

L'UTILIZZO DEI *BIG DATA* E DELL'IA PER UNA MIGLIORE QUALITÀ DELLA REGOLAMENTAZIONE*

FEDERICO GALLI , GIOVANNI SARTOR *****

Sommario

1. Introduzione: l'obiettivo della *better regulation*. – 2. Un nuovo paradigma per la *better regulation*: i *big data* e l'IA. – 3. Raccolta dei dati: costi e limiti. – 4. L'analisi e l'uso dei dati attraverso l'IA per la pianificazione delle leggi. – 5. Dati e modelli per il monitoraggio e la valutazione della regolazione. – 6. La necessità di un'etica dei dati per il settore pubblico. – 7. Conclusione.

Abstract

This paper analyses the prospects of using big data and artificial intelligence (AI) technologies to improve legislation (better regulation), in particular by means of impact assessment. After a brief background on the topic of better regulation, the new paradigm based on big data and AI will be investigated. The three phases of the data-centred regulation cycle will be tackled in succession, namely, (1) data collection and creation, (2) ex-ante impact analysis for policy planning and regulation design, and (3) ex-post monitoring and evaluation. Finally, the challenges for an acceptable use of AI and big data will be explored, focusing on the need to conform public action to a legal-ethical framework for data management and use.

Suggerimento di citazione

F. GALLI, G. SARTOR, *L'utilizzo dei big data e dell'IA per una migliore qualità della regolamentazione*, in *Osservatorio sulle fonti*, n. 3/2022. Disponibile in: <http://www.osservatoriosullefonti.it>

* Il presente contributo costituisce la rielaborazione della relazione tenuta in occasione della giornata di studi *Tecnica legislativa ed innovazione tecnologica* svoltasi a Firenze il 12 maggio 2022 e organizzata nell'ambito del progetto *Assemblee rappresentative ed innovazione tecnologica dopo la pandemia: la "reingegnerizzazione" delle procedure parlamentari e del drafting legislativo* (Legitech).

** Ricercatore a t.d. di Filosofia del diritto nell'Università di Bologna.

Contatto: federico.galli7@unibo.it

*** Professore ordinario di Filosofia del diritto nell'Università di Bologna e Professore part-time all'Istituto Universitario Europeo.

Contatto: giovanni.sartor@unibo.it

1. Introduzione: l'obiettivo della *better regulation*

Dopo il movimento di privatizzazione e liberalizzazione avvenuto negli ultimi decenni del secolo scorso e il conseguente dibattito su regolamentazione e deregolamentazione, l'attenzione si è spostata sulla qualità della regolamentazione. Il risultato pratico di questa attenzione sono stati numerosi programmi improntati alla cd. "*better regulation*" che hanno iniziato a essere adottati da organizzazioni internazionali (come l'OECD) e da molti Paesi sviluppati.

Ad esempio, l'UE ha introdotto un'agenda per una migliore regolamentazione nei primi anni 2000¹ La Comunicazione del 2002 intitolata "*La governance Europea - Un libro bianco*"² delineava quelli che sarebbero diventati i tre pilastri della strategia europea per la *better regulation*:

1. l'elaborazione di un processo standardizzato per la consultazione delle parti interessate esterne e del pubblico durante l'intero sviluppo delle proposte normative della Commissione;
2. l'introduzione di una procedura obbligatoria di valutazione d'impatto *ex ante* durante lo sviluppo di tutte le principali iniziative normative dell'UE;
3. la promozione dell'uso sistematico di valutazioni *ex post* per la revisione delle normative esistenti.

Con il pacchetto "Legiferare meglio" del 2015³, ulteriormente perfezionato nel 2017⁴, la Commissione ha rinnovato il suo impegno politico nei confronti dei medesimi obiettivi, introducendo al contempo modifiche significative al quadro normativo per la *better regulation*. Il pacchetto comprende nuovi orientamenti per la conduzione di consultazioni, valutazioni d'impatto *ex ante* ed *ex post*: tra le altre

¹ Sull'agenda europea per la "*better regulation*" e un commento critico si veda *passim*, GARBEN, GOVAERE, *The EU Better Regulation Agenda: A Critical Assessment*, Bloomsbury Publishing, New York, 2018. In dottrina italiana, si segnala uno dei primi contributi ricognitivi, CARBONE, TIBERI, 'La "*Better Regulation*" in Ambito Comunitario' in *Quaderni Costituzionali*, 2001, fasc. 3, p. 699.

² European Commission (2002), *European governance - A white paper*, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52001DC0428> (ultimo accesso Settembre 2022).

³ European Commission (2015), *Better regulation for better results - An EU agenda*, http://ec.europa.eu/smart-regulation/better_regulation/documents/com_2015_215_en.pdf (ultimo accesso Settembre 2022).

⁴ European Commission (2017), *Completing the Better Regulation Agenda: Better solutions for better results*, https://ec.europa.eu/info/sites/info/files/completing-the-better-regulation-agenda-better-solutions-for-better-results_en.pdf (ultimo accesso Settembre 2022).

modifiche, le riforme del meccanismo di supervisione della Commissione per l'esame delle valutazioni d'impatto *ex ante* ed *ex post*⁵.

Il concetto di *better regulation* incarna uno sforzo normativo che è contemporaneamente un obiettivo e un processo.

Avere una "migliore regolamentazione" significa innanzitutto essere in grado di rispondere ad una pluralità di sfide. In primo luogo, le proposte legislative dovrebbero migliorare nella formulazione linguistica, essere più semplici e chiare, così da favorire la certezza del diritto. In secondo luogo, le misure adottate dovrebbero essere adatte allo scopo che si prefiggono, nel senso che dovrebbero essere adeguate e proporzionate agli obiettivi di politica pubblica perseguiti, senza creare oneri inutili per i cittadini e le imprese⁶. Infine, la regolamentazione dovrebbe essere a prova di futuro (*future-proof*): da una parte, dovrebbe essere flessibile in modo da potersi adattare alle nuove realtà e alle evoluzioni sociali e tecnologiche. Dall'altra, dovrebbe essere guidata dall'innovazione, promuovendo la trasmissione della conoscenza ed evitando la creazione di ostacoli al progresso tecnologico e sociale.

Per rispondere a tali sfide è necessario attuare un processo normativo più efficiente ed efficace. Tale processo non è più visto come una sequenza lineare di fasi indipendenti, ma piuttosto come un ciclo di passaggi interconnessi che si alimentano a vicenda rafforzandosi. Questi passaggi contribuiscono ad un ciclo normativo virtuoso che comprende sei fasi⁷: (1) la pianificazione della regolazione, (2) la redazione legislativa, (3) l'adozione delle misure, (4) l'attuazione, (5) il monitoraggio e (6) la revisione. Elemento fondamentale in tutto il ciclo di regolazione è l'informazione: ogni decisione viene adottata sulla base di evidenze relative agli impatti previsti e osservati (in inglese si parla di "*fact-*" or "*evidence-driven*").

Per produrre queste informazioni, il processo normativo utilizza lo strumento della valutazione d'impatto in due momenti: all'inizio del ciclo, *ex ante*, fornendo previsioni sugli impatti attesi, e alla fine del ciclo, *ex post*, informando sui risultati ottenuti. L'esercizio di valutazione d'impatto *ex ante* supporta le prime tre fasi del

⁵ Sulle più recenti iniziative europee in tema di *better regulation*, si rinvia al rapporto RENDA, *Assessment of Current Initiatives of the European Commission on Better Regulation*, Policy Department for Citizens' Rights; Constitutional Affairs Directorate-General for Internal Policies, 2022.

⁶ In un rapporto commissionato dal Governo inglese, si parla di "legislazione SMART", e cioè *Specific* (specific), *Measurable* (misurabile), *Achievable* (raggiungibile), *Realistic* (realistica) and *Time-limited* (limitata nel tempo). Si veda, HURST, *The Green Book: Central Government Guidance on Appraisal and Evaluation*, HM Treasury, 2020 <www.gov.uk/official-documents>).

⁷ Commissione Europea, *Better Regulation Guidelines*, Documento di lavoro dei servizi della Commissione, 2021.

ciclo normativo (cioè, la pianificazione, l'elaborazione legislativa e l'attuazione) e fornisce ai decisori informazioni sugli impatti economici e non economici delle loro decisioni, consentendo un'analisi di vantaggi e svantaggi di ogni misura presa in considerazione. La valutazione d'impatto *ex post* intende chiarire in che misura l'intervento ha raggiunto i suoi obiettivi, se ha generato impatti non voluti o indiretti, e se sia necessario modificarlo. Nel linguaggio tecnico, si parla spesso anche di "*fitness check*", cioè quel processo di valutazione successiva che fornisce informazioni su come uno o più atti legislativi correlati hanno contribuito a un obiettivo politico stabilito in anticipo.

2. Un nuovo paradigma per la *better regulation*: i big data e l'IA

Lo sforzo di seguire il processo legislativo con la produzione di informazioni introduce un requisito fondamentale: la raccolta di dati e l'uso di strumenti per la loro analisi al fine di produrre conoscenza utilizzabile nel ciclo della regolazione.

Gli strumenti tradizionali adottati dai governi si basano su metodi qualitativi o quantitativi, o su una combinazione di essi. Ad esempio, i *focus group* o le interviste individuali possono essere utilizzati per esplorare fenomeni sociali poco conosciuti, raccogliendo esperienze, opinioni e motivazioni pertinenti, che possono poi guidare la raccolta dei dati quantitativi. Una soluzione pratica può essere quella di condurre o commissionare casi di studio e ricerche approfondite sugli obiettivi della regolazione. Altrimenti, un metodo classico è la consultazione di esperti per la raccolta o l'interpretazione delle evidenze scientifiche. A livello UE, esistono poi vari database che raccolgono dati empirici riguardanti varie tematiche. Ad esempio, Eurostat fornisce libero accesso alle statistiche a livello europeo (a partire dai dati raccolti dagli enti di rilevazione statistica degli Stati membri) utilizzando metodologie armonizzate che consentono il confronto tra Paesi e regioni. Il database Eurobarometro monitora l'opinione pubblica negli Stati membri e fornisce risultati rappresentativi delle popolazioni interessate sui temi quali, ad es., la salute, la cultura, l'ambiente, le tecnologie dell'informazione, la difesa, ecc.

Per quanto riguarda i metodi di analisi, l'approccio tradizionale è caratterizzato dalla quasi esclusiva interpretazione umana dei dati. La valutazione e la validazione critica, indipendente, competente, e trasparente delle prove raccolte garantiscono la solidità e l'affidabilità dell'analisi. La peer-review è un metodo comune di garanzia della qualità che aumenta la credibilità dei risultati e può contribuire a migliorare i modelli e i metodi.

I limiti nella raccolta e nel reperimento di dati (in qualità e quantità) e negli strumenti di analisi automatica sono state una delle principali barriere allo sviluppo

effettivo di pratiche di *better regulation* da parte dei governi. Tuttavia, oggi il contesto sociotecnico è mutato.

Come noto, negli ultimi anni abbiamo assistito a un vero e proprio “diluvio di dati”⁸. La quantità di dati disponibili è aumentata a un ritmo accelerato, poiché la vita umana si è trasferita in rete e la maggior parte dei processi economici, amministrativi e sociali sono mediati da sistemi informatici⁹. Tali sistemi permettono di acquisire e registrare una grande quantità di informazioni sul mondo fisico (es., telecamere stradali, sensori che raccolgono informazioni sul clima, dispositivi per test medici) e sulle interazioni socioeconomiche (es. sistemi di registrazione delle transazioni o di tracciamento del comportamento online). Il termine “big data”¹⁰ identifica grandi insiemi di dati che è difficile gestire con tecniche standard, a causa delle loro caratteristiche particolari, le cosiddette tre “V”: enorme volume, alta velocità e grande varietà.

Il fenomeno dei big data è stato accompagnato dallo sviluppo di tecnologie per l’archiviazione e l’elaborazione dei dati che consentono un uso efficiente di queste vaste risorse. In particolare, grazie alle tecnologie di analisi dei dati, questi vasti insiemi di dati possono essere utilizzati per estrarre informazioni. L’analisi dei dati (*data analytics*) è definita come “l’uso estensivo dei dati, dell’analisi statistica e quantitativa, dei modelli esplicativi e predittivi e della gestione basata sui fatti per guidare le decisioni e le azioni”¹¹ e include i seguenti approcci:

- L’analisi descrittiva (detta anche *business intelligence* o *performance reporting*) consente di accedere a dati storici e attuali. Offre la possibilità di informare, esplorare e creare report utilizzando dati interni ed esterni provenienti da diverse fonti.
- L’analisi predittiva utilizza tecniche quantitative (ad es., propensione, segmentazione, analisi di rete e previsioni econometriche) e tecnologie (come modelli e sistemi basati su regole) per prevedere il futuro sulla base di dati passati.
- L’analisi prescrittiva utilizza una serie di tecniche quantitative e tecnologie (ad esempio, modelli decisionali o motori di raccomandazione) per identificare i comportamenti e le azioni maggiormente vantaggiosi.

⁸ ANDERSON, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, in *Wired magazine*, 2008, p.16.

⁹ VARIAN, *Computer Mediated Transactions*, in *American Economic Review*, Vol. 100, 2010, p. 1.

¹⁰ MAYER-SCHÖNBERGER, CUKIER, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* Houghton Mifflin Harcourt, Boston, 2013.

¹¹ DAVENPORT, HARRIS, *Competing on Analytics: Updated, with a New Introduction: The New Science of Winning*, Harvard Business Press, Harvard, 2017.

Oltre ad essere analizzati per ricavare informazioni, gli insiemi di big data possono anche essere utilizzati per addestrare i sistemi di IA. In tal caso, la conoscenza nascosta nei dati sarà utilizzabile nello svolgimento di compiti in modo più o meno automatico. Ciò è possibile grazie alle tecniche di apprendimento automatico, che insegnano ai computer a costruire modelli descrittivi, predittivi o prescrittivi direttamente dai dati e ad applicarli a casi specifici, per acquisire nuove informazioni o decidere come risolvere un problema¹². A tali tecniche, si aggiungano le più recenti applicazioni nel campo dell'elaborazione automatica linguaggio naturale che permettono di estrarre, sintetizzare ed elaborare informazioni da testi redatti da essere umani.

Negli ultimi anni molte organizzazioni private hanno iniziato a sfruttare questi nuovi strumenti attraverso un cambio strategico nel condurre attività d'impresa¹³. Oggi, nel modello aziendale competitivo, le decisioni sono adottate, non più solamente sulla base di modelli decisionali teorici, bensì sulla base di evidenze empiriche (si parla di "data-driven"). L'impresa orienta prima di tutto la sua attività verso l'accumulo di dati su clienti e concorrenti, e poi verso l'adozione di tecnologie di analisi che permettono di trarre profitto da questi dati¹⁴. Si pensi alle piattaforme online. Queste sono caratterizzate da un doppio flusso di informazioni: gli individui ricevono informazioni e servizi e, allo stesso tempo, i fornitori osservano, verificano e analizzano automaticamente tutte le transazioni, utilizzando ogni carattere digitato o link cliccato. Una volta raccolte e analizzate, queste informazioni permettono di inviare agli utenti materiale mirato (es., pubblicità, notizie, offerte) e offrire servizi personalizzati (es., raccomandazioni, prodotti customizzati) contribuendo ad aumentare la fedeltà dei clienti, il traffico sulla piattaforma e l'incremento dei profitti derivanti da pubblicità.

In questo scenario di "concorrenza sull'analitica"¹⁵, il settore pubblico è rimasto evidentemente indietro. Da una parte, la digitalizzazione dei processi e dei servizi pubblici è progredita più lentamente rispetto alle attività dell'economia

¹² AGRAWAL, GANS, Goldfarb, *Prediction Machines: The Simple Economics of Artificial Intelligence* (Harvard Business Press, Harvard, 2018).

¹³ MCAFEE, BRYNJOLFSSON, DAVENPORT, *Big Data: The Management Revolution*, in *Harvard Business Review*, Vol. 90, 2012, p. 60.

¹⁴ Si veda CONSTANTIOU, KALLINIKOS, *New Games, New Rules: Big Data and the Changing Context of Strategy*, in *Journal of Information Technology*, vol. 30, 2015, p. 44, in cui gli Autori parlano di un vero e proprio nuovo contesto per le strategie aziendali, in cui l'obiettivo principale è acquisire informazioni strutturate di valore duraturo che rispondano a obiettivi organizzativi specifici e di lungo termine.

¹⁵ DAVENPORT e HARRIS (n 9).

privata. Ciò ha comportato, un conseguente ritardo nel processo di “datificazione”¹⁶ delle interazioni tra governo e individui e dunque nella possibilità di immagazzinare più agevolmente e disporre di informazioni sui cittadini¹⁷. Dall’altra, gli enti pubblici sono meno inclini ad analizzare i dati in loro possesso per trarne informazioni, per un retaggio culturale derivante dal classico modello weberiano di burocrazia¹⁸. Secondo questo modello, i dati sugli individui sono contenuti all’interno di archivi isolati, disponibili per il controllo di singole informazioni, ma senza che gli archivi generino altri dati utilizzabili per analisi successive. Questa caratteristica dell’architettura dell’informazione dei governi si è mantenuta anche nell’era dell’informatizzazione, e la mancanza di dati utilizzabili è rimasta una caratteristica di molti governi.

Tuttavia, negli ultimi anni, il crescente interesse verso le nuove tecniche di analisi dei dati e verso l’IA ha stimolato alcune amministrazioni pubbliche a riconfigurare il ciclo della regolamentazione e ad incentrarlo sul dato¹⁹. Questo processo inizia con la raccolta e la generazione di dati e prosegue con il loro utilizzo per estrarre informazioni e definire e attuare le politiche. Si conclude con una valutazione delle misure implementate per determinare se abbiano raggiunto i loro obiettivi, il che può richiedere la raccolta o la generazione di ulteriori dati, in modo che il ciclo possa ricominciare.

3. Raccolta dei dati: costi e limiti

Un settore pubblico “guidato dai dati” – nel senso che si affidi fortemente ai dati per prevedere le esigenze, modellare l’erogazione dei servizi e comprendere e rispondere ai cambiamenti – rappresenta perciò una dimensione chiave della strategia di *better regulation*. Conseguentemente, sono necessari sforzi attivi per rimuovere gli ostacoli alla raccolta, gestione e condivisione dei dati.

¹⁶ Con questo termine, si intende non semplicemente la creazione di informazioni, ma la quantificazione della vita umana attraverso l’informazione digitale, spesso caratterizzata da valore economico. Si veda per esteso, Ulises A MEJIAS, COULDRY, *Datafication in Internet Policy Review*, vol. 8, 2019.

¹⁷ LINDGREN, MADSEN, HOFMANN, MELIN, CLOSE, *Close Encounters of the Digital Kind: A Research Agenda for the Digitalization of Public Services*, in *Government Information Quarterly*, vol. 36, 2019, p. 427.

¹⁸ Così secondo MARGETTS, *Rethinking AI for Good Governance in Daedalus*, vol. 151, 2022, p. 360.

¹⁹ Si segnalano i due recenti rapporti dell’OECD, VAN OOIJEN, UBALDI, WELBY, *A Data-Driven Public Sector* <<https://www.oecd-ilibrary.org/content/paper/09ab162c-en>>; OECD, *The Path to Becoming a Data-Driven Public Sector*, 2019, <<https://www.oecd-ilibrary.org/content/publication/059814a7-en>>; e un rapporto del Parlamento europeo, MISURACA, BARCEVIČIUS, CODAGNONE, *Exploring Digital Government Transformation in the EU - Understanding Public Sector Innovation in a Data-Driven Society*, Publications Office of the European Union <<https://publications.jrc.ec.europa.eu/repository/handle/JRC121548>>.

L'uso efficiente ed efficace dei dati nel settore pubblico non deve essere dato per scontato. Richiede una struttura complessa di governance che includa, non solamente un'infrastruttura tecnica per la raccolta, conservazione e gestione di dati, ma anche la capacità di progettare una politica dei dati, di attuarla, di adottare norme appropriate, di coinvolgere attori con le competenze necessarie. I dati che possono avere rilevanza per le istituzioni pubbliche provengono da molte fonti diverse e possono essere acquisiti in molti modi diversi. Organizzazioni governative, cittadini, imprese, ricercatori e altre parti interessate della società contribuiscono tutti, consapevolmente e inconsapevolmente, alla generazione di dati, in formati aperti e chiusi, in grandi e piccole quantità, con e senza struttura, di natura personale e no.

Quando si considera l'uso dei dati da parte delle amministrazioni pubbliche, è importante sottolineare che la raccolta/generazione, la gestione e l'elaborazione dei dati comportano costi e benefici. Un'analisi costi-benefici si applica anche ai dati, poiché il costo delle attività legate ai dati è un aspetto del costo più ampio della progettazione e dell'attuazione di una politica.

Nel considerare il costo della raccolta dei dati ai fini della valutazione di una politica legislativa, un aspetto importante riguarda la possibilità di riutilizzare i dati, cioè di impiegarli per finalità non incompatibili con quelle per cui originariamente i dati erano stati raccolti. In particolare, i dati utilizzati per l'attuazione di una politica possono essere reimpiegati (senza costi di raccolta aggiuntivi) per valutare la politica ed eventualmente riformarla. Inoltre, per rendere la raccolta dei dati efficace dal punto di vista dei costi e sostenibile, si rende necessaria la progettazione di politiche di raccolta e creazione dei dati, in base alle quali i dati vengono prodotti automaticamente come effetto collaterale dei processi amministrativi, derivanti dalle azioni delle amministrazioni pubbliche e dei cittadini.

Nell'analisi degli impatti occorre anche considerare le questioni relative alla privacy e alla protezione dei dati. In questo ambito, un ruolo fondamentale è svolto dalla legge sulla protezione dei dati, tra cui il Regolamento Generale sulla Protezione dei Dati (GDPR)²⁰, la direttiva ePrivacy²¹, la direttiva sull'applicazione dei

²⁰ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE.

²¹ Direttiva 2002/58/CE del Parlamento europeo e del Consiglio, del 12 luglio 2002, relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche

dati e l'applicazione della legge), presto sostituita dal Regolamento e-Privacy²², il Regolamento sulla circolazione dei dati non personali²³ nonché dalle nuove leggi emergenti sui dati, tra cui la proposta di legge sulla governance dei dati (Data Governance Act)²⁴ e la legge sui dati (Data Act)²⁵. In particolare, queste ultime due normative affrontano anche il riutilizzo dei dati, sia quelli raccolti da enti pubblici (da rendere accessibili anche a soggetti privati) che quelli raccolti da soggetti privati (da rendere accessibili anche a enti pubblici).

Un aspetto cruciale riguarda la privacy by design e by default²⁶, in base alle quali le considerazioni relative alla privacy degli individui dovrebbe essere prese in considerazione fin dalla progettazione delle politiche di raccolta dati e nella selezione delle tecnologie impiegate. Dovrebbero essere adottate misure di riduzione del rischio²⁷ con particolare attenzione all'anonimizzazione o pseudonimizzazione dei dati, specialmente quando i dati utilizzati dalle pubbliche amministrazioni per valutare l'impatto delle politiche non richiedano l'identificazione degli individui persone fisiche.

In taluni casi, la pubblica amministrazione dovrà svolgere anche una valutazione d'impatto sulla protezione dei dati²⁸, allorché, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, l'utilizzo dei dati possa comportare un rischio elevato per gli interessati. Nel contesto che ci interessa, ciò può avvenire quando i governi effettuano valutazione su larga scala sui cittadini, specialmente quando si ricorra a tecniche di profilazione, ma anche quando vengono svolte valutazioni su categorie di dati sensibili (come nel caso della modellazione dell'erogazione dei servizi ospedalieri basate sui dati sanitari dei cittadini). Un soggetto chiave nella valutazione d'impatto è il responsabile della protezione dei dati (DPO) che un organismo pubblico è sempre e necessariamente tenuto a designare²⁹. Questi dovrebbe svolgere un ruolo chiave nella definizione delle politiche di raccolta e

²² Proposta di Regolamento del Parlamento europeo e del Consiglio relativo al rispetto della vita privata e alla tutela dei dati personali nelle comunicazioni elettroniche e che abroga la direttiva 2002/58/CE, COM/2017/010 final.

²³ Regolamento (UE) 2018/1807 del Parlamento europeo e del Consiglio, del 14 novembre 2018, relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea.

²⁴ Proposta di Regolamento del Parlamento europeo e del Consiglio relativo alla governance europea dei dati, COM/2020/767 final.

²⁵ Proposta di Regolamento del Parlamento europeo e del Consiglio riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo, COM/2022/68 final.

²⁶ Articolo 25 GDPR

²⁷ Articolo 32 GDPR.

²⁸ Articolo 35 GDPR.

²⁹ Articolo 37 GDPR.

gestione dei dati, assicurando il rispetto delle norme giuridiche a tutela della privacy.

4. L'analisi e l'uso dei dati attraverso l'IA per la pianificazione delle leggi

Uno studio del 2020 sui big data e l'IA nel governo federale degli Stati Uniti ha rilevato che quasi la metà delle agenzie studiate ha sperimentato strumenti di apprendimento automatico³⁰. Tra le applicazioni dell'IA elencate, vi sono le seguenti: l'estrazione di informazioni da rapporti testuali su eventi avversi ai farmaci; l'analisi dei reclami dei consumatori, l'elaborazione di racconti di infortuni dei lavoratori.

Più in generale, le tecniche di analisi dei dati basate sull'IA possono essere utilizzate anche dal settore pubblico per l'elaborazione delle politiche aiutando il governo a svolgere tre compiti chiave: (1) l'analisi e l'individuazione, (2) la previsione dei comportamenti e (3) la simulazione delle interazioni³¹.

Il compito di analisi e individuazione riguarda la comprensione dei comportamenti, delle tendenze e dei modelli sociali ed economici, come condizione preliminare per formulare le politiche pubbliche. Grazie ai big data e all'analitica avanzata, i governi possono analizzare le preferenze e le esigenze politiche dei cittadini per capire meglio quali incentivi funzioneranno, in quali circostanze³². L'estrema velocità del flusso dei dati consente inoltre ai responsabili politici di riorganizzare rapidamente le strategie politiche e di incorporare nuove informazioni collettive provenienti da una varietà di fonti, compresi i canali al di fuori dell'ambito politico formale come i social media, le consultazioni online e i municipi virtuali³³. Un aspetto collegato è rappresentato dalla possibilità di individuare i comportamenti e le preferenze di cittadini o di gruppi di cittadini per l'erogazione di servizi personalizzati. Ad esempio, l'utilizzo di dati provenienti da fascicoli sanitari elettronici e il monitoraggio dei risultati delle politiche sanitarie possono aiutare a migliorare l'accesso al sistema sanitario e sviluppare nuovi interventi per popolazioni più svantaggiate.

Un'importante preoccupazione riguarda l'individuazione di comportamenti di individui specifici, come ad esempio, quando l'IA è usata per l'identificare le frodi

³⁰ ENGSTROM, HO, SHARKEY, CUÉLLAR, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, in *NYU School of Law*, Public Law Research Paper, 2020.

³¹ MARGETTS, DOROBANTU, *Rethink Government with AI*, in *Nature*, vol. 568, 2019, p. 163.

³² HÖCHTL, PARYCEK, SCHÖLLHAMMER, *Big Data in the Policy Cycle: Policy Decision Making in the Digital Era*, in *Journal of Organizational Computing and Electronic Commerce*, vol. 26, 2016, p. 147.

³³ MERGEL, *Building Holistic Evidence for Social Media Impact*, in *Public Administration Review*, vol. 77, 2017, p. 489.

fiscali, il riciclaggio di denaro e le minacce al terrorismo. In questo caso, l'uso dei big data e dell'IA a fini di individuazione può svolgere un ruolo utile (per esempio, per prevenire fenomeni illeciti), ma i controlli dovrebbero essere applicati in modo equo a tutta la popolazione, con adeguate cautele e un'attenta valutazione umana dei casi potenzialmente illegali identificati dalla macchina. Un caso significativo di utilizzo abusivo dell'IA è emerso di recente nei Paesi Bassi³⁴, dove un progetto su larga scala finalizzato all'individuazione automatica delle frodi fiscali ha fatto sì che un gran numero di persone venisse privato ingiustamente dei sussidi e sottoposto a multe, e che alcuni gruppi venissero sottoposti a un trattamento iniquo.

Un'altra possibilità fornita dall'IA è rappresentata dalla predizione³⁵, e cioè la possibilità di anticipare eventi individuali o fenomeni aggregati. Le tecniche di previsione basate sui dati possono aiutare i governi ad anticipare gli sviluppi sociali, economici o naturali che probabilmente si verificheranno in futuro, ma consentono anche di sviluppare sistemi di allerta precoce e sistemi di supporto decisionale in tempo reale per valutare meglio la necessità di intervenire, progettare le misure politiche appropriate e anticipare con maggiore precisione gli impatti previsti³⁶. Ad esempio, piani di aiuto per luoghi terremotati potrebbero essere attivati sulla base di indici predittivi, invece che a posteriori, evitando alcuni danni. Inoltre, le predizioni possono consentire ai governi di anticiparne le richieste aggregate di servizi (come la necessità della capienza necessaria degli ospedali, i livelli di educazione, l'assistenza nelle scuole d'infanzia) e migliorarne così l'erogazione.

Anche in questo ambito, sono state sollevate preoccupazioni riguardo le predizioni individualizzate. Questa pratica, diffusa nel settore privato – ad esempio, nell'assistenza sanitaria, nell'industria assicurativa, nel *credit scoring* e nel reclutamento di personale³⁷ – può avere alcune applicazioni controverse anche nel settore pubblico. Esempi sono la valutazione della recidiva e la c.d. "polizia predittiva"³⁸,

³⁴ HEIKKILA, *Dutch Scandal Serves as a Warning for Europe over Risks of Using Algorithms*, *Politico* 2022, <https://www.politico.eu/cdn.ampproject.org/c/s/www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/amp>. (ultimo accesso Settembre 2022)

³⁵ AGRAWAL, GANS, GOLDFARB (n 12).

³⁶ ATHEY, *Beyond Prediction: Using Big Data for Policy Problems*, in *Science*, vol. 355, 2017, p. 483.

³⁷ LAGIOIA, SARTOR, *Profilazione e Decisione Algoritmica: Dal Mercato Alla Sfera Pubblica*, in *Federalismi.it*, 2020, p. 85.

³⁸ SUSSER, *Predictive Policing and the Ethics of Preemption*, in Jones, Mendieta (a cura di), *The Ethics of Policing: New Perspectives on Law Enforcement*, 2021, p. 268. Si pensi altresì all'utilizzo di sistemi predittivi nel valutare il rischio di recidiva del singolo detenuto, come il sistema COMPAS utilizzato negli Stati Uniti. Si veda il rapporto dell'agenzia di informazione ProPublica, Julia ANGWIN, *How We Analyzed the COMPAS Recidivism Algorithm*, *ProPublica*, vol 28., 2016.

nella medicina predittiva,³⁹ nell'identificazione delle famiglie a rischio di violenza o di abbandono dei bambini⁴⁰. Meno problematico, invece, è l'uso della previsione per anticipare fenomeni aggregati. Tali previsioni possono riguardare le condizioni future (ad es., le tendenze economiche e sociali) o gli esiti delle politiche. In quest'ultimo caso, è possibile determinare i risultati attesi di interventi alternativi attraverso un'analisi causale, cioè che ricollegate ad una certa misura politica il suo probabile effetto nella popolazione. Per la valutazione d'impatto *ex ante* delle normative possono essere necessarie previsioni sia sulle condizioni sociali, sia sugli esiti delle politiche; le prime forniscono il contesto per le seconde. La previsione automatica può basarsi su diversi modelli, ad esempio econometrici o statistici, eventualmente integrati con tecnologie per l'apprendimento automatico. L'attuale entusiasmo per l'IA non deve farci dimenticare che tutti i modelli sono fallibili e che i modelli di ottimizzazione statistica più tradizionali possono spesso superare gli approcci dell'IA.

Infine, uno sviluppo particolare, reso possibile dalla disponibilità di tecnologie avanzate e di tecniche di calcolo appropriate, è la cd. "simulazione sociale". Questa si basa su una modellazione ad agenti, piuttosto che su una formalizzazione analitica attraverso equazioni matematiche. I modelli basati sugli agenti sono in linea di principio sperimentali: il modello specifica le caratteristiche e il comportamento dei singoli agenti in modo che la dinamica aggregata del sistema emerga attraverso le interazioni di tali agenti. Ad esempio, per prevedere i modelli di traffico stradale, si può specificare attraverso appositi algoritmi il comportamento di ciascun veicolo (dato il tipo, la proprietà, ecc.) e le caratteristiche dell'ambiente (luoghi, strade, ecc.) in cui i veicoli operano. Il traffico (con congestioni, incidenti, ecc.) risulterà dalle interazioni dei singoli veicoli. Allo stesso modo, creando agenti digitali che si comporteranno in determinati modi durante gli scambi commerciali, è possibile studiare le dinamiche dei mercati.

Un'idea interessante legata al concetto di simulazione è quella di "gemello digitale"⁴¹, cioè una rappresentazione virtuale di un'entità fisica o sociale, che è

³⁹ ADJEKUM, IENCA, VAYENA, *What Is Trust? Ethics and Risk Governance in Precision Medicine and Predictive Analytics*, in *Omics: a journal of integrative biology*, vol. 21, 2017, p. 704.

⁴⁰ HUNT, TOMLINSON, SIKANDER, SKEEN, MARLOW, DU TOIT, EISNER, *Artificial Intelligence, Big Data, and mHealth: The Frontiers of the Prevention of Violence Against Children*, in *Frontiers in artificial intelligence*, vol. 3, 2020, p. 543305.

⁴¹ FULLER, FAN, DAY, BARLOW, *Digital Twin: Enabling Technologies, Challenges and Open Research*, Disponibile su IEEE, vol. 8, 2020, n. 108952.

costantemente collegata a tale entità attraverso flussi di dati⁴². L'idea del gemello digitale ha origine nell'ingegneria (in particolare in quella spaziale) ma da allora è stata estesa non solo ai dispositivi tecnologici, ma anche a edifici, fabbriche, città e altri sistemi. Il gemello digitale viene utilizzato per anticipare i problemi del corrispondente sistema reale e testare potenziali soluzioni prima di implementarle nel mondo reale. È stato sostenuto che le tecnologie di intelligenza artificiale e i big data permettono di creare gemelli digitali non solo di sistemi fisici, ma anche di sistemi sociali e sociotecnici, cioè veri e propri, gemelli della società. Tali modelli potrebbero essere utilizzati per determinare in modo proattivo come i sistemi sociali possono rispondere alle contingenze future, identificare i problemi e valutare i possibili interventi, come l'emanazione di nuove norme⁴³.

5. Dati e modelli per il monitoraggio e la valutazione della regolazione

Anche quando basata su metodi computazionali avanzati, la valutazione d'impatto per l'adozione di misure regolatorie rimane fortemente congetturale. Ciò è dovuto alla combinazione di diversi aspetti, come la disponibilità di dati, la complessità dei contesti sociali regolati, la difficoltà delle analisi causali.

In alcuni casi, l'incertezza può essere trattata matematicamente, in quanto è possibile assegnare probabilità al verificarsi di eventi futuri. Tuttavia, esistono anche casi in cui l'incertezza riguarda queste stesse probabilità, e ciò rende impossibile stabilire quale sia la probabilità che un evento futuro si verifichi. Questo accade spesso quando occorre valutare l'impatto di fattori innovativi (come ad es., le innovazioni tecnologiche) o in caso di scenari ad alta complessità (come in molti contesti ecologici e sociali)⁴⁴,

Dunque, l'incertezza della valutazione *ex ante*, anche quando viene effettuata affidandosi alle migliori metodologie, rende necessario sottolineare l'importanza del monitoraggio intermedio e della valutazione *ex post* nel ciclo della regolazione. Entrambi hanno lo scopo di verificare in che misura le previsioni *ex ante* siano

⁴² Secondo la definizione di IBM, il gemello digitale è “una rappresentazione virtuale di un oggetto o di un sistema che copre il suo ciclo di vita, viene aggiornato da dati in tempo reale e utilizza la simulazione, l'apprendimento automatico e il ragionamento per aiutare il processo decisionale”.

⁴³ Ad esempio, César Hidalgo, direttore del Center for Collective Learning di Tolosa, ha presentato l'idea di “democrazia aumentata”, e cioè la possibilità di istituire società democratiche virtuali attraverso gemelli digitali fornendo ai cittadini avatar digitali attraverso i quali partecipare direttamente a molte decisioni democratiche. Il contenuto della presentazione è disponibile al seguente link: https://www.ted.com/talks/cesar_hidalgo_una_idea_osada_para_reemplazar_a_los_politicos.

⁴⁴ HANSSON, HADORN, *The Argumentative Turn in Policy Analysis: Reasoning about Uncertainty*, vol 10, Springer, Cham, 2016.

verificate o contraddette dai fatti successivi, nonché in che misura siano emersi effetti collaterali inattesi. In questo modo, il monitoraggio e la valutazione possono fornire prove per aggiustamenti e revisioni delle politiche. Più in generale, attraverso il monitoraggio e la valutazione, i responsabili politici possono essere chiamati a rispondere dei risultati reali (piuttosto che di quelli attesi) delle loro politiche, e il processo stesso di valutazione *ex ante* (i dati raccolti, la metodologia utilizzata, ecc.) può essere sottoposto a un esame critico.

Il monitoraggio e la valutazione successiva delle leggi possono basarsi su modelli diversi, che richiedono informazioni diverse.

Attualmente, un modello popolare è l'analisi costi-benefici, che misura i potenziali benefici (vantaggi) di una misura in esame rispetto alle sue potenziali perdite (svantaggi). I benefici e le perdite sono quantificati da numeri espressi nella stessa unità, di solito corrispondente a una valuta comune (ad esempio, euro o dollari). Queste quantità - la misura del beneficio o della perdita complessiva in gioco - sono spesso determinate facendo riferimento alla disponibilità degli individui a pagare per ottenere un beneficio o alla loro disponibilità ad accettare un pagamento in cambio di una perdita, tenendo conto anche della probabilità che tale beneficio o perdita si verifichi. Sommando tutti i benefici attesi e sottraendo tutte le perdite attese, per tutti gli individui, si ottiene un unico numero che indica il merito complessivo della misura considerata, che può essere confrontato con il numero corrispondente assegnato a ciascuna delle alternative in esame.

L'analisi costi-benefici è un approccio consolidato, che ha il vantaggio di fornire risultati numerici comparabili per tutte le opzioni prese in considerazione. Tuttavia, in molti ambiti questo metodo incontra delle difficoltà, legate alla difficoltà di renderlo operativo e ad alcune questioni normative. In particolare, l'analisi costi-benefici richiede, in linea di principio, la raccolta di dati sulla disponibilità a pagare o ad accettare, per tutti gli individui interessati. Tale esigenza informativa può essere soddisfatta solo indirettamente, affidandosi a *proxy*. Ulteriori problemi riguardano il fatto che l'analisi costi-benefici non tiene conto della distribuzione dei benefici e dei costi tra i diversi individui e gruppi (si limita a considerare la loro somma), e tende a sottovalutare i beni collettivi e i beni a cui non è possibile assegnare un valore monetario significativo (come la vita umana o il rispetto dei diritti fondamentali)⁴⁵. In generale, quindi, si ritiene che l'analisi costi-benefici possa

⁴⁵ HANSSON, *Philosophical Problems in Cost-Benefit Analysis*, in *Economics & Philosophy*, vol. 23, 2007, p. 163.

contribuire ad una corretta valutazione dell'impatto di alcune politiche, ma raramente può fornire una valutazione complessiva.

Criteri alternativi o complementari per la valutazione delle misure legislative sono l'analisi dei costi minimi e l'analisi del rapporto costo-efficacia. La prima considera solo i costi, al fine di selezionare l'opzione alternativa che comporta il costo più basso; pertanto, non affronta adeguatamente i casi in cui una misura con costi più elevati fornirebbe maggiori benefici (e i maggiori benefici giustificano pienamente i costi più elevati). L'analisi del rapporto costo-efficacia considera la relazione tra unità di beneficio e unità di spesa (ad esempio, dividendo il numero di vite salvate per l'importo in euro necessario per le misure sanitarie); pertanto, non affronta adeguatamente i casi in cui una misura di portata più ampia fornirebbe un maggiore beneficio netto complessivo (beneficio-costi), anche se una misura più ristretta potrebbe avere un maggiore rapporto costo-efficacia (per es., la vaccinazione di tutti gli individui, piuttosto che solo quelli a maggior rischio, durante una pandemia, può portare un maggior beneficio aggregato)⁴⁶.

Esistono poi modelli di valutazione multicriterio in cui si considerano separatamente il raggiungimento di diversi obiettivi, come il contributo che una politica può dare al PIL, all'ambiente o ai diritti individuali, eventualmente utilizzando metodi e scale diverse per misurare tali impatti. Una misura può infatti essere più favorevole in base ad alcuni criteri e meno favorevole in base ad altri rispetto ad altre misure (ad esempio, un'opzione normativa è più protettiva di alcuni diritti individuali, come la privacy, e più costosa per le aziende). In questi casi, è necessario determinare se i vantaggi della misura per certi aspetti (ad es., i diritti alla privacy) siano più o meno importanti dei vantaggi della misura alternativa per altri aspetti (ad es., la riduzione dei costi per le aziende). Pertanto, può essere necessaria una valutazione umana per stabilire quale misura (quale pacchetto di benefici e perdite) sia complessivamente preferibile, oppure deve essere definito un calcolo che trasformi ogni beneficio o perdita in una valuta comune (es., unità di utilità, dollari, euro, ecc.), o che renda altrimenti possibile confrontare misure alternative⁴⁷.

⁴⁶ Per una discussione sui diversi modelli di valutazione della regolazione, si veda RENDA, *Regulatory Impact Assessment and Regulatory Policy*, in *Regulatory Policy in Perspective: A Reader's Companion to the OECD Regulatory Policy Outlook*, 2015, p. 35. Per alcune considerazioni critiche, si veda anche MICKLITZ, *The Measuring of the Law Through EU Politics*, in Bartl, Lawrence (a cura di), *The politics of european legal research*, Edward Elgar Publishing, Cheltenham, 2022.

⁴⁷ Esiste una vasta gamma di approcci al processo decisionale multicriterio. Per un contributo approfondito sul tema si veda KEENEY, RAIFFA, MEYER, *Decisions with Multiple Objectives: Preferences and*

Il processo decisionale multicriterio sembra essere stato adottato nelle valutazioni normative dell'UE, almeno per i regolamenti più significativi. Ad esempio, la valutazione d'impatto del Digital Services Act⁴⁸ considera separatamente gli impatti economici, gli impatti sociali, gli impatti sui diritti fondamentali e gli impatti ambientali, dettagliando ulteriormente tali impatti e cercando di quantificarne alcuni (in particolare, gli impatti economici e ambientali).

Di recente è cresciuto l'interesse per gli Obiettivi di sviluppo sostenibile delle Nazioni Unite⁴⁹ che consistono in diciassette obiettivi: (1) nessuna povertà; (2) zero fame; (3) buona salute e benessere; (4) istruzione di qualità; (5) uguaglianza di genere; (6) acqua pulita e servizi igienici; (7) energia pulita e accessibile; (8) lavoro dignitoso e crescita economica; (9) industria, innovazione e infrastrutture; (10) riduzione delle disuguaglianze; (11) città e comunità sostenibili; (12) consumo e produzione responsabili; (13) l'azione per il clima; (14) la vita sotto l'acqua; (15) la vita sulla terra; (16) la pace, la giustizia e le istituzioni forti; (17) le partnership per questi obiettivi. Il quadro degli Obiettivi di sviluppo sostenibile prevede anche sotto-obiettivi per ogni obiettivo e indicatori per il loro raggiungimento. Sebbene siano stati proposti metodi per calcolare il merito delle politiche rispetto a questo quadro⁵⁰, si tratta di un compito impegnativo, data la diversità degli obiettivi e i molteplici indicatori con cui sono quantificati.

6. La necessità di un'etica dei dati per il settore pubblico

Una sfida cruciale che i governi devono affrontare per massimizzare il ruolo dei dati nella progettazione delle politiche, nell'erogazione dei servizi e nell'efficienza delle operazioni riguarda la legittimità e l'affidabilità delle iniziative di utilizzo dei dati. Questo sforzo va oltre il rispetto delle leggi sulla protezione dei dati e degli altri strumenti giuridici applicabili, e prende in considerazione l'etica dei dati. Quest'ultima indica i valori morali da rispettare e perseguire nell'utilizzo dei dati

Value Trade-Offs Cambridge University Press, Cambridge, 1993; per un resoconto più recente ISHIZAKA, NEMERY, *Multi-Criteria Decision Analysis: Methods and Software*, John Wiley & Sons, 2013.

⁴⁸ Commission Staff Working Document - Impact Assessment Accompanying the document Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, Brussels, 15.12.2020 SWD(2020) 348 final.

⁴⁹ SCHMIDT-TRAUB, DE LA MOTHE, KAROUBI, ESPEY, *Indicators and a Monitoring Framework for the Sustainable Development Goals - Launching a Data Revolution for the SDGs*, United Nations, 2015.

⁵⁰ Si veda, ad esempio, la proposta di GUERRERO, CASTAÑEDA, *Policy Priority Inference: A Computational Framework to Analyze the Allocation of Resources for the Sustainable Development Goals*, in *Data & Policy*, vol. 2, 2020.

e le corrispondenti condotte moralmente adeguate⁵¹. La moralità nella gestione dei dati pubblici è preziosa per sostenere lo stato di diritto e per raggiungere un equilibrio tra azione governativa e bene sociale. Inoltre, contribuisce a conservare la fiducia dei cittadini e il loro sostegno alle politiche, che a loro volta facilitano la raccolta e il riutilizzo dei dati.

Oltre a garantire la privacy dei cittadini, i governi dovrebbero assicurare che la gestione dei dati e il loro utilizzo avvengano in modo responsabile, trasparente e sicuro, proporzionato ai benefici attesi per la società.

Un aspetto importante riguarda l'allocazione di responsabilità nei processi decisionali⁵². Gli analisti di dati, siano essi interni o esterni ai governi, non sono decisori dotati di diretta legittimazione democratica. Quando si tratta di progettare e attuare una politica, l'utilizzo di strumenti di *analytics* non deve perciò sostituirsi ai processi democratici di elaborazione e discussione delle politiche pubbliche, né alle procedure parlamentari di adozione delle leggi. Occorre dunque tenere ben distinta la funzione dei *data analyst* nei cicli regolatori da quella dei decisori politici: i primi debbono fornire informazioni adeguate e possibili soluzioni per l'adozione di politiche basate su evidenze empiriche; i secondi debbono adottare la decisione finale (e dunque assumersene la responsabilità), tenendo in considerazione il dibattito pubblico e le aspettative dei cittadini.

Il tema della responsabilità assume un'importanza speciale quando nel ciclo regolatorio siano coinvolte tecnologie di IA che possono prevedere le strategie politiche più efficaci e prefigurare un possibile scenario nella società collegata all'adozione di una certa misura legislativa. In questo caso, è importante che gli esseri umani rimangano in grado di decidere se applicare una determinata politica suggerita dal sistema oppure discostarsi dalla raccomandazione e adottarne una diversa. Inoltre, gli esseri umani devono poter essere in grado di stabilire i criteri utilizzati dai sistemi per selezionare le migliori strategie politiche che desiderano vedere attuati attraverso le misure selezionate.

L'esigenza di *accountability* pone anche un problema di trasparenza dei processi decisionali⁵³. I governi devono essere trasparenti sui tipi e sulla provenienza dei dati che vengono utilizzati all'interno di un processo di *better regulation*, e devono

⁵¹ FLORIDI, TADDEO, *What Is Data Ethics?* in *Philosophical Transactions of the Royal Society*, 2016.

⁵² Sul tema si veda due recenti contributi, COGLIANESE, LAMPMANN, *Contracting for Algorithmic Accountability*, in *Administrative Law Review Accord*, vol. 6, 2021, p. 175 e ENGSTROM, HO, *Algorithmic Accountability in the Administrative State*, in *Yale J. on Reg.*, 37, 2020, p. 800.

⁵³ COGLIANESE, LEHR, *Transparency and Algorithmic Governance*, in *Admin. L. Rev.*, vol. 71, 2019, p. 9.

garantire che il loro utilizzo avvenga entro confini etici e secondo principi democratici. La trasparenza deve riguardare anche l'uso di sistemi di IA: ciò richiede che i governi comunichino la provenienza dei sistemi utilizzati (specialmente se l'analisi venga affidata a soggetti privati), come vengono addestrati, quali sono gli obiettivi target massimizzati dai sistemi e quali i *proxy* (ad es., in quale modo vengono calcolati i costi e i benefici di una certa misura politica) e quale sia il peso dato alle previsioni automatizzate all'interno dell'intero processo decisionale. Il settore pubblico dovrebbe perciò orientare le proprie scelte verso strumenti tecnologici che permettano la spiegabilità (*explainability*) delle predizioni, per capire quale sia la logica dietro la raccomandazione di una certa politica.

Infine, un altro elemento di un'azione governativa improntata all'etica dei dati è costituito dalla sicurezza⁵⁴. Le attività di politica pubblica incentrate sui dati possono esporre informazioni sensibili dei cittadini ad attacchi informatici esterni, ma anche ad eventuali abusi da parte degli stessi governi. Per rispondere a queste minacce è necessario che vi siano sforzi proattivi da parte degli enti pubblici per garantire la sicurezza dei loro dati. Una maggiore responsabilizzazione all'interno delle organizzazioni pubbliche verso la sicurezza dei dati rappresenta un passo fondamentale per l'accettabilità da parte dei cittadini di politiche di raccolta e produzione di dati, e di utilizzo per migliorare le politiche pubbliche.

Gli Stati e gli enti pubblici devono sviluppare quadri etici per l'utilizzo dei big data e dell'IA nel perseguimento delle politiche pubbliche. Esistono già esempi virtuosi. Ad esempio, la OECD ha sviluppato un quadro etico per la gestione dei dati nel settore pubblico⁵⁵. I principi mirano a promuovere l'implementazione dell'etica dei dati nei progetti, prodotti e servizi di governo digitale. Tra i governi nazionali europei, il Regno Unito è stato uno dei primi ad adottare un quadro etico per i dati. Il *Data Ethics Framework*⁵⁶ rappresenta una guida l'uso appropriato e responsabile dei dati nel governo e nel settore pubblico in generale, al fine di incoraggiare un'innovazione responsabile. Altri paesi, come la Francia e la Germania, hanno invece avviato iniziative volte a istituire commissioni nazionali per lo sviluppo di un quadro etico delle tecnologie legate ai big data. Benché non riguardino specificamente il settore pubblico, queste iniziative hanno contribuito ad una discussione su un utilizzo etico dei dati nel settore pubblico. Si ricorda in Francia,

⁵⁴ TADDEO, *Is Cybersecurity a Public Good?*, in *Minds and Machines*, 2019, vol. 29, p. 349.

⁵⁵ OECD, *Good Practice Principles for Data Ethics in the Public Sector*, 2020, <<https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.pdf>>.

⁵⁶ <https://www.gov.uk/government/publications/data-ethics-framework>.

l'istituzione del Comitato nazionale per l'etica digitale (CNPEN) e in Germania, la creazione della Commissione per l'etica dei dati presso il Governo federale.

7. Conclusioni

Lo sfruttamento strategico dei big data e dell'IA presenta un enorme potenziale per sostenere programmi di *better regulation*. In questo nuovo contesto, i governi e gli enti pubblici possono lavorare in modo più efficiente, diventare più produttivi e avere la capacità di trasformare la progettazione e l'erogazione di politiche e servizi pubblici in modi che migliorano il benessere dei cittadini e della società in generale. A sua volta, questo può avere un impatto positivo sulla fiducia dei cittadini nei confronti dei decisori pubblici, sul presupposto che lo sviluppo di politiche sociali sia basato su evidenze reali e prenda in considerazione i bisogni effettivi della società. Per realizzare questo potenziale, è fondamentale che i governi esplorino e comprendano il valore innovativo e trasformativo dei dati e creino le condizioni tecniche e organizzative necessarie per sfruttarlo.

In questo contributo abbiamo tratteggiato le sfide connesse a questa riconversione tecnologica, in particolare per la valutazione di impatto *ex ante* ed *ex post* delle strategie regolatorie.

Per quanto riguarda la raccolta e la produzione dei dati, abbiamo osservato come il settore pubblico debba sviluppare le competenze tecniche e servirsi delle architetture informatiche adeguate a raccogliere e gestire ampie quantità di dati. A tal fine, si dovrebbero compiere sforzi per favorire una definitiva transizione verso la digitalizzazione dei servizi al fine di accelerare il processo di datificazione delle interazioni tra cittadini e apparati pubblici. In questo processo è necessario, da una parte, considerare i vantaggi della raccolta, della conservazione e dell'utilizzo dei dati, adottando soluzioni efficaci dal punto di vista dei costi (che, ad es., prevedano il riutilizzo dei dati per finalità connesse). Dall'altra, occorre ricorrere a soluzioni rispettose della privacy degli individui, che rispettino i principi sulla protezione dei dati e favoriscano il principio di responsabilizzazione degli attori pubblici.

In secondo luogo, quanto agli strumenti di analisi dei dati, si è analizzato come i governi possano beneficiare delle opportunità offerte dall'IA. Le tecnologie di IA possono aiutare gli apparati governativi a svolgere i compiti chiave per promuovere il benessere dei cittadini. Esse possono consentire l'analisi di grandi masse di dati in tempo reale per: rilevare e comprendere il comportamento dei target regolatori; costruire modelli predittivi e supportare il processo decisionale; simulare scenari dove sperimentare interventi politici che ottengano gli obiettivi perseguiti ed evitino conseguenze indesiderate. In tutte queste attività, è importante tenere distinto

l'utilizzo di IA a fini statistici, finalizzato cioè a fornire informazioni descrittive o prescrittive aggregate sui cittadini e sulla società, rispetto al trattamento dei dati personali finalizzato a fornire predizioni e trattamenti individualizzati. Il primo non crea di per sé problemi nella misura in cui i dati vengano raccolti e trattati conformemente alla normativa sulla privacy; il secondo, invece, può comportare importanti problemi etico-giuridici per i cittadini che includono un'eccessiva intrusione nella sfera persona, la discriminazione nell'accesso ai servizi, e lo sfruttamento delle vulnerabilità.

Infine, si è affrontato l'utilizzo di nuovi modelli di dati per il monitoraggio e la valutazione *ex post* delle leggi. Considerata la natura congetturale delle valutazioni d'impatto *ex ante*, è estremamente importante che il ciclo della regolazione includa una valutazione della normativa rispetto agli obiettivi preposti, rendendo possibile l'adeguamento delle politiche adottate. Questo processo dovrebbe contribuire al dibattito democratico riguardo la selezione delle migliori politiche, e rendere più responsabile il governo quando le politiche selezionate non siano risultate efficaci. Si dovrebbe poi considerare la necessità di supportare le valutazioni con nuovi modelli di analisi, che non si limitano ad adottare una prospettiva incentrata sui costi, ma che permettano di calcolare anche obiettivi non direttamente quantificabili.

Il passaggio al nuovo paradigma di *better regulation* incentrato sui big data e sull'IA può avvenire solamente se l'utilizzo delle nuove tecnologie nel ciclo regolatorio avvenga nel rispetto del diritto e dell'etica. Occorre compiere tutti gli sforzi per scongiurare una deriva antidemocratica dello stato della sorveglianza. In questo scenario, i comportamenti dei cittadini vengono monitorati in massa e in modo indiscriminato e vengono analizzati attraverso tecnologie sofisticate per prevenire il dissenso politico, per manipolare l'opinione pubblica, per discriminare nell'accesso ai servizi. Affinché ciò non avvenga, e invece l'utilizzo dei dati sia proporzionato e necessario ad attuare politiche pubbliche che aumentino il benessere dei cittadini, è fondamentale che i governi sviluppino, adottino e applichino linee guida e quadri di riferimento etici per l'utilizzo dei dati nel settore pubblico.