



Il contenuto dell'AI Act riprende i principî etici più importanti in tema di intelligenza artificiale?

Adriano Fabris

Professore di Etica e deontologia dell'intelligenza artificiale, Università di Pisa. Mail: adriano.fabris@unipi.it

La risposta alla domanda che viene formulata nel titolo dev'essere opportunamente articolata. Lo scopo dell'AI Act, come viene detto all'inizio del documento, è quello di regolamentare l'uso di determinati prodotti tecnologici nell'ambito dell'Unione Europea. Più precisamente, è quello di "migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale [...] nell'Unione, in conformità dei valori dell'Unione" (1). Ciò viene fatto, com'è noto, attraverso una specifica classificazione dei rischi da affrontare e mediante esplicativi divieti, giustificati dalla necessità di evitare tali rischi. L'approccio del documento, dunque, è un approccio *risk-based*. Con riferimento a questa impostazione, cioè a partire dalla "piramide del rischio" illustrata nel testo, si procede a una categorizzazione dei sistemi basati sull'intelligenza artificiale, sono discussi i livelli di rischio che possono venire ammessi e sono analizzate le possibili eccezioni.

A seguito di tale approccio il riferimento alla dimensione dell'etica non può che essere messo in secondo piano. Non è un caso, quindi, il fatto che il termine "etica" ricorre nel testo solo 2 volte e l'uso dell'aggettivo "etici" 8. In compenso, a conferma delle intenzioni del legislatore, il verbo "vietare" e i termini a esso collegati si trova 36 volte, mentre la parola "divieto" compare 23 volte.

Ciò segnala anzitutto la differenza d'interesse e d'impostazione fra il modo di affrontare determinate questioni – in questo caso lo sviluppo dell'intelligenza artificiale e le sue conseguenze – da parte di una disciplina come l'etica, per un verso, e quello che anima una volontà di regolamentazione, per altro verso. L'etica, infatti, non si limita a definire i termini di un problema da uno specifico punto di vista – in questo caso il punto di vista del rischio –, per poi cercare di evitare eventuali effetti negativi attraverso l'individuazione di norme atte allo scopo. L'etica intende, da un lato, stabilire principî condivisi che consentono all'essere umano di orientarsi nelle proprie azioni, e insieme giustificare tali principî in questa loro funzione. Si tratta di principî generali, che riguardano l'essere umano come tale, e che valgono indipendentemente dalla sua appartenenza a uno Stato o a una Comunità, oppure dall'attività che svolge. Dall'altro lato, proprio a partire da questa giustificazione, l'approccio etico ha anche il compito di motivare gli esseri umani a comportarsi in conformità con i principî indicati. Non si tratta solo di operare in un certo modo per evitare d'incorrere in sanzioni, ma è in gioco un progetto di vita e un'idea ben precisa di ciò che s'intende realizzare.

Se teniamo presenti queste considerazioni, ci rendiamo conto che non possiamo chiedere all'AI Act di svolgere una funzione che non è la sua. Tale documento, però, con l'etica comunque ha a che fare, dal momento che, in un mondo sempre più pervaso da sistemi di intelligenza artificiale, intende regolamentare i nostri comportamenti a partire da determinati valori. Vi si raccorda nella misura in cui riprende, implicitamente ed esplicitamente, alcuni dei principî condivisi che proprio riguardo all'uso di tale tecnologia sono stati stabiliti e giustificati in ambito sia europeo che internazionale. Nel corso del testo, in altre parole, si può riscontrare sia un



esplicito rinvio a tali principî, sia un implicito riferirsi a essi.

Il rinvio esplicito all'ambito etico all'interno dell'*AI Act* riguarda i sette principî, ovvero i valori ineludibili, enucleati nel documento europeo sull'intelligenza artificiale affidabile (<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>) che è stato elaborato nel 2019 da un High-Level Expert Group indipendente su mandato dell'Unione Europea. Essi – lo ricordiamo – comprendono: l'intervento e la sorveglianza umani nei confronti dei programmi e dei dispositivi basati sull'intelligenza artificiale, il fatto che essi debbono essere tecnicamente robusti e sicuri, la necessità di rispettare la vita privata degli utenti e di garantire sempre una governance dei dati a disposizione, la trasparenza dei processi tecnologici, l'assenza di discriminazioni e l'equità nell'accesso ai benefici che tali tecnologie comportano, la promozione del benessere sociale ed ambientale, la responsabilità nell'uso dell'intelligenza artificiale. Si tratta di criteri "non vincolanti", come viene detto nell'*AI Act*, che forniscono un orientamento per agire bene negli ambienti digitali. Essi, insieme ai requisiti giuridicamente vincolanti precisamente elencati nello stesso *AI Act*, "contribuiscono all'elaborazione di un'IA coerente, affidabile e antropocentrica, in linea con la Carta e con i valori su cui si fonda l'Unione" (27).

Tali principî etici sono dunque considerati lo sfondo all'interno del quale viene inquadrata la normativa generale del documento che stiamo esaminando. E indubbiamente essi costituiscono un punto di riferimento valoriale che lo ispira nei suoi vari passaggi. Va sottolineata però la diversità d'impianto fra i due testi che ho appena menzionato. Gli *Orientamenti etici per un'AI affidabile* affrontano il loro tema a partire dalla prospettiva della fiducia. L'interesse, sostanzialmente costruttivo, è volto a stabilire le

condizioni affinché l'interazione con i sistemi di intelligenza artificiale possa avvenire sulla base dell'affidabilità che a essi noi siamo disposti ad attribuire. L'approccio a partire dal rischio, che invece è quello prescelto nell'*AI Act*, presuppone che da certi programmi e dispositivi possano venire conseguenze negative, e mira dunque a salvaguardare da tali conseguenze gli esseri umani. Inoltre nell'*AI Act*, insieme ai valori che sono indicati negli *Orientamenti* e ai quali viene fatto nel testo esplicito riferimento, sono rintracciabili anche altri principî etici, o più in generale altri riferimenti valoriali. Sono i "valori dell'Unione" menzionati già all'inizio del documento. In esso tali valori risultano richiamati implicitamente, consentendo di comprendere perché certe norme debbono essere introdotte proprio nella forma in cui vengono presentate.

Si tratta soprattutto di tre valori. Sono anzitutto il principio della gestione democratica delle decisioni e poi quello della salvaguardia dei diritti individuali e delle libertà di base, non solo di ogni cittadino dell'Unione Europea, ma più in generale di tutti gli esseri umani. Vi è infine il riferimento alla necessità, che emerge da qui, di assumere determinate regole, condivise dalla comunità di riferimento, quale strumento fondamentale per il raggiungimento e il mantenimento di tali valori.

Il principio del rispetto della privacy, più volte richiamato nell'*AI Act*, e il divieto del controllo da parte di uno, o di alcuni, su tutti gli altri, ad esempio usando le tecnologie di riconoscimento biometrico, contrastano infatti con l'esercizio dei diritti democratici. Non c'è democrazia se tale controllo non è svolto in maniera equa, condivisa e trasparente: cioè, appunto, salvaguardando i diritti e la libertà delle persone. Certo, tali diritti e tale libertà debbono essere non solo riconosciuti e ribaditi, ma anche e soprattutto regolamentati, per far sì che la loro attuazione non vada a





discapito di una coesione comunitaria. Emerge così la necessità di elaborare determinate normative, come quella dell'AI Act. A volte tale necessità si esprime in maniera fin troppo dettagliata e puntuale. In ogni caso, però, una regolamentazione resta compito imprescindibile per lo sviluppo e il mantenimento dei principî democratici, e per la salvaguardia della stessa democrazia nella sua concreta attuazione. Ciò va ribadito con forza in un momento storico come quello presente, nel quale in molte parti del mondo la tendenza alla *deregulation* sembra prevalere.

C'è comunque un ultimo aspetto che bisogna affrontare, se si vuole rispondere adeguatamente alla domanda espressa nel titolo di questo contributo. Bisogna vedere se davvero "il contenuto dell'AI Act riprende i principî etici più importanti in tema d'intelligenza artificiale", sebbene in quella forma duplice, esplicita e implicita, che ho indicato. Il principio etico più importante sul quale si basa la legislazione europea è indubbiamente quello della centralità dell'essere umano nella gestione dei sistemi di intelligenza artificiale. Si parla infatti di "*human centered AI*". È una centralità che dev'essere considerata soprattutto per quanto riguarda gli effetti dell'azione di tali sistemi, allo scopo di evitare che questi effetti risultino problematici o negativi. Ma per fare in modo che ciò non avvenga, l'essere umano è chiamato ad assumersi anche la responsabilità di scrivere programmi e di progettare dispositivi in relazione ai quali possa sempre mantenere una qualche forma di controllo.

È questo il punto decisivo a partire dal quale è possibile stabilire con l'intelligenza artificiale un'interazione eticamente fondata. Si giustifica anzitutto sulla base della consapevolezza dei limiti umani, i quali devono essere sempre tenuti presenti se si vuole salvaguardare la centralità di cui si parla nei documenti europei. Ma si

legittima altresì a partire dall'assunzione del fatto che lo stesso agire dei sistemi di intelligenza artificiale, spesso più efficiente e performante di quello umano, è comunque caratterizzato anch'esso da limiti ben precisi, che ne inficiano l'affidabilità.

In conclusione, possiamo dunque dire che l'AI Act presuppone e riprende, soprattutto implicitamente, alcuni valori fondamentali dell'Unione Europea, anche in tema di intelligenza artificiale. Sono valori da tenere ben presenti, soprattutto in quanto hanno una valenza universale, cioè si estendono, nel riconoscimento della dignità di tutti gli esseri umani e del rispetto che a essi è dovuto, anche al di fuori dei confini della stessa Unione. Sono dunque valori che contrastano con altri criteri di comportamento in base ai quali vengono spesso governate le relazioni interumane: quelli che conducono all'autocrazia, quelli che favoriscono il mantenimento delle diseguaglianze e che impediscono che ciascuno possa fiorire secondo le sue capacità, quelli che, rinunciando a regole condivise, legittimano l'uso della legge del più forte.

