ApertaContrada
RIFLESSIONI SU SOCIETÀ. DIRITTO. ECONOMIA

L'executive order del presidente Biden: le politiche sull'IA da una sponda all'altra dell'oceano.

di *Fulvio Costantino* - 14 Dicembre 2023

Il 30 ottobre, attraverso l'*executive order* del Presidente degli Stati Uniti, che è un atto di indirizzo rivolto alle agenzie federali[1], si è proposta dall'altra parte dell'oceano una visione politica e strategica dell'intelligenza artificiale (IA) attraverso un quadro di interventi molto ampio, il che per un verso pone il tema dei divieti e obblighi relativi all'impiego di questo strumento, ritenuti necessari ma ancora da disciplinare, per altro verso fa capire la centralità del tema.

Mediante un approccio globale si vuole così mostrare, anche sullo scenario internazionale, di essere in grado di fornire una visione non parcellizzata del fenomeno dell'IA; si vuole comunicare inoltre come l'approccio non sia esclusivamente mirato a comandare e controllare, e ciò anche per rispondere a chi vede con preoccupazione che il fenomeno possa essere ingabbiato con delle regole stringenti e che quindi l'innovazione sia frenata.

L'ordine esecutivo non rivoluziona i principi già indicati sia dai comitati etici che, a livello giuridico, dal processo in corso nell'UE: si invocano *standard* per la sicurezza e la protezione dell'IA, protezione della *privacy*, promozione di equità e diritti civili, difesa di consumatori e lavoratori. Più in generale, si osserva come si debba investire in innovazione e concorrenza, con l'obiettivo di far avanzare la *leadership* americana nel mondo: l'intento è di essere di esempio per il settore privato e i governi di tutto il mondo.

Al centro dell'attenzione c'è l'idea di affrontare le sfide globali, in particolare lo sviluppo sostenibile, e mitigare i pericoli per le infrastrutture critiche[2]. L'intervento è collocato nella cd. strategia globale per l'innovazione responsabile.

Per un verso si adotta un approccio basato sul rischio come nell'UE, sebbene nella regolazione UE il tasso di rischio, almeno in prima approssimazione, sia chiaramente codificato; per altro verso si pone la necessità di fissare degli obblighi in capo agli sviluppatori di IA. Questo appare evidente nell'*order*, ove è previsto che gli sviluppatori dei sistemi di IA condividano i risultati dei *test* di sicurezza e altre informazioni "critiche" con il governo, laddove ci sia un "serio rischio" per la sicurezza nazionale, la sicurezza economica nazionale o la salute e l'incolumità pubblica nazionale: ciò dovrà avvenire durante la formazione, prima che gli strumenti siano resi pubblici, per garantire che questi siano sicuri, protetti e affidabili. La condivisione dei risultati è fondata sul *Defense Production Act*, e si è osservato come ad esso generalmente si ricorra in casi di emergenza (2011, crisi energetica in California, 2021, produzione di vaccini), e in ambito

tecnologico, vi si sia ricorso in passato, in presenza di timori di spionaggio da parte della Cina, quando vennero obbligate le aziende di telecomunicazioni a riferire sui componenti di provenienza estera[3].

Inoltre, si prevede lo sviluppo di *standard* e *test* consensuali, dal momento che è evidentemente attraverso il loro impiego, aggiornamento e controllo che sarà reso possibile un efficace impiego dell'IA, a fronte di rischi particolarmente gravi. L'*order* prevede infatti che questi strumenti, redatti sotto la guida del *National Institute of Standards and Technology*, siano applicati, per i settori delle infrastrutture critiche, dal Dipartimento per la sicurezza interna, il quale istituirà un *AI Safety and Security Board*; i Dipartimenti dell'Energia e della Sicurezza Nazionale affronteranno non solo le minacce alle infrastrutture critiche, ma anche i rischi chimici, biologici, radiologici, nucleari e di sicurezza energetica. In questo senso, le società dovranno fornire al governo federale, su base continuativa, informazioni, rapporti o registrazioni. In ordine alla progettazione tramite IA di materiali biologici pericolosi, le agenzie che finanziano i progetti stabiliranno degli *standard*, il cui rispetto sarà condizione per il finanziamento federale. Periodicamente ogni agenzia dovrà presentare una valutazione dei potenziali rischi nei settori delle infrastrutture critiche coinvolti, per guasti critici, attacchi fisici e attacchi informatici.

Appare evidente poi come il controllo dell'IA vada effettuata anche con lo scopo di evitare minacce da altri paesi. Si prevede infatti che il Consiglio di sicurezza nazionale e il Capo di stato maggiore della Casa Bianca dovranno sviluppare un *memorandum* sulla sicurezza nazionale, così che la comunità militare e di *intelligence* degli Stati Uniti utilizzino l'IA in modo sicuro, etico ed efficace nelle loro missioni, anche per contrastare l'uso militare nemico dell'IA. È previsto anche un *memorandum* sulla *governance* dell'IA, come componente di un sistema di sicurezza nazionale o per scopi militari e di *intelligence*. Su questo punto, peraltro, come si vedrà oltre, appare maggiore la distanza con il modello dell'Unione Europea, per limiti intrinseci della stessa. Ad ogni modo, sempre nell'ambito della sicurezza, si dovrà stabilire un programma avanzato di sicurezza informatica con l'IA per individuare e correggere le vulnerabilità nei *software* critici, per rendere essi e le reti più sicuri.

Si interviene anche contro l'uso malevolo dell'IA, sia a scopo commerciale che politico, e per tale ragione, per prevenire truffe, manipolazioni e notizie false, vengono prospettati *standard* e migliori pratiche per rilevare i contenuti generati dall'IA, e il Dipartimento del Commercio svilupperà linee guida, per autenticare i contenuti governativi ufficiali ed etichettare i contenuti generati dall'IA.

In ordine alla protezione della *privacy*, che finora non è stata al centro delle attenzioni federali rispetto all'UE, ci si preoccupa che vadano rafforzati gli strumenti crittografici, e che si valuti il modo in cui le agenzie raccolgono e utilizzano le informazioni disponibili in commercio, contenenti dati personali identificabili; si invoca inoltre una legge per la protezione dei dati personali all'interno dell'addestramento dei modelli AI.

Grande spazio è dato al tema dei diritti civili e all'equità, per evitare discriminazioni, pregiudizi e altri abusi in particolare nel campo della giustizia, dell'assistenza sanitaria e dell'edilizia abitativa: si richiamano il progetto di Carta dei diritti dell'intelligenza artificiale e l'ordine esecutivo rivolto

alle agenzie per combattere la discriminazione algoritmica (n. 14091 del 16 febbraio 2023). Si deve provvedere alla formazione, l'assistenza tecnica e il coordinamento tra il Dipartimento di Giustizia e gli uffici federali per i diritti civili. In particolare, in ordine ai diritti civili nel sistema di giustizia penale, si dovrà affrontare la discriminazione nell'uso di sistemi automatizzati, migliorare il coinvolgimento delle parti interessate esterne per promuovere la consapevolezza pubblica sui potenziali usi ed effetti discriminatori dell'IA, preoccupandosi in particolare dell'uso dell'IA da parte delle forze dell'ordine[4]. Nel settore sanitario si fa riferimento allo sviluppo di farmaci convenienti e salvavita, oltre che all'attenzione che il Dipartimento della Salute e dei Servizi Umani dovrà avere per i danni o pratiche sanitarie non sicure. Nel settore dell'istruzione, si pensa a supportare gli educatori nell'implementazione di strumenti educativi abilitati all'IA, come il tutoraggio personalizzato nelle scuole.

Legato a questo profilo è quello dell'uso responsabile ed efficace dell'IA da parte del governo per regolare, governare ed erogare benefici e in materia di appalti. In particolare, si mira alla valutazione dell'accesso ai benefici, all'avviso agli interessati circa la presenza di un sistema di IA, a forme di valutazione periodica per individuare rifiuti ingiusti, a processi per mantenere adeguati livelli di discrezionalità del personale esperto dell'agenzia, alla possibilità di appellarsi ai revisori umani, all'analisi dei sistemi algoritmici. Sul piano organizzativo, andranno aiutate le agenzie ad acquisire prodotti e servizi IA specifici in modo più rapido, economico e più efficace attraverso contrattazioni più rapide ed efficienti, si dovrà accelerare la rapida assunzione di professionisti, le agenzie dovranno fornire formazione ai dipendenti, si dovranno fornire indicazioni chiare ai proprietari, agli esecutori di programmi di benefici e agli appaltatori.

Attenzione è correttamente posta anche in ordine al tema del lavoro: si pone il tema del pericolo di una maggiore sorveglianza sul posto di lavoro, nonché dei rischi di pregiudizi derivanti dall'uso dell'IA. Si deve evitare che i lavoratori siano sottopagati, o che siano valutate ingiustamente le domande di lavoro o sia compromessa la capacità dei lavoratori di organizzarsi. È da valutare anche l'impatto dell'IA sul mercato del lavoro, in ordine al rischio di licenziamenti, di spostamento del lavoro, ma anche alle opportunità. Si promuovono misure di rafforzamento del sostegno ai lavoratori che affrontano interruzioni del lavoro, ma anche l'immigrazione e il coinvolgimento di profili altamente qualificati e con esperienza in aree critiche (e a tal fine si chiede di modernizzare e razionalizzare i criteri, i colloqui e le revisioni dei visti).

In ordine ad innovazione e concorrenza, il National *AI Research Resourc*e permetterà ai ricercatori e agli studenti di accedere a risorse e ai dati, e ci si propone di ampliare le sovvenzioni in aree come l'assistenza sanitaria e il cambiamento climatico. Si fa riferimento, inoltre, ad un ecosistema di IA equo, aperto e competitivo, con un invito alla *Federal Trade Commission* a esercitare la sua autorità, fornendo ai piccoli sviluppatori e imprenditori l'accesso all'assistenza tecnica e alle risorse, aiutando le piccole imprese a commercializzare le scoperte dell'IA.

Viene istituito, presso l'Ufficio Esecutivo del Presidente, il Consiglio per l'Intelligenza Artificiale della Casa Bianca (White House AI Council), per coordinare le attività delle agenzie di tutto il governo federale. Al fine poi di promuovere la *leadership* americana all'estero, si mira ad espandere gli impegni bilaterali, multilaterali e multilaterali, coordinati dal Dipartimento di Stato

in collaborazione con il Dipartimento del Commercio, per accelerare lo sviluppo e l'implementazione di *standard* con *partner* internazionali e organizzazioni di standardizzazione.

L'approccio degli Stati Uniti sin qui descritto è stato oggetto di rilievi sui mezzi di stampa[5]: per un verso si è riconosciuto come si sia fatto male a suo tempo a non regolare i *social network*; per altro verso si è però criticato anche l'attuale approccio, in quanto sarebbero evidenziati più rischi di quanto non faccia la stessa UE, e l'attenzione sarebbe focalizzata su rischi ipotetici, e non su rischi concreti. Si è osservato che si potrebbe sbagliare bersaglio e favorire i *big players*, che hanno i mezzi per adattarsi, rispetto alle piccole e medie imprese che lavorano con l'IA; che le agenzie federali non hanno la capacità di controllare centinaia di migliaia di sistemi; che non servono nuove regole, perché una legislazione sui diritti civili già esiste. L'ordine, inoltre, complicherebbe il ricorso all'IA, il che sarebbe particolarmente grave in riferimento ai servizi sanitari e nello sviluppo di medicinali, per le prescrizioni mediche e la selezione dei trattamenti.

Eppure, l'*order* appare un buon punto di partenza: chiede anzitutto di valutare l'impatto dell'IA in tutti i settori, con un'attenta analisi dei rischi, per cui dovranno essere stilati dei rapporti e poi redatte linee guide e *standard*. Esso inoltre cita espressamente alcuni pericoli, sui quali però già vi è casistica: basti pensare alle discriminazioni operate in materia penale con il *software* predittivo Compas[6], agli errori accertati nelle predizioni della polizia[7], nella erogazione di finanziamenti[8], nelle valutazioni scolastiche[9]. Ma, del resto anche gli usi dell'IA per scopi militari, per il sabotaggio di infrastrutture critiche, per il blocco dei sistemi informatici, per la creazione di *virus* e armi biologiche, per la manipolazione dell'opinione pubblica costituiscono ragioni di seria preoccupazione (si pensi al dibattito sull'origine del Covid o alla disinformazione a scopo politico), che appare troppo pericoloso sottovalutare.

Anche nell'Unione Europea, fin dal 2018, si è scelto di adottare un approccio ampio, asseritamente incentrato sull'eccellenza e sulla fiducia, con l'obiettivo di rafforzare la ricerca e la capacità industriale e garantire la sicurezza e i diritti fondamentali[10]. Si è puntato in particolare ad incrementare la capacità industriale e tecnologica dell'UE e l'adozione dell'IA in tutti i settori economici[11], a preparare ai cambiamenti socioeconomici[12], ad assicurare un quadro etico e giuridico adeguato[13], ad unire le forze[14]. Del resto, l'Unione Europea, con l'auspicio di potere essere un modello autorevole di regolazione a livello mondiale come è stata in materia ambientale o per il GDPR, ha bisogno di favorire anche lo sviluppo di campioni nel settore dell'IA.

Questa politica si è tradotta, nel 2021, in un piano coordinato, che ha precisato ulteriormente gli obiettivi, mirando a stabilire condizioni favorevoli allo sviluppo e all'adozione dell'IA[15], per rendere l'UE il luogo nel quale l'eccellenza prospera "dal laboratorio al mercato"[16], garantire che l'IA sia al servizio delle persone e sia un fattore positivo per la società[17], stabilire una *leadership* strategica in settori ad alto impatto[18].

Il recente accordo delle istituzioni dell'UE sul cd. *IA Act*, Regolamento sull'Intelligenza Artificiale, ne rende finalmente imminente la approvazione, e il nuovo atto costituirà il primo modello di regolazione al quale guardare: con l'idea di immettere sul mercato europeo prodotti di IA sicuri e rispettosi dei diritti fondamentali e dei valori dell'UE, la regolazione seguirà un approccio 'basato

sul rischio', e chiederà regole più rigide via via con l'aumento del rischio, al contempo vietando alcuni impieghi dell'IA (ad es. a fini di manipolazione comportamentale o *credit scoring*).

Certamente il regolamento non ha il grado di pervasività che sarebbe possibile in un ordinamento statale: non si applicherà ad ambiti che esulano dall'ambito di applicazione del diritto comunitario, lascerà agli Stati membri le competenze in materia di sicurezza nazionale, non si applicherà ai sistemi utilizzati esclusivamente per scopi militari o di difesa. Inoltre si capisce come sia difficile contenere la tentazione ad un uso molto esteso dell'intelligenza artificiale per la prevenzione e della lotta al crimine, anche laddove ci possa essere il rischio per diritti fondamentali: si pensi a come, per le azioni di contrasto delle forze dell'ordine, verrà introdotta una procedura di emergenza che consentirà alle forze dell'ordine di utilizzare uno strumento di IA ad alto rischio che non abbia superato la procedura di valutazione della conformità, anche se controbilanciato da garanzie per i diritti fondamentali. O come, nell'uso di sistemi di identificazione biometrica a distanza in tempo reale negli spazi accessibili al pubblico, si debbano stabilire i casi in cui tale uso sia strettamente necessario per scopi di contrasto e le autorità possano utilizzare tali sistemi (determinati reati, prevenzione di minacce reali, presenti o prevedibili, come gli attacchi terroristici, la ricerca di persone sospettate dei crimini più gravi).

In ogni caso, l'approccio basato sul rischio, la tutela dei diritti fondamentali, l'indicazione di misure di sorveglianza e controllo sull'uso dell'IA dell'UE non potranno che essere presi in considerazione con grande attenzione sullo scenario mondiale. L'*executive order* è già una prima reazione ai processi di regolazione in corso, e come tale va osservato con interesse.

[1] whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

[2] Le infrastrutture critiche sono definite nel Patriot Act degli Stati Uniti del 2001 come "sistemi e beni, fisici o virtuali, così vitali per gli Stati Uniti che l'incapacità o la distruzione di tali sistemi e beni avrebbe un impatto debilitante sulla sicurezza, sulla sicurezza economica nazionale, la salute o la sicurezza pubblica nazionale, o qualsiasi combinazione di queste materie". I settori coinvolti, tra gli altri, sono la difesa, i trasporti, l'energia, i servizi finanziari, la sanità, il nucleare, l'acqua.

[3] agendadigitale.eu/cultura-digitale/le-prime-regole-degli-usa-sullia-limpatto-dellexecutive-order-di-biden/

[4] Si deve garantire l'equità nelle sentenze, nella libertà condizionale e nella libertà vigilata, nel rilascio e nella detenzione preventiva, nella valutazione del rischio, nella sorveglianza, nella previsione della criminalità e nella polizia predittiva e nell'analisi forense.

[5] Biden's AI Order Is Government's Bid for Dominance e *AI Is Now Cooking, but It Shouldn't Be Overdone*, The Wall Street Journal, 7.11.2023.

[6] giurisprudenzapenale.com/2019/04/24/lamicus-curiae-un-algoritmo-chiacchierato-caso-loomis-alla-corte-suprema-del-wisconsin/

[7] sistemapenale.it/pdf_contenuti/1695830536_pietrocarlo-predictive-policing-rivtrim-form.pdf

[8] agendadigitale.eu/cultura-digitale/ai-e-finanza-tra-discriminazioni-e-imprecisioni-rischi-e-contromisure/

[9] open.online/2020/08/16/caos-maturita-regno-unito-algoritmo-calcolo-voti-penalizza-studenti-meno-abbienti/

[10] COM(2018) 237 final, COMUNICAZIONE DELLA COMMISSIONE, L'intelligenza artificiale per l'Europa.

[11] Con un aumento degli investimenti, un rafforzamento della ricerca e dell'innovazione, un sostegno ai centri di eccellenza, un trasferimento alle piccole imprese e agli utilizzatori potenziali, sostegno a prove e sperimentazioni e attrazione di investimenti privati, la messa a disposizione di più dati.

[12] *Evitare forme di esclusione, promozione del talento, della diversità e della interdisciplinarità*

[13] *Elaborazione di orientamenti etici, sicurezza e responsabilità.*

[14] *Coinvolgere gli stati membri, le parti interessate, seguire lo sviluppo e l'adozione dell'IA, coinvolgimento internazionale*.

[15] *Acquisire, mettere in comune e condividere informazioni strategiche, sfruttare il potenziale dei dati, promuovere capacità di calcolo critiche*.

[16] *Collaborare con i portatori di interessi attraverso, ad esempio, il partenariato europeo sull'IA, i dati e la robotica e i gruppi di esperti, costruire e mobilitare capacità di ricerca, Mettere a disposizione strumenti attraverso una piattaforma di IA on demand e fornire un ambiente nel quale gli sviluppatori possano effettuare prove e sperimentare e le PMI e le pubbliche amministrazioni possano adottare l'IA come i poli europei dell'innovazione digitale, finanziare e rendere scalabili idee e soluzioni innovative per l'IA*.

[17] *Coltivare i talenti e migliorare l'offerta delle competenze necessarie per consentire un fiorente ecosistema di IA, sviluppare un quadro strategico per assicurare la fiducia nei sistemi di IA, promuovere nel mondo la visione dell'UE per un'IA sostenibile e affidabile*.

[18] *Utilizzare l'IA in ambito climatico e ambientale, utilizzare la prossima generazione di IA per migliorare la salute, preservare la leadership dell'Europa: strategia per la robotica nel mondo dell'IA, rendere il settore pubblico un apripista nell'utilizzo dell'IA, applicare l'IA alle attività di contrasto, alla migrazione e all'asilo, rendere la mobilità più intelligente, più sicura e più sostenibile attraverso l'IA, sostenere l'IA per un'agricoltura sostenibile*.
