

INFORMATICA, FEDE E FIDUCIA: LA SCIENZA E I SENTIMENTI

La «spiegabilità» dell'intelligenza artificiale nodo da sciogliere Nel 'Diario di un curato dicampagna', George Bernanos scrive che «la fede non si perde, cessa di informare la vita». Da un punto di vista opposto, Jean-Paul Sartre ci ricorda che «la fiducia si guadagna goccia a goccia, ma si perde litri». Fede e fiducia sono dunque due sentimenti diversi, da non confondere tra loro. Quando si parla di scienza, il tema della fede o della fiducia si manifesta con forza: si può avere fede nella capacità della scienza di capire il mondo e migliorarlo o si deve assumere un atteggiamento fiduciario, prudente e guardingo? Non possiamo negare che le scienze mediche, tra le altre, hanno aumentato l'aspettativa di vita, ma dobbiamo anche ricordare che la fisica nucleare ha reso possibile la bomba atomica. A questo dibattito si è recentemente aggiunta l'informatica. Nata come scienza che studia il calcolo, l'informatica ha una grande peculiarità: fornisce alle macchine ciò che distingue l'essere umano dalle specie animali, la capacità di elaborare informazioni. La sua manifestazione estrema si trova nelle macchine che esibiscono un'intelligenza artificiale, in grado di imitare i comportamenti umani. L'umanità ha sempre avuto fede nella scienza, perché la sua applicazione concreta, la tecnologia, ha affrancato gli esseri umani dai lavori pesanti e ripetitivi, proteggendoli dall'ambiente esterno e dalle malattie. Il patto implicito alla base di questa fede era il dominio dell'uomo sulle macchine. L'informatica ha violato questo patto, permettendo la costruzione di macchine intelligenti che possono sostituire gli esseri umani.

La fiducia nella scienza agisce su un piano completamente diverso da quello della fede. Si basa su elementi concreti che permettono di capirne le ragioni, di prevederne gli effetti, di affrontare gli inevitabili pericoli. La fiducia nell'informatica si poggia su settant'anni di sviluppo inarrestabile, che ci ha convinto che le macchine che calcolano non ci danneggiano, ma operano a nostro favore.

Il 28 febbraio 2020, al termine di giorni di lavori che hanno coinvolto esperti di tutto il mondo, è stato reso pubblico l'appello della Santa Sede per un'«intelligenza artificiale etica», un documento articolato che parte dal presupposto che i sistemi basati sull'intelligenza artificiale devono essere concepiti, progettati e realizzati per «servire e proteggere gli esseri umani e l'ambiente in cui vivono». Per garantire il rispetto di questo vincolo, sono indicati tre requisiti: i sistemi intelligenti devono essere inclusivi, devono essere consapevoli della complessità del nostro ecosistema per prendersene cura e proteggerlo, devono rendere evidente a ogni essere umano che sta interagendo con un'entità non umana. Emerge con chiarezza, nei confronti dell'intelligenza artificiale, un approccio fiduciario che chiede con forza di mostrare in maniera trasparente le



Avvenire

caratteristiche e le finalità dei sistemi che la utilizzano. Già in passato era emerso il tema della fiducia nei confronti di questi sistemi e, in particolare, di quelli in grado di elaborare grandi molidi dati (i cosiddetti big data). Numerosi sistemi si erano mostrati discriminatori nei confronti di alcune fasce della popolazione, ad esempio rendendo difficile se non impossibile l'accesso ai mutui per l'acquisto della casa. Un'analisi attenta aveva rivelato che, pur essendo stati progettati e realizzati in maniera ineccepibile, questi sistemi erano stati 'addestrati' utilizzando un insieme di casi viziati da elementi di discrezionalità. Su questa linea di ragionamento l'Unione Europea ha finanziato alcuni progetti nell'ambito di Horizon 2020, per affrontare il tema della «spiegabilità» (la cosiddetta XAI, eXplainable Artificial Intelligence) ovvero la garanzia che i risultati forniti da un sistema basato sull'intelligenza artificiale siano spiegati in maniera comprensibile. In questi tempi di pandemia, il tema della fede o della fiducia anima e alimenta le discussioni sui vaccini, ad dimostrazione che la natura umana sia dominata dai sentimenti. Direttore del Dipartimento di Informatica dell'Università di Pisa RIPRODUZIONE RISERVATA.