

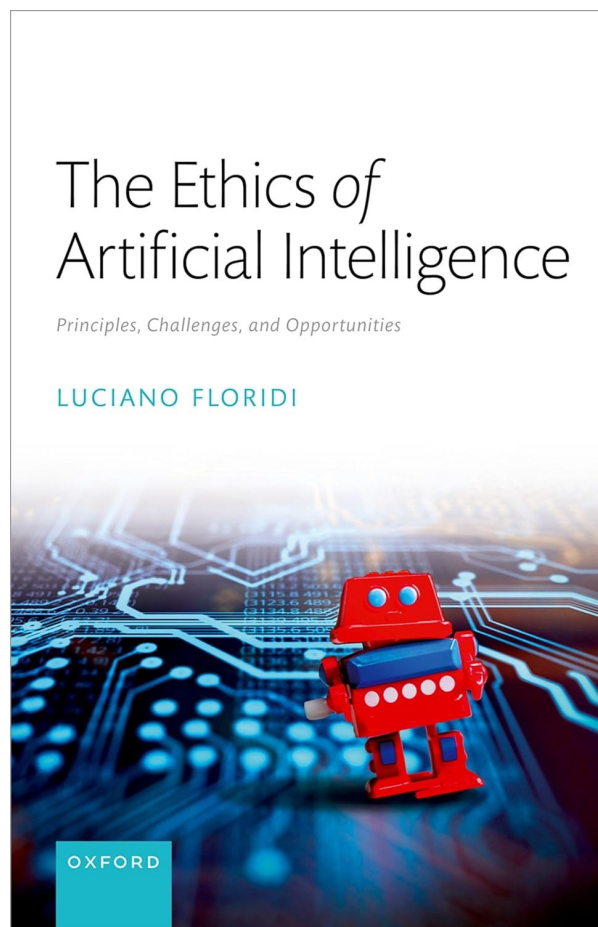
Nataša Daničić
Knjižnica Hrvatske akademije znanosti i umjetnosti
ndanicic@hazu.hr

ISSN 1845-2434
[CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)

**The Ethics of Artificial Intelligence:
Principles, Challenges, and Opportunities /
Luciano Floridi.
Oxford : Oxford University Press, 2023.**

ISBN 9780198883098

Prikaz



Naslovnica preuzeta s mrežne stranice [nakladnika](#)

Knjiga *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities* talijanskog i britanskog filozofa Luciana Floridija predstavlja nastavak njegovog dugogodišnjeg promišljanja o informacijskom društvu i etičkim pitanjima digitalnih tehnologija. Podijeljena je u dva veća dijela, u 13 poglavlja: prvi se bavi razumijevanjem prošlosti, sadašnjosti i budućnosti umjetne inteligencije, dok drugi razmatra njezine etičke dimenzije i društvene implikacije. Posebnost Floridijeva pristupa jest dosljedno inzistiranje da umjetna inteligencija nije oblik „inteligencije“ nego nova forma „agencije“, tj. sposobnosti djelovanja i postizanja ciljeva bez potrebe za sviješću ili razumijevanjem. Time autor nastoji izbjeći metaforičko i senzacionalističko tumačenje UI-ja i smjestiti ga u realističan okvir tehnoloških mogućnosti.

U prvim poglavljima autor rekonstruira kako je digitalna revolucija omogućila razdvajanje inteligencije i agencije i pokazuje da se razvoj umjetne inteligencije ne može razumjeti samo kroz tehničke inovacije, nego kroz širi proces transformacije našeg iskustva i društva. Računala, internet i algoritmi nisu samo alati, nego čimbenici promjene ontologije i načina na koji doživljavamo i strukturiramo svijet. Sadašnjost UI-ja opisuje kao rezultat razdvajanja između ljudske inteligencije i računalne agencije. Strojevi mogu djelovati uspješno bez da razumiju ono što čine, otvarajući tako pitanje tko će se kome prilagoditi: hoće li ljudi oblikovati tehnologiju ili će društvo sve više prilagođavati sebe digitalnim sustavima. U poglavljima o budućnosti Floridi ne pribjegava futurologiji već se fokusira na uočljive tendencije: korištenje sintetičkih podataka (umjetno generirani podaci koji oponašaju karakteristike i obrasce podataka iz stvarnog svijeta), razvoj generativnih modela, sve veću kompleksnost problema i potrebu za „obuhvaćanjem“ (*enveloping*) kako bi se UI sustavi uklopili u društveno prihvatljive okvire. Ovaj dio knjige ima jasnu filozofsku snagu, ali ponegdje i apstraktni karakter, što može otežati čitanje onima koji očekuju konkretnije studije slučaja.

Drugi, obimniji dio knjige bavi se etikom umjetne inteligencije. Floridi polazi od činjenice da već postoji mnoštvo etičkih smjernica, u više od 160 različitih dokumenata, i predlaže njihovu sintezu u pet temeljnih načela: dobročinstvo, neškodljivost, autonomija, pravednost i objašnjivost čime nastoji stvoriti pregledan okvir, analogan bioetičkim principima. Kritički gledano, ova sistematizacija je koristan doprinos, ali ostaje donekle na razini deklaracije - kako se ta načela prevode u praksu, ostaje otvoreno pitanje. Autor upozorava na tri vrste opasnosti: *ethics shopping*, odnosno biranje etičkih načela koja odgovaraju interesima; *bluwashing*, odnosno deklarativno isticanje etičnosti bez stvarne provedbe i *dumping*, odnosno izbjegavanje strožih normi primjenom UI-ja u manje reguliranim okruženjima, jasno pokazujući raskorak između normativnih ideala i ekonomske i političke stvarnosti.

Poseban naglasak stavljen je na pojam „meke etike“ (*soft ethics*), skup načela nadređenih pravnom okviru i regulaciji. Floridi smatra da je zakonska regulativa nužna, ali nedovoljna,

jer ne može obuhvatiti svu brzinu i složenost digitalnih promjena. Meka etika trebala bi popuniti prostor između zakonske obveze i moralne odgovornosti, pomažući organizacijama da razvijaju kulturu etičkog promišljanja i djelovanja. Ovo je jedna od originalnijih ideja knjige, iako se može prigovoriti da ostaje više programatska nego operativna.

Niz poglavlja bavi se konkretnim pitanjima algoritama i njihove primjene. Autor detektira glavne etičke probleme: od pristranosti i diskriminacije, preko netransparentnosti i manipulacije, do pitanja odgovornosti. Iako ti problemi nisu novi u literaturi, Floridi ih izlaže sustavno i pregledno, što knjizi daje značaj priručnika. U kontrastu s tim, poglavlja o „lošim praksama“: uporabi UI-ja za počinjenje kaznenih djela, financijskim prijevarama, manipulaciji tržišta ili kibernetičkim prijetnjama, više upozoravaju na prijetnje nego što nude razrađena rješenja. Autor naglašava potrebu interdisciplinarnog odgovora, ali čitatelju ostaje dojam da je riječ o tek skiciranom planu djelovanja.

Positivna protuteža tome su poglavlja o „dobrim praksama“ i UI-ju za društveno dobro (*AI4SG - AI for Social Good*). Floridi opisuje kako bi se umjetna inteligencija mogla koristiti u zdravstvu, obrazovanju, borbi protiv siromaštva ili klimatskih promjena i nudi sedam kriterija za uspješnu provedbu takvih projekata, poput transparentnosti, zaštite privatnosti i situacijske pravednosti. Dvadeset preporuka za izgradnju „dobre UI zajednice“ čine ambicioznu ideju, iako se može primijetiti da zahtijevaju visok stupanj političke volje i međunarodne suradnje, što nije lako ostvarivo. Poglavlja o UI-ju i klimatskim promjenama te o ulozi u ostvarivanju UN-ovih ciljeva održivog razvoja daju knjizi globalnu relevantnost, ali i pokazuju stanoviti nesklad između filozofskih ideala i institucionalne realnosti.

U kontekstu razmatranih argumenata, Floridi oblikuje viziju povezivanja održivosti i digitalnih tehnologija u svrhu stvaranja pravednijeg društva i zdravijeg okoliša. Filozofija, prema njemu, ima ulogu konceptualnog dizajna, oblikovanja pojmova i okvira koji omogućuju odgovorno upravljanje tehnološkim razvojem. Zaključno, knjiga staje na raskrižje između teorije i prakse: nudi sustavnu i promišljenu filozofsku interpretaciju umjetne inteligencije, ali ostavlja otvoreno pitanje kako tu interpretaciju prevesti u konkretne politike.

Iz knjižničarske perspektive knjiga je vrijedna kao pregled ključnih etičkih rasprava te kao primjer kako filozofija može oblikovati razumijevanje novih tehnologija. Ipak, ona nije priručnik s gotovim smjernicama za svakodnevnu praksu, nego intelektualni izazov i poziv na daljnje promišljanje. Floridi uspijeva povezati tehničke, društvene i filozofske aspekte u koherentnu cjelinu, ali čitatelju prepušta zadatak da tu cjelinu pretoči u vlastite profesionalne i institucionalne kontekste. Upravo u tom naporu gradnje mosta u dijalogu između teorijskih razmatranja i prektične primjene očituje se najveći doprinos ove knjige.