALI

SINTESI



3

Risultati principali e pareri della FRA

5

Salvaguardare i diritti fondamentali: ambito di applicazione, valutazioni d'impatto e responsabilizzazione

10

Non discriminazione, protezione dei dati e accesso alla giustizia: tre temi orizzontali

© Agenzia dell'Unione europea per i diritti fondamentali, 2021

La riproduzione è autorizzata con citazione della fonte.

L'uso o la riproduzione di fotografie o altro materiale non protetti dal diritto d'autore dell'Agenzia dell'Unione europea per i diritti fondamentali devono essere autorizzati direttamente dal titolare del diritto d'autore.

L'Agenzia dell'Unione europea per i diritti fondamentali, o chiunque agisca in suo nome, declina ogni responsabilità per l'uso dei contenuti che seguono.

Lussemburgo, Ufficio delle pubblicazioni dell'Unione europea, 2021

Print	ISBN 978-92-9461-204-5	doi:10.2811/167998	TK-04-20-657-IT-C
PDF	ISBN 978-92-9461-245-8	doi:10.2811/410641	TK-04-20-657-IT-N

Crediti fotografici:

Copertina: © HQUALITY/Adobe Stock

Pagina 1: © Mykola Mazuryk/AdobeStock

Pagina 3: © Mimi Potter/AdobeStock

Pagina 4: © Sikov/Adobe Stock

Pagina 7: © Gorodenkoff/Adobe Stock

Pagina 9: © VideoFlow/Adobe Stock

Pagina 11: © Monsitj/Adobe Stock

«L'IA è una gran cosa, ma dobbiamo imparare a usarla»

(Società privata, Spagna)

«C'è il rischio di riporre troppa fiducia nella macchina»

(Amministrazione pubblica, Francia)

L'intelligenza artificiale (IA) è sempre più utilizzata sia nel settore privato che nel pubblico e influenza la vita quotidiana. Alcuni vedono l'IA come la fine del controllo umano sulle macchine, mentre altri come la tecnologia che aiuterà l'umanità ad affrontare alcune delle sue sfide più pressanti. Può darsi che nessuna delle due rappresentazioni sia esatta, ma le preoccupazioni legate all'impatto dell'intelligenza artificiale sui diritti fondamentali sono evidentemente in aumento e meritano di essere esaminate da chi si occupa di diritti umani.

La relazione dell'Agenzia dell'Unione europea per i diritti fondamentali (FRA) *Preparare un giusto futuro – L'intelligenza artificiale e i diritti fondamentali nell'UE* fornisce un'istantanea dell'uso attuale delle tecnologie legate all'IA nell'Unione europea (UE) e ne analizza le implicazioni per i diritti fondamentali. La relazione si concentra su casi d'uso in quattro ambiti fondamentali: prestazioni sociali, polizia predittiva, servizi sanitari e pubblicità mirata.

La presente sintesi illustra i principali risultati emersi dalla relazione.

Definizione di intelligenza artificiale

Non esiste una definizione universalmente accettata di intelligenza artificiale. Il concetto non si riferisce tanto ad applicazioni concrete, ma piuttosto riflette i recenti sviluppi tecnologici che comprendono una varietà di tecnologie.

La ricerca della FRA non ha applicato una definizione rigorosa di IA ai casi d'uso che presenta nella relazione principale. Per le interviste, l'IA è stata definita in modo ampio, con riferimento alla definizione fornita dal gruppo di esperti ad alto livello sull'IA della Commissione europea:

«Per intelligenza artificiale (IA) si intendono quei sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici. I sistemi basati sull'IA possono consistere solo in software che agiscono nel mondo virtuale (ad esempio assistenti vocali, software per l'analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale), oppure incorporare l'IA in dispositivi hardware (per esempio in robot avanzati, auto a guida autonoma, droni o applicazioni dell'Internet delle cose)».

*Questa **definizione iniziale** del gruppo di esperti è stata oggetto di ulteriori discussioni.*

*Cfr. gruppo di esperti ad alto livello sull'IA (2019), **A definition of AI: Main capabilities and disciplines** [Una definizione di IA: principali capacità e discipline].*

Quali aspetti riguardava la ricerca?

La FRA ha condotto una ricerca sul campo in cinque Stati membri dell'UE: Estonia, Finlandia, Francia, Paesi Bassi e Spagna. Sono state raccolte informazioni su come affrontano le questioni relative ai diritti fondamentali coloro che sono coinvolti nella progettazione e nell'utilizzo di sistemi di IA in settori chiave pubblici e privati.

La ricerca, basata su 91 interviste individuali, ha raccolto informazioni sui seguenti aspetti:

- finalità e applicazione pratica delle tecnologie di IA;
- valutazioni condotte durante l'utilizzo dell'IA, quadro giuridico e meccanismi di supervisione applicabili;
- consapevolezza delle questioni relative ai diritti fondamentali e potenziali garanzie messe in atto;
- programmi per il futuro.

Inoltre, sono stati intervistati dieci esperti coinvolti nel monitoraggio o nell'osservazione di potenziali violazioni dei diritti fondamentali riguardanti l'utilizzo dell'IA, tra cui membri della società civile, avvocati e organismi di controllo.

Per una descrizione più dettagliata della metodologia di ricerca e delle domande poste nelle interviste, cfr. l'allegato 1 alla relazione principale, disponibile sul [sito web della FRA](#).

«L'aspetto più importante è trattare i casi in modo più efficiente. Si tratta di utilizzare il proprio personale, le persone che gestiscono i casi, nel modo più efficiente possibile»

(Amministrazione pubblica, Paesi Bassi)

«Quando abbiamo testato il sistema, in realtà non abbiamo considerato gli aspetti giuridici, ma ci siamo chiesti se il sistema fosse redditizio»

(Società privata, Estonia)

Il lavoro della FRA su IA, megadati e diritti fondamentali

La relazione della FRA sull'IA e i diritti fondamentali è la pubblicazione principale derivante dal **progetto sull'intelligenza artificiale, sui megadati e sui diritti fondamentali**, che mira a valutare le implicazioni positive e negative delle nuove tecnologie, tra cui l'IA e i megadati, sui diritti fondamentali.

La relazione si basa sui risultati di una serie di documenti precedenti:

- **Facial recognition technology: fundamental rights considerations in the context of law enforcement** [Tecnologia del riconoscimento facciale: considerazioni sui diritti fondamentali nel contesto delle attività di contrasto] (2019): questo documento delinea e analizza i problemi in materia di diritti fondamentali che si presentano quando le autorità pubbliche impiegano questa tecnologia in presenza per le attività di contrasto. Inoltre, presenta brevemente le misure da adottare per evitare violazioni dei diritti.
- **Data quality and artificial intelligence — mitigating bias and error to protect fundamental rights** [Qualità dei dati e intelligenza artificiale: attenuare le distorsioni e gli errori per tutelare i diritti fondamentali] (2019): questo documento sottolinea l'importanza della consapevolezza e della prevenzione nell'ambito della scarsa qualità dei dati.
- **#BigData: Discrimination in data-supported decision making** [#Megadati: discriminazione nel processo decisionale supportato dai dati] (2018): questo documento di approfondimento illustra in che modo può verificarsi tale discriminazione e suggerisce possibili soluzioni.

Nell'ambito del progetto, la FRA sta anche esplorando la possibilità di studiare esempi concreti di sfide ai diritti fondamentali quando ci si avvale di algoritmi per il processo decisionale attraverso esperimenti online o studi di simulazione.



Quadro giuridico

Il quadro generale dei diritti fondamentali* che si applica all'utilizzo dell'IA nell'UE è costituito dalla Carta dei diritti fondamentali dell'Unione europea (la Carta) e dalla convenzione europea dei diritti dell'uomo.

Diversi altri strumenti in materia di diritti umani del Consiglio d'Europa e a livello internazionale sono pertinenti, tra cui la Dichiarazione universale dei diritti umani del 1948 e le principali convenzioni delle Nazioni Unite sui diritti umani**.

Inoltre, la legislazione derivata settoriale dell'UE, in particolare l'*acquis* in materia di protezione dei dati e la legislazione dell'Unione sulla non discriminazione, contribuisce a salvaguardare i diritti fondamentali nel contesto dell'IA. Infine, si applicano anche le leggi nazionali degli Stati membri dell'UE.

* Per ulteriori informazioni, cfr. FRA (2012), *Bringing rights to life: The fundamental rights landscape of the European Union*, Lussemburgo, Ufficio delle pubblicazioni dell'Unione europea.

** Tra queste importanti convenzioni vi sono: il Patto internazionale relativo ai diritti civili e politici del 1966; il Patto internazionale relativo ai diritti economici, sociali e culturali del 1966; la convenzione internazionale sull'eliminazione di ogni forma di discriminazione razziale del 1965; la convenzione sull'eliminazione di ogni forma di discriminazione della donna del 1979; la convenzione contro la tortura del 1984; la convenzione sui diritti del fanciullo del 1989; la convenzione sui diritti delle persone con disabilità del 2006 e la convenzione internazionale per la protezione di tutte le persone dalle sparizioni forzate del 2006.

Per ulteriori informazioni sul quadro giuridico internazionale universale dei diritti umani, compresi i loro meccanismi di applicazione, cfr. ad esempio De Schutter, O. (2015), *International Human Rights Law: Cases, Materials, Commentary*, Cambridge, Cambridge University Press, 2ª edizione.

SALVAGUARDARE I DIRITTI FONDAMENTALI: AMBITO DI APPLICAZIONE, VALUTAZIONI D'IMPATTO E RESPONSABILIZZAZIONE

Considerare l'intero ambito di applicazione dei diritti fondamentali in relazione all'IA

L'utilizzo di sistemi di intelligenza artificiale coinvolge una serie di diritti fondamentali, indipendentemente dal campo di applicazione. Tra questi figurano la vita privata, la protezione dei dati, la non discriminazione e l'accesso alla giustizia, ma l'elenco non è esaustivo.

La Carta dei diritti fondamentali dell'Unione europea (la Carta) è diventata giuridicamente vincolante nel dicembre 2009 e ha lo stesso valore giuridico dei trattati UE. Essa riunisce in un unico testo i diritti civili, politici, economici e sociali. Ai sensi dell'articolo 51, paragrafo 1, della Carta, le istituzioni, gli organi e gli organismi dell'Unione devono rispettare tutti i diritti sanciti dalla Carta e gli Stati membri dell'UE devono rispettarli nell'attuazione del diritto dell'Unione. Ciò vale per l'IA come per qualsiasi altro ambito.

Il lavoro sul campo di questa ricerca mostra che nella denominazione di IA rientra una grande varietà di sistemi. Le tecnologie analizzate comportano diversi livelli di automazione e complessità, e variano anche in termini di scala e di potenziale impatto sulle persone.

I risultati della FRA dimostrano che l'utilizzo di sistemi di IA coinvolge una serie di diritti fondamentali, indipendentemente dal campo di applicazione, tra cui figurano la vita privata, la protezione dei dati, la non discriminazione e l'accesso alla giustizia, ma tale elenco non è esaustivo. Tuttavia, quando si esamina l'impatto dell'IA rispetto ai diritti fondamentali, le interviste mostrano che l'ambito di applicazione è spesso limitato a diritti specifici.

Quando ci si avvale dell'IA è necessario considerare un insieme più ampio di diritti, a seconda della tecnologia e dell'ambito di utilizzo. Oltre ai diritti relativi alla vita privata e alla protezione dei dati, all'uguaglianza, alla non discriminazione e all'accesso alla giustizia, potrebbero essere presi in considerazione altri diritti, fra cui, ad esempio, la dignità umana, il diritto alla sicurezza sociale e all'assistenza sociale, il diritto a una buona amministrazione (per lo più rilevante per il settore pubblico) e la protezione dei consumatori (particolarmente importante per le imprese). A seconda del contesto d'utilizzo dell'IA, qualsiasi altro diritto tutelato dalla Carta deve essere preso in considerazione.



PARERE 1 DELLA FRA

Nell'introdurre nuove politiche e nell'adottare una nuova legislazione in materia di IA, il legislatore dell'UE e gli Stati membri, agendo nell'ambito del diritto dell'Unione, devono garantire che sia preso in considerazione il rispetto della gamma completa dei diritti fondamentali, come sancito dalla Carta e dai trattati dell'Unione europea. Specifiche salvaguardie dei diritti fondamentali devono accompagnare le politiche e le leggi pertinenti.

Nel fare ciò, l'UE e gli Stati membri dovrebbero basarsi su solide prove relative all'impatto dell'IA sui diritti fondamentali per garantire che eventuali restrizioni di determinati diritti fondamentali rispettino i principi di necessità e proporzionalità.

Le salvaguardie pertinenti devono essere previste dalla legge per tutelare efficacemente contro l'interferenza arbitraria con i diritti fondamentali e per garantire la certezza giuridica sia agli sviluppatori di IA, sia agli utenti. I sistemi volontari per l'osservazione e la salvaguardia dei diritti fondamentali nello sviluppo e nell'utilizzo dell'IA possono contribuire ulteriormente a mitigare le violazioni dei diritti. In sintonia con i requisiti minimi di chiarezza giuridica (in quanto principio fondamentale dello Stato di diritto e presupposto per la tutela dei diritti fondamentali), il legislatore deve prestare la dovuta attenzione nella definizione dell'ambito di applicazione di un'eventuale legge in materia di IA.

Data la varietà delle tecnologie che rientrano nel concetto di IA e la scarsa conoscenza circa l'intera portata del suo potenziale impatto sui diritti fondamentali, potrebbe essere necessario valutare regolarmente la definizione giuridica dei termini relativi all'IA.

PARERE 2 DELLA FRA

Il legislatore dell'UE dovrebbe prendere in considerazione la possibilità di rendere obbligatorie valutazioni d'impatto che coprano la gamma completa dei diritti fondamentali. Tali valutazioni dovrebbero riguardare sia il settore privato che quello pubblico ed essere effettuate prima che sia utilizzato qualsiasi sistema di IA, tenendo conto della diversa natura e ambito di applicazione delle tecnologie di IA, compreso il livello di automazione e complessità, nonché dei potenziali danni. Le valutazioni dovrebbero prevedere requisiti di screening di base, che possono anche servire ad aumentare la consapevolezza sulle potenziali implicazioni in materia di diritti fondamentali.

Dovrebbero basarsi sulle buone prassi consolidate in altri settori ed essere ripetute regolarmente durante la fase di implementazione dell'IA, se del caso. Dovrebbero essere condotte in modo trasparente e i loro risultati e raccomandazioni dovrebbero essere di pubblico dominio, per quanto possibile. Per agevolare il processo di valutazione d'impatto, le aziende e le amministrazioni pubbliche dovrebbero essere tenute a raccogliere le informazioni necessarie per valutare in modo approfondito il potenziale impatto sui diritti fondamentali.

L'UE e gli Stati membri dovrebbero prendere in considerazione azioni volte a supportare coloro che sviluppano, utilizzano o intendono utilizzare sistemi di IA, per garantire l'effettivo rispetto degli obblighi di valutazione d'impatto sui diritti fondamentali. Tali azioni potrebbero comprendere finanziamenti, linee guida, formazione o sensibilizzazione e dovrebbero essere rivolte in particolare, ma non esclusivamente, al settore privato.

L'UE e gli Stati membri dovrebbero prendere in considerazione l'utilizzo degli strumenti esistenti, come le liste di controllo o gli strumenti di autovalutazione, elaborati a livello europeo e internazionale, inclusi quelli elaborati dal gruppo di esperti ad alto livello dell'UE sull'intelligenza artificiale.

Utilizzare valutazioni d'impatto efficaci per prevenire gli effetti negativi

Le valutazioni d'impatto preventive si concentrano principalmente sulle questioni tecniche e raramente si occupano dei potenziali effetti sui diritti fondamentali, poiché le conoscenze circa il modo in cui l'IA influisce su tali diritti sono scarse.

L'utilizzo di sistemi di intelligenza artificiale interagisce con un ampio spettro di diritti fondamentali, indipendentemente dal campo di applicazione. Ai sensi dell'articolo 51, paragrafo 1, della Carta, gli Stati membri dell'UE devono rispettare tutti i diritti sanciti dalla Carta quando attuano il diritto dell'Unione. In sintonia con gli standard internazionali in vigore, in particolare con i Principi guida delle Nazioni Unite su imprese e diritti umani, le imprese dovrebbero avere in atto «un processo di diligenza in materia di diritti umani per individuare, prevenire, mitigare e rendere conto del modo in cui affrontano il loro impatto sui diritti umani» (principi 15 e 17). Ciò è indipendente dalle dimensioni e dal settore, e comprende le imprese che lavorano con l'IA.

Nel perseguire i propri impegni in riferimento ai principi guida menzionati, l'UE ha adottato diversi atti legislativi che riguardano strumenti specifici per settore, in particolare nel contesto degli obblighi di diligenza in materia di diritti umani. Sono attualmente in corso discussioni sulla proposta di nuove norme di diritto derivato dell'UE, che richiederebbero alle imprese di esercitare la dovuta diligenza in relazione ai potenziali impatti sui diritti umani e sull'ambiente delle loro operazioni e catene di approvvigionamento. Tali norme sarebbero probabilmente intersettoriali e prevedrebbero sanzioni per la loro mancata osservanza, incluso in caso di utilizzo dell'IA. Cfr. la recente relazione della FRA *Business and Human rights – access to remedy* [Imprese e diritti umani: accesso alle misure correttive], che chiede un miglioramento delle norme orizzontali di diligenza in materia di diritti umani per le imprese con sede nell'UE.

Le valutazioni d'impatto sono uno strumento importante sia per le imprese che per le amministrazioni pubbliche per mitigare il potenziale impatto negativo delle loro attività sui diritti fondamentali. Il diritto dell'UE in settori specifici richiede alcune forme di valutazioni d'impatto, come le valutazioni d'impatto sulla protezione dei dati ai sensi del regolamento generale sulla protezione dei dati (RGPD). Molti intervistati hanno riferito che è stata condotta una valutazione d'impatto sulla protezione dei dati, come richiesto dalla norma, che tuttavia ha assunto forme diverse. Inoltre, le valutazioni preventive, quando vengono condotte, si concentrano principalmente sugli aspetti tecnici, mentre raramente affrontano i potenziali impatti sui diritti fondamentali. Secondo alcuni intervistati le valutazioni d'impatto sui diritti fondamentali non vengono effettuate quando un sistema di IA non ha, o non sembra avere, effetti negativi sui diritti fondamentali.

La ricerca mostra che le conoscenze degli intervistati sui diritti fondamentali (diversi dalla protezione dei dati e, in una certa misura, dalla non discriminazione) sono limitate. La maggioranza riconosce tuttavia che l'utilizzo dell'IA ha un impatto su tali diritti. Alcuni intervistati affermano che i loro sistemi non incidono sui diritti fondamentali, il che è in parte legato al tipo di compiti per i quali sono utilizzati i sistemi di IA.

Tutti gli intervistati sono consapevoli delle questioni relative alla protezione dei dati e la maggior parte si rende anche conto del fatto che, in generale, potrebbero esserci problemi di discriminazione quando ci si avvale dell'IA. Tuttavia, l'esatto significato e l'applicabilità dei diritti relativi alla protezione dei dati e alla non discriminazione non sono ancora chiari a molti intervistati.

I risultati della ricerca mostrano differenze tra il settore privato e quello pubblico: gli intervistati del settore privato sono spesso meno consapevoli dell'ampio insieme di diritti fondamentali che potrebbero essere interessati. Le questioni relative alla protezione dei dati sono note nel settore privato, ma altri diritti, come la non discriminazione o l'accesso ai diritti relativi alla giustizia, sono meno conosciuti tra i rappresentanti delle imprese che lavorano con l'IA. Alcuni erano pienamente consapevoli dei potenziali problemi, mentre altri hanno affermato che la responsabilità di verificare le questioni relative ai diritti fondamentali spetta ai clienti.



PARERE 3 DELLA FRA

L'UE e gli Stati membri dovrebbero garantire l'esistenza di sistemi di responsabilizzazione efficaci per monitorare e, se necessario, affrontare efficacemente gli eventuali impatti negativi dei sistemi di IA sui diritti fondamentali. Dovrebbero prendere in considerazione, oltre alle valutazioni d'impatto sui diritti fondamentali (cfr. parere 2 della FRA), l'introduzione di garanzie specifiche per assicurare l'efficacia del regime di responsabilizzazione. Ciò potrebbe prevedere l'obbligo giuridico di mettere a disposizione informazioni sufficienti per consentire una valutazione dell'impatto dei sistemi di IA sui diritti fondamentali, il che consentirebbe il monitoraggio esterno e la supervisione dei diritti umani da parte degli organismi competenti.

L'UE e gli Stati membri dovrebbero inoltre avvalersi maggiormente delle strutture di sorveglianza specializzate esistenti, per tutelare i diritti fondamentali nell'utilizzo dell'IA, tra cui le autorità per la protezione dei dati, gli organismi che tutelano l'uguaglianza, le istituzioni nazionali per i diritti umani, i mediatori e gli organismi per la protezione dei consumatori.

Dovrebbero essere stanziati risorse aggiuntive per istituire sistemi di responsabilizzazione efficaci attraverso lo sviluppo delle competenze e la diversificazione del personale che lavora per gli organismi di sorveglianza. Ciò consentirebbe loro di affrontare questioni complesse legate allo sviluppo e all'utilizzo dell'IA.

Analogamente, gli organismi appropriati dovrebbero essere dotati di risorse, poteri e, segnatamente, competenze sufficienti per prevenire e valutare le violazioni dei diritti fondamentali, nonché sostenere efficacemente coloro i cui diritti fondamentali sono interessati dall'IA.

Agevolare la cooperazione tra gli organismi appropriati a livello nazionale ed europeo può contribuire a condividere le competenze e le esperienze. Anche il coinvolgimento di altri attori con competenze pertinenti, come le organizzazioni specializzate della società civile, può essere d'aiuto. Nell'attuare tali azioni a livello nazionale, gli Stati membri dovrebbero prendere in considerazione l'utilizzo dei meccanismi di finanziamento dell'UE disponibili.

Garantire una supervisione efficace e una responsabilizzazione generale

Le imprese e le amministrazioni pubbliche che sviluppano e utilizzano l'IA sono in contatto con vari organismi responsabili della supervisione dei sistemi relativi all'IA nell'ambito dei rispettivi mandati e settori, incluse le autorità per la protezione dei dati. Gli utilizzatori dell'IA, però, non sanno sempre con certezza quali sono gli organismi responsabili della sorveglianza dei sistemi di IA.

In sintonia con le norme internazionali consolidate in materia di diritti umani, ad esempio l'articolo 1 della convenzione europea dei diritti dell'uomo (CEDU) e l'articolo 51 della Carta, gli Stati sono tenuti a garantire i diritti e le libertà delle persone. Per conformarsi efficacemente, gli Stati devono tra l'altro mettere in atto meccanismi efficaci di monitoraggio e di attuazione, e ciò vale anche per l'IA.

A livello di monitoraggio, i risultati evidenziano l'importante ruolo degli organismi specializzati istituiti in settori specifici che, nell'ambito dei loro mandati, sono anche responsabili della sorveglianza in materia di IA. Tra questi figurano, ad esempio, la sorveglianza nel settore bancario o le autorità per la protezione dei dati. Diversi organismi sono potenzialmente rilevanti per la sorveglianza in materia di IA dal punto di vista dei diritti fondamentali, ma le loro responsabilità per quanto riguarda la sorveglianza in materia di IA non sono ancora chiare a molti degli intervistati del settore privato e pubblico.

L'utilizzo dell'IA da parte delle pubbliche amministrazioni è talvolta oggetto di audit nell'ambito delle verifiche periodiche. Anche le imprese private di settori specifici dispongono di organismi di sorveglianza specializzati, ad esempio nell'ambito della sanità o dei servizi finanziari, i quali controllano anche l'utilizzo dell'IA e delle tecnologie correlate, ad esempio come parte dei rispettivi sistemi di certificazione. Gli intervistati del settore privato desiderano disporre di organismi che possano fornire una consulenza esperta sulle possibilità e sulla legittimità dei potenziali usi dell'IA.

Nell'UE esiste un insieme ben sviluppato di organismi indipendenti incaricati di tutelare e promuovere i diritti fondamentali, tra cui le autorità per la protezione dei dati, gli organismi che tutelano l'uguaglianza, le istituzioni nazionali per i diritti umani e i mediatori. La

ricerca mostra che coloro che utilizzano o progettano di utilizzare l'IA spesso si rivolgono a vari organismi per discutere dell'utilizzo, attuale o futuro, dell'IA, come ad esempio gli organismi per la protezione dei consumatori.

Il più delle volte gli utenti di IA hanno contattato le autorità per la protezione dei dati per ricevere orientamenti, suggerimenti o un'approvazione in caso di trattamento di dati personali. Gli esperti intervistati sottolineano l'importanza delle autorità per la protezione dei dati per la supervisione dei sistemi di IA in relazione all'uso dei dati personali. Tuttavia osservano anche che esse non dispongono di risorse sufficienti per questo compito, né di competenze specifiche in materia di IA.

Gli esperti, compresi quelli che lavorano per le autorità di sorveglianza, come gli organismi che tutelano l'uguaglianza e le autorità per la protezione dei dati, concordano sulla necessità di rafforzare le competenze degli organismi di sorveglianza esistenti affinché possano svolgere efficacemente una funzione di controllo sulle questioni relative all'IA. Riconoscono tuttavia che ciò può essere impegnativo, dato che le risorse di questi organismi sono già molto limitate e hanno anche evidenziato l'importante ruolo svolto dalle organizzazioni pertinenti della società civile specializzate nel campo della tecnologia, dei diritti digitali e degli algoritmi, che possono incentivare la responsabilizzazione nell'utilizzo dei sistemi di IA.



NON DISCRIMINAZIONE, PROTEZIONE DEI DATI E ACCESSO ALLA GIUSTIZIA: TRE TEMI ORIZZONTALI

La ricerca mostra che l'utilizzo dell'IA influisce su diversi diritti fondamentali. Oltre agli aspetti specifici legati al contesto che incidono in misura diversa sui diversi diritti, i temi dei diritti fondamentali emersi dalla ricerca che si applicano ripetutamente alla maggior parte dei casi in materia di IA includono: la necessità di garantire un utilizzo non discriminatorio dell'IA (diritto a non essere discriminati); l'obbligo di trattare i dati nel rispetto delle norme (diritto alla protezione dei dati personali); la possibilità di presentare un reclamo e di fare ricorso riguardo a decisioni basate sull'IA (diritto a un ricorso effettivo e a un processo equo).

I due principali diritti fondamentali evidenziati nelle interviste sono la protezione dei dati e la non discriminazione, inoltre sono stati spesso citati metodi efficaci per presentare un reclamo riguardo all'utilizzo dell'IA, in relazione al diritto a un processo equo e a un ricorso effettivo. I tre pareri della FRA che seguono, in linea con questi risultati, dovrebbero essere letti insieme agli altri pareri, che richiedono un riconoscimento più completo e una risposta più ampia a tutto l'insieme dei diritti fondamentali interessati dall'IA.

PARERE 4 DELLA FRA

Gli Stati membri dell'UE dovrebbero prendere in considerazione la possibilità di incoraggiare le aziende e le amministrazioni pubbliche a valutare eventuali risultati potenzialmente discriminatori nell'utilizzo dei sistemi di IA.

La Commissione europea e gli Stati membri dovrebbero prendere in considerazione la possibilità di finanziare ricerche mirate sugli impatti potenzialmente discriminatori dell'utilizzo dell'IA e degli algoritmi. Tali ricerche trarrebbero beneficio dall'adattamento di metodologie di ricerca consolidate delle scienze sociali, che vengono impiegate per individuare la potenziale discriminazione in diversi ambiti, dal reclutamento alla profilazione dei clienti.

Sulla base dei risultati di tali ricerche, dovrebbero essere elaborati linee guida e strumenti per aiutare coloro che utilizzano l'IA a individuare possibili risultati discriminatori.

Garanzie specifiche per assicurare la non discriminazione nell'utilizzo dell'IA

Raramente gli intervistati hanno dichiarato di aver svolto valutazioni dettagliate in merito alla potenziale discriminazione nell'utilizzo dell'IA. Ciò suggerisce una mancanza di valutazioni approfondite di tale discriminazione nel processo decisionale automatizzato.

L'obbligo di rispettare il principio di non discriminazione è sancito dall'articolo 2 del TUE, dall'articolo 10 del TFUE (che impone all'Unione di combattere la discriminazione fondata su una serie di motivi) e dagli articoli 20 e 21 della Carta (uguaglianza dinanzi alla legge e non discriminazione fondata su una serie di motivi). Anche disposizioni più specifiche e dettagliate in diverse direttive dell'UE sanciscono questo principio, con vari ambiti di applicazione.

L'automazione e l'utilizzo dell'IA possono aumentare notevolmente l'efficienza dei servizi ed estendersi ai compiti che gli esseri umani non sarebbero in grado di svolgere, ma è necessario garantire che i servizi e le

decisioni basati sull'IA non siano discriminatori. Riconoscendo ciò, la Commissione europea ha recentemente evidenziato la necessità di disposizioni supplementari per salvaguardare la non discriminazione nell'utilizzo dell'IA nel **piano d'azione dell'UE contro il razzismo 2020-2025**.

La maggior parte degli intervistati è in linea di principio consapevole del fatto che potrebbero verificarsi discriminazioni, ma raramente hanno sollevato questo problema in prima persona. Soltanto alcuni credono che i loro sistemi possano effettivamente fare discriminazioni.

Gli intervistati hanno anche menzionato raramente valutazioni dettagliate della potenziale discriminazione, il che indica la mancanza di una valutazione approfondita in questo senso.

Si crede comunemente che omettere informazioni su attributi protetti, come il genere, l'età o l'origine etnica, possa garantire che un sistema di IA non discrimini, però questo non è necessariamente vero. Le informazioni che potenzialmente indicano caratteristiche protette (proxy), che spesso si trovano nei set di dati, potrebbero comportare una discriminazione.

In alcuni casi i sistemi di IA possono anche essere utilizzati per testare e rilevare comportamenti discriminatori, che possono essere codificati nei set di dati. Tuttavia, pochissimi intervistati hanno menzionato la possibilità di raccogliere tali informazioni sui gruppi svantaggiati per individuare potenziali discriminazioni. In assenza di un'analisi approfondita della potenziale discriminazione nell'utilizzo effettivo dei sistemi di IA, sono praticamente assenti anche la discussione e l'analisi del potenziale effetto positivo dell'utilizzo di algoritmi per rendere più eque le decisioni. Inoltre, nessuno degli intervistati che lavora nell'ambito dell'IA ha citato l'utilizzo dell'IA volto a rilevare una possibile discriminazione come risultato positivo, nel senso che la discriminazione può essere rilevata in modo più efficace quando i dati sono analizzati per potenziali distorsioni.

Poiché individuare potenziali discriminazioni attraverso l'utilizzo dell'IA e degli algoritmi rimane una sfida, e gli intervistati hanno affrontato la questione solo brevemente, sono necessarie diverse misure in questo senso, tra cui l'obbligo di prendere in considerazione le questioni legate alla discriminazione quando si valuta l'utilizzo dell'IA e di investire in ulteriori studi sulla potenziale discriminazione avvalendosi di una gamma diversificata di metodologie, come ad esempio test di discriminazione.

Ciò potrebbe basarsi su metodologie simili e consolidate per testare le distorsioni nella vita di tutti i giorni, ad esempio per quanto riguarda le domande di lavoro, in cui il nome del candidato viene cambiato per identificarne (indirettamente) l'etnia. In relazione alle domande di IA, tali test potrebbero prevedere la possibile creazione di profili falsi per gli strumenti online, che differiscono solo per quanto riguarda gli attributi protetti. In questo modo i risultati possono essere verificati rispetto a potenziali discriminazioni. La ricerca potrebbe anche trarre vantaggio da un'analisi statistica avanzata volta a individuare le differenze nei set di dati relativi ai gruppi protetti e fungere quindi come base per studiare la potenziale discriminazione.

Infine, alcune interviste condotte nell'ambito della ricerca hanno sottolineato che i risultati di algoritmi di apprendimento automatico complessi sono spesso molto difficili da comprendere e spiegare. Pertanto, ulteriori ricerche per comprendere e spiegare meglio tali risultati (la cosiddetta «intelligenza artificiale spiegabile») possono anche aiutare a rilevare meglio la discriminazione nell'utilizzo dell'IA.



PARERE 5 DELLA FRA

Il Comitato europeo per la protezione dei dati (EDPB) e il Garante europeo della protezione dei dati (GEPD) dovrebbero prendere in considerazione la possibilità di fornire ulteriori orientamenti e sostegno al fine di attuare efficacemente le disposizioni del RGPD che si applicano direttamente all'utilizzo dell'IA per la salvaguardia dei diritti fondamentali, in particolare per quanto riguarda il significato dei dati personali e il loro utilizzo nell'IA, anche nei set di dati per la formazione sull'IA.

Vi è un elevato livello di incertezza riguardo al significato del processo decisionale automatizzato e al diritto alla revisione umana legato all'utilizzo dell'IA e al processo decisionale automatizzato. Pertanto, l'EDPB e il GEPD dovrebbero anche prendere in considerazione la possibilità di chiarire ulteriormente i concetti di «processo decisionale automatizzato» e «revisione umana», laddove sono menzionati nel diritto dell'UE.

Inoltre, gli organismi nazionali per la protezione dei dati dovrebbero fornire orientamenti pratici sul modo in cui le disposizioni in materia di protezione dei dati si applicano all'utilizzo dell'IA. Tali orientamenti potrebbero includere raccomandazioni e liste di controllo, basate su casi concreti di utilizzo dell'IA, per agevolare il rispetto delle disposizioni in materia di protezione dei dati.

Ulteriori indicazioni sulla protezione dei dati

Occorre maggiore chiarezza sull'ambito di applicazione e sul significato delle disposizioni giuridiche relative al processo decisionale automatizzato.

La protezione dei dati è fondamentale nello sviluppo e nell'utilizzo dell'IA. L'articolo 8, paragrafo 1, della Carta e l'articolo 16, paragrafo 1, del TFUE stabiliscono che ognuno ha diritto alla protezione dei propri dati personali. Il RGPD e la direttiva sull'applicazione della legge – direttiva (UE) 2016/680 – approfondiscono ulteriormente questo diritto e contengono una serie di disposizioni applicabili all'utilizzo dell'IA.

Gli intervistati hanno indicato che la maggior parte dei sistemi di IA da loro impiegati fa uso di dati personali, il che significa che la protezione dei dati è influenzata in molti modi diversi. Tuttavia, secondo gli intervistati alcune applicazioni non utilizzano dati personali, o usano solo dati anonimizzati, pertanto non si applicherebbe la normativa sulla protezione dei dati. In caso di utilizzo di dati personali si applicano tutti i principi e le disposizioni in materia di protezione dei dati.

Questa relazione evidenzia un'importante questione legata alla protezione dei dati, che è pertinente anche per altri diritti fondamentali in relazione al processo decisionale automatizzato. Secondo un sondaggio dell'Eurobarometro, solo il 40 % dei cittadini europei sa che può avere voce in capitolo in caso di decisioni automatizzate. La conoscenza di questo diritto è notevolmente più elevata tra coloro che lavorano con

l'IA, infatti la maggior parte degli intervistati ha sollevato la questione. Tuttavia molti degli intervistati, compresi gli esperti, hanno sostenuto che è necessaria maggiore chiarezza circa l'ambito di applicazione e il significato delle disposizioni giuridiche relative al processo decisionale automatizzato.

Nell'ambito delle prestazioni sociali, gli intervistati hanno citato solo un esempio di decisioni completamente automatizzate e basate su regole, mentre tutte le altre domande da loro menzionate sono riviste da esseri umani. Gli intervistati nella pubblica amministrazione hanno sottolineato l'importanza della revisione umana di qualsiasi decisione, ma raramente hanno descritto in che cosa consiste effettivamente tale revisione umana e come sono state utilizzate altre informazioni durante l'esame dei risultati dei sistemi di IA.

Sebbene ci siano opinioni divergenti tra gli intervistati sul fatto che la legislazione esistente sia o meno sufficiente, molti hanno chiesto un'interpretazione più concreta delle norme esistenti in materia di protezione dei dati per quanto riguarda il processo decisionale automatizzato, come sancito dall'articolo 22 del RGPD.

Accesso effettivo alla giustizia nei casi che comportano decisioni basate sull'IA

Per contestare efficacemente le decisioni basate sull'utilizzo dell'IA le persone devono essere a conoscenza dell'utilizzo dell'IA, e delle modalità e delle sedi in cui presentare un reclamo. Le organizzazioni che si avvalgono dell'IA devono saper spiegare il loro sistema e le decisioni basate sull'IA.

L'accesso alla giustizia è al contempo un processo e un obiettivo ed è cruciale per coloro che cercano di beneficiare di altri diritti procedurali e sostanziali. Esso comprende una serie di diritti umani fondamentali, fra cui il diritto a un processo equo e a un ricorso effettivo ai sensi degli articoli 6 e 13 della CEDU e dell'articolo 47 della Carta dei diritti fondamentali dell'Unione europea. Di conseguenza, la nozione di accesso alla giustizia obbliga gli Stati a garantire a ciascun individuo il diritto di rivolgersi al tribunale (o, in alcune circostanze, a un organismo di risoluzione alternativa delle controversie) per ottenere il ricorso qualora si accerti che i diritti della persona sono stati violati.

In conformità a questi standard, la vittima di una violazione dei diritti umani derivante dallo sviluppo o dall'utilizzo di un sistema di IA da parte di un ente pubblico o privato deve poter ricorrere dinanzi a un'autorità nazionale. Conformemente alla giurisprudenza pertinente ai sensi dell'articolo 47 della Carta e dell'articolo 13 della CEDU, il ricorso deve essere «effettivo nella prassi e nella legge».

I risultati della ricerca individuano i seguenti prerequisiti affinché il ricorso sia effettivo nella prassi nei casi che riguardano i sistemi di IA e il loro impatto sui diritti fondamentali: tutti devono essere consapevoli dell'utilizzo dell'IA e devono essere informati su come e dove presentare reclamo. Le organizzazioni che si avvalgono dell'IA devono garantire che il pubblico sia informato sul loro sistema di IA e sulle decisioni basate su di esso.

I risultati dimostrano che spiegare i sistemi di IA e il modo in cui prendono le decisioni in termini semplici può essere difficile. I diritti di proprietà intellettuale possono ostacolare la messa a disposizione di informazioni dettagliate sul funzionamento di un algoritmo. Inoltre, alcuni sistemi di IA sono complessi e ciò rende difficile condividere informazioni significative sul funzionamento di un sistema e sulle relative decisioni.

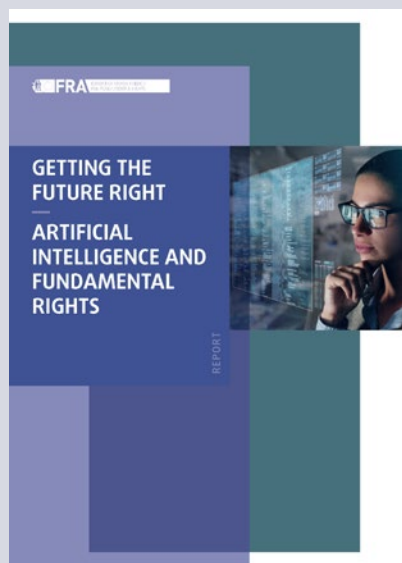
Per far fronte a questo problema, alcune aziende intervistate evitano del tutto di utilizzare metodi complessi per determinate decisioni, perché non sarebbero in grado di spiegarle. In alternativa si avvalgono di metodi di analisi dei dati più semplici per lo stesso problema, al fine di ottenere una certa comprensione dei principali fattori che influenzano determinati risultati. Alcuni degli intervistati del settore privato hanno sottolineato gli sforzi compiuti per migliorare gradualmente la loro comprensione della tecnologia dell'IA.

PARERE 6 DELLA FRA

Il legislatore dell'UE e gli Stati membri dovrebbero garantire ai singoli individui un accesso effettivo alla giustizia nei casi che comportano decisioni basate sull'IA.

Per garantire che i ricorsi disponibili siano accessibili nella prassi, il legislatore dell'UE e gli Stati membri potrebbero prendere in considerazione l'introduzione di un obbligo giuridico per le amministrazioni pubbliche e le imprese private che utilizzano sistemi di IA di fornire informazioni sul funzionamento dei loro sistemi di IA a coloro che presentano un ricorso, incluse le informazioni sul modo in cui questi sistemi di IA giungono alle decisioni automatizzate. Quest'obbligo contribuirebbe a conseguire l'uguaglianza dei mezzi in caso di ricorso alla giustizia da parte di individui e agevolerebbe l'efficacia del monitoraggio esterno e della supervisione in materia di diritti umani dei sistemi di IA (cfr. parere 3 della FRA).

Considerata la difficoltà di spiegare sistemi di IA complessi, l'UE, insieme agli Stati membri, dovrebbe prendere in considerazione l'elaborazione di linee guida per sostenere gli sforzi di trasparenza in quest'ambito. Per fare ciò si dovrebbe attingere alle competenze degli organismi nazionali per i diritti umani e delle organizzazioni della società civile attive in questo campo.



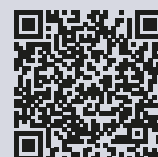
Questa sintesi presenta i risultati principali della relazione della FRA *Preparare un giusto futuro – L'intelligenza artificiale e i diritti fondamentali*. La relazione principale è disponibile sul [sito web della FRA](#).



PROMUOVERE E PROTEGGERE I DIRITTI FONDAMENTALI IN TUTTA L'UE

La relazione della FRA sull'intelligenza artificiale e i diritti fondamentali presenta esempi concreti di come le aziende e le pubbliche amministrazioni dell'UE utilizzano, o cercano di utilizzare, l'intelligenza artificiale. Si concentra su quattro ambiti fondamentali: prestazioni sociali, polizia predittiva, servizi sanitari e pubblicità mirata. La relazione esamina le potenziali implicazioni per i diritti fondamentali e analizza in che modo tali diritti sono presi in considerazione quando si utilizzano o si sviluppano applicazioni di IA.

La presente sintesi illustra i principali risultati emersi dalla relazione, che possono fornire un orientamento per gli sforzi della politica dell'UE e a livello nazionale volti a regolamentare l'utilizzo degli strumenti di IA nel rispetto dei diritti umani e dei diritti fondamentali.



FRA – AGENZIA DELL'UNIONE EUROPEA PER I DIRITTI FONDAMENTALI
Schwarzenbergplatz 11 – 1040 Vienna – Austria
Tel. +43 158030-0 – Fax +43 158030-699

fra.europa.eu

 facebook.com/fundamentalrights
 twitter.com/EURightsAgency
 linkedin.com/company/eu-fundamental-rights-agency



Ufficio delle pubblicazioni
dell'Unione europea