

---

# Data society e data governance. I big data

---

**Diritto dell'informatica, servizi informatici e sicurezza dei dati**  
**Università di Pisa**

---

**Fernanda Faini**





# DATA SOCIETY



# Centralità dei dati

- i dati costituiscono l’**“io” digitale**, sono fondamento di ogni **attività umana**, pervadono ogni aspetto dell’**esistenza**
- **il benessere e lo sviluppo umano** dipendono in modo significativo dai servizi basati sui dati
- le tecnologie applicate ai dati potenziano le capacità umane, si integrano negli oggetti (Internet of Things), danno vita a nuovi soggetti (intelligenza artificiale) → **crolla il confine tra realtà analogica e digitale**



## ***data society***

una “società di dati”, che non è soltanto fondata sui dati come nel suo avvento, ma ne è intimamente pervasa finendo per coinvolgere e plasmare l’uomo stesso come un ***data subject***

# Data governance

La **governance della società digitale** passa dalla data governance → il diritto è **deputato a regolare la vita** e quindi **è chiamato a disciplinare il “diluvio” di dati**, al fine di tutelare i diritti, sanare i conflitti e fornire certezza giuridica

*Ubi societas, ibi ius*

**Necessario e complesso equilibrio tra diritto e tecnologia**

Il diritto deve dialogare con altri insiemi di regole, la **lex informatica o digitalis** → le regole applicate dal “codice” informatico (non giuridico), che abilitano azioni e determinano effetti nella realtà digitale → condizionano ogni altra forma di regolazione, compresa quella giuridica



Ma le regole informatiche sono frutto dell'uomo che può intervenire sulle stesse attraverso il diritto → necessità di non rendere tutto ciò che è tecnologicamente possibile, solo per questo, giuridicamente legittimo → la tecnologia non deve dominare sul diritto; il diritto non deve “domare” e limitare la tecnologia

# Nuove geometrie di potere

- **Stati** → assenza di confini; dimensioni globali dell'oggetto di regolazione; indebolimento della sovranità e pretesa di potere; erosione monopolio
- **colossi tecnologici** → favoriti da dimensione sovranazionale, libertà di commercio, regole inadeguate → si presentano come spazi di libertà, ma assumono forme panottiche e esercitano un ruolo concreto nella “regolazione” → **controllori del pedaggio di accesso alla vita digitale**



**rischio Big Brother** (Big Government + Big Business) (Lessig)

es. Caso Datagate – Cambridge Analytica

- **individui** → deboli nei confronti degli uni e degli altri; apparente libertà; rischio di mercificazione dei dati, accesso mediato da forme di pagamento inedite (i propri dati, i propri interessi, i propri comportamenti)

# L'era della conoscenza o del controllo?

**“Democratizzazione” dell’informazione** → tutti possono esprimere la propria opinione al pari di chiunque altro, ma più la conoscenza è diffusa più ci sono torsioni per limitarne l'utilizzabilità



**volto oscuro della rete**

- **automazioni e “dittature dell’algoritmo”**
- **trattazioni e contrattazioni inique, discriminazioni e disuguaglianze**
- **filter bubble, echo chambers e polarizzazioni** → apparente libertà, ma in realtà vediamo chi ci somiglia con più frequenza grazie agli algoritmi
- **bufale e fake news** → difficoltà a distinguere cosa è vero



dalla società della conoscenza verso

**la società del controllo**, superpanopticon o post-panopticon?



# DEFINIZIONE E CARATTERISTICHE



# Big data: cosa sono? (1)

## Molteplici definizioni

Enormi volumi di dati detenuti da grandi organizzazioni (governi, colossi tecnologici) provenienti da diverse fonti e analizzati per mezzo di algoritmi

- tecniche di data mining (estrazione di conoscenza)
- big data analytics
- machine learning (apprendimento automatico)
- deep learning (reti neurali artificiali)
- tecnologie specifiche

Eterogenee **tracce digitali** derivanti da interazioni in rete

- dati strutturati e non strutturati
- dati generati dagli utenti
- dati personali



## Big data: cosa sono? (2)

- dati raccolti su base volontaria nelle piattaforme online (Facebook, Instagram, Amazon, LinkedIn...)
- dati “scambiati” a fronte di utilità conseguibili (raccolte punti, tessere fedeltà, app che consentono sconti)
- dati forniti in modo più o meno consapevole (GPS del telefono, sensori)
- dati “automatici” come i cookies
- dati residui o inferiti → sottoprodotto delle azioni (es. errori di digitazione)
- dati raccolti dai soggetti pubblici
- dati prodotti dagli oggetti → **Internet of Things**
- dati su cui si basa **intelligenza artificiale**

## Big data: cosa sono? (3)

# Internet of Things

oggetti dotati di sensori per ottenere informazioni, interagire tra loro e in rete

## Ampiezza degli oggetti interessati:

- dispositivi indossabili
- domotica → automazione domestica
- robotica
- industria automobilistica
- ambito biomedicale
- monitoraggio in ambito industriale
- rilevazione di eventi avversi
- etc.



I dati prodotti da Internet of Things  
sono big data

# Big data: cosa sono? (4)

## Intelligenza artificiale

Sistemi che mostrano un **comportamento intelligente** analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi

(Comunicazione della Commissione europea «L'intelligenza artificiale per l'Europa» COM(2018) 238 final del 25 aprile 2018)

- software agiscono nel **mondo virtuale**  
(es. assistenti vocali, software per analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale)
- software incorporano l'IA in **dispositivi hardware**  
(es. robot avanzati, auto a guida autonoma, droni o applicazioni dell'Internet delle cose)

## Anima di IA

- enormi quantità di **dati** (big data)
- **algoritmi** → capaci di “animare” i dati e di estrarne il valore



# Caratteristiche dei big data

## Paradigma delle 3, 4 o 5 V

- **volume** → capacità di acquisire, memorizzare e accedere a enormi quantità di dati.
- **velocità** → capacità di effettuare analisi in tempo reale o ad “alta velocità”  
Sottende **dinamicità** e **tempo** → in tempi brevi il dato diventa obsoleto.
- **varietà** → eterogenea tipologia di dati e formati, provenienti da fonti diverse (strutturate e non)

Secondo parte di dottrina anche

- **valore** → quanto i big data valgono come insieme
- **veracità** → veridicità, ossia qualità e accuratezza dell'analisi



**protagonisti gli algoritmi**

big data diventano dinamica materia prima capace di generare valore

# Finalità e utilizzi

**Premessa** → i dati sono raccolti e analizzati senza che siano sempre chiari ex ante utilizzi e finalità → risultati imprevisti frutto di correlazioni

- **conoscenza sul presente e sul futuro** → interpretazione di esigenze, profilazione, ottimizzazione della produzione, monitoraggio e supporto alle decisioni, supporto a campagne giornalistiche  
**Capacità di predizione** → elevate correlazioni indicano alte probabilità → predizioni sugli andamenti di mercato, miglioramento di diagnosi e cure, prevenzione di disastri, previsioni politiche → oracoli digitali
- **servizi e miglioramento qualità della vita** → impiego nell'erogazione di servizi e nello svolgimento di funzioni (controllo etc.), maggior efficacia ed efficienza, tempestività, risparmio



Big data possono rispondere alla **tutela di interessi generali**,  
ma anche alla realizzazione di **vantaggi economici**, obiettivi privati

# ETICA E DIRITTO





# Algoritmi e diritto

## Muta il modo di conoscere e “misurare” fatti e persone

- si possono **analizzare tendenzialmente tutti i dati** disponibili (non solo campioni)
- **inferenze e correlazioni, logica deterministica, probabilità e non causalità** → si accerta “il cosa”, senza domandarsi *ex ante* “il perché” dei fenomeni, si prende atto delle correlazioni che emergono
- **più confusione e meno esattezza** a fronte della conoscenza e comprensione garantita da tali strumenti



## Negli algoritmi criticità ontologiche per il diritto

- orientato ai valori (non ai numeri)
- prescrittivo (non descrittivo e predittivo)
- formale e “lento” (non così dinamico)
- poggia sulla scelta individuale e sulla libera volontà (non deterministico)

# Dati e conoscenza, analisi e interpretazione

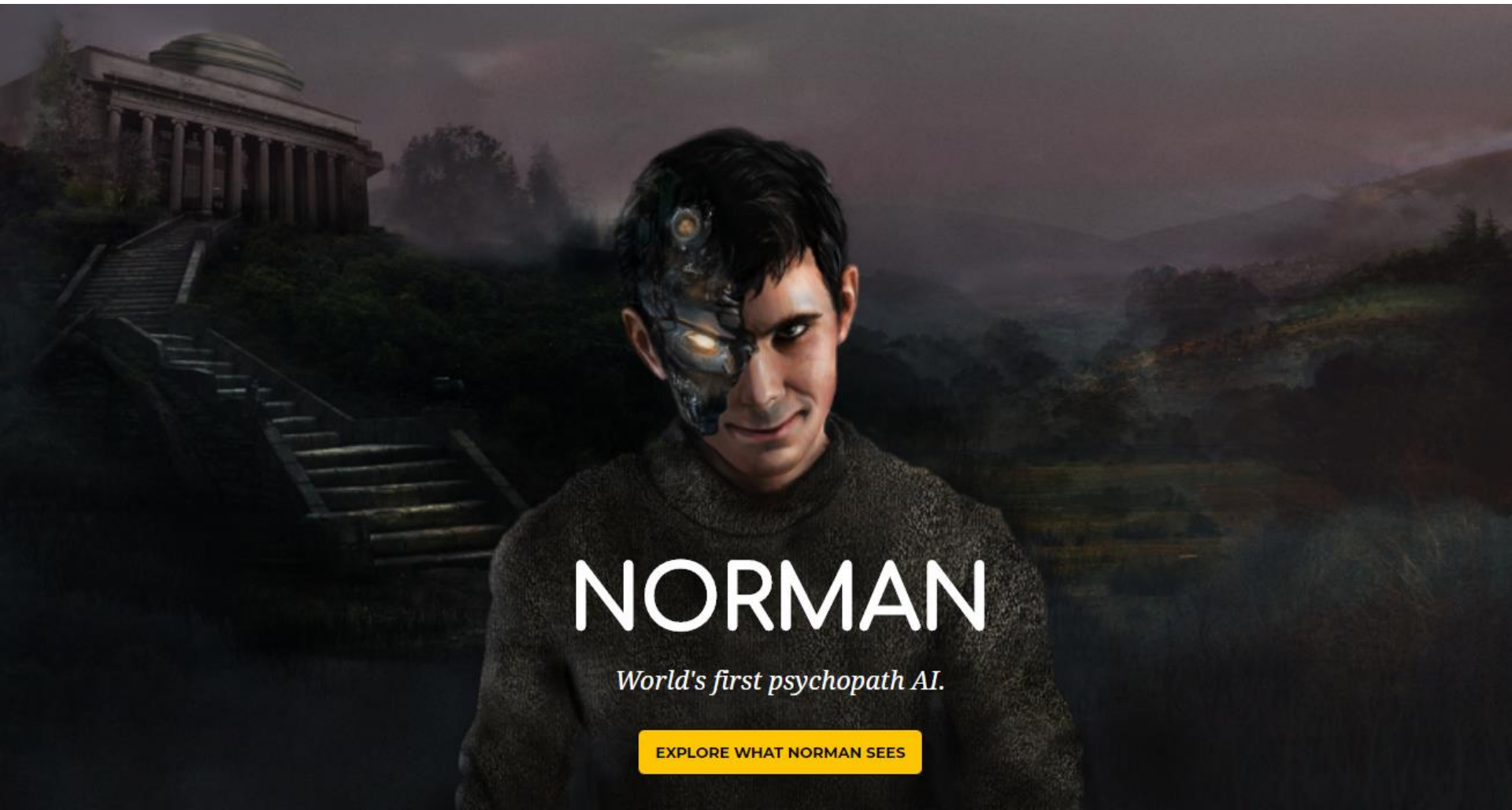
- necessità di **dati di qualità** → le **basi di dati** devono essere **costituite e annotate correttamente**, garantendo qualità, neutralità e intelligibilità, assicurando un elevato grado di sicurezza e riuscendo a prevenire i rischi
- **natura inferenziale e probabilistica delle elaborazioni** → analisi tecnica è un processo di approssimazione con il rischio di trarre conclusioni imprecise senza tradursi in conoscenza → **rischio di bias, errori tecnici e manipolazioni umane, discriminazioni**, disparità di trattamento, opacità per confusione



adeguata **contestualizzazione, analisi e interpretazione** →  
saper fare le domande giuste ai dati e saper comprendere le risposte

# Importanza dei dati

<http://norman-ai.mit.edu>





# Bias e dintorni (1)

## Caso **Eric Loomis**

sentenza State of Wisconsin v. Eric Loomis, 881 N.W.2d 749 (2016)

Software di algoritmi predittivi **Compas**, utilizzato per valutare il rischio di recidiva e pericolosità sociale → la sentenza ha legittimato l'uso del software ritenendo che rispettasse il principio del giusto processo, dal momento che costituiva solo un ausilio a disposizione del giudice per supportare la sua valutazione e non il motivo esclusivo della decisione.



sulla base della natura proprietaria del software, a Loomis non è stato concesso l'accesso al codice sorgente, utile per poter comprendere e contestare le decisioni.



Gli studi sul software Compas hanno mostrato **marginale di errore e “imparità”** nella predizione a seconda del gruppo etnico di riferimento → Compas sovrastimava alcuni fattori quali l'appartenenza ad un certo gruppo etnico.

## Bias e dintorni (2)

**Progetto sperimentale di Amazon** impiegava IA per fornire punteggi ai **candidati**. Nel 2015 emerge che il sistema non selezionava candidature per ruoli di sviluppatori di software e posti ad alto contenuto tecnologico in modo neutrale rispetto al genere: escludeva i cv delle donne



Il software aveva “imparato” in modo automatico che i candidati uomini (numericamente maggiori) erano preferibili  
e ha **penalizzato i cv con la parola donna**



Amazon è stata costretta a sciogliere il team del progetto

# Profili etici

- **individuazione e la scelta dei valori etici e delle norme giuridiche** con cui educare gli algoritmi e istruire le macchine
- **individuazione di principi comuni tra ordinamenti diversi** nel pluralismo etico che caratterizza la società contemporanea
- complessa valutazione relativa all'introduzione o all'adattamento di **categorie giuridiche e norme**



rispetto dei valori etici della società e dei correlati principi di diritto  
si traduce nella necessaria **tutela da offrire ai diversi diritti in gioco**,  
che esige il rispetto delle norme di riferimento



# Dilemmi etici

Macchine autonome, armi e veicoli a guida autonoma possono incidere su **posizioni individuali di particolare rilievo etico-giuridico**, come il diritto alla vita e all'integrità fisica



accettabile in tali casi la sostituzione della macchina all'uomo, rischiando di aprire vuoti di responsabilità in caso di danni?

In caso affermativo, difficile regolamentare il comportamento della macchina e la responsabilità di fronte a **tensioni etiche e problematiche complesse** soprattutto laddove siano in gioco interessi dello stesso peso assiologico:

- in caso di **collisione inevitabile** quale scelta dovrà compiere il veicolo a guida autonoma, laddove si tratti di recare danno a diversi potenziali soggetti o persino di sacrificare una vita piuttosto che un'altra?
- su **quali criteri etici** dovranno basarsi tali scelte e **quali istruzioni** dovrà ricevere la macchina per operare le scelte?

<https://www.moralmachine.net>

# Big data divide

- **big data divide** → asimmetria, pochi “signori” dei dati (colossi digitali e poteri pubblici) raccolgono ed elaborano big data
- **pericolo di usare le previsioni in contesti diversi** da quelli di origine con le relative discriminazioni (es. lavoro, assicurazioni, prevenzione del crimine) → dati utilizzati per finalità ulteriori rispetto ai servizi per cui sono impiegati, soprattutto laddove ciò avvenga in modo poco trasparente
- **controllo sociale, sorveglianza e monitoraggio indiretto** grazie a dati pubblici e privati → rischio di scivolare dalla società della conoscenza verso un superpanopticon o post-panopticon



**algoritmi guidano i comportamenti individuali** (es. social) →  
**apparente libertà** del singolo/utente guidato dall'automazione

# Datagate

Nel 2013 Edward Joseph **Snowden** ha rivelato le attività di **sorveglianza di massa** dei servizi di intelligence statunitensi connesse in particolare al programma governativo segreto, denominato **PRISM**.

**Sorveglianza condotta da cinque occhi** (USA, Canada, Gran Bretagna, Australia, Nuova Zelanda)

Accesso e raccolta ad opera delle agenzie governative da aziende come **Google, Microsoft, Facebook, Apple, Skype e Yahoo** di dati quali:

- email
- profili
- chat e conversazioni
- video
- foto



# Facebook – Cambridge Analytica

Nel 2018 Christopher **Wylie**, che è stato dipendente della società di consulenza Cambridge Analytica, ha fatto emergere lo spregiudicato utilizzo dei dati personali forniti dagli utenti durante l'utilizzo dell'app "This is your digital life" da parte di Cambridge Analytica (specializzata nella profilazione e micro-targeting comportamentale), per mezzo di accordi con il titolare dell'app, che trattava i dati a fini di ricerca



utilizzo dei dati per campagne di comunicazione capaci di influenzare, condizionare e manipolare gli individui in occasioni quali il referendum sulla Brexit



**manipolazione delle scelte individuali**

(in tal caso opinioni politiche,  
che hanno tutela più incisiva)



# Ownership

Chi è titolare?

Chi ha prodotto, chi detiene o chi elabora i dati?

Beni pubblici?

1. **proprietà tradizionale** → prospettiva civilistica della proprietà tradizionale, immaginando un diritto di proprietà sui dati da intendere come beni
2. **prospettiva contrattuale** → sfruttare autonomia contrattuale, al fine di tutelare i database per regolarne la cessione o la concessione di diritti di utilizzo temporaneo e al fine di regolare ulteriori aspetti afferenti a intelligenza artificiale, individuando i profili di responsabilità reciproca
3. **proprietà intellettuale** → nel caso di insieme di dati e di software interviene la normativa sul diritto d'autore e sui diritti connessi
  - **“contenitore”** → se si ritiene che il valore sia nell'insieme dei dati → diritto d'autore per le banche dati creative e diritto sui generis per le banche dati non creative (cui in genere sono ricondotti big data).
  - **software e servizi** → algoritmi hanno ruolo cruciale in tali soluzioni – se si ritiene che il valore sia nelle analisi e negli algoritmi: protezione del diritto d'autore



# Chi ne è responsabile? (1)

- **titolarità diverse** → non si tratta di un singolo bene immateriale, ma di un processo che vede proprietà dei dati, ruolo degli algoritmi, titolarità di software di analisi, contratti eventuali di fornitura di servizi → il titolare dello strumento può essere diverso dal titolare del servizio
- a seconda del caso, può cambiare la **partecipazione umana** all'azione e alla decisione che conduce a eventuali danni. Chi è responsabile in caso di errori o bias? La responsabilità è ascrivibile sempre e solo ad un soggetto umano o talvolta gli algoritmi in autonomia conducono a decisioni o soluzioni che possono provocare danni? Come avrebbe dovuto essere diversamente programmato l'algoritmo per evitare il danno?



- **“sistemi agenti”** → esigenza di imputabilità autonoma in considerazione del difficile collegamento tra le attività delle macchine e la responsabilità umana. Il Parlamento europeo affronta la questione con **sistemi di assicurazione obbligatoria** e forme di **responsabilità oggettiva** → si interroga sul riconoscimento dello status giuridico di **personalità elettronica**, seppur non sia scevro di problematiche, quali **l'attribuzione di diritti e doveri** e l'individuazione della capacità di agire

## Chi ne è responsabile? (2)

Ricostruzione della responsabilità in caso di incidenti → la dottrina indica diverse possibili strade:

- responsabilità del **produttore** (art. 114, d.lgs. 206/2005)
- responsabilità per l'**esercizio di attività pericolose** (art. 2050 c.c.)
- responsabilità dei proprietari per i **danni cagionati da animali** (art. 2052 c.c.)
- responsabilità **in vigilando e in educando** di genitori, tutori e maestri per i danni cagionati da fatti illeciti dei minori e degli allievi (art. 2048 c.c.)

Seppur meno utilizzate, non manca chi richiama:

- responsabilità per il danno cagionato da **cose in custodia** (art. 2051 c.c.)
- responsabilità per la **circolazione dei veicoli** (art. 2054 c.c.)

# Problematiche giuridiche

- **necessario rispetto dei principi dell'ordinamento**, tra cui dignità, pieno sviluppo della persona, eguaglianza e non discriminazione → necessità di valutazione su impatto giuridico, sociale ed etico e rischi per la collettività
- **concorrenza** → rischio che i colossi del web siano dominanti sul mercato con violazione delle norme antitrust (la concentrazione Microsoft/Skype 2011, concentrazione Facebook/WhatsApp 2014)
- **segreto industriale** → in specifico per proteggere gli algoritmi e le tecniche di analisi impiegate per estrarre valore dai dati
- **tutela del consumatore e dell'utente** → rischio di inefficacia alla luce di pratiche commerciali opache, black box, rischio di discriminazioni, disparità di trattamento e di prezzo, distorsioni
- **legge applicabile e giurisdizione competente** → criticità geopolitiche; raccolta e utilizzo dei dati è sovranazionale e si pongono difficoltà a stabilire la legislazione e la giurisdizione nel momento patologico del conflitto

# Problematiche giuridiche in ambito pubblico

**Rispetto norme e principi che guidano azione pubblica** → tra queste:

- **procedimento** amministrativo → principi di **imparzialità**, **trasparenza** e **partecipazione** (legge 241/1990; d.lgs. 33/2013) → es. esigenza di motivazione (art. 3, legge 241/1990), garantire forme di accesso e assicurare correlate garanzie processuali
- **amministrazione digitale** (d.lgs. 82/2005) → es. diritti digitali
- **qualità dei dati pubblici** (d.lgs. 82/2005; d.lgs. 33/2013) → quantità, incertezza, natura inferenziale e probabilistica rischiano di mettere in crisi la disciplina



esigenza di garantire **trasparenza e apertura** non solo dei dati, ma anche della **logica degli algoritmi** e del **processo di funzionamento** del servizio



esigenza di garantire **certezza del diritto**



# Il caso relativo alla scuola (1)

Caso afferente all'algoritmo usato nella procedura di mobilità dei docenti nelle sedi disponibili nell'organico della scuola.

- **TAR Lazio-Roma, sezione III bis, 21 marzo 2017, n. 3742 e 22 marzo 2017, n. 3769** → diritto di accesso documentale all'algoritmo deputato a gestire il procedimento, configurato quale atto amministrativo informatico.
- **TAR Lazio, sez. III bis, 10 settembre 2018, nn. 9224-9930; TAR Lazio, sez. III bis, 27 maggio 2019, n. 6606; TAR Lazio, sez. III bis, 13 settembre 2019, n. 10964**
- ✓ **netta chiusura** all'impiego di algoritmi per decisioni amministrative, anche laddove scaturenti da un'attività vincolata e in casi complessi e ampi
- ✓ **algoritmo incapace di assicurare le garanzie procedurali e i principi** di trasparenza, partecipazione e obbligo di motivazione con le **correlate garanzie processuali** (diritto di azione e difesa in giudizio)

## Il caso relativo alla scuola (2)

- **Consiglio di Stato, Sezione VI, 8 aprile 2019, n. 2270**
- ✓ **algoritmo** in ambito pubblico è giuridicamente **ammissibile e legittimo per attività vincolate prive di discrezionalità, ma utilizzo di procedure “robotizzate” non può essere motivo di elusione dei principi dell’ordinamento**. La regola tecnica che governa l’algoritmo resta una regola amministrativa generale, costruita dall’uomo e non dalla macchina, per essere poi solo applicata da questa.
- ✓ necessarie **conoscibilità** secondo una declinazione rafforzata del principio di trasparenza e **giustiziabilità** quale soggezione alla piena cognizione e al pieno sindacato del giudice amministrativo
- ✓ nella fattispecie violazione dei principi di imparzialità, pubblicità e trasparenza → impossibilità di comprendere le modalità con cui, attraverso l’algoritmo, sono stati assegnati i posti disponibili, costituisce di per sé un vizio tale da inficiare la procedura → gli esiti sono risultati connotati da **illogicità e irrazionalità**, essendosi verificate **situazioni paradossali**, in contrasto con le disposizioni di riferimento

# Il caso relativo alla scuola (3)

- **Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472, n. 8473 e n. 8474**

**Nuovo meta-principio di legalità algoritmica** → principi scaturenti anche dal diritto sovranazionale (es. regolamento UE 2016/679)

- ✓ **conoscibilità e comprensibilità** → necessaria motivazione e giustificazione della decisione in senso sostanziale e non rigida e meccanica applicazione delle regole procedurali → diritto a conoscere i diversi aspetti dell'algoritmo e decifrarne la logica → autori, procedimento per l'elaborazione, meccanismo di decisione, comprensivo delle priorità nella procedura valutativa e decisionale e dei dati rilevanti
- ✓ **non discriminazione algoritmica** → algoritmo non deve assumere carattere discriminatorio, garantendo la qualità dei dati, minimizzando il rischio di errori e rettificando i dati o i fattori che possano determinarli
- ✓ **non esclusività della decisione algoritmica** → deve essere assicurato intervento umano, anche solo in termini di controllo, verifica e validazione

**La censura, più che singole violazioni di legge, finisce per involgere il metodo in quanto tale per difetto di trasparenza dello stesso**

# BIG DATA E DATA PROTECTION





# Framework giuridico (1)

**Regolamento UE 2016/679 relativo alla protezione dei dati personali,  
nonché alla libera circolazione di tali dati  
General Data Protection Regulation – GDPR**

- è teso a rendere omogenea ed efficace la tutela della persona
- disciplina non riguarda **i dati anonimi** → dati che in origine o a seguito di trattamento non possono essere associati ad un interessato identificato o identificabile (considerando 26, reg. UE 2016/679) → **anonimizzazione è diversa da pseudonimizzazione**

# Framework giuridico (2)

## Regolamento (UE) 2018/1807

### sulla libera circolazione dei dati non personali nell'Unione europea

Nel rispetto dei diritti fondamentali, prevede una serie di strumenti, idonei a garantire la libera circolazione dei dati non personali e a migliorare la certezza giuridica e il livello di fiducia, avvalendosi anche della flessibilità



- superamento degli obblighi di localizzazione dei dati
- messa a disposizione e accesso di autorità competenti
- portabilità dei dati non personali, che si avvale dell'autoregolamentazione affidata a codici di condotta
- designazione di punti unici di contatto per ogni Stato e cooperazione tra Stati membri



teso a sfruttare le opportunità dell'economia basata sui dati,  
nella quale si collocano i big data

# Framework giuridico (3)

Sull'interazione tra i due regolamenti rileva la  
**Comunicazione della Commissione europea**  
**«Guidance on the Regulation on a framework for the free flow of non-  
personal data in the European Union»** del 29 maggio 2019



atto esclusivamente a titolo informativo,  
che per espressa dichiarazione non contiene alcuna interpretazione autorevole

Tali linee guida sono state adottate in conformità alle previsioni dell'art. 8, paragrafo 3, del regolamento europeo 2018/1807, che impone alla Commissione di pubblicare orientamenti sull'interazione tra i due regolamenti europei «in particolare per quanto concerne gli insiemi di dati composti sia da dati personali che da dati non personali».

Le linee guida sono consapevoli del fatto che gli insiemi di dati misti siano ricorrenti nella maggior parte delle situazioni della vita reale.

# Framework giuridico (4)

Nel caso dei **big data** i dati sono di tipologia eterogenea e provengono da fonti diverse → di norma, i big data avranno la connotazione di **insiemi di dati misti**

- dati personali e non personali non indissolubilmente legati → il regolamento europeo 2018/1807 si applica alla parte dell'insieme contenente i dati non personali e, parallelamente, il regolamento (UE) 2016/679 si applica alla parte dell'insieme contenente i dati personali.
- dati personali e non personali **indissolubilmente legati** → il regolamento europeo 2018/1807 lascia impregiudicata l'applicazione del regolamento (UE) 2016/679 → secondo le linee guida, tale disposizione si traduce nell'**applicazione piena del regolamento (UE) 2016/679 all'insieme di dati misti**, anche quando i dati personali rappresentano soltanto una piccola parte dell'insieme dei dati.

# Principi

Big data e soluzioni di IA rischiano ontologicamente di scontrarsi con i **principi applicabili al trattamento previsti dal regolamento (UE) 2016/679** (art. 5):

- **limitazione della finalità** → i dati devono essere raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità (art. 5, paragrafo 1, lett. b).
- **minimizzazione dei dati** → devono essere adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati (art. 5, paragrafo 1, lett. c).
- **esattezza e accuratezza dei dati** → dati devono essere di qualità, ossia esatti e aggiornati (art. 5, paragrafo 1, lett. d)



# Informazioni e consenso

Problemi anche per il rispetto di pilastri della disciplina:

- **informazioni da fornire all'interessato** → artt. 12-13-14
- **consenso** → art. 7 → preventivo, specifico, inequivocabile, revocabile



Su cosa saranno fornite le informazioni e su cosa sarà espresso il consenso?  
È autenticamente libero il consenso se fornito a fronte di vantaggi conseguibili in termini relazionali, professionali, sociali e a fronte dell'accesso alla propria vita digitale?

**Apparente libertà**, rileva il considerando 42, reg. UE «*il consenso non dovrebbe essere considerato liberamente espresso se l'interessato non è in grado di operare una scelta autenticamente libera o è nell'impossibilità di rifiutare o revocare il consenso senza subire pregiudizio*»

# Dati personali e dati anonimi

- esistono **dati totalmente anonimi**? Inferenze su gruppi, metadati, dati residui, dati ausiliari...
- la **definizione stessa di dato personale** può essere insufficiente → dati afferenti a gruppi o comunità sulla cui base sono prese decisioni che incidono anche sul singolo che non ha prestato il consenso e non ha ceduto dati, oltre ai metadati e ai dati residui
- **necessità di condotte difensive** → l'individuo è costretto a pagare in termini monetari e di tempo per mettere in atto strategie difensive dei propri dati

# Marketing diretto e profilazione

- **marketing diretto e connessa profilazione** → art. 21, comma 2 → può essere considerato legittimo interesse su cui basare il trattamento dei dati personali



consente il trattamento

La protezione avviene con diritto di opposizione in qualunque momento →

**meccanismo di opt-out**



è congruo con l'intenzione di elevare la tutela?

# Processo decisionale automatizzato

L'interessato ha il **diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione**, che produca effetti giuridici o che incida in modo analogo significativamente sulla sua persona.



La previsione non si applica nel caso in cui la decisione:

- a) **sia necessaria per la conclusione o l'esecuzione di un contratto** tra l'interessato e un titolare del trattamento;
- b) **sia autorizzata dal diritto dell'Unione o dello Stato membro** cui è soggetto il titolare del trattamento [...];
- c) **si basi sul consenso esplicito** dell'interessato.

Nel caso a) e c), il titolare deve attuare misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno **il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione**.

(art. 22, regolamento europeo 2016/679)

# SCENARI FUTURI E STRUMENTI DI GOVERNANCE





# Data Ethics e centralità della persona

Necessità di una nuova **etica digitale** che guidi la regolazione

**Fondamentale l'elemento umano,**

la capacità dell'uomo di orientare la tecnologia → la tecnologia da sola non basta; necessità del ruolo del diritto che deve servirsi di un approccio etico.

**Bilanciamento “mobile” tra diritti e interessi** può basarsi sulla **centralità della persona**, sui valori della dignità e dello sviluppo della stessa, condivisi fondamenti costituzionali



**persona** come **prisma i cui riflessi sono costituiti dai diversi diritti in gioco** che poggiano e convergono sul condiviso fondamento costituzionale, la clausola aperta dei diritti della personalità di cui all'art. 2 Cost.



# Diritto nella tecnica e accountability

**Approccio sistematico, preventivo, proattivo e tecnico**  
logica del regolamento UE 2016/679

- **incorporazione del diritto nella tecnica**  
richiama la funzione preventiva del diritto che si serve della tecnologia per garantire il suo rispetto → **legal protection by design** e **by default**
- **accountability, sicurezza ed approccio etico** → artt. 24 e 32
- **responsabilità e sistema sanzionatorio** → artt. 82-84 → sanzioni severe utili a garantire effettività → in linea con artt. 82-84 regolamento UE 2016/679

# Diritto alla comprensione e alla contestabilità

**Trasparenza** → ai sensi degli artt. 13, 14 e 15 del regolamento europeo 2016/679 tra le informazioni da fornire all'interessato e su cui ha accesso:

- l'esistenza di un processo decisionale automatizzato, compresa la profilazione, e, almeno in tali casi,
- **informazioni significative sulla logica utilizzata,**
- **l'importanza e le conseguenze previste** di tale trattamento per l'interessato.



## **trasparenza algoritmica**

riequilibra asimmetria, supera opacità di black box e incide sulla libertà:

**diritto alla comprensione, spiegabilità e diritto alla contestabilità,**

il singolo deve essere garantito nel diritto all'autodeterminazione informativa



## **diritto alla sindacabilità**

da parte dell'interessato e del giudice

# Apertura e diritto al riutilizzo

**Utilizzo sinergico di open data e big data**, proteggendo le libertà dei singoli e la sovranità dei dati relativi



**sanatoria al momento della diffusione,**

liberando i dati, in linea con la natura non escludibile e non rivale del bene  
(limitato dal potere economico o pubblico)



garantisce diritti e libertà degli individui,  
un accesso equo e maggiore concorrenza,  
competitività ed equità del mercato.

# Forme di tutela collettiva?

**Forme di tutela collettiva** → i dati sono beni comuni, a titolarità diffusa, con impatto collettivo



**autodeterminazione collettiva**

Accanto a possibilità per l'interessato di dare mandato a organismo o associazione per reclamo o ricorso giurisdizionale per suo conto (art. 80, comma 1), è significativo **art. 80**, comma 2 → gli Stati membri possono prevedere che un **organismo, organizzazione o associazione**, indipendentemente dal mandato conferito dall'interessato, abbia

- ✓ il **diritto di proporre un reclamo** all'autorità di controllo e
- ✓ il **diritto a proporre un ricorso giurisdizionale** qualora ritenga che i diritti di cui un interessato gode siano stati violati



**tutela collettiva ex post**  
**conforme ai valori collettivi oggetto di protezione**

# Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data

23 gennaio 2017, Comitato della Convenzione 108

- **uso etico, consapevole e socialmente responsabile** dei dati e correlata valutazione di impatto giuridico, sociale ed etico
- **approccio preventivo** di valutazione dei rischi per la collettività → diritto di controllo più ampio
- **approccio by design e by default** e rispetto dei principi di data protection
- valutazione del rischio **de-anonimizzazione** e **re-identificazione con prova a carico del titolare** del trattamento
- **ruolo dell'intervento umano** nei processi decisionali basati sui big data
- agevolare il **consenso consapevole** → es. utilizzo interfacce grafiche che simulino utilizzo dati e impatto
- **cultura** quale abilità educativa essenziale

# Atti e documenti

- **comunicazioni della Commissione europea**
  - ✓ «Verso una florida economia basata sui dati» – 02/07/2014
  - ✓ «Costruire un'economia dei dati europea»– 10/01/2017
  - ✓ «Verso uno spazio comune europeo dei dati » - 25/04/2018
  
- **documenti dell'European Data Protection Supervisor (EDPS)**
  - ✓ Opinion 7/2015 «Meeting the challenges of big data» - 19/11/2015
  - ✓ Opinion 8/2016 «on coherent enforcement of fundamental rights in the age of big data» - 23/09/2016
  - ✓ Opinion 3/2018 «on online manipulation and personal data» - 19/03/2018
  
- **pubblicazioni dell'Organisation for Economic Co-operation and Development (OECD)**, quale «Data-Driven Innovation: Big Data for Growth and Well-Being», OECD Publishing, Paris, 2015



# Strategia europea per i dati

comunicazione Commissione europea del 19 febbraio 2020

Consapevolezza che il valore dei dati risiede nel loro utilizzo e che esistono squilibri in termini di potere di mercato

4 pilastri

- ✓ **quadro di governance intersettoriale** per l'accesso ai dati e il loro utilizzo
- ✓ **abilitatori**, consistenti in investimenti nei dati e nel rafforzamento delle infrastrutture e delle capacità europee per l'hosting, l'elaborazione e l'utilizzo dei dati, l'interoperabilità
- ✓ **competenze**, ossia fornire strumenti alle persone, investire nelle competenze e nelle PMI
- ✓ **spazi comuni europei di dati** in settori strategici e ambiti di interesse pubblico.

L'ambizione consiste nel «consentire all'UE di diventare l'economia agile basata sui dati più attrattiva, sicura e dinamica del mondo, fornendo all'Europa i dati necessari a migliorare le decisioni e la vita di tutti i suoi cittadini»

# Data Governance Act

proposta di regolamento UE del 25 novembre 2020

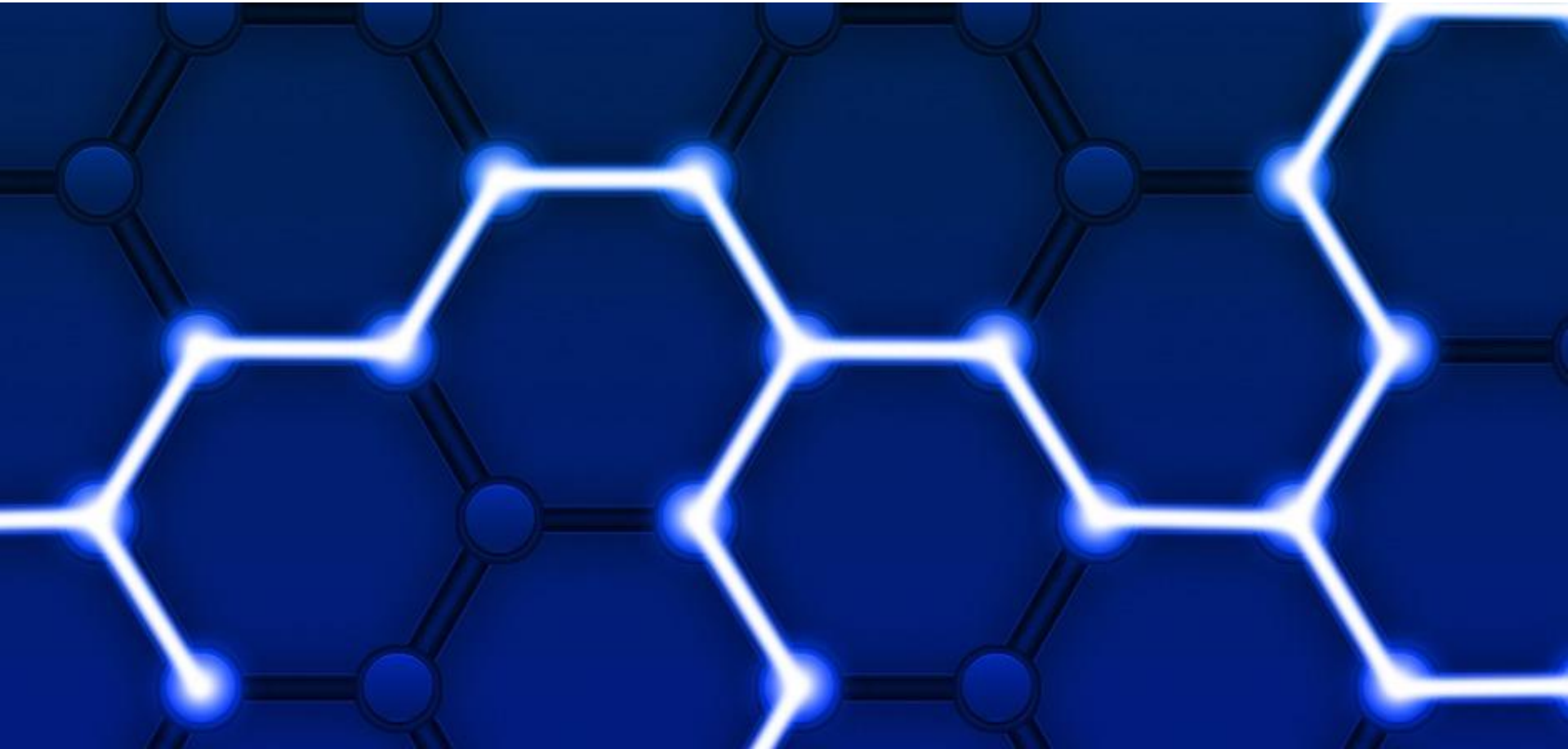
Atto volto a promuovere la **disponibilità dei dati utilizzabili** rafforzando la fiducia negli intermediari di dati e potenziando i meccanismi di condivisione dei dati in tutta l'UE, che integra la direttiva (UE) 2019/1024, dal momento che riguarda i dati oggetto di diritti di terzi detenuti da enti pubblici.

Propone:

- ✓ messa a disposizione dei dati del settore pubblico per il riutilizzo qualora tali dati siano oggetto di diritti di terzi
- ✓ condivisione dei dati tra le imprese, dietro compenso in qualsiasi forma
- ✓ consenso all'utilizzo di dati personali con l'aiuto di un "intermediario per la condivisione dei dati personali", il cui compito consiste nell'aiutare i singoli individui a esercitare i propri diritti a norma del regolamento generale sulla protezione dei dati
- ✓ consenso all'utilizzo dei dati per scopi altruistici

# *Ubi data society, ibi ius*

il governo dei dati necessita del ruolo del diritto e della forza dei diritti per riuscire a proteggere la persona nell'era digitale e disegnare il presente e il futuro della democrazia





**INFORMATICA GIURIDICA**  
Collana diretta da Giovanni Ziccardi e Pierluigi Perri

8

Fernanda Faini

## **Data society**

Governo dei dati e tutela dei diritti nell'era digitale

***Grazie per l'attenzione***



**Fernanda Faini**

Research Fellow e docente in diritto dell'informatica – Università di Pisa

**email**                      fernanda.faini@jus.unipi.it

**Linked in**                      fernandafaini

