

Dati Aperti e Algoritmi Etici

Tiziana Catarci
Direttrice ISTC-CNR

DIPARTIMENTO DI INGEGNERIA INFORMATICA
AUTOMATICA E GESTIONALE ANTONIO RUBERTI



SAPIENZA
UNIVERSITÀ DI ROMA



Dati Aperti

- La base della trasparenza
- Dati di qualità, accessibili, riutilizzabili, non proprietari.
- Benefici: controllo civico, innovazione, partecipazione, riuso.
- Esempi:
 - OpenCoesione (Italia): monitoraggio fondi europei.
 - Portali open data comunali: mobilità, edilizia, trasporti.
 - Indagini AlmaLaurea

Algoritmi nella Governance Pubblica

- Automazione nei processi decisionali pubblici.
- Alcuni esempi problematici:
 - COMPAS (USA): bias etnico nella giustizia predittiva.
 - Predictive Policing (UK/USA): criminalizzazione aree marginali.
 - SyRI (Olanda): violazione della privacy nell'individuazione di frodi.



Rischi Etici

- Bias nei dati → discriminazioni sistemiche.
- Opacità algoritmica → decisioni incomprensibili.
- Assenza di accountability → nessuno responsabile.
- Sovra-delega tecnica → illusione di oggettività.
- Esclusione digitale → disparità di accesso.

Rischi - bias

Pregiudizi nascosti nei dati



Gruppi sotto-rappresentati

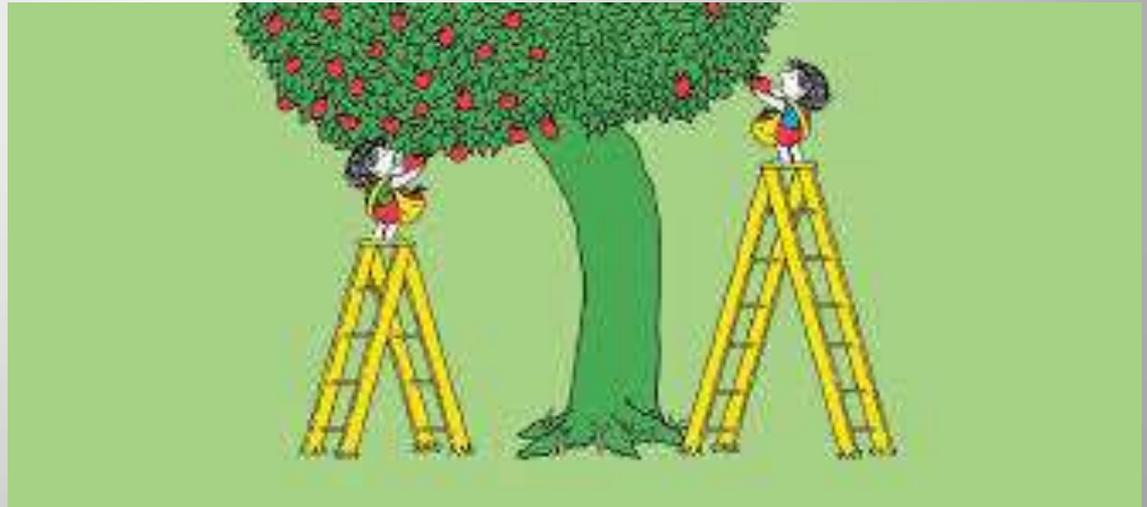


Governance Algoritmica Trasparente

- Algoritmi open source e aperti a audit esterno.
- Dati pubblici accessibili e aggiornati.
- Valutazioni di impatto algoritmico.
- Partecipazione dei cittadini ai processi digitali.
- Alfabetizzazione digitale
- Regolamenti europei: AI Act, Data Governance Act.
- La trasparenza non è solo pubblicare dati, ma rendere comprensibili e controllabili i meccanismi che influenzano la vita dei cittadini.

Uguaglianza vs. Equità

Uguaglianza significa che a ogni individuo o gruppo di persone vengono fornite le stesse risorse o opportunità. L'equità riconosce che ogni persona si trova in circostanze diverse e alloca esattamente le risorse e le opportunità necessarie per raggiungere un risultato paritario.



Responsabilità e Accountability

- Chi è responsabile se qualcosa va storto? Se un algoritmo prende una decisione sbagliata o causa danni, chi ne risponde?
- Il sistema è equo e imparziale? Il sistema è stato addestrato su dati che possono contenere bias? Potrebbe discriminare alcune persone?
- Le decisioni del sistema sono spiegabili? È importante capire *perché* un sistema ha preso una certa decisione, soprattutto in ambiti delicati (come la sanità, la giustizia, le assunzioni).
- Rispetta la privacy? I dati usati per addestrare il sistema sono stati raccolti e usati in modo corretto, rispettando le normative?
- È sicuro e affidabile? Il sistema si comporta come previsto in diverse situazioni, o può avere comportamenti imprevisti e pericolosi?



Etica

- Quell'insieme di conoscenze consolidate e di atteggiamenti profondi e radicati che portano a impostare le forme di vita con cui un gruppo sociale influenza gli individui nella loro esistenza al fine di garantire una ordinata coordinazione sociale.
- L'insieme di conoscenze e atteggiamenti profondi può essere specificato in *principî* che prescrivono comportamenti e in *valori* che indicano ciò che è buono, apprezzabile e che vale.
- A partire dalla prima metà del secondo millennio dell'era cristiana in etica sono emersi come costanti i principî o pilastri o assiomi, che informano e reggono l'intero edificio morale. La specificazione di tali principî e il loro numero variano a seconda delle diverse elaborazioni proposte, e di solito includono criteri formali di coerenza interna, di robustezza e correttezza, ma anche quelli sostanziali di beneficenza (fare il bene) e di giustizia (distribuire il bene equamente).
- Principi etici dell'IA (gruppo di esperti UE): *Rispetto dell'autonomia umana; Prevenzione del danno; Correttezza; Spiegabilità.*
- L'Etica dell'IA è un'etica integrata che segue il modello dell'Agenda 2030 dell'ONU in cui il principio *Sostenibilità* riguarda gli aspetti sociali ed ambientali, integra *benessere e giustizia.*
- **Compito specifico dell'etica dell'intelligenza artificiale è lo studio delle implicazioni culturali, sociali ed economiche derivanti dall'utilizzo dei sistemi di IA.**

Principio OMS	Descrizione	Denominazione UNESCO	Denominazione OECD	Riferimento AI Act	Riferimento GDPR
Protezione dell'autonomia	Garantire che le persone mantengano il controllo sui propri dati personali e sulle decisioni che li riguardano, con trasparenza e consapevolezza.	Rispetto, protezione e promozione dei diritti umani	Rispetto dei diritti umani, libertà e dignità	Controllo degli utenti e protezione dei dati personali (art. 8, 9, 54)	Protezione dei dati personali (artt. 6 e 7)
Benessere umano e sicurezza	Progettare e utilizzare l'intelligenza artificiale in modo da promuovere la salute e il benessere, prevenendo rischi e conseguenze indesiderate	Benessere umano e giustizia sociale	Benefici per il benessere degli individui e delle società	Sicurezza e affidabilità dei sistemi (art. 10, 15, 23)	Misure tecniche e organizzative per la sicurezza dei dati (art. 32)
Inclusività ed equità	Contrastare ogni forma di discriminazione, garantendo l'accesso equo ai benefici dell'IA e rispettando la diversità culturale, sociale e linguistica.	Inclusività e diversità culturale	Equità e giustizia	Prevenzione di bias e discriminazioni (art. 10, 13, 71)	Non discriminazione nell'uso dei dati personali (art. 5)
Trasparenza e spiegabilità	Rendere i processi decisionali del sistema comprensibili e giustificabili, assicurando che utenti e stakeholder possano comprendere le decisioni	Trasparenza e spiegabilità	Trasparenza e capacità di spiegare il funzionamento dell'IA	Documentazione dettagliata e spiegabilità delle decisioni (art. 13, 52, 53, 86)	Diritto di accesso ai dati e alle informazioni (artt. 13-15)
Responsabilità e sostenibilità	Stabilire meccanismi di supervisione continua, garantendo il rispetto delle normative, l'equità nei processi decisionali e un impatto ambientale sostenibile	Responsabilità e accountability, Sostenibilità ambientale	Responsabilità degli sviluppatori e dei decisori	Responsabilità per danni, supervisione continua e misurazione dell'impatto energetico (art. 4, 29, 62, 73, considerando 70-72)	Principio di responsabilità dei titolari e responsabili del trattamento (art. 24)

Principi Etici

Consapevolezza è Formazione

- Per fidarsi della tecnologia si deve avere familiarità con la tecnologia. Le regole del gioco dovrebbero essere note ai giocatori e i giocatori devono potersi fidare del fatto che gli altri giocatori seguano le regole (o essere in grado di capire che non lo stanno facendo).
- Consapevolezza Digitale
- Alfabetizzazione digitale
- Competenze digitali
- Sviluppo delle conoscenze fondamentali, per formare gli strumenti cognitivi necessari per discernere l'informazione attendibile e analizzarla in modo critico e responsabile
 - Apprendimento delle discipline classiche (logica, filosofia, matematica, storia, latino,...)
 - Pensiero critico (etica e morale, come in passato giuramento di Ippocrate (medicina) e filosofia morale (economia))
- Formazione multidisciplinare/interdisciplinare