FORUM
DISUGUAGLIANZE
DIVERSITÀ

MATERIALI

15 PROPOSTE PER LA GIUSTIZIA SOCIALE

Ispirate dal Programma
di Azione di Anthony Atkinson

Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza

Giorgio Resta
Università di Roma Tre

1. *Governare l'innovazione tecnologica: l'incrocio tra* big data *e* machine learning

Queste pagine intendono svolgere alcune riflessioni, dal punto di vista del giurista, sui problemi posti dal ricorso agli algoritmi quali strumenti di decisione sia in ambito pubblico sia in ambito privato[1].

È bene chiarire sin da ora che, per apprezzare correttamente la reale natura delle questioni coinvolte, è necessario concentrarsi non soltanto sul profilo dell'automazione nelle decisioni (algoritmi e *machine learning*), ma anche su quello della disponibilità di una massa enorme di dati, sulla quale si appuntano le tecniche di *data analytics* e che quindi rappresenta il presupposto essenziale per il funzionamento dei moderni algoritmi di apprendimento[2]. *Big data* e *machine learning,* in altri termini, sono i fattori fondamentali alla base delle due cruciali questioni regolatorie con le quali la società è oggi chiamata a confrontarsi:

a) come disciplinare la raccolta e l'uso dei dati fruibili per i trattamenti algoritmici (questione 'a monte');

b) come regolare il processo decisionale in quanto tale, sia nel suo *iter* procedimentale sia nei suoi effetti sociali, in modo da assicurare un equilibrato bilanciamento degli interessi collettivi coinvolti (questione 'a valle').

Su entrambi gli aspetti la discussione è accesa e non mancano esempi di interventi normativi, decisioni giurisprudenziali o prassi operative, dei quali si darà conto nelle pagine che seguono. La notevole attenzione rivolta alle suddette questioni - in ambito accademico, da parte di istituzioni ed enti di ricerca[3], nonché in alcuni casi rimarchevoli anche in sede di discussione politica, come

1 Sulla nozione di algoritmo, quale descrizione formalizzata e astratta di una procedura computazionale, e le sue implicazioni giuridiche v. W. Hoffmann-Riem, *Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht, Archiv des öffent. Rechts,* 142 (2017), 2 ss.

2 Sul punto D. Pedreschi – F. Giannotti et al., *Open the Black Box. Data-driven Explanation of Black Box Decision Systems,* in *ArXiv,* 1 (2018), 1-2; S. Barocas – A.D. Selbst, *Big Data's Disparate Impact,* 104 *California L. Rev.* 671 (2016).

3 Mi limito a ricordare i seguenti rapporti e documenti: House of Lords, Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?,* London, 2018; AI Now *Report 2018;* Council of Europe, *Discrimination, artificial intelligence, and algorithmic decision-making,* a cura di F.Z. Borgesius, Strasbourg, 2018; Art. 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation*

nell'esempio del *Koalitionsvertrag* tedesco tra CDU, CSU e SPD del 2018[4] - riflette una chiara consapevolezza della profonda ambiguità di tutti i grandi processi di innovazione tecnologica, capaci di imprimere una netta discontinuità alle dinamiche evolutive della società. Da un lato essi possono avere una valenza fortemente emancipatoria, redistribuendo il potere sociale e creando opportunità di crescita, progresso e miglioramento della condizione umana. Dall'altro, se non sono democraticamente governati, rischiano di consolidare le posizioni di privilegio, le disuguaglianze e le asimmetrie di potere esistenti in una data comunità organizzata[5].

È stato così per le innovazioni collegate alla prima rivoluzione industriale – e le lucide pagine di Karl Marx e Karl Polanyi stanno tuttora a ricordarcelo[6] – ed è così per quelle della quarta rivoluzione. Far pendere il piatto della bilancia verso l'uno o verso l'altro polo è il frutto di scelte sociali, rispetto alle quali la mediazione giuridica svolge un ruolo centrale. Ciò è ancor più vero in relazione alle innovazioni legate al mondo digitale, non foss'altro perché qui viene meno il primo e più elementare strumento di controllo e tutela dei beni, rappresentato dal possesso materiale. Rispetto ai beni intangibili, come l'informazione, qualsiasi meccanismo di allocazione esclusiva presuppone necessariamente l'intervento del diritto, che, come nel caso dei diritti di proprietà intellettuale, può creare situazioni di scarsità artificiale al fine di stimolare l'accumulazione di conoscenza e la produzione di innovazione. Misurare i caratteri e i limiti dell'intervento giuridico è allora essenziale, perché l'adozione di modelli di governo non equilibrati e troppo sbilanciati sul polo della protezione rischia – a tacer d'altro[7] - di allargare in maniera sproporzionata le sfere di proprietà, comprimendo artificialmente le sfere di libertà (civili, politiche ed economiche), con l'effetto di precludere il conseguimento di molte delle opportunità aperte dalle nuove tecnologie, specie in termini di condivisione delle conoscenze e accesso al patrimonio comune immateriale.

Di conseguenza, questo contributo dovrà prendere in considerazione non soltanto il problema della trasparenza e non discriminatorietà degli algoritmi (par. 3), ma, prima ancora, la questione del regime di appartenenza delle informazioni processate in forma automatica e usate per fini di decisione (par. 2).

2016/679, last revised February 2018; Council of Europe, *Draft Guidelines on Artificial Intelligence,* 2018; Federal Trade Commission, *Data Brokers. A Call for Transparency and Accountability,* Washington, 2014; Executive Office of the President of USA, *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights,* 2016.

4 Sono diversi passaggi dedicati agli algoritmi e al loro possibile effetto discriminatorio: v. *Ein Neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land.* Koalitionsvertrag zwischen CDU, CSU und SPD, 2018, pp. 2092, 2098, 6354.

5 In generale, ed altresì con riferimento alle tecniche di intelligenza artificiale, v. R. Baldwin, *The Globotics Upheaval. Globalization, Robotics, and the Future of Work,* Oxford, 2019; v. altresì S. Rodotà, *Tecnologie e diritti,* Bologna, 1997.

6 Si veda in proposito L. Basso, *Giustizia e potere,* in *Democrazia e diritto,* 1971, 549 ss., 555; K. Polanyi, *Per un nuovo Occidente. Scritti 1919-1958,* a cura di G. Resta – M. Catanzariti, Torino, 2013.

7 Per un'analisi d'insieme sull'importanza della regolamentazione giuridica per fini di costruzione di una sfera pubblica digitale v. P. Nemitz, *Constitutional Democracy and technology in the age of artificial intelligence, Phil. Trans. R. Soc. A* 376:20180089.

2. *Il controllo sui dati nell'economia digitale*

Quando si discorre di *big data analytics* si fa riferimento a due principali categorie di dati: *a*) i dati personali, ossia riferibili a un determinato individuo, identificato o identificabile; *b*) i dati non personali, ossia non riferibili a un determinato soggetto, per propria natura o perché sottoposti a un processo di anonimizzazione.

La gran parte dei processi decisionali automatizzati non sarebbe oggi pensabile prescindendo dall'accesso sistematico a entrambe le categorie di informazioni, sovente rese disponibili non direttamente dai *data subjects* o comunque dai *data sources,* bensì dagli intermediari professionali, i *data brokers,* i quali operano in un mercato ormai floridissimo, anche se poco conosciuto nei suoi dettagli organizzativi e regolamentato in maniera ancora frammentaria e lacunosa[8]. Quanto al rilievo di ciascuna tipologia di dati, si pensi, dal primo punto di vista, alle informazioni relative alla storia creditizia, alle propensioni di acquisto di un consumatore, oppure alle informazioni relative alla salute o alle preferenze politiche di un lavoratore: tutti dati utili ai fini della costruzione di un profilo individuale, e dunque spesso per fini di decisioni algoritmiche; e, dall'altro, alle informazioni prodotte dai macchinari intelligenti, come gli autoveicoli di ultima generazione (informazioni relative allo stato funzionale del veicolo, chilometri percorsi, anomalie rilevate, strade percorse, etc.), gli elettrodomestici interconnessi, oppure le informazioni prodotte dal settore pubblico (informazioni catastali, cartografiche, meteorologiche, etc.).

Diversi sono, ovviamente, i problemi sottesi all'utilizzo dell'una e dell'altra tipologia di informazione, le quali mettono in gioco in maniera diversa la sfera della soggettività individuale. Differenti, di conseguenza, dovrebbero essere i regimi giuridici applicabili, sia pure nella consapevolezza dell'estrema fluidità dei confini e della presenza di continue sovrapposizioni tra le due categorie di dati.

La tendenza degli ordinamenti giuridici occidentali è quella di assoggettare a garanzie e salvaguardie soprattutto il segmento delle informazioni personali, anche se ciò avviene secondo forme e con modalità molto diverse. Ad esempio, gli Stati Uniti sono ancora ben lungi dall'introdurre un regime generale di tutela dei dati personali, limitando gli statuti di protezione a specifici sotto-settori (come quelli dell'accesso al credito o delle informazioni sanitarie), per il timore di ostacolare eccessivamente lo sviluppo del mercato, frenare l'innovazione tecnologica in settori cruciali per la competitività internazionale, nonché inibire la libera diffusione delle informazioni, vista come diretta emanazione della garanzia del Primo Emendamento della

8 In proposito v. l'indagine conoscitiva della Federal Trade Commission, *Data Brokers. A Call for Transparency and Accountability,* Washington, 2014.

Costituzione federale USA[9]. Per contro, gli ordinamenti europei – che, però, sotto questo profilo possono vantare un'influenza globale sul piano dei modelli regolatori ben maggiore degli USA[10] – hanno da più di vent'anni optato per un sistema incisivo di controllo sulla circolazione dei dati personali, ispirato alla logica dei diritti fondamentali.

Nel prossimo paragrafo ci si interrogherà sui limiti posti alla raccolta e ulteriore trattamento dei dati personali per fini di profilazione e decisione algoritmica (par. 2.1.).

Nel paragrafo successivo si indagherà, invece, sul livello di tutela ascritto ai dati non personali, chiarendo in particolare se di essi possa predicarsi un regime d'appartenenza in forma esclusiva (par. 2.2.).

2.1. *I dati personali*

Il regime applicabile ai dati personali è ormai compiutamente delineato dal Regolamento UE n. 2016/679 (di seguito GDPR), al quale si aggiungerà a breve il Regolamento *e-privacy*, ancora in fase di negoziazione. Esso offre una risposta a molte delle questioni oggi sul tavolo, anche se permangono alcuni elementi d'ambiguità che potranno essere sciolti solo dalla prassi applicativa.

Quanto al profilo della raccolta dei dati, il diritto europeo muove, in termini generali, da una prospettiva opposta rispetto a quello americano. Mentre negli USA può ritenersi vigente un regime di libertà di trattamento dei dati personali, salve le specifiche ipotesi di divieto fissate a livello di leggi speciali[11], in Europa i dati personali non possono essere trattati se non in presenza di una specifica base autorizzativa rientrante tra quelle previste dall'Art. 6 GDPR per i dati personali 'comuni' e dall'art. 9 per le categorie particolari di dati (quelli che nella disciplina previgente si definivano "dati sensibili"). Tra le suddette cause di giustificazione rientra anche il consenso dell'interessato, che benché nell'economia degli artt. 6 e 9 svolga un ruolo marginale, gode sempre di preponderante attenzione sia nella prassi applicativa sia nelle analisi scientifiche.

Il riferimento al consenso permette di sciogliere subito un equivoco che aleggia spesso nella letteratura in materia. Deve ribadirsi che la previsione circa la necessità del consenso non giustifica

9 In generale, v. P.M. Schwartz – K.N. Peifer, *Transatlantic Data Privacy Law,* 106 *Georgetown Law Journal* 115 (2017).

10 G. Greenleaf, *The influence of European data privacy standards outside Europe: implications for globalization of Convention 108, Int'l Data Privacy Law,* 2 (2012), 68.

11 In generale, v. P.M. Schwartz – K.N. Peifer, *Transatlantic Data Privacy Law,* 106 *Georgetown Law Journal* 115 (2017). Tra le molte conseguenze di questa diversità di approccio, può annoverarsi anche il ricorso a maglie larghe negli USA del *political microtargeting,* su cui v. C. Bennett, *Voter databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America?*, 6 International Data Privacy Law, no. 4 (2016).

la conclusione che ciascun individuo debba ritenersi titolare di un diritto dominicale liberamente alienabile sui propri dati personali[12].

La finalità del GDPR non è quella di allocare titoli esclusivi alienabili e quindi porre le premesse per un efficiente mercato delle informazioni, bensì di bilanciare l'esigenza della circolazione intracomunitaria dei dati con il rispetto dei diritti fondamentali coinvolti (dignità, riservatezza, identità ed altre libertà civili: si noti che l'art. 8 della Carta dei Diritti UE configura il diritto alla protezione dei dati come autonomo diritto fondamentale della persona umana). Ciò si ricava dalle specifiche scelte normative compiute nel GDPR[13]: *a*) il consenso non è considerato sempre un requisito essenziale per il trattamento, anzi in molti casi esso non è necessario (per l'esecuzione di un contratto di cui è parte l'interessato, per il perseguimento del legittimo interesse di cui il titolare del trattamento è portatore, per compiti di interesse pubblico, etc.), sicché non può ritenersi vigente alcuna forma di tutela assoluta dei dati; *b*) perché sia valido è necessario provare la "libertà" della sua manifestazione (art. 4, n. 11), e dunque l'assenza di condizioni di disparità di potere sostanziale che connotano molte relazioni di mercato, specie nel quadro dei rapporti *online* che si avvalgono di condizioni di contratto standardizzate; *c*) anche qualora il consenso sia validamente inserito in un accordo contrattuale, realizzando una forma di cessione remunerata delle informazioni, esso è sempre revocabile (art. 7, c. 3), a testimonianza dell'assenza di quella stabilità delle posizioni negoziali, che è propria delle relazioni di mercato; *d)* infine esso non è di ostacolo all'esercizio del diritto alla portabilità, il quale testimonia l'esigenza di mantenere nelle mani della persona il potere di controllo sull'utilizzazione dei propri dati.

Una volta accertata la sussistenza dell'idonea base giuridica, il diritto europeo assoggetta il trattamento a importanti condizioni sostanziali e procedurali, le quali assumono un rilievo cruciale quando si sia in presenza di un'attività di "profilazione" dell'interessato[14]. Questa è definita dal GDPR come "qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica" [15]. Si deve notare che un trattamento anche

12 C. Berger, *Property Rights to Personal Data? An Exploration of Commercial Data Law,* in *Zeitschrift für geistiges Eigentum,* 9, 2017, p. 340; S. Gutwirth – G. González Fuster, *L'éternel retour de la propriété des données: de l'insistance d'un mot d'ordre,* in Degrave-de Terwangne-Dusollier-Queck (a cura di), *Law, norms and freedoms in cyberspace - Liber amicorum Yves Poullet*, Bruxelles 2018, 117.

13 Si veda in questo senso J. Drexl, *Legal Challenges of the Changing Role of Personal and Non-Personal Data in the Data Economy,* in A. De Franceschi et al., *Digital Revolution: New Challenges for Law,* forthcoming, Beck, 2019 (on file with the author).

14 Cfr. Art. 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679,* last revised February 2018, 6 e ss.

15 In generale v. F. Pizzetti, *La protezione dei dati personali e la sfida dell'Intelligenza Artificiale,* in Id., a cura di, *Intelligenza artificiale, protezione dei dati personali e regolazione,* Torino, 2018, 30 ss.

parzialmente automatizzato non esclude la sussistenza di un'attività di profilazione, contrariamente a quanto previsto dall'art. 22 GDPR in relazione alle decisioni interamente automatizzate, che soggiacciono al regime di divieto solo in quanto l'intervento umano sia del tutto escluso[16].

Quanto ai principi che disciplinano il trattamento, questo deve anzitutto essere condotto in maniera "trasparente". Ciò implica uno specifico onere informativo nei confronti dell'interessato (sia che i dati siano forniti da costui in maniera volontaria, sia che essi siano espunti da altra fonte), precisato nei suoi lineamenti dagli artt. 12-14. Si deve notare che a tal riguardo che è espressamente previsto l'obbligo di comunicare informazioni circa "l'esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'art. 22, par. 1 e 4, e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato" [17].

Si deve poi osservare il parametro della "correttezza", sicché anche un trattamento formalmente lecito potrebbe rivelarsi scorretto se ad esempio i dati sono impiegati in maniera tale da produrre effetti discriminatori, per precludere l'accesso a beni e servizi fondamentali, etc.

Inoltre, i dati in oggetto devono essere esatti ed accurati e in ogni caso devono rispettare la regola aurea della "minimizzazione": non è possibile cioè ricorrere a una massa sovrabbondante di dati, a meno che ciò non sia strettamente necessario rispetto alle finalità sottese al trattamento, né è ammissibile conservare tali dati per un lasso temporale sproporzionato (art. 5, n. 1, lett. e). Ciò si sposa, peraltro, con la regola per cui gli strumenti automatizzati di trattamento dei dati devono essere progettati sin da principio e devono operare in via predefinita in modo da ridurre al massimo la quantità di dati personali trattati e l'incidenza di tali operazioni sulla sfera della persona (*privacy by design* e *privacy by default*).

Infine, deve essere osservato il principio della finalità, sicché il trattamento validamente iniziato in relazione a un determinato scopo, come indicato nell'informativa resa al soggetto, non giustifica in linea di massima l'impiego dei dati per il conseguimento di scopi distinti, fatte salve le condizioni di compatibilità previste nell'art. 6, c. 4 GDPR. Un'applicazione importante di tale logica si rinviene nei limiti posti all'interconnessione degli archivi della p.a., la quale è subordinata a una specifica previsione di legge, ai sensi dell'art. 6, c. 1, lett. e) e art. 6, c. 3 GDPR[18].

Fra gli altri strumenti atti ad operare in chiave preventiva, qualora il trattamento presenti un rischio elevato per i diritti e le libertà delle persone fisiche, v'è la "valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali" (art. 35). Il titolare è tenuto ad adottare un

16 *Ibid.,* 7.

17 In tema M. Temme, *Algorithms and Transparency in View of The General Data Protection Regulation,* 3 *Eur. Data Prot. L. Rev.* 473 (2017), 482.

18 V. S. D'Ancona, *Trattamento e scambio di dati e documenti tra pubbliche amministrazioni, utilizzo delle nuove tecnologie e tutela della riservatezza tra diritto nazionale e diritto europeo, Riv. it. Dir. pubb. Com.,* 2018, 587 ss.

siffatto documento qualora si intenda porre in essere "una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche".

Queste ed altre regole, che non è possibile qui analizzare nel dettaglio, delineano un meccanismo di controllo abbastanza capillare, che restringe *ex ante* la quantità e la tipologia dei dati personali utilizzabili – entrando di fatto in conflitto con un altro segmento importante della legislazione comunitaria che, soprattutto in materia finanziaria e bancaria, incentiva il ricorso a forme di invasive di profilazione dei clienti al fine di valutarne la solvibilità e il merito di credito[19] - assicurando al contempo che questi abbiano un elevato grado di 'qualità', nel senso di accuratezza, esattezza e granularità[20]. In questo senso si prevede che i dati debbano essere esatti e se necessario aggiornati, non possano essere conservati per un periodo eccessivo di tempo, e in ogni caso sono suscettibili di accesso, controllo, rettifica, integrazione e persino cancellazione su istanza del soggetto interessato (artt. 13-21 GDPR).

Ciò significa, in conclusione, che l'infrastruttura regolatoria si caratterizza per una serie di filtri atti a elevare la qualità dell'ecosistema informativo, selezionando *ex ante* tipologia, volume e caratteri delle informazioni utilizzabili per fini di profilazione, analisi a scopo predittivo e decisioni algoritmiche. Ciò rappresenta una circostanza non trascurabile, perché come è ben noto, gli algoritmi funzionano secondo la logica *garbage in – garbage out*, per cui dati incongrui, inesatti o non aggiornati non possono che produrre risultati decisionali inaffidabili[21]. D'altra parte, non può dimenticarsi che *non tutti* i dati immessi nel processo automatizzato sono dati personali in senso stretto.

2.2. *I dati non personali*

19 V. ad es. il Considerando 27 e l'art. 8 della Direttiva 2008/48/CE, relativa ai contratti di credito ai consumatori, ove si prevede che "Member States shall ensure that, before the conclusion of the credit agreement, the *creditor assesses the consumer's creditworthiness on the basis of sufficient information*, where appropriate obtained from the consumer and, where necessary, *on the basis of a consultation of the relevant database*. Member States whose legislation requires creditors to assess the creditworthiness of consumers on the basis of a consultation of the relevant database may retain this requirement"; nonché il Consumer Financial Services Action Plan della Commissione del 2017. In tema v. l'indagine di V. Zeno-Zencovich, *'Smart Contracts', 'Granular Legal Norms', and Non-Discrimination,* di prossima pubblicazione.

20 Per un panorama più dettagliato v. F. Pizzetti, *La protezione dei dati personali e la sfida dell'Intelligenza Artificiale,* cit.

21 M. Temme, *Algorithms and Transparency in View of The General Data Protection Regulation,* 3 *Eur. Data Prot. L. Rev.* 473 (2017), 478.

Le informazioni rilevanti per gli algoritmi di apprendimento automatico non sono, come si è appena ricordato, soltanto quelle atte ad identificare un individuo determinato. Tutte le informazioni prodotte da macchine, o i flussi inerenti le comunicazioni elettroniche, ad esempio, costituiscono, una porzione importante dell'universo *big data,* ma non rientrano necessariamente nel novero dei "dati personali", o perché strutturalmente non riferibili a una persona determinata, o perché oggetto di un processo di anonimizzazione (che preclude l'applicabilità della disciplina in materia di protezione dei dati personali) [22].

Il regime giuridico dei "dati non personali" – come definiti, in negativo, dal recente Regolamento 2018/1807/UE sulla libera circolazione dei dati non personali nell'Unione Europea[23] – è molto meno chiaro e univoco di quello relativo a dati personali[24]. In particolare, è quanto mai accesa la discussione se il dato, in quanto tale, possa costituire il termine di riferimento di una pretesa di natura proprietaria[25]. In questo senso, in particolare, sembrava essere orientata la Commissione con la Comunicazione al Parlamento Europeo e al Consiglio del 2017 su "Building a European Data Economy", nella quale si evocava l'introduzione di un nuovo "diritto del produttore di dati" [26].

Che si ponga la questione in questi termini non sorprende più di tanto, poiché accade di continuo che, di fronte all'emersione di nuovi beni patrimonialmente rilevanti, il primo schema al quale ci si rivolge per operare un inquadramento della realtà è quello del diritto di proprietà. È stato così per l'immagine, per l'etere, per l'energia, ed è così oggi per i dati non personali. Tuttavia, applicato all'informazione, lo schema proprietario si rivela fuorviante e non è in grado di apprestare soluzioni operazionali adeguate[27].

Difatti, essendo l'informazione un bene tipicamente non rivale nel consumo e suscettibile di produrre conoscenza incrementale, introdurre barriere artificiali alla sua circolazione tramite il riconoscimento di un diritto di esclusiva significa operare una forma di etero-regolazione, la quale può giustificarsi soltanto al fine di rimediare a un fallimento del mercato. Nel caso degli ordinari diritti di proprietà intellettuale, questo è rappresentato dalla c.d. sotto-produzione del bene*:* in assenza di un monopolio legale di sfruttamento, l'autore di una creazione estetica o utile non riuscirebbe a far propri gli utili derivanti dallo sfruttamento del bene (l'informazione è bene non escludibile, oltre che non rivale non consumo), sicché non avrebbe un incentivo sufficiente a

22 Per un'analisi dettagliata v. C. Wendehorst, *Of Elephants in the Room and Paper Tigers: How to Reconcile Data Protection and the Data Economy*, S. Lohsse – R. Schulze - D. Staudenmayer (a cura di), *Trading Data in the Digital Economy: Legal Concepts and Tools*, Oxford – Baden-Baden*,* 2017, 327 ss.

23 Art. 3, Regolamento 2018/1807/UE.

24 J. Drexl, *Legal Challenges of the Changing Role of Personal and Non-Personal Data in the Data Economy,* cit.

25 D. Zimmer, *Property Rights Regarding Data?,* in S. Lohsse – R. Schulze - D. Staudenmayer (a cura di), *Trading Data in the Digital Economy: Legal Concepts and Tools*, cit., 101 ss.; M. Becker, *Rights in Data – Industry 4.0 and the IP Rights of the Future,* in *Zeitschrift für geistiges Eigentum,* 9, 2017, p. 253.

26 COM(2017) 9 final, p. 13.

27 Sulla linea qui sostenuta v. J. Drexl, *Designing Competitive Markets for Industrial Data. Between Propertisation and Access, JIPITEC,* 8 (2017), 257.

investire e produrre, con danno per l'intera collettività. Per contro, nel caso dei dati grezzi, privi cioè di un'immediata utilità estetica o industriale (quale quella sottesa al riconoscimento del diritto d'autore o del diritto di brevetto), non è la promessa di un profitto monopolistico a costituire la molla principale per innescare il processo di produzione della risorsa, bensì lo sono fattori diversi come la concorrenza e il miglioramento dell'assetto tecnologico[28]. Non a caso, nella realtà attuale, che pur è connotata dall'assenza di diritti di privativa, le informazioni 'grezze' vengono nondimeno prodotte in maniera intensiva e in diversi casi rappresentano la principale voce di capitale di un'impresa.

In altri termini, le istituzioni giuridiche esistenti, e in primo luogo gli istituti del segreto industriale e della concorrenza sleale[29], offrono strumenti di salvaguardia adeguati per proteggere il parco informativo di un'impresa, senza che sia necessario ricorrere a misure controproducenti come quelle del riconoscimento di nuovi diritti di esclusiva[30]. Controproducenti non soltanto perché rischiano di frenare il processo di sviluppo e innovazione, creando continui conflitti circa la titolarità della risorsa (si pensi alle informazioni circa le buche stradali rilevate dai sensori delle autovetture o da altri meccanismi di rilevazione) e elevando a sistema il potere di veto dei singoli detentori di spezzoni di informazioni utili soprattutto in forma aggregata per fini di apprendimento automatico (problema degli *anti-commons*); ma anche perché tali pretesi diritti di esclusiva non potrebbero non riflettersi negativamente sulla libera circolazione dei dati, e dunque sulle garanzie della libertà d'informazione, particolarmente rilevanti oggi per il funzionamento dei processi democratici in un ambiente digitale.

In conclusione qualsiasi tentativo di introdurre nuove forme di esclusiva concernenti dati non personali deve essere rigettato in quanto non giustificabile sul piano funzionale – come in parte si è rivelata essere la scelta di introdurre un diritto sui generis sulle banche di dati non creative - e foriero di conseguenze nocive sul piano dell'innovazione tecnologica e della trasparenza dei processi democratici.

3. *Le decisioni algoritmiche e le prospettive dell'uguaglianza*

28 Per un panorama sulla prassi v. C. Wendehorst, *Besitz und Eigentum im Internet der Dinge,* in H. Micklitz – L.A. Reisch et al., a cura di, *Verbraucherrecht 2.0 – Verbraucher in der digitalen Welt,* Baden-Baden, 2017, 367.

29 T. Aplin, *Trading data in the digital economy: trade secrets perspective,* in S. Lohsse – R. Schulze - D. Staudenmayer (a cura di), *Trading Data in the Digital Economy: Legal Concepts and Tools*, cit., 59 ss.

30 J. Drexl, *Designing Competitive Markets for Industrial Data. Between Propertisation and Access,* cit., 260-261.

Il ricorso a trattamenti algoritmici per finalità di previsione e/o di decisione è ormai all'ordine del giorno tanto nel settore pubblico quanto nel settore privato[31].

Si pensi, dal primo punto di vista, all'uso degli algoritmi da parte della p.a. per decidere questioni seriali o fondate su parametri predeterminati (come l'assegnazione degli insegnanti alle sedi scolastiche vacanti)[32]; per orientare la prestazione dei servizi sociali (in Pennsylvania – ricorda un articolo apparso su Nature lo scorso anno – è stato messo in atto un sistema di profilazione e *scoring,* individuale, l'*Allegheny Family Screening Tool*, finalizzato ad individuare i bambini a rischio di esclusione sociale e maltrattamento e progettare gli interventi di tutela) [33]; per operare valutazioni *data-driven* delle prestazioni dei dipendenti pubblici (come in un noto caso USA concernente la valutazione degli insegnanti) [34]; per gestire i flussi migratori e effettuare uno *screening* preventivo dei *files* dei richiedenti asilo, visto di ingresso e soggiorno, etc. [35]; per orientare le azioni di contrasto al terrorismo, sfruttando le potenzialità delle analisi predittive ma non di rado ponendo le premesse per un immenso e capillare sistema di sorveglianza occulta degli individui (è quanto è emerso con le rivelazioni di Edward Snowden) [36]; per indirizzare le operazioni di polizia e prevenire la commissione dei reati (come nel caso dell'applicativo PredPol) [37]; per assumere decisioni concernenti l'amministrazione della giustizia penale (emblematico è l'esempio di COMPAS, il software utilizzato in diverse giurisdizioni USA al fine di calcolare il rischio di recidiva e la pericolosità sociale di un soggetto sottoposto a procedimento penale, e quindi misurare l'entità e la tipologia della pena irrogabile) [38].

Si pensi, dal secondo punto di vista, al trattamento algoritmico nell'ambito dei rapporti di lavoro (algoritmi computerizzati vengono ormai di frequente utilizzati al fine di operare forme di sollecitazione selettiva alla presentazione delle domande d'impiego, per gestire in forma interamente o parzialmente automatizzata il processo di assunzione, o per operare la valutazione delle prestazioni dei dipendenti) [39]; alla vendita di beni e servizi (è da ciò che dipende l'applicazione di prezzi, e spesso anche condizioni d'offerta, differenziati nei rapporti *on line*) [40]; al mercato del

31 S.C. Olhede – P.J. Wolfe, *The growing ubiquity of algorithms in society: implications, impacts, and innovations, Phil. Trans. R. Soc. A* 376:20170364.

32 V. TAR Lazio, 10-9-2018, n. 9227; TAR Lazio, 22-3-2017, n. 3769.

33 R. Courtland, *The Bias Detectives, Nature,* 558 (2018), 357.

34 *Houston Fed. Of Teachers v. Houston Ind. School District,* 251 F. Supp. 3d 1168 (2017).

35 Per molti esempi v. M. Hu, *Algorithmic Jim Crow,* 86 *Fordham L. Rev.* 633 (2017).

36 M. Hu, *Small Data Surveillance v. Big Data Cybersurveillance*, in 42 *Pepp. L. Rev.* 773 (2015).

37 Per un'attenta analisi giuridica del problema del *predictive policing* v. T. Rademacher, *Predictive policing im deutschen Polizeirecht, Archiv des öffent. Rechts,* 142 (2017), 366 ss.

38 R. Courtland, *The Bias Detectives,* cit., 358-359.

39 P. Kim, *Data-driven Discrimination at Work,* 58 *William & Mary L. Rev.* 857 (2017); D.J. Dalenberg, *Preventing discrimination in the automated targeting of job advertisements,* 34 *Computer Law & Security Rev.* 615 (2108)*; e* B. Dzida – N. Groh, *Diskriminierung nach dem AGG beim Ansatz von Algorithmen im Bewerbungsverfahren, NJW,* 2018, 1917.

40 Sul tema dei prezzi differenziati v. ad es. T. Tillmann – V. Vogt, *Personalisierte Preise im Big-Data-Zeitalter,* in *Verbraucher und Recht,* 2018, 447; sulla segmentazione degli utenti nel servizio Airbnb, v. C. Lutz – G. Newlands,

credito (si pensi ai meccanismi di *credit scoring* al fine di valutare l'affidabilità finanziaria nel quadro delle procedure di finanziamento a singoli individui e ad imprese) [41]; alla comunicazione (dal *microtargeting* nella comunicazione politica, all'ordinamento delle notizie rese fruibili agli utenti da un *social network* quale *Facebook*) [42]; nonché ovviamente ai mercati finanziari (basti un rinvio al Regolamento 2017/589/UE in materia di *trading* algoritmico).

Il processo decisionale algoritmico, mette conto precisare, può essere interamente automatizzato, come nel caso dei filtri anti-spam che in maniera del tutto autonoma selezionano il tipo di messaggi da indirizzare nella casella della posta indesiderata, oppure può trattarsi di una delega soltanto parziale alla macchina, come nell'ipotesi in cui alla valutazione computerizzata dell'affidabilità finanziaria di un cliente faccia seguito una decisione umana definitiva circa la concessione di una linea di credito.

Come si può agevolmente intuire, l'automazione del processo decisionale permette, se ben congegnata, di conseguire notevoli vantaggi in termini di uniformità, affidabilità e controllabilità della decisione stessa. Essa appare dunque astrattamente in linea, non soltanto con le istanze di calcolabilità delle relazioni di mercato, ma anche con i valori di neutralità ed efficacia dell'azione amministrativa di cui all'art. 97 Cost. [43].

Per altro verso, però, la logica stessa delle tecniche di *big data analytics* porta con sé alcuni rischi, che devono essere attentamente considerati. Poiché la ricostruzione di tendenze predittivamente rilevanti avviene a partire dalle occorrenze empiriche esistenti, dalle quali le macchine ricavano *trend* utili ad orientare la valutazione di situazioni future, l'intero sistema ha la propensione a "codificare" il passato, ingabbiando soluzioni e predizioni all'interno delle griglie fornite dai trascorsi storici e dal set di valori che ha guidato la programmazione del sistema[44]. Ciò significa, in altri termini, che un determinato 'stato del mondo' tende a essere cristallizzato nel

Consumer Segmentation Within the Sharing Economy: The Case of Airbnb, 88 *Journ. Business Research* 187 (2018).

41 D. Keats Citron – F. Pasquale, *The Scored Society: Due Process for Automated Predictions,* 89 *Washington Law Review* 1 (2014).

42 In generale v. M. Ebers, *Beeinflussung und Manipolation von Kunden durch Behavioral Microtargeting. Verhaltenssteuerung durch Algorithmen aus der Sicht des Zivilrechts,* in *Multimedia und Recht,* 2018, 423.

43 In particolare, come si è autorevolmente notato, "i vantaggi di un'automazione dei processi decisionali amministrativi sono evidenti con riferimento a procedure seriali o standardizzate, caratterizzate da un alto tasso di vincolatezza o fondate su presunzioni, probabilisticamente significative di un certo fatto. Si pensi alle procedure di trasferimento contestuale o di prima assegnazione di sede agli insegnanti; e si pensi alla erogazione di contributi assistenziali agli aventi diritto sulla base di parametri predeterminati. Ma si pensi anche, quanto alle decisioni sfavorevoli, agli accertamenti fiscali fondati su base presuntiva i cui dati siano 'messi insieme' da una macchina o alle sanzioni amministrative (per esempio per eccesso di velocità) elaborate in via automatizzata sia quanto alla rilevazione dell'infrazione sia per la determinazione della correlativa sanzione e la 'formazione' stessa del provvedimento" (F. Patroni Griffi, *La decisione robotica e il giudice amministrativo,* accessibile all'indirizzo https://www.giustizia-amministrativa.it/documents/20142/147941/Patroni%20Griffi%20%20La%20decisione%20robotica%20e%20il%20giudice%20amministrativo%20-%2028%20agosto%202018.pdf/24218a2e-47b7-1c0a-b2eec1b670347f95/Patroni+Griffi+-+La+decisione+robotica+e+il+giudice+amministrativo+-+28+agosto+2018.pdf)

44 Per una spiegazione puntuale e accessibile dei modelli matematici e statistici sottesi alle decisioni algoritmiche C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy,* London, 2017.

processo prognostico, influenzandone i risultati ed orientando più o meno incisivamente le decisioni prese a valle della valutazione automatizzata.

Se questo può non apparire particolarmente problematico quando si prendano in esame accadimenti naturali, come l'andamento delle perturbazioni per fini di previsioni meteorologiche, ben diversa è la situazione qualora le tecniche predittive si appuntino su stati dell'uomo e su processi sociali[45]. Qui, infatti, uno dei pericoli più evidenti è che le condizioni di disparità sociale esistenti in un dato momento storico si riflettano sul giudizio prognostico tramite la costruzione di profili individuali o di gruppo, composti per inferenza da fattori come la propensione al consumo, la capacità di spesa, il luogo di residenza, i trascorsi familiari, il grado di istruzione, la storia giudiziaria, etc. [46] Se non adeguatamente monitorate e rese neutre rispetto ai rischi di *bias* già insiti nella selezione dei dati rilevanti, le decisioni algoritmiche che si basano su tali fattori sono atte a produrre effetti discriminatori e aggravare il peso delle disuguaglianze, invece che contribuire a ridurle, come pure la tecnologia potrebbe fare. Peraltro, come è ben noto, è lo stesso corretto funzionamento del processo deliberativo democratico a essere minacciato da una incontrollata proliferazione di ciò che Cathy O'Neil ha definito le "weapons of math destruction": la sollecitazione personalizzata resa possibile dalle moderne tecnologie della comunicazione tende a segmentare artificialmente (e condizionare la formazione) delle preferenze politiche, con tutti quegli effetti distorsivi sul piano dell'esercizio dei diritti democratici che da ultimo il caso *Cambridge Analytica* ha compiutamente illustrato[47]. È ovvio, poi, che qualora ci si muova in un contesto non democratico, le possibilità di accesso, aggregazione dei dati e profilazione, offerte dalle moderne tecnologie, sono tali da assicurare un controllo capillare sui comportamenti individuali, capace di reprimere qualsiasi forma di dissenso e segmentare i cittadini e le imprese in liste "rosse" e "nere" che evocano i peggiori incubi orwelliani. Che non si tratti di distopia, ma di preoccupante realtà, è dimostrato dal "social credit system" posto in atto dal governo cinese a partire dal 2014, con l'obiettivo di rafforzare la fiducia nelle istituzioni e nei mercati, e consistente nell'attribuzione di un punteggio individuale (e correlative penalizzazioni) in funzione del grado di aderenza a regole e norme sociali mostrato dai singoli nel corso della vita quotidiana[48].

Ma torniamo al tema iniziale dell'effetto discriminatorio, ricorrendo a due esempi in grado di chiarire meglio i termini del problema.

45 D. Keats Citron – F. Pasquale, *The Scored Society: Due Process for Automated Predictions,* cit., 4 ss.
46 S. Barocas – A.D. Selbst, *Big Data's Disparate Impact,* cit., 677 ss.
47 Cfr. F.J.Z. Borgesius et al, *Online Political Microtargeting: Promises and Threats for Democracy,* 14 *Utrecht L. Rev.* 82 (2018); B. Bodó – N. Helberger – C. de Vreese, *Political micro-targeting: a Manchurian candidate or just a dark horse?*, *Internet Policy Review*, *6* (2017). DOI: 10.14763/2017.4.776; W. Hoffmann-Riem, *Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht,* cit., 14-15.
48 M. Hu, *Small Data Surveillance v. Big Data Cybersurveillance*, in 42 *Pepp. L. Rev.* 773 (2015); e Y.J. Chen et al., *'Rule of Trust': The Power and Perils of China's Social Credit Megaproject,* 32 *Columbia J. Asian Law* 1 (2018).

Il più noto è senza dubbio quello relativo all'uso di un algoritmo computerizzato, noto con l'acronimo COMPAS (C*orrectional Offender Management Profiling for Alternative Sanctions*), atto a quantificare il rischio di recidiva di soggetti sottoposti a procedimento penale. Prodotto da una società commerciale, esso è stato impiegato in diverse giurisdizioni USA per calcolare la probabilità di commissione di altri reati nell'arco dei due anni successivi, e quindi per decidere sia in merito alla remissione in libertà su cauzione, sia alla stessa quantificazione della pena[49]. Diversi studi hanno testato il funzionamento del suddetto algoritmo, dimostrando la presenza di un pregiudizio sistematico a danno delle persone di colore. In particolare, un rapporto realizzato da Propublica dimostra che tra coloro i quali sono stati classificati ad alto rischio di azioni criminogene, che non hanno però nei due anni successivi compiuto atti illeciti, i bianchi sono in percentuale del 23,5%, mentre gli Afro-americani del 44,9%[50]. Per contro, in quelli classificati a basso rischio, che invece si sono resi effettivamente responsabili di atti criminali, i bianchi sono 47,7%, mentre gli afro-americani sono il 28%. È interessante capire a quali fattori sia imputabile un siffatto 'bias' discriminatorio. Un dato che emerge da diverse ricerche è che i punteggi elaborati da COMPAS sono la risultante delle risposte a 137 questioni, offerte direttamente dagli indagati o desunte da altri dati pubblici. L'origine etnica non potrebbe per legge rientrare tra le domande, ma rileva indirettamente, atteso che vengono presi in esame fattori spesso statisticamente correlati all'appartenenza 'razziale', come il luogo di residenza, i precedenti penali (personali o familiari), il consumo di stupefacenti, il livello di istruzione, etc.

Il secondo esempio attiene ai rapporti di mercato. Esso concerne il *software* utilizzato da Amazon per individuare le città e i circondari dove offrire il servizio di consegna in un giorno. Un'indagine compiuta da Bloomberg nel 2016 ha fatto venire alla luce una chiara stratificazione per fasce di reddito e in genere capitale sociale[51]. Mentre tutte le principali città statunitensi risultano coperte, al loro interno vi sono 'buchi' di copertura del servizio legati in maniera nient'affatto casuale con le aree più povere e degradate del territorio, come il Bronx a New York e Roxbury a Boston. Questo banalissimo esempio mostra come orientare i comportamenti futuri esclusivamente in base alle condizioni esistenti rischi di cristallizzare le disparità del presente, spingendo i più svantaggiati sempre più in basso nella piramide sociale.

49 Di qui la controversia decisa dalla Supreme Court del Wisconsin nel caso *State v. Loomis,* 881 N.W.2d 749 (2016), nella quale è stata rigettata la richiesta di ritenere l'irrogazione di una pena basata sulle risultanze dell'applicativo COMPAS come confliggente con le garanzie del *due process*.

50 J. Angwin – J. Larson – S. Mattu – L. Kirchner, *Machine Bias, ProPublica,* 23-5-2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

51 D. Ingold – S. Soper, *Amazon Doesn't Consider the Race of Its Customers. Should it?,* https://www.bloomberg.com/graphics/2016-amazon-same-day/.

Come già indicano questi semplici esempi, alla radice dell'effetto discriminatorio dell'algoritmo possono celarsi diversi fattori, dei quali il programmatore non sempre ha piena consapevolezza[52].

Innanzitutto, la scelta delle variabili o delle categorie in base alle quali è costruita la decisione algoritmica può tradursi in forme di discriminazione indiretta. Ad esempio, se nella programmazione di un algoritmo utilizzato per decisioni relative all'assunzione di personale si concretizza la nozione di 'buon' dipendente avendo riguardo – tra gli altri – al criterio della puntualità nel recarsi sul luogo di lavoro, ciò può finire per penalizzare sistematicamente tutti coloro che vivono in periferia ed impiegano di conseguenza maggior tempo per raggiungere ogni giorno la sede dell'impresa. E poiché in determinati contesti vivere in periferia si correla in misura prevalente con origini etniche, condizioni sociali e di reddito più disagiate, un siffatto criterio formalmente 'neutro' potrebbe finire per riprodursi a danno di categorie già svantaggiate.

In secondo luogo, può risultare fallace il set di dati sui quali si esercitano i processi di auto-apprendimento delle macchine. In particolare, tali dati potrebbero essere relativi a un campione troppo ristretto o parziale, o riflettere essi stessi una situazione discriminatoria. Se, ad esempio, si calcolasse l'attitudine a delinquere esclusivamente sulla base delle statistiche relative alla popolazione carceraria negli USA, se ne trarrebbe un risultato viziato in partenza, poiché è noto che gli Afro-americani rappresentano una quota preponderante di tale popolazione.

In terzo luogo, può rivelarsi problematica la tipologia di dati presa in considerazione. Ad esempio, se le scelte in materia di occupazione fossero fatte automaticamente in base al *ranking* delle università di provenienza, ciò avrebbe verosimilmente come risultato quello di premiare le fasce più alte, per censo ed istruzione, della popolazione.

In quarto luogo, le *proxies* utilizzate potrebbero correlarsi a fattori indici di disparità sociale: si pensi soltanto al codice di avviamento postale utilizzato come strumento predittivo del rischio di *default* rispetto alla restituzione di un credito: si tratta anche qui di un criterio formalmente neutro, ma che in realtà fotografa – specie se unito ad altri dati - una ben precisa storia di vita della persona in questione.

Infine, l'algoritmo potrebbe essere stato programmato in maniera intenzionalmente discriminatoria, come nei casi riportati sulla stampa di pubblicità *on line* richieste a *social network* in modo da escludere persone di origine ispanica, o individui con un determinato orientamento sessuale.

52 Quanto segue sintetizza l'analisi compiuta da Council of Europe, *Discrimination, artificial intelligence, and algorithmic decision-making,* a cura di F.Z. Borgesius, cit., 10-14; e da S. Barocas – A.D. Selbst, *Big Data's Disparate Impact,* cit..

3.1. *Etica, governo e regolazione degli algoritmi*

Queste brevi considerazioni inducono a sottolineare tre dati.

Il primo è che gli algoritmi possono apportare notevoli benefici non soltanto per la loro intrinseca attitudine alla razionalizzazione in senso weberiano del processo decisionale, con aumento dei livelli di rapidità ed efficienza rispetto ai costi, ma anche come strumenti di riduzione delle disuguaglianze tramite, ad esempio, l'allocazione mirata delle prestazioni sociali, il contrasto alle frodi o all'evasione fiscale, o più in generale lo stimolo ai processi partecipativi.

Il secondo è che perché i benefici attesi si traducano in effettiva prassi operativa e non siano sopravanzati dai rilevati effetti distorsivi, è necessario pre-formare le modalità di funzionamento degli algoritmi, assicurandone una sorta di *legality by design,* in modo da ridurre al minimo, e possibilmente eliminare, i rischi di impatto negativo sui diritti civili, sociali e politici delle persone[53]. Quando si parla di rischi specifici della decisione algoritmica, si fa soprattutto riferimento in letteratura a tre principali ordini di problemi:

a) la segretezza o l'inintelligibilità della logica sottesa al processo decisionale, la quale è particolarmente acuta nel caso degli algoritmi di apprendimento automatico (problema del *black box*) [54];

b) l'attitudine discriminatoria dell'algoritmo (problema del *bias*) [55];

c) la mortificazione della persona umana, resa oggetto di decisioni interamente automatizzate (problema della *dignità*)[56].

Il terzo dato è che per governare i problemi suindicati non è sufficiente affidarsi unicamente allo strumento tecnologico, quale ad esempio lo sviluppo di appositi algoritmi di auto-apprendimento volti a scovare e correggere l'esistenza di *bias* decisionali (c.d. *bias busting)* e a promuovere il valore della 'correttezza' decisionale (*fairness formulas*)[57]; né a dichiarazioni d'impegno e codici di autoregolamentazione dei soggetti professionali coinvolti (come quelli prodotti dall'organizzazione *Fairness, Accountability, and Transparency in Machine Learning*, dalla *IEEE*, dal *Future of Life Institute*) [58].

53 Vedi D. Cardon, *Le pouvoir des algortihmes, Pouvoir,* 2018, 63 ss.

54 J.A. Kroll, *The Fallacy of Inscrutability, Phil. Trans. R. Soc. A* 376:20180084; H. Shah, *Algoritmic Accountability, Phil. Trans. R. Soc. A,* 376:20170362 (2018); D. Pedreschi – F. Giannotti et al., *Open the Black Box. Data-driven Explanation of Black Box Decision Systems,* cit.

55 A. Chander, *The Racist Algorithm?,* 115 *Michigan L. Rev.* 1023 (2017).

56 Sachverständigenrat für Verbraucherfragen (SVRV), *Lösungsoptionen,* in H. Micklitz – L.A. Reisch et al., a cura di, *Verbraucherrecht 2.0 – Verbraucher in der digitalen Welt,* Baden-Baden, 2017, 25; G. Noto La Diega, *Against the De-Humanisation of Decision-Making,* 1 *JIPITEC* 3 (2018).

57 In tema v. D.R. Desai – J.A. Kroll, *Trust But Verify: A Guide to Algorithms And the Law,* 31 *Harvard J. Law & Tech.* 1 (2018), 35 ss. Per molti utili esempi sul punto v. AI NOw, 24 ss.

58 Per un utile panorama sulle principali iniziative di auto-disciplina v. Council of Europe, *Discrimination, artificial intelligence, and algorithmic decision-making,* a cura di F.Z. Borgesius, cit., 27.

Questi sono certamente strumenti utili e meritevoli di essere incoraggiati, ma che si muovono pur sempre in una logica di autodisciplina, la quale è per propria natura soggetta soltanto a quei vincoli che la cultura degli operatori e le prassi tecnologiche condivise in un dato momento storico possano suggerire[59]. Dato il rango costituzionale delle situazioni incise, è imprescindibile apprestare, prima ancora, un'adeguata infrastruttura istituzionale, composta di norme, rimedi e procedure adeguati agli interessi in gioco e in grado di assicurare un capillare controllo sociale sull'uso dell'algoritmo. Si tratta cioè di arricchire un modello di *digital ethics* con un più penetrante sistema di *digital regulation*[60].

La consapevolezza dei tre profili indicati ha ispirato la formazione di prassi innovative, l'adozione di codici di auto-regolamentazione e la formulazione di indirizzi di *policy* particolarmente impegnativi. Basti citare, dal primo punto di vista, le iniziative della *Algorithmic Justice League* e dell'*AI Now Institute*, che al pari di numerose altre organizzazioni non governative, sono impegnate nel disvelare e contrastare attraverso azioni giudiziarie i fenomeni di uso discriminatorio dell'algoritmo. O si pensi alla campagna "OpenSchufa" recentemente promossa in Germania dalla *Open Knowledge Foundation* e da *AlgorithmWatch* al fine di ottenere l'ostensione del codice sorgente, o comunque la comunicazione dei dettagli operativi, dell'algoritmo utilizzato dalla potentissima società *Schufa* (Schutzgemeinschaft für allgemeine Kreditsicherung), la quale raccoglie dati relativi alla solvibilità finanziaria di 67 milioni di persone e 5 milioni di imprese tedesche e le cui valutazioni negative possono determinare l'impossibilità di accedere al credito, stipulare un contratto di locazione, etc. [61] Dal secondo punto di vista, mette conto ricordare la *Sharing Cities Declaration* adottata a Barcellona nel 2018 e finalizzata a delineare un quadro impegnativo per assicurare condizioni di vita, produzione e sussistenza urbane che siano inclusive, aperte e sostenibili[62]. In essa trovano specifica emersione alcuni dei temi sin qui evocati, come la differenziazione delle piattaforme in collaborative e non collaborative in base, tra l'altro, al grado di inclusione sociale promosso nell'offrire servizi a condizioni identiche a differenti segmenti della popolazione e senza indulgere in discriminazioni (Principio # 1); il principio del contrasto al pregiudizio e alla discriminazione attraverso la predisposizione di condizioni eque e giuste di accesso all'occupazione per persone di qualsivoglia provenienza sociale (Principio # 4); la più ampia garanzia dei diritti digitali, specificamente inclusiva del diritto alla *accountability* algoritmica e alla portabilità dei dati personali (Principio # 4).

59 Sachverständingenrat für Verbraucherfragen (SVRV), *Lösungsoptionen,* cit., 26, 36.
60 Per un'utile tassonomia dei 3 principali modelli di governo della tecnologia, *digital ethics, digital governance* e *digital regulation*, v. L. Floridi, *Soft ethics, the governance of the digital, and the General Data Protection Regulation, Phil. Trans. R. Soc. A* 376:20180081.
61 E. Erdmann, *Schufa, öffne dich, Zeitonline,* 17-3-2018.
62 http://www.share.barcelona/declaration/.

3.2. *I modelli di disciplina emergenti: l'approccio europeo*

Oltre alle prassi e alle dichiarazioni, che richiederebbero anche uno specifico panorama sull'ampio universo del *soft la*w, non può prescindersi – come si notava pocanzi – dal ruolo del diritto in senso formale, quale strumento di mediazione tra le varie istanze sociali emergenti e soprattutto quale tecnica di controllo democratico dei nuovi poteri tecnologici[63].

Il ricorso al diritto in questo campo non è privo di problemi, né si sottrae alle obiezioni di chi abbia il timore di ingessare l'impetuoso sviluppo tecnologico attraverso regole troppo rigide e esposte a rapida obsolescenza. Non a caso esso è osteggiato negli ambienti culturali che ripongono una maggior fiducia nelle virtù di autoregolazione dei mercati, mentre esso è maggiormente incoraggiato nei contesti a più alta propensione regolatoria. Tra questi spicca lo spazio giuridico europeo. È qui che hanno trovato emersione i primi e più compiuti esempi di governo giuridico della decisione algoritmica. Se ne analizzeranno nel prosieguo due, il primo tratto dal diritto dell'Unione Europea e il secondo dall'esperienza francese.

Il primo esempio è costituito dal Regolamento generale per la protezione dei dati personali (GDPR). Esso contiene una disciplina piuttosto avanzata delle salvaguardie da adottare in caso di decisione automatizzata che coinvolga dati personali, la quale si colloca in un'immediata linea di continuità con la previgente direttiva 95/46/CE. Al tema in oggetto sono dedicati – a tacer d'altro - il Considerando n. 71, e due articoli: il 15 e il 22.

L'art. 15 configura una prima, fondamentale, garanzia di fronte a un processo decisionale automatizzato, compresa la profilazione (ai sensi dell'art. 22), che si avvalga di dati personali: il diritto di sapere. La norma, in particolare, stabilisce il diritto di ottenere informazioni circa *"la logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato*".

Per contro l'art. 22 fissa un limite sostanziale all'uso del trattamento algoritmico. Esso stabilisce che "l'interessato ha il *diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato*, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona". Si tratta di un vero e proprio divieto[64], sia pur corredato da una serie di eccezioni delle quali si farà ora cenno, alla cui violazione conseguono effetti preclusivi per il titolare del trattamento. Sottese a tale proibizione

63 Sachverständigenrat für Verbraucherfragen (SVRV), *Lösungsoptionen,* cit., 25 ss.; V. Boehme-Nessler, *Die Macht des Algorithmen und die Ohnmacht des Rechts, NJW,* 2017, 3031.

64 Art. 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679,* last revised February 2018, 19-20; P. Voigt– A. von dem Bussche, *The EU General Data Protection Regulation (GDPR). A Practical Guide,* Cham, 2017, 180.

sono diverse esigenze, tra cui quella di proteggere la dignità umana, evitando che la persona sia resa oggetto passivo di decisioni assunte in forma de-umanizzata, e di assicurare la trasparenza e la controllabilità della decisione stessa.

Tale divieto non opera qualora la decisione: *a*) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento[65]; *b*) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato; *c*) si basi sul consenso esplicito dell'interessato.

Nel caso della conclusione del contratto e del consenso esplicito, il Regolamento obbliga il titolare del trattamento ad attuare misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato. Tra queste assumono particolare rilievo il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione.

Inoltre viene fissato un limite sostanziale invalicabile, costituito dal principio per cui le decisioni algoritmiche autorizzate non possono avvalersi dei dati particolari di cui all'art. 9 (cioè i dati sulla salute, sull'orientamento sessuale, sulle opzioni ideologiche e sindacali, sulle appartenenze etniche, etc.), a meno che non sussistano le scriminanti previste dall'art. 9, par. 2, lett. *a*) (consenso esplicito della persona) o *g*) (trattamento necessario per motivi di interesse pubblico rilevante) e che non siano state adottate misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato.

Tali norme riflettono il lodevole sforzo di elaborare una disciplina che sia trasversale al settore pubblico e al settore privato, trovando applicazione in entrambi ogniqualvolta si sia in presenza di una decisione automatizzata presa a partire da dati personali. Esse soffrono di alcuni limiti intrinseci e di qualche elemento di ambiguità.

Iniziamo da questi ultimi.

Per quanto concerne l'art. 15, esso ha senza dubbio un'importanza non trascurabile quale strumento capace di rispondere almeno in parte all'esigenza indicato in precedenza come problema della "trasparenza" dell'algoritmo[66]. Se non già a livello 'proattivo' (stimolando indirettamente la leggibilità dell'algoritmo in fase di programmazione) [67], quanto meno a livello 'reattivo' esso testimonia della necessità di dotarsi di una chiave di accesso e possibilmente di comprensione della

65 Ad esempio, alcuni studi legali hanno iniziato a far uso di tecniche predittive automatizzate al fine di decidere se accettare il patrocinio nel campo dell'infortunistica (J. Croft, *Legal firms unleash office automatons, Financial Times,* 16 maggio 2016).

66 M. Temme, *Algorithms and Transparency in View of The General Data Protection Regulation,* 3 *Eur. Data Prot. L. Rev.* 473 (2017), 481.

67 G. Malgieri - G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data-Protection Regulation,* 7 *Int'l Data Privacy Law* 243 (2017).

logica funzionale dell'algoritmo, come chiaramente indicato in ambito amministrativo dalla prima decisione del TAR Lazio, che ha riconosciuto il diritto per il privato cittadino di accedere al codice sorgente del software relativo all'algoritmo usato dalla p.a. per gestire le procedure di assegnazione degli insegnanti nelle sedi vacanti[68].

Si discute, tuttavia, se il diritto ad ottenere informazioni di cui all'art. 15 si appunti sulle generali caratteristiche del modello e la logica utilizzata dal *software,* o attenga invece più specificamente al rapporto tra tale logica e i risultati per la sfera del singolo individuo della decisione adottata[69]. Si tratta cioè di un modello di controllo generale circa le conseguenze attese del trattamento, oppure di un canone conoscitivo volto a comprendere *il modo in cui la decisione è stata presa in relazione alla specifica situazione soggettiva e fattuale dell'interessato*?

Se si opta per questa seconda, più estensiva, interpretazione, vi sono due ostacoli da tener presenti. Il primo consiste nel fatto che l'applicabilità della disciplina è condizionata alla circostanza che, usati per fini di decisione, siano *dati personali*, sicché i dati non personali (si pensi ancora ai dati forniti da un'autovettura intelligente) o i dati in forma anonima (molte delle inferenze a carattere predittivo sono basate su dati anonimi, come la residenza di certi gruppi sociali in determinate aree di territorio) ne sono esclusi[70]. Il secondo è rappresentato dal considerando 63, il quale – come la legge francese che verrà di seguito discussa - fa espressamente salve le prerogative della proprietà intellettuale. Ciò significa che se l'algoritmo computerizzato sottende un *software* protetto dal diritto d'autore, o siano coinvolti segreti commerciali la richiesta di accesso potrebbe infrangersi di fronte a un siffatto scoglio e essere neutralizzata dai privilegi dominicali cristallizzati nell'ultima generazione delle regole in materia di IP[71]. Quanto ciò sia importante è dimostrato da alcune controversie statunitensi in materia di voto elettronico, là dove la richiesta di ostensione del codice operativo del *software* è stata rigettata in nome del principio dei *trade secrets*[72]. Ovviamente, qualora si tratti di un algoritmo computerizzato realizzato su commissione per conto di una pubblica amministrazione, potrebbe invocarsi il disposto dell'art. 11, primo comma, della legge sul diritto d'autore, sostenendo l'intervenuto acquisto a titolo originario del monopolio di sfruttamento in capo all'ente committente; per poi desumerne l'inopponibilità al privato dell'argomento tratto dalla tutela della proprietà intellettuale[73]. Il problema, però, rimane quanto meno per il settore privato e per i

68 Tar Lazio, 22-3-2017, n. 3769; 21-3-2017, n. 3742, in *Foro amm.,* .

69 L Edwards – M. Veale, *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?,* cit.

70 Tar Lazio, 22-3-2017, n. 3769.

71 Su questo problema v. M. Temme, *Algorithms and Transparency in View of The General Data Protection Regulation,* cit., 484.

72 D. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure,* 59 *Fla. L. Rev.* 135 (2007)*;* v. anche AI NOw, 11.

73 In questa linea v. Tar Lazio, 22-3-2017, n. 3769, che attribuisce valore assorbente ai principi della trasparenza del procedimento amministrativo, configurando l'algoritmo come un documento amministrativo.

trattamenti algoritmici condotti dalla p.a. sulla base di rapporti contrattuali con effetti obbligatori e non traslativi della titolarità; sarebbe auspicabile, a tal proposito, optare per un'interpretazione restrittiva della clausola di salvaguardia dei diritti di proprietà intellettuale e affermare la prevalenza del diritto d'accesso dell'interessato, in linea peraltro con quanto espresso nei Considerando 34 e 35 della Direttiva 2016/943/UE sulla protezione dei segreti commerciali[74].

Quanto invece all'art. 22, difficoltà derivano:

a) dall'essere il diritto in oggetto limitato all'ipotesi in cui il processo decisionale sia *integralmente* basato sul trattamento automatizzato, il che avviene in un numero limitato di casi, posto che in particolare per le decisioni che interessano il settore pubblico è generalmente previsto un intervento umano (il quale però è spesso fortemente condizionato da una previa valutazione automatizzata della fattispecie);

b) dalla ristrettezza della nozione di "decisione", la quale implica, a tacer d'altro, l'esclusione dall'ambito applicativo della norma di tutte le forme - per quanto invasive - di *microtargeting*[75]. Che si tratti di questioni rilevanti è testimoniato non solo dal problema oggi cruciale del *data-driven political microtargeting*[76], ma anche dai casi di sollecitazione pubblicitaria discriminatoria, come quello in cui si inviavano pubblicità di servizi di assistenza legale in ambito penale soltanto ai soggetti con cognomi che rivelassero l'origine afro-americana della persona[77].

c) Dal requisito dell'"effetto giuridico" come conseguenza di una decisione automatizzata. Si tratta di una limitazione alquanto rilevante dell'ambito applicativo della norma, che copre essenzialmente i casi di decisioni suscettibili di incidere situazioni giuridiche soggettive. Vi rientrano certamente le ipotesi di atti amministrativi particolareggiati, il rifiuto di una domanda di credito presentata online o scelte in materia di assunzione operate in via elettronica (cfr. Cons. 71), ma altre importanti fattispecie ne risulterebbero escluse. Si è già citato il caso del *microtargeting*, che non produce tecnicamente effetti giuridici, ma è atto a condizionare comportamenti sia di mercato sia extra-mercato, come nel caso delle *fake news* indirizzate in maniera mirata a classi di soggetti. L'interrogativo che va sollevato in relazione a tale modello regolamentare è il seguente: ha senso limitare l'impatto rilevante di cui si discorre nella norma alla posizione

74 G. Malgieri, *Trade Secrets v. Personal Data: a possible solution for balancing rights,* 6 *Int'l Data Privacy L.* 102 (2016); G. Malgieri - G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data-Protection Regulation,* cit., 262-264.

75 Martini, sub § 22, in B. Paal – D. Pauly, *Datenschutz-Grundverordnung,* München, 2017, Rn. 23.

76 F.J.Z. Borgesius et al, *Online Political Microtargeting: Promises and Threats for Democracy,* 14 *Utrecht L. Rev.* 82 (2018).

77 L. Sweeney, *Discrimination in Online Ad Delivery,* 11 *Queue* 10 (2013).

del singolo individuo, quando invece è la somma delle micro-violazioni individuali a produrre un effetto lesivo o discriminatorio per l'intero gruppo di riferimento?

Il quesito appena sollevato disvela il primo dei limiti intrinseci dell'approccio regolatorio delineato dal GDPR, consistente nella prevalente logica individualistica attraverso la quale ci si accosta a un tema di rilevanza decisamente meta-individuale e collettiva, quale è quello delle decisioni algoritmiche[78]. Non è detto, infatti, che l'assenza di lesione individualmente rilevante ai sensi della normativa sulla protezione dei dati privi la fattispecie dei caratteri di disvalore, poiché ad esempio il suddetto trattamento potrebbe produrre effetti pregiudizievoli o discriminatori per lo specifico gruppo al quale il cittadino appartenga.

Tale considerazione induce da un lato a ricordare che la normativa sulla tutela dei dati deve essere intesa come un tassello, certo al momento il più avanzato, di un più ampio mosaico regolatorio, al quale dovranno contribuire gli altri segmenti dell'ordinamento, e in primo luogo il diritto antidiscriminatorio (come delineato a partito dalle direttive 2000/43 CE, sull'uguaglianza razziale, 2000/78/CE sulla parità di trattamento in materia di lavoro, 2006/54/CE sull'uguaglianza di genere) [79], il diritto dei consumatori, il diritto amministrativo e il diritto del lavoro[80]. Dall'altro essa spinge ad affermare che anche gli strumenti di tutela, finalizzati ad assicurare un controllo esterno sulle decisioni algoritmiche, dovrebbero essere improntati ad una logica di azione *collettiva* piuttosto che individuale. È vero che il primo comma dell'art. 80 GDPR prevede la possibilità di conferire mandato ad enti del terzo settore, ma il secondo comma rimette agli stati membri la scelta discrezionale se adottare il modello dell'*opt out,* ossia dell'azione promossa direttamente dagli enti non profit, salvo il diritto di opporsi da parte del singolo[81]. Tale forma di azione collettiva sembrerebbe l'unica in grado di apprestare una tutela efficace, assieme ovviamente all'iniziativa delle autorità amministrative indipendenti competenti nel settore, che però scontano in molti casi un ritardo di mezzi e organizzazione tecnologica.

Inoltre andrebbe incoraggiato il ricorso a tecniche di controllo *ex-ante* che contribuiscano ad orientare le modalità di impiego dell'algoritmo[82]. Il GDPR contiene a tal proposito indicazioni

78 L Edwards – M. Veale, *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?,* cit.

79 Per una specifica applicazione del diritto antidiscriminatorio al caso dei *selective advertisements* in materia di lavoro, v. D.J. Dalenberg, *Preventing discrimination in the automated targeting of job advertisements,* 34 *Computer Law & Security Rev.* 615 (2108).

80 In quest'ottica v. A. Mantelero, *AI and Big Data: A Blueprint for a human rights, social and ethical impact assessment, Computer Law & Security Rev.* 34 (2018), 754.

81 In tema v. F. Casarosa, *La tutela aggregata dei dati personali nel Regolamento UE 2016/679: una base per l'introduzione di rimedi collettivi?,* in A. Mantelero – D. Poletti, a cura di, *Regolare la tecnologia: il Regolamento UE 2016/679 e la protezione dei dati personali,* Pisa, 2018, 235 ss.

82 L Edwards – M. Veale, *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?,* cit.

importanti, che potrebbero essere valorizzate nella prassi al fine di inglobare nel sistema normativo in oggetto l'ulteriore istanza di tutela contro gli effetti discriminatori dell'algoritmo. Tra queste meritano di essere ricordate la progettazione preventiva delle macchine in maniera 'privacy-enhancing' (art. 25); la valutazione di impatto da redigere qualora il trattamento presenti un rischio elevato per i diritti e le libertà delle persone fisiche (art. 35); le certificazioni adottate ai sensi dell'art. 42. Di questa attitudine espansiva della disciplina in materia di protezione dei dati v'è una traccia precisa nello stesso GDPR, il cui Considerando 71 recita al secondo comma:

"Al fine di garantire un trattamento corretto e trasparente nel rispetto dell'interessato, tenendo in considerazione le circostanze e il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato *e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o delle origini etniche, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportino misure aventi tali effetti"*.

3.3. *Algoritmi e decisioni amministrative: la riforma francese*

Le regole contenute nel GDPR hanno il pregio di guardare trasversalmente al fenomeno del trattamento automatizzato dei dati personali, al fine di prefissare garanzie minime per i diritti della persona le quali siano applicabili sia ai trattamenti nel settore pubblico sia nel settore privato. L'attitudine onnicomprensiva della normativa in oggetto ne costituisce però anche un limite, in quanto il ricorso all'algoritmo nel settore pubblico, ed in particolare nell'ambito delle procedure preordinate all'emanazione di un provvedimento amministrativo, ha indubbie particolarità di ordine funzionale e strutturale, che richiedono una disciplina più specifica e puntuale rispetto a quella del settore privato (il quale pure richiederebbe di essere disarticolato in più microsettori, come il credito, la sanità, il lavoro, etc.) [83]. Si comprende, quindi, che negli ordinamenti europei per tradizione più sensibili all'interazione tra tecnologie e diritti, come la Francia, il recepimento del GDPR abbia offerto il destro per introdurre, oltre a norme di dettaglio sulle decisioni automatizzate contemplate dal Regolamento, apposite regole sui provvedimenti amministrativi algoritmici.

83 H. Pauliat, *La décision administrative et les algorithmes: une loyauté à consacrer, Rev. Dr. Pub.,* 2018, 641 ss.

La legge n. 2018-493 del 20 giugno 2018 ha modificato l'art. 10 della celebre legge *Informatique et libertés* del 1978, prefissando i seguenti principi:

a) nessuna decisione giudiziaria che implichi la valutazione del comportamento di una persona può fondarsi su un trattamento automatizzato di dati personali preordinato a giudicare aspetti della personalità di tale persona [se ne desume che un sistema del tipo COMPAS non potrebbe trovare accoglimento nell'ordinamento francese];

b) nessuna decisione che produca effetti giuridici su una persona e basata sul trattamento automatizzato dei dati può essere assunta al di fuori delle condizioni stabilite dall'art. 22 del GDPR;

c) provvedimenti amministrativi individualizzati che si basino su un trattamento automatizzato sono ammissibili, purché rispettino le condizioni previste dal *Code des relations entre le public et l'administration* e purché il trattamento non coinvolga categorie particolari di dati (dati sulla salute, sulle preferenze politiche, dati sulla configurazione genetica, etc.).

d) tali provvedimenti devono contenere espressa menzione del trattamento automatizzato; relativamente ad essi, il responsabile del trattamento deve assicurare la piena comprensione del trattamento algoritmico e delle sue evoluzioni affinché possa spiegare alla persona interessata, nel dettaglio e in una forma intellegibile, il modo in cui il trattamento sia stato posto in opera nei suoi confronti.

Se questa disciplina è contenuta nella legge di recepimento del GDPR, già la legge sulla *République Numérique* e un successivo decreto attuativo avevano prefissato le condizioni di ammissibilità dei provvedimenti amministrativi algoritmici[84].

L'art. L 311-3-1 prevede che "una decisione individuale presa sul fondamento di un trattamento algoritmico comporta una menzione esplicita e l'informazione all'interessato. Le regole che definiscono tale trattamento, come pure le principali caratteristiche della sua messa in opera sono comunicate dall'amministrazione all'interessato che ne faccia domanda".

La norma in esame rinviava a un decreto attuativo per la definizione delle specifiche applicative.

Tale decreto, approvato il 14 marzo 2017, ha stabilito, all'art. R 311-3-1-1, che "la menzione esplicita prevista dall'art. L. 311-3-1 indica la finalità perseguita attraverso il trattamento algoritmico. Essa richiama il diritto, garantito dal suddetto articolo, di ottenere la comunicazione delle regole che definiscono tale trattamento e delle principali caratteristiche della sua messa in atto, nonché delle modalità di esercizio di tale diritto alla comunicazione e di ricorso, ove ne sussistano

84 J.-B. Duclercq, *Le droit public à l'ère des algorithmes, Rev. Dr. Pub.,* 2017, 1401 ss.

le condizioni, alla commissione di accesso ai documenti amministrativi, come definita dal presente libro". La norma successiva, l'art. R 311-3-1-2, dà specifica concretezza al diritto dell'interessato di essere informato circa il modo in cui la logica generale dell'algoritmo è stata applicata alla sua condizione particolare, comprendendo così il modo e le forme in cui essa ha inciso sui risultati della decisione. "L'amministrazione comunica alla persona destinataria di un provvedimento preso sul fondamento di un trattamento algoritmico, su istanza di parte, in forma intellegibile e a condizione di non violare segreti protetti dalla legge, le seguenti informazioni:

- il grado e il *modo in cui il trattamento algoritmico ha contribuito* alla decisione;
- i *dati trattati* e la loro origine;
- i parametri del trattamento e, se del caso, la loro ponderazione, *applicati alla situazione dell'interessato;*
- le operazioni effettuate attraverso il trattamento".

Si tratta di principi particolarmente innovativi e rilevanti, poiché da un lato prefissano una serie di garanzie procedurali e sostanziali (prima tra tutte l'impossibilità di far uso di dati sensibili nel trattamento algoritmico) che elevano la tutela della persona-cittadino rispetto alle decisioni automatizzate, ma dall'altro ammettono espressamente la validità di un provvedimento amministrativo, che incida sulla situazione soggettiva del singolo, assunto sulla base di un trattamento automatizzato di dati. Si comprende, quindi, che tale disciplina abbia sollevato anche obiezioni da parte di chi ravvisi, nella parziale delocalizzazione dello spazio deliberativo agli algoritmi computerizzati, un *vulnus* ai principi che governano il procedimento amministrativo[85].

Anche nel nostro ordinamento, peraltro, il problema era già emerso ed aveva trovato una prima valutazione giudiziaria da parte del TAR Lazio, che con pronunzia 10 settembre 2018, ha annullato i provvedimenti del Ministero dell'Istruzione conclusivi delle procedure di mobilità straordinaria degli insegnanti, in quanto assunti invece che con ordinaria istruttoria procedimentale, con valutazione demandata ad apposito algoritmo[86]. Ha osservato il collegio, in particolare, che "gli istituti di partecipazione, di trasparenza e di accesso, in sintesi di relazione del privato con i pubblici poteri non possono essere legittimamente mortificate e compresse soppiantando l'attività umana con quella impersonale, che poi non è attività, ossia prodotto delle azioni dell'uomo, che può essere svolta in applicazione di regole o procedure informatiche o matematiche. Ad essere inoltre vulnerato non è solo il canone di trasparenza e di partecipazione procedimentale, ma anche l'obbligo di motivazione delle decisioni amministrative, con il risultato di una frustrazione anche

85 In generale v. il panorama offerto da H. Pauliat, *La décision administrative et les algorithmes: une loyauté à consacrer,* cit.

86 Tar Lazio, 10-9-2018, n. 9227. In tema v. L. Viola, *L'intelligenza artificiale nel procedimento e nel processo amministrativo: lo stato dell'arte,* in *Foro amm.,* 2018, 9, 1598 ss.; P. Otranto, *Decisione amministrativa e digitalizzazione della p.a.,* in www.federalismi.it, 2018, 2, 15.

delle correlate garanzie processuali che declinano sul versante del diritto di azione e difesa in giudizio di cui all'art. 24 Cost., diritto che risulta compromesso tutte le volte in cui l'assenza della motivazione non permette inizialmente all'interessato e successivamente, su impulso di questi, al Giudice, di percepire l'iter logico-giuridico seguito dall'amministrazione per giungere ad un determinato approdo provvedimentale". Si tratta probabilmente, di un approccio eccessivamente rigido, che portato alle estreme conseguenze finirebbe per precludere l'utilizzo di algoritmi di apprendimento automatico nell'ambito dell'azione amministrativa.

Che ciò non sia un esito desiderabile, né necessitato sul piano dell'assetto giuridico vigente, è dimostrato da una recente pronunzia del *Conseil Constitutionnel* francese. La Corte nel 2018 è stata chiamata a giudicare della costituzionalità della norma interna di adeguamento al GDPR, precedentemente discussa[87]. Tra i vari profili di illegittimità denunziati, vi era anche quello del contrasto con i principi di legittimità dell'azione amministrativa, derivanti dall'ammissione delle decisioni individualizzate assunte in base a trattamento algoritmico; in particolare, i ricorrenti sostenevano che il ricorso ad algoritmi privasse l'amministrazione nel necessario di potere di apprezzamento delle situazioni individuali, e che in particolare gli algoritmi di apprendimento automatico, implicando una continua revisione delle regole di funzionamento dell'algoritmo stesso, precludesse alla stessa amministrazione la facoltà di comprendere la logica motivazionale sottesa alla decisione. La Corte ha rigettato tutte le censure, avanzando i seguenti argomenti[88]:

a) non vi è una delega del potere regolamentare allo strumento tecnologico, per ciò che i criteri e le modalità di funzionamento dell'algoritmo sono stabiliti ex ante e validati dal responsabile del procedimento;
b) il ricorso al trattamento algoritmico è subordinato alle specifiche condizioni e garanzie previste dal suddetto decreto del 2017;
c) contro il provvedimento individuale assunto sulla base di trattamento algoritmico è comunque sempre concesso ricorso amministrativo, che richiede una decisione evidentemente non basata soltanto sull'algoritmo, e se ne ricorrono le condizioni, ricorso al giudice;
d) il responsabile del trattamento deve essere sempre in grado di comprendere il funzionamento del trattamento algoritmico e le sue evoluzioni, in modo da poter spiegare alla persona interessata, nel dettaglio e in una forma intellegibile, il modo in cui il trattamento è stato posso in essere nei suoi riguardi;

87 Cons. const., déc. 12-6-2018, n. 2018-765.
88 Per alcune considerazioni in merito alla pronunzia v. E. Rulli, *Giustizia predittiva, intelligenza artificiale e modelli probabilistici. Chi ha paura degli algoritmi?,* in *Analisi giuridica dell'economia,* 2, 2018, 533 ss., 540 ss.

e) per quanto detto, devono ritenersi preclusi all'amministrazione quegli algoritmi di apprendimento automatico il cui funzionamento sfugge alla comprensione del responsabile del procedimento, ma non tutti gli algoritmi che possano semplificare e rendere più precisa e neutrale l'azione amministrativa.

4. *Gli algoritmi come opportunità e la politica (del diritto)*

Le ultime norme citate suggeriscono quale dovrebbe essere, ad avviso di chi scrive, l'approccio tecnicamente corretto al tema dei *big data* e dell'intelligenza artificiale, intorno a cui stiamo costruendo le basi della nostra convivenza futura. Non ci si dovrebbe ispirare a un *laissez-faire* tecnologico, né a un luddismo di retroguardia.

È invece necessario operare, in tutte le sedi, perché i processi in atto, i quali sono destinati a regolare segmenti crescenti della vita sociale dell'uomo, siano sottoposti a una logica di controllo democratico, che assicuri un adeguato bilanciamento tra la 'funzionalità tecnologica' e la desiderabilità sociale degli scopi perseguiti, e rispetto alla quale la mediazione giuridica svolge un ruolo centrale.

Si deve cioè lavorare all'adozione di strumenti regolatori e di governo, preordinati ad evitare che la saldatura tra potere economico e potere tecnologico produca una società della sorveglianza e della discriminazione, in cui tutti siano profilati, segmentati in gruppi e resi destinatari di effetti giuridici o sociali in funzione dell'assetto di potere esistente. Non può, infatti, ignorarsi il pericolo che, come si è bene osservato, i nuovi modelli algoritmici creino le condizioni per un nuovo medioevo digitale. Si delinea cioè il rischio "di una società connotata da una segmentazione per caste, ove lo *status* non è però dato dalla nascita o dall'appartenenza a classificazioni sociali tradizionali (quelle su cui vigilano le norme in materia di non-discriminazione), ma da algoritmi e dai valori di coloro li generano. Classificazioni che sono poi impiegate per prendere decisioni che coinvolgono una pluralità di soggetti, i quali però non hanno contezza della propria posizione" [89].

Se ciò implica rigettare tanto un modello sregolato di "capitalismo della sorveglianza", il quale finirebbe per inchiodare la società alle sue iniquità e ai suoi pregiudizi, codificandoli nel linguaggio informatico[90], quanto un sistema programmato di "sorveglianza di stato", come nel già descritto caso del "social credit system" cinese[91], è necessario stabilire quale tecnica di intervento

89 A. Mantelero, *La gestione del rischio nel GDPR: limiti e sfide nel contesto dei Big Data e delle applicazioni di* Artificial Intelligence, in A. Mantelero – D. Poletti, a cura di, *Regolare la tecnologia: il Regolamento UE 2016/679 e la protezione dei dati personali,* cit., 289 ss., 302.
90 S. Barocas – A.D. Selbst, *Big Data's Disparate Impact,* cit., 671 ss.
91 V. supra, par. 3.

sia più adeguata al controllo dei trattamenti algoritmici. La logica del divieto *tout court* non sembra percorribile, per il semplice fatto che l'innovazione tecnologica, se attentamente monitorata, può apportare immensi benefici sociali, creando le premesse per una società più aperta ed inclusiva. Del pari, sembra utopistico pensare a un meccanismo di co-decisione pubblico-privato o a un'approvazione preventiva da parte di appositi enti pubblici degli algoritmi utilizzabili anche da soggetti privati. La strada più proficua appare quella dell'intervento a geometria variabile, composto cioè da forme più *soft* di incentivazione all'adozione di tecnologie e prassi organizzative *human rights compliant* (come nella logica dell'*accountability* prescritta dal GDPR per assicurare un trattamento conforme a parametri di sicurezza da valutare caso per caso e su iniziativa del singolo titolare del trattamento, oppure del ricorso a certificazioni e marchi che attestino l'uso di *transparency enhancing technologies*) e strumenti più incisivi con funzione prettamente regolamentare (come nel caso dei limiti sostanziali alle decisioni automatizzate posti dall'art. 22 GDPR), i quali dovrebbero poi ricadere a cascata sulla fase della programmazione dell'algoritmo, incentivando in ultimo una sorta di *legality by design*) [92]. Le regole emergenti in ambito in europeo, pur perfettibili da molti punti di vista, si ispirano ad una logica siffatta. Converrà, in conclusione, ribadire sinteticamente alcuni dei tratti caratterizzanti di questo modello.

Innanzitutto è caratteristico dell'approccio europeo sottoporre a un'attenta disciplina l'ecosistema informativo che sta a monte del funzionamento degli algoritmi, fissando alcuni requisiti di qualità e quantità (esattezza, accuratezza, minimizzazione dei dati, etc.) dei dati destinati a rappresentare l'input dei processi di *machine learning*. Le prime e più efficaci garanzie partono proprio dalla 'giuridificazione' dei fenomeni di trattamento dei dati personali e dal controllo sulla profilazione individuale (par. 2.1.). Coerente, peraltro, con questa prospettiva è il divieto di sottoporre a trattamento algoritmico i dati che attengono al nucleo duro della *privacy,* e cioè i dati sulla salute, sull'orientamento politico, sulle fedi religiose, sull'origine etnica, etc. Per altro verso per sfruttare al massimo la capacità innovativa delle tecniche di *machine learning,* andrebbe tenuto fermo un principio di libera utilizzazione dei dati non personali, resistendo con forza ai tentativi di introdurre surrettiziamente diritti di proprietà sui dati grezzi (par. 2.2.).

In secondo luogo si attribuisce grande rilevanza al principio della trasparenza, il quale viene declinato in forme diverse e spesso convergenti: diritto di essere informato *ex ante* circa l'esistenza di un trattamento automatizzato; diritto di conoscere la logica di cui tale trattamento si avvale; diritto di comprendere il modo e il grado in cui il trattamento algoritmico ha influito sui risultati della decisione che coinvolga la sfera del singolo; diritto di accedere all'algoritmo in quanto parte

92 In quest'ottica v. il ricco contributo di J.A. Kroll – J. Huey – S. Barocas et al., *Accountable Algorithms,* 165 *U. Pa. L. Rev.* 633 (2017); G. Malgieri - G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data-Protection Regulation*, cit.; M. Hildebrandt, *Algorithmic regulation and the rule of law, Phil. Trans. R. Soc. A* 376:20170355.

integrante di un procedimento amministrativo. Tutto ciò può contribuire in maniera rilevante alla 'leggibilità' e alla *accountability* degli algoritmi stessi, specie là dove si opti per un'interpretazione restrittiva delle clausole di salvaguardia poste a protezione della proprietà intellettuale e dei segreti commerciali in caso di richieste volte a conoscere la logica sottesa ai trattamenti algoritmici (par. 3.2.).

In terzo luogo, si va affermando un principio di massima, per cui l'individuo non deve essere sottoposto a una decisione integralmente automatizzata, allorché questa incida in maniera significativa sulla propria sfera giuridica. Quando ciò venga ammesso dall'ordinamento, si applicano una serie di garanzie sostanziali, tra le quali risalta lo strumento del riesame della decisione attraverso un diretto coinvolgimento dell'uomo, oltre ovviamente al controllo giudiziario *ex post* sulla conformità del trattamento algoritmico ai requisiti di legge (par. 3.3.).

Perché tali rimedi acquistino maggiore efficacia, appare necessario fare un ulteriore passo in avanti e, oltre ad estendere il perimetro di applicabilità delle norme in oggetto, valorizzare la dimensione collettiva del controllo sugli algoritmi. Ciò significa non soltanto sottolineare, come si è già fatto, che l'impatto potenzialmente distorsivo degli algoritmi sul piano richiede il concorso di diversi sotto-settori dell'ordinamento, e in primo luogo del diritto anti-discriminatorio, che va declinato in funzione del nuovo ecosistema digitale. Significa soprattutto potenziare il quadro dei rimedi di natura collettiva (come le azioni collettive promosse da enti *non profit* anche senza mandato preventivo da parte dei singoli) e affiancare alle tecniche di controllo individuale le iniziative del potere pubblico, valutando ad esempio l'introduzione di un'apposita autorità amministrativa indipendente che assommi le competenze frammentate tra le altre autorità per il settore del digitale[93].

93 In questo senso è indirizzata la proposta del Sachverständigenrat für Verbraucherfragen (SVRV), *Lösungsoptionen,* cit., 38 ss.