



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Diritto Privato e Critica del Diritto

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di Laurea Magistrale in Giurisprudenza

a.a. 2021/2022

SISTEMI ESPERTI NEL DIRITTO: RIFLESSIONI SU COMPAS E SUL CASO LOOMIS

Relatore: Ch.mo prof. Claudio Sarra

Controrelatrice: Ch.ma prof.ssa Chiara Abatangelo

Laureando: Daniele Prota

Matricola: 1117717

INDICE

Introduzione	5
<u>Capitolo 1</u> : Un'introduzione all'intelligenza artificiale nel diritto	7
- 1.1 Il rapporto tra intelligenza umana e intelligenza artificiale	9
- 1.2 Che cos'è l'intelligenza artificiale?	10
- 1.3 Il test di Turing e i <i>chatbot</i>	12
- 1.4 Le implicazioni del test di Turing e l'argomento della stanza cinese	17
- 1.5 I sistemi esperti in generale	19
o 1.5.1 La loro evoluzione	20
o 1.5.2 Cosa fanno fare i sistemi esperti	21
o 1.5.3 Le due macro-tipologie di sistemi esperti	22
o 1.5.4 La struttura dei sistemi esperti	23
o 1.5.5 La scelta del modello giuridico	26
- 1.6 Esempi di sistemi esperti legali (SEL)	28
o 1.6.1 Sistemi per l'analisi giuridica: l'Automa Infortunistico e SEFIT	29
o 1.6.2 Sistemi di supporto all'attività giudiziaria: il Progetto DAEDALUS P.M. Assistant	32
o 1.6.3 Sistemi basati sui casi: HYPO	34
o 1.6.4 Altri esempi di sistemi esperti legali	35
<u>Capitolo 2</u> : Le conseguenze giuridiche dell'uso dell'intelligenza artificiale nel diritto	37
- 2.1 L'idea del giudice robot	39
- 2.2 Un <i>software</i> può interpretare la legge?	41
- 2.3 La soggettività e la responsabilità dei robot intelligenti: la normativa attuale ...	51
o 2.3.1 La Risoluzione del Parlamento europeo del 2017	56
o 2.3.2 Le critiche degli esperti alla Risoluzione	58
o 2.3.3 Cenni sulla responsabilità penale delle macchine intelligenti	61
▪ 2.3.3.1 La tesi di Gabriel Halley	64

<u>Capitolo 3: L'intelligenza artificiale nel caso Loomis</u>	67
- 3.1 Una breve introduzione ai <i>risk assessment tool</i>	69
- 3.2 COMPAS e il caso Loomis v. Wisconsin	70
o 3.2.1 Che cos'è e come funziona COMPAS	71
o 3.2.2 La posizione della difesa e la sentenza definitiva	73
- 3.3 Le critiche a COMPAS	81
o 3.3.1 L'inchiesta di ProPublica	82
o 3.3.2 Gli studi successivi	87
- 3.4 Lo sviluppo di PATTERN e i dubbi costituzionali in merito ai <i>risk assesment tool</i>	89
o 3.4.1 Il giusto processo nel diritto costituzionale interno	90
<u>Capitolo 4: Lo stato dell'IA nel diritto in Italia e nell'Unione europea</u>	93
- 4.1 La normativa sull'IA al di là della responsabilità per danni	95
o 4.1.1 L'esperienza della Cina e degli Stati Uniti	95
o 4.1.2 Il percorso dell'Unione europea	99
▪ 4.1.2.1 La proposta di Regolamento sull'IA della Commissione europea	101
▪ 4.1.2.2 La Carta etica del CEPEJ	106
- 4.2 Esperienze moderne di uso dell'IA nel diritto in Italia.....	109
o 4.2.1 Cenni su Key Crime e XLAW	111
o 4.2.2 Come sta cambiando il diritto penale	115
Conclusioni	119
Bibliografia	121
Ringraziamenti	133

INTRODUZIONE

Questo contributo vuole mettere in luce alcuni degli aspetti più interessanti dell'intelligenza artificiale nel diritto, con specifico riferimento all'ambito penalistico e alla giustizia predittiva. Si tratta di uno dei temi più interessanti degli ultimi anni, vista la sua capacità di attraversare trasversalmente molte branche del diritto (diritto costituzionale, diritto commerciale, diritto del lavoro, diritto internazionale, diritto dell'Unione europea, ecc) e intrecciarsi con discipline come l'ingegneria e la filosofia.

Si vuole capire se le macchine dotate di intelligenza artificiale siano idonee all'utilizzo da parte dei giuristi (principalmente avvocati e magistrati) nella loro attività professionale analizzando uno dei casi più emblematici in materia: *Loomis v. Wisconsin*, che nel 2016 è stato deciso in via definitiva dalla Corte Suprema del Wisconsin. È stato necessario introdurre l'argomento tramite un capitolo volto a spiegare cosa sia l'intelligenza artificiale e in cosa consistano i sistemi esperti, degli applicativi in grado di svolgere numerose mansioni di aiuto all'essere umano (e quindi anche al giurista). Uno di questi, COMPAS, è stato usato nel caso *Loomis* per determinare automaticamente il rischio di recidiva dell'imputato ed è stato analizzato dai ricercatori per verificare l'attendibilità dei suoi *output*.

Dopo aver visto brevemente il funzionamento dei sistemi esperti, ci si è spostati sulle conseguenze che l'uso di tali macchine avrebbe in materia di responsabilità civile e penale, esponendo brevemente il contenuto della Risoluzione del Parlamento europeo P8_TA(2017)0051 del 16 febbraio 2017 "recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica" e alcune tra le posizioni più note sul versante penalistico. Visto l'uso sempre più massiccio dell'intelligenza artificiale non solo nel diritto, ma in moltissimi ambiti lavorativi e quotidiani, è sembrato utile verificare quali potrebbero essere gli strumenti giuridici per delineare un sistema di responsabilità coerente con l'avanzamento tecnologico in atto: se, quindi, sia il caso di riconoscere tale responsabilità direttamente in capo agli automi o se continuare ad attribuirli ai produttori/sviluppatori/utilizzatori.

A maggior ragione, ci si è anche chiesti se, allo stato attuale, un robot possa svolgere le mansioni di un giudice e, di conseguenza, se possa interpretare la legge. A tal fine si sono analizzate non solo le norme fondamentali sull'attività dei magistrati, ma anche il modo in cui un sistema esperto possa analizzare automaticamente i testi giuridici.

Nel terzo capitolo si affronta il caso Loomis, non limitandosi a commentare criticamente la sentenza della Corte Suprema del Wisconsin e la posizione della difesa, ma capendo anche come funziona tecnicamente COMPAS, il *risk assesment tool* oggetto di dibattito nei mesi e negli anni successivi al processo per via della presunta discriminatorietà dei suoi *output* verso le persone di colore. Si è voluto prendere come esempio questo sistema per ragionare sui possibili rischi che si correrebbero se si prendessero con troppa leggerezza tali strumenti, che necessitano di un controllo costante perché possono essere affetti da pregiudizi, pregiudicando in tal modo il senso del loro utilizzo. Se, infatti, una macchina fosse meno precisa o affidabile dell'uomo, non si vedrebbe a che scopo usarli nelle aule di giustizia chiedendo loro di prevedere cose come l'esito di una lite o il rischio di reiterazione del reato. I valori che entrano in gioco sono, *in primis*, quelli che si trovano nella Costituzione e nelle fonti europee, come il principio di non colpevolezza, di ricevere un giusto processo, di uguaglianza e, se a subire le conseguenze di un *software* difettoso fosse un imputato innocente, potrebbe anche venir meno la funzione rieducativa della pena.

Nel capitolo finale si è deciso di approfondire le normative attuali in materia di intelligenza artificiale dando uno sguardo sia alla situazione negli Stati Uniti e in Cina, sia al dibattito in corso in Europa (Unione europea e CEDU) e in Italia. In Unione europea uno degli atti più importanti che ha meritato un'attenta analisi è la proposta di Regolamento sull'IA della Commissione europea dell'aprile del 2021, mentre sul versante della CEDU si è vista la Carta etica del CEPEJ del 2018. Per finire, ci si è soffermati su alcuni sistemi esperti attualmente in uso in Italia, tra cui spiccano XLAW e Key Crime.

CAPITOLO PRIMO

UN'INTRODUZIONE ALL'INTELLIGENZA ARTIFICIALE NEL DIRITTO

SOMMARIO: 1. Il rapporto tra intelligenza umana e intelligenza artificiale - 1.2 Che cos'è l'intelligenza artificiale? - 1.3 Il test di Turing e i *chatbot* - 1.4 Le implicazioni del test di Turing e l'argomento della stanza cinese - 1.5 I sistemi esperti in generale - 1.5.1 La loro evoluzione - 1.5.2 Cosa sanno fare i sistemi esperti - 1.5.3 Le due macro-tipologie di sistemi esperti - 1.5.4 La struttura dei sistemi esperti - 1.5.5 La scelta del modello giuridico - 1.6 Esempi di sistemi esperti legali (SEL) - 1.6.1 Sistemi per l'analisi giuridica: l'Automa Infortunistico e SEFIT - 1.6.2 Sistemi di supporto all'attività giudiziaria: il Progetto DAEDALUS P.M. Assistant - 1.6.3 Sistemi basati sui casi: HYPO - 1.6.4 Altri esempi di sistemi esperti legali

1.1 Il rapporto tra intelligenza umana e intelligenza artificiale

L'obiettivo di questo contributo è approfondire il tema della giustizia predittiva analizzando gli aspetti problematici e non di COMPAS, un programma molto diffuso negli Stati Uniti in grado di predire il rischio di recidiva di un imputato. Prima di fare ciò, tuttavia, è necessario partire dalle basi per poter valutare con maggior precisione l'effettiva portata delle nuove tecnologie nel diritto.

Innanzitutto: che cos'è la giustizia predittiva? Si può definire come l'insieme delle tecniche utili a predire con successo la soluzione di un certo quesito giuridico. Per fare questo, giuristi e ingegneri si affidano a delle macchine nella convinzione di poter "insegnare" loro il diritto con un livello di precisione tale da poter ritenere affidabili le conclusioni a cui giungono. La radice di tutto è l'unione della matematica col diritto attraverso programmi che potremmo definire intelligenti.

Ed è proprio qui che inizia questa indagine: l'intelligenza. Pur mancando una nozione unitaria di intelligenza, è possibile descrivere le circostanze in cui si manifesta: adattamento all'ambiente, pensiero astratto, utilizzo efficiente di risorse limitate, comunicazione, ecc. Studiare questi aspetti ha delle conseguenze dirette anche nella comprensione e nello sviluppo dell'intelligenza artificiale: la filosofia, la matematica, l'economia, la medicina, la psicologia e la linguistica sono solo alcune delle discipline che si interessano, spesso da secoli, al tema. Basti pensare all'importanza dello studio della razionalità, del pensiero, della logica, del calcolo delle probabilità o dei metodi di apprendimento per cogliere il nesso tra intelligenza umana e artificiale¹.

La ricerca su questi ultimi due fronti va di pari passo, tanto che le nozioni di una disciplina possono tornare utili all'altra. Per esempio, le scoperte in campo neuroscientifico in merito all'elaborazione degli stimoli sensoriali del cervello umano possono avere una certa rilevanza anche nello sviluppo di robot intelligenti in grado di interagire con l'ambiente circostante, e viceversa. Nonostante questo, restano ferme alcune differenze tra i due tipi di intelligenza: quella artificiale, per esempio, è più adatta a svolgere calcoli aritmetici con

¹ G. Sartor, 2016, 279; Sul tema si veda anche S. Russel-P. Norvig, 2010

grandi numeri, mentre quella umana è più adatta a tracciare delle analogie non predeterminate².

1.2 Che cos'è l'intelligenza artificiale?

Nella recente comunicazione COM/2018/237 del 25 aprile 2018, la Commissione europea ha definito l'intelligenza artificiale nel seguente modo: “Intelligenza artificiale’ (IA) indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi”.

Secondo la Commissione possiamo definire intelligenti due macro-tipologie di sistemi: i *software* che agiscono esclusivamente nel mondo virtuale (motori di ricerca, assistenti vocali, sistemi di riconoscimento facciale, ecc.) e gli *hardware* dotati di IA che agiscono nel mondo reale (auto a guida autonoma, droni, Internet delle cose, ecc.)³.

Già nel 1980, John Searle in “*Minds, Brains and Programs*”, distingueva l'intelligenza artificiale forte da quella debole. La prima, ossia quella che ci interessa più da vicino, ha come obiettivo ricreare una vera e propria mente pensante simile in tutto e per tutto a un cervello umano. La seconda, invece, si propone di sviluppare dei sistemi capaci di simulare il comportamento umano, ma non di pensare in autonomia⁴.

I programmi e le macchine dotate di intelligenza artificiale solitamente hanno una serie di caratteristiche che le distinguono dai loro corrispettivi tradizionali: conoscenze su un determinato campo o attività, capacità di ragionamento, capacità di risoluzione dei problemi (il cd. *problem solving*), abilità percettive, di apprendimento, di pianificazione e di manipolazione di oggetti esterni. L'uso dell'intelligenza artificiale porta con sé numerosi vantaggi tra cui una migliore gestione delle informazioni e la risoluzione dei problemi in modo più rapido e con meno errori. Una macchina può eseguire compiti pericolosi o ripetitivi senza difficoltà, essendo in grado di operare in luoghi nocivi o mortali per un essere umano

² G. Sartor, 2016, 280

³ COM/2018/237 final

⁴ J. Searle, 1980, 417

(basti pensare al campo dell'estrazione mineraria o dell'esplorazione spaziale). Si tratta di *software* e *hardware* privi di emozioni che possono ragionare in modo puramente razionale, non si stancano, non si annoiano, non hanno bisogno di ferie o di svago: sono entità artificiali create al solo scopo di svolgere determinate mansioni in modo efficiente⁵.

Al netto di tutto ciò, la ricerca sul tema ha da tempo evidenziato delle zone d'ombra da non poter ignorare non solo dal punto di vista tecnologico, ma anche etico. Ci si può chiedere se sia moralmente corretto sviluppare robot con l'obiettivo di ricreare la mente e i comportamenti degli esseri umani demandando loro compiti che i ricercatori stessi non saprebbero compiere in modo tanto efficiente, consegnando loro conoscenze e informazioni di tutti i tipi. Dal punto di vista tecnico, invece, una macchina, di norma, può operare solo nei margini stabiliti dai suoi creatori, il che presuppone un costante aggiornamento dei sistemi e un continuo dispendio di risorse monetarie⁶.

Sono molteplici i campi in cui l'intelligenza artificiale viene adottata e qui è bene ricordarne alcuni: riconoscimento vocale, visivo e della scrittura, robotica, elaborazione del linguaggio naturale, sviluppo di videogiochi e di sistemi esperti. È sempre più frequente vedere applicativi comunicare con l'uomo tramite il linguaggio naturale o riconoscere i volti delle persone per ragioni di sicurezza pubblica o durante delle operazioni di polizia. Esistono da decenni sistemi esperti capaci di effettuare diagnosi mediche o di assistere gli avvocati nella loro professione, come esistono intelligenze artificiali al servizio di atleti "esportivi" in grado di simulare il comportamento di un giocatore umano (un po' come accade negli scacchi)⁷.

Un esempio di quest'ultimo tipo può essere GT Sophy, un'intelligenza artificiale sviluppata da Sony AI e inserita nel noto simulatore automobilistico per PlayStation 4, Gran Turismo Sport. È stata creata per poter competere con i migliori piloti del mondo in un ambiente virtuale molto vicino alla realtà che richiede tattica e capacità di prendere decisioni strategiche rapidamente. Per ottenere questo risultato, gli sviluppatori hanno combinato *"state-of-the-art, model-free, deep reinforcement learning algorithms with mixed-scenario*

⁵ G. Naveen-M. A. Naidu-B. T. Rao-K. Radha, 2019, 1980

⁶ G. Naveen-M. A. Naidu-B. T. Rao-K. Radha, 2019, 1981

⁷ G. Naveen-M. A. Naidu-B. T. Rao-K. Radha, 2019, 1982

training to learn an integrated control policy that combines exceptional speed with impressive tactics". Il *deep reinforcement learning* è il modo in cui Sophy impara dai propri errori, alla cui base c'è un sistema di ricompense e obiettivi che indirizza le sue scelte⁸.

1.3 Il test di Turing e i chatbot

Parlando di intelligenza artificiale forte, Alan Turing, uno dei pionieri della materia, nel 1950 mise a punto la sua versione del "gioco dell'imitazione" con lo scopo di verificare se una macchina potesse davvero essere considerata intelligente. L'esperimento originale prevedeva la presenza di tre soggetti umani: un uomo (A), una donna (B) e un interrogante (C). L'interrogante doveva capire il sesso dei suoi due interlocutori ponendo loro delle domande senza poterli vedere o sentire, mentre i due soggetti potevano cercare di ingannare C mentendo nelle risposte. Turing, nella sua versione, sostituì uno dei due interrogati con una macchina, lasciando inalterato lo svolgimento dell'esperimento⁹.

Questo test segnò un punto di svolta perché spostò l'attenzione dei ricercatori sull'aspetto linguistico dell'interazione uomo-macchina, favorendo la nascita di sistemi capaci di dialogare con gli esseri umani sfruttando basi di dati sempre più consistenti.

Pensando all'attualità, *software* come Siri, Google Assistant e Alexa, ossia assistenti virtuali in grado di offrire agli utenti una serie di servizi legati principalmente (ma non esclusivamente) alla domotica, sono frutto della visione di Turing.

Sul solco delle idee di Turing, negli anni Sessanta i ricercatori iniziarono la sperimentazione di *chatbot* intelligenti capaci, quantomeno in modo astratto, di superare il gioco dell'imitazione e quindi di ingannare il proprio interlocutore. Eliza e Parry, due bot capaci di rispondere alle domande riformulando le informazioni in esse contenute, fanno parte della prima generazione. Eliza è considerato il capostipite dei *chatbot*, è stato sviluppato tra il 1964 e il 1966 da Joseph Weizenbaum ed era la parodia di uno psicoterapeuta Rogeriano. Parry, invece, è stato programmato tra il 1970 e il 1972 dallo psichiatra Kenneth Mark Colby

⁸ P. R. Wurman-S. Barrett-K. Kawamoto *et al.*, 2022, 223

⁹ A. M. Turing, 1950, 433

ed era in grado di simulare il comportamento di un paziente schizofrenico paranoide. La particolarità di questa prima ondata di *software* è che, pur usando le informazioni lessicali e linguistiche provenienti dalle domande che venivano loro rivolte, non alteravano in alcun modo il loro *database* iniziale. Questa è una grossa differenza con i sistemi moderni che, al contrario, possono imparare dal *web* e dai loro interlocutori accrescendo, così, le loro conoscenze¹⁰.

Con ogni probabilità il tipo di macchina che aveva in mente Turing mentre metteva a punto il suo test era, come le due appena citate, del tutto autosufficiente: non le sarebbero servite delle fonti esterne per migliorare progressivamente le proprie prestazioni.

Già con la seconda generazione di *chatbot* le cose cambiano. Jabberwacky e Cleverbot sono due programmi sviluppati da Rollo Carpenter capaci di apprendere dalle interazioni con l'utente: grazie a questa loro caratteristica potevano usare le frasi, i lemmi e le risposte dette da un certo soggetto nelle conversazioni successive e, cosa non da poco, potevano cambiare il tenore delle loro risposte in base alla personalità che gli si voleva dare¹¹.

Uno studio del 2015 pubblicato su *Computers in Human Behavior* mette in luce le capacità comunicative di Cleverbot: l'obiettivo era capire se ci fossero delle differenze tra le interazioni uomo-uomo e uomo-macchina. Cleverbot era particolarmente adatto alla ricerca perché nel 2011 ottenne un punteggio notevole nel test di Turing: il 59% dei 1334 interroganti pensarono che fosse un uomo e non un bot.

Lo studio ha analizzato nel dettaglio 100 conversazioni tra esseri umani effettuate tramite programmi di messaggistica istantanea e altrettante effettuate con il *chatbot*. I parametri presi in considerazione erano: le parole, i messaggi, il rapporto tra le parole uniche e il totale di parole usate nella conversazione, le parolacce, il gergo, le emoticon e le variabili di LIWC (un gruppo di quattordici fattori, tra cui le parole per frase e la percentuale di parole con più di sei lettere).

¹⁰ G. Roncaglia, 2014, 2; J. Weizenbaum, 1966, 36-45; K. M. Colby-S. Weber-F. D. Hilf, 1971, 1-25

¹¹ G. Roncaglia, 2014, 2

I risultati hanno mostrato che i messaggi inviati al bot sono più numerosi e più brevi rispetto a quelli inviati ad altre persone. Nello specifico, le parole per messaggio nel primo caso andavano da 2 a 13, mentre nel secondo andavano da 2 a 25; il numero di messaggi, invece, andava da 19 a 248 se l'interlocutore era Cleverbot, mentre da 3 a 122 se si trattava di un essere umano. Un altro dato interessante è l'uso di turpiloquio: nelle conversazioni con la macchina le parolacce sono 30 volte più frequenti rispetto alle conversazioni con altri umani (4,29% contro lo 0,16%). È stato, inoltre, evidenziato come nell'85% delle conversazioni tra persone non c'era traccia di linguaggio scurrile, mentre nelle conversazioni con il bot la percentuale scendeva al 20%¹².

Alla luce dei dati appena riportati è possibile ipotizzare che, da un lato, le persone trattino la macchina come se fosse un bambino o un non madrelingua, adattandosi ai suoi ritmi e alle sue risposte; dall'altro si può constatare come i soggetti sembrano avere meno scrupoli comunicativi con il bot, come se volessero provare a capire i suoi limiti messaggio dopo messaggio. Insomma, una volta che un soggetto sa di star parlando con una macchina, muta, spesso di molto, il suo stile comunicativo.

Un altro *chatbot* degno di nota è A.L.I.C.E. (*Artificial Linguistic Internet Computer Entity*), un programma sviluppato tra il 2000 e il 2004 e vincitore di tre Loebner Prize, un riconoscimento dato ai programmi che si avvicinano maggiormente agli standard del test di Turing. Nel 2013, con l'uscita della versione 2.0, il *software* è diventato capace di ottenere informazioni dalla rete e dai dati forniti dalle conversazioni degli altri bot, aprendo la strada agli attuali assistenti virtuali che, a differenza di quelli ipotizzati da Turing, non stanno imparando a "barare", ossia a fingersi umani con artifici comunicativi, ma stanno progressivamente imparando le caratteristiche linguistiche tipiche dell'uomo¹³.

Al pari di A.L.I.C.E., anche Kuki (noto in passato con il nome Mitsuku) è riuscito ad aggiudicarsi molteplici volte il Loebner Prize. Sviluppato nel 2005 da Steve Worswick, Kuki non solo è raggiungibile tramite il proprio sito web, ma è presente anche sulle maggiori piattaforme social: TikTok, YouTube, Instagram, Facebook, Telegram, Twitter e Discord (il

¹² J. Hill-W. R. Ford-I. G. Farreras, 2015, 248

¹³ G. Roncaglia, 2014, 2

che lo espone a una quantità di dati inimmaginabile ai tempi di Turing). Una delle sue caratteristiche principali è la capacità di scomporre le frasi e ragionare sui singoli elementi al loro interno. Alla domanda “Puoi mangiare una casa?”, Kuki risponderebbe di no perché sa che le case sono fatte di mattoni e che i mattoni non sono commestibili¹⁴.

Uno studio del 2021 l’ha preso in esame per cercare di capire se possa instaurarsi un rapporto di amicizia tra un essere umano e un robot. I 118 partecipanti (47 uomini e 71 donne di età compresa tra i 18 e i 58 anni) hanno interagito per almeno 5 minuti con il bot ogni tre giorni per tre settimane. Questo periodo di tempo, secondo i ricercatori, era sufficiente a costruire un legame significativo con la macchina visto che, mediamente, il tempo necessario a due persone per entrare in confidenza varia dalle tre alle nove settimane. Le interazioni non erano sottoposte a delle linee guida particolari: i partecipanti dovevano semplicemente chiacchierare con Mitsuku in lingua inglese. Alla fine di ogni conversazione, ogni persona doveva compilare un questionario volto a misurare i seguenti parametri: attrazione sociale, *self-disclosure*, intimità, qualità, empatia e competenza comunicativa. Dopo l’ultima interazione, dovevano compilarne uno relativo ai *feelings of friendship* instaurati con il bot¹⁵.

Tutti gli elementi venivano valutati dando un punteggio da 1 a 5 a delle affermazioni, dove 1 indica un forte disaccordo e 5 un forte accordo:

- L’attrazione sociale veniva misurata con frasi come “È piacevole stare con Mitsuku” o “Mitsuku non è molto amichevole”;
- La *self-disclosure*, ossia la capacità di aprirsi a livello personale con il bot, veniva misurata con frasi come “Mi sono sentito a mio agio a rivelare informazioni personali durante l’interazione” o “È stato facile rivelare informazioni personali nell’interazione”;
- L’intimità veniva valutata con frasi come “Mi sono sentito coinvolto nell’interazione” o “Ho vissuto l’interazione come superficiale”;

¹⁴ E. A. J. Croes-M. L. Antheunis, 2021, 286

¹⁵ E. A. J. Croes-M. L. Antheunis, 2021, 286

- La qualità veniva misurata con frasi come “Ho apprezzato l’interazione con Mitsuku” o “Vorrei interagire con Mitsuku di nuovo”;
- L’empatia veniva misurata con affermazioni come “Mitsuku era una buona ascoltatrice” o “Mitsuku ha detto la cosa giusta per farmi sentire meglio”;
- La competenza comunicativa veniva valutata con frasi come “Mitsuku ha comunicato correttamente” o “Mitsuku si è risultata competente”;
- I *feelings of friendship*, ossia i sentimenti che provano reciprocamente due amici, venivano misurati con affermazioni come “Sarei sconvolto se non potessi più interagire con Mitsuku” o “Continuerò a interagire con Mitsuku in futuro”.

Eccezion fatta per l’intimità, tutti i fattori diminuivano nel tempo: i partecipanti erano sempre meno interessati al bot e valutavano sempre più negativamente la qualità delle loro interazioni sia da un punto di vista empatico che di semplice competenza comunicativa. L’intimità, dal canto suo, si aggirava su punteggi che andavano da 2,38 e 2,60 su 5. Quindi, nonostante le persone stessero progressivamente perdendo interesse verso Mitsuku, percepirono un livello di intimità stabile, anche se medio¹⁶.

La conclusione che si può trarre da questi dati è che probabilmente il bot non ha mai saputo approfondire la conoscenza del proprio interlocutore umano, mantenendo le conversazioni su un livello troppo superficiale. Di conseguenza, le persone, non notando un reale coinvolgimento dall’altra parte, hanno gradualmente perso interesse verso il *chatbot* e non sono state invogliate ad aprirsi con lui (come, al contrario, farebbero due amici).

Per istaurare e mantenere un rapporto di amicizia sono necessari molteplici elementi che, per ora, sono assenti in Mitsuku e in qualunque altro *chatbot*: complicità, linguaggio gergale, un certo coinvolgimento emotivo e la condivisione di esperienze significative. I robot possono al massimo far finta di essere interessati all’altra persona o di avere qualcosa di cui parlare, il che mina alla base il rapporto che si vorrebbe creare (per quanto, almeno inizialmente, i partecipanti si siano dimostrati molto curiosi di conoscere Mitsuku).

¹⁶ E. A. J. Croes-M. L. Antheunis, 2021, 292

Citando lo studio in esame: *“The main implication of the current study is that this particular chatbot Mitsuku is not humanlike enough for humans to develop feelings of friendship toward the bot. Based on the results of this study and feedback from the participants we can suggest some reasons, namely that the chatbot has no memory, lacks humor and empathy, and is too superficial.”*¹⁷.

Comprendere il modo in cui l’uomo si rapporta con le intelligenze artificiali ha una grande rilevanza anche nel mondo del diritto, visto il loro crescente utilizzo dai professionisti legali di tutto il mondo. Sapere come un giudice percepisce il programma che sta usando per decidere una controversia o come un avvocato interagisce con il suo *software* di giurimetria a livello emotivo prima ancora che tecnico, può servire a capire chi dei due ha davvero il controllo della situazione: il professionista o il robot, con tutte le conseguenze che ciò ha dal punto di vista della responsabilità delle scelte adottate.

1.4 Le implicazioni del test di Turing e l’argomento della stanza cinese

Ad oggi nessun robot ha mai superato il test di Turing ma, nonostante la meta sembri ancora lontana, è bene iniziare a riflettere sulle possibili implicazioni che porterebbe il superamento del test: cosa dimostra, davvero, il gioco dell’imitazione?

A ben vedere, alla macchina viene chiesto di fingersi intelligente e di simulare i comportamenti umani. Non si tenta di indagare il funzionamento interno della mente robotica, ossia ciò che realmente accade quando recepisce le informazioni o formula una risposta. Di conseguenza, ai fini del test, non serve che il computer sia autocosciente o che sia capace di pensiero: basta che sappia ingannare il suo interlocutore simulando una conversazione naturale.

Uno dei più noti critici dell’intelligenza artificiale forte è John Searle che, per avvalorare le sue tesi, in cui mette in dubbio l’affidabilità del test di Turing, ha ideato un esperimento mentale chiamato “argomento della stanza cinese”. In questo esperimento, una persona che

¹⁷ E. A. J. Croes-M. L. Antheunis, 2021, 293

non conosce il cinese viene messa in una stanza contenente dei fogli con degli ideogrammi e un manuale di istruzioni in lingua inglese (che il soggetto capisce perfettamente). Attraverso una fenditura riceve un foglio in cui è scritta una frase in cinese: il suo compito è rispondere scegliendo i foglietti adatti in base alle regole scritte nel manuale. Di conseguenza, pur non conoscendo il cinese e, quindi, non potendo capire ciò che sta dicendo, la persona può comunicare senza problemi con chi è al di là della porta¹⁸.

Tracciando un parallelismo con il gioco dell'imitazione così come modificato da Turing, il manuale è il programma, le persone che l'hanno scritto sono i programmatori, il soggetto è il computer, i foglietti con gli ideogrammi sono la base di dati, i fogli che il soggetto riceve dall'esterno sono le domande e il gruppetto di ideogrammi che fa uscire dalla fenditura sono le risposte. La stanza cinese, seguendo la ricostruzione di Searle, altro non sarebbe che la macchina sottoposta al test di Turing, la quale, con queste premesse, riuscirebbe a superarlo brillantemente pur non "pensando". Il soggetto, infatti, non comprende nulla di ciò che sta dicendo o delle domande che gli vengono sottoposte: si limita a eseguire le istruzioni del manuale in modo meccanico. Dall'altra parte, coloro che ricevono le risposte non sanno dire se provengono da una persona capace di comprendere il cinese o meno. Secondo Searle, così come la persona dell'esperimento non ricollega alcun significato ai simboli che utilizza, anche i calcolatori digitali non comprendono realmente cosa stanno dicendo o facendo: manipolano simboli formali secondo le regole contenute in un programma e questo non è sufficiente a stabilire se siano in qualche modo dotati di intelligenza, percezione, pensiero, ecc¹⁹.

L'argomento di Searle rimette al centro della discussione concetti come la coscienza di sé stessi, la capacità di pensare e di comprendere le proprie azioni, tutti elementi che sembrano mancare anche alle macchine che solitamente vengono definite intelligenti. In tal modo, si mette da parte l'intelligenza artificiale forte in favore di quella debole, che forse riesce a descrivere con maggiore efficacia ciò che realmente possono fare i robot moderni: imitare il comportamento umano in modo incredibilmente efficiente.

¹⁸ J. R. Searle, 1990, 26

¹⁹ J. R. Searle, 1990, 26

Lo stesso Turing ammette la complessità del problema della coscienza, ma al tempo stesso lo liquida sostenendo che: *“I do not wish to give the impression that I think there is no mystery about consciousness [...] But I do not think these mysteries necessarily need to be solved before we can answer the question with which we are concerned in this paper”*²⁰.

Questa affermazione suona paradossale perché da un lato Turing ha messo a punto un test per dimostrare che le macchine possono essere dotate di pensiero e di una mente in tutto e per tutto simile a quella umana, ma dall'altro sostiene che, ai fini della sua ricerca, non è necessario risolvere un problema come quello della coscienza, che è strettamente legato al funzionamento della mente degli esseri umani. Quindi, se un robot può pensare senza essere cosciente, in cosa si differenzerebbe esattamente il concetto di intelligenza artificiale forte da quello di intelligenza artificiale debole?

1.5 Cenni sui sistemi esperti

Una volta capito, seppur per sommi capi, che cos'è l'intelligenza artificiale e quali sono le questioni aperte sul tema, è possibile iniziare a vedere una delle sue applicazioni attualmente più usate nel mondo del diritto: i sistemi esperti.

Negli anni Settanta e Ottanta, dopo circa vent'anni dalla celebre conferenza di Dartmouth che diede il via agli studi su quella che ai tempi veniva chiamata “intelligenza automatica”, la ricerca iniziò a concentrarsi in ambiti specialistici con l'obiettivo di costruire dei sistemi esperti capaci di risolvere problemi che richiedessero una particolare competenza e l'elaborazione di ampie basi di dati²¹.

Al fine di cogliere al meglio la portata di questo particolare modo di sfruttare l'intelligenza artificiale, può essere d'aiuto paragonare un programma tradizionale con un sistema esperto. Ciò che differisce maggiormente è la loro struttura e, di conseguenza, il loro approccio ai problemi: i primi sono fondati su algoritmi, i secondi sull'esperienza e su regole euristiche. Questo li porta, di fronte allo stesso quesito, a operare in modo divergente, in

²⁰ A. M. Turing, 1950, 447

²¹ G. Sartor, 2016, 290

quanto i programmi applicheranno uno schema predeterminato di istruzioni al fine di raggiungere una conclusione certa, mentre i sistemi esperti forniranno un ventaglio di possibili soluzioni. In più, i programmi tradizionali non possono in alcun modo spiegare i loro *output* o esporre i dati su cui si basano, al contrario dei sistemi esperti, che fungono da veri e propri consulenti²²: i loro ragionamenti sono verificabili perché è possibile identificare i passaggi in cui si articolano accedendo a un *record* nel quale viene registrata la conoscenza utilizzata durante l'elaborazione in modo che sia comprensibile all'utente²³.

Se in campo medico si videro i primi risultati con lo sviluppo di Mycin, un sistema esperto capace di eseguire diagnosi mediche²⁴, lo stesso non si può dire con riferimento al diritto. Già in quegli anni emersero delle difficoltà nella programmazione di *software* in grado di sostituire il giurista umano poiché l'attività di un professionista legale non si esaurisce nell'applicazione di conoscenze pregresse, ma richiede competenze ulteriori: l'analisi del caso di specie, la ricerca di informazioni rilevanti, la valutazione delle analogie con dei casi precedenti, la formulazione di ipotesi ragionevoli, la comprensione delle coerenze e delle incoerenze in un ragionamento, ecc. Queste sono tutte abilità difficilmente replicabili da una macchina, ragion per cui si è dimostrato più agevole sviluppare dei sistemi di supporto al giurista piuttosto che dei programmi in grado di sostituirlo integralmente.²⁵

1.5.1 La loro evoluzione

Le prime due generazioni di sistemi esperti introdotte tra la fine degli anni Sessanta e gli anni Ottanta mostravano notevoli lacune e limitazioni che le rendevano poco adatte all'uso al di fuori dei laboratori di ricerca.

²² A. Contaldo-F. Campara, 2020, 21

²³ P. L. M. Lucatuorto, 2006, 4; Le spiegazioni non solo devono presentarsi in linguaggio naturale, ma devono anche adattarsi al grado di conoscenza del dominio posseduto dall'utente, impiegando termini e forme espressive a lui comprensibili: le spiegazioni consistenti in un mero *reply* del ragionamento dell'esperto nel suo gergo sono inutili. Lo stesso discorso vale per le spiegazioni troppo sintetiche o troppo dettagliate.

²⁴ E. H. Shortliffe, 1976, xiii

²⁵ G. Sartor, 2016, 290

La prima generazione era basata su una rigida logica booleana secondo uno schema “vero/falso”, mentre la seconda seguiva un modello probabilistico che analizzava le cause di un problema e ne indicava i possibili effetti. In generale, entrambi i modelli si sono dimostrati o troppo limitati rispetto agli esperti umani o poco utilizzabili nella realtà. È dalla terza generazione, a partire dagli anni Ottanta, che le cose iniziano a migliorare. La svolta arrivò grazie allo sviluppo della cd. *fuzzy logic* (logica sfumata e approssimata) applicata ai processi inferenziali. I nuovi sistemi esperti basati su questa tecnologia presero il nome di *decision-support system* (o DSS) e, come dice il nome, aiutano l’utente durante il processo decisionale. Al contrario dei sistemi classici, quindi, non danno soluzioni nette ed è proprio da questo dettaglio che si vede il cambio di prospettiva: non è più la macchina a decidere, ma è l’uomo che usa le informazioni e le valutazioni della macchina per poter prendere una decisione più circostanziata²⁶.

Un esempio di sistema esperto di questo tipo è COMPAS, il *software* statunitense in grado di calcolare il rischio di recidiva di un imputato citato all’inizio di questo contributo, attorno al quale è nato un acceso dibattito tra gli esperti. Grazie alla *fuzzy logic* si esce completamente dal paradigma classico del “vero/falso” e si abbraccia un’ottica più sfumata e rispondente alla realtà dei problemi sottoposti alla macchina, per cui una cosa può essere, per esempio, parzialmente vera e falsa allo stesso tempo. Questi sistemi possono al massimo cercare di avvicinarsi al “vero” e al “falso”, al contrario di ciò che provavano a fare le generazioni precedenti (con risultati poco promettenti).

1.5.2 Cosa sanno fare i sistemi esperti

Pur non potendo sostituirsi agli esseri umani nei processi decisionali, i sistemi esperti si sono dimostrati adatti a svolgere un gran numero di mansioni.

Come riportato in uno studio del 2019 pubblicato nell’*International Research Journal of Engineering and Technology* (IRJET), i sistemi più moderni possono dare consigli, dimostrare e spiegare le conclusioni a cui arrivano, effettuare diagnosi, interpretare dati, giustificare le

²⁶ A. Contaldo-F. Campara, 2020, 37

soluzioni che reputano corrette e persino predire le ripercussioni delle scelte che si intendono adottare. In aggiunta, i sistemi esperti possono lavorare in ambienti pericolosi per l'uomo, non provano fatica e stress, sbagliano generalmente meno di un professionista umano e sono molto veloci nello svolgimento delle loro mansioni. Questo non significa che possano risolvere ogni problema con facilità o che possano operare senza la supervisione di un soggetto umano, anzi. Non bisogna dimenticare i limiti della tecnologia utilizzata per sviluppare i sistemi, la difficoltà di reperire le informazioni da usare come base di conoscenza, gli elevati costi di sviluppo (che col passare del tempo diventano sempre più contenuti) e la generale difficoltà di mantenere i sistemi aggiornati ed efficienti nel tempo²⁷.

Attualmente, i sistemi esperti vengono utilizzati con regolarità, oltre che nel diritto, nel campo delle diagnosi mediche, in quello petrolchimico al fine di individuare anomalie e guasti in grandi impianti e in quello finanziario per valutare i rischi di determinate attività, soprattutto in materia di mutui e tributi²⁸.

1.5.3 Le due macro-tipologie di sistemi esperti

Esistono due macro-tipologie di sistemi esperti:

- Sistemi esperti basati su regole
- Sistemi esperti basati su alberi

I primi adottano uno degli schemi più utilizzati nell'informatica, *if-then*, per cui "se" si realizza una certa condizione, l'antecedente, "allora" avrà luogo una certa azione, la conseguente. Un esempio in ambito medico può aiutare la comprensione del meccanismo: se un soggetto è affetto da vari sintomi e vuole scoprire quale sia il suo problema di salute, fornisce alla macchina tutta una serie di informazioni in merito alla propria sintomatologia (mal di testa, raffreddore, febbre) e, in automatico, il sistema svolgerà una serie di passaggi per ottenere una diagnosi. Sommando tutte le informazioni che gli sono state date, deduce che, con buona probabilità, il soggetto ha l'influenza. I secondi usano i dati a loro

²⁷ G. Naveen-M. A. Naidu-B. T. Rao-K. Radha, 2019, 1983

²⁸ A. Contaldo-F. Campara, 2020, 39

disposizione per creare un albero composto da tutte le alternative possibili. Il sistema riesce a riconoscere e risolvere i problemi in base a una sequenza di informazioni: partendo da una situazione iniziale, ogni scelta ramifica il ragionamento fino al raggiungimento della soluzione²⁹.

In entrambi i casi, ci si trova davanti a una concatenazione di inferenze che generano dei *matching*: il sistema confronta l'antecedente di ciascuna regola e i fatti presenti nella sua memoria temporanea. Questo permette all'utente di ricostruire i passaggi logici che la macchina ha effettuato per raggiungere una certa conclusione e quindi capirne le motivazioni³⁰.

Le catene inferenziali si generano in due modi:

- Concatenamento in avanti (*forward chain*)
- Concatenamento all'indietro (*backward chain*)

La *forward chain* parte dai dati iniziali del problema e applica le regole di concatenazione in avanti, fino al raggiungimento di una conclusione. La *backward chain*, al contrario, parte da un obiettivo e cerca di dimostrare la sua coerenza con i dati iniziali utilizzando delle regole di produzione all'indietro³¹.

1.5.4 La struttura dei sistemi esperti

Generalmente i sistemi esperti sono composti da tre elementi:

- Base di conoscenza (*knowledge base*)
- Motore inferenziale (*inference engine*)
- Interfaccia utente (*user interface*)

La base di conoscenza consiste in un particolare *database* al cui interno sono contenute tutte le informazioni relative al dominio (ossia l'ambito) di cui si deve occupare il sistema: in

²⁹ A. Contaldo-F. Campara, 2020, 25

³⁰ L. Ferrari, 1994, 6

³¹ L. Ferrari, 1994, 6

questo modo la macchina entra in possesso dei dati necessari per ragionare e risolvere i problemi che le vengono sottoposti³². È un modulo del tutto distinto e separato dalle altre conoscenze di cui è in possesso il *software*, come quella per comprendere e dare risposta ai quesiti espressi con un certo formalismo o quella contenente le informazioni su come comunicare efficacemente con l'utente³³.

Pensando al diritto, una *knowledge base* deve contenere tutte le fonti di un certo ordinamento o settore (normativa interna, sovranazionale e internazionale), così come la dottrina e la giurisprudenza. Deve tener conto anche della tradizione giuridica dello Stato in cui opera: in un paese di *common law* la giurisprudenza avrà una rilevanza maggiore rispetto a un paese di *civil law*, che, al contrario, darà più importanza a elementi come la normativa vigente.

Una legge può contenere molteplici informazioni che il sistema deve sapere e riconoscere: enunciati normativi, definizioni, rinvii, enunciati procedurali, criteri gerarchici, enunciati politici o valutativi in generale e altri tipi di enunciati difficilmente classificabili³⁴. Da questo si può intuire la delicatezza della base di conoscenza che, se viene composta in modo scorretto, può portare a dei risultati paradossali, magari attribuendo valore definitorio a una regola processuale o invertendo la gerarchia delle fonti del diritto.

Solitamente la *knowledge base* è suddivisa in due blocchi: la memoria temporanea e la memoria a lungo termine. Citando le parole di Laura Ferrari: "La memoria temporanea contiene la conoscenza dichiarativa o asserzionale su un particolare problema da risolvere e la situazione corrente del tentativo di soluzione. Nella memoria a lungo termine sono contenute le dipendenze o relazioni tra i fatti elementari, espresse sotto forma di regole o di altro tipo di rappresentazione della conoscenza". In altre parole, la prima si occupa del quesito che gli viene sottoposto in un certo momento, mentre la seconda collega le informazioni di cui è in possesso³⁵.

³² P. L. M. Lucatuorto, 2006, 4

³³ A. Contaldo-F. Campara, 2020, 27

³⁴ A. A. Martino, 1988, 11

³⁵ L. Ferrari, 1994, 6

Del motore inferenziale si è già parlato incidentalmente nel paragrafo precedente introducendo i concetti di concatenazione in avanti e all'indietro, ossia le due modalità in cui la macchina può eseguire le deduzioni con cui risolvere i problemi. Più in generale, il motore inferenziale è identificabile come il cuore del sistema esperto, la componente con cui il programma sceglie quali azioni eseguire, in quale ordine disporle e quali informazioni della base di conoscenza utilizzare³⁶. Come ricordato nelle pagine precedenti, è grazie a queste deduzioni logiche che è possibile ricostruire i ragionamenti della macchina³⁷.

È composto da due parti: uno schedatore e un interprete. Il primo sceglie quali azioni eseguire e in quale ordine (secondo un criterio di priorità), mentre il secondo mette in atto le istruzioni dello schedatore attingendo alla *knowledge base* di cui dispone³⁸.

La struttura del motore inferenziale viene adattata alle esigenze che il sistema deve soddisfare: un sistema esperto per professionisti legali sarà influenzato sia dal linguaggio di programmazione con cui è stato creato che dai ragionamenti giuridici usati come punto di riferimento per le inferenze. Un esempio può aiutare a capire questo passaggio: la logica deduttiva tipicamente rigida del programma non gli permetterebbe di procedere oltre la sua base di conoscenza, lasciando insoddisfatti eventuali quesiti che superassero tale limite, ma grazie all'apporto del ragionamento giuridico, questa lacuna può essere superata con l'applicazione di concetti come l'analogia e l'interpretazione estensiva³⁹.

Il terzo e ultimo elemento che compone i sistemi esperti è l'interfaccia utente, ciò che permette all'utente di interagire con il programma ponendo delle domande e ricevendo delle risposte⁴⁰. Queste interazioni, con cui è possibile fornire nuovi dati al sistema, possono avvenire o tramite menù o con linguaggio naturale: il primo metodo è più semplice da mettere e punto perché si evita di dover "insegnare" alla macchina un linguaggio che non

³⁶ P. L. M. Lucatuorto, 2006, 4

³⁷ L. Ferrari, 1994, 4

³⁸ L. Ferrari, 1994, 4

³⁹ A. Contaldo-F. Campara, 2020, 29

⁴⁰ M. Fini-P. Milani, 2005, 16

conosce, mentre il secondo è decisamente più immediato per l'utente, che può dialogare con il sistema in modo diretto senza dover usare dei menù⁴¹.

1.5.5 La scelta del modello giuridico

La costruzione di un sistema esperto prevede lo svolgimento di numerosi passaggi: si inizia con l'identificazione, la concettualizzazione e la formalizzazione di un determinato dominio e delle relative informazioni, per poi arrivare alle fasi di realizzazione, collaudo e valutazione dei risultati che tale sistema produce⁴².

Uno dei momenti cruciali si colloca all'inizio di questo processo: la formalizzazione, ossia la trascrizione delle conoscenze acquisite dall'esperto del dominio. Ciò è possibile mediante il ricorso a un formalismo concordato e tramite l'uso di un *tool* di sviluppo che traduce e inserisce le informazioni nella macchina in un linguaggio a lei comprensibile, il che può generare molti problemi se viene fatto in modo poco preciso.

Uno dei presupposti della formalizzazione è la scelta di un modello giuridico da seguire, ossia la struttura di cui dotare i contenuti del sistema applicando concetti tipici della filosofia e della teoria del diritto. Ma non solo. Applicare un modello piuttosto che un altro, influisce anche nella selezione dei contenuti da formalizzare nella base di conoscenza, e a tal fine è necessario stabilire quali informazioni siano necessarie e sufficienti per risolvere i problemi giuridici che il sistema deve affrontare. Questa operazione identifica cosa debba ritenersi "diritto" al fine dello sviluppo del sistema e, in aggiunta, quali conoscenze extragiuridiche siano connesse con le prime (e quindi debbano essere inserite nel *software*)⁴³.

Tra i modelli più risalenti troviamo i normativistici e decisionistici. I primi riprendono la tradizione positivista del diritto assumendo che l'ordinamento giuridico consista in un insieme di norme generali e astratte e che, di conseguenza, le basi di conoscenza debbano

⁴¹ L. Ferrari, 1994, 7

⁴² A. Contaldo-F. Campara, 2020, 31

⁴³ G. Sartor, 1996, 57

essere composte da enunciati generali del tipo “se A allora B”, dove A è la condizione (la fattispecie) e B è la conseguenza (l’effetto giuridico)⁴⁴.

I limiti del normativismo sono sostanzialmente gli stessi dell’antica impostazione positivista del diritto: la scarsa attinenza con la realtà a cui si aggiunge la dubbia utilità di un sistema progettato senza tenere conto di elementi ulteriori rispetto alle semplici leggi. Se il diritto non si esaurisce nelle disposizioni legislative, non si vede perché una macchina che dovrebbe essere più efficiente e precisa dell’uomo debba prendere in considerazione solo (o quasi) fonti di quel tipo.

In aperta polemica con i modelli normativistici, troviamo i modelli decisionistici ispirati al realismo giuridico statunitense. Il diritto non è più visto come un insieme di norme generali e astratte, ma come un insieme di decisioni individuali e concrete. La base di conoscenza è, quindi, formata principalmente da sentenze, mentre i ragionamenti della macchina sono volti alla ricerca e all’applicazione di casi simili a quello che si sta risolvendo. Un sistema di questo tipo è progettato per applicare l’analogia, l’estensione al nuovo caso della soluzione già adottata nei precedenti⁴⁵.

Ci si trova davanti ad applicativi utili allo svolgimento di mansioni molto specifiche, come la ricerca e la selezione di materiale giurisprudenziale, e ancorati agli ordinamenti di diritto anglosassone, con la conseguenza che avrebbero poco terreno fertile in quelli di diritto romano-germanico.

Tra i modelli tradizionali troviamo anche quelli istituzionalistici che, come dice il nome, adottano come punto di partenza teorico le dottrine istituzionalistiche del diritto, le quali vedono il fenomeno giuridico come qualcosa di complesso che non si può ridurre alla mera elencazione di norme di comportamento. Conseguentemente, i sistemi di questo tipo non contengono solo dati strettamente giuridici, ma anche elementi che derivano dal contesto socio-istituzionale nel quale le regole vengono applicate e le liti vengono risolte. Alcuni, per esempio, considerano le qualità degli attori implicati in una certa controversia (le parti, i

⁴⁴ G. Sartor, 1996, 59

⁴⁵ G. Sartor, 1996, 64

giudici e gli avvocati) al fine di anticipare l'esito della lite. Il grande difetto di questi modelli è, paradossalmente, la loro eccessiva complessità: ad oggi è quasi impossibile formalizzare una tale mole di dati realizzando sistemi esperti performanti⁴⁶.

Tra i modelli di recente teorizzazione ci sono quelli sincretistici: il loro obiettivo è combinare più orientamenti informatici e teorico-giuridici, andando a mutuare le proprie caratteristiche dai modelli precedenti. Una delle proposte più interessanti è l'ibridazione tra i sistemi a base normativistica e decisionistica, costruendo in tal modo dei programmi capaci di usare i precedenti giurisprudenziali quando vengono meno le regole di origine legislativa. Oltre a questi, negli ultimi anni sono stati realizzati dei modelli cd. dell'argomentazione giuridica, che, al contrario di quelli visti fin qui, non sono volti allo sviluppo di sistemi assiomatici, ma di macchine capaci di produrre e comparare argomentazioni e contro argomentazioni. Infine, i cd. modelli dell'intuizione giuridica operano senza avere una base di conoscenza manipolabile in base a regole predefinite: i loro comportamenti non sono razionalizzabili poiché muovono da delle intuizioni frutto dell'esperienza. Volendone trovare un'origine filosofica, questi ultimi sembrano attingere alla corrente di pensiero *fact sceptic* nordamericana o ad alcuni filoni dell'ermeneutica giuridica della seconda metà degli anni Novanta⁴⁷.

1.6 Esempi di sistemi esperti legali (SEL)

Esistono molti tipi di sistemi esperti legali: da quelli a supporto dell'attività giudiziaria a quelli per il reperimento di informazioni giuridiche, da quelli per la redazione di atti a quelli per l'analisi giuridica. Indipendentemente dal loro impiego, tutti presuppongono due condizioni: il processo di elaborazione giuridica deve essere sufficientemente razionalizzabile e le fonti del diritto devono essere individuabili e traducibili in un linguaggio manipolabile dalla macchina⁴⁸.

⁴⁶ G. Sartor, 1996, 66

⁴⁷ G. Sartor, 1996, 67

⁴⁸ P. L. M. Lucatuorto, 2006, 5

Gli assiomi appena citati sono tutt'altro che pacifici, vista la complessità non solo degli ordinamenti giuridici ma anche (e forse soprattutto) del linguaggio a volte eccessivamente articolato di sentenze, leggi, regolamenti e, più in generale, testi normativi, che richiedono uno sforzo interpretativo notevole anche ai giuristi umani.

La difficoltà di queste operazioni aumenta considerando le caratteristiche degli elementi che compongono il patrimonio giuridico del sistema: da un lato abbiamo le conoscenze pubbliche, ossia le fonti scritte, mentre dall'altro troviamo le conoscenze private, ossia le informazioni derivanti dall'esperienza. Accanto a queste, ci sono anche le cd. conoscenze inconsce, ossia tutte quelle nozioni che non sono né pubbliche né private, come quelle linguistiche: tutti sono in grado di parlare ma non è stata ancora sviluppata una grammatica completa di alcun linguaggio naturale. Rappresentano uno dei più grandi problemi dell'intelligenza artificiale poiché sono difficilmente formalizzabili e rappresentabili⁴⁹.

La macchina deve possedere tutte queste conoscenze perché ciò che le viene chiesto non è solamente rappresentare su schermo un certo documento presente nella sua memoria, ma ci si aspetta che sia in grado di eseguire calcoli e deduzioni logiche sulle informazioni di cui è dotata⁵⁰.

1.6.1 Sistemi per l'analisi giuridica: l'Automa Infortunistico e SEFIT

In Italia i sistemi per l'analisi giuridica sono stati sviluppati a partire dalla metà degli anni Settanta e servono a qualificare giuridicamente un fatto, per esempio determinando se un contratto sia nullo o se sussistano i presupposti per dichiarare l'inabilitazione di una persona fisica. Sono per loro natura interattivi, funzionano dialogando con l'utente che deve porre il quesito nel modo più preciso possibile e sono indirizzati tanto ai giuristi "generici" privi di una specializzazione nel settore in cui rientra il caso, quanto alle pubbliche

⁴⁹ G. Sartor, 1990, 126; Tra le conoscenze private si trovano la maggior parte delle regole di tipo euristico, che permettono all'esperto di prendere decisioni anche in situazioni con poche informazioni.

⁵⁰ P. L. M. Lucatuorto, 2006, 5

amministrazioni, che possono usarli per calcolare i tributi, salari, pensioni, ecc⁵¹. La base di conoscenza è generalmente costituita dalla normativa di un certo settore, anche se capita che venga integrata da concetti giuridici e casistiche giurisprudenziali⁵².

Tra gli esempi che si possono fare (Methodus, Sefit, Lexis, Iri, Remida, Daedalus-Cassazione, ecc.), l'Automa Infortunistico merita un particolare approfondimento. Realizzato tra il 1974 e il 1976 presso l'Istituto per la Documentazione Giuridica di Firenze, l'Automa aveva come obiettivo offrire consulenze giuridiche in materia di risarcimento del danno patrimoniale da sinistro stradale direttamente al cittadino: attraverso l'interazione uomo-macchina, il sistema riusciva a capire gli aspetti di fatto del caso e, conseguentemente, applicare le relative norme. Essendo tecnologicamente molto limitato (l'IA al suo interno operava in modo piuttosto meccanico), ciò che ci si poteva aspettare da una sua consulenza era un mero aiuto a orientarsi con più consapevolezza sul da farsi in caso di sinistro stradale, non era di certo in grado di competere con l'attività professionale di un giurista umano⁵³.

I ricercatori volevano costruire un programma che partisse da dei presupposti opposti rispetto a quelli portati avanti fino a quel momento nel resto del mondo. Nell'articolo in cui viene presentato il progetto, Elio Fameli, per dimostrare questo punto, fa riferimento a quattro *software* già esistenti all'epoca: uno sviluppato dall'Accademia cecoslovacca delle scienze capace di quantificare il mantenimento che i genitori dovevano versare ai figli, uno messo a punto dall'Accademia delle scienze dell'URSS usato in materia di licenziamento su richiesta del lavoratore, uno progettato tra il 1966 e il 1968 dai ricercatori giapponesi dell'Università Tohoku a Sendai con lo scopo di quantificare le quote ereditarie in capo ai successori e, per finire, uno nato negli Stati Uniti grazie al quale gli avvocati potevano fornire delle consulenze più circostanziate in materia di divorzio grazie a un sistema che, tramite un questionario elettronico, filtrava e organizzava le informazioni provenienti dai clienti⁵⁴.

⁵¹ P. L. M. Lucatuorto, 2006, 6; Si veda anche C. Cevenini, 2004 nella parte in cui afferma che: "L'introduzione di sistemi esperti nell'ambito di funzioni amministrative [...] rappresenta una profonda innovazione rispetto ai modelli di *e-government* ancor oggi prevalenti nel nostro paese, consentendo di conciliare l'applicazione distribuita (periferica) di norme giuridiche con l'uniformità, il rigore e l'efficienza dei processi amministrativi".

⁵² A. Contaldo-F. Campara, 2020, 51

⁵³ A. Contaldo-F. Campara, 2020, 51

⁵⁴ E. Fameli, 1976, 2

In tutti questi progetti i ricercatori si mettevano dal punto di vista del professionista con l'esigenza di migliorare la qualità del suo lavoro tramite delle macchine capaci di applicare la legge in modo più razionale, limitando quanto più possibile la discrezionalità del professionista (giudice, avvocato, ecc.). Un altro aspetto da notare è che all'epoca, nello sviluppo di tali sistemi, non si mettevano in conto eventuali problemi di interpretazione della norma o di verifica dei presupposti per la sua applicazione: gli studiosi cecoslovacchi, nel formulare l'equazione per quantificare l'obbligo dei genitori di mantenere i figli, presuppongono tanto l'esistenza dell'obbligazione quanto la titolarità delle situazioni attive e passive in capo alle parti. Lo stesso discorso si può estendere, *mutatis mutandis*, alle altre tre esperienze citate⁵⁵.

Con riferimento a quest'ultimo aspetto, l'Automa e i sistemi a lui analoghi (soprattutto i più risalenti e meno performanti), lavoravano meglio con un linguaggio giuridico talmente formalizzato da risultare quasi completamente scevro di elementi non razionali. L'idea, di nuovo, è porre un argine che potremmo definire "matematico" alla discrezionalità del professionista legale, tuttavia, come ricorda Fameli: "Contraddittoriamente e inevitabilmente, però, esperienze del genere sono tanto più valide quanto più tengono conto dell'imprescindibile esigenza di non ridurre il diritto, e quindi anche il linguaggio giuridico, alla sola dimensione della razionalità"⁵⁶. Questo monito aiuta a valutare con più consapevolezza l'effettiva portata dei sistemi moderni e, in un certo senso, dovrebbe aiutare i ricercatori a evitare di inseguire modelli di diritto irrealistici. Ciononostante, sviluppare sistemi come questi, in cui la base di conoscenza è in larga parte formata da fonti normative, permette di mettere in risalto e affrontare tutte le ambiguità e le lacune di un certo settore dell'ordinamento.

Più o meno nello stesso periodo dell'Automa Infortunistico, fu sviluppato anche SEFIT, un sistema esperto messo a punto dall'Istituto di Studi sulla Ricerca e Documentazione Scientifica del CNR allo scopo di facilitare l'accesso al Fondo per l'Innovazione Tecnologica (FIT) per gli operatori economici, in particolare le piccole-medie imprese. Anche in questo

⁵⁵ E. Fameli, 1976, 5

⁵⁶ E. Fameli, 1976, 7

caso i destinatari non erano i professionisti legali, ma gli utenti “comuni” che volevano soddisfare un’esigenza molto precisa: orientarsi agevolmente nelle diverse procedure previste per l’accesso al Fondo ed essere assistiti nella valutazione delle opzioni che la normativa rendeva loro disponibili⁵⁷.

Ben lungi dall’essere un mero programma di consultazione normativa, SEFIT permetteva alla sua utenza di orientarsi nell’ambito in questione in breve tempo, in maniera completa, interattiva e potendo avere tutti i chiarimenti necessari per poter prendere una decisione circostanziata.

Più nello specifico, i ricercatori si erano posti quattro obiettivi:

- Valutare le condizioni di ammissibilità per accedere ai finanziamenti;
- Indicare tutte le formalità da adempiere nella procedura;
- Valutare la compatibilità dell’impresa del cliente con la disciplina sul FIT;
- Valutare quale delle due forme di finanziamento (mutuo agevolato e contributo a fondo perduto) fosse più conveniente per l’impresa⁵⁸.

È proprio quest’ultimo punto ad essere uno dei più interessanti perché il sistema non solo disponeva dell’intera normativa di settore (leggi, decreti, ecc.), ma aveva conoscenze economico-finanziarie sufficientemente complete per permettergli di effettuare una scelta extravagante rispetto alla stretta applicazione del diritto.

1.6.2 Sistemi di supporto all’attività giudiziaria: il Progetto DAEDALUS P.M. Assistant

I cosiddetti *Decision-Support Systems* o DSS sono strumenti pensati per aiutare PM e giudici senza dare loro delle risposte definite come farebbero i sistemi di analisi, ma fornendo loro assistenza durante l’intero processo decisionale senza indicarne previamente l’esito⁵⁹.

⁵⁷ A. Valente, 1988, 225

⁵⁸ A. Valente, 1988, 228

⁵⁹ P. L. M. Lucatuorto, 2006, 7

Uno degli esempi più illustri in materia di DSS è il Progetto DAEDALUS ideato dal magistrato dott. Carmelo Asaro e in grado di assistere il pubblico ministero durante la sua attività processuale e preprocessuale. Tra le numerose attività che può eseguire, durante la fase delle indagini preliminari può controllare la legittimità degli atti tipici posti in essere dal PM verificando la sussistenza di quattro presupposti:

- L'atto deve essere emanato nei confronti di una persona sottoposta alle indagini;
- L'atto deve essere emanato nei confronti di una persona iscritta nel registro degli indagati;
- L'atto deve essere utile all'accertamento del reato;
- L'atto deve essere ammissibile per l'accertamento di quel particolare reato⁶⁰.

Ancora, DAEDALUS può essere utile anche nella formulazione dell'imputazione al termine delle indagini e nella verifica del materiale probatorio a disposizione del PM durante l'udienza preliminare. Il sistema, in tal modo, si configura come un vero e proprio assistente digitale del magistrato requirente, un programma capace di agire nei momenti più delicati della vicenda processuale verificando la correttezza di determinate scelte, calcolando termini, aiutando nella redazione di atti e organizzando informazioni. Può anche aiutare a prendere delle decisioni limitandone gli errori avvisando l'utente della loro presenza, sarà poi il PM a valutare se procedere comunque con la scelta ritenuta errata dalla macchina o se seguirne il consiglio che sarà, come accade di norma nei sistemi esperti, argomentato e analizzabile punto per punto. Può anche rappresentare l'interpretazione di una determinata normativa applicando le informazioni di cui dispone la sua base di conoscenza o tenendo in considerazione anche quelle fornite dal PM⁶¹.

Vista la versatilità di DAEDALUS, dopo qualche anno dalla sua presentazione al pubblico anche i giudici iniziarono a sperimentarne l'utilizzo, portando alla nascita del Progetto DAEDALUS-Cassazione: un applicativo nato dall'esigenza, avanzata dalle Sezioni Penali della

⁶⁰ C. Asaro-A. Martino-E.Nissan, 2002, 5

⁶¹ M. Iaselli, 2001

Corte di Cassazione, capace di calcolare i termini di scadenza delle misure cautelari e di prescrizione dei reati, nonché per la gestione dei dati afferenti ai reati⁶².

1.6.3 Sistemi basati sui casi: HYPO

Diffusi soprattutto nei paesi di *common law*, i sistemi basati sui casi sono abili nella ricerca di materiale giurisprudenziale da usare in casi simili a quelli già decisi come strumento argomentativo per dimostrare un'analogia, come base utile a giustificare una certa interpretazione di un concetto vago o, se ci si trova in un ordinamento di *civil law*, come mezzo per colmare una lacuna legislativa⁶³.

Il sistema HYPO, realizzato alla fine degli anni Ottanta da Kevin D. Ashley e E. L. Rissland presso l'università del Massachusetts, rappresenta uno degli esempi più celebri di applicativi capaci di ricercare precedenti da utilizzare nell'argomentazione giuridica, specie nel campo dei segreti commerciali nei settori della tutela del *software* e del diritto dei contratti⁶⁴.

Una delle cose più sorprendenti è che non solo è in grado di suggerire argomenti che possono essere usati a sostegno della propria tesi, ma può fare anche il contrario, ossia indicare quali potrebbero essere gli argomenti adottati dalla controparte, fornendo delle indicazioni per indebolirli. Per arrivare a questo risultato, una volta descritto il caso, HYPO lo analizza per determinare quali sono gli aspetti giuridicamente rilevanti sulla base delle motivazioni dei precedenti: dopo aver trovato i casi simili, li ordina a seconda che siano a favore o contro e, su queste basi, produce lo schema di un'argomentazione⁶⁵.

Un altro aspetto interessante del *software* è la sua capacità di generare casi ipotetici, individuando ulteriori elementi di fatto potenzialmente rilevanti o ricostruzioni alternative dei fatti utilizzabili dall'utente nella messa a punto delle proprie argomentazioni. Durante il processo, infatti, il giurista è chiamato a considerare tutte le ricostruzioni plausibili dei fatti,

⁶² M. Iaselli, 2001

⁶³ D. Tiscornia, 2002, 137

⁶⁴ G. Sartor, 1990, 93

⁶⁵ G. Sartor, 1990, 94

non solo quella corrispondente al vero: un conto è la verità di un'affermazione e un altro è la possibilità di provarla⁶⁶.

1.6.4 Altri esempi di sistemi esperti legali

Accanto ai progetti sopracitati, esistono altre tipologie di sistemi esperti giuridici che è bene ricordare brevemente per esigenze di completezza espositiva.

Nei sistemi di reperimento concettuale di informazioni giuridiche, un particolare settore del diritto viene ricondotto a un certo modello concettuale in modo da poter costruire una base di dati contenente le informazioni relative a quello stesso settore. Ciò che il sistema fa è confrontare le informazioni inserite dall'utente con quelle presenti nella base di dati: il risultato prodotto da questo confronto, è l'informazione che serve all'utilizzatore. Di recente sono stati messi a punto dei modelli semantici che consentono al *software* di comprendere il significato dei testi e persino i ragionamenti che hanno portato l'utente a formulare la domanda⁶⁷.

I sistemi di redazione assistita sono molto elementari, non richiedono un'IA particolarmente raffinata e sono utilizzati per preparare documenti giuridici come contratti, atti amministrativi, ecc. A differenza dei programmi tradizionali che consentono di automatizzare l'inserimento di alcune informazioni (per esempio, i dati sulle parti in citazioni, ordinanze, precetti, ecc), i sistemi di redazione assistita possono aggregare automaticamente le informazioni sul contenuto dei documenti⁶⁸.

L'ultima categoria da citare è rappresentata dai programmi impiegati a fini didattici usati in ambito scolastico, universitario e per l'aggiornamento di soggetti già inseriti o da inserire nel mondo del lavoro. Il loro sviluppo è assai delicato, in quanto devono poter contare su un gran numero di strategie didattiche, buone capacità di interazione con gli allievi e su una

⁶⁶ G. Sartor, 1990, 95

⁶⁷ P. L. M. Lucatuorto, 2006, 8

⁶⁸ D. Tiscornia, 2002, 154

flessibilità tale da potersi adattare al grado di preparazione del singolo studente, il che presuppone la capacità di valutare gli errori e il livello dell'utente⁶⁹.

⁶⁹ P. L. M. Lucatuorto, 2006, 9

CAPITOLO SECONDO

LE CONSEGUENZE GIURIDICHE DELL'USO DELL'INTELLIGENZA ARTIFICIALE NEL DIRITTO

SOMMARIO: 2.1 L'idea del giudice robot - 2.2 Un *software* può interpretare la legge? - 2.3 La soggettività e la responsabilità dei robot intelligenti: la normativa attuale - 2.3.1 La Risoluzione del Parlamento europeo del 2017 - 2.3.2 Le critiche degli esperti alla Risoluzione - 2.3.3 Cenni sulla responsabilità penale delle macchine intelligenti

2.1 L'idea del giudice robot

In questo capitolo si vogliono analizzare alcuni tra i temi più interessanti che legano l'intelligenza artificiale al diritto, tra cui l'idea del giudice robot (e la sua conformità con l'ordinamento attuale) e la responsabilità di tali artefatti robotici.

A ben vedere, l'idea che sta alla base di alcuni sistemi esperti, non è nuova nel panorama giuridico. Già Leibniz, nel 1666, aveva ipotizzato un cambio di paradigma nel sistema processuale: secondo il filosofo tedesco, il diritto deve far proprio un linguaggio semplice e immediato grazie al quale risolvere agilmente le controversie calcolando le ragioni e i torti. Il punto focale della sua visione sta nel "calcolo" dell'esito della lite: non si parla più di interpretazione della legge, ma di operazioni matematiche che portano a una decisione⁷⁰. Come riporta Luigi Viola: "le parti, affermava, un giorno, di fronte ad una disputa, potranno sedersi e procedere ad un calcolo; ipotizzava una simbolizzazione del pensiero con cui operare calcoli logico-matematici. Si dovrebbe introdurre, cioè, nelle scienze umane, la certezza di quelle matematiche."⁷¹.

La differenza rispetto al passato è che oggi disponiamo di macchine capaci di coniugare diritto e tecnologia. Quello che un tempo era perlopiù un'ingenua utopia, forse oggi può diventare realtà: prevedere l'esito di un processo grazie all'informatica. Un pensatore illuminista che si allinea con questa visione è Montesquieu che, ne *L'Esprit des Lois*, descrive il giudice come *bouche de la loi*. In breve, secondo il filosofo francese, il giudice non dovrebbe permettersi di attribuire alla legge un significato diverso da quello letterale: la norma, dunque, andrebbe semplicemente dichiarata perché già evidente dalla normale lettura della disposizione. Se così non fosse, si concretizzerebbe un arbitrio del giudice, "la cui decisione non sarebbe più anticipatamente prevedibile dagli altri consociati"⁷². Oltre a sostituirsi al legislatore, quindi, il giudice violerebbe il principio di certezza del diritto.

Proseguendo su questa linea di pensiero, si arriva ad affermare che affidarsi agli algoritmi per decidere le liti sarebbe una scelta non solo possibile, ma auspicabile poiché in linea con

⁷⁰ G. W. von Leibniz, 1666; Si veda anche G. Iommi Amunátegui, 2015

⁷¹ L. Viola, 2018, 168

⁷² S. Gasparini, 2002, 126

il principio di certezza del diritto e con le idee illuministe che si trovano ancora oggi negli ordinamenti giuridici.

Leggendo l'articolo 12 delle Disposizioni sulla legge in generale, al primo comma, ne troviamo un chiaro segno: "Nell'applicare la legge non si può ad essa attribuire altro senso che quello fatto palese dal significato proprio delle parole secondo la connessione di esse, e dalla intenzione del legislatore". Il primo passo che deve compiere l'interprete, quindi, è quello descritto da Montesquieu, che riletto ai giorni nostri assomiglia sempre di più a un calcolo matematico.

L'articolo 101 della Costituzione, al secondo comma, richiama nuovamente l'impostazione illuminista del giudice *bouche de la loi*. Come riportano i lavori preparatori dell'Assemblea costituente: "Sul secondo comma («I magistrati dipendono soltanto dalla legge, che interpretano e applicano secondo coscienza») rilevò [l'on. Ruini] che era stato proposto di sostituire alla parola «dipendono» le parole «sono vincolati» oppure «obbediscono» oppure «sono soggetti»; a nome della Commissione accettò quest'ultima espressione, «la quale presenta minori inconvenienti, e comunque rientra piuttosto nella revisione stilistica». A proposito della seconda parte del secondo comma («che interpretano secondo coscienza») il presidente della Commissione osservò che essa presentava un dilemma: «o è una dichiarazione generica di ovvio significato, e allora se ne può fare a meno, o apre la via a una interpretazione che sarebbe pericolosa, e allora vi è una ragione di più per abbandonarla. Io non credo – aggiunse l'on. Ruini – che, parlando di coscienza del giudice, si possa intendere la tendenza e l'ammissione del cosiddetto *diritto libero*, costruzione teorica per me inammissibile; ma non discara, fra gli altri, all'hitlerismo. Ad ogni modo, poiché è stato manifestato un dubbio, ed il togliere l'inciso non nuoce – anzi, abbrevia il testo – il Comitato acconsente alla soppressione»⁷³.

Il risultato dei lavori dell'Assemblea è l'attuale testo dell'articolo 101 secondo comma Cost. "I giudici sono soggetti soltanto alla legge", che, oltre a essere espressione dell'indipendenza dei magistrati dagli altri poteri dello Stato, cristallizza e ci consegna i principi su cui si basa

⁷³ V. Falzone-F. Palermo-F. Cosentino (a cura di), 1948, 187

l'idea del diritto illuminista. La scelta di togliere ogni riferimento alla "coscienza" del giudice, infatti, può essere letta come un tentativo di affermare nuovamente la centralità del testo di legge sopra ogni possibile interpretazione⁷⁴, come se la disposizione parlasse da sola e non ci fosse bisogno di un soggetto capace di coglierne il significato. Oggi sappiamo che, nei fatti, l'attività del magistrato giudicante va ben oltre la mera interpretazione letterale, potendo arrivare, ove possibile, ad adeguare un testo di legge alla Costituzione⁷⁵; le tracce dell'impostazione illuminista sono ancora presenti nel nostro ordinamento e potrebbero essere il grimaldello per poter introdurre i sistemi esperti nei processi.

Se quindi questo è ciò che, sulla carta, si chiede ad un giudice umano, è ben possibile che lo stesso si possa fare con un programma, magari permettendogli di imparare autonomamente e di migliorare col tempo. Se tali sistemi fossero perfettamente funzionanti, scevri da ogni difetto e pienamente controllabili dalle parti (temi su cui si tornerà nel terzo capitolo), potrebbero addirittura garantire una prevedibilità maggiore dell'esito dei processi, un rispetto più puntuale del principio di certezza del diritto e una maggior rapidità dei tempi processuali, in ossequio all'articolo 111 secondo comma Cost. e all'articolo 6 primo comma CEDU.

2.2 *Un software può interpretare la legge?*

Se la tradizione giuridica illuminista ci permette di dire che è possibile utilizzare *software* basato sull'IA nelle aule di giustizia, bisogna ora mettere a fuoco la questione da un punto di vista più tecnico.

Alla legge manca sovente l'inequivocità, l'analiticità e la completezza a cui ambivano i pensatori del passato, tanto che i giuristi faticano a interpretarla e coordinarla con le altre norme. Questo accade perché spesso il legislatore deve trovare delle soluzioni di compromesso poiché mosso da forze politiche instabili e ideologicamente diverse tra loro: il risultato sono disposizioni scritte in maniera evasiva o elastica proprio perché, altrimenti,

⁷⁴ U. Vincenti, 2020, 43

⁷⁵ A. Pisaneschi, 2016, 72; Per approfondire si segnalano i lavori di G. Sorrenti, 2006; F. Sorrentino, 1987

non si sarebbe raggiunto un accordo. La giurisprudenza, nel tentativo di sopperire a tali mancanze, ha visto aumentare l'importanza della propria attività interpretativa. Mediante l'interpretazione della legge, come di qualsiasi altro testo scritto in linguaggio naturale, infatti, "si può integrare o ridurre o comunque modificare e, in taluni casi limite, capovolgere la portata di una norma"⁷⁶.

Due esempi possono aiutare a comprendere il fenomeno: l'applicazione dell'art. 2059 c.c. e della disciplina sulle espropriazioni per pubblica utilità. Il primo dispone tassativamente, in un solo comma, che "Il danno non patrimoniale deve essere risarcito solo nei casi determinati dalla legge". La disposizione sembra molto chiara, soprattutto se affiancata all'articolo 185 c.p., il quale prevede che "ogni reato, che abbia cagionato un danno patrimoniale o non patrimoniale, obbliga al risarcimento il colpevole", configurando così l'ipotesi di cui all'articolo 2059 c.c. Nonostante la limpidezza delle disposizioni, si è prodotto un generale malcontento tra la popolazione, che trovava ingiusta tale impostazione normativa. In questo contesto, è stata la giurisprudenza a modificare in via interpretativa la portata dell'articolo 2059 c.c. ritenendo che il patrimonio di una persona fosse costituito non solo dai suoi beni, ma anche dalle sue stesse potenzialità fisiche. Questo cambio di paradigma ha portato alla creazione del cd. danno biologico, riconoscibile anche quando non causato da un fatto costituente reato⁷⁷, senza che il legislatore abbia cambiato di una virgola le disposizioni coinvolte.

Il secondo esempio, riguardante le espropriazioni per pubblica utilità, mostra di nuovo quanto possa essere incisiva l'attività giurisprudenziale. Originariamente la P.A., per poter espropriare una proprietà privata, doveva prima emanare un decreto di esproprio e determinare il relativo indennizzo, ma nella pratica ciò avveniva di rado. Col passare degli anni, l'Autorità ha iniziato a occupare i terreni ben prima di adempiere agli obblighi di legge, costruendoci sopra opere pubbliche, così da rendere irrealistica la prospettiva di demolizioni e di restituzioni. La giurisprudenza, trovandosi davanti una situazione di illegalità così manifesta, ha fatto ricorso alla nozione di "accessione invertita" prevista dall'art. 938 c.c.

⁷⁶ R. Borruso-R. M. Di Giorgi-L. Mattioli-M. Ragona, 2004, 323

⁷⁷ R. Borruso-R. M. Di Giorgi-L. Mattioli-M. Ragona, 2004, 323; C. Salvi, 1989; F. Romanazzi, 2021

(che non riguardava le espropriazioni) per permettere, in via eccezionale, alla P.A. di occupare una porzione di un fondo altrui, costruirci sopra un edificio e acquisirne la proprietà, salvo il riconoscimento del danno arrecato all'ex proprietario⁷⁸. Più precisamente, la norma stabilisce che "Il costruttore è tenuto a pagare al proprietario del suolo il doppio del valore della superficie occupata, oltre il risarcimento dei danni". Anche in questo caso non è servito modificare alcun testo di legge per permettere ai giudici di individuare una soluzione creativa a un problema divenuto insostenibile.

A tutto ciò si aggiunge l'endemica ambiguità della lingua, che porta a dare significati diversi alle stesse espressioni in base al contesto storico-culturale in cui vengono usate. Com'è noto, il senso complessivo di ogni frase è dato dai seguenti fattori: la sintassi, cioè il modo con cui le parole sono correlate, la grammatica, cioè la forma data a ogni singola parola, e la semantica, cioè il significato proprio di ogni parola. Uno dei problemi derivanti dalla semantica è il significato da attribuire ai giudizi di valore, concetti astratti usati anche a livello normativo: atti osceni, pudore, opera d'arte, buon costume, ecc. Secondo alcuni, tali espressioni sarebbero addirittura una parziale abdicazione del legislatore rispetto al potere giudiziario, chiamato non più ad applicare la legge, ma a crearla e a cambiarla a seconda dei tempi e dei luoghi⁷⁹. Senza arrivare a questo estremo, è comunque da sottolineare la problematicità in sede interpretativa di disposizioni come l'articolo 527 c.p., che portano i giudici a dover capire, in quell'esatto momento, cosa sia un atto osceno e cosa no in base a ciò che ritengono essere la sensibilità sociale corrente. Un altro esempio è il concetto di opera d'arte o dell'ingegno così come prevista dal diritto d'autore, che è cambiato in modo così radicale da poter vedere ormai da qualche anno mostre d'arte dedicate ai videogiochi anche in contesti in cui tradizionalmente non ce lo si aspetterebbe, come la mostra Play istituita a La Venaria Reale.

È quindi possibile affidare l'attività interpretativa a una macchina? Secondo una prima teoria, che potremmo definire oggettiva, la risposta sarebbe affermativa. Questa visione mette al centro il principio di certezza del diritto, che prevede chiarezza, conoscibilità e

⁷⁸ R. Borruso-R. M. Di Giorgi-L. Mattioli-M. Ragona, 2004, 324

⁷⁹ R. Borruso-R. M. Di Giorgi-L. Mattioli-M. Ragona, 2004, 339; Sul tema si sono espressi anche A. Trabucchi, 1959; D. Dalfino, 2020; M. Luciani, 2020

univocità delle norme a cui ricollega una coerente e rapida applicazione delle stesse da parte dei giudici. In tal modo si renderebbero prevedibili le conseguenze a cui portano determinati comportamenti; la tecnologia verrebbe usata solo al momento dell'applicazione della legge e non per l'accertamento dei fatti, che continuerebbe a essere svolto dagli esseri umani. Questo perché, secondo tale ottica, la legge sarebbe di per sé un algoritmo perfettamente interpretabile e applicabile anche da una macchina (alla luce dell'articolo 12 preleggi). Si può portare a sostegno di questa tesi anche una certa giurisprudenza della Corte di Giustizia, la quale, nel 2005⁸⁰ e nel 2007⁸¹, ponendo l'accento proprio sulla ragionevole prevedibilità delle decisioni, ha criticato apertamente le sentenze in contrasto con l'interpretazione giurisprudenziale più diffusa in un dato momento storico di una certa disposizione. Questa visione riprende la tradizione positivista del diritto: la legge, che si identifica con l'ordinamento giuridico (e viceversa), non può che essere completa e priva di lacune e il compito del giurista deve essere analizzarla come farebbe uno scienziato con il proprio oggetto di studio. Non può avere spazio alcun tipo di arbitrio da parte del giudice, che deve approcciarsi al diritto con l'atteggiamento dell'uomo di scienza⁸², prediligendo l'applicazione "automatica" dell'interpretazione letterale delle disposizioni di legge. Quest'ultima, al

⁸⁰ CGCE, 28/6/2005, C-189/02, C-202/02 in cui si legge: "In tale contesto, occorre ricordare la giurisprudenza della Corte europea dei diritti dell'uomo relativa all'art. 7, n. 1, della CEDU, peraltro richiamata da più ricorrenti (v., in particolare, Corte europea dei diritti dell'uomo, sentenze 22 novembre 1995, S.W. e CR. c. Regno Unito, serie A nn. 335-B e 335-C, § 34-36 e § 32-34; 15 novembre 1996, Cantoni e. Francia, Recueil des arrêts et décisions 1996-V, § 29-32, e 22 giugno 2000, Coëme e a./Belgio, Recueil des arrêts et décisions 2000-VII, § 145).

Da tale giurisprudenza deriva infatti che la nozione di «diritto» ai sensi del detto art. 7, n. 1, corrisponde a quella di «legge» utilizzata in altre disposizioni della CEDU ed include il diritto di origine sia legislativa che giurisprudenziale. Tale disposizione, che sancisce in particolare il principio di previsione legale dei reati e delle pene (*nullum crimen, nulla poena sine lege*), benché non possa essere intesa come un divieto di graduale chiarimento delle norme sulla responsabilità penale, può, secondo tale giurisprudenza, opporsi all'applicazione retroattiva di una nuova interpretazione di un divieto di legge.

Ai sensi della stessa giurisprudenza, ciò avviene, in particolare, nel caso si tratti di un'interpretazione giurisprudenziale il cui risultato non era ragionevolmente prevedibile nel momento in cui l'infrazione è stata commessa, alla luce, in particolare, dell'interpretazione vigente a quell'epoca nella giurisprudenza relativa alla disposizione legale in questione".

⁸¹ CGCE, 8/2/2007, C-3/06 in cui si legge: "In particolare, l'art. 7, n. 1, della Convenzione europea di salvaguardia dei diritti dell'uomo e delle libertà fondamentali, stipulata a Roma il 4 novembre 1950, che consacra in particolare il principio di previsione legale dei reati e delle pene (*nullum crimen, nulla poena sine lege*), può opporsi all'applicazione retroattiva di una nuova interpretazione di una norma che descrive un'infrazione (v., in tal senso, sentenza Dansk Rørindustri e a./Commissione, cit., punto 217).

Ciò avviene, in particolare, nel caso si tratti di un'interpretazione giurisprudenziale il cui risultato non era ragionevolmente prevedibile nel momento in cui l'infrazione è stata commessa, alla luce, in particolare, dell'interpretazione vigente a quell'epoca nella giurisprudenza relativa alla disposizione legale in questione (v. sentenza Dansk Rørindustri e a./Commissione, cit., punto 218)".

⁸² M. La Torre, 2013

giorno d'oggi, pare aver perso l'importanza che aveva in passato e si è soliti formulare due ipotesi per spiegarne i motivi. Secondo una parte della dottrina, infatti, si è ormai raggiunto un livello di elaborazione teorica talmente elevato in materia di interpretazione letterale che sarebbe rimasto troppo poco spazio per ulteriori ricerche. Secondo altri, il tema viene sempre più trascurato perché sarebbero emersi dei fattori "antiletteralistici" di carattere ideologico, istituzionale o professionale che, di fatto, rendono difficile la ricerca⁸³.

Al di là degli aspetti filosofici, il campo del trattamento automatico del linguaggio naturale (NLP o *Natural Language Processing*) e dei sistemi di apprendimento automatico del linguaggio (MLL o *Machine Language Learning*) ha compiuto enormi passi avanti a partire dagli anni Ottanta e l'area giuridica, nello specifico, è in grado di generare moltissimi dati derivanti dalla produzione normativa, amministrativa, giurisprudenziale, defensionale e negoziale. Tutte queste informazioni sono analizzabili automaticamente dai sistemi esperti grazie a dei meccanismi di "codificazione", ossia tramite l'annotazione di "etichette" che permettono di identificare i testi da un punto di vista linguistico in rapporto a vari aspetti: quello fonetico, lessicale, morfologico, grammaticale, sintattico, semantico, pragmatico, ecc. Quando tutti questi dati vengono classificati (manualmente o automaticamente), il sistema può gestirli e utilizzarli in modo più efficiente⁸⁴. Per esempio, a partire dall'analisi automatica dei testi attraverso il cd. *parsing* (grazie al quale viene ricostruita la struttura sintattica delle frasi), la macchina può svolgere autonomamente funzioni particolarmente sofisticate, come il recupero delle sentenze, la stesura di una sintesi del loro contenuto, della loro struttura logica, ecc. Da un punto di vista più tecnico, i testi vengono suddivisi in unità linguistiche, ossia le singole parole e segni di punteggiatura, a cui si associa la corrispondente categoria morfosintattica per poter ricostruire e rappresentare la struttura del segmento di periodo secondo un sistema di relazioni binarie, per esempio tra soggetto e oggetto. Gli strumenti intelligenti vengono addestrati su uno o più domini, il che li rende poco adatti all'applicazione su testi di natura diversa: è quindi necessario che vengano sviluppati avendo in mente in partenza quale sarà la loro area di utilizzo futura⁸⁵.

⁸³ P. Chiassoni, 2000, 5

⁸⁴ E. Cicconi, 2020, 76; N. Calzolari-A. Lenci, 2004, 56

⁸⁵ E. Cicconi, 2020, 76

Ci sono, infatti, delle differenze tra un testo giuridico e un testo in lingua comune. In uno studio esposto da Giulia Venturi nel 2012 vengono comparati proprio questi due tipi di linguaggio confrontando un *corpus* di testi giuridici in tema di responsabilità dello Stato per mancata attuazione di direttive europee e uno rappresentativo della lingua comune. Il primo era diviso in due parti: una composta da provvedimenti normativi interni ed europei e una composta da sentenze emesse da vari giudici (giudici ordinari e amministrativi di merito, Corte di cassazione, Corte costituzionale, Corte EDU). Il secondo era composto da articoli di giornale provenienti da La Repubblica e Due Parole, volendo diversificare il più possibile il tenore dei testi: si passa da un quotidiano a diffusione nazionale a uno più ponderato e di facile lettura, scritto da esperti linguisti⁸⁶. Dopo la loro sottoposizione alle analisi automatiche degli strumenti sopra descritti, sono emerse alcune differenze di cui è bene dare conto. I testi giuridici presentano frasi più lunghe, usano con minor frequenza i lemmi del cd. “vocabolario italiano di base” e con maggior frequenza i lessemi di “alto uso” e di “alta disponibilità”, ossia i lemmi di uso immediatamente inferiore a quelli fondamentali (per esempio, abbassare, alimento o zampa) e quelli che sono poco scritti ma legati a oggetti, fatti ed esperienze quotidiane (per esempio, abbraccio, accavallare o zampogna)⁸⁷. Ancora, i testi giuridici sono ricchi di sostantivi, aggettivi, preposizioni e numeri, ma hanno pochi verbi e avverbi, il che si spiega con il loro carattere tendenzialmente informativo, e hanno una complessità sintattica maggiore degli articoli di giornale⁸⁸.

I sostenitori della cd. teoria soggettiva sono del parere esattamente opposto. Il punto focale, in questa tesi, è il postulato secondo cui la legge non ha un significato oggettivo, di conseguenza l'interprete e la sua sensibilità sono essenziali e non sostituibili da una macchina. Il motivo è che, come ricordato in precedenza, il diritto è permeato da disposizioni valoriali (buona fede, ragionevolezza, buon costume, ecc.), che di per sé non possono essere oggettive: il significato che un determinato giudice dà al buon costume può essere diverso dal significato che gli attribuisce un altro giudice, senza che ciò si sostanzi in una errata interpretazione delle disposizioni di legge. A sostegno di questa idea possono essere portati

⁸⁶ G. Venturi, 2012, 139 ss.

⁸⁷ L. Lorenzetti, 2010

⁸⁸ E. Cicconi, 2020, 77

alcuni esempi di disaccordi interpretativi profondi (DIP), ossia “quelle divergenze particolarmente radicali che occorrono nell’attività interpretativa di giudici e giuristi nei casi in cui essi hanno a che fare con espressioni e locuzioni valutative contenute in disposizioni esplicite, ovvero in norme implicite (costruite per via dogmatica o giurisprudenziale, e poi nuovamente suscettibili di interpretazione), disposizioni e norme che, nella maggior parte dei casi (e certamente nei casi più interessanti), fanno parte della categoria dei principi costituzionali a carattere fondamentale, che riconoscono e garantiscono diritti, appunto, fondamentali”⁸⁹.

I DIP possono essere diretti, quando hanno ad oggetto il contenuto sostanziale delle disposizioni, o indiretti, quando riguardano l’approccio interpretativo da adottare; in entrambi i casi non si tratta di questioni attinenti solo e soltanto a principi costituzionali fondamentali, visto che tali divergenze possono riscontrarsi anche in disposizioni di legge come l’art. 529 c.p. (comune sentimento del pudore): ciò che più conta è che l’espressione usata sia di tipo valutativo o valoriale (rispetto della persona umana, diritto alla vita, pari dignità sociale, famiglia come società naturale, libertà come autodeterminazione, rieducazione del condannato, diritto alla salute) per via della loro natura intrinsecamente compromissoria e indeterminata. Due esempi classici di DIP sono quelli riguardanti il rapporto tra:

- Diritto alla vita e diritto all’autodeterminazione in materia di “fine vita”;
- Diritto di manifestazione del pensiero e principio di dignità in materia di stampa e informazione⁹⁰.

Nonostante il loro numero sia limitato rispetto ai casi “facili” o più ripetitivi, i DIP richiedono spesso l’intervento di principi costituzionali fondamentali, il che solleva questioni importanti in ordine al rapporto fra diritto e morale negli stati di diritto, alla correttezza o meno delle decisioni giudiziali, alla natura e al ruolo dei giudizi di valore coinvolti nell’interpretazione,

⁸⁹ V. Villa, 2017, 2

⁹⁰ V. Villa, 2017, 6

ecc. Questioni così delicate, per quanto siano quantitativamente marginali, meritano la giusta attenzione sia nel dibattito dottrinale che nelle aule di giustizia⁹¹.

Si è soliti spiegare l'esistenza dei DIP, a livello generale, identificando almeno due cause. Una prima motivazione è che nelle società moderne vi è la presenza di un ormai consolidato pluralismo etico, ossia di "concezioni comprensive, ragionevoli ma incompatibili"⁹². Ciascuna di esse esprime una certa prospettiva del valore della vita umana e sui suoi fini, che non potrà mai essere condivisa da tutti i cittadini poiché figlia di un determinato modo di intendere gli aspetti filosofici, religiosi e morali dell'esistenza. Ciò porta all'elaborazione di testi costituzionali come quello attualmente in vigore in Italia, che è nato grazie all'incontro di culture politiche profondamente diverse: una di ispirazione cattolica, una di ispirazione laico-liberale e una di ispirazione marxista, che hanno portato alla formulazione di espressioni generiche e indeterminate che concedono ampi spazi di discrezionalità⁹³.

Un'altra causa di carattere generale è la presenza di più livelli di riconoscimento e tutela dei diritti. Esistono, infatti, molteplici sedi sovranazionali e internazionali di produzione, riconoscimento e di garanzia dei diritti fondamentali (CEDU, Carta di Nizza, convenzioni e trattati internazionali) e questa incidenza non dipende solo dal riconoscimento interno di fonti del diritto ulteriori rispetto a quelle "locali", ma anche dal fatto che ci sono frequenti richiami, nelle sentenze delle Corti nazionali, di argomenti utilizzati da Corti sovranazionali e internazionali (il c.d. "dialogo fra le corti"). Vi è quindi, secondo Massimo Luciani, il sovrapporsi di cinque sfere di legalità che portano all'aumento dei DIP:

- la legalità legale, custodita dalla Corte di cassazione;
- la legalità costituzionale, custodita dalla Corte costituzionale;
- la legalità comunitaria, custodita dalla Corte di giustizia;
- la legalità convenzionale a livello europeo, custodita dalla Corte EDU;

⁹¹ V. Villa, 2017, 9; L. Ferrajoli, 2001, 328 ss.

⁹² V. Villa, 2017, 10; J. Rawls, 2005

⁹³ V. Villa, 2017, 10

- la legalità internazionale, custodita da varie Corti con competenze e caratteri diversi⁹⁴.

Come spesso accade in questi casi, esiste anche una visione di compromesso che potremmo definire mediana. Questa, nel rispondere affermativamente al quesito esposto all'inizio del paragrafo, introduce dei temperamenti rispetto alla teoria oggettiva. Il presupposto su cui si fonda è la centralità dell'articolo 12 preleggi, che deve poter risolvere ogni conflitto interpretativo poiché "se tutti sono soggetti alla legge, ivi compresi i giudici ex art. 101 comma 2 Cost., allora ogni conflitto – anche interpretativo – deve trovare la sua risposta nella legge stessa"⁹⁵.

Secondo la teoria mediana, la matematica può aiutare (e non sostituire) i giuristi nella pratica del diritto perché sono le stesse norme e principi a contenere alcuni meccanismi che potremmo definire algoritmici:

- L'articolo 3 Cost., che recita "Tutti i cittadini hanno pari dignità sociale e sono eguali davanti alla legge, senza distinzione di sesso, di razza, di lingua, di religione, di opinioni politiche, di condizioni personali e sociali. È compito della Repubblica rimuovere gli ostacoli di ordine economico e sociale, che, limitando di fatto la libertà e l'eguaglianza dei cittadini, impediscono il pieno sviluppo della persona umana e l'effettiva partecipazione di tutti i lavoratori all'organizzazione politica, economica e sociale del Paese.", altro non è che una regola matematica: a parità di variabili, il risultato deve essere uguale. Se F1 è il fatto compiuto da una determinata persona e F2 è lo stesso fatto compiuto, però, da un'altra persona e T è il trattamento giuridico che ne deriva, otteniamo la seguente proporzione: $F1 : T = F2 : T$. Su questa scia si colloca anche l'articolo 12 delle preleggi, il quale fissa una sequenza di azioni volte a risolvere eventuali conflitti interpretativi, attribuendo, per esempio, maggiore importanza alla lettera della legge piuttosto che all'analogia. Anche questo ricorda il funzionamento di un algoritmo, che può essere definito come un insieme di istruzioni volte alla risoluzione di un problema. Se ciò non bastasse, anche il processo in sé

⁹⁴ M. Luciani, 2006

⁹⁵ L. Viola, 2018, 38

assomiglia a un algoritmo: la sentenza o il provvedimento giudiziale (PG) altro non sarebbe che il risultato di una serie di fasi e azioni stabilite dalla legge: si uniscono i fatti provati (FP) alle norme così come interpretate (IP). Matematicamente potremmo dire $PG = FP \wedge IP$ ⁹⁶;

- Per quanto riguarda le clausole valoriali, bisogna ricordare che anche queste dovrebbero essere interpretate secondo i criteri dell'articolo 12 preleggi alla luce delle norme costituzionali. Quindi qualunque interpretazione data a suddette clausole non potrà che essere *intra ius* e non *extra legem*, con la conseguenza che non si dovrebbe mai sovvertire il dato letterale⁹⁷;
- Passando al valore della certezza del diritto, possiamo vedere come venga richiamato sia dal Codice di procedura civile che dal Regio decreto sull'Ordinamento giudiziario. L'articolo 348 bis c.p.c. richiede espressamente, ai fini dell'appello, la ragionevole probabilità del suo accoglimento, postulando quindi la prevedibilità della decisione e, quindi, la sua tendenziale certezza. Ancora, l'articolo 65 del Regio decreto n.12 del 30 gennaio 1941 stabilisce che la Corte di cassazione "assicura l'esatta osservanza e l'uniforme interpretazione della legge, l'unità del diritto oggettivo nazionale"⁹⁸.

Tutto ciò non significa che gli algoritmi possono essere d'aiuto sempre e comunque. Ci sono dei casi, infatti, in cui la legge affida espressamente alla sensibilità del giudice il compito di risolvere una controversia: ciò accade, per esempio, ex art 316 c.c. in materia familiare ed ex art 702 ter comma 5 c.p.c. in materia di rito sommario di cognizione. Vista la delicatezza delle questioni, l'obiettivo del legislatore è dare più importanza alle scelte discrezionali dell'interprete, rendendo quindi inutilizzabili i modelli matematici. L'ottica mediana, quindi, da un lato riconosce e ammette l'applicazione dei sistemi esperti durante i processi facendo leva, tra gli altri, sull'articolo 12 preleggi, ma dall'altro lato riconosce alcuni limiti legislativi da non oltrepassare per evitare di frustrare la *ratio* delle disposizioni che riconoscono in capo al giudice umano un'ampia discrezionalità. In conclusione, citando Luigi Viola: "soggettività e oggettività non vanno viste come alternative in contrapposizione tra loro, ma come

⁹⁶ L. Viola, 2018, 43

⁹⁷ L. Viola, 2018, 44; A. Di Porto, 2017, 119

⁹⁸ L. Viola, 2018, 46

coesistenti in costante tensione dialettica, in un processo continuo di rimandi e implicazioni: oggettività e soggettività fanno parte dello stesso procedimento ermeneutico”⁹⁹.

2.3 *La soggettività e la responsabilità dei robot intelligenti: la normativa attuale*

Se, quindi, parte della dottrina riconosce la possibilità di affidare le sentenze e l’interpretazione della legge ai robot, ci si può iniziare a chiedere se tali macchine siano anche responsabili dei danni che causano e se siano qualificabili come “persone”. Uno dei temi più spinosi e interessanti legati all’intelligenza artificiale, infatti, è il riconoscimento di una sorta di “personalità o soggettività giuridica” in capo alle macchine intelligenti da cui discendano dei diritti e dei doveri come accade per le persone fisiche e giuridiche. Per affrontare il discorso è necessario mettere in chiaro cosa si intende per “macchina intelligente”.

Nel capitolo precedente si sono già espone le diverse tesi in materia, ossia quella dell’intelligenza artificiale forte e dell’intelligenza artificiale debole, due modi opposti per vedere lo stesso fenomeno a seconda che si sostenga o meno la possibilità di ricreare in tutto e per tutto un sistema equiparabile all’essere umano o una sua semplice simulazione. Ciò su cui pare esserci un sostanziale accordo tra gli esperti è che per iniziare a poter parlare di “personalità elettronica” le macchine devono essere dotate di autoapprendimento, ossia della capacità di imparare da sole senza che gli esiti delle loro azioni siano prevedibili dai loro programmatori o costruttori. Non occorre quindi arrivare agli estremi visti in opere di fantasia come Pluto di Naoki Urasawa: è sufficiente dare uno sguardo al mondo che ci circonda per rendersi conto, come si è visto nel capitolo precedente, che macchine di questo tipo esistono già. Rispetto a quelle normali, queste operano partendo da degli *input* immessi nel sistema dall’uomo e si distinguono per la loro capacità non solo di processarli al fine di raggiungere un risultato (*output*), “bensì anche di correggere il proprio comportamento, di imparare dalla propria esperienza e di migliorare le proprie *performance*”¹⁰⁰.

⁹⁹ L. Viola, 2018, 48

¹⁰⁰ F. Caroccia, 2020, 219

La normativa attuale, specie a livello europeo, viene spesso ritenuta lacunosa e dispersiva poiché incapace di delineare un quadro giuridico unitario in materia. Se si considerano i robot come oggetti, si possono citare la Direttiva 2006/42/CE e la Direttiva 2001/95/CE, due testi che fissano degli standard comuni sulla sicurezza e sulla progettazione delle macchine¹⁰¹. A seconda dello specifico ambito in cui tali dispositivi vengono adoperati, la normativa si arricchisce ulteriormente:

- Nel caso di dispositivi medici come protesi, nanocapsule robotiche, esoscheletri e robot chirurgici, è attualmente in vigore il Regolamento 2017/745/UE che ne disciplina l'impiego;
- Parlando di *privacy* e decisioni automatizzate si deve far riferimento al noto Regolamento 2016/679/UE, il cd. GDPR;
- In materia consumeristica esiste la Direttiva 99/44/CE sulla vendita e sulle garanzie dei beni di consumo, che disciplina la responsabilità del venditore che consegna al consumatore un bene non conforme a quanto stabilito nel contratto di vendita¹⁰².

Il settore della responsabilità è ad oggi uno dei più accesi, ma, al netto del dibattito pubblico in corso, rappresentano un punto fermo la Direttiva 85/374/CEE e la Direttiva 99/34/CE sulla responsabilità dei danni da prodotti difettosi (confluite nel D.P.R. n. 224/1988 e nel codice del consumo, D.lgs. 206/2005), che stabiliscono il principio della responsabilità oggettiva, o senza colpa, del produttore.¹⁰³ Secondo quanto previsto da queste norme, per "prodotto" si intende "ogni bene mobile anche se fa parte di un altro bene mobile o immobile"¹⁰⁴, precisando quindi che si deve trattare di un oggetto tangibile; per "difettoso", invece, si intende qualcosa che "non offrirà la sicurezza che ci si può legittimamente attendere, tenuto conto della presentazione, dell'uso al quale è destinato e del momento della messa in

¹⁰¹ L. Coppini, 2018, 718; Ai sensi dell'art. 2 Dir. 2001/95/CE: "È considerato sicuro quel prodotto che, in condizioni di uso normali o ragionevolmente prevedibili, compresa la durata e se del caso, la messa in servizio, l'installazione e le esigenze di manutenzione, non presenta alcun rischio, oppure presenta unicamente rischi minimi, compatibili con l'impiego del prodotto e ritenuti accettabili nel rispetto di un elevato livello di tutela della salute e della sicurezza delle persone".

¹⁰² L. Coppini, 2018, 719

¹⁰³ R. Trezza, 2020, 51; Per approfondire si rimanda a G. De Cristofaro-A. Zaccaria, 2013; G. D'Amico, 2015

¹⁰⁴ Direttiva 85/374/CEE, Art. 2

circolazione”¹⁰⁵. Seguendo tale impostazione, come detto poc’anzi, l’unico responsabile dei danni causati dal bene difettoso sarebbe il produttore, ossia il “fabbricante di una materia prima, del prodotto finito o di una parte di esso”¹⁰⁶, che si vedrà imputata una cd. responsabilità oggettiva o “senza colpa” priva degli elementi del dolo o della colpa previsti in materia di responsabilità extracontrattuale ex art. 2043 c.c.

Parlando di onere della prova, questo spetta al danneggiato, che dovrà dimostrare tre cose: il difetto, il danno e il nesso di causalità tra i due. Ciò non vuol dire che il produttore non si dovrà preoccupare di dimostrare di non essere responsabile: egli potrà provare, per esempio, che il difetto che ha determinato il danno non esisteva al momento della messa in circolazione del prodotto oppure che le conoscenze tecnico-scientifiche al momento della creazione della macchina non permettevano di prevedere e sistemare il difetto¹⁰⁷.

La domanda a questo punto è se questa normativa possa o meno essere applicata, allo stato attuale, anche ai sistemi intelligenti. Uno dei problemi principali è il concetto di tangibilità, in quanto i dispositivi dotati di IA sono considerati “ibridi” poiché posseggono sia una componente materiale, l’*hardware*, che una immateriale, il *software* (e a volte solo quest’ultima). Secondo Remo Trezza, solo i danni causati da un *hardware* possono rientrare nella normativa in vigore, mentre a quelli provocati da un *software* si dovranno applicare le regole previste in materia di proprietà industriale¹⁰⁸, generando quindi possibili fraintendimenti in sede di applicazione. I prodotti dotati di intelligenza artificiale, infatti, pur essendo “mobili” e quindi tangibili (come richiesto dalla direttiva 374/1985/CE) vedono la comparsa dell’elemento immateriale, il *software*, che può avere un’importanza preminente sull’attività del prodotto, arrivando a essere autonomo rispetto alla sua componente materiale. In questi casi si tratta di verificare se il prodotto rimanga ancora tale dopo l’implementazione di un elemento di programmazione oppure se diventi qualcos’altro che non rientri necessariamente nella direttiva citata. Se tale componente immateriale avesse il

¹⁰⁵ Direttiva 85/374/CEE, Art. 6

¹⁰⁶ Direttiva 85/374/CEE, Art. 3

¹⁰⁷ R. Trezza, 2020, 51; V. D’Antonio, 2009, 608

¹⁰⁸ R. Trezza, 2020, 52; Sullo stesso punto si possono citare anche L. Coppini, 2018, 722; M. Bassini-L. Liguori-O. Pollicino, 2018, 334; E. Palmerini, 2016, 1826 ss.

sopravvento su quella tangibile, la macchina diventerebbe qualcosa in più di un semplice prodotto (così come identificato dall'art. 2) perché sarebbe autonomo e in una certa misura anche imprevedibile, ma al tempo stesso sarebbe qualcosa in meno di un essere vivente dotato di razionalità e volontà, rendendo difficile la sua catalogazione ai fini della determinazione della responsabilità (che è stata pensata per oggetti con caratteristiche differenti)¹⁰⁹.

Un ulteriore elemento di difficoltà è rappresentato dalla qualificazione dei soggetti coinvolti nella creazione delle macchine intelligenti, ossia il programmatore e il produttore. Il primo è colui che sviluppa gli algoritmi e i programmi attraverso cui il dispositivo agisce; il secondo sviluppa materialmente la macchina per la diffusione sul mercato. Può capitare che queste due figure non coincidano ed è proprio in questi casi che le cose si complicano, poiché individuare tutti i responsabili non è un'operazione sempre agevole, specie nelle catene di produzione su larga scala¹¹⁰.

La situazione sembra essere più chiara nei paesi extra europei come il Giappone, la Cina e la Corea del Sud, che si stanno sempre più dotando di testi legislativi più o meno settoriali frutto di una visione unitaria della robotica e dell'intelligenza artificiale. Il Giappone, nell'ambito della *Robot Revolution Initiative*, nel 2015 ha predisposto un piano quinquennale per la regolamentazione dei robot e della responsabilità dei danni da essi causati, con particolare attenzione all'aspetto risarcitorio. La Cina nel 2017 ha avviato un piano per divenire, entro il 2030, il punto di riferimento mondiale del settore, analizzando le problematiche giuridiche della robotica e mettendo a punto delle riforme adatte a regolarla in modo efficace, mentre in Corea del Sud stanno lavorando, dal 2012, a un documento (*Robot Ethics Charter*) per fissare degli standard minimi di produzione e funzionamento dei robot, i diritti e i doveri degli utilizzatori e dei proprietari, nonché i diritti e le responsabilità dei robot medesimi¹¹¹.

¹⁰⁹ R. Trezza, 2020, 52

¹¹⁰ R. Trezza, 2020, 53

¹¹¹ L. Coppini, 2018, 720

Dopo aver dato uno sguardo alla situazione normativa eurounitaria ed estera, ci si può chiedere se anche nell'attuale Codice civile siano rinvenibili delle disposizioni applicabili alle macchine intelligenti. Gli articoli più interessanti su cui riflettere sono il 2050 e il 2051 c.c. dedicati alla responsabilità per danni causati da entità terze non per forza umane. Nell'art. 2050 il pregiudizio nasce da un'attività pericolosa (per sua natura o per i mezzi utilizzati) condotta dall'uomo, la cui prova liberatoria sta nella dimostrazione di aver adottato tutte le misure idonee a prevenire il fatto, mentre nel 2051 ci si riferisce a un danno provocato direttamente dalla cosa, senza che ci sia alcun intervento umano¹¹².

Pensando alla robotica, queste norme potrebbero trovare terreno fertile nel campo dell'esplorazione spaziale, si pensi ai robot astronauti di cui da tempo si avvale la NASA per compiere operazioni pericolose o ai robot chirurghi usati nelle sale operatorie. Un esempio del primo tipo è *Justin*, un robonauta controllato da remoto da un operatore umano: se danneggiasse un satellite, si potrebbe applicare, in astratto, l'articolo 2050 c.c. visto che il robot non possiederebbe alcuna capacità decisionale, di conseguenza l'attività pericolosa sarebbe ascrivibile solo all'essere umano che lo guida. Discorso diverso si potrebbe fare pensando a *Robonaut2*, un robot astronauta capace di sostituire l'uomo durante lo svolgimento delle operazioni sul campo. I danni da lui causati potrebbero ricadere nella disciplina dell'articolo 2051 c.c., dovendo individuare, quindi, la figura del custode che in quel momento avrebbe potuto evitare il danno, con l'unica eccezione del caso fortuito¹¹³. Il punto è proprio capire in che misura una persona possa evitare il danno di una macchina come *Robonaut2*, che potrebbe trovarsi nella situazione di dover scegliere se provocare un danno minimo o un danno grave per portare a termine il compito assegnato in modo efficace: in questo caso si potrà parlare di caso fortuito?

¹¹² L. Coppini, 2018, 723; È da citare anche C. Salvi, 1998, 176 in cui viene precisato che la pericolosità, apprezzabile *ex ante*, è la potenzialità lesiva di grado superiore al normale, da valutare in base sia al profilo qualitativo dei pregiudizi minacciati che al profilo quantitativo-statistico con riguardo al numero dei sinistri. In giurisprudenza si distinguono attività pericolose tipiche, previste dalle leggi e regolamenti, e atipiche, la cui pericolosità viene valutata dal giudice in concreto.

¹¹³ L. Coppini, 2018, 723

2.3.1 La Risoluzione del Parlamento europeo del 2017

Con la Risoluzione P8_TA(2017)0051 del 16 febbraio 2017 “recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica”, il Parlamento europeo ha proposto una serie di norme giuridiche, etiche e deontologiche con l’obiettivo di gettare le basi di una futura regolamentazione comune e organica della materia a livello europeo. Il documento è composto da due parti: la raccomandazione vera e propria e un allegato contenente la “Carta sulla Robotica”, un testo di carattere etico.

L’assioma di partenza è che “non solo oggi i robot sono in grado di svolgere attività che tradizionalmente erano tipicamente ed esclusivamente umane, ma lo sviluppo di determinate caratteristiche autonome e cognitive — ad esempio la capacità di apprendere dall’esperienza e di prendere decisioni quasi indipendenti — li ha resi sempre più simili ad agenti che interagiscono con l’ambiente circostante e sono in grado di alterarlo in modo significativo” e che “in tale contesto, la questione della responsabilità giuridica derivante dall’azione nociva di un robot diventa essenziale”¹¹⁴. Subito dopo viene definito cosa si intende per robot autonomi: “l’autonomia di un robot può essere definita come la capacità di prendere decisioni e metterle in atto nel mondo esterno, indipendentemente da un controllo o un’influenza esterna”, precisando che “tale autonomia è di natura puramente tecnologica e il suo livello dipende dal grado di complessità con cui è stata progettata l’interazione di un robot con l’ambiente”¹¹⁵.

Il Parlamento ha riconosciuto l’importanza di elaborare un sistema di responsabilità pensato per sistemi intelligenti bilanciando “i valori intrinsecamente europei, universali e umanistici che caratterizzano il contributo dell’Europa alla società” e affermando che “tali regole non devono influenzare il processo di ricerca, innovazione e sviluppo nel settore della robotica”¹¹⁶.

Come si dovrebbe fare quando ci si appresta a regolare un certo ambito della realtà, bisogna partire da una necessità concreta, un problema a cui serve una soluzione giuridica: in questo

¹¹⁴ P8_TA(2017)0051, 2017, Z

¹¹⁵ P8_TA(2017)0051, 2017, AA

¹¹⁶ P8_TA(2017)0051, 2017, U

caso, il Parlamento lo ha rinvenuto nella difficoltà sempre maggiore di individuare i responsabili delle azioni o omissioni che non possono essere ricondotte a uno specifico soggetto umano quando a operare è un robot intelligente. Al tempo stesso riconosce l'insufficienza dell'attuale disciplina in materia di danno da prodotto difettoso e di responsabilità contrattuale o perché manchevoli sul piano dell'onere probatorio o perché risulterebbero inapplicabili alle nuove macchine "progettate per scegliere le loro controparti, negoziare termini contrattuali, concludere contratti e decidere se e come attuarli"¹¹⁷.

Venendo ai contenuti della raccomandazione, questa suggerisce di istituire un sistema di assicurazione obbligatorio, analogo a quello in vigore per le autovetture, in cui iscrivere tutti i robot al momento dell'immissione sul mercato. Ogni macchina dovrebbe essere dotata di una "scatola nera" contenente i dati di ogni operazione effettuata e, soprattutto, i passaggi logici che hanno condotto alla messa in atto delle operazioni stesse. Ma non solo. Oltre a imporre il regime assicurativo ai produttori e ai proprietari, dovrebbe essere istituito un fondo di garanzia che assicuri il risarcimento dei danni in caso di assenza di tale copertura assicurativa. In sostanza ogni robot verrebbe dotato di un numero di matricola individuale inserito in un registro comune pubblico europeo che consentirebbe di associarlo al suo fondo, in modo che chiunque entri in contatto con la macchina possa sapere in ogni momento tutto ciò che riguarda i limiti della sua responsabilità in caso di danni alle cose, i nomi dei contributori e tutte le informazioni utili in materia¹¹⁸.

La seconda proposta, la più criticata dell'intera raccomandazione, è logicamente collegata alla prima e attiene all'istituzione di uno *status* giuridico specifico per i robot, in modo che "possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato"¹¹⁹. Questo particolare status, com'è facilmente intuibile, sarebbe applicabile solo alle macchine dotate di un certo margine di autonomia e, per questo motivo, in grado di assumere decisioni non prevedibili dall'uomo, senza con ciò togliere l'eventuale responsabilità del costruttore o di chi abbia l'obbligo di custodirle: secondo il Parlamento

¹¹⁷ P8_TA(2017)0051, 2017, AG

¹¹⁸ F. Caroccia, 2020, 224

¹¹⁹ P8_TA(2017)0051, 2017, 59, lett. f)

europeo, è immaginabile un caso in cui sia l'automa che l'uomo siano chiamati a rispondere del medesimo danno.

Il motivo per cui questa raccomandazione è tanto dibattuta è che in un certo senso consacra e riconosce un margine di libertà decisionale in capo alle macchine intelligenti, arrivando a proporre di costruire uno status giuridico ad hoc pensato per loro, distinto da quelli già esistenti nel diritto civile. È chiaro che se non fossero così avanzate, non si giustificerebbe l'applicazione di un regime giuridico distinto, per esempio, da quello dei beni. A questo punto il problema torna a essere la definizione di "intelligenza artificiale" che si intende accogliere a livello normativo, individuando a priori il livello di autodeterminazione che la macchina deve avere per poter applicare il regime di responsabilità adatto.

A tutto ciò si aggiunge un altro elemento: la rapidità a cui procede lo sviluppo delle IA. Il rischio di pubblicare un ipotetico "Codice della Robotica" che regoli in modo preciso il livello di sofisticatezza che le macchine devono avere per poter essere considerate responsabili dei danni da loro causati, rischia di "nascere inattuale", in quanto durante la discussione precedente alla promulgazione, la tecnologia potrebbe aver fatto un ulteriore passo avanti e reso parzialmente inattuali le norme appena entrate in vigore: ciò che veniva considerato avanzato nel 2017 potrebbe non esserlo più nel 2022. Ciò si tradurrebbe nella necessità di modificare le disposizioni con una certa rapidità, generando forse più confusione che chiarezza. Il punto, quindi, non è evitare di affrontare il problema anche sul piano giuridico, ma farlo attraverso scelte legislative abbastanza flessibili da poter stare al passo con la ricerca.

2.3.2 Le critiche degli esperti alla Risoluzione

Le critiche alla Raccomandazione sono numerose: da un lato ci sono quelle relative al regime di responsabilità elettronica (alcuni ritengono che porterebbe all'elusione dell'art. 2740 c.c.)¹²⁰, dall'altro ci sono quelle indirizzate all'introduzione della "personalità elettronica". Queste ultime sono state riassunte in una lettera aperta alla Commissione

¹²⁰ F. Caroccia, 2020, 226; Obiezione superata da R. Borruso-G. Ciacci, 2005, 76

europea presentata da quasi trecento esperti di IA e robotica, industriali, giuristi, filosofi e medici. Secondo i firmatari, il Parlamento avrebbe sopravvalutato le capacità delle macchine intelligenti a causa di una percezione “*distorted by Science-Fiction and a few recent sensational press announcements*”¹²¹. Dal punto di vista etico e legale, la creazione di una personalità elettronica sarebbe sbagliata indipendentemente dal modello di riferimento:

- Se si usasse il modello delle persone fisiche, si riconoscerebbero in capo ai robot diritti come quello alla dignità, all’integrità, alla cittadinanza e alla giusta retribuzione, andando contro quanto stabilito dalla Carta dei diritti fondamentali dell’Unione europea e dalla Convenzione EDU;
- Non si potrebbe adottare nemmeno il modello delle persone giuridiche perché si dovrebbe partire dal presupposto che ci siano delle persone fisiche che rappresentino e dirigano l’attività della persona giuridica (mentre, nel caso dei robot autonomi, ciò non succede per definizione);
- Se, infine, si usasse il modello anglosassone del Trust o quello tedesco del Treuhand, non solo si dovrebbe gestire un assetto normativo molto complesso che non risolverebbe tutti i problemi legati alla responsabilità, ma anche in questo caso si presupporrebbe l’apporto di un essere umano (quantomeno del fiduciario)¹²².

Queste, in breve, sono le critiche sollevate dagli esperti al Parlamento e alla Commissione europea che, dal 2017, nei documenti successivi alla Raccomandazione, non hanno più menzionato la personalità elettronica. Ciò si può spiegare, secondo Francesca Carocchia, per due motivi: la mancanza di competenza dell’Unione europea in materia e la difficoltà di inserirsi su un terreno parzialmente già regolato. *In primis*, le istituzioni europee non hanno il potere di determinare il concetto di “persona”: solo gli Stati membri possono farlo, pur nel rispetto dei diritti umani stabiliti nei trattati internazionali e delle libertà fondamentali garantite dal diritto eurounitario. *In secundis*, alcuni ordinamenti europei hanno già affrontato il tema, per esempio ammettendo la creazione di società controllate da logiche

¹²¹ Open Letter to the European Commission: Artificial Intelligence and Robotics

¹²² Open Letter to the European Commission: Artificial Intelligence and Robotics

algoritmiche in cui i sistemi autonomi agiscono come se fossero dotati di soggettività, potendo porre in essere decisioni contrattualmente vincolanti¹²³.

Andando oltre alle critiche e al contenuto della Risoluzione, si è già detto che attribuire una forma di personalità ai robot, quale che sia il modello di riferimento, li farebbe diventare dei centri autonomi di imputazioni di interessi e/o di rapporti giuridici, rendendoli quindi titolari di obblighi extracontrattuali e contrattuali. Ma non solo. All'insorgenza di questi obblighi in capo alle macchine, è bene chiedersi se corrisponderebbe l'insorgenza di diritti, il che sottende due interrogativi distinti: stabilire quale sia lo statuto ontologico delle IA e se sia utile riconoscere loro alcune situazioni soggettive attive. Dal punto di vista giuridico, un esempio che può nuovamente tornare utile è la disciplina degli animali: pur essendo dei beni, possono apprendere attraverso l'esperienza e assumere comportamenti non controllabili dal proprietario. Se anticamente non sono mancati processi e condanne a topi, cavallette e maiali, a partire dal XIX secolo gli Stati hanno rivisto il sistema della responsabilità in chiave antropocentrica, riprendendo in tal modo la tradizione romanistica. Il fondamento dell'obbligazione risarcitoria diventa, quindi, la rimproverabilità della condotta, per questo si è scelto di concentrarsi su chi aveva in custodia l'animale piuttosto che sull'animale stesso (art. 2052 c.c.)¹²⁴. Quest'ultimo, di conseguenza, diventa una *res*, un bene mobile, per quanto sia riconosciuto come essere senziente: ciò non impedisce all'ordinamento di tutelarlo e di condannare i proprietari per i danni da lui causati.

Una soluzione ottimale, aldilà della scelta di quale modello adottare, può essere data cambiando la prospettiva con cui si guarda il problema: la domanda non dovrebbe essere se sia possibile attribuire la soggettività alle macchine, ma a che cosa servirebbe. Se il punto di partenza fosse questo, si potrebbero dotare le macchine di una soggettività giuridica limitata in base al ruolo che svolgono concretamente, costruendo, quindi, un apparato di norme volto a ridurre o a neutralizzare i rischi derivanti dalla loro attività, rifacendosi di volta in volta a modelli diversi (noti o inediti) in base alle fattispecie che si vogliono regolare¹²⁵.

¹²³ F. Caroccia, 2020, 228

¹²⁴ L. Coppini, 2018, 725

¹²⁵ F. Caroccia, 2020, 249; Sul tema si segnala anche G. Teubner, 2019

2.3.3 Cenni sulla responsabilità penale delle macchine intelligenti

Il tema della responsabilità e della soggettività dei robot tocca da vicino anche il diritto penale, forse complicando ulteriormente le cose. Uno dei principi cardine della materia, il principio di legalità, non permette di applicare per analogia una norma penale a un reato commesso da un robot perché, allo stato attuale, non sono considerati “persone” (basti pensare alle difficoltà già riscontrate in tal senso con riferimento alle persone giuridiche). Oltre a ciò, affinché un soggetto possa essere condannato per la commissione di un reato, serve che la sua condotta sia causalmente diretta al compimento del fatto e che alla base ci sia un nesso psichico, una sorta di appartenenza psichica dell’atto. All’articolo 85 c.p., infatti, è stabilito che “È imputabile chi ha la capacità d'intendere e di volere”, ossia chi comprende il significato delle proprie azioni e delle disposizioni di legge, potendo quindi scegliere quale condotta tenere¹²⁶.

Alla luce di queste disposizioni, una macchina intelligente può essere considerata capace di intendere e volere? Alcuni ritengono che si potrebbe applicare l’articolo 111 c.p.¹²⁷, rubricato “Determinazione al reato di persona non imputabile o non punibile”, in cui si stabilisce che, in ragione della non imputabilità del robot, risponderebbe del reato chi lo ha prodotto/fabbricato/utilizzato. L’IA, in questo caso, diventerebbe come un figlio e l’uomo diventerebbe come un genitore, la *ratio* che si vuole usare è la stessa: il figlio/robot non può essere responsabile dei reati ai quali è stato indotto dal suo genitore/programmatore. Se, al contrario, si escludesse del tutto questa possibilità, quindi si decidesse di non far rientrare questi casi nella fattispecie dell’articolo 111, la responsabilità penale resterebbe comunque in capo al costruttore/proprietario (anche nei casi più complessi in cui le macchine sono talmente avanzate da accarezzare gli standard dell’IA forte).

A questo punto sorge un problema: fino a che punto l’atto illecito della macchina è imputabile al suo produttore/proprietario? Dalle norme codicistiche emerge che la responsabilità è attribuibile a una persona se è configurabile almeno il profilo della colpa e difficilmente ciò potrà avvenire se a commettere il reato è un’entità artificiale che ha deciso

¹²⁶ A. Berti Suman, 2020, 261

¹²⁷ S. Cedrola, 2017

di agire senza che vi sia stata alcuna manomissione da parte di un essere umano. A tal proposito, andrebbero distinte le ipotesi di condotta illecita “autonoma” del robot e di condotta illecita “indotta”. Nel primo caso, imputando il fatto solo all’uomo, gli si riconoscerebbe una vera e propria responsabilità oggettiva; nel secondo, accanto alla responsabilità dell’essere umano, bisognerebbe interrogarsi sulla possibilità di riconoscerne una concorrente in capo alla macchina, chiedendosi anche quale sia il criterio di imputazione più adatto. Ci possono essere casi in cui una persona decide di usare una macchina come esecutrice della propria volontà omicida o come strumento di minaccia senza che si concretizzi un danno¹²⁸: si tratterebbe di ipotesi di malfunzionamento o l’automa aveva una qualche autonomia decisionale? Fino a che punto la condotta è “indotta”?

Da questo discorso emerge la possibilità, per il diritto penale, di concepire i robot come meri strumenti del reato, ossia come oggetti dipendenti dall’essere umano e incapaci di mettere in atto azioni puramente autonome. I comportamenti, quindi, sarebbero sempre attribuibili all’uomo sia oggettivamente che soggettivamente: la macchina non potrebbe eseguire nessuna azione senza un precedente apporto umano e tale azione, per quanto possa essere complessa, appartarrebbe all’agente (che la realizza con coscienza e volontà)¹²⁹.

La sentenza n. 26257 del 25 maggio 2017 della Corte di cassazione ha posto in essere alcune linee guida sul punto. Il Tribunale di Bergamo aveva confermato la decisione del Giudice di Pace di assolvere una madre dal reato di lesioni colpose (art. 590 c.p.) ai danni del figlio minore, al quale erano state amputate due dita della mano destra a seguito di un incidente domestico: il bambino, mentre la madre stava preparando una torta, ha infilato la mano nel frullatore. Secondo il Tribunale, il comportamento della donna non sarebbe rimproverabile perché avrebbe tenuto una condotta diligente raccomandando al figlio di non toccare nulla. Al contrario, la Cassazione ha ritenuto che “la particolarità della vicenda in esame è costituita dal fatto che si tratta di valutare la condotta (omissiva) del genitore rispetto ad un evento lesivo che il minore si è autoinflitto a seguito di un comportamento evidentemente frutto della sua incapacità di percepire il pericolo ed i rischi conseguenti all’inserimento dell’arto

¹²⁸ A. Berti Suman, 2020, 262

¹²⁹ S. Riondato, 2017, 87; T. Delogu, 1974

all'interno del contenitore di un frullatore acceso e funzionante"¹³⁰. Secondo i giudici, se la preparazione della torta comportava l'allontanamento della madre dal luogo in cui sussisteva il rischio, quest'ultima non poteva limitarsi a raccomandare al figlio di non toccare nulla, ma doveva adottare delle precauzioni più incisive, come spegnere il frullatore¹³¹. Alla luce di questa sentenza, le azioni di una macchina (in questo caso un semplice frullatore) sono da considerarsi strumentali all'azione criminosa e interamente attribuibili all'essere umano. Secondo alcuni, "la natura robotica dello strumento può avere un particolare rilievo, tra l'altro, quale elemento per misurare l'intensità della *mens rea* così come la pericolosità o la dannosità dell'azione complessiva, il che può essere valutato anche come fattore aggravante o attenuante dal giudice in fase di condanna"¹³².

Una vicenda ancora più spinosa è quella che ha riguardato un profilo Twitter di nome Tay, un *chatbot* creato da Microsoft nel 2016 a fini di ricerca. Tale programma era capace di rispondere in modo automatico a chiunque decidesse di scrivergli: l'obiettivo di Microsoft era acquisire dati e informazioni dalle conversazioni dirette con altri utenti e dalle conversazioni tra utenti, ma dopo poco l'esperimento è stato interrotto. Tay, dopo appena 16 ore dal lancio, aveva assorbito e iniziato a usare un linguaggio osceno e diffamatorio (da qui la rilevanza penale) nei confronti di un'attivista femminista, il che lo rendeva, di fatto, uno strumento fuori controllo. Se questo genere di applicativi venisse usato con l'obiettivo di screditare la reputazione di una persona, provocandole in tal modo un danno, si potrebbe facilmente configurare la fattispecie di cui all'art. 595 c.p.¹³³

Se considerando i robot come strumenti del reato non si vedono particolari problemi applicativi, questi ultimi, al contrario, sorgono quando si scardina la concezione classica secondo cui *machina delinquere non potest*. Anche se ciò avvenisse, secondo il Professor Silvio Riondato, resterebbero fermi tre presupposti fondamentali:

- L'attribuzione alle macchine della personalità sarebbe puramente fittizia (come lo è, del resto, l'attribuzione alle persone giuridiche). Gli automi, allo stato attuale, non

¹³⁰ Corte di cassazione, sentenza n. 26257, 25 maggio 2017

¹³¹ Corte di cassazione, sentenza n. 26257, 25 maggio 2017

¹³² S. Riondato, 2017, 89

¹³³ M. Curtotti, 2020, 524

hanno le stesse qualità morali, psicologiche e intellettuali dell'uomo e si potrebbero comunque predisporre delle regole che ne impediscono lo sviluppo oltre un certo limite di somiglianza all'essere umano, analogamente a ciò che accade con le cd. chimere (art. 13 legge n. 40/2004);

- Il riconoscimento sarebbe giustificato da necessità e finalità puramente umane;
- Sarebbero sempre gli uomini a decidere cosa sia giusto o sbagliato per i robot¹³⁴.

2.3.3.1 La tesi di Gabriel Hallevy

Seppur la dottrina assolutamente prevalente non voglia rimettere in discussione l'assioma citato poc'anzi, è utile citare una delle tesi più accreditate tra le voci più innovative: quella di Gabriel Hallevy, professore di diritto penale presso la Facoltà di Giurisprudenza israeliana dell'Ono Academic College. Secondo l'autore, non ci sarebbero dei veri ostacoli concettuali e normativi alla perseguibilità dei robot intelligenti. La responsabilità penale, nella sua ricostruzione, richiederebbe la sussistenza di due presupposti: l'*actus reus*, l'elemento oggettivo, e la *mens rea*, l'elemento soggettivo. Se una macchina riuscisse a replicare entrambi gli elementi, non si potrebbe negare la sua diretta responsabilità della commissione di un reato¹³⁵.

L'*actus reus*, nell'ottica tipicamente materialistica degli ordinamenti di diritto anglosassone, può agevolmente concepirsi con riferimento a un'IA non solo pensando alle condotte attive, come il movimento di un braccio robotico, ma anche alle condotte omissive integrate dall'inerzia della macchina. Ma c'è di più: secondo il Professor Hallevy, tale requisito si può ammettere anche quando la macchina si trovi a operare al di fuori dell'ambiente tangibile, per esempio nei reati informatici. La *mens rea* è decisamente più insidiosa. Per la teoria generale del diritto penale, è necessario che il soggetto sia munito di coscienza e volontà, le quali, a seconda della loro intensità, sono declinabili nel dolo, della colpa e della preterintenzione. L'autore, facendo leva sulla definizione di coscienza, asserisce che essa

¹³⁴ S. Riondato, 2017, 92; Per approfondire il tema si segnala G. Hallevy, 2010; Id. 2013

¹³⁵ G. Hallevy, 2015, 35

consista nella capacità di ricevere e comprendere informazioni sulla realtà di fatto e che sia presente nei sistemi esperti moderni¹³⁶.

Così come gli umani sono capaci di acquisire dati dal mondo esterno grazie ai cinque sensi e di rielaborarli attraverso processi cognitivi, allo stesso modo i sistemi intelligenti più recenti possono riprodurre tutto ciò usando degli algoritmi, il che permetterebbe loro di prevedere e desiderare un certo risultato. Ciò che manca alla macchina sarebbe solo la sfera emotiva, ma questa lacuna, almeno nell'ordinamento interno, non sarebbe un ostacolo perché è la stessa legge, all'art. 90 c.p., a tenerla distinta dalla sfera cognitiva, stabilendo che "Gli stati emotivi o passionali non escludono né diminuiscono l'imputabilità". Di conseguenza, il fatto che i sistemi intelligenti non siano in grado di provare delle vere e proprie emozioni o sentimenti, è del tutto irrilevante ai fini del giudizio di colpevolezza. L'autore, quindi, suggerisce di trattare i robot come le persone giuridiche per ragioni pratiche e di ordine sociale, ricreando una *fictio iuris* analoga a quella esistente per le società, le associazioni, ecc (di cui sono già state dette le criticità)¹³⁷.

Nonostante la suggestività della proposta, si può notare un punto debole capace di metterla in discussione: l'assenza di rimproverabilità delle condotte delle macchine intelligenti. Il concetto di colpevolezza delineato dall'art. 27 Cost. richiede la compresenza di quattro elementi: l'imputabilità, il dolo o la colpa, la conoscenza o la conoscibilità della legge penale violata e l'assenza di cause di esclusione della colpevolezza o scusanti. Un robot, per quanto avanzato, non possiede la libertà di autodeterminazione, la condizione imprescindibile ai fini dell'esigibilità del comportamento alternativo lecito. Di conseguenza non può scegliere se tenere una condotta conforme o meno all'ordinamento, nemmeno i sistemi in grado di agire in modo imprevedibile possono farlo: non sono in grado di prefissarsi degli obiettivi egoistici per il cui raggiungimento decidono di sacrificare i beni giuridici altrui¹³⁸.

Un altro punto debole della ricostruzione di Hallevy si può riscontrare sul versante delle sanzioni penali. Se è vero ciò che è stabilito dall'art. 27 co. 3 Cost., ossia che "le pene [...]

¹³⁶ M. Curtotti, 2020, 531

¹³⁷ M. Curtotti, 2020, 532

¹³⁸ M. Curtotti, 2020, 533

devono tendere alla rieducazione del condannato”, non si spiega come una macchina possa venire rieducata. La rieducazione, infatti, presuppone la capacità del reo di comprendere l’errore commesso traendo degli insegnamenti dalla pena, ma ciò non può accadere a meno che la macchina non venga programmata appositamente per questo scopo, il che sarebbe poco sensato poiché per modificare il comportamento di un robot è sufficiente incidere direttamente sull’algoritmo attraverso meccanismi di *machine learning* diretti a ottimizzarne il funzionamento¹³⁹. Avrebbe poco senso, insomma, agire attraverso delle misure sanzionatorie di tipo afflittivo o repressivo così come pensate per l’uomo.

¹³⁹ M. Curtotti, 2020, 534

CAPITOLO TERZO

L'INTELLIGENZA ARTIFICIALE NEL CASO LOOMIS

SOMMARIO: 3.1 Una breve introduzione ai *risk assessment tool* - 3.2 COMPAS e il caso Loomis v. Wisconsin - 3.2.1 Che cos'è e come funziona COMPAS - 3.2.2 La posizione della difesa e la sentenza definitiva - 3.3 Le critiche a COMPAS - 3.3.1 L'inchiesta di ProPublica - 3.3.2 Gli studi successivi - 3.4 Lo sviluppo di PATTERN e i dubbi costituzionali in merito ai *risk assesment tool* - 3.4.1 Il giusto processo nel diritto costituzionale interno

3.1 Una breve introduzione ai risk assessment tool

Nel continuo tentativo di migliorare la qualità della giustizia e di garantire sentenze sempre più “giuste”, nel corso degli anni ha preso piede, tanto tra i ricercatori quanto nelle aule di tribunale, una particolare categoria di sistemi esperti: i *risk assessment tool* o sistemi di valutazione del rischio di recidiva, programmi capaci di determinare la pericolosità di una persona e la probabilità che questa commetta un reato negli anni successivi alla misurazione. Ne esistono più di quattrocento sparsi in tutto il mondo e il loro sviluppo è iniziato nei primi anni Settanta, quando le valutazioni si basavano esclusivamente su delle semplici interviste rivolte ai soggetti interessati. Col passare dei decenni sono stati migliorati grazie all’inserimento di numerose variabili, come quelle statiche (età, genere, ecc) e dinamiche (lavoro, livello di istruzione, ecc), e alla capacità di stabilire un percorso ottimale di trattamento e reinserimento sociale dell’imputato¹⁴⁰. I più evoluti vedono l’implementazione del *machine learning*¹⁴¹, un meccanismo di autoapprendimento che, da un lato, per via della sua tendenziale segretezza, non consente di capire fino in fondo come la macchina arriva alle sue valutazioni, e che, dall’altro lato, permette al sistema di imparare autonomamente e di migliorarsi senza bisogno del supporto costante dei programmatori.

Nel sistema processuale statunitense i *risk assessment tool* vengono usati per:

- Valutare la sussistenza dei presupposti per mantenere una carcerazione preventiva;
- Valutare il rischio di recidiva o l’ammissibilità di misure alternative alla detenzione¹⁴².

Nonostante l’obiettivo dietro al loro utilizzo sia di evitare il più possibile la “*mass incarceration*”, negli Stati Uniti si vede un aumento del numero dei detenuti, tanto che la dottrina si domanda ormai da tempo se sia davvero il caso di continuare a usare tali strumenti. Le perplessità maggiori sono due: l’attendibilità degli *output* e la loro

¹⁴⁰ D. Zingales, 2021, 3; B. L. Garrett-J. Monahan, 2020, 451; A. S. Niederman Y. J. Choi-G. V. Broeck, 2019, 711

¹⁴¹ A. S. Niederman-Y. J. Choi-G. V. Broeck, 2019, 712

¹⁴² D. Zingales, 2021, 5

compatibilità con i principi del giusto processo sanciti nel 5° e nel 14° emendamento della Costituzione¹⁴³.

Il dibattito in corso oltremare interessa anche l'ordinamento italiano almeno sotto il profilo dell'affidabilità dei sistemi, che fin troppo spesso hanno applicato *bias* di natura discriminatoria basati sull'etnia e sulle condizioni socioeconomiche degli imputati. Queste storture possono portare a conseguenze non auspicabili, come la moltiplicazione di *output* affetti da *bias* a causa di errori di calcolo o di codificazione.

3.2 COMPAS e il caso *Loomis v. Wisconsin*

Uno dei casi più discussi in materia di giustizia predittiva è senza dubbio quello che vede come protagonista il signor Eric L. Loomis. Nel 2013, in seguito a una sparatoria in cui era coinvolta anche la polizia, gli vengono addebitati cinque capi d'accusa, tutti in recidiva:

- a. messa in pericolo della sicurezza;
- b. tentativo di fuga o elusione di un ufficiale del traffico;
- c. guida di un veicolo senza consenso del proprietario;
- d. possesso di arma da fuoco da parte di un pregiudicato;
- e. possesso di un fucile a canna corta o pistola¹⁴⁴.

Nel determinare la pesante pena a sei anni di reclusione e cinque di *extended supervision*, il Tribunale circondariale di La Crosse ha tenuto conto dei risultati elaborati da COMPAS, un *software* in grado di prevedere il rischio di recidiva dell'imputato¹⁴⁵.

Loomis, attraverso un'istanza di revisione della pena, sosteneva che l'utilizzo di COMPAS violasse il suo diritto a un equo processo¹⁴⁶. Dopo il rigetto dell'istanza e il conseguente

¹⁴³ D. Zingales, 2021, 6; XIV Emendamento, Sezione 1, Costituzione degli Stati Uniti d'America: "[...] *nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws*".

¹⁴⁴ S. Carrer, 2019, 4

¹⁴⁵ S. Quattrocolo, 2020, 279

¹⁴⁶ S. Carrer, 2019, 4

appello, la questione è passata alla Corte Suprema del Wisconsin, che ha emesso la sentenza definitiva il 13 luglio 2016¹⁴⁷.

3.2.1 Che cos'è e come funziona COMPAS

Prima di proseguire con l'analisi del caso, è necessario fare chiarezza su che cosa sia COMPAS e su quale sia il suo funzionamento. COMPAS sta per *Correctional offender management profiling for alternative sanctions*, è di proprietà della società californiana Equivant (fino al 2017 nota come Northpointe) ed è un programma basato su algoritmi informatici¹⁴⁸ o, volendo usare le categorie introdotte nel primo capitolo, un sistema esperto di supporto all'attività giudiziaria (DSS). È stato sviluppato nel 1998 e dal 2000 viene usato per calcolare il rischio di recidiva di un imputato nei 2 anni successivi al test: da allora ha valutato più di 1 milione di persone¹⁴⁹.

Per poterlo usare, il giudice del Tribunale circondariale aveva ordinato un *Pre-sentencing Investigation Report* (PSI), una relazione utilizzata per modulare la severità della pena composta dai risultati delle indagini fatte sulla storia personale di Eric Loomis, al cui interno erano contenute anche le determinazioni di COMPAS¹⁵⁰. Il programma interpreta le informazioni offerte dall'imputato tramite un questionario di 137 domande relative ai suoi dati anagrafici, alla sua vita sociale, alle sue opinioni personali e, più in generale, al suo rapporto con le droghe e la criminalità¹⁵¹.

Ma non solo. COMPAS può attingere a un archivio di dati "basato sulla raccolta per molti anni di dati statistici sui condannati, sui fattori di rischio presenti nei vari casi (alcolismo, droga, prostituzione e simili) e sulle recidive"¹⁵²: incrociando il risultato del questionario con i dati statistici di cui è già in possesso, è in grado di calcolare il rischio di recidiva dell'imputato. Queste risultanze, come ricordato in precedenza, sono utilizzate dai tribunali

¹⁴⁷ M. C. Falchi, 2020

¹⁴⁸ G. Contissa-G. Lasagni-G. Sartor, 2019, 621

¹⁴⁹ J. Dressel-H. Farid, 2018

¹⁵⁰ S. Carrer, 2019, 4

¹⁵¹ R. Celella, 2019

¹⁵² A. Santosuosso, 2020, 92

statunitensi, per esempio, per “emettere decisioni sull’assegnazione di misure alternative o sull’applicazione di istituti assimilabili alla sospensione condizionale della pena”¹⁵³.

Il concetto giuridico di recidiva è stato diviso in tre macrocategorie comprensibili e analizzabili dal sistema:

- Recidiva preprocessuale. Consiste nel rischio di mancata comparizione al processo e di nuovo arresto dell’imputato durante il periodo di *pretrial release*. I fattori che vengono attualmente presi in considerazione sono: il reato con la pena maggiore, casi pendenti, precedenti mancate comparizioni, precedenti arresti su cauzione, precedenti condanne al carcere, abuso di droga in passato, status occupazionale e durata della residenza;
- Recidiva generale. È usata per prevedere la commissione di un reato in recidiva dopo la valutazione di COMPAS. In questo caso gli elementi valutati sono: arresti precedenti, condanne precedenti al carcere, precedenti casi di libertà vigilata, problemi educativi e occupazionali, uso di droga in passato, età al momento della valutazione ed età al primo arresto;
- Recidiva violenta. Come suggerisce il nome, mira a calcolare la probabilità che l’imputato commetta reati violenti dopo il rilascio. I fattori soppesati dal programma sono: episodi di violenza e di *non-compliance* passati, problemi educativi e occupazionali, età al momento della valutazione ed età al primo arresto¹⁵⁴.

L’esito delle valutazioni di COMPAS consiste in un grafico a tre barre che, con un punteggio da uno a dieci per ciascuna recidiva, permette di osservare le corrispondenti probabilità di rischio: la fascia bassa va da uno a quattro, la fascia media va da cinque a sette e la fascia alta va da otto a dieci¹⁵⁵: Eric Loomis era stato giudicato ad alto rischio in tutti e tre i settori¹⁵⁶.

¹⁵³ G. Contissa-G. Lasagni-G. Sartor, 2019, 621

¹⁵⁴ Practitioner’s Guide to COMPAS Core, 2019, 31

¹⁵⁵ Practitioner’s Guide to COMPAS Core, 2019, 4

¹⁵⁶ S. Carrer, 2019, 4

Resta da capire come interpretare correttamente questi dati. La risposta viene fornita, almeno per sommi capi, dalla guida del *software*. Vengono distinti tre livelli di interpretazione, dal più superficiale al più profondo:

- Primo livello. Ci si limita a valutare il punteggio ottenuto dal soggetto, senza mettere in correlazione i vari fattori presi in considerazione da COMPAS. Si tratta di un'interpretazione semplicistica che non tiene conto delle possibili cause del fenomeno criminoso, ciononostante può essere utile a identificare le aree su cui proseguire l'analisi;
- Secondo livello. Si studiano le connessioni tra i vari fattori di rischio e si valutano gli interventi da adottare nelle aree di bisogno più delicate;
- Terzo livello. Grazie all'utilizzo di teorie criminologiche utili a comprendere il comportamento dell'imputato, si studiano le possibili cause sottostanti al reato. Si tratta di un "*mix of explanatory theories that help 'connect the dots' of need and other influencing factors to paint a picture of the individual's pathway to crime*"¹⁵⁷.

Nel PSI era stabilito che le determinazioni di COMPAS non dovessero essere utilizzate dal giudice per modulare la severità della pena, sottolineando, in questo modo, la natura meramente accessoria del *software*¹⁵⁸.

3.2.2 La posizione della difesa e la sentenza definitiva

Una volta capito cosa sia COMPAS e quali siano i fatti relativi al caso Loomis, è utile vedere quali sono le critiche avanzate dalla difesa nel corso del processo conclusosi davanti alla Corte Suprema del Wisconsin.

Possono essere riassunte in tre punti:

- a. COMPAS è il programma di una società privata coperto dal segreto commerciale, di conseguenza la difesa non può conoscere nel dettaglio il modo in cui decide e valutare

¹⁵⁷ Practitioner's Guide to COMPAS Core, 2019, 4

¹⁵⁸ S. Carrer, 2019, 4

- se i suoi *output* sono corretti. Questo, secondo la difesa, porta alla violazione del diritto costituzionale al giusto processo;
- b. Le determinazioni del sistema non possono essere prese in considerazione perché non portano a una sentenza individualizzata (come sancito dai principi del giusto processo), essendo pensato per valutare dati di gruppo;
 - c. COMPAS fa un uso controverso del genere ai fini della decisione, portando, anche in questo caso, alla violazione del diritto a un giusto processo¹⁵⁹.

In sostanza, ciò a cui mira la difesa è mettere in discussione l'affidabilità del *software* cercando di dimostrare l'opacità delle sue conclusioni ponendo l'accento sulla natura privatistica dello strumento, sul modo in cui elabora le informazioni e sulla presunta discriminatorietà dei risultati. Citando un passo della sentenza che riassume efficacemente questa posizione: "*Loomis contends that unless he can review how the factors are weighed and how risk scores are determined, the accuracy of the COMPAS assessment cannot be verified*"¹⁶⁰.

La difesa, nell'argomentare il primo punto, traccia un'analogia tra il caso corrente e due precedenti importanti in materia: il caso Gardner v. Florida del 1977 e il caso State v. Skaff del 1989. Nel primo, l'imputato era stato condannato a morte in base alle informazioni contenute in un PSI che gli erano parzialmente ignote: il giudice Stevens, nel redigere il suo parere, ha scritto che questo comportamento da parte della Corte costituisce una violazione del diritto al giusto processo dell'imputato. Nella sentenza verrà poi scritto che "[t]he defendant ha[d] a legitimate interest in the character of the procedure which leads to the imposition of sentence even if he may have no right to object to a particular result of the sentencing process"¹⁶¹, cristallizzando così un principio che troverà spazio anche nel caso del 1989.

In State v. Skaff, infatti, il PSI era stato consegnato al difensore con l'avvertenza, da parte del tribunale, di non farlo leggere al proprio assistito (condannato in primo grado per spaccio

¹⁵⁹ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 127

¹⁶⁰ State v Loomis, 2016, 21

¹⁶¹ Gardner v Florida, 1977, 430 U. S. 358

e possesso di cocaina). Secondo la Corte d'Appello negare a Skaff la possibilità di conoscere il contenuto del suo *report* significa negargli pregiudizialmente il diritto di ricevere una decisione basata su informazioni corrette: qualsiasi errore del PSI potrebbe influenzare la sentenza, incidendo in tal modo sul futuro dell'imputato senza che egli possa dire o fare alcunché. Nel caso Skaff si è nuovamente chiarito che l'imputato deve avere i mezzi necessari per esaminare l'accuratezza della sentenza e i fattori che la influenzano: impedirglielo significherebbe violare i principi del giusto processo¹⁶².

Loomis, nell'evidenziare l'analogia tra la sua posizione e quella di Gardner e Skaff, afferma che Northpointe, non rivelando come COMPAS determini i fattori di rischio, non permetterebbe alla difesa di accedere a informazioni fondamentali per il giudizio, violando in tal modo il proprio diritto a subire un giusto processo¹⁶³. In aggiunta, per confutare i risultati del *software*, a detta della difesa non sarebbe sufficiente basarsi sui grafici conclusivi, ma *"unless he can review how the factors are weighed and how risk scores are determined, the accuracy of the COMPAS assessment cannot be verified."*¹⁶⁴. Insomma, ciò che Loomis chiede è di poter analizzare nel dettaglio i processi logico-computazionali che stanno alla base delle valutazioni del *software*.

La Corte Suprema, dopo aver ammesso l'impossibilità di comprendere fino in fondo il modo in cui COMPAS utilizza i dati a sua disposizione per calcolare i punteggi di rischio¹⁶⁵, ha affermato che, come scritto nella guida del sistema, i punteggi sono basati perlopiù su dati statistici come la condotta criminale, con un uso limitato di variabili dinamiche come l'assunzione di stupefacenti. In più, il *report* fornito da COMPAS contiene una lista di 21 domande e risposte riguardanti fattori di cui Loomis poteva verificare l'accuratezza, eventualmente fornendo informazioni e dati contrastanti¹⁶⁶. Di conseguenza, citando nuovamente la sentenza: *"to the extent that Loomis's risk assessment is based upon his answers to questions and publicly available data about his criminal history, Loomis had the*

¹⁶² K. Freeman, 2016, 87

¹⁶³ State v Loomis, 2016, 21

¹⁶⁴ State v Loomis, 2016, 22

¹⁶⁵ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 127

¹⁶⁶ S. Carrer, 2019, 4

*opportunity to verify that the questions and answers listed on the COMPAS report were accurate*¹⁶⁷. In sostanza, visto che molte delle informazioni usate del *software* sono pubbliche o fornite dallo stesso imputato, la Corte ritiene che Loomis abbia già avuto l'opportunità di verificarle.

Per quanto riguarda il secondo punto, i giudici, nel ribadire l'importanza essenziale del diritto a un giudizio individualizzato, precisano che i calcoli di COMPAS sono solo uno dei tanti elementi che compongono la valutazione finale del giudice, che deve tener conto dei risultati del sistema con riguardo ad ogni specifico individuo, considerando anche tutti gli altri elementi a disposizione¹⁶⁸. Hanno inoltre tracciato un parallelismo tra la sentenza *State v Gallion* del 2004 e la funzione del *software*, affermando la necessità, da parte dei giudici, di avere quante più informazioni possibili prima della sentenza di condanna¹⁶⁹. In tal senso, COMPAS permetterebbe ai decisori di avere un quadro più esaustivo per giudicare al meglio.

La terza e ultima critica di Loomis riguarda l'uso scorretto del genere nelle valutazioni del sistema poiché, essendo gli uomini quelli statisticamente ritenuti più inclini alla recidiva, i giudici avrebbero basato la sentenza su un'illegittima discriminazione di genere¹⁷⁰. Anche in questo caso, la Corte si è allineata con le scelte del Tribunale di La Crosse perché, a detta dei giudici, la difesa non è riuscita a dimostrare un'incidenza così determinante del genere nella condanna¹⁷¹. Al contrario, secondo la Corte, il dato è stato usato per migliorare l'accuratezza della decisione: *"if the inclusion of gender promotes the accuracy, it serves the interests of institutions and defendants, rather than a discriminatory purpose"*¹⁷². Di conseguenza *"COMPAS's use of gender promotes accuracy that 'ultimately inures to the benefit of the justice system including defendants'"*¹⁷³.

Nel corso del processo sono stati toccati anche altri punti importanti con riferimento all'affidabilità del programma. Nonostante la difesa avesse prodotto in giudizio una serie di

¹⁶⁷ *State v Loomis*, 2016, 23

¹⁶⁸ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 127

¹⁶⁹ *State v Loomis*, 2016, 30

¹⁷⁰ S. Carrer, 2019, 4

¹⁷¹ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 128

¹⁷² *State v Loomis*, 2016, 34

¹⁷³ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 128

studi che ne contestavano la precisione, i giudici hanno ritenuto più corretti alcuni test effettuati in altri stati americani, i quali avevano concluso che, seppur con qualche riserva, COMPAS rappresentasse un mezzo di calcolo affidabile. Per quanto concerne il rischio che il sistema attribuisca troppa importanza ad alcuni fattori, come ad esempio l'ambiente familiare, il livello di educazione o l'appartenenza ad un certo gruppo etnico, la Corte ha ribadito la necessità di un monitoraggio e un aggiornamento costante del programma sulla base dei cambiamenti sociali più recenti, a cui deve far seguito un uso attento e cauto da parte dei giudici¹⁷⁴.

Nonostante la Corte Suprema ammetta le lacune e le falle che affliggono COMPAS, ha deciso, a tratti controintuitivamente, di ammetterne comunque l'utilizzo all'interno dei tribunali del Wisconsin. Ciò su cui i giudici supremi hanno posto l'accento sono le cautele che devono essere adottate quando si prendono in considerazione le valutazioni dei programmi predittivi. Sono tre le occasioni in cui ciò può avvenire:

- per la comminazione di misure alternative alla detenzione, nel caso di imputati a basso rischio di recidiva;
- per valutare la possibilità di reinserire un condannato in modo sicuro all'interno della società, anche con l'affidamento in prova;
- per l'imposizione di termini e condizioni per la libertà vigilata, per la supervisione e per le eventuali sanzioni alle violazioni delle regole previste dai regimi alternativi alla detenzione¹⁷⁵.

Di conseguenza, uno strumento come COMPAS può, per esempio, aiutare a stabilire quale sia il livello di supervisione adatto ai condannati a basso rischio in regime di libertà vigilata, favorendo in tal modo programmi di reinserimento sociale più efficaci. Come stabilito anche nella sentenza *Malenchik v. State* del 2010, i *risk assesment tool* possono essere usati per “*enhance a judge's evaluation, weighing, and application of the other sentencing evidence*”

¹⁷⁴ S. Carrer, 2019, 4

¹⁷⁵ *State v Loomis*, 2016, 36

*in the formulation of an individualized sentencing program appropriate for each defendant*¹⁷⁶.

Oltre a ribadire il fatto che COMPAS sia uno dei tanti elementi che i giudici sono tenuti a valutare nel processo decisionale, la Corte sottolinea l'importanza di non usare in alcun modo questi *tool* né per determinare la severità o l'entità della pena, né per decidere in merito all'incarcerazione dell'imputato. Il suo scopo, infatti, è individuare le esigenze del reo e valutare il suo rischio di reiterazione del reato¹⁷⁷. Se è stato necessario stabilirlo in modo così netto, evidentemente i giudici supremi hanno riscontrato il rischio di un appiattimento delle sentenze sulle valutazioni (difficilmente spiegabili appieno) di tali algoritmi. Ma non solo. Per rafforzare l'idea che i sistemi sono degli strumenti delicati e lontani dall'essere perfetti, ogni PSI contenente il report di COMPAS, secondo i giudici supremi del Wisconsin, deve avere all'interno una nota con tutte le limitazioni al suo utilizzo stabilite dalla sentenza¹⁷⁸.

La Corte, nonostante questi timori, ha ritenuto che nel caso di specie il punteggio di rischio sia stato solo uno dei molteplici fattori considerati dal Tribunale circondariale di La Crosse nel determinare la pena, fattori che Loomis non aveva contestato. È stato inoltre osservato che, sebbene i giudici avessero usato tali valutazioni per determinare la pena, abbiano attribuito loro un peso minimo, affermando addirittura che la sentenza non sarebbe stata diversa in assenza dei dati forniti da COMPAS¹⁷⁹. Questa decisione stride con il ragionamento portato avanti fin qui: se il programma non può essere usato per determinare la gravità della pena, non si spiega come mai i giudici abbiano deciso di allinearsi con il Tribunale di La Crosse dopo aver constatato la violazione della regola appena espressa. Poco importa se sia o meno stato influente nell'insieme dei fattori che hanno portato alla condanna: l'uso scorretto, alla luce dei fatti, c'è comunque stato. In caso contrario, la forza di tale disposizione risulta fortemente limitata, vista la possibilità di aggirarla con relativa agilità.

¹⁷⁶ Malenchik v State, 2010, 11

¹⁷⁷ S. Carrer, 2019, 4

¹⁷⁸ State v Loomis, 2016, 40

¹⁷⁹ S. Carrer, 2019, 4

Alcuni commentatori si spingono ancora oltre, affermando che la Corte ha eretto una protezione troppo debole o addirittura inutile contro le valutazioni di COMPAS e, a cascata, dei sistemi a lui analoghi¹⁸⁰. Se a questo aggiungiamo il motivo per cui vengono utilizzati, ossia poter avere il parere di qualcosa che sulla carta dovrebbe sbagliare meno dell'uomo, si rischia di fare troppo affidamento su degli strumenti di cui si ignorano le criticità¹⁸¹. Nel corso della storia statunitense, si è visto che servono generalmente a poco le raccomandazioni ai giudici come quelle formulate dalla Corte Suprema del Wisconsin. Un caso classico per spiegare questo concetto è Stati Uniti v Rodriguez del 1978, in cui è stata usata l'espressione "*cannot unring a bell*": è difficile riuscire a far dimenticare alla giuria o al giudice qualcosa che ha visto o sentito, sarebbe come cercare di dimenticare una ferita dopo aver subito un fendente di spada o aver sentito il suono di un campanello¹⁸². Insomma, una semplice dichiarazione non ha l'influenza che la Corte Suprema del Wisconsin presume di avere quando decide di emettere dichiarazioni cautelative al posto di riforme effettive dell'uso di COMPAS in aula.

In questi casi va preso in considerazione il cd. effetto ancoraggio che i giudici subiscono quando usano strumenti scientifici o tecnologici. Numerosi studi hanno dimostrato che i giudici (e, più in generale, le persone) sono altamente influenzabili dai numeri e dai risultati generati da un computer o da uno strumento che ritengono affidabile, col risultato di vedere magistrati aderire, come detto prima, a ciò che dice la macchina solo perché lo dice la macchina. Gli avvertimenti scritti, ancora una volta, non sembrano in grado di arginare questo fenomeno, soprattutto quando chi decide non possiede le competenze tecniche utili a comprendere il funzionamento di ciò che usa. Eppure la Corte si è dimostrata insensibile a

¹⁸⁰ K. Freeman, 2016, 89

¹⁸¹ K. Freeman, 2016, 99; Si veda anche L. J. Skitka *et al.*, 1999, 992 in cui si afferma: "*In fact, there is some evidence to suggest that the introduction of automated decision aids does not unilaterally lead to a reduction in human error, but instead often creates opportunities for simply a different class of errors. For example, various problems with automated decision aids have been identified, including mode misunderstandings and mode errors; failures to understand automation behavior; confusion or lack of awareness concerning what automated aids are doing and why; and difficulty associated with tracing the reasoning processes and functioning of automated agents*".

¹⁸² K. Freeman, 2016, 99

questa realtà, ritenendo verosimile la tesi secondo cui il Tribunale avrebbe imposto la stessa pena anche senza i punteggi di COMPAS¹⁸³.

A gettare ulteriori dubbi sull'utilizzo del programma sono le ricerche e le inchieste che attestano la discriminatorietà dei suoi risultati, che sovrastimerebbero il rischio di recidiva degli imputati di colore, sottostimando, al contempo, quello degli imputati bianchi. Il razzismo insito nel *software* e la discriminazione di genere sollevata da Loomis potrebbero far scattare un allarme per il sistema penale statunitense. Come ribadito in un recente caso deciso dalla Corte Suprema degli Stati Uniti, *Buck v Davis*: "*our criminal law punishes people for what they do, not who they are*"¹⁸⁴.

Leggendo la sentenza, sembra quasi che COMPAS venga considerato alla stregua di un *expert witness*, una figura tipica del sistema processuale statunitense che assomiglia ai periti e ai consulenti tecnici presenti nell'ordinamento italiano. Tuttavia, a differenza dei testimoni esperti, il programma non può essere sottoposto a controinterrogatorio per le opinioni che esprime, nonostante vengano prese in grande considerazione dai giudici, tanto che possono arrivare a influenzare la sentenza finale più di quanto farebbe un testimone esperto. La Corte, insomma, ha considerato i punteggi di rischio come dei dati di fatto, senza dare la possibilità alla difesa di esaminare come sono stati calcolati.

È utile volgere lo sguardo a un'esperienza diversa da quella statunitense, in cui il problema dei *tool* predittivi è stato affrontato in modo più cauto e, forse, corretto. Nel caso *Director of Public Prosecutions for Western Australia v Mangolamara* del 2007 la questione riguardava due perizie psichiatriche effettuate a una persona, appartenente alla comunità aborigena, condannata per numerose violenze sessuali. La Corte Suprema dell'Australia Occidentale, nel parlare di questi strumenti di valutazione, ha sottolineato che "*the research data and methods underlying the assessment tools are assumed to be correct but this has not been established by the evidence*"¹⁸⁵ e ha proseguito sostenendo che i *tool* "*were not*

¹⁸³ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 130

¹⁸⁴ *Buck v Davis*, 2017, 4

¹⁸⁵ *Director of Public Prosecutions for Western Australia v Mangolamara*, 2007, 165

devised for and do not necessarily take account of the social circumstances of indigenous Australians in remote communities"¹⁸⁶.

È piuttosto evidente la diversa prospettiva adottata dalla Corte australiana rispetto a quella adottata dalla Corte statunitense: nel primo caso prevale un atteggiamento di cautela, nel secondo un atteggiamento fiducioso. I giudici australiani arrivano a sostenere che hanno *"grave reservations as to whether a person of the respondent's background can be easily fitted within the categories of appraisal presently allowed for by the assessment tools"*¹⁸⁷, dubitando in tal modo degli stessi presupposti che hanno portato alle valutazioni peritali.

3.3 Le critiche a COMPAS

ProPublica, un'organizzazione senza scopo di lucro statunitense operante nel campo del giornalismo investigativo di interesse pubblico, nel 2016 ha pubblicato un'analisi (poi ripresa da numerosi studi scientifici) che evidenzia in modo piuttosto netto la discriminatorietà delle valutazioni di COMPAS a discapito dei cittadini afroamericani¹⁸⁸.

Per spiegare i motivi per cui un programma può essere affetto da questi difetti, si è soliti distinguere tra pregiudizi indiretti e diretti. I primi nascono dalle premesse: se queste sono fallaci o errate, lo saranno anche le conclusioni. Generalmente nascono dall'attività (consapevole o meno) dei programmatori che si occupano dello sviluppo dei *software*: è la stessa architettura a essere compromessa. In astratto, se esistessero solo pregiudizi indiretti, si potrebbero arginare con una certa facilità esercitando un controllo più rigoroso sulle premesse¹⁸⁹.

I secondi, invece, sono più insidiosi perché interessano il campione di dati a disposizione dell'algoritmo, che è legato al contesto sociale e storico in cui è elaborato ed è destinato a mutare con esso per farsi portatore di nuovi e diversi pregiudizi. Basti pensare al giudice che sia più propenso ad affidare il figlio alla madre piuttosto che al padre: la scelta non è

¹⁸⁶ Director of Public Prosecutions for Western Australia v Mangalamara, 2007, 166

¹⁸⁷ Director of Public Prosecutions for Western Australia v Mangalamara, 2007, 166

¹⁸⁸ J. Angwin-J. Larson-S. Mattu-L. Kirchner, 2016

¹⁸⁹ C. Cavaceppi, 2020, 106

necessariamente frutto di un pregiudizio nei confronti del padre, ma può essere il riflesso dell'esistenza di fattori psicosociali, economici e culturali del territorio in cui opera, come "la presenza o meno di servizi sociali, il fatto che il bambino vada a scuola oppure no, l'orario di lavoro di ciascun genitore, o anche semplicemente la mancanza di interesse da parte di uno dei due genitori a farsi carico del figlio piccolo"¹⁹⁰.

Di conseguenza, la questione non è arginare il pregiudizio diretto, poiché si rischierebbe di rendere la norma giuridica totalmente avulsa dalla realtà e dal sentire sociale, ma capire come l'IA possa evolversi pur essendo stata creata con dati e valori superati¹⁹¹. Possono essere d'aiuto le moderne tecnologie di *machine* e *deep learning*: le prime permettono ai sistemi di apprendere nuove informazioni partendo dai dati di cui sono già in possesso; le seconde, decisamente più raffinate delle prime, consentono alla macchina di imparare cose nuove simulando il funzionamento di una rete neurale, "*enabling machines to learn more about a given set of observations and make accurate predictions*"¹⁹².

3.3.1 L'inchiesta di ProPublica

Stando all'inchiesta di ProPublica, che merita di essere citata per il ruolo decisivo che ha avuto nel dibattito sul tema, COMPAS sarebbe del tutto inadatto al ruolo che, di fatto, ricopre nelle aule di giustizia, ossia influenzare la qualificazione e la quantificazione delle sanzioni penali inflitte dai giudici statunitensi. Si è visto che, in generale, le persone con la pelle scura vengono considerate più ad alto rischio di recidiva rispetto alle persone con la pelle chiara, e talvolta la differenza è notevole. Per citare un caso emblematico, Brisha Borden, una ragazza diciottenne di colore della Florida, è stata valutata con un rischio 8 dopo aver commesso qualche reato minore, mentre Vernon Prater, un uomo bianco con alle spalle delle condanne per rapina a mano armata, è stato valutato con un rischio 3. Dopo due

¹⁹⁰ C. Cavaceppi, 2020, 106

¹⁹¹ C. Cavaceppi, 2020, 107

¹⁹² G. Naveen-M. A. Naidu-B. T. Rao-K. Radha, 2019, 1985

anni, Prater è stato condannato a otto anni di reclusione per un furto, mentre Borden non ha commesso alcun reato¹⁹³.

Ma non solo. ProPublica ha analizzato i punteggi di rischio assegnati a più di settemila persone arrestate nella contea di Broward, in Florida, dal 2013 al 2014 e ha controllato quanti di costoro sono stati imputati in nuovi processi nei due anni successivi. Per quanto riguarda i delitti violenti, i punteggi si sono dimostrati poco precisi: solo il 20% delle persone identificate come ad alto rischio hanno effettivamente commesso altri reati. Se si estende l'analisi senza limitarsi a una singola categoria di delitti, l'algoritmo si è dimostrato poco più preciso del lancio di una moneta¹⁹⁴.

COMPAS sembra, quindi, avere dei pregiudizi nei confronti di specifiche etnie, tanto che gli imputati di colore vengono erroneamente segnalati come ad alto rischio il doppio delle volte rispetto agli imputati bianchi e viceversa: gli imputati bianchi vengono catalogati erroneamente come a basso rischio più spesso degli imputati di colore. Questo stesso fenomeno lo si può vedere anche isolando alcuni specifici reati, come quelli violenti. In questo caso gli imputati di colore hanno il 77% di probabilità in più di venir considerati ad alto rischio rispetto agli imputati bianchi¹⁹⁵.

Se a questo aggiungiamo una certa opacità dei calcoli che effettua il sistema per determinare i punteggi, ci si trova davanti a un quadro abbastanza problematico: non solo la macchina è affetta da pregiudizi, ma le persone interessate non ne sanno nemmeno il motivo, pur subendone le conseguenze durante i processi. Quelli appena citati sono tutti dati che vanno presi con le dovute precauzioni in quanto frutto di un'inchiesta giornalistica e non di uno studio accademico. Nonostante questo, hanno il pregio di delineare la direzione che ha preso il dibattito sul tema nel 2016, subito dopo la sentenza Loomis: si è iniziato a guardare con occhio più critico l'opportunità di usare sistemi come COMPAS in giudizio.

¹⁹³ J. Angwin-J. Larson-S. Mattu-L. Kirchner, 2016

¹⁹⁴ Han-Wei Liu-Ching-Fu Lin-Yu-Jie Chen, 2019, 131

¹⁹⁵ J. Angwin-J. Larson-S. Mattu-L. Kirchner, 2016

Il paradosso è che se, al contrario, il sistema funzionasse correttamente, la giustizia penale sarebbe più equa perché scevra di tutti quei pregiudizi tipicamente umani che influenzano, anche inconsciamente, le sentenze dei giudici.

Com'è ampiamente noto in psicologia, quando un essere umano fa delle valutazioni o prende delle decisioni, entrano frequentemente in gioco le cosiddette euristiche di giudizio. Si tratta di procedure che, da un lato, il sistema cognitivo adotta per lavorare in modo più efficiente (come se prendesse delle scorciatoie di pensiero), ma che, dall'altro lato, espongono a errori di valutazione: i cd. *bias*. Le tre principali euristiche sono quella della disponibilità, della rappresentatività e di ancoraggio¹⁹⁶. L'euristica della disponibilità o evocabilità porta a delle distorsioni di giudizio quando si prende in considerazione un evento saliente, temporalmente vicino e vivido: le persone valuteranno come maggior rischio viaggiare in aereo piuttosto che in automobile perché gli incidenti aerei hanno un forte impatto emotivo e perché i media vi riservano molta attenzione, nonostante i dati dimostrino l'oggettiva erroneità di questa idea. Si sarà portati a sovrastimare la pericolosità dei voli anche nel caso in cui l'incidente si sia appena verificato o se questo venisse raccontato con dovizia di particolari¹⁹⁷.

Tutto ciò, inutile dirlo, ha degli effetti: le correlazioni illusorie e l'illusione di controllo. Le correlazioni illusorie portano a pensare che, per esempio, assumere droga ed essere delinquenti siano due caratteristiche concomitanti solo perché ricordiamo un certo numero di individui tossicodipendenti con precedenti penali. Tuttavia, per poter tracciare una vera correlazione tra i due elementi, bisognerebbe sapere quante sono le persone che assumono droga e sono delinquenti, che assumono droga e non sono delinquenti, che non assumono droga e sono delinquenti, che non assumono droga e non sono delinquenti¹⁹⁸. L'illusione di controllo, invece, consiste nella tendenza degli esseri umani a ritenere di poter esercitare un qualche tipo di controllo su eventi casuali. Questo conduce all'applicazione di svariati *bias*, tra cui quello ottimistico che, in breve, porta a una sovrastima delle proprie capacità nelle situazioni di pericolo, come in sala operatoria e nei reparti di malattie infettive,

¹⁹⁶ C. Bona-R. Rumiati, 2013, 151

¹⁹⁷ C. Bona-R. Rumiati, 2013, 152

¹⁹⁸ C. Bona-R. Rumiati, 2013, 155

sottostimando la probabilità di subire le conseguenze negative derivanti da determinati comportamenti inadeguati¹⁹⁹.

L'euristica della rappresentatività valorizza il grado di similarità tra l'evento da stimare e il processo che l'ha generato o la categoria di riferimento. Kahneman e Tversky hanno spiegato questo fenomeno con un esperimento: a un gruppo di persone sono state sottoposte due sequenze di lanci di una moneta dove T sta per testa e C sta per croce, la prima era T-T-T-T-C-C-C-C, mentre la seconda era T-T-C-T-C-C-T-C. La maggioranza ha ritenuto la seconda sequenza più probabile nonostante, da un punto di vista statistico, le due serie abbiano le stesse possibilità di realizzarsi poiché tutti i lanci sono eventi indipendenti. La spiegazione a queste risposte è che la seconda serie corrisponde all'idea di "caso" che si associa al lancio di una moneta. Le conseguenze pratiche di questa euristica sono molteplici e in questa sede è bene ricordarne una: la cd. fallacia del giocatore d'azzardo, ossia la convinzione che dopo una serie di eventi dello stesso segno (come il risultato "testa" nel lancio della moneta), quegli eventi diventeranno meno probabili nella prova successiva. Una sua variazione è l'illusione della "mano calda", un fenomeno che si può riscontrare tra i tifosi di pallacanestro che associano un canestro appena segnato da un giocatore a una più alta probabilità, per quello stesso giocatore, di segnarne un altro in poco tempo²⁰⁰.

L'euristica di ancoraggio e aggiustamento si nota quando un individuo formula una prima valutazione di un fatto, l'ancoraggio, e dopo l'acquisizione di ulteriori informazioni, modifica di poco (gli aggiustamenti) l'idea iniziale fino a emettere una decisione definitiva. La peculiarità di questo fenomeno si nota soprattutto quando le nuove informazioni in possesso del soggetto sono talmente tante e talmente incisive che ci si aspetterebbe una valutazione diversa rispetto a quella iniziale, ma ciò non accade, mentre l'aspetto più stupefacente è che talvolta vengono assunti come ancoraggi degli elementi del tutto arbitrari che non hanno alcuna relazione con l'oggetto della decisione²⁰¹. Ci sono state molteplici sperimentazioni di questa euristica in ambito giudiziale, tanto in campo civile

¹⁹⁹ C. Bona-R. Rumiati, 2013, 156

²⁰⁰ C. Bona-R. Rumiati, 2013, 166

²⁰¹ C. Bona-R. Rumiati, 2013, 178

quanto in campo penale. In ambito risarcitorio, negli Stati Uniti è stato rilevato che l'ammontare della domanda influisce sulla liquidazione del danno: maggiore è la somma richiesta a titolo di risarcimento dalla parte attrice, più elevata sarà la somma liquidata dal giudice. In ambito penale, invece, in un esperimento condotto da English e Mussweiler nel 2001 si è visto che, sottoponendo ai partecipanti dei casi realistici di violenza sessuale e facendogli assumere il ruolo di giudici, si sono dimostrati fortemente influenzati dalle conclusioni dell'accusa: maggiore era l'entità della pena richiesta dai magistrati requirenti, maggiore era la pena irrogata agli imputati²⁰².

Questi processi mentali, che in un modo o nell'altro affliggono le sentenze dei giudici umani, devono essere noti agli ingegneri e non devono essere presenti quando i sistemi sono operativi, specie se vengono demandate loro funzioni di supporto all'attività dei magistrati giudicanti. Per poter fare affidamento su una macchina, questa non deve sbagliare o comunque deve farlo meno dell'uomo e, soprattutto, deve essere chiaro agli occhi di un esperto il motivo per cui sbaglia applicando, come nel caso di COMPAS, discriminazioni del tutto sconnesse dalla realtà²⁰³. Per questo è necessario che il funzionamento del sistema non sia coperto da segreto industriale e, soprattutto, che sia possibile instaurare un contraddittorio sui risultati a cui giunge. Risultati che a volte sorprendono gli stessi imputati. È il caso di James Rivelli, un uomo di 54 anni di Hollywood, in Florida, arrestato per un furto in una farmacia. Pur avendo alle spalle numerosi precedenti che includono aggressioni, furti, traffico di droga e una condanna a cinque anni di reclusione nel Massachusetts, COMPAS lo valutò con un rischio 3: meno di un anno dopo fu accusato nuovamente di due furti dal valore complessivo di mille dollari²⁰⁴.

Nonostante il sistema non dovesse essere usato per modulare la pena in base al punteggio, ma solo per valutare l'idoneità dell'imputato alla libertà vigilata o a programmi di trattamento, nel 2012 il Wisconsin Department of Corrections ne ha esteso l'applicazione a tutte le fasi del processo: dalla sentenza alla libertà vigilata. Questo ha portato a processi come il caso Loomis, in cui, come detto in precedenza, la Corte Suprema ha dovuto fissare

²⁰² C. Bona-R. Rumiati, 2013, 180

²⁰³ S. Vinciguerra, 2013, 90

²⁰⁴ J. Angwin-J. Larson-S. Mattu-L. Kirchner, 2016

delle regole per l'utilizzo di detti programmi. Come ha dichiarato l'assistente del Procuratore generale del Wisconsin, Christine Remington, durante le discussioni della Corte nel caso Loomis: *"We don't want courts to say, this person in front of me is a 10 on COMPAS as far as risk, and therefore I'm going to give him the maximum sentence"*²⁰⁵.

3.3.2 Gli studi successivi

L'inchiesta di ProPublica ha trovato conferma in numerosi studi successivi, tra i quali spicca quello condotto nel 2018 da Julia Dressel e Handy Farid pubblicato su Science Advances. L'obiettivo dei ricercatori era comparare le valutazioni di COMPAS con quelle di un gruppo di persone, stabilendo in tal modo se il programma (il cui tasso di precisione, secondo ProPublica, si aggira intorno al 67% per gli imputati bianchi e intorno al 63,8% per gli imputati di colore) fosse davvero più accurato degli esseri umani²⁰⁶.

A ciascuno dei mille volontari che hanno partecipato (e a COMPAS) sono state mostrate le descrizioni di alcuni imputati che includevano informazioni sul sesso, l'età e i precedenti penali senza fare alcun cenno all'etnia. In tutto anche gli imputati valutati erano mille e le descrizioni erano formulate nel seguente modo: *"The defendant is a [SEX] aged [AGE]. They have been charged with: [CRIMECHARGE]. This crime is classified as a [CRIMINAL DEGREE]. They have been convicted of [NON-JUVENILE PRIOR COUNT] prior crimes. They have [JUVENILE- FELONY COUNT] juvenile felony charges and [JUVENILE-MISDEMEANOR COUNT] juvenile misdemeanor charges on their record"*. Nella seconda versione dell'esperimento è stata esplicitata l'etnia e, in ogni caso, è stato spiegato in cosa consiste il reato ascritto all'imputato²⁰⁷.

I risultati parlano chiaro: COMPAS non è molto più preciso dei volontari. La loro accuratezza media è del 62,8%, mentre quella del programma è del 65,2%. Ma c'è di più. I volontari avevano a disposizione solo 7 fattori per farsi un'idea sulla probabilità di recidiva, mentre

²⁰⁵ J. Angwin-J. Larson-S. Mattu-L. Kirchner, 2016

²⁰⁶ J. Dressel-H. Farid, 2018

²⁰⁷ J. Dressel-H. Farid, 2018

COMPAS durante i processi ne ha 137: questo significa che un gruppo di persone con una manciata di informazioni è efficiente quasi quanto uno dei *software* intelligenti più usati negli Stati Uniti. Questo è ciò che si può affermare con riferimento alla precisione, ma il test ha misurato anche l'equità delle valutazioni. In assenza di indicazioni sull'etnia degli imputati, sia i partecipanti che COMPAS si sono dimostrati, dati alla mano, ugualmente discriminatori, dimostrando un tasso di falsi positivi (persone ritenute a rischio senza che abbiano commesso altri reati) e falsi negativi (persone ritenute non a rischio ma che hanno commesso altri reati) molto simile. In generale, il tasso di falsi positivi degli imputati di colore è sempre più alto rispetto a quello degli imputati bianchi, i quali hanno un tasso di falsi negativi più alto rispetto a quelli di colore²⁰⁸.

Un altro dato di assoluto rilievo è fornito da un esperimento successivo in cui un gruppo di 400 volontari ha dovuto valutare il rischio di recidiva degli imputati conoscendo il dato sull'etnia. Nonostante questo, i risultati non hanno mostrato un significativo inasprimento delle discriminazioni, anzi: allo stato attuale *“there is no sufficient evidence to suggest that including race has a significant impact on overall accuracy or fairness. The exclusion of race does not necessarily lead to the elimination of racial disparities in human recidivism prediction”*²⁰⁹.

Equivant, a causa del segreto industriale, non divulga apertamente l'esatto funzionamento di COMPAS ma è comunque possibile fare delle ipotesi in tal senso. Lo studio in esame ha scoperto che un semplice predittore lineare a conoscenza delle stesse informazioni dei volontari nella versione del test senza il dato sull'etnia, quindi con 7 fattori a disposizione, produce più o meno gli stessi risultati di COMPAS: quest'ultimo ha una precisione del 65% circa, mentre il predittore lineare si attesta al 66,6%²¹⁰. La prima considerazione che si può fare è che nonostante COMPAS operi con molte più informazioni del predittore, non risulta affatto più efficace, anzi: si può arrivare a supporre che COMPAS si comporti, a tutti gli

²⁰⁸ J. Dressel-H. Farid, 2018

²⁰⁹ J. Dressel-H. Farid, 2018

²¹⁰ J. Dressel-H. Farid, 2018

effetti, come il *software* utilizzato da Julia Dressel e Handy Farid nell'esperimento, il che è abbastanza impressionante.

3.4 Lo sviluppo di PATTERN e i dubbi costituzionali in merito ai *risk assesment tool*

Il *First Step Act* (o *Formerly Incarcerated Reenter Society Transformed Safely Transitioning Every Person Act*), una riforma federale statunitense nata nel 2018 che ha come obiettivo il contrasto alla *mass incarceration*, alle discriminazioni etniche e che vuole incentivare il ricorso alle misure alternative alla detenzione tramite i programmi di reinserimento elaborati dal Dipartimento di Giustizia²¹¹, ha affidato proprio a dei sistemi algoritmici un ruolo determinante per il conseguimento di tali finalità.

Il Congresso ha autorizzato lo sviluppo di un nuovo *risk assesment tool* destinato al *Bureau of Prisons* (BOP) per svolgere valutazioni in materia di *early release* e di riduzioni di pena a tutti quei soggetti detenuti negli istituti penitenziari federali con un rischio di recidiva minimo o basso²¹². Nel 2019 vede quindi la luce PATTERN, un nuovo programma di natura istituzionale non coperto, per questo motivo, dal segreto industriale. La trasparenza pare essere uno dei punti su cui il Dipartimento di Giustizia ha voluto puntare, tanto che durante lo sviluppo si sono tenute tre *listening session* allo scopo di sentire le opinioni e i contributi degli esperti del settore e delle organizzazioni rappresentative delle vittime di reati e di interessi pubblici²¹³.

Se è vero che PATTERN rappresenta un passo in avanti consistente rispetto a quanto sperimentato con COMPAS, è allo stesso modo vero che non è esente da difetti. Nello specifico, è stato evidenziato l'utilizzo da parte del sistema di dati provenienti da BRAVO (*Bureau Risk and Verification Observation*), un *software* afflitto dai problemi di trasparenza citati nelle pagine precedenti²¹⁴, e sono state riscontrate delle marcate differenze nei punteggi assegnati in base all'etnia: il 53% degli imputati di colore è stato valutato come ad

²¹¹ H.R. 5682 First Step Act, 2018

²¹² First Step Act, 2018, Sec. 101, §3632

²¹³ The First Step Act of 2018: Risk and Needs Assessment System Update, 2020, 1

²¹⁴ Comment Letter to the Department of Justice on PATTERN First Step Act, 2019

alto rischio, contro il 29% degli imputati bianchi; specularmente, il 7% degli imputati di colore è stato classificato a basso rischio, contro il 29% degli imputati bianchi. Nonostante tali perplessità, il sistema si è dimostrato il 15% più preciso degli altri *tool*, diventando uno dei più efficaci nel predire il rischio di recidiva²¹⁵.

Quanto ai dubbi di natura costituzionale, le critiche si concentrano sul significato da attribuire all'espressione "giusto processo" in relazione al 5° e al 14° emendamento della Costituzione. Questi ultimi, infatti, costituiscono la cd. *due process clause* che prevede, tra i tanti, due principi cardine: la *accuracy* e la *transparency*. La prima consiste nel diritto di ogni imputato a vedersi giudicato in base a fatti precisi e puntuali e, di conseguenza, di poter impugnare la decisione che non rispettasse tale parametro. È evidente come il punto di partenza per poter applicare questa regola sia la capacità della difesa di comprendere l'erroneità dei fatti alla base della decisione, che nel caso di un sistema esperto saranno gli *input* forniti alla macchina: se questo manca, secondo l'ordinamento statunitense, non ci si trova davanti a un giusto processo. Lo stesso si può dire anche con riferimento alla *transparency*, ossia il diritto di conoscere i ragionamenti alla base degli *output* del programma (o delle sentenze del giudice)²¹⁶. Se, come accade con i *tool* privati, questo non è possibile fino in fondo, saltano entrambe le garanzie costituzionali che stanno alla base del sistema processuale statunitense, ma gli stessi concetti sono riscontrabili anche nell'ordinamento processuale italiano all'articolo 111 Cost.

3.4.1 Il giusto processo nel diritto costituzionale interno

Dopo aver visto le ripercussioni dei programmi predittivi nell'ordinamento statunitense e aver capito che uno dei temi più delicati è il loro rapporto con i principi del giusto processo, è utile dare uno sguardo più da vicino al contenuto di tale diritto

²¹⁵ D. Zingales, 2021, 11; Tale è la stima riportata nel The First Step Act of 2018: Risk and Needs Assessment System, 2019, p. 57, fig.3; Si veda anche C. D. Stimson, 2020, 14, che riporta i risultati di studi secondo i quali PATTERN ha registrato una capacità predittiva pari all'80% per gli uomini e al 70% per le donne, a fronte del 67% di quella di COMPAS.

²¹⁶ D. Zingales, 2021, 14; Nella dottrina penalistica italiana si vedano F. Basile, 2019; M. Caterini, 2020; M. Gialuz, 2019; S. Quattrocchio, 2019.

nell'ordinamento italiano, cercando di mettere a fuoco gli aspetti più problematici che un uso massiccio di sistemi come COMPAS porterebbe sul piano delle tutele costituzionali.

Con la legge costituzionale n.2 del 23 novembre 1999, l'articolo 111 della Costituzione fu arricchito di cinque nuovi commi recanti le regole che oggi sono note come principi del giusto processo, applicabili in tutti i processi in cui si ravvisi l'esercizio del potere giurisdizionale. Il primo comma, che recita "La giurisdizione si attua mediante il giusto processo regolato dalla legge", oltre a sancire una riserva di legge in materia, pone diversi dubbi circa il significato dell'espressione "giusto processo": alla sua forte portata evocativa si accompagna una certa vaghezza. La disposizione sembra far riferimento a un ideale di giustizia che preesiste rispetto alla legge e collegato ai diritti inviolabili di tutte le persone coinvolte nel processo. Se ciò fosse corretto, tale clausola renderebbe aperto e non tassativo l'elenco delle garanzie stabilite nei commi successivi²¹⁷.

Un diritto che entra in gioco quando si pensa al rapporto di questa norma con i *risk assessment tool* è quello al contraddittorio. Al secondo comma, infatti, l'articolo 111 sancisce il fatto che l'imputato debba essere messo nelle condizioni di difendersi in modo pieno ed effettivo esponendo le proprie ragioni prima che venga emanato un provvedimento a suo carico. Tale garanzia può essere letta con l'accezione di "parità delle parti" sia in campo civile che penale, pur avendo delle differenze²¹⁸. Nel processo civile è più facile dare ad attore e convenuto gli stessi strumenti di difesa, mentre nel processo penale bisogna fare i conti con le inevitabili differenze che intercorrono tra PM e imputato: ecco perché, in questo caso, non si parla di parità delle parti *tout court*, ma di parità delle parti mitigata dal principio di adeguatezza. Il PM è senza dubbio una parte del processo, ma essendo una parte pubblica ha delle prerogative che non trovano corrispondenza nei diritti dell'imputato: la possibilità di avvalersi della Polizia giudiziaria, di avere dei poteri autoritativi come la possibilità di limitare (sia pur momentaneamente) la libertà personale, di disporre l'accompagnamento coattivo, le intercettazioni, il sequestro, ecc.

²¹⁷ C. Conti, 2022; Sul tema del giusto processo si veda anche G. Ubertis, 2008

²¹⁸ C. Conti, 2022

Detto questo, è evidente che il diritto al contraddittorio può dirsi pienamente rispettato se, a monte, l'imputato sa con precisione quali sono i fatti che l'autorità gli contesta e quali sono gli elementi che hanno portato a tale contestazione: ci si può domandare, anche in questo caso, come ciò possa avvenire se uno degli elementi su cui si basa l'accusa è un *report* di una macchina non pienamente spiegabile nemmeno dagli stessi programmatori che hanno costruito il sistema. Anche la tutela posta dall'articolo 6 comma 3 della CEDU può entrare in conflitto con l'utilizzo dei *risk assessment tool* nella parte in cui stabilisce che "Ogni accusato ha diritto soprattutto a essere informato [...] della natura e dei motivi dell'accusa elevata a suo carico" perché non è sempre agevole capire nel dettaglio perché il programma decide in un modo piuttosto che un altro, nonostante i ragionamenti dei sistemi esperti siano, almeno sulla carta, controllabili.

CAPITOLO QUARTO

LO STATO DELL'IA NEL DIRITTO IN ITALIA E NELL'UNIONE EUROPEA

SOMMARIO: 4.1 La normativa sull'IA al di là della responsabilità per danni - 4.1.1 L'esperienza della Cina e degli Stati Uniti - 4.1.2 Il percorso dell'Unione europea - 4.1.2.1 La proposta di Regolamento sull'IA della Commissione europea - 4.1.2.2 La Carta etica della CEPEJ - 4.2 Esperienze moderne di uso dell'IA nel diritto in Italia - 4.2.1 Cenni su Key Crime e XLAW - 4.2.2 Come sta cambiando il diritto penale

4.1 La normativa sull'IA al di là della responsabilità per danni

Questo contributo ha voluto mettere in luce alcuni dei temi più rilevanti in materia di giustizia predittiva e intelligenza artificiale nel diritto, il cui naturale epilogo è un approfondimento sulle normative attualmente in vigore o in discussione a livello interno e sovranazionale, non dimenticando di dare uno sguardo a ciò che sta accadendo al di fuori del continente europeo. Si vedranno, inoltre, alcuni dei sistemi esperti più importanti in uso in Italia.

Nelle pagine precedenti si è già parlato di leggi, direttive e disposizioni che a vario titolo regolano o vorrebbero regolare uno o più settori dell'informatica giuridica, tra cui la Raccomandazione del Parlamento europeo sulla soggettività dei robot e sulla responsabilità per i danni da loro causati, ma il lavoro ricognitivo di questo capitolo sarà su più larga scala e consentirà di individuare i punti focali di ciascun approccio regolatorio. Le differenze sono spesso dovute alle differenze storiche, geopolitiche, giuridiche e culturali di ciascun paese o alla relativa novità di un fenomeno che fino a qualche decennio fa sembrava di dominio esclusivo degli scrittori di fantascienza come Isaac Asimov che, con le “Tre leggi della robotica”, rappresenta ancora oggi una stella polare per gli esperti:

1. Un robot non può recar danno a un essere umano né può permettere che, a causa del suo mancato intervento, un essere umano riceva danno.
2. Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non vadano in contrasto alla Prima Legge.
3. Un robot deve proteggere la propria esistenza, purché la salvaguardia di essa non contrasti con la Prima o con la Seconda Legge.²¹⁹

4.1.1 L'esperienza della Cina e degli Stati Uniti

Come si è accennato nel secondo capitolo, la Cina è uno dei paesi più attivi nel campo dell'intelligenza artificiale, avendo predisposto, nel 2017, un piano di sviluppo con il quale

²¹⁹ I. Asimov, 1942

mira diventare entro il 2030 il paese più avanzato nel settore: il *New Generation Artificial Intelligence Plan* (AIDP). Nello stesso anno si è dotata anche di una legge sulla cybersicurezza, la *Cybersecurity Law*, mentre nel 2021 ha pubblicato il Libro Bianco del *National Industrial Information Security Development Center* e la *Personal Information Protection Law*, contenenti numerosi principi e norme in materia di intelligenza artificiale.

Come si legge nell'AIDP, il Governo vuole ottenere risultati avanguardistici già entro il 2025 e, entro il 2030, vuole che "*China's AI theories, technologies, and applications should achieve world-leading levels, making China the world's primary AI innovation center, achieving visible results in intelligent economy and intelligent society applications, and laying an important foundation for becoming a leading innovation-style nation and an economic power*"²²⁰. Il progetto, sicuramente molto ambizioso, è gestito dal Comitato di Consulenza Strategica, un ente coordinato dal Ministero della Scienza e della Tecnologia, e pone al centro della propria strategia il settore privato e i governi locali. A tal fine, sono state identificate alcune società da sostenere economicamente, tramite sovvenzioni pubbliche, per migliorare lo sviluppo tecnologico nei rispettivi settori: Baidu nella guida autonoma, Alibaba nelle *smart cities*, Tencent in campo medico, ecc²²¹.

Per quanto riguarda l'ampio tema della *governance* sociale, si possono ricordare alcuni interventi molto discussi negli ultimi anni che permettono di inquadrare l'approccio che l'esecutivo intende seguire. Nel 2015 la Corte Suprema, nel richiedere un attento esame dei precedenti giurisprudenziali, ha raccomandato il ricorso agli strumenti intelligenti in base al principio "giudizi simili in casi simili": ciò avrebbe portato alla realizzazione di una *common law sui generis* guidata dagli algoritmi, in cui sono questi ultimi a indicare quale soluzione applicare al caso concreto. Ancora più recente è il cosiddetto *Social Credit System*, il sistema di crediti sociali che analizza i comportamenti delle persone attribuendo o togliendo punti in base a un metro di giudizio stabilito dalle autorità basato sulle tecnologie di riconoscimento facciale o vocale e sulle moderne tecniche predittive. Il Piano, inoltre, propone di sviluppare l'IA per migliorare di sistemi di apprendimento, sanare le disparità

²²⁰ G. Webster-R. Creemers-P. Triolo-E. Kania, 2017, 6

²²¹ A. Malaschini, 2022, 109

educative tra le zone rurali e le zone urbane del paese, sviluppare nuove tecniche di sorveglianza, di contrasto al terrorismo e raggiungere quel “minimo standard morale” della società cui ha fatto riferimento Xi Jinping nella sua relazione al 19° Congresso del Partito nel 2017²²².

Nonostante venga detto che l’intelligenza artificiale debba servire al benessere dell’umanità (facendo anche riferimento alla tutela dei diritti umani)²²³, appare evidente che, in uno Stato autoritario come la Cina, tale principio etico sia facilmente ridimensionabile citando, per esempio, l’articolo 51 della Costituzione: “L’esercizio da parte dei cittadini della Repubblica Popolare Cinese delle loro libertà e dei loro diritti non può prevalere sugli interessi dello Stato, della società e della collettività, o sulle legittime libertà e diritti degli altri cittadini”. Il riferimento ai diritti umani, quindi, va valutato alla luce dei presupposti etici che caratterizzano la società cinese, che sono diversi da quelli occidentali: gli interessi statali avranno sempre la precedenza rispetto ai diritti individuali. La stessa adesione al piano predisposto dall’AIDP, pur non essendo formalmente obbligatoria e priva di valore normativo, è fortemente incoraggiata (fino a diventare, di fatto, inevitabile) perché potrà essere collegata, per esempio, al rispetto di determinati standard di certificazione la cui mancata osservanza potrà determinare sanzioni come limiti all’accesso al mercato o multe²²⁴. Il risultato è la creazione di un enorme *database* di dati personali in mano a un governo che può disporne più o meno liberamente.

Anche gli Stati Uniti, come la Cina, non hanno ancora una normativa organica e completa sull’intelligenza artificiale, tuttavia, a partire dal 2016, sono stati pubblicati numerosi documenti e atti di provenienza prevalentemente governativa volti a indirizzare le future leggi sul tema. L’attenzione è aumentata a tal punto che nel 2019 sono state presentate al 116° Congresso statunitense 51 proposte di legge contenenti il termine “intelligenza artificiale”, mentre nel 2015-2016 erano solo due²²⁵. A differenza dell’approccio assertivo e concentrato sul controllo dei dati e delle tecnologie del governo cinese, negli Stati Uniti il

²²² A. Malaschini, 2022, 110

²²³ R. Arcesati, 2021, 4

²²⁴ A. Malaschini, 2022, 112; H. Yang, 2019

²²⁵ Y. Chae, 2020, 17

focus, almeno inizialmente, era indirizzato al settore militare e industriale, nel tentativo di cementificare il proprio primato nel settore.

Dopo l'emanazione dei primi documenti e l'istituzione delle prime commissioni sotto le presidenze Obama e Trump, l'11 febbraio 2019 viene pubblicato l'Ordine Esecutivo 13859 intitolato "*Maintaining American Leadership in Artificial Intelligence*" in cui viene indicato, tra le altre cose, che gli Stati Uniti devono porsi all'avanguardia dell'innovazione tecnologica, aumentare la fiducia dei cittadini nell'IA, proteggere le libertà civili, la *privacy* e i valori americani²²⁶, ma il documento più recente su cui soffermarsi è quello contenente le conclusioni della *National Security Commission on Artificial Intelligence* (NSCAI) del marzo del 2021. Questa commissione, istituita con la legge di bilancio della Difesa per il 2019, aveva il compito di fornire alla presidenza e al Congresso delle raccomandazioni "*to advance the development of artificial intelligence, machine learning, and associated technologies by the United States to comprehensively address the national security and defense needs of the United States*"²²⁷. È formata da 15 membri con grande esperienza nel settore: il presidente è stato CEO di Google, mentre tra i commissari si possono trovare il CEO di Oracle, di In-Q-Tech, di Amazon Web Services, il capo della divisione IA di Google, presidi e docenti di facoltà scientifiche come il California Institute of Technology, il *Center of Security and Emergency Technologies della Georgetown University* e la *National Academy of Sciences*, oltre a parlamentari esperti del settore tecnologico tra cui il presidente *pro tempore* della *Federal Communication Commission*²²⁸.

Le conclusioni sono articolate in due sezioni: *Defending America in the AI era* e *Winning the technology competition*. Nella prima si parla delle minacce emergenti, dei problemi della Difesa, del ruolo dell'IA negli scenari di guerra, dei rischi dei sistemi d'arma autonomi, della relazione tra Intelligence e IA e del rapporto uomo-macchina nell'utilizzo di strumenti potenzialmente capaci di assumere decisioni autonome. La seconda si concentra sulla sfida tecnologica con la Cina e il resto del mondo, sulla gestione dei finanziamenti per la ricerca sull'IA, sulle misure da adottare in materia di tutela della proprietà intellettuale. Si ribadisce,

²²⁶ E.O. 13859, 2019, 3967

²²⁷ H. R. 5515, 2018, sec. 1051

²²⁸ A. Malaschini, 2022, 123

inoltre, l'importanza di sviluppare una iniziativa coordinata con i Paesi alleati e si affronta il tema delle tecnologie legate all'IA: microelettronica, biotecnologie, quantum computing, 5G, produzione additiva, sistemi autonomi, robotica e nuovi modelli energetici²²⁹.

Ciò che stupisce del report sono le parole del presidente e del vicepresidente della NSCAI: *"America is not prepared to defend or compete in the AI era. This is the tough reality we must face"*²³⁰. Il tenore complessivo delle analisi è schietto e diretto, lontano dai toni trionfali che si notano leggendo i documenti cinesi. Gli Stati Uniti, quindi, si preparano ad affrontare l'IA con un atteggiamento competitivo avendo già dichiaratamente identificato l'avversario principale: la Cina.

4.1.2 Il percorso dell'Unione europea

Il dibattito in seno alle istituzioni europee in materia di intelligenza artificiale è sfociato nella proposta di Regolamento sull'AI della Commissione europea dell'aprile del 2021, tuttavia, in via preliminare, è utile osservare il modo in cui si è arrivati a tale risultato dando conto dei documenti e degli atti più importanti che si sono avvicinati negli anni precedenti.

Nel giugno del 2018 venne istituito l'*High Level Expert Group on Artificial Intelligence* (o Gruppo di Esperti ad Alto Livello sull'Intelligenza Artificiale), un ente designato dalla Commissione formato da 52 esperti di estrazione accademica, civile e industriale con l'obiettivo di coordinare lo sviluppo di una strategia europea per l'IA. Uno dei suoi apporti più importanti sono senza dubbio gli "Orientamenti etici per un'IA affidabile" presentati nell'aprile del 2019 e contenenti dei principi che fungono da linee guida ancora oggi. Nello specifico, viene stabilito che un'IA affidabile si deve basare su tre componenti da mantenere durante l'intero ciclo di vita del sistema:

- Legalità, ossia l'IA deve "ottemperare a tutte le leggi e ai regolamenti applicabili";
- Eticità, nel senso che "l'IA deve assicurare l'adesione a principi e valori etici";

²²⁹ A. Malaschini, 2022, 124

²³⁰ NSCAI, Final Report, 2021, 1

- Robustezza “dal punto di vista tecnico e sociale poiché, anche con le migliori intenzioni, i sistemi di IA possono causare danni non intenzionali”²³¹.

Il documento, nel concentrarsi sui profili dell’eticità e della robustezza, individua quattro principi cardine:

- Rispetto dell’autonomia umana;
- Prevenzione dei danni;
- Equità;
- Esplicabilità.

In considerazione di ciò, il Gruppo ha individuato sette indicazioni chiave che i sistemi di IA dovrebbero soddisfare:

- Intervento e sorveglianza umana;
- Robustezza tecnica e sicurezza;
- Riservatezza e governance dei dati;
- Trasparenza;
- Diversità;
- Non discriminazione ed equità;
- Benessere sociale e ambientale;
- *Accountability*.

Ciò che viene auspicato è la creazione di macchine antropocentriche da mettere al servizio dell’umanità e del bene comune, massimizzandone i benefici e prevenendone o minimizzandone i rischi. L’ottica europea, quindi, è focalizzata sugli aspetti etici e strutturali dell’IA “poiché gli esseri umani e le comunità riusciranno ad avere fiducia nello sviluppo della tecnologia e nelle sue applicazioni solo quando esisterà un quadro di riferimento chiaro e completo per conseguire l’affidabilità”²³².

²³¹ AI HLEG, 2019, 2

²³² AI HLEG, 2019, 5

4.1.2.1 La Proposta di Regolamento sull'IA della Commissione europea

Su questo solco si colloca il Libro Bianco sull'intelligenza artificiale pubblicato dalla Commissione europea il 19 febbraio 2020 contenente numerose indicazioni usate per redigere la proposta di Regolamento che la stessa Commissione ha presentato al Parlamento europeo il 21 aprile 2021, in piena adesione ai principi stabiliti nel 2019: tutela della privacy e della dignità umana, affidabilità e trasparenza, tracciabilità e controllo umano, libertà di espressione, rispetto della dignità, del genere, dell'etnia, ecc²³³.

Quanto alla proposta di Regolamento, si tratta di un tentativo piuttosto audace di delineare una normativa armonizzata sull'intelligenza artificiale a livello europeo che prende le mosse dal GDPR (Regolamento 2016/679) del 2016²³⁴. Nel Titolo Secondo, all'articolo 5, vengono elencate una serie di pratiche vietate volte a creare una cornice entro la quale le nuove tecnologie possono o meno essere sviluppate. È proibita la diffusione e l'uso di: pratiche che utilizzino tecniche subliminali per distorcere il comportamento delle persone in modo da poter provocare a sé stessi o ad altri dei danni fisici o psicologici, i cd. *dark pattern*, come l'uso di suoni non percepibili idonei a influenzare i comportamenti umani; sistemi che sfruttino le vulnerabilità di uno specifico gruppo di persone dovute all'età o alla disabilità fisica o mentale, al fine di distorcerne il comportamento in modo da provocare a sé stessi o ad altri un danno fisico o psicologico (come, per esempio, dei giocattoli interattivi che spingano i bambini a comportamenti dannosi); tecnologie usate per un periodo di tempo determinato dalle autorità al fine di valutare e giudicare un soggetto in base alle sue caratteristiche sociali o personali con l'obiettivo di stabilire se applicare dei trattamenti favorevoli o sfavorevoli in un contesto diverso da quello in cui i dati sono stati raccolti o dei trattamenti che siano ingiustificati o sproporzionati rispetto al comportamento sociale o alla sua gravità (il cd. *social scoring* applicato in Cina); pratiche biometriche di identificazione a fini di polizia, ossia il cd. *facial recognition* già in uso in Cina, a meno che non si rendano necessari per svolgere delle ricerche mirate di vittime di un crimine, per prevenire attentati o attacchi terroristici, per scoprire, localizzare, identificare o fermare un soggetto accusato

²³³ A. Malaschini, 2022, 129

²³⁴ F. Donati, 2021, 457

di gravi crimini, dovendo comunque valutare preventivamente la natura della situazione da affrontare, le conseguenze che le misure stesse possono avere sui diritti e sulle libertà di tutti gli interessati e la loro proporzionalità alle necessità da soddisfare (previa autorizzazione dell'autorità giudiziaria o di un'autorità amministrativa indipendente dello Stato membro)²³⁵.

Le politiche che la Commissione europea intende mettere in atto sono, quindi, circondate da una serie di garanzie che, almeno in questa fase, le contraddistinguono rispetto a quanto visto negli Stati Uniti e in Cina, dove l'attenzione sembra spostata in prevalenza sugli aspetti operativi dell'IA e non sulle questioni squisitamente etiche (che comunque non mancano). Un altro passaggio importante della proposta è il Titolo Terzo, intitolato "Sistemi ad alto rischio", dedicato alle macchine intelligenti per le quali non è pensabile l'applicazione delle norme già esistenti a livello europeo. Secondo la proposta, i sistemi a cui si dovrebbero applicare tali norme più stringenti sono quelli che possono creare un rischio elevato per la salute, la sicurezza e i diritti delle persone e che vedono il concorso di due condizioni²³⁶:

- Il sistema "è destinato a essere utilizzato come componente di sicurezza di un prodotto, o è esso stesso un prodotto, disciplinato dalla normativa di armonizzazione dell'Unione elencata nell'allegato II";
- Il prodotto, "il cui componente di sicurezza è il sistema di IA, o il sistema di IA stesso in quanto prodotto è soggetto a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio di tale prodotto ai sensi della normativa di armonizzazione dell'Unione elencata nell'allegato II"²³⁷.

Nell'allegato II richiamato dalla disposizione si fa riferimento a una serie di normative dedicate a macchinari, strumenti medici, giocattoli, mezzi personali di protezione, automobili, ascensori, aerei, ecc²³⁸, mentre nell'allegato III, a cui fa riferimento il comma successivo, si elencano una serie di sistemi considerati sempre ad alto rischio, indipendentemente dalle condizioni citate, poiché capaci di influenzare l'esercizio dei diritti

²³⁵ COM(2021) 206 final, 47

²³⁶ A. Malaschini, 2022, 137; Per un approfondimento puntuale si veda anche COM(2021) 206 final, 13

²³⁷ COM(2021) 206 final, 49

²³⁸ COM(2021) 206 final ANNEXES 1 to 9, 2

fondamentali. Tra questi si possono ricordare: quelli che fanno uso di tecnologie di identificazione biometrica e profilazione, quelli che gestiscono infrastrutture critiche (acqua, gas, elettricità, controllo del traffico), quelli dedicati all'ordine pubblico (dove possono essere usate anche tecnologie predittive) e quelli applicati per l'amministrazione della giustizia, proprio come i *risk assesment tool* di cui si è parlato nel capitolo precedente²³⁹. È, inoltre, importante sottolineare che tale lista è aggiornabile dalla Commissione quando necessario, il che le permette di adattarsi con l'avanzare degli sviluppi tecnologici²⁴⁰.

Negli articoli successivi viene articolata una complessa normativa volta a regolare tali macchine: in questa sede è bene darne conto, seppur per sommi capi. Innanzitutto, ex art. 9, devono essere dotate di modelli di gestione dei rischi che permettano "l'identificazione e l'analisi dei rischi noti e prevedibili associati a ciascun sistema di IA ad alto rischio" e che rendano possibile prevedere eventuali criticità future analizzando i dati raccolti dal sistema di monitoraggio stesso. I produttori, quindi, devono mirare all'eliminazione o alla minimizzazione dei rischi attraverso processi costanti di stima e valutazione.

Deve, inoltre, essere prodotta e mantenuta per tutto il ciclo di vita del sistema una documentazione tecnica idonea a dimostrare il rispetto delle norme e, in aggiunta, i sistemi devono essere disegnati e sviluppati in modo da "garantire che il loro funzionamento sia sufficientemente trasparente da consentire agli utenti di interpretare l'output del sistema e utilizzarlo adeguatamente"²⁴¹ e "in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui il sistema di IA è in uso" al fine di prevenire o contenere rischi alla salute, alla sicurezza o ai diritti fondamentali²⁴². A tal proposito, la persona fisica deve: essere capace di comprenderne le capacità e i limiti per controllarne il funzionamento, essere consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema intelligente, essere in grado di interpretare correttamente i suoi risultati e di decidere, in qualsiasi

²³⁹ COM(2021) 206 final ANNEXES 1 to 9, 4

²⁴⁰ COM(2021) 206 final, 49

²⁴¹ COM(2021) 206 final, 54

²⁴² COM(2021) 206 final, 55

situazione, di non usare il sistema o altrimenti di ignorare, annullare o ribaltare i suoi output. Deve, infine, essere in grado di intervenire sul suo funzionamento o di interromperlo²⁴³.

Sempre secondo la proposta, ai fini della *accountability*, i sistemi devono essere registrati in un *database* dell'Unione europea, mentre con riferimento alla trasparenza, è necessario che gli utenti sappiano se si stanno interfacciando con uno strumento intelligente (per esempio con un *chatbot*) per poter decidere se proseguire o meno l'interazione e, nel caso in cui vengano generati o manipolati dei contenuti che somiglino a persone o a cose esistenti, deve essere chiaro che si tratta di creazioni artificiali²⁴⁴.

Le macchine a basso rischio sono, al contrario, tutte quelle non considerate come proibite o a rischio elevato e, per questo motivo, non sono sottoposte a una disciplina diversa da quella già esistente. Si possono trovare facilmente in Europa da diversi anni (si passa dai videogiochi ai filtri antispam) e la Commissione europea auspica la creazione di codici di condotta ad adesione volontaria con cui anche tali sistemi si possano adeguare alle norme previste per quelli ad alto rischio al fine di aumentare la fiducia dei cittadini verso l'IA e, di conseguenza, aumentarne la diffusione sul mercato²⁴⁵.

Per finire, prima di analizzare le critiche rivolte alla proposta, è bene capire quali sono le norme formulate in materia di *governance*, monitoraggio e sanzioni. L'articolo 56 istituisce il Consiglio europeo per l'intelligenza artificiale composto dai responsabili delle diverse autorità nazionali di vigilanza, dal Garante europeo della protezione dei dati e presieduto dalla Commissione stessa per "contribuire all'efficace cooperazione delle autorità nazionali di controllo e della Commissione per quanto riguarda le materie disciplinate dal presente regolamento", "coordinare e contribuire agli orientamenti e all'analisi [...] sulle questioni emergenti nel mercato interno in relazione alle materie disciplinate dal presente regolamento" e "assistere le autorità nazionali di controllo e la Commissione nel garantire l'applicazione uniforme del presente regolamento". Uno dei suoi obiettivi, quindi, è incentivare l'interpretazione uniforme delle disposizioni proposte, cercando di limitare le

²⁴³ COM(2021) 206 final, 55

²⁴⁴ A. Malaschini, 2022, 139; K. C. Halm-K. Sheriff-J. Mark-K. Berry, 2021

²⁴⁵ A. Malaschini, 2022, 140

differenze e le zone d'ombra che si possono formare all'interno dell'UE. Un altro dettaglio che emerge dall'articolo in esame è l'istituzione, in ogni Stato membro, di "autorità nazionali competenti al fine di garantire l'applicazione e l'attuazione del presente regolamento" e di "un'autorità nazionale di controllo tra le autorità nazionali competenti". Quest'ultima "agisce in qualità di autorità di notifica e di autorità di vigilanza del mercato", contribuendo a formare, in questo modo, una struttura (forse eccessivamente stratificata) di enti nazionali e sovranazionali interconnessi e capaci di intervenire qualora i sistemi generino rischi inattesi.

Passando alle sanzioni, l'articolo 71 prevede che la violazione di quanto previsto dagli articoli 5 e 10, dedicati rispettivamente alle pratiche proibite e alla gestione dei dati, porti a una sanzione amministrativa pecuniaria fino a 30 milioni di euro o, se l'autore è una società, fino al 6% del fatturato mondiale totale annuo. Al contrario, se la violazione riguarda delle disposizioni diverse da quelle di cui agli articoli 5 e 10, la sanzione amministrativa pecuniaria può arrivare fino a 20 milioni di euro o, se l'autore è una società, fino al 4% del fatturato mondiale totale annuo. Il dato più rilevante è la mancanza di un sistema di denuncia da parte dei singoli cittadini, relegando tale compito alle sole autorità pubbliche e generando, così, una lacuna potenzialmente capace di compromettere l'effettivo rispetto della normativa proposta.

Nei mesi successivi alla pubblicazione del documento da parte della Commissione europea, sono emerse numerose critiche da parte di esperti e addetti ai lavori. Una parte di queste è rivolta al linguaggio usato, ritenuto da alcuni poco preciso e ambiguo, che renderebbe difficile capire quali sono i criteri con cui applicare le sanzioni²⁴⁶; altre si concentrano sulle ricadute economiche che comporterebbe la sua eventuale entrata in vigore. I costi per l'intera economia europea determinati dagli adeguamenti richiesti, infatti, potrebbero ammontare, nei prossimi 5 anni, a 31 miliardi di euro, rallentando così il processo di informatizzazione e favorendo l'esodo di ricercatori e imprenditori verso paesi con minori limiti normativi²⁴⁷. Sul lato tecnico, alcuni sostengono che sia difficile identificare con

²⁴⁶ G. Olvi-C. Bocchi, 2021

²⁴⁷ A. Malaschini, 2022, 144

precisione quale sia il *provider* ai fini dell'attribuzione della responsabilità, dal momento che i sistemi di IA sono il frutto di un assemblaggio di strumenti diversi²⁴⁸; inoltre, più in generale, le nuove norme in materia di trasparenza potrebbero pregiudicare la tutela della proprietà intellettuale e dei segreti commerciali²⁴⁹.

Per concludere, uno dei punti più interessanti è quello relativo alla “spiegabilità”, un problema già visto nel capitolo precedente con riferimento all'opacità delle determinazioni di COMPAS: spesso nemmeno gli sviluppatori sanno perché la loro IA ha compiuto un'azione piuttosto che un'altra, quindi stabilire per regolamento che deve essere possibile capire esattamente i motivi dietro un determinato *output* è, ad oggi, utopistico²⁵⁰.

4.1.2.2 La Carta etica del CEPEJ

Anche il Consiglio d'Europa, la cd. grande Europa, ha mosso i primi passi verso la definizione di una cornice valoriale e normativa in materia di intelligenza artificiale che, a differenza di quanto visto nel paragrafo precedente, si focalizza esclusivamente sui sistemi usati per l'amministrazione della giustizia.

Il 3-4 dicembre 2018 la Commissione Europea per l'Efficienza della Giustizia (CEPEJ) ha adottato la “Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi”, che per ragioni di brevità d'ora in poi verrà chiamata “Carta etica”, contenente cinque principi di cui devono tenere conto tanto il legislatore nella fase di elaborazione normativa, quanto i giudici durante i processi. Tale atto si colloca al fianco della “Convenzione per la salvaguardia dei Diritti dell'Uomo e delle Libertà fondamentali” e alla “Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale”: insieme, pur regolando ambiti diversi non per forza riguardanti la tecnologia, tratteggiano una serie di regole che coinvolgono sia l'attività legislativa che quella giudiziaria in materia di intelligenza artificiale nei processi. Ogni avanzamento

²⁴⁸ A. Malaschini, 2022, 144; Sullo stesso tema si veda anche J. Buyers *et al.*, 2021

²⁴⁹ J. Buyers *et al.*, 2021

²⁵⁰ G. Olvi-C. Bocchi, 2021

tecnologico deve tenere in considerazione non solo l'articolo 111 Cost. ma anche, per esempio, l'articolo 6 CEDU: indipendentemente dall'uso di sistemi intelligenti, non deve essere pregiudicato il diritto ad essere giudicato da un giudice terzo, indipendente e imparziale e di avere un giusto processo²⁵¹.

I cinque principi stabiliti dalla Carta etica sono:

- Principio del rispetto dei diritti fondamentali: assicurare che l'elaborazione e l'attuazione di strumenti e servizi di intelligenza artificiale siano compatibili con i diritti fondamentali;
- Principio di non-discriminazione: prevenire specificamente lo sviluppo o l'intensificazione di qualsiasi discriminazione tra persone o gruppi di persone;
- Principio di qualità e sicurezza: in ordine al trattamento di decisioni e dati giudiziari, utilizzare fonti certificate e dati intangibili con modelli elaborati multidisciplinarmente, in un ambiente tecnologico sicuro;
- Principio di trasparenza, imparzialità ed equità: Rendere le metodologie di trattamento dei dati accessibili e comprensibili, autorizzare verifiche esterne;
- Principio del "controllo da parte dell'utilizzatore": Precludere un approccio prescrittivo e assicurare che gli utilizzatori siano attori informati e abbiano il controllo delle loro scelte²⁵².

Il primo dato da evidenziare è che la proposta della Commissione esaminata in precedenza si colloca in tendenziale continuità con questi principi, ma la loro analisi può portare a galla altri spunti interessanti. Innanzitutto, le macchine devono essere programmate a monte per rispettare i diritti umani in materia processuale e non devono discriminare i soggetti che ne subiscono le decisioni, dovendo prestare attenzione alla gestione dei dati sensibili come l'origine etnica, le condizioni socioeconomiche, le opinioni politiche, la fede religiosa o filosofica, l'appartenenza a un sindacato, i dati genetici, i dati biometrici, i dati sanitari o i dati relativi alla vita sessuale o all'orientamento sessuale. Se si dovesse prospettare il rischio della violazione di tali diritti, come è avvenuto nel caso Loomis, spetta ai programmatori

²⁵¹ C. Cavaceppi, 2020, 102

²⁵² CEPEJ, 2018, 6

correggere il problema nel modo che ritengono più idoneo al fine di limitare il rischio di una sua reiterazione o, se possibile, neutralizzarlo. Non solo: il CEPEJ stabilisce che “Dovrebbe [...] essere incoraggiato l’utilizzo dell’apprendimento automatico e delle analisi scientifiche multidisciplinari, al fine di contrastare tali discriminazioni”²⁵³, quindi individua egli stesso il modo che ritiene più efficace per evitare problemi discriminatori, tuttavia, come è stato più volte ricordato in questo contributo, sono proprio le macchine ad apprendimento automatico a destare maggiori perplessità a causa dell’opacità del loro funzionamento anche agli occhi dei loro stessi creatori, il che può inficiare anche la portata del quarto principio. Quest’ultimo, infatti, prevede che i sistemi di elaborazione dati debbano essere trasparenti, imparziali, equi e potenzialmente soggetti a un controllo esterno che, nelle macchine ad apprendimento automatico, risulterà inevitabilmente più complesso.

Il quarto principio, tuttavia, ha un’ulteriore ricaduta sul versante della proprietà intellettuale e dei brevetti in quanto, secondo il CEPEJ, “deve essere raggiunto un equilibrio tra la proprietà intellettuale di alcune metodologie di trattamento e l’esigenza di trasparenza (accesso al processo creativo), imparzialità (assenza di pregiudizi), equità e integrità intellettuale (privilegiare gli interessi della giustizia) quando si utilizzano strumenti che possono avere conseguenze giuridiche, o che possono incidere significativamente sulla vita delle persone”²⁵⁴. Di qui l’auspicio di ottenere la completa trasparenza dei codici sorgente e della documentazione sulla quale si fondano, oltre che una descrizione chiara e completa del loro funzionamento e la creazione di Autorità indipendenti incaricate di verificare e certificare le metodologie di trattamento nella misura in cui assicurino effettivamente il rispetto di tali principi²⁵⁵. Si tratta di questioni delicate e oggetto di dibattito, specie alla luce della sentenza Loomis, dove uno dei punti sollevati dalla difesa era proprio l’impossibilità di analizzare i ragionamenti e il codice sorgente di COMPAS.

Anche il quinto e ultimo principio, che parla del cd. *under user control*, merita un breve approfondimento. Dal dato letterale si comprende che si tratta di una disposizione rivolta ai professionisti del sistema giudiziario, ossia avvocati e magistrati, che “dovrebbero essere in

²⁵³ CEPEJ, 2018, 8

²⁵⁴ CEPEJ, 2018, 11

²⁵⁵ C. Cavaceppi, 2020, 103

grado, in qualsiasi momento, di rivedere le decisioni giudiziarie e i dati utilizzati per produrre un risultato e continuare ad avere la possibilità di non essere necessariamente vincolati a esso alla luce delle caratteristiche specifiche di tale caso concreto”²⁵⁶. In sostanza, ciò che auspica il CEPEJ è un uso consapevole degli operatori del diritto degli strumenti intelligenti, avendo ravvisato il pericolo, citato capitolo precedente, di un eccessivo conformismo degli stessi negli *output* delle macchine. Al contrario, ciò che l’uomo dovrebbe avere è lo spirito critico e di analisi per comprendere le caratteristiche peculiari del caso che si trova di fronte e stabilire se sia il caso di assecondare o meno i risultati del sistema esperto. Il controllo della decisione investe anche i destinatari della stessa che dovrebbero poterla comprendere in tutti i suoi aspetti con un linguaggio chiaro e comprensibile e “deve avere il diritto di opporvisi, al fine di far giudicare il suo caso direttamente da un tribunale ai sensi dell’articolo 6 della CEDU”²⁵⁷.

4.2 Esperienze moderne di uso dell’IA nel diritto

A partire dal 2018, nell’ambito del Piano Coordinato sull’intelligenza artificiale della Commissione europea, anche l’Italia ha iniziato un percorso programmatico verso l’adozione di una normativa in materia che, tuttavia, ha iniziato a dare i suoi frutti solo nell’ottobre del 2021, quando fu presentato dal Governo il “Programma strategico Intelligenza Artificiale 2022-2024”. Pur essendo indicati numerosi obiettivi, priorità e aree di intervento da implementare entro il 2024, si tratta di un progetto ancora in fase embrionale, da alcuni ritenuto insufficiente e poco definito, che intende solo indicare la direzione da seguire nei prossimi anni²⁵⁸. Il documento, in generale, è concentrato sulla necessità di dare una forte spinta alla ricerca e all’adozione dell’IA in tutti i principali settori: industria e manifatturiero, sistema educativo, agroalimentare, cultura e turismo, salute e benessere, ambiente,

²⁵⁶ CEPEJ, 2018, 12

²⁵⁷ CEPEJ, 2018, 12

²⁵⁸ A. Malaschini, 2022, 157

infrastrutture e reti, banche, finanza e assicurazioni, Pubblica Amministrazione, città, aree e comunità intelligenti, sicurezza nazionale e tecnologia dell'informazione²⁵⁹.

Oltre a questo, il suo pregio maggiore è fare il punto sullo stato dell'avanzamento tecnologico in Italia, proponendo alcune politiche da adottare in tre aree di intervento: talenti e competenze, ricerca e applicazioni. Si suggerisce, per esempio, di attivare tre nuovi cicli di corsi di dottorato specificamente progettati per le esigenze delle amministrazioni pubbliche; promuovere l'integrazione, all'interno dei curricula di tutti i cicli scolastici, di attività, metodologie e contenuti volti allo sviluppo delle competenze STEM; ampliare corsi di programmazione e inclusione di corsi di IA applicata e stage in tutti i curricula ITS; promuovere la creazione di start-up tra professionisti e studenti ad alto potenziale; definire di un sistema di governance nazionale a supporto della certificazione dei prodotti di IA immessi sul mercato (con profili di rischio più elevati, in particolare per la salute, la sicurezza o i diritti) con la definizione di chiari strumenti armonizzati in linea con la proposta di Regolamento sull'intelligenza artificiale della Commissione europea già vista nei paragrafi precedenti²⁶⁰.

Se questo è ciò che si può registrare con riferimento all'aspetto programmatico/normativo in Italia, la situazione sembra più vivace guardando alle moderne sperimentazioni di sistemi esperti nel diritto. Un primo esempio da citare è Toga, un applicativo pensato per i professionisti che si occupano di diritto penale che raccoglie tutte le fattispecie di reato previste dall'ordinamento italiano (che, secondo il sito del programma, sono più di 5000) permettendo di verificare, apparentemente senza margine di errore, "la competenza, la procedibilità, l'ammissibilità ai riti alternativi, i termini prescrizionali del reato e della pena, la specie delle misure cautelari applicabili anche nelle fattispecie più complesse e la durata delle stesse nonché la pena per ciascun tipo di reato"²⁶¹: è come se si trattasse di un aiutante artificiale del giurista, uno strumento che permette di digitalizzare molti aspetti professionali della sua attività.

²⁵⁹ Programma strategico, 2021, 17

²⁶⁰ Programma strategico, 2021, 22 ss.

²⁶¹ C. Cavaceppi, 2020, 109

Un altro sistema attualmente operativo è Watson di IBM. L'Ufficio I della Direzione generale affari internazionali del Ministero della Giustizia ha avuto la necessità di dotarsi di un programma di rilevazione statistica delle estradizioni, di attuazione dei mandati di arresto europeo e delle richieste di assistenza giudiziaria. Per fare ciò, IBM e la DGSIA (Direzione generale dei sistemi informativi automatizzati) hanno realizzato un *software* basato su Watson Explorer capace di estrarre dai fascicoli aperti presso la Direzione generale affari internazionali e cooperazione giudiziaria del Ministero, elementi e dati relativi ai Paesi richiesti o richiedenti, alle fattispecie di reato per le quali si procede, agli esiti delle richieste e ai tempi medi di esaurimento delle procedure. Tutto ciò consente il calcolo e la presentazione di alcuni indicatori, come il numero fascicoli per reato e per Paese, e di fornire tempestivamente le informazioni statistiche richieste a livello nazionale e internazionale²⁶².

4.2.1 Cenni su Key Crime e XLAW

I prossimi due programmi si discostano da quelli appena visti poiché si focalizzano sull'attività di prevenzione dei reati attraverso tecnologie predittive: hanno, infatti, l'obiettivo di calcolare la probabilità di consumazione dei reati stessi al fine di prevenirli tramite un'adeguata azione di pattugliamento. Il primo si chiama Key Crime, è stato sviluppato nel 2004 dall'assistente capo della Polizia di Stato Mario Venturi²⁶³, ed è in grado di analizzare i dati di indagine dei reati commessi per fornire un'indicazione probabilistica sulle future azioni criminali segnalando dove potranno consumarsi con specifico riferimento a quelle "seriali": il programma, infatti, "nulla può per prevenire o accertare fatti criminali occasionali o fatti anche potenzialmente seriali all'atto della loro insorgenza"²⁶⁴. È particolarmente indicato per le rapine, ma può essere adattato, per esempio, anche alle truffe e ai furti in appartamento o alle violenze carnali, visto che il meccanismo alla base è

²⁶² C. Cavaceppi, 2020, 110; L'estrazione dei dati opera sui fascicoli presenti nel sistema Calliope (un protocollo informatico che costituisce l'infrastruttura di base tecnico funzionale su cui avviare il processo di ammodernamento e di trasparenza del Ministero della Giustizia), proponendo un'interfaccia di analisi utilizzabile anche ai fini statistici, il tutto quasi senza ricorrere ad attività manuali di inserimento dati da parte dei funzionari dell'Ufficio.

²⁶³ G. Taurisano, 2016, 3

²⁶⁴ C. Parodi-V. Sellaroni, 2019, 56

sostanzialmente lo stesso²⁶⁵. È stato sperimentato *in primis* dalla Questura di Milano a partire dal 2007 producendo risultati notevoli: la soluzione dei casi trattati con Key Crime è passata dal 27% nel 2007 al 54% nel 2013, con picchi dell'80% (nell'anno 2013) in casi di rapine ai danni delle farmacie. Si è generato anche un significativo effetto di deterrenza, riducendo il numero di rapine alle farmacie nei primi 10 mesi del 2014 del 25%²⁶⁶.

Il buon esito di questo primo impiego ha fatto sì che, nel 2009, la Questura decidesse di estenderne l'utilizzo al contrasto alle rapine ai danni di istituti di credito e ampliando l'area di impiego dalla città all'intera provincia di Milano. Anche in quest'ambito i risultati sono stati piuttosto netti, sino a giungere nel 2013 all'individuazione delle responsabilità per il 74% degli eventi commessi. Inoltre, nei primi 10 mesi nel 2014 si è registrata una riduzione del 48% di tali reati²⁶⁷.

Il *database* è formato da due macro-tipologie di dati:

- le informazioni fornite dalle persone offese tramite la compilazione, ad opera della polizia giudiziaria, di alcuni moduli;
- tutti gli altri elementi oggettivi relativi al fatto, come le immagini delle telecamere, le tracce biologiche, gli accertamenti diretti della polizia giudiziaria, ecc²⁶⁸.

Successivamente i dati vengono trasferiti nel sistema e confrontati costantemente tra loro, "al fine di localizzare gli eventi e classificare gli stessi in base alle caratteristiche personali dell'autore dei fatti"²⁶⁹: in questo modo, rispetto a ogni nuovo episodio vengono identificate automaticamente similitudini e dissonanze rispetto ai precedenti. Il programma è inoltre in grado di verificare le abitudini degli autori dei reati, come le espressioni utilizzate, gli orari dei fatti, le modalità di allontanamento del luogo dell'evento, ecc., e l'esistenza di elementi ricorrenti nei vestiti e negli oggetti presenti nelle immagini del *database*²⁷⁰. In sintesi,

²⁶⁵ C. Parodi-V. Sellaroni, 2019, 57

²⁶⁶ M. Venturi, 2014, 7

²⁶⁷ M. Venturi, 2014, 7

²⁶⁸ C. Parodi-V. Sellaroni, 2019, 57

²⁶⁹ C. Parodi-V. Sellaroni, 2019, 57

²⁷⁰ C. Parodi-V. Sellaroni, 2019, 57; L'efficacia del programma potrebbe aumentare se il confronto dei dati fosse programmato per avvenire anche su aree limitrofe, o addirittura su macroaree. In tali casi emergerebbe un problema di coordinamento gestibile con direttive generali degli organi centrali, forme di coordinamento tra uffici giudiziari già previste in termini generali dal codice di procedura penale agli articoli 371 e seguenti o precisi e puntuali interventi

citando Mario Venturi, “la filosofia analitica del *software* parte dall’osservazione del particolare per arrivare al generale per poi ridiscendere dal generale al particolare, ottenendo così una capacità previsionale basata, da ultimo, su modelli matematici in grado di predire un’azione futura”²⁷¹.

L’aspetto problematico non è tanto la previsione della fascia oraria e del luogo in cui probabilmente si verificherà il reato, ma la possibilità, tramite i dati in possesso di Key Crime, di ricondurre a un certo soggetto una serie di reati di cui prima si ignorava l’autore. Riportando le parole di Cesare Parodi e Valentina Sellaroni: “in che termini l’insieme di elementi tratti dal data base del programma possono “entrare” – in esito all’elaborazione – in un’annotazione di p.g. finalizzata – al contrario – a una richiesta di misura cautelare o, poniamo, di una richiesta di intercettazione o quale elemento di prova al fine dell’esercizio dell’azione penale? In che modo potranno essere, tali elementi e soprattutto le interconnessioni individuate, utilizzate in sede dibattimentale? Che tipo di valutazione è “accettabile” per il sistema penale in relazione a tali esiti?”²⁷²

In primo luogo, tutti gli elementi raccolti nel *database* devono essere acquisiti secondo ciò che stabiliscono le norme codicistiche. Pur potendo apparire come una considerazione scontata e superficiale, in realtà è la linea di discriminazione fondamentale tra ciò che può essere introdotto e ciò che non può essere introdotto nel processo. Infatti, se nell’ottica della polizia di prevenzione, ferme restando eventuali responsabilità per le singole condotte dirette ad acquisire elementi al di fuori dei termini di legge, tutto il materiale è diretto alla prevenzione di reati, mentre nell’ottica della polizia giudiziaria potranno essere utilizzati solo gli elementi acquisiti in ossequio alla disciplina codicistica contenuta nel Libro Terzo del Codice di procedura penale²⁷³.

In secondo luogo, gli elementi che sono posti alla base delle comparazioni devono poter essere singolarmente portati all’attenzione del giudice: le singole immagini che ritraggono,

normativi. Si tratta di capire se e in che termini l’interesse generale corrisponda alla massima diffusione possibile di un programma simile.

²⁷¹ M. Venturi, 2014, 6

²⁷² C. Parodi-V. Sellaroni, 2019, 58

²⁷³ C. Parodi-V. Sellaroni, 2019, 58

per esempio, il medesimo cappellino indossato durante le rapine o le dichiarazioni delle persone informate che descrivono la corporatura di un autore di molteplici rapine potrebbero essere oggetto di “ostensione” al Tribunale, nel primo caso, e oggetto di testimonianza (o di valutazione nell’ambito di un verbale, in fase di indagine) nel secondo caso? Se la risposta è affermativa, si giunge a un problema ulteriore: in base a quali criteri la comparazione ha portato, tra migliaia di elementi, a evidenziare una serie di relazioni che si assumono significative sul piano della ricostruzione delle responsabilità? In che modo è possibile affermare che tali relazioni siano idonee a costituire il fondamento di responsabilità penali? La risposta a tali quesiti non può che poggiare sulle future disposizioni in materia di intelligenza artificiale che, tra le altre cose, devono stabilire i parametri con cui si può considerare affidabile l’*output* di una macchina predittiva.

Il secondo *software* su cui è bene soffermarsi è XLAW, un programma simile a Key Crime usato dapprima in via sperimentale nella Questura di Napoli, mentre ora in uso anche a Prato, Salerno e Venezia, che viene descritto come “l’algoritmo-poliziotto che prevede furti e rapine”. Come riportato in un’intervista al suo stesso ideatore in occasione del conferimento del “Premio innovazione SMAU”, XLAW è stato sviluppato dall’Ispettore Superiore Elia Lombardo con l’obiettivo di prevenire la commissione dei reati di tipo predatorio, come furti, furti con strappo, furti commessi con destrezza, rapine e truffe, usando l’intelligenza artificiale, il machine learning e le conoscenze criminologiche. Questo tipo di delitti sono tipicamente realizzati da soggetti professionali che hanno bisogno di un profitto veloce e che usano le stesse strategie indipendentemente dal centro urbano in cui si trovano: ci devono essere delle prede appetibili e delle vie di fuga, il che rende alcune zone delle vere e proprie “riserve di caccia”. Nello specifico, ciò che l’autore del reato cerca è un luogo in cui si susseguano delle fasi o delle operazioni ricorrenti: entrata e uscita da abitazioni, uffici, scuole, mercati, esercizi commerciali, arrivo di treni, autobus, navi da crociera ecc. Sovrapponendo i crimini a tali dinamiche socioeconomiche ci si accorge che

avvengono in concomitanza a queste fasi o operazioni regolari, il che permette di prevedere efficacemente i singoli delitti²⁷⁴.

Da un punto di vista tecnico, XLAW è dotato di un algoritmo di tipo euristico che, in base alle caratteristiche del territorio in esame, ai delitti denunciati dai cittadini e ad altre informazioni in possesso della Polizia, ricerca e mostra i modelli criminali ricorrenti riuscendo a prevedere quando e dove verranno commessi i successivi reati. In tal modo il poliziotto ha a disposizione una mappa costantemente aggiornata che gli comunica, oltre che il momento e il luogo di consumazione del delitto, il tipo di crimine, il modus operandi dell'autore e il tipo di preda. Dopodiché si passa alla fase operativa in cui le pattuglie vengono inviate nei posti segnalati dal sistema²⁷⁵.

Da quando, nel 2004, è iniziata la sperimentazione a Napoli, i crimini predatori sono diminuiti del 22%, mentre le denunce e gli arresti in flagranza di reato sono aumentati del 24%. In aggiunta, dice l'Ispezzore Lombardo: "Abbiamo [...] tagliato i costi di gestione della sicurezza, riducendo lo stress di uomini e mezzi a disposizione per assicurarla. Per esempio, se prima un'auto pattuglia percorreva mediamente circa 180Km, grazie all'algoritmo che perlustra il territorio ed evita di girare a vuoto, i Km percorsi sono scesi mediamente a 24/30 al giorno"²⁷⁶.

4.2.2 *Come sta cambiando il diritto penale*

È ora possibile tirare le somme per comprendere la direzione che potrebbe prendere il diritto penale se si utilizzassero in modo più pervasivo le tecnologie predittive e l'IA. Per quanto riguarda l'ambito repressivo, in dottrina si ipotizza una loro applicazione²⁷⁷:

- In materia cautelare ex art. 274 lett. c) c.p.p. come fattore da cui desumere la possibile reiterazione del reato. La disposizione, infatti, prevede che le misure

²⁷⁴ G. Di Gennaro-E. Lombardo, 2020, 677; Per approfondire la basi teoriche di XLAW si veda P. M. Mayhew-R. V. Clarke-A. Sturman-J. M. Hough, 1976; D. B. Cornish-R. V. Clarke (eds.), 2014

²⁷⁵ Smau, 2018, 2

²⁷⁶ Smau, 2018, 2

²⁷⁷ C. Cavaceppi, 2020, 112

possano essere disposte quando “per specifiche modalità e circostanze del fatto e per la personalità della persona sottoposta alle indagini o dell'imputato, desunta da comportamenti o atti concreti o dai suoi precedenti penali, sussiste il concreto e attuale pericolo che questi commetta gravi delitti [...] della stessa specie di quello per cui si procede” e l'ipotetico *tool* predittivo potrebbe essere posto a fondamento della decisione, posto che possa dimostrare ciò che richiede la legge;

- Nella valutazione della pericolosità sociale dell'imputato ex art. 203 c.p., essendo richiesta come condizione per disporre le misure di sicurezza. Secondo la norma in esame, “la qualità di persona socialmente pericolosa si desume dalle circostanze indicate nell'articolo 133”, che, oltre a indicare i fattori relativi alla gravità del reato, elenca quattro caratteristiche da cui si desume la capacità a delinquere dell'imputato: i motivi a delinquere e il carattere del reo; i precedenti penali e giudiziari e, in genere, la condotta e la vita del reo, antecedenti al reato; la condotta contemporanea o susseguente al reato; le condizioni di vita individuale, familiare e sociale del reo”;
- Nel procedimento di prevenzione ex art. 6 d.lgs 159/2011, visto il richiamo alla pericolosità per la sicurezza pubblica come presupposto per l'applicazione della sorveglianza speciale;
- In sede dibattimentale per valutare la concessione del beneficio della sospensione condizionale della pena ex art. 163 e 164 c.p., posto che tale misura “è ammessa soltanto se, avuto riguardo alle circostanze indicate nell'articolo 133, il giudice presume che il colpevole si asterrà dal commettere ulteriori reati”.

Un possibile punto di frizione nell'adozione dei sistemi predittivi nel processo penale è dato dall'articolo 220 co. 2 c.p.p.²⁷⁸, che vieta le perizie volte a “stabilire l'abitualità o la professionalità nel reato, la tendenza a delinquere, il carattere e la personalità dell'imputato e in genere le qualità psichiche indipendenti da cause patologiche”, salvo quanto previsto in materia di esecuzione di pena e misure di sicurezza, il che limita molto (almeno allo stato attuale) l'ingresso di tali tecnologie nelle aule di giustizia.

²⁷⁸ C. Cavaceppi, 2020, 112; Si veda anche L. Quattrocchio, 2019, 1761 ss.

Un'altra criticità emersa e discussa più volte è la verifica della qualità, della trasparenza e dell'accessibilità dei dati elaborati dall'algoritmo²⁷⁹: chi dovrebbe controllare la sussistenza di questi elementi? Il Ministero della Giustizia, una commissione ad hoc di giudici non coinvolti nel processo in corso o degli altri soggetti? I controlli andrebbero fatti su base periodica al rilascio di ogni nuovo aggiornamento, ad ogni utilizzo o a campione? Questi sono tutti interrogativi a cui va data una risposta, tra le altre cose, per aumentare la fiducia dei consociati verso gli strumenti intelligenti e per giustificare la vincolatività degli *output* per il giudice.

Il *quid pluris* offerto dall'IA potrebbe essere la riduzione delle pendenze giudiziarie nei casi più "semplici" o ripetitivi in cui ci si può permettere di sveltire i processi affidando in più o meno larga parte, con il consenso dell'imputato, la decisione a un sistema esperto. Si potrebbe obiettare che non sarebbe corretto far decidere una questione così delicata, ossia la scelta se farsi giudicare da un robot o da un essere umano, alle sole parti processuali, ma forse è bene ricordare che, attualmente, l'imputato può scegliere liberamente di rinunciare alle garanzie date dall'impianto tendenzialmente accusatorio del processo ordinario e optare per un processo di stampo inquisitorio come il rito abbreviato, la cui *ratio* è ridurre il carico giudiziario²⁸⁰. A parità di necessità, non sarebbe peregrino affidare a un robot che rispetti tutti i parametri di affidabilità, trasparenza e controllo previsti dalla Carta etica del CEPEJ, il compito di decidere, su richiesta della difesa, i casi di competenza del Giudice di Pace, ossia quelli che hanno ad oggetto i cd. reati bagatellari.

Questa scelta farebbe riemergere uno dei concetti alla base delle proposte legislative in discussione in questi anni: la neutralità dell'intelligenza artificiale. Questa, infatti, dipende dagli *input* dati dall'utilizzatore, dal programmatore o dal processo di autoapprendimento frutto del *machine learning* e, come si è visto, può portare a sistemi affetti da pregiudizi che mal si conciliano con i principi stabiliti a livello europeo. È tuttavia evidente che la scelta di prediligere un'IA neutrale dipende dall'obiettivo politico e sociale che si vuole raggiungere: "l'imparzialità o l'obiettività dell'algoritmo non è un qualcosa di assoluto ma dipenderà dagli

²⁷⁹ C. Cavaceppi, 2020, 112

²⁸⁰ C. Giovanni, 2000; Per approfondimenti sul tema si segnalano, tra i tanti, L. Degl'Innocenti-M. De Giorgio, 2006; P. Tonini, 2017, 800 ss.

obiettivi politici e, quindi, dalle scelte normative che si faranno in un dato momento storico”²⁸¹. Quantomeno nel contesto occidentale, acquisisce notevole importanza la trasparenza della macchina tanto a valle quanto a monte, dovendo essere garantita la conoscibilità dei suoi processi decisionali sia in capo all’indagato/imputato che in capo al legislatore, che dovrà verificare se gli obiettivi che si è prefissato sono stati soddisfatti o meno.

Si potrebbe assistere alla minimizzazione della lesione dei beni giuridici protetti dalle norme penali grazie all’anticipazione della soglia di tutela assicurata da applicativi come XLAW e Key Crime. Ciò ribalterebbe l’assetto liberale del diritto penale, consacrando la diversa visione di stampo securitario e portando all’inevitabile sacrificio della libertà di agire dei consociati, che sarebbero costantemente valutati dal sistema. Sul fronte prettamente applicativo, delle macchine ben sviluppate potrebbero assicurare un’applicazione delle norme il più possibile oggettiva attraverso l’uso di dati neutrali, spogliati cioè di elementi discriminatori²⁸².

Se si volesse mantenere l’impianto liberale e controfattuale del diritto penale usando al contempo l’IA, non si potrebbe che accettare di tutelare i beni giuridici di alcuni *ex post*, evitando di sacrificare la libertà dell’intera collettività per ragioni di prevenzione: delle due l’una. Se si assicurassero i diritti costituzionali e si rispettassero i principi europei, si potrebbe delineare una prospettiva mediana, in cui l’IA verrebbe usata in ambiti circoscritti e sempre sotto il controllo tanto dell’autorità quanto dell’imputato. Al contrario, l’estrema conseguenza di un diritto penale “intelligente” è un diritto penale autoritario, dove le esigenze del singolo cedono fatalmente e inevitabilmente il passo a quelle dello Stato²⁸³.

²⁸¹ C. Cavaceppi, 2020, 118

²⁸² C. Cavaceppi, 2020, 119

²⁸³ C. Cavaceppi, 2020, 120

CONCLUSIONI

L'interrogativo a cui si voleva dare risposta con questo contributo, ossia se le macchine dotate di intelligenza artificiale siano idonee all'utilizzo nel diritto, non può ancora trovare una risposta definitiva. Ciò che si è visto è lo stato attuale della giustizia predittiva sia da un punto di vista tecnologico che giuridico, usando come fulcro argomentativo COMPAS e il caso Loomis e osservando i problemi più importanti in materia.

Nonostante il dibattito sia ancora in corso e la tecnologia sia ancora in fase di sviluppo, i sistemi esperti sono già una realtà che trova applicazione in alcune Corti e studi di avvocati, di conseguenza l'approfondimento del tema non può più essere rimandato. Si è visto che un uso troppo ottimistico di tali strumenti (che possono essere una risorsa preziosissima, se sviluppati correttamente) può portare a sentenze come quella della Corte Suprema del Wisconsin, in cui si è deciso di usare un programma poi verificatosi discriminatorio e affetto da pregiudizi.

L'UE ha deciso di occuparsene circondando le norme in discussione di garanzie per i diritti fondamentali dei cittadini, ma ciò che preoccupa è se sia davvero possibile creare dei sistemi esperti trasparenti, sicuri e affidabili o se questi parametri rappresentino solamente un disincentivo per le società, che potrebbero trovare troppo oneroso produrli e commercializzarli nei paesi europei.

Il diritto penale, e più in generale l'ordinamento giuridico, nei prossimi anni potrebbe subire un mutamento per accogliere programmi come COMPAS, XLAW e Key Crime, sistemi capaci di prevedere la commissione di reati e la probabilità di recidiva degli imputati: già solo questo mette in discussione il modo in cui si è concepito il diritto fino a questo momento, obbligandoci a ripensare alcuni principi che si davano ormai per certi, come il principio di non colpevolezza, il diritto a un giusto processo e il principio di uguaglianza.

BIBLIOGRAFIA

Adriana Valente, *SEFIT: la progettazione di un sistema esperto per l'accesso al fondo per l'innovazione tecnologica*, in *Informatica e diritto*, XIV Annata, Vol. XIV, n. 3, 1988, pp. 223-234

Alan Mathison Turing, *Computing Machinery and Intelligence*, in *Mind*, Volume LIX, Issue 236, 1950, pp. 433-460

Alessandra Carleo (a cura di), *Calcolabilità giuridica*, il Mulino, 2017

Alicia Solow-Niederman, YooJung Choi, Guy Van den Broeck, *The Institutional Life of Algorithmic Risk Assessment*, in 34 *Berkeley Tech. L.J.* 705, 2019, pp. 705-744

Amedeo Santosuosso, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Mondadori Università, 2020

Andrea Pisaneschi, *Diritto costituzionale*, Giappichelli, 2016

Antonio Anselmo Martino, *Introduzione ai sistemi esperti nel diritto*, in *Informatica e diritto*, XIV Annata, Vol. XIV, n. 3, 1988, pp. 5-14

Antonio D'Aloia, *Intelligenza artificiale e diritto: come regolare un mondo nuovo*, Franco Angeli, Milano, 2020

Antonio Felice Uricchio, Giuseppe Riccio, Ugo Ruffolo (a cura di), *Intelligenza Artificiale tra etica e diritti. Prime riflessioni a seguito del libro bianco dell'Unione europea*, Cacucci Editore, 2020

Brandon L. Garrett, John Monahan, *Judging Risk*, in *California Law Review*, Vol. 108, 2020, pp. 439-493

Carlo Bona, Rino Rumiati, *Psicologia cognitiva per il diritto. Ricordare, pensare e decidere nell'esperienza forense*, Il Mulino, 2013

Carlotta Conti, *Giusto processo*, in *ilPenalista*, ilpenalista.it/bussola/giusto-processo, ultima visita il 15 settembre 2022

Carmelo Asaro, Antonio Martino, Ephraim Nissan, *The Daedalus System: A Tool For The Italian Investigating Magistrate And The Prosecutor. A Sample Session: Investigating An Extortion Case*, in *Computing and Informatics*, Vol. 20, 2001, pp. 1-40, V 15.3

Cesare Parodi, Valentina Sellaroli, *Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*, in *Diritto Penale Contemporaneo*, Fascicolo 6/2019, pp. 47-71

Cesare Salvi, *La responsabilità civile*, Giuffrè Francis Lefebvre, 1998

Charles D. Stimson, *The First Step Act's Risk and Needs Assessment Program: A Work in Progress*, in *Heritage Foundation Legal Memorandum*, No. 265, 2020

Claudia Cevenini, *I servizi delle società dell'informazione. Profili informatico-giuridici*, Gedit, 2004

Comment Letter to the Department of Justice on PATTERN First Step Act, 2019

Commissione europea per l'efficienza della giustizia (CEPEJ), *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, CEPEJ(2018)14

Commissione europea, *Allegati della Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, COM/2021/206 final ANNEXES 1 to 9

Commissione europea, *Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, COM/2021/206 final

Comunicazione della Commissione, *L'intelligenza artificiale per l'Europa*, COM/2018/237, 25/04/2018

Corte di cassazione, Sez. Quarta Penale, Sentenza n.26257 del 25/05/2017

Corte di Giustizia UE, Sentenza 28/6/2005 - Cause Riunite C-189/02 P, C-202/02 P, da C-205/02 P a C-208/02 P e C-213/02 P

Corte di Giustizia UE, Sentenza 8/2/2007 - Causa C-3/06 P

Derek B. Cornish, Ronald V. Clarke (eds.), *The Reasoning Criminal: Rational Choice Perspectives on Offending*, Transaction Publishers, 2014

Diana Zingales, *Risk assessment: una nuova sfida per la giustizia penale?*, in *Diritto Penale e Uomo*, Fascicolo 12/2021, pp. 2-22

Direttiva 85/374/CEE del Consiglio del 25 luglio 1985

Domenico Dalfino, *Norme "vaghe", poteri del giudice e sindacato di legittimità*, in *Questione Giustizia*, Fascicolo 1/2020

Edward Hance Shortliffe, *Computer-Based Medical Consultations: MYCIN*, Elsevier Computer Science Library, 1976

Elio Fameli, *L'"automa infortunistico": un esperimento di consulenza giuridica automatica*, in *Informatica e diritto*, *Il Annata*, Vol. II, n. 1, 1976, pp. 1-50

Emmelyn A. J. Croes, Marjolijn L. Antheunis, *Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot*, in *Journal of Social and Personal Relationships*. 2021;38(1), pp. 279-300

Equivant, *Practitioner's Guide to COMPAS Core*, 2019

Erica Palmerini, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Responsabilità civile*, 6, 2016, pp. 1826 ss.

Ermanno Vitale (a cura di), Luigi Ferrajoli, *Diritti fondamentali. Un dibattito teorico*, Laterza, 2001

Fabio Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, Fascicolo 10/2019, pp. 1-34

Federica Romanazzi, *Il danno non patrimoniale contrattuale*, in *Diritto.it*, Gruppo Maggioli, 2 aprile 2021

Federico Sorrentino, *I principi generali dell'ordinamento giuridico nell'interpretazione e nell'applicazione del diritto*, 1987

Filippo Donati, *Diritti fondamentali e algoritmi nella proposta di regolamento sull'intelligenza artificiale*, in *Il Diritto dell'Unione Europea*, Fascicolo 3-4, 2021, pp. 453-466

First Step Act (Public Law 115-391), 2018

Franco Pizzetti (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, 2018

Gabriel Hallevy, *Liability for Crimes Involving Artificial Intelligence Systems*, Springer, 2015

Gabriel Hallevy, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, in *Akron Intellectual Property Journal*, Vol. 4, Iss. 2, 2010, pp. 171-201

Giancarlo Taddei Elmi, Alfonso Contaldo (a cura di), *Intelligenza artificiale. Algoritmi giuridici lus condendum o "fantadiritto"?*, Pacini Giuridica, 2020

Giangiaco Olivi, Chiara Bocchi, *Regulating artificial intelligence in the EU: Top 10 issues for businesses to consider*, in *Dentos*, 2021

Gino Roncaglia, *Computer che copiano: test di Turing, web corpora e filtraggio collaborativo*, in *UnitusOpen*, 2014

Giovanni Canzio, *Giudizio abbreviato* [agg. IV, 2000], in *Enciclopedia del diritto*, Giuffrè Francis Lefebvre, enciclopediadeldiritto.giuffrefrancislefebvre.it/#/detail/298061/2719563, ultima visita il 29 ottobre 2022

Giovanni D'Amico, *La Riforma del Codice del Consumo*, CEDAM, 2015

Giovanni De Cristofaro, Alessio Zaccaria, *Commentario breve al diritto dei consumatori. Codice del consumo e legislazione complementare*, CEDAM, 2013

Giovanni Sartor, *Intelligenza artificiale e diritto. Un'introduzione*, Dott. A. Giuffrè Editore, 1996

Giovanni Sartor, *L'informatica giuridica e le tecnologie dell'informazione*, Giappichelli Editore, 2016

Giovanni Sartor, *Le applicazioni giuridiche dell'intelligenza artificiale*, Dott. A. Giuffrè Editore, 1990

Giulia Venturi, *Investigating legal language peculiarities across different types of Italian legal texts: an NLP-based approach*, in Proceedings, IAFL Porto, 2012, pp. 138-156

Giuseppe Contissa, Giulia Lasagni, Giovanni Sartor, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in Diritto di Internet n. 4/2019, Pacini Giuridica, pp. 619-634

Giusi Sorrenti, *L'interpretazione conforme a Costituzione*, Dott. A. Giuffrè Editore, 2006

Glicerio Taurisano, *KeyCrime: il "conclusive reasoning" nell'attività anticrimine della Polizia di Stato*, in Profiling. I profili dell'abuso, PrimoPiano, N. 2, Anno 7, giugno 2016

Godofredo Iommi Amunátegui, *Symbolic Languages and Ars Combinatoria*, in Studia Leibnitiana, Bd. 47, H. 2 (2015), pp. 233-239

Gottfried Wilhelm von Leibniz, *Dissertatio de Arte Combinatoria*, 1666

Gouda Naveen, M Ashish Naidu, Dr. B. Thirumala Rao, K Radha, *A Comparative Study on Artificial Intelligence and Expert Systems*, in International Research Journal of Engineering and Technology, Volume 6, Issue 2, Feb 2019, pp. 1980-1986

Graham Webster, Rogier Creemers, Paul Triolo, Elsa Kania, *A Next Generation Artificial Intelligence Development Plan*, New America, 2017

Gruppo di Esperti ad Alto Livello sull'Intelligenza Artificiale, *Orientamenti etici per un'IA affidabile*, 2019

Gruppo di lavoro sulla Strategia Nazionale per l'Intelligenza Artificiale, Ministero dell'Università e della Ricerca, del Ministero dello Sviluppo Economico, del Ministro per l'Innovazione tecnologica e la Transizione Digitale, *Programma Strategico Intelligenza Artificiale 2022-2024*, 2021

Guido Alpa (a cura di), *Diritto e intelligenza artificiale*, Pacini Giuridica, 2020

Gunther Teubner, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, Edizioni Scientifiche Italiane, 2019

H.R. 5682-First Step Act, 2018

H.R.5515, John S. McCain National Defense Authorization Act for Fiscal Year 2019

Han-Wei Liu, Ching-Fu Lin, Yu-Jie Chen, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, in *International Journal of Law and Information Technology*, 2019, 27, 122–141

Hongquan Yang, *The Privacy, Data Protection and Cybersecurity Law Review: China*, *The Law Reviews*, 2019

Indiana Supreme Court, *Anthony Malenchik v. State of Indiana*, 928 N.E.2d 564, No. 79S02-0908-CR-365, 2010

Isaac Asimov, *Io, Robot*, Mondadori, 2018

Jennifer Hill, W. Randolph Ford, Ingrid G. Farreras, *Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations*, in *Computers in Human Behavior*, Volume 49, 2015, pp. 245-250

John Buyers, Tamara Quinn, David Cubitt, Catherine Hammon, *Debate continues over the pros and cons of regulating Artificial Intelligence*, in *Osborne Clarke/Lexology*, 2021

John Rawls, *Political Liberalism Expanded Edition*, Columbia University Press, 2005

John Rogers Searle, *Is the Brain's Mind a Computer Program?*, in *Scientific American*, vol. 262, no. 1, 1990, pp. 25–31

John Rogers Searle, *Minds, brains and programs*, in *Behavioral and Brain Sciences*, Volume 3, Issue 3, 1980, pp. 417-424

Joseph Weizenbaum, *ELIZA - a computer program for the study of natural language communication between man and machine*, in *Communications of the ACM*, Volume 9, Issue 1, 1966, pp. 36-45

Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, *Machine Bias*, ProPublica, 2016

Julia Dressel, Hany Farid, *The accuracy, fairness, and limits of predicting recidivism*, in Science Advances Vol 4, Issue 1, 17 Jan. 2018

K.C. Halm, Katherine Sheriff, Jonathan Mark, Kate Berry, *European Commission Proposes Sweeping New Regulatory Schema for "High-Risk" AI Systems*, in Lexology, 2021

Katherine Freeman, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, in North Carolina Journal of Law & Technology, Volume 18, Issue 5, 2016, pp. 75-106

Kenneth Mark Colby, Sylvia Weber, Franklin Dennis Hilf, *Artificial Paranoia*, in Artificial Intelligence 2, 1971, pp. 1-25

Laura Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in Politica del diritto, Fascicolo 4, dicembre 2018, pp. 713-740

Laura Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in Politica del diritto, 4, 2018, pp. 713-740

Laura Ferrari, *I sistemi esperti e il mondo degli impianti complessi: un esempio di applicazione*, in Liuc Papers, 10, Serie Tecnologia 2, 1994

Leonardo Degl'Innocenti, Mario De Giorgio, *Il Giudizio Abbreviato*, Giuffrè Editore, 2006

Linda J. Skitka, Kathleen L. Mosier, Mark Burdick, *Does automation bias decision-making?*, in Int. J. Human-Computer Studies 51, 1999, pp. 991-1006

Luca Lorenzetti, *Lessico*, in Enciclopedia dell'Italiano, Treccani, 2010, [treccani.it/enciclopedia/lessico_\(Enciclopedia-dell'Italiano\)](http://treccani.it/enciclopedia/lessico_(Enciclopedia-dell'Italiano)), ultima visita il 12 novembre 2022

Luigi Viola, *Interpretazione della legge con modelli matematici. Processo, a.d.r., giustizia predittiva*, Diritto Avanzato Edizioni, 2018

Maintaining American Leadership in Artificial Intelligence, Executive Order 13859 of February 11, 2019

Maria Caterina Falchi, *Intelligenza Artificiale: se l'algoritmo è discriminatorio*, in *Ius in itinere*, 2020

Mario Caterini, *Il Giudice Penale Robot*, in *La legislazione penale*, 2020

Mario Venturi, *KeyCrime. La chiave del crimine*, in *Profiling. I profili dell'abuso*, PrimoPiano, N. 4, Anno 5, dicembre 2014

Massimo La Torre, *Positivismo giuridico* [Annali VI, 2013], in *Enciclopedia del diritto*, Giuffrè FrancisLefebvre, enciclopediadeldiritto.giuffrefrancislefebvre.it/#/detail/4012821/21629745, ultima visita il 13 novembre 2022

Massimo Luciani, *Certezza del diritto e clausole generali*, in *Questione Giustizia*, Fascicolo 1/2020

Massimo Luciani, *Costituzionalismo irenico e costituzionalismo polemico*, in *Giurisprudenza costituzionale*, Giuffrè Editore Spa, Vol. IV, 2006, pp. 1643-1668

Matteo Fini, Paola Milani, *Intelligenza e Coscienza. L'IA tra Searle e Dennett: sviluppi dell'Intelligenza Artificiale*, Department of Economics University of Milan Italy, Departmental Working Papers, 2005

Michele Iaselli, *Le nuove prospettive nel campo dell'informatica giudiziaria*, in *Diritto&Diritti*, maggio 2001

Mitja Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto Penale Contemporaneo*, 29 maggio 2019

National Security Commission on Artificial Intelligence, *Final Report*, 2021

Nicoletta Calzolari, Alessandro Lenci, *Linguistica computazionale. Strumenti e risorse per il trattamento automatico della lingua*, in *Mondo digitale*, n.2, giugno 2004

Open Letter to the European Commission: Artificial Intelligence and Robotics, robotics-openletter.eu

- P. Mayhew, R. V. G. Clarke, A. Sturman, J. M. Hough, *Crime As Opportunity*, Home Office Research Study No. 34, 1976
- Paola Severino (a cura di), *Intelligenza artificiale. Politica, economia, diritto, tecnologia*, LUISS University Press, 2022
- Paolo Moro, Claudio Sarra (a cura di), *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, FrancoAngeli, 2017
- Paolo Tonini, *Manuale di Procedura Penale*, Giuffrè Editore, 2017
- Pasquale Stanzone, Antonio Musio (a cura di), *La tutela del consumatore*, Giappichelli, 2009
- Peter R. Wurman, Samuel Barrett, Kenta Kawamoto, et al., *Outracing champion Gran Turismo drivers with deep reinforcement learning*, *Nature* 602, 2022, pp. 223–228
- Pier Luigi M. Lucatuorto, *Artificial Intelligence and Law: Judicial Applications of Expert Systems (Intelligenza Artificiale e Diritto: Le Applicazioni Giuridiche dei Sistemi Esperti)*, in *Cyberspace and Law (Ciberspazio e Diritto)*, Vol. 7, No. 2, 2006, pp. 219-242
- Rebecca Arcesati, *Lofty Principles, Conflicting Interests. AI ethics and governance in China*, in *MERICS China Monitor*, 2021
- Remo Trezza, *Diritto e intelligenza artificiale. Etica, Privacy, Responsabilità, Decisione*, Pacini Giuridica, 2020
- Renato Borruso, Gianluigi Ciacci, *Diritto civile e informatica*, in *Trattato di Diritto Civile del Consiglio Nazionale del Notariato*, Edizioni Scientifiche Italiane, 2005
- Renato Borruso, Rosa Maria Di Giorgi, Leonello Mattioli, Mario Ragona, *L'informatica del diritto*, Dott. A. Giuffrè Editore – Milano, 2004
- Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, P8_TA(2017)0051
- Roberta Nannucci (a cura di), *Lineamenti di informatica giuridica. Teoria, metodi, applicazioni*, Edizioni Scientifiche Italiane, 2002

Rosanna Celella, *L'algoritmo che condanna: i limiti della giustizia predittiva*, in Data Protection Law, 2019

Salvi Cesare, *Risarcimento del danno* [XL, 1989], in Enciclopedia del diritto, Giuffrè Francis Lefebvre, enciclopediadeldiritto.giuffrefrancislefebvre.it/#/detail/297568/2647151, ultima visita il 24 novembre 2022

Serena Quattrocolo, *Equo processo penale e sfide della società algoritmica*, in BioLaw Journal – Rivista di BioDiritto, n. 1, 2019, pp. 135-144

Serena Quattrocolo, *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in Cass. Pen, 2019, 1761 ss.

Sergio Vinciguerra, *Principi di criminologia*, Cedam, 2013

Silvia Gasparini, *Appunti minimi di storia del diritto 2: Età moderna e contemporanea*, Imprimerie, Padova, 2002

Simone Cedrola, *Robot e responsabilità penale: prospettive attuali e future*, in IusInItinere, 15/05/2017, aggiornato il 15/01/2019

Smau, *XLAW: L'algoritmo-poliziotto che prevede furti e rapine*, 2018

Stefania Carrer, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in Giurisprudenza Penale Web, 2019, pp. 4

Stuart Russel, Peter Norvig, *Artificial Intelligence. A Modern Approach*, Pearson Education, Inc., 2010

Supreme Court of Wisconsin, *State of Wisconsin v. Eric L. Loomis*, Case no. 2015AP157-CR, 5 April – 13 July 2016

The Constitution of the United States, 1789

Trabucchi Alberto, *Buon costume* [V, 1959], in Enciclopedia del diritto, Giuffrè Francis Lefebvre, enciclopediadeldiritto.giuffrefrancislefebvre.it/#/detail/296332/2457703, ultima visita il 24 novembre 2022

Tullio Delogu, *Lo “strumento” nella teoria generale del reato*, in Riv. it. dir. proc. pen., 1974

U.S. Department of Justice, Office of the Attorney General, *The First Step Act of 2018: Risk and Needs Assessment System*, 2018

U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System - UPDATE*, 2020

U.S. Supreme Court, *Buck v. Davis*, 580 U.S. ___, 15-8049, 2017

U.S. Supreme Court, *Gardner v. Florida*, 430 U.S. 349, No. 74-6593, 1977

Ubertis Giulio, *Giusto processo* (dir. proc. pen.) [Annali II-1, 2008], in Enciclopedia del diritto, GiuffrèFrancisLefebvre, enciclopedia.deldiritto.giuffrefrancislefebvre.it/#/detail/295873/2422450, ultima visita il 29 novembre 2022

Umberto Vincenti, *Introduzione all’etica pubblica*, Edizioni Libreria Progetto Padova, 2020

Vito Velluzzi (a cura di), *Significato letterale e interpretazione del diritto*, G. Giappichelli Editore, 2000

Vittorio Falzone, Filippo Palermo, Francesco Cosentino (a cura di), *La Costituzione della Repubblica Italiana illustrata con i lavori preparatori*, Colombo in Roma Via Campo Marzio, 1948

Vittorio Villa, *Disaccordi interpretativi profondi. Saggio di metagiurisprudenza ricostruttiva*, G. Giappichelli Editore, 2017

Western Australian Supreme Court, *Director of Public Prosecutions v. Mangolamara*, WASC 71; 169 A Crim R 379, 2007

Yoon Chae, *U.S. AI Regulation Guide: Legislative Overview and Practical Considerations*, in Robotics, Artificial Intelligence & Law, Vol. 3, No. 1, Gennaio-Febbraio 2020, pp. 17-40

RINGRAZIAMENTI

Per concludere questo contributo e, con esso, il mio percorso universitario, è doveroso procedere con dei ringraziamenti. *In primis*, ringrazio il mio relatore, il professor Claudio Sarra, per avermi assistito e aiutato in questi mesi di ricerca e scrittura.

Ringrazio, inoltre, tutte le persone che mi sono state vicine sia dentro che fuori dalle aule del Bo per il loro insostituibile supporto: i miei genitori, mio fratello e i miei amici più stretti.

In questi anni ho dovuto dar fondo a tutta la caparbia e l'impegno che possedevo, a volte scoprendo lati di me che non conoscevo, ed è anche grazie a queste persone se sono riuscito a raggiungere il tanto desiderato traguardo.

“Draco dormiens nunquam titillandus”