

Luciano Floridi

# Etica dell'intelligenza artificiale

Sviluppi, opportunità, sfide

Edizione italiana a cura di  
Massimo Durante



*Raffaello Cortina Editore*

---

## INDICE

Prefazione	11
------------	----

### PRIMA PARTE

#### *Comprendere l'intelligenza artificiale*

1. Passato: l'origine dell'intelligenza artificiale	21
1.1 Introduzione: la rivoluzione digitale e l'intelligenza artificiale	22
1.2 Il potere di scissione del digitale: tagliare e incollare la modernità	26
1.3 Nuove forme dell'agire	32
1.4 IA: un ambito di ricerca in cerca di una definizione	34
1.5 Conclusione: etica, governance e design	35
2. Presente: IA come nuova forma dell'agire e non dell'intelligenza	39
2.1 Introduzione: che cos'è l'IA? "La riconosco quando la vedo"	40
2.2 IA come controfattuale	44
2.3 Le due anime dell'IA: ingegneristica e cognitiva	48
2.4 IA: un divorzio riuscito nell'infosfera	53
2.5 L'uso umano degli esseri umani e delle interfacce	58
2.6 Conclusione: chi si adatterà a chi?	61
3. Futuro: lo sviluppo prevedibile dell'IA	65
3.1 Introduzione: scrutare nei semi del tempo	66
3.2 Dati storici, ibridi e sintetici e il bisogno di ludicizzazione	67
3.3 Problemi difficili, problemi complessi e il bisogno di avvolgimento	77
3.4 Il design come futuro dell'IA	83
3.5 Conclusione: l'IA e le sue stagioni	85

SECONDA PARTE

*Valutare l'intelligenza artificiale*

4. Un quadro unificato di principi etici per l'IA	91
4.1 Introduzione: troppi principi?	92
4.2 Un quadro unificato di cinque principi per l'IA etica	93
4.3 Beneficenza: promuovere il benessere, preservare la dignità e sostenere il pianeta	96
4.4 Non maleficenza: privacy, sicurezza e "cautela della capacità"	97
4.5 Autonomia: il potere di "decidere di decidere"	98
4.6 Giustizia: promuovere la prosperità, preservare la solidarietà, evitare l'iniquità	99
4.7 Esplicabilità: rendere possibili gli altri principi tramite l'intelligibilità e la responsabilità	100
4.8 Una visione sinottica	101
4.9 L'etica dell'IA: da dove e per chi?	102
4.10 Conclusione: dai principi alle pratiche	103
5. Dai principi alle pratiche: i rischi di comportamenti contrari all'etica	107
5.1 Introduzione: traduzioni rischiose	107
5.2 Lo shopping etico	109
5.3 Il "bluewashing" etico	111
5.4 Il lobbismo etico	113
5.5 Il dumping etico	115
5.6 L'elusione dell'etica	118
5.7 Conclusione: l'importanza di conoscere meglio	120
6. Etica soft e governance dell'IA	123
6.1 Introduzione: dall'innovazione digitale alla governance del digitale	123
6.2 Etica, regolazione e governance	127
6.3 La compliance nei confronti delle norme: necessaria ma insufficiente	129
6.4 Etica hard e soft	130
6.5 L'etica soft come quadro etico	134
6.6 Analisi dell'impatto etico	138
6.7 Preferibilità digitale e cascata normativa	140
6.8 Il duplice vantaggio dell'etica digitale	141
6.9 Conclusione: l'etica come strategia	143

7. La mappatura dell'etica degli algoritmi	145
7.1 Introduzione: una definizione operativa di algoritmo	145
7.2 La mappa dell'etica degli algoritmi	148
7.3 Prove inconcludenti che portano ad azioni ingiustificate	150
7.4 Prove imperscrutabili che portano all'opacità	153
7.5 Prove fuorvianti che portano a pregiudizi ( <i>bias</i> ) non voluti	159
7.6 Risultati ingiusti che portano alla discriminazione	163
7.7 Effetti trasformativi che sollevano sfide per l'autonomia e la privacy informativa	166
7.8 Tracciabilità come presupposto della responsabilità morale	171
7.9 Conclusione: l'uso buono e cattivo degli algoritmi	175
8. Cattive pratiche: l'uso improprio dell'IA per il male sociale	177
8.1 Introduzione: l'uso criminale dell'IA	177
8.2 Preoccupazioni	181
8.3 Minacce	189
8.4 Possibili soluzioni	205
8.5 Sviluppi futuri	215
8.6 Conclusione: dagli usi malvagi dell'IA all'IA socialmente buona	219
9. Buone pratiche: l'uso dell'IA per il bene sociale	221
9.1 Introduzione: l'idea di IA per il bene sociale	222
9.2 Una definizione di AI4SG	227
9.3 Sette fattori essenziali per il successo dell'AI4SG	231
9.4 Conclusione: fattori di bilanciamento per l'IA per il bene sociale	255
10. Macchine ultraintelligenti, singolarità e altre distrazioni fantascientifiche	259
10.1 Introduzione: l'aggiornata paura ancestrale dei mostri	259
10.2 Credenti e miscredenti nella vera IA: un dibattito sulla fede	263
10.3 Gli adepti della singolarità: la fine è vicina, la vera IA sta arrivando	264
10.4 IA-teismo dell'IA: quello che i computer non possono fare, presumibilmente	270
10.5 Adepti della singolarità e atei dell'IA: una diatriba inutile	273
10.6 Conclusione: il problema non è HAL ma l'umanità nel suo complesso	276

11. La società per la buona IA	279
11.1 Introduzione: quattro modi per realizzare una società per la buona IA	279
11.2 Chi possiamo diventare: rendere possibile l'umana realizzazione di sé, senza svalutare le capacità umane	282
11.3 Cosa possiamo fare: migliorare l'agire umano, senza rimuovere la responsabilità umana	283
11.4 Cosa possiamo conseguire: incrementare le capacità della società, senza ridurre il controllo umano	284
11.5 Come possiamo interagire: coltivare la coesione sociale, senza erodere l'autodeterminazione umana	285
11.6 Raccomandazioni per una società della buona IA	287
11.7 Conclusione: la necessità di politiche concrete e costruttive	295
12. Il gambetto: l'impatto dell'IA sul cambiamento climatico	297
12.1 Introduzione: il potere duplice dell'IA	297
12.2 L'IA e le "transizioni gemelle" dell'Unione Europea	303
12.3 AI e cambiamento climatico: sfide etiche	304
12.4 IA e cambiamento climatico: l'impronta ecologica	307
12.5 Tredici raccomandazioni a favore dell'IA contro il cambiamento climatico	312
12.6 Conclusione: una società più sostenibile e una biosfera più sana	316
13. L'IA e gli obiettivi di sviluppo sostenibile delle Nazioni Unite	319
13.1 L'IA per il bene sociale e gli obiettivi di sviluppo sostenibile delle Nazioni Unite	320
13.2 Valutare le evidenze dell'IAxOSS	321
13.3 L'IA per promuovere l'"azione per il clima"	326
13.4 Conclusione: un programma di ricerca per l'IAxOSS	329
14. Conclusione: il verde e il blu	331
14.1 Introduzione: dal divorzio tra agire e intelligenza al matrimonio tra il verde e il blu	331
14.2 Il ruolo della filosofia come design concettuale	334
14.3 Il bellissimo errore di natura	336
Ringraziamenti	341
Bibliografia	347