



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

La colpa del Golem.

Il Responsibility Gap e l'impatto sulle categorie
filosofico-giuridiche

A.A. 2021/2022

Relatore

chiar.mo prof. Stefano Fuselli

Correlatore

chiar.mo prof. Claudio Sarra

Laureando

Michele Zanellato

Matricola n. 1169955

A mio nonno Lanfranco

Indice

Introduzione	ii
1 Da sogno a realtà	1
1.1 Il sogno del rabbino	1
1.2 I passi del Golem: verso le Brain-Computer Interfaces	3
2 Il rabbino si fa Golem: le Brain-Computer Interfaces	7
2.1 Classificazione delle BCI	8
2.2 Ambiti di impiego delle interfacce. Le <i>Assistive technologies</i> .	11
3 Il Responsibility Gap	17
3.1 Introduzione	17
3.2 Responsibility Gap	18
3.3 <i>Please, mind the gap</i> . Responsabilità e interfacce	19
3.4 Il Golem a processo. Il penale	22
3.4.1 Un caso limite	22
3.4.2 Il Golem tra Arizona e Italia	22
3.5 L'intenzione	25
3.5.1 La disciplina dell'Arizona	33
3.5.2 L'intenzione nell'ordinamento penale italiano. Il dolo .	34
3.5.3 L'intenzione nell'ordinamento penale italiano. La colpa	39
3.6 L'azione	41
3.6.1 Cosa è azione?	42
3.6.2 La condotta negli ordinamenti penali	44
3.6.3 Il movimento corporeo	46
3.6.4 La volontarietà	48
3.6.5 Il controllo	48
3.7 Il Golem a processo. Il civile	51
4 L'uomo, il Golem	55

4.1	Introduzione	55
4.2	L'uomo o il suo cervello?	57
4.2.1	Il libero arbitrio	63
4.3	L'uomo o il suo algoritmo?	66
4.3.1	Identità e continuità psicologica	77
4.4	Umanità e individualità	81
5	Il costo della responsabilità	83
5.1	Introduzione	83
5.2	Integrità mentale	84
5.2.1	Privacy neurale	88
5.3	Autodeterminazione e libertà cognitiva	92
5.3.1	Tra cause e motivi	98
6	Conclusioni	103

Introduzione

Lo scopo del presente lavoro è indagare l'impatto sulle categorie filosofico-giuridiche di un particolare tipo di neurotecnologia rappresentato dalle c.d. *Brain- Computer Interfaces*. Più nello specifico, la trattazione verterà sul problema noto come *Responsibility Gap*: per la particolare dinamica di interazione che connette i neurodispositivi ai propri utenti, possono sorgere particolari tensioni con alcune tradizionali categorie giuridiche.

Le ricadute sulle medesime verranno analizzate attraverso la presentazione di un caso fittizio che vede come protagonista una donna in grado di muovere le protesi robotiche di cui è dotata attraverso l'attivazione di una particolare interfaccia installata nel suo cervello. L'analisi verrà condotta prevalentemente sul versante penalistico, dove si individua con maggiore evidenza la portata dirompente del *Responsibility Gap*.

Sotto un profilo più filosofico, l'attenzione sarà posta successivamente sul soggetto utente delle BCI, per comprendere in che modo l'interazione con queste può trasformare sia il concetto di uomo e di corpo umano, sia il concetto di individuo, con le conseguenze che ne derivano sotto il profilo della responsabilità. A questo proposito, dunque, si è preso come punto di riferimento il confronto tra varie concezioni culturali e antropologiche. Lungi dal ripercorrere esaustivamente ogni piega di tale dibattito ormai consolidato, è comunque utile assumerlo come sfondo sul quale si inserisce la questione.

A seconda del modo di qualificare l'individuo utente di BCI, infatti, cambia anche il modo di rapportarsi a questi nell'utilizzo dei mezzi a disposizione per giungere a formulare un effettivo giudizio di responsabilità.

Prima di intraprendere il nostro percorso, tuttavia, è necessario compiere un passo indietro e andare all'origine della questione del *Responsibility Gap*, ripercorrendo la storia che ha portato all'introduzione delle BCI nella realtà di tutti i giorni.

Il punto di partenza, tuttavia, non è la realtà stessa, ma un sogno.

Capitolo 1

Da sogno a realtà

1.1 Il sogno del rabbino

Secondo la leggenda¹, nella soffitta della Staronová synagoga, la sinagoga Vecchia-Nuova, situata nel cuore del ghetto ebraico di Praga, riposa, in un sonno eterno, coperta da vecchi mantelli per la preghiera e resti di libri, una creatura. È il Golem, essere antropomorfo creato con una zolla di argilla del fiume Moldava. Artefice di tale prodigio è il leggendario rabbino Judah Loew ben Bezalel, il quale, in una notte del 1580, ebbe un sogno.

Guida e punto di riferimento dell'intero ghetto ebraico praghese, proprio a lui la comunità ebraica si rivolse per difendersi dalle sempre più violente accuse rivolte agli Ebrei da parte di delatori animati da sentimenti anti-semiti. “Rabbi” Loew non si sottrasse alla sua responsabilità. Rivolse, infatti, all’Altissimo una richiesta di aiuto. La risposta dell’Eterno non si fece attendere: in quella faticosa notte del 1580, Egli apparve in sogno al rabbino.

Le parole furono poche, concise, e, secondo i racconti, ordinate alfabeticamente: “Crea un Golem di argilla e annienta la malvagia canaglia divoratrice di Ebrei”².

A Loew fu chiaro il compito. Insieme al genero Jizchak Ben Simson e al suo allievo Jakob ben Chajim Sasson si diresse sulle rive del fiume Moldava. A mezzanotte, alla luce di fiaccole e accompagnati da specifici salmi, creò il Golem plasmandolo con l’argilla, seguendo il comando divino.

Proprio in quella notte di 442 anni fa, sulla riva di un fiume dell’Europa Centrale, si affaccia per la prima volta, seppure in un contesto leggendario ed

¹Salfellner, *Il golem di Praga: Leggende ebraiche dal ghetto*, 2019.

²“*Ata Bra Golem Dewuk Hachomer W’tigzar Zedim Chewel Torfe Jisrael.*” *ivi*, pag. 45.

esoterico, una Intelligenza prodotta dall'uomo. Per la creazione del Golem fu necessario l'incontro dei quattro elementi: il fuoco, evocato da Ben Simson, l'acqua, evocata da Ben Chajim Sasson e l'aria, simboleggiata da Rabbi Loew. Il quarto elemento fu offerto dal fiume: era la terra, l'argilla.

Con essa i tre uomini modellarono un uomo: un essere completamente formato, benchè privo della vita. Per donargli quest'ultima compirono, a turno, sette giri intorno al corpo. Il fuoco evocato da Ben Simson colorò il corpo di rosso fuoco. Con le evocazioni del secondo dal corpo uscì acqua: comparvero capelli e unghie su mani e dita. Rabbi Loew compì a sua volta i sette giri.

Infine, piegandosi in direzione dei quattro punti cardinali, pronunciarono parole ben precise, con le quali conclusero l'opera di imitazione della natura, ponendosi al pari di Dio:

“Allora il Signore Dio plasmò l'uomo con polvere del suolo e soffiò nelle sue narici un alito di vita e l'uomo divenne un essere vivente”³.

Ecco dunque che con l'aria, lo spirito, il quarto elemento, la terra, divenne vivo. Da ultimo, il rabbino scrisse sulla fronte del Golem la parola *emèt*, “verità”. L'opera fu così compiuta. L'essere fu creato. Una intelligenza “artificiale”.

Grazie alla creazione del Golem, la comunità ebraica fu protetta e servita per lungo tempo. Il Rabbino, infatti, subito dopo aver instillato la vita nel *monstrum*, gli disse:

“Devi sapere che ti abbiamo formato da una zolla di terra. Il tuo compito sarà quello di difendere gli Ebrei dalle persecuzioni. Ti chiamerai Josef e dormirai nella stanza del Rabbinato. Tu, Josef, dovrai eseguire i miei ordini, in qualsiasi momento e in qualsiasi luogo io ti invierò, nel fuoco e nell'acqua, dovrai ubbidire anche se ti dovessi ordinare di gettarti dal tetto, anche se ti inviassi negli abissi”⁴.

Il Golem dunque, anche se intelligente, era un mero strumento nelle mani del Rabbino. Una macchina antropomorfa, al servizio del padrone.

Un giorno, tuttavia, il Golem cominciò a violare il primo comando fondamentale impartitogli dal rabbino. Venne meno al comando di difendere gli Ebrei. In preda alla furia, entrò in una spirale di violenza, giungendo,

³Genesi 2,7-22.

⁴Salfellner, *Il golem di Praga: Leggende ebraiche dal ghetto*, 2019, pag. 46.

secondo alcuni⁵, addirittura a uccidere. Tale condotta, posta in essere in contrasto con i comandamenti del Rabbino, segnò la fine del Golem. Rabbi Loew, infatti, ancora una volta non venne meno alla sua responsabilità. Raggiungendo la creatura, le comandò di fermarsi e di seguirlo. Questa ubbidì.

Nella soffitta della Sinagoga, il rabbino, consapevole della necessità di porre fine alla minaccia, decise di privare l'essere della vita. Chiamò il genero e l'allievo. Seguendo il medesimo rituale che aveva segnato l'inizio di tutto, lo attuarono in senso inverso. Le stesse parole del Libro della Creazione furono lette al contrario. Il rabbino cancellò dalla parola *èmet*, scritta in fronte al Golem, la lettera *alef*, trasformando dunque la “verità” in *met*, “morte”. Il Golem, dunque, si ritrasformò in una sagoma d'argilla. I creatori lo coprirono con vesti e vecchi libri, relegando la loro stessa creatura a un sonno eterno. Judah Loew ben Bezalel, quasi come monito per i posteri, fece distruggere le scale che portavano alla soffitta. Era, secondo la tradizione, il 1593⁶.

Oltre quattro secoli dopo la conclusione della vicenda, la sinagoga Vecchia-Nuova è ancora attiva. Ora il Golem è divenuto una semplice statuetta venduta a turisti e appassionati di racconti del ghetto ebraico. La stessa soffitta dopo molti anni è stata svuotata e ristrutturata, lasciando al posto del Golem un vuoto che viene colmato solo dai racconti che ancora si condividono. Qui la leggenda, dunque, muore.

Se si presta attenzione, tuttavia, si possono scorgere ancor oggi le orme del Golem. Escono dalla soffitta della sinagoga e superano il XVI secolo. Se continuiamo a seguire le orme, queste attraversano l'Europa per condurre a uno specifico anno e a uno specifico luogo. La creatura non nasce più per il sogno di un rabbino, su una spiaggia di un fiume.

Il *monstrum* ha cambiato forma.

1.2 I passi del Golem: verso le Brain-Computer Interfaces

Luglio 1924. All'università di Jena il dottor Hans Berger, grazie a uno strumento da lui stesso chiamato Elettroencefalogramma, inventato *ad hoc* per osservare e misurare l'attività elettrica nel cervello umano, riuscì a captare e scoprire le onde alfa e beta, delle oscillazioni neurali nella gamma di

⁵ *Il Golem: un uomo di fango nel ghetto di Praga*, <https://www.storicang.it/a/il-golem-uomo-di-fango-nel-ghetto-di-praga-15184>, Accessed: 2021-11-04.

⁶ Salfellner, *Il golem di Praga: Leggende ebraiche dal ghetto*, 2019, pag. 46.

frequenza 8-12 Hz. Si registrò così la prima elettroencefalografia (EEG) umana. Cinque anni dopo, pubblicò l'esperimento in un articolo⁷ apparso nell'*Archive für Psychiatrie und Nervenkrankheiten*. Nonostante il suo contenuto rivoluzionario, la ricerca di Hans Berger fu accolta dalla comunità scientifica con “incredulità e scetticismo schiacciante”⁸. I fatti ben presto mostrarono la portata della ricerca di Berger.

Il 12 maggio 1934, il fisiologo Edgar Douglas Adrian, già premio Nobel per la medicina nel 1932, e il biochimico Brian Matthews, in una riunione della *Physiological Society* a Cambridge, confermarono i risultati della ricerca di Hans Berger, dando a quest'ultimo pieno credito per quanto scoperto⁹.

Il primo passo per la connessione uomo-macchina era stato compiuto. Il Golem cominciò ad avvicinarsi al rabbino.

Nel 1973, Paul C. Lauterbur, chimico statunitense, pubblicò sulla rivista *Nature* un articolo¹⁰ nel quale erano mostrate immagini di tubi.

Ciò che fu innovativo non fu tanto il contenuto di tali immagini, bensì la tecnica impiegata: per la prima volta vennero utilizzati gradienti di un campo magnetico, applicati unitamente a una tecnica di retroproiezione, per ottenere immagini di corpi fisici. La straordinarietà del risultato fu tale che a Lauterbur fu assegnato il premio Nobel per la Medicina nel 2003¹¹ insieme a Peter Mansfield. La tecnica fu denominata Risonanza Magnetica (RM).

Non si dovette attendere molto prima che si mappasse attraverso RM lo stesso organo che aveva ideato tale tecnica: il cervello umano.

Nel 1982, alla General Electrics, nello stato di New York, il medico John Schenck si sottopose volontariamente alla prima risonanza magnetica del cervello. Schenck aveva intuito che il campo magnetico prodotto dalle macchine per la RM poteva interagire con le molecole d'acqua presenti all'interno dei vari organi corporei e far loro emettere un segnale radio in grado di viaggiare fuori dal corpo. La digitalizzazione e l'applicazione di algoritmi poteva portare alla ricostruzione di un'immagine degli organi interni¹².

⁷Berger, “Über das elektroencephalogramm des menschen”, 1929.

⁸Berger e Wälchli, *Hans Berger*, 1956.

⁹Stone e Hughes, “Early history of electroencephalography and establishment of the American Clinical Neurophysiology Society”, 2013.

¹⁰Lauterbur, “Image formation by induced local interactions: examples employing nuclear magnetic resonance”, 1973.

¹¹L'assegnazione fu così motivata dall'Accademia: “for their discoveries concerning magnetic resonance imaging”. Vd Leach, “Nobel prize in physiology or medicine 2003 awarded to Paul Lauterbur and Peter Mansfield for discoveries concerning magnetic resonance imaging”, 2004.

¹²Ienca, “Intelligenza²: per un'unione di intelligenza naturale e artificiale”, 2019.

L'esperimento fu un completo successo. Il cervello fu "fotografato" per la prima volta. La mappatura anatomica del cervello umano fu solo l'inizio.

Il passo successivo sarebbe stato la ricostruzione della funzione dello stesso: l'attività cerebrale.

Già negli anni Ottanta del XIX secolo, il medico e fisiologo italiano Angelo Mosso elaborò la prima tecnica di *neuroimaging* funzionale, denominandola "Bilancia della circolazione umana"¹³. Lo straordinario strumento confermò le tesi di Mosso: il flusso di sangue all'interno del cervello variava a seconda dell'attività cerebrale, a seconda, cioè, della funzione svolta dall'organo. Là dove i soggetti sottoposti all'esperimento si impegnavano in determinate azioni cognitive, si registrava un aumento del flusso sanguigno cerebrale verso le aree di maggiore attività neuronale¹⁴.

Circa un secolo dopo, nel 1990, Seiji Ogawa, ricercatore presso l'Università di Tokyo, intuì e dimostrò che è possibile sfruttare le proprietà magnetiche dell'emoglobina nel flusso sanguigno all'interno del cervello¹⁵. Tale metodo è oggi conosciuto come BOLD (*Blood Oxygen Level-Dependent*)¹⁶. Conducendo dunque una Risonanza Magnetica, Ogawa dimostrò l'idoneità della tecnica elaborata da Lauterbur per ottenere una mappatura non solo anatomica, ma anche funzionale del cervello umano.

I risultati a cui pervenne Ogawa, unitamente alla ricerca di John Belliveau¹⁷, permisero di derivare dalla RM una risonanza magnetica funzionale. Si tratta della fMRI, la *Functional Magnetic Resonance Imaging*, che permette di misurare l'attività elettrica del cervello utilizzando risposte emodinamiche come marcatori indiretti¹⁸.

Magnetismo ed elettricità rappresentano i fenomeni fisici alla base di tali tecniche. Nello stesso periodo, un'altra tecnica analoga si sarebbe rivelata fondamentale per lo sviluppo di dispositivi di connessione uomo-macchina: la fNIRS *Functional Near-Infrared Spectroscopy*¹⁹, Spettroscopia funzionale

¹³Sandrone, Bacigaluppi, Galloni e Martino, "Angelo Mosso (1846–1910)", 2012.

¹⁴Per una descrizione completa dell'esperimento, si veda Sandrone, Bacigaluppi, Galloni, Cappa et al., "Weighing brain activity with the balance: Angelo Mosso's original manuscripts come to light", 2014.

¹⁵Ogawa et al., "Brain magnetic resonance imaging with contrast dependent on blood oxygenation", 1990.

¹⁶L'ideazione del metodo BOLD valse a Seiji Ogawa l'assegnazione della medaglia d'oro, nel 1995, da parte della International Society of Magnetic Resonance.

¹⁷Ricercatore presso il Massachusetts General Institute, decise di utilizzare una risonanza magnetica ad alta risoluzione per studiare l'attività cerebrale, confrontando poi i volumi di sangue nelle aree del cervello stimolate con quelli delle aree non stimolate. Si veda Belliveau, *Functional NMR imaging of the brain*, 1990.

¹⁸Ienca, "Intelligenza²: per un'unione di intelligenza naturale e artificiale", 2019.

¹⁹Jobsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen

nel vicino infrarosso. La fNIRS impiega luce diffusa nella banda spettrale del vicino infrarosso per indagare l'attività emodinamica della corteccia cerebrale e la conseguente capacità funzionale ad essa associata²⁰.

La porta dell'edificio della scatola cranica, che dalla nascita dell'umanità separava ciò che è interno da ciò che è esterno, è stata, alla fine, aperta.

Il Golem può fare il suo ingresso.

sufficiency and circulatory parameters", 1977.

²⁰*ivi*. Simile alla fMRI in quanto tecnica non invasiva di neuroimaging funzionale, presenta comunque dei notevoli vantaggi in termini di *performance* e di precisione nell'indagine e, soprattutto, una maggiore flessibilità per l'utilizzo di dispositivi portabili. Si veda anche Pinti et al., "The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience", 2020.

Capitolo 2

Il rabbino si fa Golem: le Brain-Computer Interfaces

Giugno 1973. Tre mesi dopo la pubblicazione su *Nature* delle straordinarie immagini dei tubi da parte di Paul Lauterbur, Jacques Vidal, professore di Informatica presso l'University of California a Los Angeles, pubblicò nell'*Annual review of Biophysics and Bioengineering* un articolo¹ nel quale si coniò l'espressione *Brain-Computer Interfaces* (BCI), gettando le fondamenta per il relativo, emergente ambito di ricerca².

Vidal nel suo elaborato teorizzò un prototipo di un canale di comunicazione diretto tra cervello e computer: uno strumento di connessione che avrebbe permesso di controllare dispositivi esterni tramite segnali elettrici cerebrali captati tramite EEG.

Quattro anni dopo, lo stesso Vidal condusse un esperimento nel quale un soggetto riuscì a muovere un cursore su schermo tramite EEG non invasiva³. Dopo il 1977, gli esperimenti volti a stabilire una connessione cervello-macchina si moltiplicarono.

Nel 1998, i neurologi Philip Kennedy e Roy Bakay installarono una interfaccia invasiva su un essere umano. La persona in questione era un paziente che soffriva della c.d. sindrome di *Locked-In*⁴. Il tentativo ebbe successo.

¹Jacques J Vidal, "Toward direct brain-computer communication", 1973.

²*Cerebral connections: UCLA engineers tap into rich legacy of brain-computer interface technology, howpublished = <http://samueli.ucla.edu/brain-computer-interface/>, note = Accessed: 2021-11-08.*

³L'esperimento è descritto in Jacques J. Vidal, "Real-time detection of brain events in EEG", 1977.

⁴Si tratta di un raro disturbo neurologico nel quale il soggetto è affetto da una paralisi totale del Sistema nervoso somatico, con eccezione dei movimenti codificati degli occhi.

L'interfaccia riuscì a simulare i movimenti desiderati dal paziente. Da allora, l'applicazione delle BCI caratterizza molteplici ambiti della società, non più esclusivamente quello medico-terapeutico.

2.1 Classificazione delle BCI

Procedere a individuare una definizione e classificazione univoca delle interfacce uomo-macchina⁵ sotto il profilo dei dispositivi risulterebbe inutile, se non problematica. L'eterogeneità degli ambiti di applicazione e la numerosità dei dispositivi impiegabili vanificherebbero ogni tentativo di ricondurre il fenomeno a una definizione astratta e rigorosa.

L'esperienza delle BCI, più che in termini di dispositivi, si può inquadrare più propriamente sotto il profilo delle *dinamiche* di connessione e delle funzioni espletate.

In primo luogo, per quanto riguarda la pervasività degli allacciamenti con il cervello dell'utente, si è soliti distinguere tra interfacce invasive e non invasive.

Le BCI invasive comportano la registrazione di segnali cerebrali attraverso l'innesto di elettrodi all'interno del cervello. I primi esperimenti condotti negli anni '70 negli Stati Uniti su primati⁶ hanno portato a scoprire, infatti, una relazione tra gli impulsi elettrici nel cervello di tali scimmie e la direzione verso la quale queste muovevano i loro arti superiori⁷. Attualmente, cinquanta anni dopo il compimento di questi primi passi, le ricerche sulle interfacce invasive sono volte ad assicurare nuove funzionalità a soggetti affetti da disabilità⁸.

La sindrome di *Locked-In*, tuttavia, lascia impregiudicata l'attività cerebrale, con la conseguenza che quest'ultima può essere rilevata tramite EEG, che risulta simile a quella di un paziente sano. Si veda *Locked-In syndrome*, <https://rarediseases.info.nih.gov/diseases/6919/locked-in-syndrome/>, Accessed: 2021-11-10.

⁵Ricordiamo che accanto all'interfaccia *brain-computer* si conoscono anche le c.d. *Brain to brain Interfaces*. Sono, queste, delle interfacce applicate in via sperimentale su topi, che permettono di instaurare una connessione diretta tra i segnali cerebrali dei cervelli di due roditori, per permettere una cooperazione reciproca in vista di una ricompensa comune. La trattazione, tuttavia, verterà sulle BCI per le ricadute etico e soprattutto giuridiche della loro applicazione. In ogni caso, per una disamina delle BBI si veda Pais-Vieira et al., "A brain-to-brain interface for real-time sharing of sensorimotor information", 2013.

⁶Schmidt, Bak e McIntosh, "Long-term chronic recording from cortical neurons", 1976.

⁷Georgopoulos et al., "Mental rotation of the neuronal population vector", 1989.

⁸MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017, pag 51. L'autore rileva comunque che l'impianto di tali dispositivi può portare a un accumulo del tessuto cicatriziale, che potrebbe diminuire la qualità dei segnali cerebrali o annullarli. Il cervello, infatti, reagirebbe in maniera sempre

Le interfacce non invasive, invece, prevedono l'analisi dell'attività cerebrale attraverso le tecniche di *neuroimaging* tramite EEG. In questo caso, la non-invasività è data dal fatto che gli impulsi cerebrali vengono captati attraverso elettrodi installati all'esterno del cranio: non si compie alcun intervento chirurgico e di conseguenza si riducono anche i rischi di danni e infezioni al cervello. Si rende possibile, ad esempio, analizzare l'attività cerebrale di soggetti affetti da paralisi⁹.

Si individua anche una terza categoria: le interfacce parzialmente invasive. Se, da un lato, non prevedono l'installazione di elettrodi direttamente sul cervello, ma a livello della scatola cranica, dall'altro, comunque, richiedono un intervento chirurgico, con annessi rischi¹⁰.

Sia le interfacce invasive che le interfacce non invasive si fondano su una interazione diretta tra il cervello dell'utente e un dispositivo informatico.

L'interazione è scomponibile in quattro fasi¹¹.

La prima è la fase di *input*. Il cervello dell'utente, in risposta a uno stimolo, genera un'attività cerebrale specifica. Una delle tecniche più conosciute è la c.d. *DBS*, la *Deep Brain Stimulation*. Attraverso l'impianto di elettrodi in una determinata zona del cervello, è possibile controllare artificialmente l'attività alla quale è adibita. Varie sono le possibilità applicative a fini terapeutici. Si è rivelata particolarmente efficace, infatti, tanto per la riduzione del tremore e della rigidità di soggetti affetti dalla sindrome di Parkinson, quanto per il controllo di epilessia, emicrania, fino ad arrivare addirittura a recare notevoli benefici per la cura dell'Alzheimer e per la lotta all'obesità¹². Alcuni impieghi di tale tecnica comportano, tuttavia, delle complesse ricadute etico-giuridiche¹³. È divenuto possibile, infatti, l'impiego della *DBS* nei confronti di colpevoli di crimini violenti, ad esempio a sfondo sessuale¹⁴, per inibire l'attività cerebrale proveniente dall'amigdala,

più forte alla presenza di un dispositivo estraneo al suo interno.

⁹Birbaumer e Cohen, "Brain-computer interfaces: communication and restoration of movement in paralysis", 2007.

¹⁰Per gli esperimenti condotti e il loro sviluppo, si veda Cahn e Polich, "Meditation states and traits: EEG, ERP, and neuroimaging studies.", 2006.

¹¹Riprendo, qui, la suddivisione delineata in Ienca, "Intelligenza²: per un'unione di intelligenza naturale e artificiale", 2019, pagg. 111-112.

¹²Kleiner-Fisman et al., "Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes", 2006.

¹³Per ripercorrere il dibattito e le problematiche connesse all'impiego a fini terapeutici della *DBS*, si veda il cap. 4 del presente lavoro, in particolare il paragrafo "Identità e continuità psicologica."

¹⁴In riferimento alla c.d. *neuropunishment*, per un'analisi più approfondita, si veda Buyx e Birks, "Neuroscience and social problems: The case of neuropunishment", 2018.

considerata come l'origine degli impulsi sessuali¹⁵.

La seconda fase di questa interazione consiste nella misurazione e registrazione dell'attività cerebrale. Qui, l'attività cerebrale è rilevata e misurata durante il compimento di un processo cognitivo. La misurazione registrata può, poi, essere implementata a seconda del tipo di BCI impiegata, in particolare con EEG, spettroscopia funzionale nel vicino infrarosso (*fNIRS*) o altre invasive come l'elettrocorticografia (*ECOG*). I dati ottenuti nella seconda fase devono essere codificati e conseguentemente classificati da parte.

In questa successiva terza fase, questi sono elaborati per isolare i dati ricevuti dall'attività cerebrale di fondo che non è diretta a uno specifico compito, quale quella derivante dal movimento dato dallo sbattere delle palpebre o coinvolta nella percezione dei colori. La fase di pre-processamento e di classificazione permette di trasformare il segnale in dati computabili e traducibili in comandi per i dispositivi connessi. Il protagonista di questo *step* è un algoritmo, che rappresenta il “cuore” della BCI, in quanto permette tutta l'attività di decodificazione dei dati per trasformarli in *outputs* idonei a controllare tali dispositivi esterni¹⁶.

Nella quarta e ultima fase di interazione, infatti, i segnali sono tradotti in *output*: si esegue l'azione desiderata attraverso il controllo dei *devices* interfacciati dalla BCI¹⁷. A ogni conclusione della quarta fase di interazione corrisponde la fine di un ciclo, o *loop*. Una volta completato ogni ciclo, il soggetto percepisce il *feedback* derivante dal ciclo precedente e il ciclo successivo ha inizio. L'utente, tuttavia, sulla base del *feedback* ricevuto, può rimodulare l'attività cerebrale secondo la propria preferenza, così da perfezionare il processo in modo da conseguire l'esito prescelto. Il risultato è la formazione di quello che si definisce un *closed loop system*, un circuito a ciclo chiuso¹⁸.

¹⁵Baird et al., “The amygdala and sexual drive: insights from temporal lobe epilepsy surgery”, 2004.

¹⁶Hassanien e Azar, “Brain-computer interfaces”, 2015, pag 98.

¹⁷Le applicazioni possono essere le più varie, comprendendo dispositivi software e hardware come sedie a rotelle, protesi, dispositivi di sensore.

¹⁸In alcuni casi, il *feedback* non necessariamente deriva dal segnale inviato dal dispositivo al cervello. Può, infatti, essere ricevuto anche attraverso un canale visivo. In questo caso, si ha un *open-loop system*, o circuito a ciclo aperto.

2.2 Ambiti di impiego delle interfacce. Le *Assistive technologies*

Per comprendere l’eterogeneità dei settori di applicazione delle *Brain-Computer Interfaces*, è possibile ricordare una dei più rilevanti ambiti di utilizzo delle interfacce: l’ambito medico-terapeutico. Tale utilizzo è particolarmente evidente nelle c.d. *Assistive technologies*.

Lo *U.S. Technology-Related Assistance for Individuals With Disabilities Act of 1988* assegna tale definizione a “*any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities*”¹⁹. In riferimento a tali tecnologie, le interfacce vengono in rilievo là dove vi sono utenti dotati di un cervello sano in grado di espletare le normali funzioni, ma che, tuttavia, sono affetti da paralisi in alcune parti del corpo²⁰. In questi casi, rimanendo intatte le funzionalità cerebrali, le interfacce sono in grado di allacciarsi alla rete neuronale e ottenere dati al fine di connettersi a un dispositivo che, successivamente, permette di svolgere alcune attività precluse all’arto paralizzato che, in condizioni normali, dovrebbe essere mosso per il suo compimento²¹.

Uno dei primi esempi dell’utilizzo delle interfacce neuronali in questa direzione è dato dal caso di Matthew Nagle. Nel 2004, Nagle, rimasto tetraplegico dopo essere stato accoltellato nel 2001 all’uscita di un locale, si sottopose come volontario a un esperimento clinico per l’innesto di *BrainGate*²², una speciale interfaccia neuronale che, inserita nel suo cervello a livello della regione della corteccia motoria responsabile del movimento della mano sinistra e del braccio, gli avrebbe permesso di controllare questi ultimi letteralmente con la mente, *bypassando* la paralisi. In particolare, il sistema prevedeva, e prevede tutt’ora, l’inserimento di un sensore impiantato nel cervello e un dispositivo esterno di decodificazione che si connette a una protesi. Più nello specifico, il sensore rileva l’attività dei neuroni in determinate aree

¹⁹Bryant e Seay, “Republication of The Technology-Related Assistance to Individuals with Disabilities Act: Relevance to Individuals with Learning Disabilities and Their Advocates”, 2020.

²⁰MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017. La casistica è varia. Si includono soggetti che hanno subito lesioni al midollo spinale, amputazioni o ictus. Si veda Brain, “Brain Waves Module 1: Neuroscience, society and policy”, 2011.

²¹Donoghue et al., “Development of neuromotor prostheses for humans”, 2004.

²²È questo un sistema di impianto neuronale originariamente costruito e progettato dalla Cyberkinetics, una azienda americana che operava in collaborazione con la *University of Utah*, ora acquisita dalla BrainGate Inc.

del cervello, le invia al *device* che, a sua volta, induce il movimento della protesi. I risultati sorprendenti dell'esperimento furono pubblicati nel 2005 sulla rivista *Nature*²³. Il soggetto non solo riusciva a muovere un braccio robotico e un cursore sullo schermo di un computer, ma anche ad accendere le luci e a interagire con una televisione.

Matthew Nagle e l'interfaccia entrarono, dunque, in un vero e proprio rapporto di simbiosi²⁴. Un rapporto che, grazie alla dinamica di *feedback* che caratterizza tali dispositivi, permette al soggetto di avere una risposta circa lo sviluppo dell'azione e del movimento, così da correggerla o mantenerla. Non solo. Anche l'interfaccia può "giovarsi" della risposta dell'utente, modificando la propria risposta a seconda della reazione del soggetto²⁵. Come nella più efficiente delle relazioni simbiotiche, detta *mutualismo*, dunque, entrambi possono cooperare per il conseguimento di un beneficio comune, in questo caso il compimento dell'azione²⁶.

L'utilizzo di tali strumenti quali *assistive technology* si rivede anche in un'altra vicenda particolarmente interessante. In particolare, riguarda un uomo di nome Neil Harbisson, che fin dalla nascita soffre di acromatopsia²⁷. L'impianto di una antenna *cyborg* ossointegrata nel suo cranio permette ad Harbisson di percepire e "ascoltare" i colori, che vengono trasformati in apposite vibrazioni acustiche²⁸. Non solo. L'antenna, infatti, permette di percepire colori invisibili all'occhio umano, come segnali infrarossi e ultravioletti²⁹, fino a giungere, addirittura, a disporre di una connessione internet che permette di trasformare in colori i segnali provenienti dal satellite. L'interazione uomo macchina, rabbino e golem, è tale in questo caso che Neil Harbisson è stato definito come il primo *cyborg*³⁰. Lo stesso Har-

²³Hochberg et al., "Neuronal ensemble control of prosthetic devices by a human with tetraplegia", 2006.

²⁴La nozione di simbiosi è stata al centro di un lungo dibattito nella letteratura. Attualmente, si adotta la c.d. definizione "De Bary", secondo la quale la simbiosi è "convivenza di organismi di specie diversa." Vd Douglas, *The symbiotic habit*, 2021.

²⁵Mak e J. R. Wolpaw, "Clinical applications of brain-computer interfaces: current state and future prospects", 2009.

²⁶C'è comunque da chiedersi se, o quando, tale rapporto si possa trasformare in un'altra delle dinamiche simbiotiche: il *parassitismo*. In questo caso, con una sempre più intensa interazione uomo-macchina, sapere chi sia l'ospite e chi il parassita può divenire sempre più problematico.

²⁷Consiste in una patologia dei coni della retina, trasmissibile ereditariamente, che comporta l'incapacità totale di distinguere i colori. Se parziale, è il comune daltonismo. Vd Alexander et al., "Restoration of cone vision in a mouse model of achromatopsia", 2007

²⁸Harbisson, "Painting by ear", 2008.

²⁹Ronchi, *eCulture: cultural content in the digital age*, 2009.

³⁰Jeffries, "Neil Harbisson: the world's first cyborg artist", 2014. La definizione di Cy-

bisson ha accolto tale identità, fondando, nel 2010, la *Cyborg Foundation*, con la finalità di promuovere progetti e ricerche volte a estendere e creare nuove forme di interazione tra corpo umano e tecnologia. L'antenna di cui è dotato è divenuta parte integrante della persona di Harbisson a tal punto che quest'ultimo ha istituito, nel 2017, la *Transpecies Society*³¹.

Applicazioni non terapeutiche delle interfacce

L'impiego delle *Brain Computer-Interfaces* oggi non è più limitato solo all'ambito medico-terapeutico. Al contrario, sempre più numerosi sono gli esempi di impianti di BCI su soggetti sani, senza perseguire alcuna finalità di cura o di riabilitazione.

Nel 2002, l'ingegnere Kevin Warwick, attualmente vice rettore presso la *Coventry University*, si sottopose volontariamente a una serie di esperimenti da lui stesso supervisionati che furono condotti nel contesto di quello che venne definito *Project Cyborg*. Lo scopo del progetto era quello di divenire, letteralmente, un cyborg³². In particolare, nella seconda fase³³ del progetto, allo stesso Warwick fu impiantata una interfaccia neuronale dotata di un sensore *BrainGate*, collegata a un dispositivo al quale erano connessi vari allacciamenti. L'interfaccia recepiva gli stimoli provenienti dal sistema nervoso tramite il collegamento con un vettore introdotto a livello del nervo mediano nel polso sinistro. I segnali di output si rivelarono sufficientemente forti e precisi da muovere un braccio robotico imitando i movimenti del

borg si fa risalire a due ricercatori, l'austriaco Manfred Clynes e l'americano Nathan Kline, che nel 1960 hanno definito "cyborg" un essere umano potenziato con organi artificiali per sopravvivere in contesti extraterrestri inospitali. Nonostante tali contesti extraterrestri siano ancora, forse ancora per poco, fantascienza, il potenziamento di essere umani con organi artificiali è diventato realtà. Cfr. Clynes e Kline, "Cyborgs and space", 1960.

³¹Sono emblematiche le parole degli stessi membri: "*we think that Transpecies Society members are tinkering with ambiguities about what is natural and what is artificial, incorporating these organs in order to reveal new environmental stimuli.*" vd López e Padrós, "Designing organs at the Transpecies Society: hybrid practices between cybernetics and artificial intelligence", 2019, pag. 3.

³²Warwick, Gasson, Hutt, Goodhew et al., "The application of implant technology for cybernetic systems", 2003. Lo si chiami Cyborg, lo si chiami Golem, il risultato non cambia. Creatore e creatura si confondono sempre di più.

³³La prima fase, svoltasi quattro anni prima, nel 1998, consisteva nell'inserire un piccolo trasmettitore sotto la superficie cutanea, utilizzato per accendere luci, controllare porte automatiche e in generale interagire con l'ambiente circostante. Lo scopo ultimo di questa prima fase era verificare qual era il livello massimo di interazione prima che il corpo rigettasse un dispositivo estraneo. Vd. Magazine, 8.02 (February 2000), 'Cyborg 1.0: Interview with Kevin Warwick'.

braccio di Warwick. Il progetto cyborg riuscì³⁴. Vi è una peculiarità da notare. Il professore Kevin Warwick e il suo braccio “naturale” si trovavano presso la *Columbia University* a New York. Il braccio robotico artificiale, si trovava, invece, in un laboratorio della *University of Reading*, nella città di Reading in Inghilterra.

Il legame creatore-creatura non è più dovuto a una formula magica pronunciata da un rabbino sulle sponde di un fiume, ma da una semplice connessione internet.

Neurogaming e vita quotidiana

Le interfacce neuronali sono ormai reperibili nel mercato e a disposizione di chiunque. In questo caso il carattere non invasivo di molti dei dispositivi permette un utilizzo privo delle problematiche riconnesse alle interfacce invasive. La loro applicazione non richiede particolari capacità e possono essere utilizzate e dismesse a piacimento dell'utente. Il termine utilizzato per definire tali tecnologia non invasiva e “portabile” è “neurotecnologia pervasiva”³⁵. Attraverso le registrazioni tramite *EEG* è possibile monitorare l'attività cerebrale per vari scopi, come il neuromonitoraggio, l'esercizio neurocognitivo e il controllo di dispositivi con interfacce neuronali³⁶.

Tra i principali utilizzi di questi dispositivi vi è il supporto alle attività quotidiane, nonchè, il *gaming*, o, meglio, il *neurogaming*, nel quale è possibile giocare senza il supporto di un tradizionale controller, ma semplicemente attraverso le onde cerebrali.

Nel 2013, ad esempio, Adam Gazzaley e il suo team di ricercatori della *University of California* a San Francisco hanno sviluppato *Neuroracer*, un gioco di corsa nel quale le macchine sono controllate direttamente dal cervello del giocatore, attraverso EEG³⁷. *Neuroracer* è stato solo il primo passo.

L'azienda americana Emotiv ha trasformato nel suo prodotto di punta un casco, connesso a un computer, dotato di quattordici elettrodi che si adagiano in determinate zone del cuoio capelluto dell'utente. Questo, sulla base dell'attività cerebrale, può modulare la difficoltà di un videogioco a seconda

³⁴Warwick, Gasson, Hutt e Goodhew, “An attempt to extend human sensory capabilities by means of implant technology”, 2005.

³⁵Ienca, “Intelligenza²: per un'unione di intelligenza naturale e artificiale”, 2019, pag. 120.

³⁶*ibidem*

³⁷Anguera et al., “Video game training enhances cognitive control in older adults”, 2013.

della stanchezza, del livello di eccitazione e di attenzione del soggetto, in una relazione adattiva ed evolutiva, “leggendo” la sua mente³⁸.

L’industria dei videogiochi rappresenta solo la punta dell’iceberg. Le applicazioni commerciali di tale tecnologia con finalità di supporto all’attività di ogni giorno sono sempre più numerose, in particolare in riferimento al settore delle comunicazioni mobili.

L’azienda Apple, ad esempio, dopo una fase di sperimentazione ha aperto alla possibilità di connettere al proprio *iPhone* dei gadget che prevedono l’utilizzo di *EEG*. La sperimentazione verte sulla possibilità di utilizzare *XWave*, creato dall’azienda americana Neurosky. In questo caso l’utente, utilizzando delle apposite cuffie dotate di un cavo audio riesce a controllare le onde cerebrali del proprio cervello sullo schermo del proprio *smartphone*, nel quale è riprodotta la sua attività cerebrale³⁹. Il dispositivo ha anche un costo particolarmente modesto. In base a quanto riportato nello *store* online del sito della Neurosky, infatti, il prezzo si aggira intorno ai cento dollari.

L’esperienza delle BCI, dunque, caratterizza trasversalmente la vita quotidiana di soggetti affetti da disabilità e soggetti sani, per scopi terapeutici e scopi commerciali. Si sta creando una galassia di dispositivi estremamente complessa e variegata.

Come è stato evidenziato⁴⁰, infatti, in tale ambito convergono almeno tre macro aree tecnico-scientifiche.

La prima, com’è intuibile, è rappresentata dalle neuroscienze. Una scienza multidisciplinare avente ad oggetto lo studio di neuroni, cellule gliali e rete neurali: in altre parole, il sistema nervoso⁴¹.

La seconda importante macro area tecnologica che converge è l’area del c.d. *Internet of things*, ossia la dinamica sempre più forte di interconnessione tra dispositivi, unitamente al balzo in avanti dell’Informatica e delle Tecnologie dell’informazione, o *ICT*⁴², nonché al progresso delle ricerche nell’ambito della intelligenza artificiale e degli algoritmi ad apprendimento automatico, che si rivela fondamentale nella sua applicazione nelle BCI.

³⁸MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017.

³⁹Blondet et al., “A wearable real-time BCI system based on mobile cloud computing”, 2013.

⁴⁰Fuselli, “Brain-Computer Interface e soggettività agente. Considerazioni etico-giuridiche.”, 2021, in Moro, *Etica, diritto e tecnologia. Percorsi dell’informatica giuridica contemporanea*, 2021, pag. 153.

⁴¹Shulman, *Brain imaging: What it can (and cannot) tell us about consciousness*, 2013.

⁴²Un imprescindibile punto di riferimento nell’analisi delle tecnologie dell’informazione è Floridi, *La quarta rivoluzione: come l’infosfera sta trasformando il mondo*, 2017, in particolare il secondo capitolo: “Spazio:infosfera”.

Infine, la terza grande area che interviene è quella della robotica che, nei suoi sviluppi più recenti, fa sì che le interfacce diventino sempre più versatili attraverso un uso relativamente facile da parte del suo utilizzatore. Alla fine, la ragione ultima dello sviluppo e della diffusione di queste neurotecnologie si deve a tre loro caratteristiche: “*small, cheap and out of control*”⁴³. Piccole, economiche e, soprattutto, fuori controllo. Proprio in riferimento a quest’ultima qualità si individuano alcune rilevanti problematiche.

È stata notata, infatti⁴⁴, una doppia vulnerabilità regolativa.

Da un lato tali dispositivi, dal momento che possono costituire prodotti commerciali destinati all’immissione nel mercato, non necessitano di ricetta medica e non sono soggetti alla normativa applicabile ai dispositivi medici.

Dall’altra, i dati raccolti ed elaborati dalla tecnologia in questione non appartengono ad alcuna categoria di dati in relazione alla quale è accordata tutela dalla legge. Di conseguenza, le aziende produttrici delle interfacce non sono soggette alle rigide regole previste a tutela della raccolta e della conservazione, ad esempio, dei dati genetici⁴⁵.

Il suo impatto, tuttavia, è ben più radicale. Non è più solo limitato a una semplice evoluzione del mercato o della scienza medica. Le ricadute possono essere molto più profonde e ampie. Per capire la portata delle problematiche è necessario, ancora una volta, seguire i passi del Golem.

⁴³Ienca, “Intelligenza²: per un’unione di intelligenza naturale e artificiale”, 2019, pag. 123.

⁴⁴Ienca, Haselager e Emanuel, “Brain leaks and consumer neurotechnology”, 2018.

⁴⁵L’articolo 4.1, n. 13 del Regolamento UE 2016/679 definisce i dati genetici “i dati personali relativi alle caratteristiche genetiche ereditarie o acquisite di una persona fisica che forniscono informazioni univoche sulla fisiologia o sulla salute di detta persona fisica, e che risultano in particolare dall’analisi di un campione biologico della persona fisica in questione” Nel considerando 34 dello stesso regolamento si specifica che “è opportuno che per dati genetici si intendano i dati personali relativi alle caratteristiche genetiche, ereditarie o acquisite, di una persona fisica, che risultino dall’analisi di un campione biologico della persona fisica in questione, in particolare dall’analisi dei cromosomi, dell’acido desossiribonucleico (DNA) o dell’acido ribonucleico (RNA), ovvero dall’analisi di un altro elemento che consenta di ottenere informazioni equivalenti”. Il successivo articolo 9.1 individua i dati genetici come “categorie particolari di dati personali”, vietandone il trattamento. Infine, il D.lgs. 101/2018, all’art. 22 comma 2 riconduce tali dati alla categoria dei “dati sensibili” e alla relativa tutela. Per le problematiche relative alla privacy, si veda il cap. 5, in particolare il paragrafo “Privacy neurale”.

Capitolo 3

Il Responsibility Gap

3.1 Introduzione

Il Golem è stato creato. La creatura muove ora i primi passi ed entra nel ghetto. Gli abitanti sono interdetti e incerti su chi o cosa stanno guardando. È stato loro detto che quest'uomo dall'espressione impassibile e di pietra è uno straniero, parente del Rabbino. È giunto nel paese per aiutare e difendere gli abitanti da tutti coloro che vogliono fare loro del male. Ed effettivamente, il Golem ubbidisce ai comandi impartiti e svolge con fredda precisione la sua missione. A un certo punto, tuttavia, questo stesso guardiano cambia. Si rivolge contro gli stessi indifesi che doveva proteggere. Con furia e violenza distrugge, assale e uccide. Dopo la devastazione e i danni subiti, gli abitanti chiedono giustizia. Nel *Beth din*¹ durante l'esame dei testimoni e gli interrogatori, al giudice cominciano a sorgere molte domande. Così come tutti gli abitanti del ghetto, infatti, si chiede *come* e *perché*. Non capiscono perché costui ora attacca le stesse persone che si rivolgevano a lui per ottenere protezione e aiuto. Non capiscono perché il Golem, dopo l'intervento del rabbino, si sia immediatamente calmato e abbia placato la sua furia. Il giudice, tuttavia, intuisce che la questione è molto più complicata. Non riesce ancora a cogliere con certezza e precisione in che modo, ma capisce che il rabbino è indissolubilmente legato all'essere. L'unico modo per fare chiarezza e giustizia e soddisfare le pretese degli abitanti danneggiati e feriti è interrogare il rabbino. Chiedergli chi è il Golem, quale sia il loro legame, se è al suo comando. Il religioso risponde raccontandogli del suo sogno e della creazione della creatura sulle rive del fiume Moldava. A questo punto al giudice si pone un fondamentale problema. Chi è respon-

¹Il *Beth din* è il del tribunale rabbinico nel sistema giudiziario ebraico.

sabile? Il creatore o la creatura che ha deviato dagli ordini del padrone? Come distinguere i due? Il dubbio non appartiene solo alla leggenda. Non è finzione, immaginazione o fantascienza. È realtà. Sempre più forti sono sentite queste domande, e sempre più necessarie le risposte. Prima di porsi quesiti e darsi risposte, tuttavia, bisogna inquadrare il problema.

3.2 Responsibility Gap

Nel 2004, nella rivista *Ethics in information technology*, sesta edizione, venne pubblicato un articolo da parte del ricercatore Andreas Matthias.

Nell'elaborato, l'Autore pone l'attenzione sulle problematiche connesse all'utilizzo di una determinata tecnologia basata su algoritmi c.d. ad apprendimento automatico, o *machine learning algorithms*. Attraverso la raccolta di dati campioni, i c.d. *training data*, detti algoritmi riescono a fare previsioni e a prendere decisioni, migliorando e implementando la propria *performance* senza ricevere istruzioni implicite².

Proprio per questa loro caratteristica Matthias rileva come si crea una nuova situazione nella quale il costruttore e programmatore non può prevedere il comportamento futuro di tali macchine. Non può, cioè, essere considerato legalmente o moralmente responsabile. Questo perchè, in ultima analisi, “*nobody has enough control over the machine’s actions to be able to assume the responsibility for them*”³.

Ecco allora che la società si trova innanzi a quello che è definito, con una fortunata espressione, il *Responsibility Gap*, una distanza tra la nuova realtà degli algoritmi e il tradizionale paradigma legale per attribuire la responsabilità.

È una distanza tanto profonda e crescente, che lo stesso Matthias riconosce l'esistenza di un *aut-aut*. O la società desiste dall'utilizzare questa tecnologia, scelta che comunque è riconosciuta come impossibile, o si deve affrontare il *gap*. La scelta di voltarsi dall'altra parte e ignorare la questione o di sottovalutare il problema pone il pericolo di mettere in crisi lo stesso *moral framework*, il sistema morale della società e, addirittura, le fondamenta stesse del concetto di responsabilità legale⁴.

²Bishop e Nasrabadi, *Pattern recognition and machine learning*, 2006.

³Matthias, “The responsibility gap: Ascribing responsibility for the actions of learning automata”, 2004, pag. 177.

⁴*ivi*.

3.3 *Please, mind the gap.* Responsabilità e interfacce

L'espressione *Responsibility Gap* nella letteratura è diventata un termine “ombrello”, indicativa di una problematica che trascende l'ambito originario per esser riferita anche, tra le altre, all'esperienza delle interfacce neuronali. Non è una operazione casuale. Anche in rapporto a questa neurotecnologia possono sorgere dei *gaps* di responsabilità.

I dispositivi neurali, infatti, assistono il soggetto nel compimento di un'azione finalizzata a un determinato obiettivo⁵. È, questa, una relazione nella quale tanto l'utente quanto l'interfaccia si influenzano reciprocamente per generare e indirizzare il movimento. In questo caso, può risultare difficile capire chi delle due parti del rapporto ha esercitato effettivamente una maggiore influenza sull'altro e, di conseguenza, chi ha avuto il controllo sull'azione.

Per comprendere le difficoltà che sorgono, è possibile ricordare un'utile distinzione tra BCI attive, passive e reattive⁶.

Nelle BCI attive⁷ l'algoritmo alla base del funzionamento delle interfacce rileva determinati segnali dell'attività cerebrale che è prodotta nel momento in cui il soggetto compie intenzionalmente un determinato compito mentale. L'utente, cioè, *immagina* di compiere il movimento senza effettivamente svolgerlo: a seconda del movimento scelto, si registrano diverse attivazioni delle aree della corteccia cerebrale adibite al movimento. È stato anche rilevato⁸ come sia possibile *leggere* l'intenzione del soggetto di compiere un determinato movimento proprio dalla corteccia parietale posteriore prima che effettivamente sia inviato un segnale alla corteccia motoria. Il segnale poi viene decodificato e tradotto nel movimento desiderato, ad esempio del braccio robotico⁹.

⁵Si pensi, ad esempio, a un algoritmo di una *brain-computer interface* che deve rilevare l'attività cerebrale, i *neural patterns*, per controllare un braccio robotico che funge da protesi per un soggetto paralizzato o privo dell'arto.

⁶Zander et al., “Enhancing human-computer interaction with input from active and passive brain-computer interfaces”, 2010. È opportuno tenere a mente che tale tripartizione non è accolta da alcuni autori e da una parte della letteratura. Per una opinione diversa, cfr. J. R. Wolpaw e E. W. Wolpaw, “Brain-computer interfaces: something new under the sun”, 2012.

⁷Per la descrizione delle tre tipologie di interfacce, seguo Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

⁸Affalo et al., “Decoding motor imagery from the posterior parietal cortex of a tetraplegic human”, 2015.

⁹Gli stessi ricercatori hanno anche aperto alla possibilità di tradurre in segnali idonei a provocare il movimento anche intenzioni non motorie, come ad esempio, il desiderio di

Nelle BCI reattive è lo stesso sistema dell'interfaccia che, attraverso l'invio di uno stimolo, fa sì che l'attività cerebrale sia diversamente modulata a seconda della natura dell'impulso stesso¹⁰. Quest'ultimo, infatti, può consistere in lettere o simboli che compaiono su uno schermo. L'utente focalizza l'attenzione su uno di questi facendo sì che l'algoritmo delle BCI lo selezioni: l'attenzione selettiva, dunque, rappresenta il mezzo per il soggetto per interagire con l'ambiente circostante.

Infine, le BCI passive monitorano semplicemente l'attività cerebrale, nel momento in cui il soggetto non è impegnato in nessuna attività mentale. Una delle applicazioni più utili delle interfacce passive è volta a far sì che l'utente non si distraiga là dove sia impegnato in lavori faticosi o pericolosi, monitorando il livello e rilevando eventuali cali di attenzione¹¹.

Queste tre categorie di interfacce possono essere analizzate sotto il profilo dei risultati finali che si raggiungono¹². Si individuano, in particolare, due principali *outputs*.

Nel c.d. *goal-selection* l'utente indirizza il dispositivo verso un risultato, o *goal*, impartendogli il comando di realizzarlo. Il soggetto potrebbe, ad esempio, comandare al braccio robotico di prendere un oggetto o interagire con un pulsante. In questo caso i singoli passaggi dell'esecuzione (la velocità, la direzione, l'apertura e la chiusura delle dita robotiche) sono svolti dall'algoritmo dell'interfaccia, mentre l'utente determina solo il risultato finale. Proprio per questo il procedimento di *goal-selection* risulta più agevole per il soggetto. Il ruolo più impegnativo è svolto dall'interfaccia, alla quale è richiesta una notevole capacità computazionale per il conseguimento di quanto voluto dal suo utilizzatore.

Nel caso del *process-control*, invece, l'utente esercita un maggiore controllo sull'intero processo. Egli non deve più impartire un solo comando, ma, al contrario, deve guidare senza soluzione di continuità il dispositivo in tutti gli *steps* dell'esecuzione, fino all'obiettivo desiderato. La dinamica si rivela più impegnativa sia per la BCI che per l'utente, al quale è continuamente richiesto di impartire comandi per assicurare che il processo raggiunga il fine

movimento.

¹⁰Höhne et al., "A novel 9-class auditory ERP paradigm driving a predictive text entry system", 2011.

¹¹Per la sperimentazione delle BCI passive in questo caso, vd Martel, Dähne e Blankertz, "EEG predictors of covert vigilant attention", 2014. Proprio questa tipologia di interfaccia è utilizzata nell'ambito del *neurogaming* che abbiamo descritto in precedenza. Cfr. Mühl et al., "A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges", 2014, pag. 68.

¹²In questo caso, il punto di riferimento è J. R. Wolpaw e E. W. Wolpaw, "Brain-computer interfaces: something new under the sun", 2012.

desiderato. Ad esempio, là dove sia necessario muovere una sedia a rotelle, la persona disabile dotata di BCI dovrebbe impegnarsi in una continua e faticosa attività cerebrale per guidarla e per concentrarsi, pensando a ogni singolo movimento. L'utente non solo dovrebbe pensare alla direzione nella quale si muove la sedia, ma anche la velocità, verificando la necessità o meno di frenare a seconda dell'ambiente circostante.

Tali tecnologie, per quanto ancora soggette ad alcuni limiti strutturali, rappresentano solo il primo passo. La nuova frontiera è l'interazione di queste dinamiche di interazione con algoritmi ad apprendimento automatico.

Questi ultimi possono leggere le intenzioni e i risultati a cui tende l'utente a tal punto da predire il comportamento futuro del soggetto. L'avvento di questa nuova strumentazione certamente porterà con sé nuove problematiche, ma è in grado anche di colmare eventuali deficit di controllo da parte dell'utente sulle interfacce¹³. Alla fine, quello che si può formare è un *gap* che “*cannot be bridged by traditional concepts of responsibility ascription*”¹⁴.

È interessante il riferimento all'immagine dell'attraversare gettando un ponte (*to bridge*): è un gettare un ponte sopra il fiume del progresso, dato dal sempre più vasto e pervasivo utilizzo di queste tecnologie “intelligenti”, che inevitabilmente si infrange sulle categorie giuridiche, filosofiche ed etiche che si conoscono.

Ecco, allora, che non si può più rimandare il momento di verificare se tali categorie sono idonee a reggere l'urto dell'onda e unire le due sponde, o se l'onda si è ormai trasformata, con il tempo, in uno *tsunami* a seguito del passaggio del quale tutto è distrutto e tutto deve essere ricostruito.

È arrivato così il momento di uscire dal laboratorio, senza chiudere la porta, e aprire le porte di un altro luogo, non meno importante. Il Tribunale. È giunta l'ora di verificare, effettivamente, quali siano le possibili ricadute giuridiche di questo fenomeno, facendo la conoscenza, innanzitutto, di una donna di nome Mary.

¹³Per approfondire la questione anche tecnicamente, vd. Glaser et al., “Machine learning for neural decoding”, 2020.

¹⁴Matthias, “The responsibility gap: Ascribing responsibility for the actions of learning automata”, 2004, pag. 175.

3.4 Il Golem a processo. Il penale

3.4.1 Un caso limite

Mary¹⁵ è una donna sposata che conduce una vita ordinaria. L'unica, piccola, peculiarità che distingue Mary dal resto delle persone altrettanto ordinarie è che riesce a muovere gli oggetti con il pensiero. La donna, infatti, sceglie, in piena coscienza, di sostituire le proprie braccia biologiche con due arti robotici che riesce a muovere attraverso una piccola interfaccia invasiva installata nel suo cervello. Grazie alla tecnologia BCI, Mary guida e utilizza le sue nuove braccia con il proprio pensiero¹⁶, in una modalità che non sembra apparentemente diversa, per un osservatore esterno, dal modo in cui muoveva i suoi vecchi arti. Fino a questo momento non si sono verificati incidenti. Un giorno, tuttavia, dopo un litigio con il marito per questioni di gelosia, Mary *immagina*, per un istante, di accoltellare il coniuge con un coltello che ha visto sul tavolo. Prima ancora di poter reagire, il suo braccio destro afferra il coltello e lo pianta nel petto del marito.

Dopo il fatto, la donna chiama l'ambulanza.

3.4.2 Il Golem tra Arizona e Italia

La vicenda di Mary e del suo sfortunato marito presenta molteplici profili rilevanti in una prospettiva legale. Per comprendere in che modo la donna sia responsabile, una volta accertato il fatto è necessario procedere a una sua qualificazione giuridica. La prima domanda che ci si deve porre è se la donna ha voluto o meno accoltellare il marito: ci si deve chiedere, cioè, a che titolo il fatto possa esserle imputato. La ricostruzione nell'uno o nell'altro senso espone l'utente di BCI a diverse conseguenze.

Se il fatto ricadesse, ad esempio, nella giurisdizione dello Stato dell'Arizona¹⁷, Mary potrebbe essere accusata innanzitutto di *second-degree murder*. In particolare, l'*Arizona criminal code* stabilisce che commette omicidio di secondo grado la persona che, senza premeditazione, "*intentionally causes the death of another person, including an unborn child or, as a result of*

¹⁵È qui presentato l'immaginario ma realistico caso descritto in Gurney, "Killer robot arms: A case-study in brain-computer interfaces and intentional acts", 2018.

¹⁶Può essere questa una interazione *goal-selection* basata su una interfaccia attiva che prevede l'utilizzo di quell'algoritmo ad apprendimento automatico, a cui si fa riferimento nel paragrafo precedente, che è in grado di predire le intenzioni dell'utente.

¹⁷Il riferimento è ancora a Gurney, "Killer robot arms: A case-study in brain-computer interfaces and intentional acts", 2018. Gurney, dopo aver illustrato il caso, ne espone le conseguenze alla luce della legislazione dell'Arizona. Nonostante sia un contesto di *common law*, vale comunque la pena di seguire l'autore in questa ricostruzione.

*intentionally causing the death of another person, causes the death of an unborn child*¹⁸. Dunque, l'autore del reato causa *intenzionalmente* la morte di un altro soggetto. Se riconosciuta colpevole, in base a tale qualificazione Mary rischierebbe dai dieci ai venticinque anni di carcere.

Al contrario, se si ammettesse che la donna è stata provocata dal marito, ecco che ciò potrebbe portare a conseguenze diverse. Il fatto verrebbe qualificato non già come *second-degree murder*, ma come *manslaughter*. Ancora una volta è la legge¹⁹ che ne fornisce una definizione. Si commette *manslaughter* nel “*committing second degree murder as prescribed in section 13-1104, subsection A on a sudden quarrel or heat of passion resulting from adequate provocation by the victim.*” La cornice edittale prevede la pena della reclusione fino a dieci anni. Nella vicenda in questione, se da un lato Mary effettivamente era in preda a un forte stato di alterazione emotiva (*heat of passion* a cui fa riferimento la norma), dall'altro non vi è stata provocazione da parte della vittima, con la conseguenza ultima che Mary dovrà rispondere inevitabilmente di *second-degree murder*.

Altre considerazioni devono essere fatte se Mary fosse cittadina italiana e il fatto ricadesse nella giurisdizione penale italiana. In questo caso il pubblico ministero, concluse le indagini preliminari, formula l'imputazione.

Ancora una volta si deve ricostruire l'elemento psicologico del reato: capire, cioè, se il fatto deve essere attribuito a titolo di colpa o di dolo.

Alla luce dell'ordinamento penale italiano, il primo indice normativo di riferimento è l'articolo 43 del codice penale, rubricato *Elemento psicologico del reato*²⁰. La disposizione distingue tra delitto doloso, preterintenzionale e colposo. Una volta accertati i fatti, a Mary sarebbe contestato il delitto di omicidio. Quest'ultimo trova come norme incriminatrici gli articoli 575 e seguenti del codice penale.

In primo luogo, il fatto in questione potrebbe essere qualificato come omicidio doloso, e, dunque, alla luce del disposto dell'art 43, *secondo l'intenzione*. Il punto di riferimento è l'art. 575 c.p.²¹. Mary, tuttavia, ha ucciso

¹⁸Ariz. Rev. Stat. Ann. § 13-1104(a)(1).

¹⁹Ariz. Rev. Stat. Ann. § 13-1103(a)(2).

²⁰Recita così l'art. 43: “*Il delitto: è doloso, o secondo l'intenzione, quando l'evento dannoso o pericoloso, che è il risultato dell'azione od omissione e da cui la legge fa dipendere l'esistenza del delitto, è dall'agente preveduto e voluto come conseguenza della propria azione od omissione; è preterintenzionale, o oltre l'intenzione, quando dall'azione od omissione deriva un evento dannoso o pericoloso più grave di quello voluto dall'agente; è colposo, o contro l'intenzione, quando l'evento, anche se preveduto, non è voluto dall'agente e si verifica a causa di negligenza o imprudenza o imperizia, ovvero per inosservanza di leggi, regolamenti, ordini o discipline.*”

²¹Art. 575 c.p.: “*Chiunque cagiona la morte di un uomo è punito con la reclusione non*

il marito: il rapporto di coniugio è considerato dalla legge²² una circostanza aggravante speciale, con la conseguenza ultima, ed estrema, che la donna, se riconosciuta colpevole di omicidio doloso aggravato dal rapporto di coniugio, può essere sottoposta alla pena perpetua dell'ergastolo.

Ad altra sanzione sarebbe esposta l'imputata se il fatto fosse qualificato come omicidio colposo, o, sempre in combinato disposto con l'art. 43, *contro l'intenzione*²³. Si potrebbe infatti sostenere che Mary avrebbe semplicemente fatto un uso "scorretto" delle braccia, per imprudenza o imperizia nell'utilizzo degli arti e dunque, in definitiva, nell'inviare input all'interfaccia²⁴. Muta la sanzione e muta anche il riferimento normativo²⁵. In base a tale qualificazione, dunque, Mary sarebbe esposta a una pena detentiva che va da un minimo di sei mesi al massimo di cinque anni²⁶.

La ricostruzione nell'uno o nell'altro senso comporta dunque delle conseguenze radicali e irreversibili. Mary, come ogni altro utente di BCI, si troverebbe esposta in ogni caso a gravi sanzioni penali, con la possibilità, addirittura, di trascorrere il resto della sua vita in carcere.

L'elemento decisivo, anche per costruire una eventuale strategia di difesa da parte del difensore, risiede proprio in quella intenzione che determina l'elemento psicologico del reato e permette la qualificazione giuridica del fatto a seconda che il delitto sia *secondo*, *contro* o *oltre* questa.

inferiore ad anni ventuno."

²²Art. 577 c.p. primo comma, n. 1: "*Si applica la pena dell'ergastolo se il fatto preveduto dall'articolo 575 è commesso: 1) contro l'ascendente o il discendente anche per effetto di adozione di minorenni o contro il coniuge, anche legalmente separato, contro l'altra parte dell'unione civile o contro la persona stabilmente convivente con il colpevole o ad esso legata da relazione affettiva.*" Nella vicenda che si sta analizzando non vengono in rilievo le altre circostanze aggravanti previste negli altri numeri di questo primo comma dell'art 577, ossia l'utilizzo di sostanze venefiche, la premeditazione e il concorso delle circostanze previste all'art 61 c.p., numeri 1 e 4.

²³Come è noto, la colpa consisterebbe nella violazione di regole cautelari scritte o non scritte, come regole di diligenza, prudenza o perizia, poste a tutela di beni considerati dal legislatore meritevoli di tutela. Per una autorevole dottrina, si veda Fiandaca, *Diritto penale. Parte generale*, 2019.

²⁴In questo caso, non si configurerebbe una condotta negligente, se intesa come la "violazione di una regola di condotta che prescrive un'attività doverosa". vd Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 274.

²⁵Art 589 c.p., "*Chiunque cagiona per colpa la morte di una persona è punito con la reclusione da sei mesi a cinque anni [...]*".

²⁶Non si riscontra nel caso fittizio la presenza delle circostanze aggravanti previste per omicidio colposo ai commi 2 e 3 dell'art. 589: rispettivamente l'aver commesso il fatto "con violazione delle norme per la prevenzione degli infortuni sul lavoro" e "nell'esercizio abusivo di una professione per la quale è richiesta una speciale abilitazione dello Stato o di un'arte sanitaria."

È necessario chiedersi in che modo possa essere descritta tanto in termini giuridici, quanto in termini neuroscientifici, iniziando a conoscere proprio questi ultimi.

3.5 L'intenzione

Il punto di partenza per poter rispondere a questo quesito è verificare effettivamente in cosa consista l'intenzione di compiere un'azione, in questo caso l'intenzione di accoltellare l'uomo.

Tra i principali problemi che sono emersi nella letteratura quello che più rileva ai fini della trattazione è individuare la relazione che lega l'azione che si può definire intenzionale e la più sfuggente intenzione di agire²⁷. Si incontra subito un primo ostacolo: è molto difficile tracciare una definizione univoca e pacifica.

Un primo approccio, definito *desire-belief model*, riconduce l'intenzione ai desideri (*desires*) e alle credenze (*beliefs*) del soggetto²⁸. È questa una ricostruzione che, lungi dall'essere definitiva, è stata oggetto di varie critiche²⁹.

Alcuni³⁰ ritengono addirittura che sia più agevole darle una definizione in “negativo”: ecco allora che diviene intenzionale quella azione che *non* è compiuta “*by accident, by mistake, unwittingly, inadvertently [...]*”³¹.

In ogni caso, qualunque sia la definizione che si può adottare, è comunque possibile tracciare delle utili distinzioni generali, tenendo sempre in mente

²⁷Bratman, “Two faces of intention”, 1984, pag. 375. Alla base vi è la consapevolezza che l'intenzione caratterizza tanto le azioni quanto i c.d. stati mentali. In ogni caso, anche il concetto di azione sarà ricostruito alla luce delle problematiche delle interfacce. Per ora è sufficiente considerare l'azione come il movimento del braccio robotico della donna che afferra il coltello e pugnala la vittima.

²⁸Tra i sostenitori di questo modello, si veda, senza pretese di completezza, Audi, “Intending”, 1974 e P. M. Churchland, “The logical character of action-explanations”, 1970.

²⁹In particolare, è stato osservato come nella vita quotidiana molte volte ci si trova innanzi all'alternativa tra varie scelte, tutte egualmente condivisibili e attuabili alla luce dei desideri e delle credenze. Ciò che orienta la persona nell'intraprendere l'una o l'altra strada è proprio l'intenzione, che dunque non si può ridurre a mero desiderio o a mera credenza. Si veda, in particolare e senza pretese di completezza, Bratman, “Intention and means-end reasoning”, 1981, pp. 262-265.

³⁰Gurney, “Killer robot arms: A case-study in brain-computer interfaces and intentional acts”, 2018.

³¹Bratman, “Two faces of intention”, 1984, pag. 384.

che le intenzioni possono essere di diversa natura e di diversa intensità³², e che la complessità dell'attività cerebrale ha portato a evidenziare come le varie intenzioni possono formarsi influenzandosi reciprocamente, con una interazione che non è ancora compresa³³ del tutto e della quale, dunque, non se ne può tenere conto nella programmazione dell'algoritmo delle interfacce³⁴.

Nonostante queste problematiche, una prima distinzione utile da poter ricordare corre tra le intenzioni c.d. distali e prossimali.

Le prime³⁵ si formano là dove il soggetto è deciso di compiere una determinata azione nel futuro.

Le seconde, invece, emergono come intenzioni volte a iniziare un'azione in contemporanea con il sistema motorio, stimolato dalle intenzioni motorie³⁶.

Alla luce poi dell'utilizzo delle BCI, è necessario chiedersi se l'utente ha agito a seguito dell'una o dell'altra, anche per poter verificare effettivamente se l'azione è premeditata o no. È ormai pacifico nel dibattito neuroscientifico che il *pensiero* di compiere un'azione e la sua esecuzione sono due momenti diversi dell'agire³⁷.

Altri ancora hanno evidenziato la distinzione tra *P-intentions* o *present-directed intentions* e *future-directed intentions* o *F-intentions*³⁸. Queste ultime si possono ricondurre essenzialmente alla deliberazione dell'azione, mentre le prime permettono di strutturare la deliberazione in modo più idoneo e conforme alla situazione specifica basandosi su una rappresentazione più concreta dell'obiettivo che si vuole perseguire.

A prescindere dalla qualificazione nell'uno o nell'altro senso, l'analisi della traduzione dell'intenzione in azione alla luce delle BCI si rivela ardua

³²Vi possono essere intenzioni di compiere una determinata, singola azione o di agire in modo continuativo, di agire con una determinata modalità o di astenersi dall'agire.

³³Rainey, Maslen e Savulescu, "When thinking is doing: Responsibility for BCI-mediated action", 2020.

³⁴Cosa siano e in che modo si manifestino le intenzioni è comunque oggetto di una multidisciplinare ricerca. Si veda, ad esempio, Pacherie, "Towards a dynamic theory of intentions", 2006, nonché Pacherie, "Towards a dynamic theory of intentions", 2006.

³⁵Riprendo Rainey, Maslen e Savulescu, "When thinking is doing: Responsibility for BCI-mediated action", 2020.

³⁶A. R. Mele, "Acting for reasons and acting intentionally", 1992.

³⁷"Those with chips implanted in their brains can think about executing a bodily movement, and that thought alone can cause the movement [...] but forming an intention or plan and executing it are two separate mental acts [...] one can form an intention to act but not executing that intention [...] by changing one's mind at the last moment." Glannon et al., *Bioethics and the brain*, 2007, pag 142.

³⁸Pacherie, "The sense of control and the sense of agency", 2007. L'autrice poi individua anche una terza specie, le c.d. *motor-intentions*, o *M-intentions*, che conducono al movimento corporeo.

e problematica sotto almeno tre profili.

Innanzitutto è necessario analizzare quale sia la dinamica che lega intenzione e azione e in che modo se ne può tenere conto nella programmazione delle interfacce.

In secondo luogo, occorre interrogarsi circa la natura e la rilevanza della c.d. funzione di veto.

Infine, la questione si complica ulteriormente per il fatto che l'attività cerebrale del soggetto può essere anche involontaria, non “voluta”, con la conseguenza che l'utente delle interfacce può aver inviato non consciamente³⁹ l'impulso cerebrale alla BCI, la quale, tuttavia, ha eseguito il movimento.

Iniziando proprio dal primo aspetto, si deve subito prendere atto di una prima criticità. È ancora sconosciuto sotto ogni profilo, anche sotto un punto di vista neuroanatomico, il meccanismo, il fattore che “traduce” lo stato mentale del soggetto in una effettiva azione. Questo è definito *I factor*⁴⁰. La sua natura è tanto ignota che alcuni sostengono che forse sia impossibile addirittura *in principio* cogliere la reale essenza di questo fattore esecutivo. Inevitabilmente questo mistero si ripercuote sull'impiego delle BCI.

La strada per attraversare il *gap* è resa ancora più tortuosa, dunque, per il fatto che al sistema delle interfacce è richiesto di rilevare o meno l'attivarsi di questo sconosciuto fattore, in modo da distinguere il piano della mera immaginazione e il piano dell'intenzione volta a dare esecuzione all'azione. Una operazione che comporta, allora, che l'algoritmo stesso debba essere in grado di “distinguere non solo la sintassi dei segnali cerebrali, [...] ma anche la loro semantica intenzionale o [...] la loro forza illocutoria”⁴¹. Quest'ultima deve essere intesa come la forza, la rilevanza di ciò che si dice e che si *vuole intendere* comunicando oralmente. Le BCI devono cogliere il significato dietro al significante, ossia l'effettiva intenzione che si cela dietro l'attività cerebrale: ciò rappresenta tuttavia un limite quasi insuperabile per

³⁹Ci si potrebbe anche chiedere cosa voglia dire “consciamente” e in che rapporto sia la coscienza con l'intenzione. È una indagine che tuttavia esulerebbe dal perimetro della trattazione in esame per entrare in sterminati ambiti di ricerca. Per il fine perseguito in questa sede, si può intendere semplicemente come “con consapevolezza”.

⁴⁰Rainey, Maslen e Savulescu, “When thinking is doing: Responsibility for BCI-mediated action”, 2020. Il fattore “I” è da questi autori considerato proprio come l'elemento che fa la differenza nello studio delle azioni intenzionali.

⁴¹È questa l'interessante ricostruzione di Stefano Fuselli in Fuselli, “Brain-Computer Interface e soggettività agente. Considerazioni etico-giuridiche.”, 2021, pag. 166, in Moro, *Etica, diritto e tecnologia. Percorsi dell'informatica giuridica contemporanea*, 2021. Il riferimento è alla teoria degli atti linguistici che, com'è noto, è stata elaborata da John Langshaw Austin nel 1975. Senza pretese di completezza, il rimando è a Austin, *How to do things with words*, 1975.

potenziare e raffinare l'algoritmo, che condiziona anche quanto controllo può esercitare l'utente.

Tutta questa incertezza inevitabilmente si ripercuote nell'accertamento della responsabilità del soggetto, dal momento che si incrina anche l'illusione di poter trovare risposte alla questione attingendo alla conoscenza neuroscientifica.

Il secondo profilo da prendere in considerazione attiene alla c.d. funzione di veto. Per comprendere questo secondo aspetto, occorre in primo luogo chiarire cosa sia il veto. Com'è noto, il neurofisiologo Benjamin Libet⁴² ha rilevato in via sperimentale⁴³ che l'attività cerebrale coinvolta nel compimento di un atto volontario, situata nella corteccia motoria secondaria, ha inizio 300 millisecondi prima che il soggetto sottoposto all'esperimento fosse consapevole della volontà di agire. Rimangono poi 150 millisecondi prima che il movimento sia effettivamente eseguito, e, negli ultimi 50 millisecondi, quest'ultimo non può essere più fermato⁴⁴. Nei millisecondi che sono impiegati nell'eseguire il movimento, trasportando l'impulso all'arto, è possibile comunque impedirne l'esecuzione o sviarlo: è possibile cioè porre quello che è stato definito un veto sul movimento. Il tempo necessario per bloccare l'impulso è di 200 millisecondi⁴⁵: conclusa questa finestra di tempo, che non a caso è stata definita un punto di non ritorno⁴⁶, non è più possibile incidere sul movimento in alcun modo. In altre parole, qualunque movimento corporeo è preceduto da un'attività preparatoria composta da tutti i processi che permettono di trasmettere il segnale dalla corteccia motoria all'arto: il cervello è più veloce della nostra consapevolezza. Gli esperimenti di Benjamin Libet hanno scopercchiato un vero e proprio vaso di Pandora, giungendo a mettere in discussione l'esistenza stessa del libero arbitrio⁴⁷. Se i movimenti non sono frutto di scelte libere e consapevoli, ma semplicemente di una precedente, involontaria e inconsapevole attività cerebrale, le scelte che si

⁴²Benjamin Libet (1916-2007) è stato un neurofisiologo e psicologo statunitense, considerato un pioniere nello studio della coscienza umana.

⁴³Per una descrizione completa dell'esperimento, si veda Libet, *Mind time: The temporal factor in consciousness*, 2009.

⁴⁴Libet et al., "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential)", 1993.

⁴⁵Schultze-Kraft et al., "The point of no return in vetoing self-initiated movements", 2016.

⁴⁶*ivi*.

⁴⁷Così ad esempio Wegner, *The illusion of conscious will*, 2017. Altri hanno sostenuto invece che la coscienza della scelta sia troppo lenta a formarsi per essere considerata effettivamente responsabile del movimento. Vd. Blackmore, "Why psi tells us nothing about consciousness", 1998.

compiono sono effettivamente libere? Lo stesso Libet ha, comunque, ritenuto i risultati del lavoro compatibili con la nozione di libero arbitrio, facendo leva proprio sulla possibilità per il soggetto di esercitare il potere di veto⁴⁸, influenzando la concreta esecuzione dell'azione. Non verrebbe meno il libero arbitrio, la libertà di agire verrebbe semplicemente declinata al negativo, ossia come libertà di *non* agire⁴⁹. La presenza del veto e della facoltà del soggetto di esercitarla diviene centrale per poter tracciare una distinzione tra due categorie di intenzioni rilevanti per analizzare la situazione in cui si trova un utente di BCI: le c.d. intenzioni *post-veto* e le intenzioni c.d. *pre-veto*. Riprendendo la definizione di veto che ne danno alcuni autori("*To veto a conscious proximal intention or urge to do something is to decide not to act on it and to refrain, accordingly, from acting on it*"⁵⁰), in riferimento alle due classi di intenzioni suddette, si può ritenere che se in risposta a uno stimolo si forma una intenzione *pre-veto* e a seguito di riflessione è mantenuta l'intenzione, allora questa diviene *post-veto*. Ancora, se dopo la prima, *pre-veto*, se ne forma un'altra di diversa, è l'ultima a divenire *post-veto*⁵¹. Riconducendo quanto detto finora all'esperienza delle interfacce neuronali, si deve rilevare come la possibilità di implementare in queste tecnologie il potere di veto, significherebbe garantire un maggiore controllo per l'utente delle BCI⁵². È, tuttavia, una soluzione solo apparente. Anche se si riuscisse effettivamente a garantire il riconoscimento da parte dell'interfaccia del ve-

⁴⁸Scrive Libet: "Non è esclusa la possibilità che i fattori su cui si basa la decisione di veto si sviluppino attraverso processi inconsci che precedono il veto [...] La consapevolezza di quella decisione di veto può richiedere di essere preceduta da processi inconsci, ma il contenuto di quella consapevolezza è una caratteristica separata che non necessita del medesimo processo." Libet, *Mind time: The temporal factor in consciousness*, 2009, pag 151. Per una disamina filosofica dell'esperimento, si veda Chiereghin, "La coscienza: un ritardato mentale?", 2008.

⁴⁹In letteratura alcuni autori, con un interessante gioco di parole, hanno trasformato il libero arbitrio, *free will*, in *free won't*. Vd. Meixner, "New perspectives for a dualistic conception of mental causation", 2008.

⁵⁰A. Mele, "Proximal intentions, intention-reports, and vetoing", 2008, pag. 5.

⁵¹Riprendo ancora una volta Gurney, "Killer robot arms: A case-study in brain-computer interfaces and intentional acts", 2018.

⁵²Nota giustamente Stefano Fuselli: "Questo significa, però, in ultima analisi, *delegare* [enfasi presente nel testo originario] la funzione di veto a quello stesso algoritmo che, magari, è stato addestrato ad anticipare i *pattern* dell'attività cerebrale per rendere più veloce la risposta motoria dell'apparecchiatura. Anche quella che, secondo una certa lettura, poteva sembrare essere l'ultima traccia di una volontà cosciente irriducibile a una attività neuronale deve essere trasformata in un set di dati digitalizzabili." Vd. Fuselli, "Brain-Computer Interface e soggettività agente. Considerazioni etico-giuridiche.", 2021, in Moro, *Etica, diritto e tecnologia. Percorsi dell'informatica giuridica contemporanea*, 2021, pag. 167.

to⁵³ e delle intenzioni *pre* e *post*, si sposterebbero semplicemente i termini del problema del *Responsibility Gap* sull'effettivo controllo dell'utente sul "filtro" che è assicurato dalle interfacce là dove permettano il veto, allo stesso modo in cui il corpo con i suoi tempi permette di intervenire bloccando il movimento.

Il terzo e ultimo aspetto da prendere in considerazione, che presenta alcuni punti in comune con gli altri due profili, attiene all'intenzionalità o meno dell'attività cerebrale dell'utente, ossia il volere gli impulsi che giungono alle interfacce e che sono da queste tradotti in movimenti dell'arto o della protesi. Tra una attività cerebrale involontaria, quale è quella che controlla, ad esempio, il battito cardiaco, e una attività invece che si può definire intenzionale o volontaria, vi è una zona grigia che non è facile sondare⁵⁴. Emblematica è l'attività neurale là dove vengano alla memoria dei ricordi episodici⁵⁵. Il ricordo involontario di ricordi episodici può essere, ad esempio, distinto dal ricordo intenzionale sotto il profilo dell'attivazione della regione dorsale frontale⁵⁶. Ancora, le regioni frontali del cervello si attivano quando, ad esempio, si cerca mentalmente una parola⁵⁷. Proprio questa rilevazione ha portato alcuni a sostenere, dunque, che durante l'attività neurale in una certa misura le persone attivano, volontariamente, il proprio cervello⁵⁸. Tale "utilizzo" del cervello permette l'interazione tra BCI e utente e fa sì che il primo possa controllare la prima esercitando comunque un controllo, anche se parziale. Si trasforma così in un'abilità⁵⁹ che può essere affinata ed esercitata dall'utente: *"we can exercise control over our brains, and we can act neurally, and so we can establish voluntariness about some types of neural*

⁵³Operazione comunque resa assai ardua dalla natura comunque incerta di questo veto: *"It is not clear how a subject who intends to perform a movement that is predictable on the basis of neural EEG signals could change her mind, cancel the intention and refrain from performing it."* Glannon, "Ethical issues in neuroprosthetics", 2016, pag. 11.

⁵⁴Rainey, Maslen e Savulescu, "When thinking is doing: Responsibility for BCI-mediated action", 2020.

⁵⁵Com'è noto, l'espressione memoria episodica è stato coniato da Endel Tulving nel 1972: si può definire come l'insieme di esperienze personali passate che si sono verificate in determinati luoghi e momenti. Per un maggiore approfondimento, cfr. Tulving, "Elements of episodic memory", 1983.

⁵⁶Hall et al., "The neural basis of involuntary episodic memories", 2014.

⁵⁷È questo il fenomeno che si rivede quando si cerca una parola che si ha "sulla punta della lingua." Baars, "How brain reveals mind neural studies support the fundamental role of conscious experience", 2003.

⁵⁸Kirmayer e Gold, "Re-socializing psychiatry", 2012.

⁵⁹J. R. Wolpaw, Birbaumer et al., "Brain-computer interfaces for communication and control", 2002.

activity”⁶⁰. Altri ancora si spingono a paragonare l’utilizzo delle BCI nell’apprendimento di un nuovo linguaggio da parte dell’utente⁶¹, tanto da far emergere in letteratura, in riferimento a utenti che non riescono a integrarsi con l’interfaccia di cui sono dotati, l’espressione *BCI illiterates*: degli analfabeti “neuronal” che falliscono nell’obiettivo di far funzionare correttamente i dispositivi⁶².

Sulla base di queste considerazioni, si pone un fondamentale problema in riferimento alla necessità di far sì che l’interfaccia riesca a distinguere tra l’attività voluta e quella non voluta, tra il “rumore di fondo” dell’attività neurale involontaria e gli impulsi inviati consapevolmente dalla persona.

La questione sorge dal momento che, da un punto di vista neuroanatomico, il *covert-speech*, o *inner-speech*, ossia il dialogo interiore che ogni soggetto ha con se stesso⁶³, può fungere da fattore di attivazione (*trigger*), nei confronti dell’interfaccia, non diversamente da quanto è in grado di fare l’*overt-speech*, il flusso di parole che sono verbalizzate o dette ad alta voce⁶⁴. Questo perchè i dispositivi neurali possono in primo luogo rilevare il dialogo interiore, per poi, in secondo luogo, decodificarlo nei medesimi segnali estraibili dal dialogo esteriore⁶⁵, giungendo, in conclusione di questa dinamica di rilevazione, al processamento, alla ricostruzione e alla esternalizzazione del primo come se si trattasse del secondo⁶⁶.

⁶⁰Rainey, Maslen e Savulescu, “When thinking is doing: Responsibility for BCI-mediated action”, 2020, pag. 51. Gli autori nello stesso punto usano come esempio per analogia l’abilità di guidare una macchina: un guidatore esperto può essere considerato più moralmente responsabile di un guidatore meno esperto in caso di incidente. Dice Rainey: “*The untrained driver does not know about appropriate gears. In having such knowledge available to them, and failing to act upon it, the trained driver omits to correct a worse scenario and is responsible for that omission in a way the untrained driver is not. In knowing that they are untrained, however, the untrained driver still ought to be held responsible for their actions as they ought to have been able to foresee problems would arise from getting behind the wheel untrained.*”

⁶¹Farahany, “A neurological foundation for freedom”, 2012.

⁶²Per un approccio critico all’espressione, vd Thompson, “Critiquing the concept of BCI illiteracy”, 2019.

⁶³È la c.d. lettura endofasica, che si considera tra le altre cose anche come elemento fondamentale nella regolazione dell’attività cognitiva e del comportamento, tanto nell’infanzia quanto nell’età adulta. vd Alderson-Day e Fernyhough, “Inner speech: development, cognitive functions, phenomenology, and neurobiology.”, 2015, che richiama R. M. Diaz, Berk e R. Diaz, *Private speech: From social interaction to self-regulation*, 1992.

⁶⁴Rainey, Maslen, Mégevand et al., “Neuroprosthetic speech: The ethical significance of accuracy, control and pragmatics”, 2019.

⁶⁵Bocquelet et al., “Key considerations in designing a speech brain-computer interface”, 2016.

⁶⁶Chakrabarti et al., “Progress in speech decoding from the electrocorticogram”, 2015.

Per evitare che una interfaccia troppo sensibile all'attività cerebrale traduca i pensieri e le intenzioni sbagliate, portando all'esterno ciò che dovrebbe rimanere nel foro interno del soggetto, si è pensato ancora una volta di recuperare la possibilità di incorporare nelle BCI quella stessa funzione di veto della quale prima si è parlato. Ad esempio, è stato sostenuto⁶⁷ che se l'utente ricevesse un *feedback* verbale dal sistema del neurodispositivo con l'indicazione dell'azione che si sta per eseguire, si potrebbe esercitare un veto sulla stessa interfaccia. In questo caso la stessa funzione di veto verrebbe trasformata. La persona non ferma o modifica un'azione iniziata a seguito di attività cerebrale involontaria, ma un'azione che trova principio nella stessa interfaccia. Si crea cioè, un cerchio nel quale l'interfaccia può rilevare il veto sull'azione posto dal soggetto e quest'ultimo può esercitare una funzione di veto sull'azione avviata dell'interfaccia. Alla fine, tuttavia, anche nel caso irrealistico nel quale si riuscisse a compiere tale operazione, si richiederebbe all'utente uno sforzo cognitivo che, alla lunga, può divenire stancante e sfibrante, vanificando l'utilità effettiva del sistema interfaccia-eventuale protesi.

Ecco allora che, in definitiva, se cogliere l'essenza e la natura dell'intenzione appare assai problematico, ancora di più complesso da un punto di vista tecnico è far sì che le interfacce ne tengano conto per la corretta traduzione e individuazione della stessa.

In tutto ciò, la cara Mary sta ancora aspettando il verdetto. Il giudice prende atto della rilevanza e della frammentarietà di queste considerazioni e deve capire in che modo queste considerazioni possono orientarlo nella qualificazione del fatto.

Innanzitutto Mary, immaginando vividamente di accoltellare il marito, avrebbe potuto provocare il movimento della protesi attraverso la rilevazione da parte dell'interfaccia dell'attivazione della zona del cervello che ha riprodotto l'*immagine* di accoltellare il marito nella mente della donna.

In secondo luogo, è possibile che la BCI abbia male *interpretato* l'intenzione di Mary, rilevando un impulso cerebrale che era *quasi* il segnale cerebrale per muovere intenzionalmente il braccio. Il "*quasi*" si dà per quella incertezza circa l'effettiva distinzione tra attività volontaria e non⁶⁸. Inoltre, in questo caso viene alla luce anche quel misterioso *I-factor*⁶⁹ che traduce

⁶⁷Rainey, Maslen e Savulescu, "When thinking is doing: Responsibility for BCI-mediated action", 2020.

⁶⁸Attualmente, anche se in futuro si potrebbero fare dei miglioramenti, il tasso di errore delle interfacce è del 5%: cinque segnali su cento non sono individuati e tradotti correttamente. Cfr. Haselager, "Did I do that? Brain-computer interfacing and the sense of agency", 2013.

⁶⁹vd *supra*.

in azione l'intenzione. In un ulteriore scenario, l'algoritmo dell'interfaccia impiantata nel cervello della donna potrebbe non essere così sensibile da rilevare e distinguere le intenzioni *pre* e *post* veto, evitando di portare ad esecuzione le prime. Di conseguenza, darebbe inizio al movimento in entrambi i casi: ciò si ripercuoterebbe anche sulla stessa percezione di Mary, l'utente finale, dal momento che potrebbe vedere le braccia robotiche muoversi "da sole" senza sapere che in realtà sta comunque traducendo in azione una propria intenzione.

Infine, nel caso più semplice da interpretare ed affrontare, l'interfaccia è semplicemente incorsa in un malfunzionamento, con la conseguenza ultima di sollevare l'utente da ogni responsabilità. In quest'ultima ottica sarebbe da considerare la posizione di un ulteriore soggetto, il costruttore o programmatore, in capo al quale può sorgere una responsabilità civilistica per prodotto difettoso⁷⁰. Come si vede, varie sono le possibilità.

3.5.1 La disciplina dell'Arizona

Ancora una volta è possibile iniziare dalla legislazione dell'Arizona.

L'elemento discriminante per poter ricondurre la condotta di Mary ai c.d. *crimes of intent*, ossia ai delitti intenzionali⁷¹ o ai *crimes of neglect*⁷², i delitti colposi, risiede nella distinzione tra *mens rea* e *actus reus*. La prima si può ricondurre alla forma dell'elemento soggettivo del reato data del dolo, la seconda invece si può invece ricondurre alla condotta illecita della persona. Nel caso di Mary, si giunge alla paradossale conclusione nella quale la donna manifesta la *mens rea*, perchè una qualche forma di intenzione, anche se non bene qualificabile giuridicamente, si è formata, ma non procede a porre in essere l'*actus reus*, dal momento che è stato l'algoritmo dell'interfaccia ad avviare il movimento. Alla fine, tuttavia, ciò non toglie che il marito è comunque morto con un coltello nel petto⁷³. In altre parole: il solo *pensiero*

⁷⁰Il profilo della responsabilità civilistica è altrettanto problematico da considerare e verrà analizzato più avanti. Qui ci si può limitare a ricordare il tema della prevedibilità del danno: non solo è arduo individuare l'intenzione dietro l'azione, ma gli algoritmi ad apprendimento automatico possono errare in modo imprevedibile, rendendo ancora più difficile la ricostruzione della posizione del programmatore. Un primo rimando è a Larus et al., *When computers decide: European recommendations on machine-learned automated decision making*, 2018.

⁷¹Tra questi, il *second-degree murder*, che può essere imputato alla donna.

⁷²Rientra in questa categoria il delitto di *manslaughter* che prima è stato menzionato. Vd. *Ariz. Rev. Stat. Ann.* § 13-201.

⁷³Gurney, "Killer robot arms: A case-study in brain-computer interfaces and intentional acts", 2018. A seguito della manifestazione del pensiero, avviene comunque l'omicidio: *mens rea* ma non *actus reus*.

dell'omicidio ha comportato l'omicidio stesso. A questo esito paradossale si giunge anche per la definizione di intenzione che dà la legislazione di questo Stato americano. L'*Arizona Revised Statutes* afferma che “‘*Intentionally*’ or ‘*with the intent to*’ means, with respect to a result or to conduct described by a statute defining an offense, that a person’s objective is to cause that result or to engage in that conduct”⁷⁴. Il problema è che, così come da un punto di vista tecnico delle interfacce, anche sotto il profilo giuridico non si ritrova la distinzione tra intenzione *pre* e *post* veto: è lasciato allora all’interprete capire se questa definizione si avvicina di più all’una o all’altra⁷⁵, ma in base a una rigida interpretazione letterale la donna effettivamente ha formato una intenzione e questo è sufficiente per l’imputarle il delitto intenzionale.

La questione, tuttavia, porta con sé una molteplicità di problematiche, in primis la persistente incertezza, più volte menzionata in questo scritto, che caratterizza sotto il profilo neuroscientifico la ricostruzione dell’elemento intenzionale.

3.5.2 L’intenzione nell’ordinamento penale italiano. Il dolo

Attraversando l’oceano e giungendo nel diverso contesto di *civil law* qual è quello italiano, l’analisi qui intrapresa prosegue per una diversa strada per verificare se anche in questo caso si potrebbe giungere ad analoghe considerazioni rispetto a quelle effettuate nel paragrafo precedente.

Anche alla luce della disciplina italiana ci si può chiedere se il delitto commesso sia riconducibile ai delitti dolosi o intenzionali o colposi o non intenzionali. Per rispondere al quesito si può iniziare proprio dall’analisi dei primi.

Com’è noto, l’art 43 c.p. equipara l’intenzione al dolo. Così come affermato anche nella Relazione sul Progetto definitivo del codice penale, n. 59, l’intenzione, e con essa dunque il dolo, “è la direzione della volontà, di cui,

⁷⁴Ariz. Rev. Stat. Ann. § 13-110(10)(a).

⁷⁵Gurney, ad esempio, ritiene che sia più vicina alla definizione di intenzione *post-veto*. L’autore, poi, fa leva sul significato del termine *objective* presente nella definizione di intenzione che fornisce la legislazione dell’Arizona. In particolare, ricordando che l’*Oxford English Dictionary* definisce il termine come “*a thing aimed at or sought; a goal*”, afferma che l’inclusione di questa parola “*is indicative of the kind of intent that leads directly to an action, barring the interference of some outside force*”. Vd. Gurney, “Killer robot arms: A case-study in brain-computer interfaces and intentional acts”, 2018, pag. 783. Non potendosi sostenere che Mary abbia cercato (*sought*) di uccidere il coniuge, ecco che viene meno anche l’obiettivo (*objective*) richiesto della definizione normativa di intenzione. Mary, allora, da un punto di vista giuridico potrebbe non aver avuto alcuna intenzione di compiere *second-degree murder*.

pertanto, è una qualificazione”⁷⁶. Occorre, cioè, la direzione della volontà verso l’evento⁷⁷. La più recente dottrina aderisce essenzialmente alle tesi per le quali l’azione dolosa, intesa come azione volontaria, si risolve in un “tendere a”, in un cercare di ottenere qualcosa con una serie di atti⁷⁸. Tale ricostruzione è condivisa anche dalle Sezioni Unite della Suprema Corte: il dolo è un “pensiero elaborante, motivato da un obiettivo che si risolve in un’intenzione”⁷⁹.

L’attuale qualificazione e ricostruzione del dolo è solo il punto di arrivo di un confronto risalente tra la c.d. teoria della volontà e la teoria della c.d. rappresentazione, che ha diviso la dottrina tedesca dei primi anni del XX secolo. In particolare, i sostenitori della prima, che trova in Franz Von Liszt il proprio caposcuola, sostenevano, in estrema sintesi, che ciò che è voluto non è l’evento, bensì il movimento corporeo che conduce all’evento⁸⁰. In altre parole, “l’atto di volontà e il movimento corporeo corrispondente sono una sola ed identica cosa, appresa da un doppio aspetto”⁸¹.

Al contrario, secondo i sostenitori della teoria della volontà, tra i quali si deve ricordare Robert Von Hippel, il dolo deve essere ricostruito come la volizione di ciò che si rappresenta. Nel codice penale vigente è stata effettuata una sintesi tra le due scuole di pensiero: “v’ha delitto doloso solo quando l’evento sia stato non solo preveduto, ma anche voluto; dal che discende, come corollario logico, che il reato è colposo o contro l’intenzione, quando l’evento, anche se preveduto non è voluto dall’agente”⁸². Si ha delitto doloso, dunque, se l’autore del reato si rappresenta e vuole tutti gli elementi costitutivi del reato, non solo l’evento naturalistico⁸³.

Poste queste considerazioni generali, ci si può chiedere se e quale tipologia di dolo si può individuare nel caso di Mary. Due sono le forme

⁷⁶ 1: *Relazione sul Libro 1. del Progetto / Ministero della Giustizia e degli Affari di culto*, 1929.

⁷⁷ Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 219.

⁷⁸ *ivi*, pag 220.

⁷⁹ Vd. C. S.U. 38343/14

⁸⁰ Liszt, *Lehrbuch des deutschen Strafrechts*, 1888, pag. 119, nota 5. Von Liszt nello stesso scritto, a pag. 165 giudica come imprecisa (*ungenau*), pericolosa (*gefährlich*) e scorretta (*unrichtig*) la tesi allora dominante che descriveva il dolo in termini di “coscienza e volontà degli elementi del reato.” Questa opinione tuttavia è superata, dalla quarta edizione fino alla undicesima, dall’idea che, effettivamente, il dolo è “realizzazione consapevole di tutti gli elementi del reato”.

⁸¹ De Marsico, *Coscienza e volontà nella nozione del dolo*, 1930, pag. 9.

⁸² 1: *Relazione sul Libro 1. del Progetto / Ministero della Giustizia e degli Affari di culto*, 1929. Si può fare riferimento a Ronco, *Scritti patavini: Due tomi indivisibili*, 2017, pag. 578.

⁸³ Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 220.

che potrebbero caratterizzare questa fattispecie, due sono le distinzioni che possono venire qui in rilievo: quella che intercorre tra dolo c.d. d'impeto, così come contrapposto al dolo di proposito, e la distinzione tra dolo c.d. intenzionale e c.d. dolo non intenzionale.

Partendo dalla prima, una volta definito l'impeto come “improvviso moto dell'animo che, acceso da una passione o da un forte sentimento, determina la volontà e spinge ad agire senza dar tempo alla riflessione”⁸⁴, ecco che il conseguente dolo d'impeto verrebbe ad essere là dove la condotta o azione criminosa sia la conseguenza di “un improvviso impulso [...] che si manifesta nei casi in cui la spinta ad agire ha radici affettive, come l'ira o la gelosia.”⁸⁵. Effettivamente, seguendo questa ricostruzione, in questo caso si riscontrano sia l'impulso improvviso sia le radici affettive della spinta ad agire. Inoltre, anche seguendo un'ulteriore interpretazione, per la quale si ha dolo d'impeto “quando la decisione criminosa è improvvisa ed immediatamente eseguita, esplodendo repentinamente nell'atto criminoso”⁸⁶, si riscontrano comunque i suoi estremi. Nel caso invece del dolo di proposito, che trova nella premeditazione la sua massima espressione, occorre comunque una cesura temporale rilevante nella quale si forma l'intenzione criminosa che successivamente viene portata ad esecuzione.

Tutte queste considerazioni e ricostruzioni si scontrano con i caratteri peculiari della vicenda in esame. La prima cosa che potrebbe fare l'utente, infatti, è affermare che egli *non voleva* fare ciò che ha fatto, tenere quella condotta, causare quell'evento. È stata stato il suo cervello che ha inviato senza che egli volesse il segnale all'interfaccia. Un'affermazione del genere se detta da un soggetto sano, privo di interfaccia, non costituirebbe un'efficace difesa: equivarrebbe a dire “non sono stato io, è stato il mio cervello”. Se si aprisse a una tale affermazione verrebbe meno ciò che fonda lo stesso concetto di responsabilità: il libero arbitrio. Si potrebbe, allora, sostenere, ammettendo l'assenza di volontà, che Mary abbia previsto l'evento come possibile ma, nonostante ciò, abbia agito comunque. Questa ricostruzione ricondurrebbe l'elemento psicologico alla seconda distinzione accennata, in particolare al dolo non intenzionale, al quale si può ricondurre il c.d. dolo eventuale, una declinazione dell'elemento doloso di controversa ricostruzione. Alla base vi è l'assunto per il quale la previsione dell'evento può essere equivalente alla volontà, il prevederlo è come volerlo.

Per inquadrare con maggiore precisione il dolo eventuale, nonché per per-

⁸⁴Così l'Enciclopedia Treccani sotto la voce “impeto”.

⁸⁵Marinucci, *Manuale di diritto penale. Parte generale*, 2015, pag 317.

⁸⁶Mantovani, *Diritto penale. Parte generale*, 2015, pag. 324.

mettere di distinguerlo dalla c.d. colpa con previsione, sono state presentate numerose ricostruzioni, ora dottrinali, ora giurisprudenziali, tenendo sempre come punto di riferimento normativo gli artt. 43 c.p.⁸⁷ e 61 n.3. c.p.⁸⁸. A prescindere dalla fondatezza e condivisibilità di ciascuna di esse, ciò che qui rileva ed è centrale è capire come debba essere intesa la previsione e in che modo deve essere calata concretamente nella vicenda in questione. Appare condivisibile l'opinione per la quale è imprescindibile distinguere tra la previsione *in concreto* dell'evento, presupposto ineliminabile della volontà dell'evento⁸⁹ e la previsione *in astratto* dello stesso. Solo nel primo caso si rinverrebbero gli estremi del dolo eventuale, là dove, cioè, “può rimproverarsi l'agente di avere avuto ben chiaro davanti agli occhi, ammonitore, l'evento tipico e di non aver desistito dal procedere”⁹⁰, configurandosi invece, nel secondo caso, una colpa con previsione⁹¹. Tale previsione, tuttavia, riguarda solo l'elemento della rappresentazione, richiedendosi sempre, per integrare il dolo, la “piena accettazione ovvero volontà, seppure in forma eventuale, non del rischio, ma della lesione o dell'evento”⁹². Da un lato previsione in concreto dell'evento, che è considerato come concretamente verificabile, dall'altro l'accettazione del rischio del suo verificarsi⁹³.

Applicando queste elaborazioni nel caso di specie la questione, tuttavia, si complica. Occorre chiedersi in che modo l'utente possa prevedere l'evento e cosa vuol dire che abbia accettato il suo verificarsi. Prima ancora di poter affermare che Mary abbia previsto l'evento in concreto o in astratto, si deve ricordare la peculiarità della tecnologia implementata nelle interfacce. Si è sempre in presenza di un algoritmo ad apprendimento automatico: il suo comportamento può essere impreveduto. Il problema dunque si rinviene proprio nel significato da attribuire alla previsione dell'evento. Si dovrebbe sostenere che l'utente ha previsto quel determinato comportamento in quel determinato contesto. Stabilire, poi, se vi sia stata accettazione o meno è altrettanto arduo: in che momento l'utente accetta che si verifichi l'evento?

⁸⁷Art. 43 c.p.: “il delitto è doloso [...] quando l'evento [...] è preveduto e voluto dall'agente [...]”

⁸⁸Tra le circostanze aggravanti comuni, al numero 3 è indicato “l'aver, nei delitti colposi, agito nonostante la previsione dell'evento”.

⁸⁹Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 229.

⁹⁰*Ivi*, pag. 230. Si veda anche Padovani, *Diritto penale*, 2008, pag. 202.

⁹¹Questa tesi trova conferma in alcune pronunce della Cassazione: in particolare, vd. C. II, 1.9.1994, C. I, 24.2.1994.

⁹²Padovani, *Diritto penale*, 2008, pag. 202.

⁹³È stato comunque rilevato che l'accettazione del rischio da parte dell'agente inevitabilmente presuppone che il verificarsi dell'evento sia possibile in concreto e comunque caratterizzato da un'alta probabilità statistica. Vd. C. V, 17.9-1.12.2008.

Nell'istante in cui si *immagina* di compiere l'azione o nel momento in cui vede le protesi muoversi e non interrompe il movimento? Qualunque sia la risposta, si deve rilevare come sia presente, tuttavia, un *quid* che può minare alla base il tentativo di ricondurre l'intenzione del soggetto dotato di interfaccia all'una o all'altra tipologia di dolo. Ciò che qui viene messo in discussione, infatti, è l'essenza stessa del dolo così come inteso finora. Se davvero è richiesta la volontà e rappresentazione da parte di chi agisce, anche in questo caso si giungerebbe all'esito per il quale vi è rappresentazione, dal momento che comunque l'utente immagina l'evento, ma diventa problematico e incerto capire se lo vuole. In soccorso alla prova della presenza dell'elemento doloso potrebbero venire in rilievo i c.d. contrassegni del dolo, o indici rivelatori, di carattere oggettivo, che, partendo dal fatto, prevedono un percorso di ricostruzione dell'elemento psicologico, attraverso l'utilizzo di massime d'esperienza. In aggiunta, ci si può rifare a indicatori di carattere soggettivo, attinenti al soggetto agente⁹⁴. Alla luce della realtà delle BCI, tuttavia, ci si può domandare, anche in modo provocatorio, se il soggetto agente al quale le massime possono riferirsi sia l'utente o l'interfaccia. Si potrebbe arrivare ad enucleare regole d'esperienza di carattere soggettivo che non si riferiscono al soggetto, ma alla "personalità" dell'algoritmo, alla "personalità" del Golem, che da "cosa" strumentale all'azione diviene "chi" agisce. Nell'ambito delle interfacce potrebbero, comunque, essere individuati segnali come l'esperienza pregressa dell'utente, eventuali funzionamenti "anomali", nonché la stessa percezione di agire da parte del soggetto: dovrebbero essere considerati degli elementi indicativi oggettivi e soggettivi, interni ed esterni al soggetto.

Considerare in modo isolato la singola interazione dalla quale deriva l'evento significa ricostruire solo parzialmente l'intera dinamica di simbiosi tra algoritmo e utente, con possibili ricadute sulla responsabilità del soggetto, rifiutando di prendere atto che la realtà delle BCI, come ormai si è capito, è diversa e impone, pretende di tenere conto che qui ci si trova innanzi a qualcosa di artificiale, che si slega tuttavia dal suo artefice per interagire quasi *personalmente* con il suo utente. Certo è che, inevitabilmente, si deve indagare il funzionamento dell'interfaccia e il rapporto con il suo utilizzatore. Il difensore dell'utente imputato avrebbe gioco facile a costruire la propria strategia difensiva proprio sulla incertezza che caratterizza la triplice relazione tra BCI-intenzione-attività cerebrale della quale prima si è parlato. In

⁹⁴Alcuni indicatori di questo tipo sono individuati dalle Sezioni Unite nella sentenza 38343/2014 (Sentenza *Thyssenkrupp*): tra questi, ad esempio, la storia e le precedenti esperienze dell'agente, la sua personalità, il fine della condotta, la sua motivazione di fondo, i tratti di scelta razionale che sottendono la condotta.

attesa di ulteriori progressi nello sviluppo delle interfacce, che forse possono aiutare l'interprete, si deve prendere atto che il *Responsibility Gap* porta con sé un altro, profondo *gap*: un *gap* epistemologico⁹⁵, nel quale cade ogni tentativo di giungere alla verità con gli strumenti tradizionali. Dovrebbe essere possibile, ad esempio, ricostruire la “cronologia” degli impulsi cerebrali inviati dal soggetto, e conservarla in appositi server, con evidenti ricadute sul piano della privacy. Il tema della conservazione dei dati personali ricavabili dalla decodificazione dell'attività neurale sarà ripreso nel capitolo cinque.

3.5.3 L'intenzione nell'ordinamento penale italiano. La colpa

Per quanto riguarda un possibile elemento di colpa, non si giungerebbe a considerazioni tanto diverse.

Perché un delitto sia colposo l'art. 43 richiede che l'evento, anche se preveduto, non sia voluto. Ci si deve chiedere ancora cosa può prevedere e cosa può volere l'utente di BCI. Sotto il profilo della previsione, valgono le stesse considerazioni proposte in riferimento al dolo eventuale. La previsione dell'evento si dovrebbe tradurre nella previsione del funzionamento delle protesi in quella determinata modalità con quella determinata finalità: se ad esempio da un utente con esperienza pluriennale nell'utilizzo delle interfacce ci si potrebbe aspettare che egli effettivamente possa prevedere l'esito di rappresentarsi mentalmente quella immagine, è innegabile che vi è un elemento di imprevedibilità che esula dall'*id quod plerumque accidit*. Il modello di causalità recepito dall'art. 40⁹⁶ del codice penale, fondato sulla teoria della *condicio sine qua non*, si scontra con l'imprevedibilità dell'attività di *machine-learning* degli algoritmi delle interfacce.

Altro elemento richiesto dalla disposizione citata è la derivazione causale dell'evento da un comportamento di negligenza, imprudenza, imperizia o posto in essere in inosservanza di leggi, regolamenti, ordini e discipline. Anche questo profilo presenta alcune criticità. Qualificare la condotta come colposa significa riconoscere la violazione di una o più regole cautelari. La dottrina maggioritaria ritiene che per ricavare una regola cautelare bisogna

⁹⁵C. Bublitz et al., “Legal liabilities of BCI-users: Responsibility gaps at the intersection of mind and machine?”, 2018.

⁹⁶Art. 40 c.p.: “Nessuno può essere punito per un fatto preveduto dalla legge come reato, se l'evento dannoso o pericoloso, da cui dipende la esistenza del reato, non è conseguenza della sua azione od omissione. Non impedire un evento, che si ha l'obbligo giuridico di impedire, equivale a cagionarlo.”

far riferimento al parametro del c.d. agente modello⁹⁷. Quest'ultimo sarebbe da individuare non già tenendo conto del “concreto modello di agente” ma su quello “dell'azione modello che avrebbe dovuto essere tenuta”⁹⁸.

Il problema di tenere conto di questo parametro nel valutare la condotta dell'utente di BCI risiede proprio nell'individuare l'azione modello da tenere. Si ripensi alla vicenda descritta. Domandarsi quale azione doveva o non doveva, essere tenuta è fuorviante. In questo caso, infatti, la vera domanda da porsi sarebbe: quale *pensiero* doveva essere riprodotto nella mente dell'utente? Qui, paradossalmente, non si dovrebbe andare alla ricerca dell'azione modello di cui parla la dottrina, ma del *pensiero modello* che doveva essere pensato. Anche ammesso che si possa utilizzare comunque il parametro suddetto, quale regola cautelare è stata violata dall'utente? La regola di *non pensare* determinate immagini in uno stato di alterazione emotiva? Esiste effettivamente un pensare imprudente o dato dalla imperizia? Certo, anche in questo caso da un utente esperto si potrebbe aspettare un maggiore controllo, ma qui non si parla del controllo sulle proprie azioni, che può venire meno nell'impeto emotivo dello stato di alterazione psicofisica. Si parla del controllo sui propri *pensieri*. I soggetti dotati di BCI, infatti, dovrebbero sopprimere determinati pensieri, evitare che il flusso di immagini mentali e il discorso interiore possano indurli a entrare in stati emotivi tali da influenzare i segnali cerebrali che giungono alle interfacce. È una attività che presenta molte criticità, a partire dal piano pratico. È stato dimostrato che evitare intenzionalmente di pensare a qualcosa è assai difficile per la persona⁹⁹. Il controllo che si dovrebbe esercitare sulla propria mente ha anche delle ricadute normative dirimpenti: si dovrebbero presidiare dei veri e propri doveri mentali, sanzionando eventuali violazioni. In futuro, alla fine, la freddezza emotiva potrebbe divenire un dovere giuridico¹⁰⁰.

Ancora più paradossale sarebbe l'affermazione che la condotta dell'utente ha violato leggi, regolamenti, ordini o discipline. Ci si deve chiedere quale norma giuridica, scritta o meno, sia stata violata, e in che modo. Anche ammesso che vi sia una disposizione normativa violata, si arriverebbe all'ipotesi estrema nella quale la violazione è avvenuta con il semplice *pensiero*,

⁹⁷È il criterio del *homo eiusdem professionis et condicionis*. Si veda, tra gli altri, Fian-daca, *Diritto penale. Parte generale*, 2019, pag. 579, Marinucci, *Manuale di diritto penale. Parte generale*, 2015, pag. 300.

⁹⁸Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 275.

⁹⁹Il soggetto porrebbe in essere un attività di controllo riconducibile agli *ironic processes of mental control*. Si veda Wegner, “Ironic processes of mental control.”, 1994, ma anche Wenzlaff e Wegner, “Thought suppression”, 2000.

¹⁰⁰Così C. Bublitz et al., “Legal liabilities of BCI-users: Responsibility gaps at the intersection of mind and machine?”, 2018.

con buona pace del principio di materialità, cardine del rapporto di causalità e del moderno diritto penale liberale.

Alla fine, il fiume di questioni relative all'intenzione dell'utente non fa altro che portare a valle un problema che si individua in primo luogo a monte. L'interiorità diviene esteriorità. Il mero pensiero cambia la realtà esterna. È una ricaduta radicale che travolge molto di più delle teorie dell'elemento psicologico del reato. Siamo d'innanzi a qualcosa che non ha un precedente nella storia, che rende molto arduo individuare una esperienza pregressa che può guidare l'interprete nella qualificazione giuridica del fatto e nell'enucleare delle utili massime d'esperienza, di qualunque natura esse siano. Il Golem sta compiendo rapidi passi verso di noi, ma il percorso non è ancora così chiaro e battuto da poter costruire una segnaletica che evita di perdersi nella selva del progresso.

Fin qui si è parlato solo dell'elemento psicologico del reato, sotto un profilo doloso e colposo. Le considerazioni presentate fino a questo punto si fondano su un presupposto implicito e scontato, che tuttavia potrebbe non essere così ovvio alla luce dell'esperienza delle interfacce: dolo o colpa caratterizzano la condotta di colui o colei che dunque deve essere un soggetto *agente*. Così come nel più articolato sistema di scatole cinesi, il problema dell'individuazione dell'intenzione altro non è che una grande scatola che racchiude in sé un'altra problematica. Sorge infatti un'altra domanda: l'utente di BCI è effettivamente un soggetto agente che pone in essere una condotta, o anche lo stesso concetto di condotta si erode innanzi alla marea dell'esperienza delle interfacce?

Il *gap* sta diventando ancora più profondo.

3.6 L'azione

Il difensore di Mary, giurista particolarmente attento ai più recenti sviluppi delle neuroscienze afferma, dinanzi al giudice che la donna non ha tenuto alcuna condotta illecita per il semplice, quanto paradossale, fatto che non c'è stata azione. Certo, le braccia si sono mosse, il coltello è stato piantato, il marito è morto. Mary, tuttavia, non ha compiuto alcuna azione. Queste affermazioni, che potrebbero essere giudicate come espressione di relativismo forense o di stravolgimento della realtà per meri fini di difesa, potrebbero non essere così infondate come appaiono.

In letteratura, infatti, è sorto con sempre maggiore rilevanza il quesito circa le ricadute dell'interazione tra utente e macchina, o, meglio, tra cervello dell'utente e algoritmo, sul concetto di azione sia sotto un profilo

neuroscientifico e filosofico sia sotto il profilo giuridico. Ancora una volta il punto di partenza è l'analisi del primo.

3.6.1 Cosa è azione?

Alla base della questione vi è una considerazione tanto facile da rilevare quanto gravida di conseguenze: le parti del rapporto che caratterizza la dinamica di connessione tra uomo e BCI, in questo caso in quella tra l'interfaccia di Mary e le protesi robotiche, sono da un lato l'algoritmo, dall'altro il cervello con i suoi impulsi più o meno volontari. I grandi esclusi dall'interazione sono i muscoli degli arti che, in condizioni di normalità, dovrebbero muoversi. A ben vedere, infatti, l'intera sequenza può *bypassare* il corpo e l'apparato muscolare. Ecco allora che, fondamentalmente, un “*motionless body, the epitome of inaction, might be acting*”¹⁰¹. Un corpo che non si muove, cioè, può provocare effetti nell'ambiente esterno circostante nell'ambito di quelli che in letteratura sono stati definiti *BCI-mediated events*¹⁰².

Prima di scendere più in profondità nell'analisi, è possibile innanzitutto ricordare che accanto alle azioni condotte attraverso un movimento corporeo, si conoscono anche i c.d. *mental acts*, ossia degli “atti” che non sono svolti attraverso l'intervento del corpo, ma solo mentalmente¹⁰³, generalmente determinano una disposizione mentale o uno stato mentale¹⁰⁴. Gli utenti sarebbero in grado allora di provocare del movimento compiendo proprio questi atti mentali utilizzando interfacce attive o reattive¹⁰⁵, con ciò diventando dei veri e propri *agenti incorporati*¹⁰⁶.

¹⁰¹Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019, pag. 457. Nella trattazione successiva si riprenderà più volte questo articolo.

¹⁰²*ivi*. Gli autori utilizzano questa espressione in luogo dei termini “azione” o “movimento” in quanto permette di includere al suo interno sia i vari cambiamenti nell'attività degli algoritmi delle BCI sia i movimenti corporei causati da queste ultime.

¹⁰³In altre parole: “*there is no motor output to be controlled and no sensory input vector that could be manipulated by bodily movement.*” Vd. Metzinger, “The problem of mental action”, 2017, pag. 1. Altri tuttavia sostengono che queste “azioni mentali” non sono altro che dei meri accadimenti che riguardano il soggetto. Così in particolare Strawson, “XI—Mental ballistics or the involuntariness of spontaneity”, 2003.

¹⁰⁴Proust, “A plea for mental acts”, 2001.

¹⁰⁵Le passive pongono problemi ancora più complessi, in quanto il soggetto non compirebbe nessuna azione, né mentale né fisica, dal momento che, com'è stato evidenziato nel capitolo precedente, si limitano a tradurre in output la mera attività cerebrale, sia essa volontaria o meno.

¹⁰⁶Himmelreich, “Agency and embodiment: Groups, human-machine interactions, and virtual realities”, 2018. Come è stato sostenuto, “[...]BCIs change the interactions between the CNS and its environment. The new BCI outputs that the BCI creates are fundamentally

Al di là della posizione di coloro che considerano questa interazione “extra-muscolare” una liberazione dai limiti imposti dalla natura all’uomo¹⁰⁷, indiscutibili problemi si pongono in relazione alle tradizionali ricostruzioni dell’azione.

Il punto di partenza riprendere un’interessante distinzione tra i c.d. *basic acts*, e i *non-basic acts*. Ai primi¹⁰⁸ si riconducono quelle azioni che non vengono eseguite attraverso un’altra azione, ad esempio muovere un braccio. La seconda categoria, invece, si riferisce a tutte quelle azioni che inevitabilmente sono condotte e portate a compimento facendo dell’altro: un esempio si può ritrovare nell’azione di guidare una macchina, che viene condotta attraverso l’atto di muovere il volante, cambiare le marce e premere i pedali. Questa distinzione porta al primo punto utile per poter ricostruire il problema dell’azione alla luce dell’esperienza delle BCI perché, com’è stato notato¹⁰⁹, una medesima azione volta al medesimo scopo e che produce la medesima conseguenza potrebbe essere considerata basica in un contesto “normale”, non basica se eseguita mediante interfaccia. Il soggetto, infatti, esegue comunque l’azione compiendone un’altra, ossia il *pensare* di muovere la BCI, così da controllarla e inviare a quest’ultima i relativi segnali cerebrali¹¹⁰.

Detto ciò, cercando di capire se quelle poste in essere da un utente siano effettivamente azioni, le tensioni si colgono in relazione alla c.d. teoria standard dell’azione, o teoria causale dell’azione, che, come si sa, trova in Donald Davidson il suo autore¹¹¹. In breve, in base a tale teoria si considera azione un avvenimento che è causato, secondo determinate modalità, da un

different from natural CNS outputs, which come from spinal motoneurons.” Vd.J. R. Wolpaw e E. W. Wolpaw, “Brain-computer interfaces: something new under the sun”, 2012, pag. 11.

¹⁰⁷Nicolelis, *Beyond boundaries: The new neuroscience of connecting brains with machines—and how it will change our lives*, 2011, pag.315. Le ricadute etico-filosofiche di tale concezione saranno esaminate più avanti, in particolare nel capitolo 4.

¹⁰⁸Cfr, Goldman, “Theory of human action”, 2015 e Danto, *Analytical philosophy of action*, 1973. Per una opinione più critica circa la loro esistenza, invece, si veda Lavin, “Must there be basic action?”, 2013. Per la loro analisi seguono Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

¹⁰⁹Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

¹¹⁰Si pensi ad esempio all’azione di muovere una mano. In un contesto normale una persona esegue il movimento. Se dotato di BCI, attiva o reattiva che sia, invece, l’utente dovrebbe pensare al movimento, così che questo sia eseguito ad esempio dalla protesi robotica. In questo modo, dunque, compie l’azione non basica solo compiendone un’altra, basica, ossia l’azione mentale di inviare i segnali cerebrali all’interfaccia.

¹¹¹Senza pretese di completezza e senza entrare in ambiti che esulano dall’ambito di ricerca di questo elaborato, si rimanda a Davidson, *Essays on Actions and Events: Philosophical Essays Volume 1*, 2001.

determinato stato mentale, in base a una determinata intenzione, indipendentemente dal fatto che la persona abbia o meno coscienza di quest'ultima. Riprendendo un'utile interpretazione¹¹², tale concezione si può scomporre essenzialmente in due concetti. Il primo attiene alla natura dell'azione. In particolare, un avvenimento è un'azione quando è causato in modo “corretto” da “corretti” stati mentali. In forza del secondo, invece, essendo lo stato mentale del soggetto la causa dell'azione, questo è idoneo a fornire una “ragione” all'azione, ossia a razionalizzarla.

Riferendo tali considerazioni alla dinamica delle interfacce neuronali, ecco che il “corretto” stato mentale (leggi: intenzione) per il compimento dell'azione si coglie solo nell'utente dotato di interfacce attive o reattive, e non, invece, anche in colui che è dotato di quelle passive.

La conseguenza ultima, dunque, è che solo il primo effettivamente ha compiuto un'azione, mentre i *BCI-mediated events* nel secondo caso non possono qualificarsi come tali¹¹³.

Nel versante giuridico il discorso si amplia ulteriormente, dal momento che la concezione filosofica si intreccia sia con la concezione normativa di condotta, con tutti i suoi elementi costitutivi, sia con la ricostruzione di azione alla luce dei progressi nelle neuroscienze.

3.6.2 La condotta negli ordinamenti penali

Gli ordinamenti di *common law* e *civil law* si incontrano, seppure con le debite differenze sistematiche, su una ricostruzione molto simile di condotta giuridicamente rilevante.

Il *Model Penal Code* americano, che si può prendere come primo punto di riferimento, definisce “azione” un “*bodily movement whether voluntary or involuntary*”¹¹⁴. Tuttavia, solo in caso di azione o condotta volontaria o di omissione di un'azione di cui è fisicamente capace è possibile considerare l'agente responsabile¹¹⁵. Si rende necessario, allora, specificare quando un'azione si può definire volontaria o quando non lo sia. In negativo, infatti, non si possono intendere come azioni volontarie giuridicamente rilevanti convulsioni e riflessi motori automatici, movimenti corporei eseguiti durante

¹¹²Schlosser, “Agency, ownership, and the standard theory”, 2011.

¹¹³Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

¹¹⁴*US Model Penal Code*, §1.13,2.

¹¹⁵In particolare: “A person is not guilty of an offense unless his liability is based on conduct which includes a voluntary act or the omission to perform an act of which he is physically capable.” *Ivi*, § 2.01,1.

il sonno o in stati di ipnosi¹¹⁶. Perché vi sia un'azione volontaria, dunque, vi deve essere un movimento corporeo voluto, o, in altre parole, *causato* da una volontà¹¹⁷. Nel caso di un soggetto dotato di BCI, tuttavia, vi è stato un movimento. ma questo può non essere corporeo (come nel caso di movimento di una carrozzella o di una protesi esterna) e non sempre volontario.

In un ordinamento di *civil law* quale quello italiano, la disciplina non muta così radicalmente. L'art. 42, co. 1, del codice penale subordina ogni addebito di responsabilità a una azione o omissione commessa con “coscienza e volontà”¹¹⁸. Da un lato, dunque, azione e omissione, dall'altro coscienza e volontà.

Com'è noto, azione e, in negativo, l'omissione devono essere intese innanzitutto “come un moto e/o una inerzia corporei coscienti e voluti [...] il movimento corporeo dell'uomo in tanto è azione, in quanto è voluto ed in quanto è cosciente; e uno stato d'inerzia corporea, in tanto è una inazione in quanto è voluto e cosciente”¹¹⁹. Vi deve essere, in altre parole, un comportamento umano dominato o dominabile dalla volontà¹²⁰. La qualificazione della condotta in tal senso rispecchia il concetto classico di condotta, che è recepito dallo stesso codice penale. Tale concezione, tuttavia, deve essere letta alla luce del concetto c.d. sociale di azione, per il quale a caratterizzare la condotta giuridicamente rilevante è un ineliminabile elemento di socialità, inteso come “risposta dell'uomo ad una situazione, conosciuta e conoscibile, attraverso l'attuazione di una possibilità di reazione a sua disposizione

¹¹⁶“*The following are not voluntary acts within the meaning of this Section: (a) a reflex or convulsion; (b) a bodily movement during unconsciousness or sleep; (c) conduct during hypnosis or resulting from hypnotic suggestion; (d) a bodily movement that otherwise is not a product of the effort or determination of the actor, either conscious or habitual.*”
ivi, § 2.01,2.

¹¹⁷M. S. Moore, *Act and crime: The philosophy of action and its implications for criminal law*, 2010.

¹¹⁸Art. 42 c.p.: “*Nessuno può essere punito per un'azione od omissione preveduta dalla legge come reato, se non l'ha commessa con coscienza e volontà. Nessuno può essere punito per un fatto preveduto dalla legge come delitto, se non l'ha commesso con dolo, salvi i casi di delitto preterintenzionale o colposo espressamente preveduti dalla legge. La legge determina i casi nei quali l'evento è posto altrimenti a carico dell'agente, come conseguenza della sua azione od omissione. Nelle contravvenzioni ciascuno risponde della propria azione od omissione cosciente e volontaria, sia essa dolosa o colposa.*”

¹¹⁹Rocco e Italia, *Lavori preparatori del codice penale e del codice di procedura penale. vol. 5: progetto definitivo di un nuovo codice penale con la relazione del guardasigilli Alfredo Rocco*, 1929.

¹²⁰Cocco, *Trattato breve di diritto penale. Parte generale, 2018*, pagg. 41 ss. L'Autore precisa nella stessa pagina che il comportamento umano deve essere inteso come *actus humanus*, non *actus hominis*: un atto nel quale la volontà è lo “*specifico umano, quale esternazione o obiettivazione di una persona nella società*”.

secondo la sua libertà”¹²¹. Non solo. Perché vi sia un’effettiva azione o omissione, è necessaria la *suitas* della condotta, ossia il “dominio o controllo potenziale” sulla stessa da parte dell’agente¹²². È questa un’impostazione ormai assodata e conosciuta da ogni interprete, così come confermato anche sul versante giurisprudenziale: “[...] La coscienza e volontà come dominio anche solo potenziale dell’azione od omissione rappresentano il livello iniziale di considerazione della condotta umana rilevante per il diritto penale”¹²³. È ormai pacifico che il difetto di *suitas* è difetto di un elemento essenziale del fatto e non dell’elemento psicologico del reato¹²⁴.

In definitiva, riprendendo le ricostruzioni giuridiche di condotta penalmente rilevante, anche in riferimento alla disciplina di *common law* d’oltreoceano prima descritta, si possono individuare degli elementi costitutivi dai quali iniziare l’indagine.

3.6.3 Il movimento corporeo

La condotta è un comportamento umano inteso come movimento corporeo, che deve essere volontario (o cosciente) e sul quale l’agente esercita un controllo (*suitas*) anche potenziale. Ognuno di questi profili deve essere valutato alla luce dell’impatto delle BCI.

Innanzitutto, vi sono due possibili ipotesi ricostruttive che possono essere individuate e ciascuna di esse apre a due distinti ordini di conseguenze.

Per analizzare la prima si deve fare un passo indietro per ricordare di come prima si sia fatto riferimento ai c.d. agenti scorporati, o *disembodied agency*. Per mantenere la tradizionale concezione di condotta si dovrebbe rigettare questa definizione per abbracciare un’altra interpretazione: se effettivamente alla base si allontanasse l’idea che le BCI siano qualcosa di altro dal soggetto, per accogliere, invece, l’opinione secondo la quale l’utente non è un soggetto dotato di interfacce, ma è tale *insieme* ai neurodispositivi, si può parlare di agenti *re-incorporati*, che pongono in essere quella che si può definire una *re-embodied agency*¹²⁵. In questa diversa prospettiva la

¹²¹Così Jescheck e Weigend, *Lehrbuch des Strafrechts: allgemeiner Teil*, 1996, così come riportato in Cocco, *Trattato breve di diritto penale. Parte generale*, 2018, pag. 40. Per una opinione critica della teoria sociale dell’azione, si veda Pagliaro, *Il reato*, 2007, pag. 71. Nel suo scritto Pagliaro ritiene che la teoria sociale dell’azione “disperde il *quid propri* dell’azione”.

¹²²Romano e Grasso, *Commentario sistematico del Codice penale: Art. 1-84*, 1987.

¹²³Così C, I, 1-17.7.2008.

¹²⁴Si veda, tra gli altri, Pagliaro, *Principi di diritto penale: parte generale*, 2003, nonché Bettiol, *Diritto penale: parte generale*, 1982.

¹²⁵Così C. Bublitz et al., “Legal liabilities of BCI-users: Responsibility gaps at the

tradizionale concezione di condotta sarebbe mantenuta. Non vi è più un dualismo tra uomo e macchina, ma le due entità sono riunite, sia in forza di una *fictio iuris*, sia fisicamente, in una unica persona giuridica¹²⁶.

Vi sono, tuttavia, delle rilevanti implicazioni che sorgono da questa conclusione. È stato infatti notato¹²⁷ che se da un lato si salverebbe l'impianto tradizionale della categoria della condotta penalmente rilevante, dall'altro questo verrebbe comunque deformato a tal punto da portare a poter affermare che l'utente si re-incorpora nel cursore mosso su uno schermo, almeno là dove il funzionamento dell'interfaccia preveda l'utilizzo di un software, come avviene, appunto, in caso di movimento del cursore attraverso il controllo di un'applicazione. Inoltre, lo stesso concetto di "movimento corporeo" andrebbe soggetto a una radicale evoluzione per il semplice fatto che sarebbe trasformata la stessa idea di "corpo". Il nuovo corpo umano si estenderebbe oltre i limiti imposti dalla natura, potendo andare al di là del *qui* e dell'*ora*, ponendo anche alcuni potenziali conflitti tra giurisdizioni. Si pensi ad esempio all'esperimento nel quale una scimmia rinchiusa in un laboratorio della *Duke University* della North Carolina ha mosso, attraverso l'ausilio di una BCI, un robot situato a Kyoto, in Giappone¹²⁸. Sorge il quesito: dove ha agito la scimmia? La risposta, se si mantiene la tradizionale concezione di condotta e si abbraccia l'idea di un agente re-incorporato nelle interfacce, è che l'animale ha agito (ha cioè posto in essere un movimento corporeo volontario) in America e in Giappone. Sarebbe un esemplare di scimmia molto particolare in quanto dotato di un rivoluzionario dono dell'ubiquità.

Un secondo approccio prevede, invece, il superamento della tradizionale concezione di condotta. In particolare, in letteratura già molti sostengono il bisogno di superare il concetto di azione come movimento corporeo che si estrinseca nel mondo esterno per considerare azione, invece, un "atto mentale volontario che incide sul mondo esterno"¹²⁹. L'azione dunque diventa pensiero. Anche questa ricostruzione presenta dei notevoli profili

intersection of mind and machine?", 2018. Per una disamina in chiave etico-filosofica, nonchè per le ricadute sul senso di *agency*, si rimanda ai capitoli successivi.

¹²⁶In questo caso alcuni definiscono le interfacce come *quasi-bodily devices*. Così Steinert et al., "Doing things with thoughts: Brain-computer interfaces and disembodied agency", 2019.

¹²⁷*ivi*

¹²⁸Per l'esperimento, vd. Nicoletis, *Beyond boundaries: The new neuroscience of connecting brains with machines—and how it will change our lives*, 2011.

¹²⁹Cfr Rainey, Maslen e Savulescu, "When thinking is doing: Responsibility for BCI-mediated action", 2020, e Stefano Fuselli, *Brain-Computer Interface e soggettività agente. Considerazioni etico-giuridiche* in Moro, *Etica, diritto e tecnologia. Percorsi dell'informatica giuridica contemporanea*, 2021.

critici. Se da un lato, infatti, si terrebbe conto dell'interazione interfaccia-cervello dell'utente, dall'altro sono evidenti le frizioni con il principio di materialità richiesto dal moderno diritto penale liberale. Il brocardo latino *nemo cogitationis poena patitur* verrebbe trasportato via dalla corrente del progresso.

3.6.4 La volontarietà

Altro elemento da considerare in relazione alla nozione di condotta attiene al suo necessario requisito di volontarietà. Anche qui si potrebbero proporre due considerazioni tra loro alternative, comunque strettamente interconnesse con quelle presentate subito sopra.

Se si mantiene la concezione tradizionale di condotta, considerando interfaccia e utente come un *unicum*, riemergono i problemi evidenziati nel paragrafo precedente in riferimento all'intenzione. Dovrebbe essere considerata volontaria, infatti, l'attivazione della BCI da parte del soggetto: sorgerebbe la necessità di indagare la dinamica di interazione tra i due, conducendo una indagine che si scontrerebbe con l'incertezza che caratterizza la rilevazione dei segnali cerebrali da parte dell'algoritmo.

Se, al contrario, si arrivasse a definire l'azione come un atto mentale volontario, ecco che i nodi si individuerebbero sul piano epistemologico o, in una prospettiva giuridica, sul versante probatorio, nonché sulle ricadute etiche. Si dovrebbe cioè accertare la volontarietà del pensare che diviene agire. Anche là dove effettivamente si riuscisse a dimostrare che il pensiero è stato voluto dalla persona, la società dovrebbe essere pronta ad aprirsi alla possibilità di considerare un soggetto penalmente responsabile sulla base di una condotta letteralmente immaginata.

3.6.5 Il controllo

Infine, l'ultimo profilo da illuminare attiene all'individuazione della *svitas*, alla possibilità, cioè, che il soggetto effettivamente eserciti un controllo sulla propria condotta.

Nell'analisi di questo terzo e ultimo aspetto un possibile aiuto lo si trova, ancora una volta, nella letteratura neuroscientifica. In questo modo, infatti, è possibile confrontarsi con il problema evidenziando le differenze di interazione che intercorrono tra le varie tipologie di BCI. Un approccio che, al contrario, ignorasse o comunque ridimensionasse eccessivamente il

contributo scientifico porterebbe a una riflessione giuridica che, alla lunga, può rivelarsi grossolana e anacronistica¹³⁰.

Più nello specifico, dunque, ci si può rifare a una distinzione, emersa in letteratura¹³¹, in base alla quale si dovrebbero distinguere tre possibili “classi” di controllo che può porre in essere l’utente: il controllo sull’avvio dell’azione, o *executory control*, quello sull’azione nel momento del suo svolgimento *guidance control* e, infine, il controllo “in negativo” dato dal controllo esercitato attraverso la funzione di veto¹³².

Partendo dal primo, tipico delle BCI attive e reattive è il fatto che l’utente dà, o dovrebbe dare, avvio al movimento impegnandosi nella relativa attività mentale, come può essere quella necessaria per controllare il movimento di una carrozzina o concentrarsi sull’obiettivo di muovere le protesi. Ciò agevolerebbe la ricostruzione dell’azione da parte dell’interprete, dal momento che appare ragionevole ritenere che l’utente abbia agito: troppe energie sono necessarie per poter affermare che il soggetto non ha avuto alcun ruolo nel compimento dell’azione. È stato rilevato allora come, paradossalmente, gli atti compiuti esercitando questo tipo di controllo si avvicinano di più all’idea di azione secondo la teoria causale di Davidson rispetto alle azioni “normali”. Queste interfacce tradurrebbero, infatti, solo i “corretti” stati mentali richiesti dalla concezione standard dell’azione. Di conseguenza si individuerrebbe il legame causale tra stato mentale e movimento delle BCI¹³³. In definitiva, dunque, un soggetto dotato di interfaccia attiva o passiva potrebbe aver avuto un controllo sufficiente sulla condotta per poter addebitargli l’azione. Si ritiene che l’*executory control* ha una rilevanza tale che là dove non vi sia, non si può affermare che vi sia azione secondo la teoria standard¹³⁴. Ad esempio, tornando al caso di Mary e delle sue braccia, questo tipo di comando si potrebbe affermare che sia “sostituito”

¹³⁰Certo è che comunque è necessario chiedersi fino a che punto può spingersi la centralità della conoscenza scientifica, in questo caso neuroscientifica, prima di doversi arrestare innanzi all’ambito proprio della scienza giuridica.

¹³¹Ad esempio, Shepherd, “Conscious control over action”, 2015, Pacherie, “The sense of control and the sense of agency”, 2007 e Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

¹³²In relazione a quest’ultimo si possono enucleare delle riflessioni simili a quelle già portate alla luce in precedenza nell’ambito dell’analisi dell’intenzione. Negli attuali dispositivi neurali, se è implementata la possibilità di avviare l’azione, dovrebbe essere possibile anche esercitare la funzione di veto, nonostante le difficoltà che sono state evidenziate nelle pagine che precedono. Proprio per questo, dunque, per la ricostruzione del veto e della sua funzione si rimanda a queste ultime.

¹³³Riprendo ancora Steinert et al., “Doing things with thoughts: Brain-computer interfaces and disembodied agency”, 2019.

¹³⁴*ivi*.

dall'emozione e dallo stato emotivo in cui si trovava. Verrebbe a mancare il controllo, verrebbe a mancare l'azione¹³⁵. Certo è che la questione si complicherebbe ulteriormente nel caso in cui l'utente sia dotato di interfacce nelle quali è implementata la più recente tecnologia basata su algoritmi predittivi che, affinando ulteriormente la sensibilità delle BCI ai segnali neurali, possono pregiudicare e rendere assai più arduo esercitare, e capire se è stato esercitato, un controllo da parte del soggetto¹³⁶. In questo caso, infatti, l'interfaccia *bypasserebbe* il controllo sull'inizio dell'azione causando comunque il movimento¹³⁷.

Il secondo tipo di controllo con il quale ci si può confrontarsi è, come detto, esercitato sull'azione *in fieri*. Questo si potrebbe rivelare particolarmente importante ai fini della trattazione dal momento che è stato sostenuto, infatti, che proprio la presenza o meno del controllo sull'azione in svolgimento permette di distinguerla da un mero evento. In particolare, si ha azione “[...] when we are somehow in touch with the mechanisms [...] by which our movements are guided and their course guaranteed”¹³⁸. Per qualificare tale declinazione del controllo, alcuni ritengono che si individua là dove un atto è causato da propri “meccanismi”¹³⁹ interni o se è dovuto a quelli che sono stati definiti “*moderate reason-responsive mechanisms*”¹⁴⁰. Decisivo è chiarire proprio il rapporto dell'utente con questi meccanismi interni. Da una parte, certamente, non tenere in considerazione il carattere artificiale e imprevedibile delle BCI risulterebbe eccessivo e ormai inadeguato alla situazione attuale, dall'altra vi è comunque chi ritiene che, alla fine, il controllo sulle azioni esercitato da una persona priva di interfaccia non si discosta tanto da quello esercitato da una che, invece, ne è dotata¹⁴¹. Così come nei

¹³⁵ “*There is an agent’s mental state or event, R, that rationalizes the performance of an action of type A. R causes an intermediary state, N, which causes a movement of type M. This movement would constitute or realize an instance of A-ing, were it not caused by N. Typically, N is a state that appears to undermine the agent’s control over the movement, such as a state of severe nervousness, and so it seems clear that the agent does not perform an action at all. But according to a causal theory, the agent performs an A-ing, as M is both caused and rationalized by R.*” Così Schlosser, “Bending it like Beckham: movement, control and deviant causal chains”, 2010, pag. 299.

¹³⁶Truccolo, Hochberg e Donoghue, “Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes”, 2010.

¹³⁷Si veda anche Wang e Jung, “A collaborative brain-computer interface for improving human performance”, 2011.

¹³⁸Così Frankfurt, “The problem of action”, 1978, pag. 158.

¹³⁹Fischer, “Responsibility and control”, 1982.

¹⁴⁰J. C. Bublitz e Merkel, “Autonomy and authenticity of enhanced personality traits”, 2009.

¹⁴¹Schönau, “The Spectrum of Responsibility Ascription for End Users of

movimenti di tutti i giorni non ci si concentra su ogni singolo passaggio delle azioni compiute, lasciando che alcuni di questi *steps* siano compiuti, per così dire, in automatico, concentrandosi solo sul risultato finale, allo stesso modo un utente di BCI può concentrarsi sul risultato finale: quello stesso *goal-selection* al quale si è fatto riferimento nella descrizione degli *outputs* delle interfacce attive, reattive e passive, lasciando che sia l'interfaccia a compiere i passaggi intermedi¹⁴². Dunque quest'ultima forma di controllo, essendo esercitabile sia in presenza sia in assenza di BCI, in quest'ultimo caso sul movimento che è corporeo in tutti i sensi, permetterebbe un maggiore dominio sulla condotta, rendendola più fluida¹⁴³. Ecco allora che nel caso in cui un soggetto sia dotato di interfaccia attiva o reattiva, il controllo sullo svolgimento dell'azione non viene svolto in maniera così innovativa e diversa da richiedere una rivisitazione del concetto di *suitas*. Ciò, tuttavia, dipende dalla tipologia di interfaccia impiantata: in futuro, in caso di moltiplicazione dei modelli di BCI potrebbe essere necessaria una evoluzione del concetto di controllo. Nel caso di specie nel quale Mary ha accoltellato il marito, bisognerebbe individuare quale tipologia di interfaccia le è stata impiantata, per capire quale tipologia di controllo ha potuto, o dovuto, esercitare.

3.7 Il Golem a processo. Il civile

L'impatto del *Responsibility Gap* così come descritto non riguarda solo l'ambito penalistico e le sue categorie giuridiche fin qui ricostruite, ma estende la sua portata sul fronte civilistico. Diviene inevitabile, infatti, una volta arrivati a questo punto della trattazione, scomporre la questione per ricostruirla alla luce della prospettiva propria del diritto civile. Per procedere in questa indagine si sottolinea l'ingresso nella questione di un secondo soggetto, o di un terzo, se si arriva all'estremo di considerare lo stesso algoritmo come *soggetto* di diritto, oltre all'utente finale dei dispositivi: il produttore dell'algoritmo delle BCI. Il punto di partenza è proprio la disamina della sua posizione giuridica.

Iniziando dalla disciplina europea, ci si può chiedere se possa sorgere una eventuale responsabilità del produttore per prodotto difettoso. La direttiva del Consiglio Europeo 85/374/CEE all'articolo 7, lett. e, infatti, nega la responsabilità del produttore se costui prova “*che lo stato delle conoscenze*

Neurotechnologies”, 2021.

¹⁴²*ivi*. Nota l'autore che l'obiezione per la quale tale ragionamento non può essere applicato a qualcosa di artificiale come le BCI può essere superata dal fatto che queste mantengono pur sempre un rapporto interattivo con l'intenzione della persona.

¹⁴³*ivi*.

scientifiche e tecniche al momento in cui ha messo in circolazione il prodotto non permetteva di scoprire l'esistenza del difetto.”

Sotto questo profilo due sono gli aspetti critici da rilevare. In entrambi i casi ancora una volta torna con prepotenza il tema della prevedibilità del “comportamento” dell’algoritmo: com’è stato evidenziato sopra, in base all’attuale stato di conoscenza l’imprevedibilità è connaturata agli stessi *machine-learning algorithms*.

In primo luogo, è stato notato¹⁴⁴ che si dovrebbe considerare anche la possibilità che si verifichi un danno imprevedibile da uso “corretto” dell’interfaccia. Si immagini ancora il caso fittizio di Mary. L’interfaccia ha funzionato correttamente sotto ogni profilo: ha rilevato gli impulsi cerebrali e li ha tradotti in azione portando al movimento delle protesi. L’interazione cervello-interfaccia-protesi è stata perfettamente coerente con lo scopo per il quale è stata ideata e programmata. Al contrario, proprio se la donna non fosse riuscita a far muovere le braccia robotiche e ad accoltellare il marito, la tecnologia potrebbe essere stata difettosa, facendo sorgere un’eventuale responsabilità per il costruttore. In caso contrario, infatti, si potrebbe quasi sostenere che quest’ultima sorga per un funzionamento *troppo* corretto della BCI.

In secondo luogo, i danni imprevedibili potrebbero essere prodotti da un malfunzionamento della tecnologia, ma lo stesso malfunzionamento potrebbe essere imprevedibile¹⁴⁵. La programmazione di un sistema ad apprendimento automatico prevede che questo sia testato per reagire a diverse situazioni e compiti che corrispondono a predeterminati *sets* di dati. L’efficienza dell’interfaccia si individua proprio nella sua capacità di relazionarsi con situazioni nuove senza uscire dai “binari” posti nella fase di programmazione. Tuttavia, una deviazione anche minima dai percorsi delineati nella fase di *training* può portare a comportamenti imprevedibili difficili da analizzare o da spiegare, sia in termini tecnici che in termini giuridici¹⁴⁶. Questa incertezza incide¹⁴⁷ inevitabilmente anche sul nesso di causalità tra danno e dispositivo: questa tensione emerge in particolar modo alla luce del regolamento europeo 2017/745. All’art. 87, infatti, si prevede che “*I fabbricanti segna-*

¹⁴⁴Sarra, “Relevant Legal Issues for Hybrid Human-Robotic Assistive Technologies: A First Assessment”, pag. 11.

¹⁴⁵Larus et al., *When computers decide: European recommendations on machine-learned automated decision making*, 2018.

¹⁴⁶Burrell, “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”, 2016.

¹⁴⁷Così Sarra, “Relevant Legal Issues for Hybrid Human-Robotic Assistive Technologies: A First Assessment”, pag. 12.

lano ogni incidente grave di cui al paragrafo 1, lettera a), immediatamente dopo aver stabilito il nesso causale, anche solo ragionevolmente possibile, fra l'incidente e il loro dispositivo, e non oltre 15 giorni dopo aver avuto conoscenza dell'incidente." È evidente che i più recenti sviluppi nella tecnologia degli algoritmi e la conseguente complessità di ricostruire il nesso causale rende assai arduo conformarsi a questa disciplina.

Per evitare di giungere dinnanzi alla situazione di *aut-aut* nella quale la responsabilità è attribuita o all'utente o al costruttore, alcuni hanno pensato di "distribuire" la responsabilità civilistica tra utente, produttore e programmatore, sulla base di modelli di responsabilità predefiniti in base ai quali delineare il concorso di ciascuno alla produzione del danno¹⁴⁸. Altri, invece, individuano come possibile alternativa la previsione di un'assicurazione obbligatoria che deve essere sottoscritta dall'utente, che copre almeno i comportamenti più anomali e "abnormi" delle BCI¹⁴⁹.

Fin qui si è parlato del concetto di azione, di condotta, di movimento corporeo, di volontarietà e di controllo sul versante penalistico, nonché di alcune ricadute sulla configurazione della responsabilità civilistica. Tuttavia, resta ancora un punto da considerare. Perché vi possa essere effettivamente un'azione, una condotta penalmente rilevante, vi deve essere un "chi" al quale riferirla. In linea con la visione antropocentrica che caratterizza l'esperienza giuridica tradizionale, parrebbe scontato individuare in questo "chi" l'uomo. Si ripensi, tuttavia, al caso della donna e delle sue protesi robotiche: quel comportamento, del quale dottrina e giurisprudenza parlano, non è più solo umano. Le protesi si sono mosse dietro iniziativa dell'algoritmo che ha decodificato gli impulsi cerebrali dell'utente. L'umano incontra il non umano. Già è messa in discussione, allora, la natura di colui al quale riferire la condotta, qualunque essa sia. Se il diritto è un sistema solare che vede l'uomo come il suo grande Sole, nel prossimo futuro l'uomo potrebbe divenire solo uno dei tanti pianeti che orbitano attorno a quello che potrebbe divenire un astro nascente della personalità giuridica: il Golem che si fa algoritmo.

¹⁴⁸Zornoza et al., "Robots Liability: A Use Case and a Potential Solution", 2017.

¹⁴⁹Bertolini et al., "On robots and insurance", 2016.

Capitolo 4

L'uomo, il Golem

4.1 Introduzione

Il giudice sta cominciando a ricostruire, passo dopo passo, l'intera vicenda. Le difficoltà di ricostruire l'intenzione, l'azione e il controllo sulla stessa sono ormai chiare, così come è chiaro il legame che lega Golem e rabbino: sia questa connessione magia o tecnologia, si è ormai capito che l'interazione è intensa e lo diventerà sempre di più. Eppure, nella mente del giudice che è chiamato a pronunciarsi sull'intera vicenda, si affaccia con sempre maggiore forza e prepotenza un dubbio, una domanda che non può più restare priva di risposta. La domanda che si pone è molto semplice quanto enormemente complessa. *Chi* ha davanti?

L'essere è immobile e silente davanti a lui, ma non capisce su cosa si stanno posando i suoi occhi. Mentre rivolgeva le domande al rabbino, era chiaro che il suo interlocutore era un uomo. Per quanto capace di padroneggiare la magia, il giudice aveva innanzi pur sempre un essere umano. L'umanità del rabbino e l'umanità del giurista si toccano e si rispecchiano l'una nell'altra. Adesso che è giunto il momento di rivolgersi all'essere, tuttavia, l'umanità del giudice incontra il buio. Il giurista coglie la profondità del baratro che li separa. E, pure, questa distanza che appare incommensurabile non fa venire meno un atteggiamento di meraviglia e stupore. Al contrario, più il baratro è profondo, maggiore è la forza con la quale si affaccia la domanda divenuta ormai imprescindibile.

Chi è il creatore e chi la creatura? Non è un quesito fine a se stesso. A seconda della prospettiva, infatti, cambia anche colui che da *oggetto* di indagine diviene *parte* del giudizio. È una nuova luce che inevitabilmente porta a chiedersi fino a che punto l'uomo si sia fatto Golem e viceversa.

Questa indagine avanza su strade ancora poco battute dagli interpreti del diritto, che conducono non solo a comprendere meglio la natura dell'essere che ci sta di fronte, ma anche a mettere in discussione la natura dell'uomo stesso. Il problema che si pone, infatti, è stabilire se vi è o meno un confine che separa l'algoritmo e il suo utente. La dimensione più antropologica del *Responsibility Gap*, che ora ci si accinge ad analizzare, porta con sé un quesito fondamentale e decisivo per rapportarsi al problema. Al centro della trattazione si deve porre, ora, il soggetto dotato di interfaccia, colto nel suo essere *individuo* e dunque, per definizione, un essere non diviso e non frammentato in una pluralità di parti. La domanda che bisogna porsi ora attiene proprio alla individualità del soggetto, al suo essere *in-diviso*. Bisogna chiedersi, cioè, se dalla interazione con le interfacce emerge ancora un solo individuo, anche se radicalmente diverso dagli altri, o se nascono due individualità: l'utente da un lato, l'interfaccia dall'altra. Tale scissione comporterebbe dunque la frammentazione dell'utente in un *dividuo* con profonde ricadute sul piano filosofico e giuridico per il fatto che “solo ciò che è, nella sua individualità, non disarticolato può avere, nel suo rapporto a sé o un altro, un'identità e un'autonomia”¹.

Fin qui è stata descritta una delle due possibili in-dividualità, il primo grande “attore” nel problema noto come *Responsibility Gap*: l'interfaccia. L'attenzione ora deve spostarsi sulla seconda protagonista: l'uomo. Per capire in che modo l'entrata in scena delle interfacce può cambiare il concetto stesso di uomo, nonché la consapevolezza che l'umanità ha di sé stessa, si possono muovere i primi passi da come si considera, e come si considererà in futuro, l'uomo. Sarà qui ripercorso un dibattito risalente e consolidato, il quale, non potendo essere questa la sede per una sua ricostruzione integrale, può comunque un'utile cornice all'interno della quale inserire le problematiche fin qui affrontate.

Il punto di partenza è capire se quel *chi* che agisce è l'uomo o il suo cervello. Si sarà notato, infatti, che fino a questo momento la prospettiva giuridica è sempre stata integrata con una prospettiva neuroscientifica. Sotto certi profili, quest'ultima da un lato ha rappresentato un vero e proprio limite a determinate evoluzioni della scienza giuridica, dall'altro lo sviluppo di nuovi strumenti in ambito neuroscientifico ha rappresentato un “innesco” per nuove riflessioni circa l'opportunità e l'urgenza di cambiamento per alcune categorie giuridiche. Questo si dà proprio per il fatto che, come ormai è chiaro, con le più recenti tecniche di *neuroimaging* “accade che lo sguardo

¹Fuselli, “Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto”, 2020.

del neuroscienziato si spinga sino a osservare fenomeni che poi hanno a che fare con ‘coscienza’, ‘identità’, ‘libertà’, ‘responsabilità’”².

Il punto di riferimento non è più l'essere umano nel suo complesso, l'essere dietro alla *persona*, ma una precisa sua parte, un preciso suo organo: il cervello.

Una volta analizzata questa dimensione, si passerà a un altro profilo rilevante. L'individuo non solo può essere *dividuo* scindendosi dal suo cervello, ma anche attraverso una disarticolazione del sé nella sua interfaccia. In questo caso, allora, per cercare di capire quale sia il confine che separa il primo dalla seconda, è utile anche passare in rassegna alcuni orientamenti filosofici e antropologici che si collocano a metà tra presente e futuro.

Si può procedere alla luce di tre importanti “lanterne”: il diritto, la filosofia e la “magia” del nostro tempo, le neuroscienze. Iniziare questo cammino significa chiedersi anche in che modo questi tre saperi si intersecano reciprocamente, delineando un modello epistemologico che inevitabilmente incide sul punto di arrivo; il ritmo dei passi, infatti, è scandito da quel legame “magico” che lega artefice e artefatto: la nuova frontiera delle neuroscienze. Proprio da questa occorre partire.

4.2 L'uomo o il suo cervello?

Il 13 settembre 1848, durante i lavori di costruzione di una ferrovia nel New England, gli operai presenti nel cantiere assistettero a un incidente molto particolare che avrebbe segnato la storia della ricerca scientifica.

Il manovale addetto agli esplosivi, tale Phineas Gage, avrebbe infatti consegnato il suo nome al progresso e al futuro. A causa di un errore nella tempistica di detonazione per livellare un terreno irregolare, una sbarra di acciaio, spessa 3 centimetri e lunga 109, gli perforò completamente il cranio, per poi fuoriuscire. Non solo sopravvisse, riuscendo a muoversi e a farsi accompagnare in ospedale, ma sopravvisse come un uomo completamente diverso. L'uomo che era sul letto d'ospedale non era lo stesso del giorno precedente. Tanto quest'ultimo era educato, gentile, paziente e responsabile, quanto il soggetto ricoverato, la stessa persona che poi avrebbe vissuto ancora a lungo, era capricciosa, irriverente e irrispettosa.

Fin da subito medici e chirurghi si chiesero com'era possibile un cambiamento tanto radicale nel comportamento di un uomo, anche per l'assenza di danni cerebrali o malattie, almeno visibili.

²Gaiani, “Riduzionismo e neuroscienze: il dibattito filosofico recente”, 2014, pag. 48.

Vent'anni dopo, nel 1868, il medico John Harlow attribuì, in via ipotetica, senza gli strumenti di mappatura cerebrale oggi presenti e senza la possibilità di eseguire un'autopsia, il cambio di personalità di Gage al perforamento del cranio e ai conseguenti danni cerebrali³. Harlow non era l'unico in quel periodo ad aver intuito una correlazione tra conformazione cerebrale e personalità.

Quasi in contemporanea alla pubblicazione degli studi di Harlow, due medici, il francese Paul Broca e il tedesco Carl Wernicke, avevano sostenuto⁴ la riconducibilità delle azioni e dell'abilità di parlare dei propri pazienti a "centri" motori e del linguaggio presenti del cervello.

Le conclusioni dei due medici furono accolte con maggiore entusiasmo⁵ e meraviglia rispetto ai risultati, analoghi, di Harlow per due motivi.

In primo luogo, le loro ricerche erano sostenute dalle risultanze di una autopsia sui cadaveri dei pazienti. In secondo luogo, alcuni ritengono che le conclusioni di Harlow, in forza delle quali vi è una causalità biologica alla base delle azioni e del comportamento sociale degli uomini, fossero contrastanti con i valori morali e religiosi degli Stati Uniti di fine XIX secolo: le spiegazioni causali dell'agire umano non potevano essere nient'altro che fondate su motivi di ordine etico e religioso⁶.

Al di là delle conclusioni del caso di specie, che pure sono state soggette ad alcune critiche⁷, la prima metà del XX secolo è stata segnata profondamente da questi primi passi nella conoscenza dell'anatomia del cervello.

Sessantasette anni dopo, nel 1935, i medici Antonio Egas Moniz e Almeida Lima eseguirono un'operazione chirurgica sui lobi frontali, recidendo le connessioni della corteccia prefrontale dell'encefalo di pazienti psichiatriche per cercare di lenire le loro sofferenze, compiendo quella è conosciuta come la prima lobotomia, allora chiamata leucotomia. Il cervello umano non solo veniva osservato, ma anche modificato. Da una dimensione meramente "contemplativa" della complessità di tale organo, si passò a una attiva, di

³In particolare, si segnalano Harlow, "Passage of an iron rod through the head.", 1848, nonché Harlow, "Recovery after severe injury to the head", 1868.

⁴Vd. Broca, "Sur le siège de la faculté du langage articulé", 1865 e Wernicke, *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*, 1874.

⁵La portata di queste scoperte è stata tale che due zone della corteccia cerebrale hanno preso il nome dei due neurologi: sono l'Area di Wernicke, presente nel lobo temporale, che permette la comprensione del linguaggio, e l'Area di Broca, nell'emisfero cerebrale dominante, la cui funzione è l'elaborazione del linguaggio.

⁶Questa motivazione è ipotizzata da Hannah Damasio. Si veda Damasio et al., "The return of Phineas Gage: clues about the brain from the skull of a famous patient", 1994.

⁷Tra gli altri, vd. Macmillan, *An odd kind of fame*, 2000.

modificazione e intervento sul medesimo⁸. A Egas Moniz la prima lobotomia valse il premio Nobel nel 1949, per aver dimostrato il valore terapeutico di tale intervento.

Solo un anno dopo l'operazione, nel 1936, Walter Freeman e James Watts pubblicarono un articolo in cui, oltre a coniare il termine *lobotomia*, delinearono la procedura standard per l'operazione, d'ora in poi chiamata procedura "Freeman-Watts"⁹.

Da quel momento in poi, la tecnica della lobotomia avrebbe avuto un'ampia diffusione. Nel 1949 venne compiuto il numero più alto di lobotomie, ben 5.074, fino ad arrivare a un totale, nel 1951, di 18.608 operazioni¹⁰.

Ben presto, tuttavia, la fiducia e la convinzione nell'utilità terapeutica della lobotomia lasciarono il posto a diffuse e profonde inquietudini e paure, sentite a più livelli nella società. Una tecnica ideata per alleviare sofferenza e dolore e "migliorare"¹¹ la persona si stava rivelando un aberrante strumento di alterazione e, in certi casi, di distruzione dell'identità e della personalità del paziente¹².

Il primo Paese a bandire tale tecnica fu l'Unione Sovietica nel 1950¹³, e anche negli Stati Uniti crebbe in via esponenziale un crescente disagio

⁸Un contributo interessante per percorrere la storia della c.d. *psychosurgery* si trova in Mashour, E. E. Walker e Martuza, "Psychosurgery: past, present, and future", 2005.

⁹Freeman e Watts, "Prefrontal lobotomy in agitated depression. Report of a case", 1936.

¹⁰Edward, *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*, 1997.

¹¹Sui problemi connessi alla tendenza al miglioramento dell'uomo si veda il paragrafo successivo, dedicato al c.d. *neuroenhancement*.

¹²Una delle pazienti più ricordate, ma per troppo tempo dimenticate, è stata Rosemary Kennedy, sorella del ben più conosciuto John Fitzgerald, trentacinquesimo presidente degli Stati Uniti. Considerata mentalmente instabile per via di comportamenti lunatici, agitati e talvolta violenti, fu sottoposta a lobotomia nel 1941, all'età di 23 anni, affinché fosse "curata" così da non provocare pubblico imbarazzo e danneggiare le carriere politiche dei fratelli. Il risultato dell'operazione fu tragico. Una ragazza sensibile e intelligente, forse affetta solo da dislessia, perse l'uso della parola, subì una paralisi permanente di un braccio e non fu mai più in grado di camminare autonomamente. Le sue capacità cognitive diminuirono fino a essere paragonabili a quelle di un bambino di due anni. Rosemary Kennedy fu subito internata, all'oscuro dei suoi fratelli, nella *St. Coletta School for Exceptional Children* nel Wisconsin. Solo nel 1961 i familiari vennero a conoscenza del luogo del suo internamento e solo nel 1987 il pubblico venne a conoscenza della sua lobotomia. Per il racconto della sua storia e la descrizione delle conseguenze dell'operazione, si veda Kessler, *The sins of the father: Joseph P. Kennedy and the dynasty he founded*, 1996, pag. 243.

¹³La lobotomia venne infatti considerata come una pratica contraria ai principi di umanità, in quanto "attraverso la lobotomia un malato psichiatrico è trasformato in un idiota". Così in Diefenbach et al., "Portrayal of lobotomy in the popular press: 1935-1960", 1999.

nel consentire la pratica, tanto che nel 1977 il presidente Jimmy Carter creò un apposito Comitato Nazionale per investigare circa possibili utilizzi della lobotomia basati su pregiudizi razziali contro esponenti di minoranze etniche statunitensi. Non solo. Le critiche alla lobotomia crebbero a tal punto che una consistente parte dell'opinione pubblica chiese la revoca del Premio Nobel assegnato al chirurgo Antonio Egas Moniz¹⁴.

Il tentativo di modificare l'uomo attraverso la lobotomizzazione del cervello umano rappresenta uno dei più grandi fallimenti della medicina recente¹⁵. La storia della lobotomia è addirittura considerata da qualcuno come un monito a guardarsi dal suo utilizzo distorto e profondamente lesivo della dignità umana, facendo bene attenzione che i fattori che hanno comportato l'avvento della *psychosurgery* sono ancora presenti: diverse teorie circa la qualificazione delle malattie mentali, il confronto acceso tra neurologi e psichiatri, il bisogno di cura e l'entusiasmo immediato e acritico dell'opinione pubblica¹⁶.

Negli anni più recenti¹⁷, nonostante il ricorso a determinati approcci terapeutici sia ormai tramontato, il fascino che la nostra materia grigia esercita non è calato. Ha preso avvio, infatti, un'indagine volta a "illuminare" il "lato oscuro" del cervello, aprendo alla possibilità di una "radicale modificazione nell'antropologia di senso comune e, in particolare, nel generale senso di giustizia che il diritto presupporrebbe nelle sue regolazioni [...]"¹⁸. La scoperta di questo "nuovo mondo" non ha lasciato indifferente i filosofi più accorti. Il momento in cui effettivamente anche l'ambito filosofico ha cominciato ad imbarcarsi per navigare verso i nuovi orizzonti delle neuroscienze ha una

¹⁴Jansson, "Controversial psychosurgery resulted in a Nobel Prize", 2007.

¹⁵*ivi*, pag. 367.

¹⁶Così Valenstein, *Great and Desperate Cures: the rise and decline of psychosurgery and other radical treatments for mental illness*. 1986, pagg. 5-6. Secondo l'autore "*Psychosurgery was not a medical aberration, spawned in ignorance. In a real sense, the history of psychosurgery is a cautionary tale: these operations were very much a part of the mainstream medicine of their time, and the factors that fostered their development and made them flourishing are still active today [...]*."

¹⁷Per un autorevole, anche se risalente, sguardo ad ulteriori sviluppi del tema si veda, tra gli altri, Fuselli, *Diritto, neuroscienze, filosofia. Un itinerario*, 2014.

¹⁸Così Sarra, "Questioni pregiudiziali: una prospettiva epistemologica sui rapporti tra neuroscienze e diritto", 2014, pag. 85. L'Autore si riferisce al contributo in Zeki et al., "For the law, neuroscience changes nothing and everything", 2004, nel quale si sostiene che le neuroscienze sono in grado di incidere profondamente sull'evoluzione della scienza giuridica. Il punto di riferimento è dato dal tema della responsabilità penale e della funzione retributiva della pena. Per una opinione più scettica circa la portata delle neuroscienze si veda S. J. Morse, "New neuroscience, old problems: legal implications of brain science.", 2004, nonché S. Morse, "Lost in translation?: An essay on law and neuroscience", 2011.

data precisa.

Nel 1986 Patricia Churchland pubblica *Neurophilosophy*¹⁹, uno scritto che è tutt'ora considerato uno spartiacque²⁰ che segna la distinzione tra un *avanti* e un *dopo* l'impatto della ricerca neuroscientifica negli ambienti di ricerca.

Da quell'anno in poi, infatti, le neuroscienze si sono articolate in nuovi campi del sapere, nuovi rami della speculazione filosofica²¹. Ambiti come "filosofia delle neuroscienze", "neurofilosofia" si affacciano per la prima volta nel dominio proprio della filosofia, ai confini del territorio proprio del diritto.

Più nello specifico, la filosofia delle neuroscienze si può intendere come un tentativo di sottoporre a un giudizio critico le scoperte delle neuroscienze: in altre parole, una "revisione critica e teorica sui presupposti filosofici e sulle implicazioni filosofiche della ricerca neuroscientifica"²².

La neurofilosofia, invece, comporta una ricerca e un confronto: capire in che modo le nuove conquiste conoscitive circa il nostro cervello interagiscono e mutano le tradizionali categorie filosofiche, quali, ad esempio, "emozione", "desiderio", "azione morale"²³.

È all'interno della neurofilosofia che si inserisce la c.d. neuroetica, che rappresenta la chiave di accesso a un'indagine molto più profonda e ad ampio spettro, avente ad oggetto nozioni come la libertà, l'"identità umana" e la "responsabilità giuridica". Un'interessante indicazione circa la definizione di "neuroetica" e riguardo alle forme nelle quali si può sviluppare è fornita dalla neurofilosofa Adina Roskies.

L'Autrice distingue le etiche delle neuroscienze dalle neuroscienze dell'etica.

Le prime si occupano di questioni che sorgono dal confronto e scontro tra le conoscenze neuroscientifiche e il loro impatto etico e sociale. Le seconde, invece, approfondiscono quelli che sono i fondamenti fisici e biologici di tali questioni etiche e antropologiche²⁴.

¹⁹Patricia Smith Churchland, *Neurophilosophy*, 1986.

²⁰Bickle, Mandik e Landreth, "The Philosophy of Neuroscience, in «The Stanford Encyclopedia of Philosophy»", 2019.

²¹Per seguire l'evoluzione delle neuroscienze, mi rifaccio a Gaiani, "Riduzionismo e neuroscienze: il dibattito filosofico recente", 2014. Lo stesso Autore, comunque, è consapevole delle eccezioni e dei limiti a cui è soggetta tale divisione.

²²*ivi*, pag. 49.

²³*ibidem*

²⁴Vedi Roskies, "Neuroethics for the new millenium", 2002. Per un'altra definizione, ripresa da un convegno tenuto a San Francisco il 13-14 Giugno 2002, dal titolo "*Neuroethics: mapping the field*", si rimanda a Santosuosso e Bottalico, "Neuroscienze e categorie giuridiche: quale impatto", 2009.

La nascita di tali branche della Filosofia è solo il punto di partenza, che si inserisce in un cammino molto più lungo.

Jake Dunagan, direttore dell'*American Institute for the Future*, ha annunciato, nel 2010, la venuta di quella che è stata definita una *Neurocentric Age*²⁵. È un'epoca nella quale “*we are looking to the brain sciences as the explanatory touchstone for diverse questions and disciplines, and we are already witnessing neuro-becoming the prefix du jour, generating an array of neurologisms*”²⁶.

Alcuni autori non solo hanno riconosciuto il ruolo sempre più centrale degli studi sulla materia grigia umana, ma hanno anche evidenziato come questi siano stati oggetto di grande interesse da parte di operatori economici e oggetto di forti interessi di mercato, tanto da coniare il termine “neurocapitalismo”²⁷.

La portata evocativa e l'ascendente di tali scoperte sulla cultura di massa è tale da aver portato alla luce una vera e propria “neropolitica.”

Il teorico politico William Connolly ha attribuito un significato molto preciso a tale termine: “*the politics through which cultural life mixes into the composition of brain-body processes. And vice-versa*”²⁸.

In definitiva, sempre più confuso e sempre più sbiadito è il *gap* che separa ciò che è interno da ciò che è esterno all'uomo. Accorciandosi la distanza, si mette in discussione anche lo stesso ruolo dell'uomo nel mondo, il quale cade in una sua progressiva “naturalizzazione”. Nel 1998, Domenico Parisi, allora membro e direttore dell'ISTC-CNR ha interpretato la “naturalizzazione del soggetto” in due sensi: “prima di tutto significa non considerarlo come ‘speciale’. Gli esseri umani hanno la tendenza e l'esigenza di considerare se stessi come ‘speciali’, e la base di questa ‘specialità’ viene trovata nella mente, nella cultura, ecc. Invece la Scienza non considera nulla come ‘speciale’. Ogni cosa ha almeno alcune delle sue caratteristiche che sono diverse da quelle di ogni altra cosa, ma nulla è ‘speciale’. L'altra cosa che significa naturalizzare l'uomo è che l'uomo viene visto e studiato dall'esterno. La naturalizzazio-

²⁵Dunagan, “Politics for the neurocentric age”, 2010, pag. 55.

²⁶*ibidem*. Dunagan si richiama a un suo precedente scritto per il tema dei *neurologisms*: si veda Dunagan, “Neuro-futures: the brain, politics and power”, 2004. Il corsivo è nel testo originario.

²⁷Questa è la posizione degli intellettuali marxisti Ewa Hess e Henrich Jokeit. Così i due autori, che individuano nelle neuroscienze, intese come l'insieme degli studi sul cervello, una “*usurpatory tendency to become not only the humanities of science, but the leading science of the twenty-first century. The legitimacy, impetus and promise of this claim derive from the maxim that all human behaviour is determined by the laws governing neuronal activity and the way it is organised in the brain.*” Hess e Jokeit, “Neurocapitalism”, 2009, pag. 1.

²⁸Connolly, *Neropolitica: Thinking, culture, speed*, 2002, pag. 3.

ne è l'eso-scienza dell'uomo. L'uomo tende a vedere se stesso dall'interno, anche quando fa scienza dell'uomo. Invece la naturalizzazione annuncia un nuovo tipo di scienze dell'uomo in cui l'uomo viene visto e studiato da fuori, come vediamo e studiamo ogni altro fenomeno della realtà"²⁹.

L'uomo, il rabbino, dunque, perde la sua specialità e diviene nient'altro che un *oggetto*. La *persona* viene spogliata della sua maschera e rimane nuda sul letto di una sala operatoria. Le ricadute di questa perdita di originalità dell'essere umano non si fermano qui.

4.2.1 Il libero arbitrio

Dal 1986 ad oggi sempre più forti sono emerse le frizioni tra una concezione riduzionistica dell'uomo, che comporta come automatico corollario il riconoscimento della fondatezza di una visione deterministica dell'agire umano, e la possibilità di formulare un fondato giudizio di responsabilità, che richiede, invece, come presupposto imprescindibile, un certo livello di libertà di giudizio e di autodeterminazione. Come considerare responsabile un soggetto che non ha agito, ma che "è agito" dal suo cervello?

Sull'annosa questione circa la sussistenza o meno della libertà di giudizio molte le voci che si sono elevate, nell'uno e nell'altro senso. Molteplici sono stati gli autori che hanno elevato gli scudi a difesa del riconoscimento del libero arbitrio.

Alcuni, ad esempio, hanno evidenziato l'inutilità dell'esame di un singolo cervello per determinare o meno la responsabilità del soggetto: è l'interazione tra più soggetti, e, dunque, più cervelli, che fa *emergere* delle proprietà ulteriori non ricavabili sulla base di singoli studi anatomici o fisiologici del cervello³⁰. La stessa Churchland, nonostante riconosca che l'impatto delle neuroscienze sia ormai inevitabile e imprescindibile, nega che effettivamente queste ultime siano in grado di individuare, da sole, le ragioni profonde dell'agire morale dell'uomo³¹.

Altri, invece, sostengono la completa superiorità delle neuroscienze su altri saperi, quali la Filosofia: conoscere in modo sempre più approfondito il funzionamento del cervello significa riuscire a spiegare in modo sempre più completo complessi modi di essere e di comportarsi dell'uomo. Tale posizione è inamovibile: la filosofia deve attendere, passiva, che le neuro-

²⁹Parisi, *La naturalizzazione della cultura*, 1998.

³⁰Così, riassumendo, Gazzaniga e S. Inglese, *Chi comanda? Scienza, mente e libero arbitrio*, 2017.

³¹Afferma la filosofa che "la piattaforma è solo la piattaforma". Vd. Patricia S Churchland, *Neurobiologia della morale*, 2012, pag. 15.

scienze aprano i “cancelli” e percorrano le “strade” della mente senza alcuna possibilità di intervento, essendo privata di ogni utilità epistemologica. Il *mainstream* delle neuroscienze risiede nella fisiologia cellulare e nella biologia molecolare³².

Accanto ai due fronti se ne può, tuttavia, individuare un terzo: è la posizione dei c.d. neurocivilizzatori moderati o riformisti³³.

Tra questi, di particolare pregio è l'opinione di Stephen Morse, portavoce di una posizione c.d. “compatibilista”.

Da un lato il giurista riconosce i presupposti naturalistici delle neuroscienze, dall'altro, tuttavia, non esclude comunque la compatibilità di un determinismo universale con il *free-will*, il libero arbitrio. In particolare, egli evidenzia un punto tanto scontato quanto fondamentale: dal momento che ogni azione umana è prodotta dal cervello, l'unica situazione nella quale effettivamente non è possibile riconoscere come responsabile l'agente è quando agisce in uno stato alterato. In caso contrario, la possibilità di formulare un giudizio di responsabilità è innegabile. Il giurista presenta il caso di quella che è definita la *Brain Overclaim Syndrome*: si rinviene là dove si è consapevoli che si è comunque responsabili delle proprie azioni, ma attribuire tali azioni esclusivamente al cervello, in forza di un legame causale esclusivamente biochimico, permette di sentirsi assolti da ogni colpa, in un'ottica deresponsabilizzante³⁴.

L'Autore, comunque, non ritiene che le neuroscienze siano in grado di modificare radicalmente gli ordinamenti giuridici, non potendo incidere su quello che egli riconosce essere il “nucleo” del diritto, in particolar modo del diritto penale, nonchè suo fondamento: una psicologia di senso comune, o *folk-psychology*. Il comportamento umano, secondo questa teoria, è causato in via principale da stati mentali quali progetti, desideri, convinzioni e solo in parte da fattori biologicamente determinati. Certo, la stessa definizione di “stato mentale” è controversa sia tra giuristi che tra scienziati e filosofi, ma ciò comunque non fa venire meno la loro centralità nella teoria dell'azione. La stessa dinamica relazionale che intercorre tra due o più persone dà una dimostrazione della validità di tale teoria. Nel momento di scontro e di incontro tra due soggetti, lo stesso linguaggio utilizzato è incompatibile con una concezione determinista o riduzionistica dell'esistenza, rimandando all'insieme di emozioni, convinzioni, credenze incompatibile con una conce-

³²John Bickle, *Philosophy and neuroscience: A ruthlessly reductive account*, 2003, pag.2. È lo stesso Autore a definirlo un *ruthless reductionism*, un riduzionismo estremo.

³³Riprendo qui Paolo Sommaggio, “Neurocivilizzazione”, 2014.

³⁴Per una disamina più approfondita, si veda Kolber, “Will there be a neurolaw revolution”, 2014, pag. 821.

zione riduzionistica dell'uomo: “*brains do not convince each other, people do*”³⁵. La psicologia di senso comune, dunque, comporterebbe che le azioni umane sarebbero più razionalmente spiegate attraverso il richiamo alle predisposizioni mentali e i soggetti agenti più sensibili a fornire motivazioni se vi sono le giuste condizioni³⁶.

Ora, al di là della effettiva fondatezza o meno delle posizioni a favore o contro la sussistenza del libero arbitrio e a prescindere dalle prove a sostegno dell'una o dell'altra, ci si deve chiedere quale sia l'*utilità* di riconoscere o meno la presenza di quello che si può definire il libero arbitrio.

A tal proposito, si può condividere l'idea per la quale, anche se la presenza o meno del libero arbitrio rimane un'incognita, comunque “*the characterisation of human persons and their responsibility in this world confers on them a value and dignity that cannot simply be reduced to biology. Even though humans are physical beings, they cannot be explained by mere scientific concepts, since they can transcend the concept of physicality*”³⁷. Riconoscere la capacità degli esseri umani a compiere decisioni “libere”, qualunque sia la nozione di libertà, permette di mantenere in piedi l'intero sistema morale che è alla base della stessa convivenza degli esseri umani. Alla domanda “esiste il libero arbitrio?” bisogna rispondere che essa non ha importanza o è mal posta. Molto più utile sarebbe chiedersi se bisogna o meno “far finta” che vi sia. Ecco allora che la risposta non può che essere positiva.

Bisogna agire *come se* ci fosse, in forza di una “scommessa morale”: i benefici di questa “finzione” sono enormemente superiori ai costi. Per conservare e proteggere la stabilità dell'intero sistema normativo e morale, appare allora necessario l'abbandono di quella che è stata definita una “*paper ethics*”³⁸, un'etica “di carta”, fragile, che non vede nulla oltre il cervello umano, considerato origine e fine dell'intera personalità umana, per abbracciare una neuroetica più autentica, la quale “*rises beyond its nature and reflects a wisdom not entirely its own*”³⁹.

Un'etica, dunque, che sia consapevole del fatto che la conformazione anatomica del cervello non esaurisce la sua complessità, allo stesso modo in cui la più complessa tra le figure origami non esaurisce la sua forma più

³⁵S. J. Morse, “The status of neurolaw: a plea for current modesty and future cautious optimism”, 2011, pagg. 598-599.

³⁶*ivi*. Per una panoramica più completa, si veda, dello stesso autore, S. Morse, “Avoiding irrational neurolaw exuberance: a plea for neuromodesty”, 2011.

³⁷MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017, pag. 127.

³⁸Cheshire, “The origami brain: from neural folds to neuroethics”, 2011.

³⁹*ivi*

propria in una delle sue piegature: così come ogni piegatura della carta punta oltre se stessa, in una retta finita verso una direzione infinita, così una neuroetica più genuina deve puntare oltre se stessa, in una dimensione che va al di là del mero piano descrittivo⁴⁰.

4.3 L'uomo o il suo algoritmo?

La questione fin qui ricostruita verte, essenzialmente, sul rapporto tra persona e cervello, tra libertà e determinismo, tra un agire morale, libero, e un agire meccanicistico, dovuto: compatibilità, determinismo, libertà, naturalizzazione. Nella complicata storia tra l'uomo e il suo cervello si inserisce una terza parte: l'interfaccia neurale. Uno dei profili problematici per cui si dà il problema del *Responsibility Gap*, infatti, attiene anche alla difficoltà di cogliere il *confine* ontologico tra utente e interfaccia, il limite al di qua del quale un neurodispositivo può essere considerato come parte dell'essere umano che ne dispone.

È necessario, adesso, compiere un passo avanti. La domanda che bisogna porsi, infatti, non è più semplicemente se l'uomo è il suo cervello, bensì se l'uomo è la sua interfaccia, o, più in generale, in che modo uomo e neurotecnologia condizioneranno a vicenda la propria esistenza. I due termini del rapporto sono l'interfaccia, da un lato, e l'uomo, dall'altro. Se la costituzione e le caratteristiche della prima sono conosciute, capire quale sia effettivamente la natura del secondo presenta notevoli difficoltà. Non esiste, infatti, una definizione univoca di "umanità". Il problema relativo alla domanda "chi" o "cosa" è l'uomo si dà essenzialmente per il fatto che non si rinviene una sua categorizzazione in termini generali e astratti né su un piano normativo né su un piano morale. La crisi emerge, in particolar modo, con riguardo al fenomeno del potenziamento umano, che, in quanto espressione dell'aspirazione dell'uomo di andare *oltre*, di superare i propri limiti, permette anche di guardare cosa potrà essere l'umanità in futuro.

Il potenziamento umano, o *enhancement*, si può definire, secondo un autorevole contributo, come l'insieme interventi biomedici che sono posti in essere al fine di migliorare la forma umana e le sue prestazioni al di là di ciò che è necessario per ripristinare o mantenere la salute⁴¹. Ciò, tuttavia,

⁴⁰*ivi*. Cheschire trae spunto dall'opera dell'origamista Brian Chan, il quale ha costruito una replica in origami del cervello umano. Così afferma l'Autore in riferimento a una neuroetica incentrata esclusivamente sul cervello: "A brain-based neuroethics ultimately is a paper ethics, a morally thin construction that tears under the stress and collapses under pressure".

⁴¹Juengst e Moseley, "Human Enhancement", 2019.

presuppone la distinzione tra potenziamento, che opera su individui sani, e terapia, che opera su individui malati, e, di conseguenza, la distinzione tra ciò che è qualificabile come salute e ciò, invece, che è definito come malattia, e la determinazione di cosa è normale e cosa non lo è⁴². Le scelte nell'uno o nell'altro senso non si basano solo su nozioni medico-scientifiche, ma presuppongono determinate prese di posizione in ambito valoriale, invadendo con prepotenza il campo delle relazioni sociali⁴³.

L'impianto di interfacce neurali si colloca a metà tra una funzione di potenziamento e una di terapia. È qui che si coglie la portata totalizzante dell'esperienza di interazione con le BCI. Un loro impiego sempre più massiccio può portare a ridefinire il concetto stesso di salute e malattia, di normalità e anormalità. Là dove, ad esempio, gli utenti potenziati dalle BCI siano in numero superiore ai soggetti privi, ecco che potrebbe essere ridefinito anche il concetto di "normale": i nuovi mostri, i nuovi esseri anormali potrebbero essere gli stessi esseri umani che rifiutano l'interazione con le neurotecnologie, rimanendo fedeli al tradizionale sostrato biologico di cui è dotato l'uomo.

I possibili settori in cui si può avere un potenziamento umano sono molteplici.

Uno dei più importanti, dal quale possono derivare anche le conseguenze più inquietanti, è quello militare, nel quale un possibile pregiudizio del bene giuridico in questione è finalizzato al potenziamento psicofisico dell'utente. Degna di considerazione è la ricerca condotta nell'ambito del *Defence Advanced Research Projects Agency* (DARPA), un progetto di ricerca statunitense lanciato nel settembre del 2018 e volto a sperimentare sui soldati americani l'impiego sia di psicofarmaci che di neurotecnologie ai fini di potenziamento delle capacità di combattimento e resistenza, ridefinendo i limiti di resistenza imposti dalla natura⁴⁴. Attraverso connessioni neurali con apposite interfacce si sperimentano fasci di microonde per cogliere e trasmettere l'attività cerebrale al fine ultimo di influenzare e controllare il comportamento. La connessione *wireless* consente di creare una "memoria estesa"⁴⁵ per assicurare che l'attenzione del soggetto rimanga costante, tenendo sotto controllo

⁴²Fuselli, "Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell'uomo", 2015, pag. 155. Il contributo di Stefano Fuselli sarà più volte trattato in questo paragrafo.

⁴³Così T.H Murray, *Enhancement*, in Steinbock, *The Oxford Handbook of Bioethics*, 2007, pag. 494.

⁴⁴Non a caso lo slogan del programma di reclutamento del DARPA recita: "*Be all that you can be, and a lot more.*" Vd. C. L. Annas e G. J. Annas, "Enhancing the fighting force: medical research on American soldiers", 2008, pag. 283.

⁴⁵Amato, "Neuroscienze e utilizzazione militare delle tecniche di potenziamento umano", 2014, pag. 182.

gli impulsi emotivi per permettere una maggiore lucidità nel momento decisionale, anche attraverso la ricezione di un flusso continuo di dati, in forza di una sorta di “telepatia sintetica”⁴⁶. Il supersoldato del futuro è in grado di mantenere l’operatività in combattimento per un periodo ininterrotto nel tempo senza provare stanchezza, fame, dolore, cancellando ogni forma di inibizione e qualunque ricordo doloroso. Non a caso un programma del DARPA che porta il nome di *Persistence in Combat* mira a creare militari “*unstoppable because pain, wounds and bleeding*”.

All’interno poi del programma denominato N3 o *Next-Generation Non-Surgical Neurotechnology* è stata progettata una interfaccia neurale bidirezionale che permette ai militari di connettersi a droni o robot a distanza con il mero pensiero, attraverso un continuo flusso di informazioni tra cervello del soldato e dispositivo esterno.

Brian E. Moore, tenente colonnello dell’Aeronautica Militare Americana, ha affermato che la portata delle BCI in campo militare è paragonabile, sotto il profilo dell’impatto sulla tecnologia militare, a quello derivante dallo sviluppo e dall’utilizzo degli armamenti nucleari⁴⁷.

All’applicazione militare dei dispositivi si accompagna anche un loro utilizzo nell’ambito dell’ingegneria della memoria, così come dimostrato dai ricercatori dell’Università di San Diego, i quali attraverso un laser ottico hanno stimolato determinate zone del cervello della cavia, nel caso di specie un ratto, per rafforzare o indebolire determinate connessioni neurali, incidendo dunque sulle loro capacità mnemoniche⁴⁸.

Come si evince anche dagli ambiti di impiego, il superamento della condizione umana e la ridefinizione del limite che separa ciò che è umano da ciò che non lo è dipende in gran parte dal modo in cui è considerato il rapporto che lega persona e corpo⁴⁹. L’esperienza della lobotomia ha reso ormai pacifico che profondi interventi sulla dimensione fisica e corporale dell’individuo comportano profonde ricadute sull’individuo stesso nel suo complesso, modificando “*the very core of identity and personality, our mental powers of memory, and will*”⁵⁰. I limiti del corpo sono i limiti dell’individuo. Alterare tali confini significa alterare qualcosa di ben più profondo. Si pensi, ad

⁴⁶ *ivi*, pag. 184.

⁴⁷ Vd. B. E. Moore, *The brain computer interface future: time for a strategy*, 2013.

⁴⁸ L’esperimento completo è descritto in Nabavi et al., “Engineering a memory with LTD and LTP”, 2014.

⁴⁹ Per la trattazione successiva, il rimando è a MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017, pagg. 157 ss.

⁵⁰ Così Cole-Turner, *Transhumanism and transcendence: Christian hope in an age of technological enhancement*, 2011, pag. 7.

esempio, all'installazione di una protesi costituita da un cuore meccanico. Se l'artefatto è installato all'interno del corpo, sostituendo un cuore "naturale", non si porrebbero problemi rilevanti: l'inviolabilità dell'individuo è affermata e riconosciuta pacificamente. Problematiche di ordine morale, etico e antropologico sorgono, invece, nel caso in cui vi sia un'estensione del corpo *oltre* il corpo stesso: il cursore mosso su uno schermo attraverso impulsi cerebrali è considerabile come parte della dimensione fisica dell'individuo?

Nella fusione tra uomo e macchina, infatti, la soggettività umana può scomparire.

Gli uomini diventano i loro stessi prodotti, l' "Io" che decide si avvicina sempre di più all' "Io" che è il risultato della decisione⁵¹.

Ciò che è artificiale tende ad avvicinarsi sempre di più a ciò che è naturale, giungendo a "fondersi e a con-fondersi"⁵². In questa tendenza si può rinvenire, come sostrato filosofico, la tendenza a una svalutazione del corpo, in linea con una visione materialistica e non finalistica della natura.

La connessione è tale che in futuro, a una decrescente perdita di capacità cognitive per l'invecchiamento dell'utente di BCI, si potrà accompagnare un incremento delle funzionalità dell'algoritmo: "*ultimately, your brain would die, and your mind would find itself entirely in the computer*", afferma il futurologo e scienziato canadese Hans Moravec⁵³. Sempre più parti del mondo esterno condivideranno lo status, morale e psicologico, di componenti del cervello umano⁵⁴.

È una condivisione, o, meglio, come è già stata chiamata nel corso della trattazione, una simbiosi, che potrebbe portare addirittura a una vera e propria "evoluzione" della macchina a partire dall'interazione con l'utente. L'algoritmo sotteso, ad esempio, al funzionamento della BCI potrebbe trasformarsi in modo tale da risultare "umano" in quanto evolutosi lentamente da un essere che è inevitabilmente umano. La conseguenza sarebbe la creazione di una realtà che è ancora umana, ma che trascende le radici biologiche dell'umanità⁵⁵.

⁵¹Ivi, pag. 7-8. Dice Cole-Turner: "*as we turn technology on ourselves so that we change our own bodies and brains, the 'I' is swept up in the change and modified through its own action. When these technologies of human enhancement get inside us, they become part of us, turning us into our own products and blurring the lines we once drew between subjects and object, agent and effect.*"

⁵²Palazzani, "La condizione tecno-umana e le tecnologie convergenti: percorsi scientifici e filosofici "oltre" l'umano", 2019.

⁵³H. Moravec, *Mind children: The future of robot and human intelligence*, 1988, pag. 4.

⁵⁴Così Clark e Cyborgs, "Minds", 2003.

⁵⁵Questa l'opinione di Ray Kurzweil in Kurzweil, *The singularity is near: When humans transcend biology*, 2005, pag. 9.

Una parte della società accoglie con fiducia e ottimismo questa convivenza. Tra questi, un ruolo preminente è ricoperto dai c.d. transumanisti.

Il Transumanesimo si può descrivere come un'ideologia del progresso volta ad emancipare gli esseri umani dai limiti che la natura impone loro, rappresentati, *in primis*, da un corpo mortale. Il mezzo scelto per il superamento della natura umana è, come si può capire, il progresso tecnologico e una maggiore interazione uomo-tecnologia⁵⁶.

È interessante riprendere le parole del biologo Julian Huxley (1887-1975). In un articolo del 1957 egli afferma, in riferimento al Transumanesimo, che *“the present limitations and miserable frustrations of our existence could be in large measure surmounted [...] The human species can, if it wishes, transcend itself not just sporadically [...] but in its entirety, as humanity. [...] It will at last be consciously fulfilling its real destiny”*⁵⁷. Una umanità, dunque, che si emancipa dalla mortalità dell'esistenza grazie all'ingerenza totalizzante del progresso tecnologico, variamente inteso. I fedeli di questa “religione senza rivelazione”⁵⁸ credono fermamente nella possibilità di miglioramento dell'umanità, in attuazione di ideali di giustizia e libertà più autenticamente intesi⁵⁹, con ciò presentandosi dunque come i successori degli umanisti “tradizionali”. Nel “Manifesto dei Transumanisti Italiani”, così si afferma: “l'idea cardine del Transumanesimo può essere riassunta in una formula: è possibile ed auspicabile passare da una fase di evoluzione cieca ad una fase di evoluzione autodiretta consapevole. Noi siamo pronti a fare ciò che oggi la scienza rende possibile, ovvero prendere in mano il nostro destino di specie. Siamo pronti ad accettare la sfida che proviene dai risultati delle biotecnologie, delle scienze cognitive, della robotica, della nanotecnologia e dell'intelligenza artificiale, portando detta sfida su un piano politico e filosofico, al fine di dare al nostro percorso un senso e una direzione. Si badi che questo progetto non ha molto a che fare con l'eugenetica negativa e autoritaria predicata nel XIX secolo e messa in pratica dagli Stati Uniti d'America, dalla Germania nazionalsocialista e dalle socialdemocrazie scandinave nel XX secolo [...] In altre parole, è pura mistificazione identificare l'eugenetica negativa e autoritaria del passato con l'attuale modello transumanista di evoluzione autodiretta, che è proteso a garantire in positivo la salute e il potenziamento degli individui e della loro prole, tenendo sempre

⁵⁶Tirosh-Samuelson, “Transhumanism as a secularist faith”, 2012.

⁵⁷Per l'articolo completo, vd. Huxley, “Transhumanism”, 2015.

⁵⁸È lo stesso Huxley a fornire questa definizione. Vd. Huxley, *Religion without revelation*, 2016.

⁵⁹Tirosh-Samuelson, “Transhumanism as a secularist faith”, 2012.

ferma la libertà di scelta e il diritto alla felicità del nascituro”⁶⁰.

Alcuni autori⁶¹, tuttavia, rilevano come questo sia un accostamento improprio e ingenuo. Gli umanisti considerano l'umanità come un valore da proteggere e l'uomo come una creatura privilegiata rispetto agli altri esseri della natura. L'umanità e la sua conservazione, la sua protezione, sono il fine, non il mezzo. Al contrario, i transumanisti considerano l'umanità semplicemente come un gradino nel processo evolutivo. L'uomo *deve* evolversi se vuole raggiungere un livello di sviluppo e di espressione della propria personalità più alto. L'umanità dunque è solo un mezzo per tendere a questo obiettivo superiore⁶².

Ancora più radicale è l'ideologia del c.d. Postumanesimo. A questa etichetta si sono ricondotte un'eterogeneità di movimenti e di correnti di pensiero, accomunate, tuttavia, da un lato dal persistente tema della perdita di differenze nette tra meccanismo artificiale e organismo biologico. Lo sfondo etico e filosofico è rappresentato, in questo caso, dalla “caduta delle coordinate umanistiche”⁶³. Se i transumanisti auspicavano il miglioramento dell'umanità attraverso una maggiore integrazione tra uomo e macchina, considerando comunque l'umanità come onnipresente in questo cambiamento, i postumanisti esasperano a tal punto il superamento dei limiti dell'uomo da abbandonare completamente, e non solo in una dinamica evolutiva, la concezione e le sembianze dell'uomo per come si conoscono oggi⁶⁴.

Un contributo fondamentale per comprendere chi sia, o, meglio, *cosa* sia l'essere “postumano” è stato dato dall'autrice americana Katherine Hayles, nel suo libro del 1999 intitolato “*How we become posthuman*”.

L'essere postumano è un amalgama, un insieme di componenti eterogenee che si rimodulano in una continua dinamica di costruzione e ricostruzione⁶⁵: in questo *monstrum* la coscienza non è altro che il prodotto del sostrato fisico, il corpo non è altro che una primitiva forma materiale originale che può essere migliorata o sostituita da qualcosa di nuovo e diverso.

⁶⁰Campa, “Manifesto dei transumanisti italiani”, gen. 2009, pag. 76. Nel Manifesto l'Autore difende l'ideologia transumanista dalle principali critiche e accuse mosse da vari esponenti dell'opinione pubblica italiana, facendo anche una sintesi dei pregiudizi che molti nutrono nei confronti di questa tendenza.

⁶¹È di questa opinione MacKellar, *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*, 2017, pag. 162.

⁶²Waters, *From Human to Posthuman: Christian Theology and Technology in a Postmodern World*, 2006, pag. 78.

⁶³Marchesini, *Post-human: verso nuovi modelli di esistenza*, 2002, pag. 528.

⁶⁴Riprendo qui Savulescu, “Moral Status of Enhanced Beings: What Do We Owe the Gods?”, 2009, pag. 214.

⁶⁵Hayles, “Katherine: How we became posthuman”, 1999, pag. 3.

Gli esseri umani, nel nuovo mondo postumano, non solo possono essere paragonati a macchine intelligenti, ma i due sono, addirittura, interscambiabili⁶⁶.

Lo stesso concetto di unitarietà del corpo nella nuova realtà postumana viene travolto. L'essere postumano si rinviene nel flusso di dati che connette le sue componenti. Non vi è più distinzione tra uno e molti: il carattere eterogeneo dell'essere umano comporta la distribuzione della capacità cognitiva umana, che viene “distribuita” in più parti dell'essere postumano, che non necessariamente sono in comunicazione tra di loro⁶⁷. La natura, nello specifico umana, così per come è oggi conosciuta, è considerata come una “costrizione”, il prodotto di una “normativizzazione” forzata, imposta dalla società⁶⁸. L'individuo diviene “performatività”, essendo libero di “costruire-decostruire, fare e dis-fare liberamente la propria identità”⁶⁹. L'identità, allora, subisce un perenne mutamento non già sulla base di influenze della società, ma in base a “desideri, pulsioni e impulsi interni all'individuo”⁷⁰. L'“io” si spoglia della fissità del corpo, che ne costituisce espressione, per divenire un'identità fluida, frammentata, dinamica. “Il soggetto unico, reale, razionale, universale, si infrange in frammenti senza centro, senza sequenzialità, senza ordine”⁷¹.

Il post-umano si trasforma in *data made flesh*, un flusso di dati in carne e ossa⁷².

Il tema della simbiosi sempre crescente tra intelligenza naturale e intelligenza artificiale si colloca ormai al confine tra scienza e fantascienza, tra interpretazioni del presente e previsioni del futuro.

Tra queste ultime si colloca l'ipotesi, sostenuta da alcuni⁷³, che in futuro non vi saranno più due “specie”, l'uomo e la macchina, a confrontarsi, ma solo una. Algoritmi e tecnologie intelligenti sostituiranno l'uomo e la sua intelligenza. L'era meccanica, così come definita da alcuni autori, prenderà

⁶⁶ *ivi*

⁶⁷ *ivi*

⁶⁸ Palazzani, “La condizione tecno-umana e le tecnologie convergenti: percorsi scientifici e filosofici “oltre” l'umano”, 2019.

⁶⁹ *ivi*, pag. 67.

⁷⁰ *ibidem*

⁷¹ *ibi*

⁷² Il riferimento è a Gibson et al., *Neuromancer*, 1984, pag. 16.

⁷³ Si veda, in particolare, Kurzweil, *The age of spiritual machines: When computers exceed human intelligence*, 2000, H. Moravec, *Mind children: The future of robot and human intelligence*, 1988, Garis, “The Artilect War: Cosmists vs. Terrans. A Bitter Controversy Concerning Whether Humanity Should Build Godlike Massively Intelligent Machines”, 2008.

inizio con il raggiungimento di un punto di non ritorno in cui l'accelerazione dello sviluppo tecnologico sarà tale da causare quella che è definita, in letteratura, la Singolarità.

Il saggista Robert Geraci definisce quest'ultima come il punto più alto nella curva del progresso, che subisce un'impressionante crescita esplosiva, in cui le macchine diverranno abbastanza intelligenti da iniziare ad apprendere da sole⁷⁴. In questo picco dell'evoluzione umana avverrà l'irreversibile sostituzione del mondo naturale e biologico con quello meccanico, espressione di un nuovo "regno", un regno virtuale in cui l'uomo non sarà più la specie dominante⁷⁵. Secondo l'informatico e saggista Raymond Kurzweil la Singolarità ha una data precisa: il 2045.

Nel 2045 l'uomo trascenderà definitivamente i suoi limiti, la sua natura mortale, il suo corpo limitato, controllando il proprio destino, andando oltre la stessa mortalità umana⁷⁶.

L'ultimo nemico a essere sconfitto sarà, infine, la morte.

Lo sguardo di filosofi, scienziati e saggisti si spinge, infatti, tanto in là nella fantascienza da immaginare una realtà futura nella quale sarà possibile effettuare un *upload* della coscienza e della mente degli uomini in supercomputers, ai quali sarà affidato il compito di prendersi cura dei bisogni più materiali e fisici degli esseri umani. La stessa umanità potrebbe essere trasformata in un'entità pensante collettiva⁷⁷. L'era della mente e del cervello sostituirà l'era delle macchine e l'uomo per come si è sempre conosciuto si convertirà in un flusso di dati codificati da algoritmi ad apprendimento automatico. Solo la computazione resterà⁷⁸. L'umanità, dunque, raggiungerà, finalmente, dopo secoli di sofferenza, la c.d. *cyberimmortalità*⁷⁹.

Si può capire, allora, come il postumano sia una tendenza dell'umanità al superamento di se stessa, il tentativo di rinnegare il limite, di costruire un mondo nuovo, migliore, privo di sofferenza, di dolore, in cui regna una libertà assoluta, nella quale ciascuno ha opportunità illimitate per lo sviluppo personale e sociale⁸⁰.

⁷⁴Geraci, *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*, 2012, pag. 149.

⁷⁵Tirosh-Samuelson, "Transhumanism as a secularist faith", 2012.

⁷⁶così in Kurzweil, *The singularity is near: When humans transcend biology*, 2005, pag.9.

⁷⁷È di questa idea H. Moravec, *Mind children: The future of robot and human intelligence*, 1988, pag. 116.

⁷⁸H. P. Moravec, *Robot: mere machine to transcendent mind*, 2000, pag. 163.

⁷⁹Per la descrizione di questa futura conquista dell'uomo, si veda Bainbridge, "Religion for a galactic civilization 2.0", 2009.

⁸⁰Waters, *From Human to Posthuman: Christian Theology and Technology in a Postmodern World*, 2006, pagg. 77-78.

Nonostante la visione utopica, molti sono i profili critici di questa “emanipazione” dell'uomo.

In primo luogo, Laura Palazzani rileva⁸¹ un “errore” di fondo nell'ideologia transumana e postumana. La tendenza e l'aspirazione a vivere esperienze svincolate dai limiti della corporeità, per poter compiere attività e vivere esperienze che trascendono le tradizionali coordinate spaziali e temporali, presuppone comunque l'utilizzo dei propri sensi. L'uomo postumano e transumano non trascenderebbe la sua corporeità, ma la presupporrebbe, con la conseguenza ultima che non vi è un superamento dei limiti corporali, ma, piuttosto, un loro potenziamento. Nel tentativo di andare oltre la natura umana, transumani e postumani la affermano come imprescindibile, con buona pace, dunque, della convinzione affermata per la quale l'uomo è senza natura, e “quest'ultima si riduce a storia, prassi, libertà”⁸². Alla fine, “non si può divenire, se non si è: l'essere è il presupposto e il postulato del divenire. Il divenire senza essere si esaurisce nella fluidità caotica e casuale del processo, che non ha un inizio e nemmeno una fine, non ha un centro e una direzione, nel quale non si possono distinguere fasi o identificare percorsi”⁸³.

In secondo luogo, Katherine Hayles rileva come la costruzione di una realtà in cui le persone, chiunque o qualunque cosa siano diventate, sono completamente libere e quasi immortali annullandosi nella tecnologia, potrebbe comportare il sacrificio di molti dei valori tipici di una moderna società liberale, in primis il valore rappresentato dal rispetto dell'individuo e della sua identità⁸⁴.

La realizzazione di una società postumana, e, forse, anche transumana, inevitabilmente conduce al superamento del confine tra l'individuo, delineato dalla sua conformazione biologica personale, e l'ambiente che lo circonda⁸⁵. Non ci saranno più barriere tra chi pensa e l'oggetto del pensiero, tra pensiero e materia. Mente e computer si fonderanno in un'unica entità⁸⁶. All'interno del flusso di dati la centralità del cervello umano viene meno⁸⁷. La rete di informazioni, che ne permette il flusso, è preminente rispetto alla conservazione dell'identità del singolo. Non vi sarà più alcuna intimità in quanto ogni pensiero diviene un dato registrabile e trasmissibile in un tutto

⁸¹Palazzani, “La condizione tecno-umana e le tecnologie convergenti: percorsi scientifici e filosofici “oltre” l'umano”, 2019.

⁸²*ivi*, pag. 74.

⁸³*ivi*, pag. 75.

⁸⁴Hayles, “Katherine: How we became posthuman”, 1999.

⁸⁵Waters, *From Human to Posthuman: Christian Theology and Technology in a Postmodern World*, 2006, pag. 78.

⁸⁶Pepperell, *The post-human condition*, 1995, pag. 34.

⁸⁷Tirosh-Samuelson, “Transhumanism as a secularist faith”, 2012.

collettivo.

Le implicazioni possono essere molto profonde sul piano normativo. Se si intende il potenziamento umano come un andare oltre i limiti, considerando i limiti come delle mancanze, si deve prendere atto che “l’idea stessa di ‘correggere dei difetti’ [...] implica avere un modello di riferimento - in questo senso un ‘ideale’ regolativo - rispetto a cui quelli sono valutati come difetti la cui correzione non solo è opportuna, ma anche moralmente doverosa”⁸⁸. L’installazione di interfacce neurali per migliorare le prestazioni di arti o articolazioni, o per assicurare un utilizzo più immediato e funzionale di computer e altri dispositivi potrebbe essere, in futuro, un obbligo giuridico fondato su un’intuizione meramente ideologica: la condizione successiva è migliore di quella precedente. L’evoluzione della natura umana verso il postumano o verso il transumano, un futuro raggiunto anche attraverso l’impiego delle neurotecnologie, è il bene al quale tendere. Alcuni hanno addirittura sostenuto la necessità e l’utilità di rendere obbligatori i miglioramenti biomedici al fine di rendere le persone più “moralì”, affermando dunque una forte adesione alla pratica del c.d. *moral enhancement* o potenziamento morale⁸⁹. Alla fine, la conseguenza più radicale potrebbe essere il riconoscimento sociale e giuridico solo “di quegli esseri che sono *mutanti* non in quanto sinonimo di ‘già modificati’, ma di modificantesi”⁹⁰.

Tuttavia, “interpretare l’evoluzione nel senso di un miglioramento o di un progresso, implica fare intervenire un criterio valutativo che essa, in quanto dinamica biologica, non ha affatto”⁹¹. Si pongono a fondamento di questa evoluzione determinati valori dominanti. In che modo relazionarsi, allora, con coloro che non considerano tali valori come preminenti? Il tema del potenziamento dell’uomo, del superamento dei suoi limiti, del raggiungimento della perfezione transumana o postumana non è infatti una questione di autodeterminazione del singolo.

Affermare che ognuno è sovrano del proprio destino e padrone del proprio corpo, e che il potenziamento umano, ad esempio attraverso BCI, è una scelta dettata dalla libertà del singolo, non dovendo e non potendo essere una decisione soggetta al giudizio della collettività o a una politica di controllo

⁸⁸Fuselli, “Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell’uomo”, 2015, pag. 173.

⁸⁹Vd. Persson e Savulescu, “Unfit for the future? Human nature, scientific progress, and the need for moral enhancement”, 2011.

⁹⁰Così P. Sommaggio, *Filosofia del biodiritto: una proposta socratica per società postumane*, 2016, pag. 48. Il corsivo è presente nel testo originario.

⁹¹Fuselli, “Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell’uomo”, 2015, pag. 174. L’Autore riprende qui Chiereghin, *L’eco della caverna: ricerche di filosofia della logica e della mente*, 2004, pagg. 329-330.

da parte dello Stato, porta a una contraddizione profonda. È una convinzione che nasconde in sé un sistema antitetico al complesso di valori liberali del quale si afferma figlia: è il sistema meno liberale “perché esclude ogni altro soggetto dalla sfera argomentativa e, quindi, si pone in tendenziale opposizione con l’idea che il diritto legittimo sia frutto della più ampia mobilitazione possibile della libertà comunicativa dei cittadini”⁹². Inoltre, partendo dalla premessa che ogni atto di autodeterminazione soggiace all’unico limite di non arrecare danno agli altri consociati, potrebbe essere considerata dannosa la stessa scelta di non sottoporsi a una tecnica di potenziamento, in quanto implicante la rinuncia ai benefici di quest’ultimo⁹³.

Rapportando tutte le considerazioni fin qui enucleate al problema della responsabilità, le incertezze relative alla modificazione della natura umana che possono essere di ostacolo all’attribuzione di responsabilità potrebbero essere ridimensionate riprendendo lo stesso senso della domanda “chi è l’uomo”⁹⁴. Il quesito, infatti, acquista un carattere “autoriflessivo” dal momento che l’uomo diviene sia interrogato che interrogante. Ecco allora che “la domanda pare diventare impraticabile proprio da colui che ha il massimo interesse, ossia l’uomo stesso”⁹⁵. La domanda in esame, nella sua portata totalizzante, da un lato coinvolge l’essere uomo di colui che la pone, dall’altra, tuttavia, fa sì che lo stesso soggetto divenga “opaco a se medesimo”⁹⁶. La natura umana non può presentarsi come il possibile oggetto della conoscenza e si “manifesta proprio nell’inapplicabilità del rapporto soggetto-oggetto e, quindi, anche nella sua indisponibilità alle varie modalità della ‘oggettivazione’”⁹⁷. La stessa natura umana, tuttavia, fa sì che “nulla di ciò che l’uomo è o fa, nulla di ciò che diviene, nemmeno progettare un futuro post o transumano, è estraneo a quell’insieme di possibilità che si dischiudono all’uomo per il fatto di essere uomo”⁹⁸.

Nel caso di un utente di interfaccia, dunque, il confine tra algoritmo

⁹²Amato, “La riproduzione assistita come diritto. Riflessioni biogiuridiche”, 2001, qui pag. 116.

⁹³Vd. Fuselli, “Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell’uomo”, 2015. Nella oceanica letteratura sul tema del potenziamento “obbligato”, si rimanda senza pretese di completezze a Mehlman et al., “Ethical and legal issues in enhancement research on human subjects”, 2011 e al più recente Sandel e Galli, *Contro la perfezione. L’etica nell’età dell’ingegneria genetica*, 2022.

⁹⁴Seguo qui Fuselli, “Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell’uomo”, 2015.

⁹⁵*ivi*, pag. 181.

⁹⁶*ivi*, pag. 182.

⁹⁷*ibidem*.

⁹⁸*ibidem*.

e utente si dissolve. Quest'ultimo, infatti, dotandosi di arti robotici, non farebbe altro che esercitare la sua capacità di sfruttare le possibilità di cui dispone in linea con il “principio di scelta”⁹⁹ insito nell'uomo. In questo senso, alla fine, l'algoritmo non è un possibile soggetto agente: al contrario, è l'espressione della capacità del soggetto di conservare la sua umanità. Il *gap* di responsabilità inizierebbe allora a ridursi, eliminando dall'equazione la difficile questione dell'umanità dell'individuo.

4.3.1 Identità e continuità psicologica

Le potenzialità modificative dell'identità personale fanno sorgere un altro grande paradigma che deve essere ricostruito per comprendere quale sia la dimensione dell'utente di BCI che è maggiormente esposta a possibili sconvolgimenti e alterazioni¹⁰⁰. Il difensore di Mary potrebbe presentare come argomento a sostegno dell'assenza di responsabilità un'alterazione nella sua identità personale tale da causare la sua condotta. In questo caso, tuttavia, ancora una volta ci si trova innanzi a un *gap* epistemico, dal momento che è necessario verificare sia in che modo può essere declinato il concetto di identità personale sia l'impatto su tale categoria dell'utilizzo di BCI.

La questione si porrebbe con una prepotenza ancora più forte se l'utente avesse sviluppato i sintomi di una malattia degenerativa. In questa situazione verrebbe in rilievo¹⁰¹ l'analisi dell'applicazione a fini terapeutici dei neurodispositivi nei confronti di soggetti malati di Parkinson.

Attraverso il loro utilizzo, infatti, è possibile applicare la tecnica di stimolazione cerebrale profonda o *DBS*: mediante l'allaccio di elettrodi è possibile inviare impulsi elettrici in una determinata area del cervello, andando ad agire direttamente sull'origine di disturbi del movimento. A livello tecnico l'applicazione degli elettrodi è seguita dall'innesto di un generatore di impulsi, installato sotto la superficie della cute, da qualcuno paragonato a una sorta di *pacemaker* del cervello¹⁰².

L'utilizzo di tale tecnica di neuromodulazione ha impieghi terapeutici molto vasti, che spaziano dalla cura dell'epilessia a quella dell'Alzheimer, passando per il trattamento di disturbi ossessivo compulsivi e anoressia, fino

⁹⁹*ivi*, pag. 183.

¹⁰⁰L'inquadramento filosofico e psicologico del concetto di identità personale è sempre stato al centro di un ampio dibattito. Non potendo essere questa la sede per una approfondita e multidisciplinare sua analisi, si rinvia, nel *mare magnum* di contributi, a Reichlin, *Etica e neuroscienze: stati vegetativi, malattie degenerative, identità personale*, 2012.

¹⁰¹Il punto di riferimento qui, che sarà più volte ripreso, è Fuselli, “Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto”, 2020.

¹⁰²Vd. Jansson, “Controversial psychosurgery resulted in a Nobel Prize”, 2007.

a giungere a configurarsi come strumento di potenziamento e di modificazione dell'umore. Quest'ultima applicazione si rivela particolarmente rilevante in quanto il controllo degli impulsi mediante DBS è affidato, guarda caso, a un algoritmo¹⁰³. L'esecuzione della tecnica è affidato a dispositivi c.d. di seconda generazione o ad anello chiuso anche detti dispositivi *closed-loop*, che si distinguono dai *devices* di prima generazione sul piano dell'interazione con il cervello.

Nei primi, il flusso del segnale è unidirezionale, dal dispositivo al cervello, nei secondi, invece, il flusso è bidirezionale: inviano segnali al cervello e registrano quelli derivanti dall'attività cerebrale. Il risultato ultimo è che l'algoritmo riesce a regolare l'intera dinamica automaticamente, senza l'intervento del paziente.

Per comprendere la portata del problema si possono ricordare due casi clinici molto discussi in letteratura.

Il primo dei due vede come protagonista un signore olandese di 62 anni, affetto da Parkinson. Dopo essersi sottoposto volontariamente alla DBS, recuperò la mobilità, ma sviluppò uno squilibrio psichico tale da menomare profondamente la sua capacità decisionale. Spegnendo la DBS, il paziente recuperava la sua personalità, ma perdeva il controllo motorio. Il comitato etico dell'ospedale scelse di rimettere a lui la decisione di conservare o meno i dispositivi che permettevano la DBS: in un momento di lucidità il signore optò per la loro conservazione, mantenendo il controllo sui suoi muscoli, ma perdendo per sempre la propria autonomia, in quanto fu necessario il suo internamento in una struttura psichiatrica¹⁰⁴.

Nel secondo caso, invece, un soggetto affetto dalla sindrome di Tourette, sottoposto a DBS, ha sviluppato episodi di dissociazione della personalità, dovuti, probabilmente, al venire in superficie di ricordi traumatici dell'infanzia, che puntualmente venivano meno con l'interruzione della stimolazione¹⁰⁵.

È così emersa in letteratura l'annosa questione circa la conservazione o meno del tradizionale paradigma dell'identità personale. È una domanda che sorge anche alla luce di una possibile sua alterazione da parte della DBS,

¹⁰³Un contributo circa gli impieghi psichiatrici della DBS si trova in Goering et al., "Staying in the loop: Relational agency and identity in next-generation DBS for psychiatry", 2017, nonché Synofzik, Schlaepfer e Fins, "How happy is too happy? Euphoria, neuroethics, and deep brain stimulation of the nucleus accumbens", 2012.

¹⁰⁴Il caso è descritto in Klaming e Haselager, "Did my brain implant make me do it? Questions raised by DBS regarding psychological continuity, responsibility for action and mental competence", 2013, pag. 534.

¹⁰⁵Vd. Goethals et al., "Brain activation associated with deep brain stimulation causing dissociation in a patient with Tourette's syndrome", 2008.

che fa venire meno la continuità psicologica che sperimenta l'individuo. Ci si chiede, in particolare, se sia ancora possibile aderire a una concezione c.d. statica dell'identità personale o se, al contrario, sia necessario adottare una ricostruzione dinamica della stessa. Percorrendo questa seconda via secondo alcuni Autori si ridimensionerebbe la "pericolosità" della DBS: riquilificando la nozione di identità personale si riquilificherebbero anche le sue modificazioni. Non più attentati alla sua integrità, ma semplici sue trasformazioni¹⁰⁶. La questione, cioè, non verterebbe sull'identità, ma sull'autonomia: di conseguenza l'unico punto di vista degno di nota sarebbe quello del paziente, il solo in grado di affermare se si tratta di una modifica positiva o negativa¹⁰⁷.

Accanto a queste problematiche, appare comunque utile passare in rassegna alcune ricostruzioni del concetto di identità personale.

Una autorevole voce da ascoltare per comprendere i possibili significati da attribuire a questa categoria si rinviene nell'*Handbook of Neuroethics* del 2015¹⁰⁸. Catriona Mackenzie e Mary Walker forniscono una panoramica di tre possibili concezioni della identità personale, nel tentativo di capire quale delle tre si addice maggiormente all'esperienza dei soggetti sottoposti a DBS¹⁰⁹.

La prima forma di identità personale si può definire come "numerica", e presenta due articolazioni teoriche rilevanti: la teoria della continuità psicologica e quella della continuità biologica. L'identità numerica, infatti, stabilisce i c.d. "criteri ontologici che consentono di distinguere tra l'essere persona e l'essere non persona o di distinguere due diversi individui o di stabilire cosa fa sì che essa continui a essere numericamente una nonostante i cambiamenti o di stabilire quando una persona cessa di esistere"¹¹⁰. A un cambiamento nell'identità numerica corrisponde una trasformazione dell'individuo, che diventa letteralmente "qualcun altro"¹¹¹. Non è que-

¹⁰⁶Per questa posizione si veda Focquaert e De Ridder, "Direct intervention in the brain: ethical issues concerning personal identity", 2009, così come Schermer, "Changes in the self: The need for conceptual research next to empirical research", 2009.

¹⁰⁷A tal proposito. cfr. Synofzik e Schlaepfer, "Stimulating personality: ethical criteria for deep brain stimulation in psychiatric patients and for enhancement purposes", 2008.

¹⁰⁸Clausen, Levy et al., *Handbook of neuroethics*, 2015.

¹⁰⁹Vd. Mackenzie, M. Walker et al., "Neurotechnologies, personal identity, and the ethics of authenticity", 2015. Per una descrizione delle tre forme di identità, seguono anche Goddard, "Deep brain stimulation through the "lens of agency": Clarifying threats to personal identity from neurological intervention", 2017.

¹¹⁰Così Stefano Fuselli in Fuselli, "Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto", 2020.

¹¹¹Schermer, "Ethical issues in deep brain stimulation", 2011.

sta, tuttavia, secondo le due Autrici, la trasformazione che hanno subito i pazienti nei due casi precedenti descritti.

La seconda forma si fonda su una concezione pratica dell'identità, che a sua volta presuppone una domanda esistenziale circa il soggetto che agisce. Quel "chi" che agisce, infatti, è mosso sia da ragioni personali che riflettono la propria esperienza di vita, sia dalle condizioni sociali o culturali nelle quali è inserito. Questo perché "l'identità pratica offre un modo per spiegare in terza persona le nostre azioni e un modo in prima persona per giustificarle. La centralità che assume la decisione, tuttavia, fa sì che l'autoidentificazione sia più sincronica che diacronica, e tenda a costituirsi attorno a un nucleo assiologico pressochè immodificabile"¹¹².

Infine, la terza e ultima forma, che secondo Mackenzie e Walker si riscontra nei casi suddetti, è la c.d. concezione narrativa dell'identità: si inciderebbe direttamente sulla capacità di autonarrazione, intrinsecamente relazionale e strutturalmente dinamica¹¹³, che permette all'individuo di reinventarsi e di ricostruirsi grazie all'interazione con altri individui. I pazienti soggetti alla DBS dunque vedrebbero modificata radicalmente questa capacità.

Se dunque queste sono le possibili qualificazioni della categoria di identità personale, c'è da chiedersi quale siano le sue ricadute sul piano giuridico.

Come è stato evidenziato¹¹⁴, l'identità personale che viene in rilievo nell'ambito dei neurodispositivi che permettono l'esecuzione di varie tecniche, tra le quali la DBS, non è sovrapponibile perfettamente al concetto di identità personale, così come ricostruito sul piano più squisitamente legale. Questo perché "l'identità personale rispetto alla quale si vanta un diritto, non è né l'immagine che il soggetto ha di sé (verità personale), che può, in casi estremi, essere scorrelata dalla realtà, né l'insieme dei dati oggettivi riferibili al soggetto (verità storica), ma l'immagine, socialmente mediata e oggettivata, del soggetto stesso"¹¹⁵.

In una prospettiva *de iure condendo* un possibile strumento a difesa dell'identità personale così come intesa dal punto di vista neurologico è individuato da Marcello Ienca e Roberto Andorno nel riconoscimento della posizione giuridica soggettiva definita come *diritto alla continuità psicologi-*

¹¹²Fuselli, "Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto", 2020.

¹¹³Vd Schechtman, "Getting our stories straight", 2009.

¹¹⁴Fuselli, "Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto", 2020.

¹¹⁵Cocuccio, "Il diritto all'identità personale e l'identità "digitale"", 2016.

ca¹¹⁶. Più precisamente, i due Autori sono guidati dall'idea di “guidare l'integrazione responsabile dell'intelligenza artificiale nell'intelligenza umana in modo da preservare il senso di identità personale - definito come la continuità della propria vita mentale - ed evitare una manipolazione incosciente”¹¹⁷. La tutela di questo *neurodiritto* agisce in due direzioni.

Da un lato, fa sì che ciascuno possa mantenere un livello di controllo sul proprio comportamento e, cosa più importante, sulla propria esperienza soggettiva, evitando la discontinuità che invece hanno attraversato i pazienti neurologici nelle vicende sopra descritte.

Dall'altro, impedisce anche che terzi non autorizzati possano intervenire pericolosamente alterando o manipolando il senso di identità del soggetto, intervenendo abusivamente nella sfera psicologica dell'individuo: si protegge la mente per proteggere la persona.

Secondo i due bioeticisti, il diritto alla continuità psicologica è una “istanza neurofocalizzata” del diritto a sviluppare una personalità sancito dall'art. 29 della Dichiarazione Universale dei Diritti dell'Uomo, là dove afferma che “ogni individuo ha dei doveri nei confronti della società, nella quale soltanto è possibile il libero e pieno sviluppo della sua personalità”.

4.4 Umanità e individualità

L'interazione di connessione con le interfacce, come è stato visto in questo capitolo, può portare a chiedersi chi è l'uomo, o la donna, utente delle interfacce, e in che modo la natura dell'uomo e l'identità del singolo vengono investite dal rapporto con l'algoritmo. A prescindere dai caratteri propri della natura umana, che, forse, sono insondabili, comunque si capisce quali possibili risposte si possono dare alla domanda “chi ha agito?”.

I riduzionisti individuerebbero il “chi” che ha agito nel cervello. È tutto lì: l'utente è il cervello.

I compatibilisti risponderebbero ugualmente facendo riferimento all'attività cerebrale, ma senza tuttavia negare la possibilità di esercitare la propria capacità di autodeterminazione e, più in generale, l'esistenza del libero arbitrio.

I transumanisti e postumanisti non avrebbero dubbi: non si pone neanche la distinzione tra utente, cervello e algoritmo. Ad agire è stata una

¹¹⁶Ienca e Andorno, “Towards new human rights in the age of neuroscience and neurotechnology”, 2017.

¹¹⁷Così Ienca in Ienca, “Intelligenza²: per un'unione di intelligenza naturale e artificiale”, 2019, pag. 163.

entità psicofisica potenziata in modo invidiabile dalla BCI.

Tutti questi orientamenti inevitabilmente prevedono un impatto sull'individualità dell'utente.

Se si aderisce a una posizione riduzionistica, la scissione tra uomo e cervello comporta automaticamente la scissione dell'individualità tra uomo e algoritmo. Riconoscendo il cervello come “altro” dall'utente, i segnali neurali non sono riconducibili alla persona, ma esclusivamente al cervello. Non riconducendo l'attività neurale all'individuo, l'utente che abbiamo di fronte si divide, disarticolandosi in tre parti: uomo, algoritmo, cervello. Ognuna di queste tre parti è un centro di imputazione autonomo e a sé stante: l'interfaccia in quanto ha rilevato e tradotto nell'*output* richiesto l'attività neurale, il cervello in quanto ha attivato l'interfaccia, ciò che rimane dell'utente in quanto titolare di una “tradizionale” soggettività giuridica. E' proprio seguendo tale ricostruzione riduzionistica che si pone il dubbio circa l'identità del responsabile, alla base della questione del *Responsibility Gap*. A quale delle tre parti ricondurre l'evento?

La stessa individualità, al contrario, sarebbe mantenuta aderendo a una ricostruzione sia compatibilista sia, paradossalmente, transumana o postumana. Non vi sarebbe nessuna divisione: cervello, algoritmo, utente si riconducono a un tutto indiviso e unitario. Se il cervello non è altro dall'utente, ogni segnale neurale inviato è riconducibile a quest'ultimo. L'interfaccia, dunque, non sarebbe altro che un canale attraverso il quale si estrinseca una sola e indivisa individualità, la quale si esprime anche attraverso impulsi neurali involontari. Seguendo questo orientamento si conserverebbe un paradigma di integrità che fa sì che venga riconosciuta una unica identità che, forse, permette di fondare un giudizio di responsabilità, riconducendo ad unità l'utente.

In questo caso, tuttavia, se si mantiene la sua individualità, essa ne risulterebbe comunque trasformata dalla presenza dell'interfaccia e, in un certo senso, evoluta. Il capitolo successivo è dedicato proprio a questo: capire in che modo questa nuova individualità 2.0 deve essere considerata nella ricerca della responsabilità e in che modo può essere pregiudicata dai mezzi utilizzabili per la sua determinazione.

Capitolo 5

Il costo della responsabilità

5.1 Introduzione

Il giudice si rende che conto che il Golem che è innanzi a lui inevitabilmente segnerà per sempre la storia del ghetto e la vita dei suoi abitanti. Dal giorno in cui la creatura ha compiuto il primo passo dentro il villaggio, tutto cambia.

Nonostante il notevole sforzo, non si è colta l'essenza ultima del mostro, occultata in profondità, dietro ad un velo di magia quasi intoccabile.

E, pure, il Golem è lì. È davanti al giudice, è davanti al giurista, è davanti a ognuno di noi. I suoi passi pesanti si sentono sempre più forti.

Prima di procedere con la trattazione, si faccia un passo indietro. Nelle pagine precedenti è stata messa alla prova la nozione di intenzione: la distinzione, emersa nella letteratura neuroscientifica, tra segnali neurali “volontari” e “non volontari” comporta alcune difficoltà di inquadramento anche sul piano giuridico. Se l'interfaccia è stata attivata da impulsi involontari sui quali l'utente non ha potuto esercitare alcun controllo, in che modo considerare dolosa la sua condotta? Sul piano di una possibile condotta colposa la questione non cambia. Come poter affermare che questi avrebbe dovuto esercitare un controllo sui propri pensieri, quasi in forza di un obbligo giuridico alla freddezza mentale? Le difficoltà in questo senso emergerebbero anche per le caratteristiche che dovrebbe presentare sotto un profilo tecnologico l'interfaccia, per permettere all'utente di esercitare un potere di veto sulla rilevazione degli impulsi neurali da parte della stessa.

Ancora, oltre alla qualificazione dell'intenzione e alla sua riconducibilità all'utente di BCI, in crisi è lo stesso concetto di condotta, dolosa o colposa che sia. Il movimento delle braccia robotiche del caso di scuola esaminato è ascrivibile alla categoria della condotta giuridicamente rilevante? In questo

caso, infatti, il *gap* si pone per la qualificazione del mero pensare dell'utente, rilevato dall'interfaccia, come azione e dunque come condotta. Le ricadute sul piano giuridico sono evidenti: se si desse una risposta affermativa, tramonterebbe quel principio di materialità cardine del moderno diritto penale liberale.

A ben vedere, le ricadute sono ancora più profonde della mera trasformazione di alcune categorie giuridiche.

Ora, tra un passato riduzionista o compatibilista e un futuro transumano, si colloca il presente. Al di là di questa eterogeneità di approcci, che oscillano tra un rimanere ancorati a un tradizionale paradigma umano e un disperato tentativo di superarlo, si incontra l'individuo. Occorre prestare particolare attenzione agli strumenti attualmente a disposizione per tentare di saltare il *gap*. La trattazione, in particolare, verterà su due beni giuridici, dei quali può essere titolare questo "nuovo" individuo, che devono essere tenuti in considerazione nella ricerca della responsabilità per evitare un loro possibile sacrificio e pregiudizio: l'integrità mentale, anche nella sua accezione di "privacy neurale", e la libertà di autodeterminazione dell'utente, la tutela della quale, in un prospettiva *de iure condendo* è possibile, secondo alcuni, attraverso il riconoscimento del diritto alla c.d. libertà cognitiva.

Verificare l'opportunità o meno della tutela dell'autodeterminazione attraverso la difesa del nuovo "neurodiritto" alla libertà cognitiva potrebbe rappresentare, alla fine, l'occasione per impostare il problema del "*Responsibility Gap*" su basi diverse, quantomeno sotto il profilo dell'intenzione, e di giungere così a una possibile conclusione.

Il punto di partenza è, ancora una volta, una domanda. Se l'obiettivo è il superamento del *gap*, qual è il costo della responsabilità?

5.2 Integrità mentale

Il primo bene giuridico di questo nuovo individuo che potrebbe essere messo a rischio in vari modi ¹ è definito, da alcuni Autori², come integrità mentale.

¹Uno di questi potrebbe essere, ad esempio, l'installazione di elettrodi a livello intracranico per potenziare la memoria episodica dell'utente. L'applicazione delle BCI in questo senso è stata dimostrata in Burke et al., "Brain computer interface to enhance episodic memory in human participants", 2015. In particolare, attraverso il loro utilizzo è possibile stimolare un *boost* nelle capacità mnemoniche dei pazienti. L'installazione delle interfacce in questo senso permetterebbe, ad esempio, a un soggetto che presenta primi sintomi della sindrome di Parkinson di ottenere un potenziamento nella memoria in riferimento a singoli eventi.

²Vd. Ienca e Andorno, "Towards new human rights in the age of neuroscience and neurotechnology", 2017.

Le difficoltà di inquadramento e di ricostruzione della fattispecie dell'integrità mentale in esame sorgono per l'assenza di riferimenti di diritto positivo all'ondata di "neurociviltà" degli ultimi anni nell'enunciazione di diritti fondamentali, in primis, ad esempio, nella formulazione dell'art. 3 della Carta, che risale al 2000, un'era in cui lo sviluppo delle interfacce era ancora in una fase embrionale.

Solo recentemente si stanno cogliendo, in alcuni ordinamenti, segnali di una crescente sensibilità a temi che si presentano per l'impiego delle neuroscienze. In particolare, sotto il profilo della tutela di nuove situazioni giuridiche "neurofocalizzate", uno dei primi esempi in questo senso è presentato dal Cile. Attraverso due progetti di riforma della Costituzione, infatti, il legislatore cileno ha puntato a reprimere ogni ingerenza esterna nella mente degli individui, tutelando l'integrità mentale e psicologica, anche attraverso la protezione dei dati ottenibili da tali interferenze³. Ancora, in Spagna il Segretario di Stato per l'Intelligenza Artificiale ha recentemente pubblicato una "Carta dei diritti digitali" che annovera i neurodiritti tra i diritti dei cittadini per la nuova era digitale. In Italia, l'autorità garante per la protezione dei dati personali ha promosso il tema dei neurodiritti nella Giornata europea della protezione dei Dati Personali del 28 gennaio 2021. Il Comitato direttivo sulle Neurotecnologie dell'OCSE nel dicembre del 2019 ha approvato la "Raccomandazione sull'Innovazione Responsabile nelle Neurotecnologie" che considera i due neurodiritti della privacy neurale e della libertà cognitiva. Il Consiglio d'Europa, nel frattempo, ha presentato un piano d'azione strategico quinquennale per verificare l'effettiva riconducibilità dei neurodiritti ai diritti umani già previsti dalle Convenzioni Europee. Non da ultimo, una proposta di "Dichiarazione Universale sulle Neuroscienze e i Diritti Umani" è stata avanzata presso l'UNESCO.

All'interno di questo panorama si è cercato di fornire una definizione di "integrità mentale", nonché di delineare gli elementi essenziali di un attentato a questa⁴. Sotto il primo profilo, ci si può rifare alla ricostruzione per la quale l'integrità mentale coincide con l'idea di "intangibilità dei propri stati mentali rispetto a interventi esterni"⁵, esprimibile in un "*interest in not having at least some of his mental states intentionally altered by others in*

³Bosoer, "Chile at the Forefront of Neurorights Protection", 2021. La riforma costituzionale è stata tuttavia bocciata nel referendum del 4 Settembre 2022.

⁴Ancora una volta, un prezioso contributo lo si ritrova in Ienca, "Intelligenza²: per un'unione di intelligenza naturale e artificiale", 2019, pag. 161.

⁵Fuselli, "Neurotecnologie e tutela dell'integrità psichica. Profili filosofico giuridici di un mutamento in atto", 2020.

certain ways”⁶. Sotto il secondo profilo, tre sono i requisiti di un’azione per poter essere qualificata come minaccia all’integrità mentale: deve presupporre l’accesso e la manipolazione dei segnali neurali recepiti, deve avvenire in assenza del consenso informato del soggetto e provocare danni fisici e/o psicologici.

Il riconoscimento di un nuovo neurodiritto all’integrità mentale non è, tuttavia, una soluzione indiscussa in letteratura per vari motivi.

In primo luogo, Andrea Lavazza rileva come la mera tutela giuridica dell’integrità mentale sia insufficiente: pur essendo idonea a punire eventuali sue violazioni, non è in grado di prevenirle. L’elaborazione teorica di un possibile diritto deve essere accompagnata dall’implementazione, sul piano tecnico, di una limitazione alle funzionalità dei neurodispositivi idonei a incidere sul bene giuridico in esame⁷. Questa soluzione tecnica si dovrebbe tradurre, nella pratica, nell’incorporazione all’interno dei *devices* di un sistema in grado di rilevare e segnalare ogni monitoraggio dell’attività cerebrale e diffusione dei dati ricavabili non autorizzati, garantendone l’interruzione⁸.

L’effettiva realizzazione di questa soluzione incontra numerosi ostacoli.

Innanzitutto, se si vuole evitare che il sistema blocchi automaticamente ogni interferenza, l’utente dovrebbe riuscire a capire quali siano le attività “non autorizzate” che possono arrecargli pregiudizio. Ciò comporta, a sua volta, che si riesca a installare una funzione di controllo attivabile dall’interessato: un’attività difficile da realizzare sia sotto un profilo tecnico sia sotto un profilo economico e commerciale, in quanto presuppone un’interazione e una collaborazione a più livelli tra legislatore, produttore e programmatore⁹.

In secondo luogo, un altro ordine di problemi legato al possibile riconoscimento del diritto all’integrità mentale sorge in riferimento all’estensione della tutela concessa, che porta con sé l’interrogativo circa il grado di paternalismo che deve assumere un possibile intervento del legislatore. In altre parole, gli strumenti che possono impedire di sacrificare l’integrità mentale nella ricerca della responsabilità presuppongono il superamento del *gap* stesso: presuppongono, cioè, che si sia individuato con sufficiente certezza il punto in cui finisce l’interfaccia e inizia l’uomo, e viceversa. Le questioni sorgono in particolar modo a seguito dell’emersione in letteratura della

⁶Birks e Buyx, “Punishing intentions and neurointerventions”, 2018, pag. 133.

⁷Lavazza, “Freedom of thought and mental integrity: The moral requirements for any neural prosthesis”, 2018.

⁸Silvia Inglese e Lavazza, “What Should We Do With People Who Cannot or Do Not Want to Be Protected From Neurotechnological Threats?”, 2021.

⁹*ivi*.

c.d. tesi della mente estesa¹⁰. Secondo questa ricostruzione, così come un quaderno in cui sono annotate le informazioni più utili per vivere la propria quotidianità rappresenta, per un soggetto affetto da Alzheimer, una vera “estensione” della propria mente, allo stesso modo un neurodispositivo che rallenta una possibile malattia degenerativa può essere qualificato come un “veicolo” per le capacità cognitive della mente, configurandone una estensione, anche se ciò potrebbe comportare un pregiudizio all’integrità mentale del soggetto, specialmente se installato quando la patologia non è ancora in uno stato avanzato¹¹.

Uno strumento individuato da Andrea Lavazza per far fronte a questi nuovi problemi consiste nella stipulazione di una convenzione internazionale, così da ridurre i rischi di un possibile raggiramento, che regoli tanto la posizione del produttore e programmatore quanto quella del consumatore. Il primo potrebbe essere tenuto a implementare nei dispositivi il sistema di monitoraggio e di rilievo dei segnali non rilevati, così che il secondo possa bloccarli o dare il proprio consenso, sempre assicurando, tuttavia, la protezione dei dati ricavabili¹². Ancora una volta questa prospettiva potrebbe essere una soluzione. Essa, tuttavia, si colloca a monte del *Responsibility Gap*, nel momento della costruzione dei dispositivi, non riuscendo a trovare una via di uscita a valle del problema. Il “chi” che ha agito rimarrebbe ancora senza nome.

Se questo è il sostrato tecnico e neuroscientifico sul quale si inserisce la ricostruzione all’integrità mentale, si deve procedere ora a verificare il suo impatto sulle categorie giuridiche, mettendo alla prova la sua effettiva portata innovativa con la nota categoria dell’integrità psichica.

Ci si deve interrogare, in particolare, se l’aggettivo “psichico” che qualifica l’integrità psichica possa essere considerato come un concetto “ombrello” al quale ricondurre anche l’integrità mentale intesa come protezione della dimensione neurale da possibili interventi per mezzo delle nuove neurotecnologie.

Se l’aggettivo psichico racchiude tutto ciò che è attiene alla sfera interiore, se ricondotto all’integrità psichica si può intendere essenzialmente come correlato alla salute¹³, intesa, come si rinviene nella stessa definizione fornita dall’OMS, come uno “stato di completo benessere fisico, psichico e menta-

¹⁰Clark e Chalmers, “The extended mind”, 1998.

¹¹Ritchie et al., “CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer’s disease dementia and other dementias in people with mild cognitive impairment (MCI)”, 2017.

¹²*ivi*

¹³Nisco, *La tutela penale dell’integrità psichica*, 2012, pag. 5.

le.”. La Consulta italiana ha poi ricondotto tale diritto all’alveo dei diritti costituzionalmente garantiti nella sentenza 233/2003, considerandolo come corollario del diritto alla salute sancito all’art. 32, trovando anche un riferimento comunitario nell’art. 3, co. 1, della Carta dei Diritti Fondamentali dell’Unione Europea, che sancisce che “ogni individuo ha diritto all’integrità fisica e psichica”. In questo senso, qualificandosi l’integrità mentale come assenza di interventi esterni, e dunque come permanenza del soggetto in uno stato di salute, non apparirebbe necessaria l’introduzione di questo nuovo neurodiritto, potendosi infatti ricondurre la tutela della mente al tradizionale paradigma dell’integrità psichica, e proteggere allo stesso tempo l’organo cerebrale attraverso il ricorso al concetto di integrità fisica.

L’integrità psichica potrebbe risultare pregiudicata sotto un altro rilevante profilo.

5.2.1 Privacy neurale

Nel momento in cui è possibile conservare la “cronologia” degli impulsi neurali inviati dal soggetto nel corso dell’utilizzo della BCI, questi subiscono un fondamentale processo di “datificazione” che comporta il sorgere di problematiche relativamente alla qualificazione giuridica di tali dati e alla loro conservazione. Il costo della ricerca della responsabilità e del superamento del *gap* può consistere, infatti, anche in una risemantizzazione del concetto di privacy, o, quantomeno, nell’estensione della protezione accordata ai dati c.d. sensibili.

Il punto di partenza è dato dalla tutela dei “tradizionali” dati sensibili e la disciplina del loro trattamento è data da più fonti normative.

L’articolo 12 della Dichiarazione universale dei diritti umani afferma: “nessun individuo potrà essere sottoposto ad interferenze arbitrarie nella sua vita privata, nella sua famiglia, nella sua casa, nella sua corrispondenza, né a lesione del suo onore e della sua reputazione. Ogni individuo ha diritto ad essere tutelato dalla legge contro tali interferenze o lesioni”.

Similmente, l’art. 8, co. 1, della CEDU dispone che “ogni persona ha diritto al rispetto della propria vita privata e familiare, del proprio domicilio e della propria corrispondenza”.

Ancora, a livello comunitario il punto di riferimento per il riconoscimento e la tutela del diritto alla privacy è la direttiva 95/46, “relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati”. Più recentemente, il regolamento 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016, divenuto operativo dal 25 maggio 2018, relativo alla protezione delle perso-

ne fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati ha abrogato la direttiva 95/46/CE: come si sa, la *ratio* è quella di rafforzare la protezione dei dati personali di cittadini dell’Unione Europea, unificando e rendendo omogenea la normativa sulla privacy all’interno dell’Unione. In Italia la sua entrata in vigore ha comportato l’abrogazione degli articoli del Codice per la protezione dei dati personali (d.lgs. n. 196/2003) con esso incompatibili. L’art. 9 GDPR, infatti, include tra le “categorie particolari di dati personali” oggetto di tutela rafforzata i dati già specialmente protetti dal D. Lgs. 196/2003 in quanto sensibili.

Il sistema normativo così delineato fornisce la tradizionale disciplina di riferimento in tema di privacy e trattamento di dati personali. Una volta ricordato questo fondamentale quadro legale, per comprendere¹⁴ in che modo concetti quale quello di “privacy”, o “dato personale” vengono ridefiniti dalla simbiosi tra uomo e macchina si può portare all’attenzione uno dei primi saggi sul tema, redatto nel 1890 dai giuristi statunitensi Samuel Warren e Louis Brandeis¹⁵, nel quale si rinviene il paradigma originale di privacy inteso come “diritto a non far parlare di sé”. È un paradigma che poi è stato ampliato da Alan Westin nella nozione di “privacy informativa”, ossia il diritto di ciascuno di determinare le modalità, le quantità e la natura dei dati che vengono comunicati ad altri¹⁶. In altre parole, il diritto a controllare le informazioni attraverso cui le persone “presentano sé stesse agli altri”¹⁷: tale controllo farebbe sì che ciascuno abbia diritto anche di presentarsi “falsamente” agli altri, assumendo atteggiamenti che dall’esterno potrebbero apparire ipocriti, ma che, tuttavia, sono volutamente mostrati dall’individuo in quanto conformi all’immagine che vuole dare agli altri. Come affermato in letteratura, l’onestà totale, alla quale si può giungere attraverso strumenti di *brain reading*, è nemica della privacy¹⁸.

È proprio sotto il profilo del controllo di questa condivisione da parte del soggetto che si colgono le frizioni con le dinamiche delle connessioni delle BCI.

I segnali cerebrali ottenuti tramite EEG o fMRI, infatti, possono contenere la codificazione di dati sensibili dell’utente, quale lo stato di salute

¹⁴Come punto di riferimento per la trattazione è stato preso Ienca e Andorno, “Towards new human rights in the age of neuroscience and neurotechnology”, 2017.

¹⁵Warren e Brandeis, “The right to privacy”, 1989.

¹⁶Westin, “Privacy and freedom”, 1968.

¹⁷Gometz, “La privacy della mente: alcune riflessioni sul rapporto tra protezione dei dati personali e libertà di pensiero”, 2020, che riprende Marmor, “What Is the Right to Privacy?”, 2015.

¹⁸Così Marmor, “What Is the Right to Privacy?”, 2015, pag. 38.

o la sua condizione psicofisica¹⁹ e nessuno specifico strumento, normativo o tecnologico, protegge i dati cerebrali dall'essere estratti alla stessa maniera degli altri dati sensibili²⁰, anche alla luce del *quid* che distingue i primi dai secondi: l'individuo, infatti, non può applicare un controllo cosciente sull'attività cerebrale, stabilendo quali segnali possono essere registrati e quali no, dal momento che le tecniche di registrazione suddette si collocano “al di sotto” della soglia di consapevolezza.

Ciò nonostante, si potrebbe comunque pensare di estendere la tutela concessa ai dati sensibili “tradizionali”, data dalle fonti suddette, anche a queste nuove informazioni personali: alcuni hanno visto la *ratio* di questa tutela in quella che è stata definita come una “ragionevole aspettativa di privacy”, già presente e tutelata nell'ordinamento statunitense in relazione al prelievo di campioni di sangue²¹.

L'estensione della tutela, tuttavia, incontra un ostacolo rilevante. Come è stato notato, infatti, tutelare queste informazioni personali deve comportare anche la protezione della fonte di queste informazioni, dal momento che non è possibile scindere il dato dalla sua fonte di produzione, ossia l'attività neurale²².

Ecco, allora, che per proteggere le onde cerebrali, intese come dati, e le informazioni ricavabili a partire da queste, è stato delineato come nuovo, possibile, “neurodiritto” il diritto alla privacy mentale.

In questo modo si accorderebbe uno strumento di tutela contro la “fuga” dei dati neurali sensibili, registrati in neurodispositivi, nel grande flusso di dati che ogni giorno attraversa la nostra società: un fiume in piena di dati che si svolge in tempo reale e che non può essere monitorato dal paziente. Se lo si colloca poi in un'epoca di datificazione e commercializzazione dei c.d. *big data*, sono evidenti le necessità di porre degli argini al suo straripamento²³.

¹⁹In letteratura si parla di *functional connectome fingerprinting*, delle vere e proprie “impronte digitali neurali” che permettono il riconoscimento dell'individuo. Vd. Finn et al., “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity”, 2015.

²⁰Cfr art. 9 del GDPR, che contiene la definizione tanto di “dato personale” quanto di “trattamento”.

²¹Si veda Shen, “Neuroscience, mental privacy, and the law”, 2013.

²²Ienca, “Intelligenza²: per un'unione di intelligenza naturale e artificiale”, 2019, pag. 156. Ienca lo definisce il problema di “Inception”, richiamando la nota opera cinematografica, e auspica una futura e più ampia tutela nei confronti dell'attività neurale.

²³Per un contributo più approfondito circa i *big data* e i dati cerebrali, si veda Kellmeyer, “Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices”, 2021. L'Autore, in particolare, auspica l'avvio di un dibattito che interessi ogni livello della società dai programmatori delle IA all'utente finale. È questo un modo per garantire da un lato una maggiore conoscenza di queste nuove tecnologie,

Riconoscendo questa nuova situazione giuridica soggettiva, si accorderebbe a Mary una protezione anche contro attacchi di *neurohackers*, sia diretti, che indiretti.

Sono indiretti gli attacchi volti a carpire le banche dati in cui si conservano le informazioni ottenute tramite tecniche c.d. di *brain-reading*: sotto questo profilo non si discosterebbero da un tradizionale attacco hacker a una tradizionale banca dati²⁴.

Al contrario, molto più insidioso è un possibile attacco diretto all'attività neurale della vittima: in questo, caso, infatti, oltre alla sottrazione di dati sensibili, si può pregiudicare la stessa integrità mentale del soggetto, causando danni diretti al cervello. Un *brain-hacking*, o hackeraggio del cervello²⁵, che avviene nelle fasi di misurazione, decodifica e feedback della rilevazione dei dati da parte dell'interfaccia²⁶, infatti, non solo è in grado di carpire le informazioni, ma anche di influenzare il controllo dell'utente BCI sull'interfaccia, diminuendolo o aumentandolo di intensità, giungendo addirittura a poterlo "dirottare".

Vi è, poi, un altro aspetto da prendere in considerazione, legato alla protezione dei dati ricavabili dalla registrazione dell'attività neurale. Se l'interfaccia è attivata dai pensieri dell'utente, e questi ultimi hanno comportato l'attivazione, ad esempio, degli arti robotici lo stesso pensiero potrebbe essere considerato come "offensivo" e dannoso. Captando i segnali neurali dell'utente, quest'ultimo potrebbe essere tenuto responsabile sulla base di un mero pensiero, con buona pace dei principi liberali di materialità espressi dal brocardo latino *cogitationis poena nemo patitur*. Si potrebbe allora pensare, anche alla luce di una certa letteratura, in particolare anglosassone, che così come è sanzionato la detenzione abusiva di armi a prescindere dal compimento di determinati reati per loro mezzo, così dovrebbero essere sanzionati determinati pensieri offensivi e dannosi, a prescindere dall'effettivo comportamento materiale²⁷.

sia da parte del legislatore che dell'utente, così da permettere un uso più consapevole alla luce delle ricadute etico e giuridiche che ne derivano, dall'altro una riduzione del rischio di strumentalizzazione dell'utente per divenire un semplice mezzo per lo sviluppo delle IA.

²⁴In Italia il tema della *Cybersecurity* è particolarmente rilevante. Secondo il *Global Cybersecurity Outlook 2022* diffuso dal *World Economic Forum*, quattro aziende su dieci hanno dichiarato di aver subito attacchi informatici negli ultimi due anni.

²⁵Vd. Ienca e Haselager, "Hacking the brain: brain-computer interfacing technology and the ethics of neurosecurity", 2016.

²⁶Sono queste tre delle quattro fasi del ciclo di connessione dell BCI. Si veda il primo capitolo, in particolare il paragrafo "Classificazione delle BCI".

²⁷Così ad esempio Frederick Schauer, il quale critica fortemente la tutela accordata dal Primo Emendamento della Costituzione americana al c.d. *free speech* in quanto impedisce

Per evitare la distopica realtà nella quale “quando l’ultimo bit di informazione cerebrale sarà stato assorbito nell’infosfera digitale, niente sarà più considerabile privato”²⁸, il capo bioeticista della NASA Paul Root Wolpe ha esortato a considerare il cranio come un dominio di privacy assoluta, nel quale non si può entrare, senza la volontà della persona, né con una ordinanza del tribunale, né per motivi di pubblica sicurezza militare o nazionale, neanche se in gioco vi è il bene pubblico²⁹, considerando dunque il diritto alla privacy mentale come un diritto assoluto.

Tuttavia, se l’elemento discriminante è dato dal consenso dell’interessato, resta ancora velato da molte incertezze il momento in cui è prestato il consenso da parte di quest’ultimo: i neurodispositivi, come già visto, potrebbero comportare un mutamento nella personalità del soggetto tale da compromettere la sua autonomia decisionale³⁰.

L’emersione della nuova categoria della privacy neurale e la questione della qualificazione dei dati ricavabili dall’attività cerebrale si collega a doppio filo, infatti, a un’altra importante categoria.

Il salto nel vuoto del *Responsibility Gap* potrebbe anche costare il sacrificio della capacità di autodeterminazione.

5.3 Autodeterminazione e libertà cognitiva

Il brocardo latino *nemo tenetur se detegere* esprime il principio di diritto processuale penale in forza del quale nessuno può essere obbligato ad affermare la propria responsabilità penale e, dunque, ad autoincriminarsi. La regola trova cittadinanza in vari ordinamenti, sia di *civil law* che di *common law*³¹.

Nell’esperienza giuridica italiana è considerata l’espressione, come si sa, del diritto al silenzio, posto a presidio della libertà morale e personale dell’indagato e dell’imputato, il quale ha diritto di scegliere come e se difendersi, anche se colpevole, in attuazione del più generale diritto alla difesa affermato sul piano costituzionale dall’art. 24 della Costituzione e in armonia con

di reprimere gli *harmful thoughts*, pensieri nocivi per la società, che potrebbero risolversi nel compimento di atti altrettanto nocivi. Vd Schauer, “On the Distinction Between Speech and Action”, 2014.

²⁸Ienca, “Intelligenza²: per un’unione di intelligenza naturale e artificiale”, 2019, pag. 143.

²⁹Wolpe, “Is my mind mine? Neuroethics and brain imaging”, 2009.

³⁰Sul punto si veda Macioce, “Informed consent procedures between autonomy and trust”, 2019.

³¹Come noto, il principio è sancito dal V emendamento della Costituzione americana.

i principi del c.d. giusto processo, sanciti a livello comunitario dall'articolo 6 della CEDU.

In un'impostazione liberale del processo penale, l'interesse alla repressione dei reati cede innanzi all'interesse di proteggere la libertà di autodeterminarsi del singolo. L'interesse dell'ordinamento ad accertare la responsabilità deve cedere di fronte alla capacità di autodeterminarsi dell'individuo. Lo scopo dell'istituto è “[...] quello di rafforzare la libertà morale dell'imputato per sollevarlo dallo stato di soggezione psicologica in cui possa venire a trovarsi a cospetto dell'autorità e per porlo a riparo da eventuali pressioni che su di lui possano essere esercitate”³².

Tuttavia, alla domanda fondamentale: “sei colpevole?”, il soggetto dotato di interfaccia potrebbe rispondere senza proferire parola. Al quesito potrebbe rispondere, infatti, direttamente il suo cervello. Monitorando la sua attività cerebrale, l'interessato non potrebbe tacere la risposta. L'individuo nei confronti del quale è rivolto il quesito potrebbe non sapersi proteggere dalla nudità del suo stesso cervello. Il silenzio potrebbe portare a un'autoincriminazione così come una parola pronunciata ad alta voce, attentando alla sua libertà di autodeterminazione. In ambito italiano, ciò potrebbe configurare una violazione degli articoli 2,3,24 e 11 della Costituzione, che tutelano a livello costituzionale il diritto al silenzio, nonché dell'art. 188 del codice di procedura penale, secondo il quale “non possono essere utilizzati, neppure con il consenso della persona interessata, metodi o tecniche idonei a influire sulla libertà di autodeterminazione o ad alterare la capacità di ricordare e di valutare i fatti”.

In una dimensione processualpenalistica, il problema si pone in primo luogo per la difficoltà nel ricondurre le tecniche di lettura dell'attività cerebrale, quali la EEG o la fMRI, a mezzi di cui all'art. 188. Vari sono i fattori che devono essere tenuti in considerazione³³: la validità teoretica della sua applicazione, l'affidabilità per la ricostruzione del fatto in esame, la possibilità di servirsi di un adeguato criterio per interpretare i risultati. Non da ultimo, l'eventuale sottoposizione dell'interessato alle tecniche in esame deve avvenire sempre nel rispetto della sua integrità fisica e mentale, anche là dove questi dia il suo consenso, potendosi configurare una violazione dell'art. 5 della Costituzione, che vieta atti di disposizione del proprio corpo che “cagionino una diminuzione permanente della integrità fisica”.

Inoltre, da non dimenticare è il disposto dell'art. 224-bis c.p.p secondo il

³²È questa la posizione della Corte costituzionale nella risalente ma imprescindibile sentenza n.34 del 1973.

³³Seguo qui Folgieri, “Brain computer interface and transcranial magnetic stimulation in legal practice and regulations”, 2020.

quale “non possono in alcun modo essere disposte operazioni che contrastano con espressi divieti posti dalla legge o che possono mettere in pericolo la vita, l’integrità fisica o la salute della persona o del nascituro, ovvero che, secondo la scienza medica, possono provocare sofferenze di non lieve entità”. I mezzi per superare quel *gap* di responsabilità possono incontrare essi stessi dei limiti, che non fanno altro che renderlo più profondo.

In secondo luogo, anche ammessa la loro applicabilità, rimane ancora incerto il significato da attribuire alle risultanze di tali tecniche: secondo il filosofo Daniel Dennett, ad esempio, nei segnali neurali è possibile leggere solo il “veicolo” del pensiero, ma non anche il suo contenuto; in altre parole, si coglierebbe solo il significante, non riuscendo a carpirne il significato³⁴.

La questione non riguarda solo il problema dell’autoincriminazione per un fatto commesso, ma anche per un fatto illecito che si *commetterà* in futuro, aprendo a molteplici scenari fino ad ora considerati fantascientifici³⁵.

In un articolo del 2013 lo psicologo Eyal Aharoni e il neuroscienziato Michael Gazzaniga, unitamente ad altri scienziati e ricercatori, hanno pubblicato un articolo in cui illustravano i risultati di uno programma di sperimentazione su un gruppo di 96 ex-detenuti maschi. Nel corso dei quattro anni successivi alla loro scarcerazione furono sottosti a periodiche fMRI durante l’espletamento di determinate attività informatiche, appositamente ideate per verificare la loro impulsività e la loro capacità di prendere decisioni rapide. Al termine del periodo di studio il team di scienziati ha dimostrato che i soggetti che mostravano una bassa attività nella zona della corteccia cingolata anteriore, detta ACC, sono esposti a un maggior rischio di recidiva, quantificabile statisticamente nel doppio della probabilità di commettere ulteriori reati rispetto ai soggetti che presentavano una più intensa attività³⁶.

I risultati ai quali sono approdati Aharoni e Gazzaniga sono innegabilmente gravidi di conseguenze sia sul piano giuridico che sul piano etico.

Sotto un profilo giuridico, possibili ricadute si danno rispetto alla portata applicativa dell’art. 133 del codice penale italiano che impone al giudice di tenere conto, ai fini dell’applicazione della pena, anche della “capacità a

³⁴Per un approfondimento della questione, si veda Hofstadter et al., *L’io della mente. Fantasie e riflessioni sul sé e sull’anima*, 1985.

³⁵Nel 1956 uscì il romanzo *The Minority Report*, che narra la storia di una particolare divisione di polizia, la divisione *Precrimine*, adibita alla repressione di illeciti penali molto particolari: reati che sarebbero stati commessi nel *futuro*. La previsione della loro prossima realizzazione avviene grazie ai particolare poteri psichici di precognizione di tre individui, chiamati *Precog*.

³⁶Gli studi descritti qui sono presentati in Aharoni et al., “Neuroprediction of future rearrest”, 2013.

delinquere”, che deve essere desunta, tra le altre cose, anche dai “motivi a delinquere e dal carattere del reo”³⁷. Là dove il giudice dovesse tenere conto dei risultati degli esperimenti suddetti nell’esercizio del suo potere discrezionale di applicazione della pena, si negherebbe alla radice quella stessa capacità di autodeterminazione che è tutelata su più fronti dall’ordinamento giuridico. All’imputato sarebbe comminata una pena che terrebbe conto dell’attitudine a commettere ulteriori reati sulla base di una attività involontaria e inconsapevole, assumendo come giustificazione la presenza di un “potenziale biomarcatore neurocognitivo per un comportamento antisociale persistente”³⁸.

Sotto il profilo etico, invece, ciò che potrebbe essere fonte di inquietudine non sono le implicazioni etiche dei test condotti, comunque riconducibili alla sempre presente e aperta questione dell’esistenza del libero arbitrio, ma la natura dei mezzi che possono essere scelti per farvi fronte. Se vi è la conoscenza che un soggetto commetterà un reato sulla base del monitoraggio della sua attività cerebrale del cervello, si potrebbe agire direttamente su questo per far venire meno la fonte di tale “devianza”.

Alla base di questa pretesa vi è, tradizionalmente, l’equiparazione dei comportamenti socialmente riprovevoli, quali possono essere quelli criminali, a stati di malattia: così come da un punto di vista medico è non solo naturale, ma doveroso, applicare una terapia per curare una malattia, così potrebbe divenire doveroso combattere le devianze dei criminali, o futuri tali. In questo senso, l’impiego di BCI, ad esempio attraverso l’applicazione di dispositivi che permettono una DBS costante e sottratta al controllo del soggetto o l’installazione di un dispositivo non invasivo che monitora e inibisce determinate onde cerebrali, si potrebbe rivelare come l’inizio non già di un percorso di miglioramento dell’uomo, come auspicato dai transumanisti, ma di distruzione dello stesso.

Le conseguenze di quella che si può definire una “neurociviltà forzata”³⁹ possono essere degne della migliore, o peggiore, distopia.

Per scongiurare scenari di interventismo eugenetico ispirate a un perfezionismo, lesivo della libertà e della dignità umana, è stata prospettata l’introduzione della c.d. libertà cognitiva, considerato come l’espressione

³⁷ Art. 133 co. 2, c.p.: “Il giudice deve tener conto, altresì, della capacità a delinquere del colpevole, desunta: 1) dai motivi a delinquere e dal carattere del reo; 2) dai precedenti penali e giudiziari e, in genere, dalla condotta e dalla vita del reo, antecedenti al reato; 3) dalla condotta contemporanea o susseguente al reato; 4) dalle condizioni di vita individuale, familiare e sociale del reo.”

³⁸ Aharoni et al., “Neuroprediction of future rearrest”, 2013.

³⁹ Paolo Sommaggio, “Neurociviltà”, 2014.

di due diritti complementari: il diritto di ogni individuo a usare le nuove neurotecnologie e il diritto a non subire un uso imposto e non consensuale, garantendo, così, da un lato il diritto ad alterare volontariamente i propri stati mentali, dall'altro il diritto a rifiutare una loro modificazione⁴⁰.

La libertà cognitiva dovrebbe essere così assunta come un inalienabile diritto umano, che spetta a ogni uomo in quanto uomo⁴¹, nonché un principio cardine che deve ispirare il legislatore nel regolamentare le neurotecnologie.

In altre parole, si può definire come “*the right and freedom to control one’s own consciousness and electrochemical thought processes is the necessary substrate for just about every other freedom*”⁴²: è la condizione *sine qua non*, il sostrato che vi deve essere perché possa esistere la libertà di pensiero, la libertà di religione, la libertà di espressione.

Due sono le possibili sue declinazioni: la libertà cognitiva declinata in una dimensione *negativa* e in una *positiva*⁴³.

Nella sua accezione negativa, la libertà cognitiva significa tutelare l'individuo da ogni possibile interferenza esterna di terzi, siano questi rappresentati tanto da una pubblica autorità quanto da privati malintenzionati. Nella sua formulazione in negativo, infatti, la libertà cognitiva dovrebbe essere ricondotta alle situazioni giuridiche che impongono agli altri consociati, nonché a pubbliche istituzioni, di astenersi da ogni intervento forzato che possa arrecare pregiudizio al suo titolare. Sotto questa luce il problema risiederebbe nel fatto che non tutti gli attentati alla libertà cognitiva sono necessariamente posti in essere coattivamente⁴⁴. L'individuo, infatti, potrebbe subirne uno senza averne consapevolezza.

La libertà cognitiva declinata in positivo, invece, permette a chiunque di autodeterminarsi garantendo la possibilità di dotarsi di BCI, conseguendo il loro potenziale cognitivo desiderato. Sotto questo profilo, sorge tuttavia la questione circa quali operazioni siano tecnicamente ed eticamente possibili e quali no, anche in un'ottica di sicurezza e di riduzione del rischio di incidenti a lungo termine.

Se questo è il quadro teorico nel quale si colloca la libertà cognitiva, ci si deve chiedere in che modo questo “neurodiritto” si rapporta alle categorie e

⁴⁰Così J.-C. Bublitz, “My mind is mine!? Cognitive liberty as a legal concept”, 2013, pag. 234.

⁴¹Così in Van Banning et al., “Human Rights Reference Handbook”, 2004.

⁴²J.-C. Bublitz, “My mind is mine!? Cognitive liberty as a legal concept”, 2013, pag. 235. Si ritrova questa posizione anche in Sententia, “Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition”, 2004.

⁴³Seguo qui Paolo Sommaggio e Mazzocca, “Cognitive liberty and human rights”, 2020.

⁴⁴Ienca e Andorno, “Towards new human rights in the age of neuroscience and neurotechnology”, 2017, pag. 24.

alle tutele fornite dall'ordinamento giuridico. Questa nuova situazione può essere ricondotto a una tutela già fornita dal diritto positivo, o è necessaria una evoluzione e rivoluzione dell'ordinamento per permetterne una tutela?

Sono state formulate tre possibili risposte⁴⁵: la libertà cognitiva potrebbe essere in primo luogo considerata come un “prerequisito” per tutti i diritti posti a tutela dell'attività neurale. In secondo luogo, la sua tutela potrebbe essere data da una interpretazione estensiva dei diritti attualmente contemplati dall'ordinamento giuridico. Infine, la libertà cognitiva potrebbe essere ricostruita come un diritto a sé stante.

Per rispondere a questo quesito vengono qui in rilievo alcuni riferimenti normativi.

L'articolo 13 della Costituzione Italiana, anche se non si riferisce esplicitamente alla libertà cognitiva, afferma che “è punita ogni violenza fisica e morale sulle persone comunque sottoposte a restrizioni di libertà”. La disposizione ricordata deve essere letta in combinato disposto con l'articolo 27 della Carta costituzionale, là dove vieta l'applicazione di pene contrarie al senso di umanità, affermandone la funzione rieducativa. Infine, sempre sul piano costituzionale, sono di interesse anche gli articoli 2, 3 e l'art. 32, che, come si sa, dispone, al suo secondo comma, che “nessuno può essere obbligato a un determinato trattamento sanitario se non per disposizione di legge. La legge non può in nessun caso violare i limiti imposti dal rispetto della persona umana”.

La disposizione che più rileva ai nostri fini è la prima menzionata, l'articolo 13. Ci si può chiedere, infatti, se la libertà cognitiva che viene qui descritta possa o meno essere ricondotta alla libertà morale e quali possono essere le possibili ricadute relativamente al problema del *Responsibility Gap* di cui si tratta.

In un imprescindibile saggio del 1960, Giuliano Vassalli definisce la libertà morale come “la libertà di volere, di poter determinare il proprio comportamento senza esterne imposizioni, distinta dalla libertà di agire in uno od altro rapporto della vita di movimento o di relazione [...]”, nonché come la libertà “della sfera psichica dell'uomo (coscienza, intelletto, volontà) da frodi o da ingiuste imposizioni esterne”⁴⁶.

Alla libertà morale intesa nel senso di libertà di volontà si è ricondotta poi, oltre alla libertà di pensiero e di coscienza, la c.d. libertà psichica, ricostruita da Vassalli come l'“inviolabilità della persona psichica”⁴⁷.

⁴⁵Paolo Sommaggio e Mazzocca, “Cognitive liberty and human rights”, 2020.

⁴⁶Vassalli, *Scritti giuridici - Volume III, Il processo e le libertà*, 1997, pag. 253 ss.

⁴⁷*ivi*, pag. 305. Se da un lato a Vassalli, secondo parte della dottrina (Vd.Nisco, *La tutela penale dell'integrità psichica*, 2012, pag. 38), si deve riconoscere l'intuizione di aver

Tornando alla questione della riconducibilità della libertà cognitiva alla libertà di cui all'art. 13, se si assume la libertà morale come assenza di imposizioni esterne che possono condizionare la sfera psichica, l'impiego coattivo di tecniche neuroscientifiche in grado di incidere su tale posizione configurerebbe un attentato alla libertà morale dell'individuo: non sarebbe necessaria l'introduzione del nuovo "neurodiritto" alla "*cognitive liberty*". L'individuo dotato di BCI sarebbe già tutelato dall'attuale quadro normativo, essendo sufficiente una estensione della tutela della libertà morale anche ai processi neuronali dell'utente di BCI data sia dall'articolo 13, co.2, della Costituzione, sia dall'articolo 188 del c.p.p.

L'introduzione della libertà cognitiva come diritto "neurofocalizzato" della libertà morale, invece, permetterebbe una protezione giuridica più aderente al progresso neuroscientifico in tema di rapporto con il proprio cervello, ma si baserebbe anche su una implicita quanto precisa presa di posizione ideologica. Il cervello, infatti, verrebbe visto come la "sede" dell'autodeterminazione dell'uomo, aprendo le porte, così, a un "tacito" riduzionismo "legale".

Considerare la libertà cognitiva come un diritto a sé stante, infatti, potrebbe precludere ogni tentativo di superare il *gap*.

5.3.1 Tra cause e motivi

Il problema di *chi* ha attivato l'interfaccia ancora una volta non troverebbe soluzione: separando il soggetto dalla sua attività cerebrale, escludendo, cioè, quest'ultima dalla sfera di libertà morale di cui gode la prima, per ricondurla a un diritto *ad hoc*, ancora una volta si imporrebbe la domanda: l'utente o il suo cervello? Ancora una volta, dunque, la conseguenza ultima consisterebbe nella frammentazione dell'individuo in due soggettività: cervello da un lato, utente dall'altro. L'interfaccia, infatti, è stata attivata dall'attività cerebrale: quest'ultima sarebbe considerata come una entità, un vero e proprio soggetto agente, distinto dal resto. Eppure non è *altro* dall'utente.

distinto la "libertà psichica" sia dall'integrità fisica sia dall'integrità morale, dall'altro si è comunque rilevata la persistente incertezza da attribuire alla categoria della libertà, in primis per la complicata relazione tra neuroscienze e diritto. In particolare, secondo la stessa dottrina, il punto di partenza fondamentale è considerare che la libertà si poggia su un elementare atteggiamento psichico, ossia la consapevolezza di essere padroni di se stessi. Inoltre, questa libertà positiva consiste nell'essere consapevole che "io" penso, voglio, agisco e sono responsabile delle mie scelte riuscendo a spiegarle facendo riferimento alle mie idee e finalità. Vd. Berlin, "Two concepts of liberty", 2017.

Non a diverse conclusioni potrebbe portare considerare la libertà cognitiva come un prerequisito, una chiave di lettura per tutti i diritti posti a tutela dell'attività neurale. Anche in questo caso, infatti, si deve procedere con cautela per non cadere nella stessa tendenza riduzionistica: se la libertà cognitiva è il sostrato sul quale si fonda l'esercizio di ogni altro diritto, ancora una volta si presupporrebbe che il cervello sia l'origine della capacità dell'uomo di essere libero, abbracciando un implicito riduzionismo.

La posizione preferibile, che forse è utile anche per impostare il problema del *Responsibility Gap*, oggetto della trattazione, potrebbe essere quella di considerare la libertà cognitiva, nella sua accezione di libertà di autodeterminazione, come già racchiusa nella libertà morale, ma non già sulla base dei processi fisici e neuronali, ma per il fatto stesso di poter considerare l'uomo come privo di costrizione: tutelare la libertà cognitiva significherebbe tutelare l'assenza di costrizione.

Il bene tutelato non sarebbe più il cervello, la tutela del quale comunque potrebbe essere ricondotta alla tutela dell'integrità fisica dell'individuo, essendo il cervello nient'altro che un organo del corpo umano, ma l'utente che rimane unito, che rimane *in-dividuo*, allontanandosi dunque da possibili ombre di riduzionismo.

La visione che si dovrebbe abbracciare, infatti, è compatibilista, seguendo la dottrina per la quale la libertà di volere non coincide con la totale assenza di cause, ma con la capacità di agire seguendo motivazioni proprie⁴⁸. Si dovrebbe allora recuperare la distinzione tra “cause” e “ragioni” di un comportamento: libertà è agire in conformità a ragioni, e non cause, proprie⁴⁹. L'agire di Mary non è stato causato dalla BCI, ma motivato dalla gelosia. Così facendo, dunque, secondo una determinata impostazione filosofica, si recupererebbe anche il concetto di “persona”, dal momento che è impossibile “interrogarsi sulla natura della volontà senza interrogarsi su quella dell'agente volontario in quanto soggetto di volizioni e azioni”⁵⁰.

Il tema del riconoscimento o meno della libertà cognitiva quale espressione della libertà morale fornisce anche una possibile chiave di lettura della questione del *Responsibility Gap*. Può essere l'occasione, infatti, se non di superare il *gap*, quantomeno di impostare un metodo nella ricerca della responsabilità. Evitando di considerare come limite la tutela della libertà cognitiva, che nasconde in sé il seme del riduzionismo, da un lato si tutelerebbe comunque la capacità di autodeterminazione, non potendo essere

⁴⁸Nisco, *La tutela penale dell'integrità psichica*, 2012, pag. 45.

⁴⁹*ibidem*. L'Autore riprende la dottrina tedesca, in particolare Schockendoff, “Der freie Wille – ein problemgeschichtlicher Abriss”, 2010, pag. 3 ss.

⁵⁰De Monticelli, *La novità di ognuno: Persona e libertà*, 2012, pag. 143.

intaccata da tecniche che permettono una possibile autoincriminazione.

Dall'altro, tuttavia, la stessa autodeterminazione che verrebbe a proteggersi potrebbe, così qualificata, permettere di iniziare a considerare l'utente come potenzialmente responsabile. Se ci si allontana dalla visione determinista insita nella libertà cognitiva e si abbraccia invece la libertà di autodeterminazione così come ricostruita, si inizia a ragionare non già in termini di cause, ma di motivi. Bisogna chiedersi: quale motivazione aveva il soggetto dotato di neurodispositivi?

Questa può essere l'occasione per cambiare la prospettiva di riferimento. Le incertezze che circondano il problema noto come *Responsibility Gap* si danno per il presupposto implicito sul quale si presenta il problema. Nell'*impasse* circa l'individuazione di chi ha agito il giurista cade solo a patto di confermare la premessa della questione, secondo la quale vi sarebbero, cioè, tre individui in gioco: utente, il suo cervello, ossia la sua attività neurale, e la sua interfaccia. È un assunto quasi autoevidente che comporta come conseguenza necessitata che ci debba chiedere non già quale sia la ragione dell'agire, ma la sua causa naturalistica. Ciò parte, tuttavia, da una premessa tanto implicita quanto discutibile: per capire chi ha agito è necessario guardare e tutelare il cervello, che è *altro* dalle altre due parti.

Nonostante permangano comunque innegabili incertezze antropologiche, giuridiche e tecnologiche circa la possibilità o meno di considerare il movimento degli arti robotici della donna come "azione", sotto il profilo dell'agente la domanda non deve più essere *chi* ha agito tra colui che è dotato di BCI e il suo cervello e *come* dal punto di vista tecnico, a meno di abbracciare ancora una volta una impostazione riduzionistica. Certo, l'interfaccia è stata attivata dall'attività cerebrale, ma la prima non è altro dalla seconda, a prescindere dalla volontarietà o meno degli impulsi cerebrali: questi sono pur sempre prodotti dall'attività cerebrale del cervello. La domanda che ci si deve porre, allora, è *perché* ha agito l'individuo. L'interfaccia, in questo senso, non farebbe altro che tradurre in movimento i "motivi" del pensare-agire dell'utente. Recuperando l'unità tra quest'ultimo e il suo cervello, il quesito del *Responsibility Gap* certamente non verrebbe risolto, ma potrebbe essere ridimensionato. Colui che è dotato di BCI è libero di attivare l'interfaccia secondo motivi propri per il solo fatto che il suo cervello era libero da condizionamenti, a meno di interventi in senso modificativo della sua integrità mentale, e l'interfaccia ha rilevato questa libertà traducendola in movimento.

Questa potrebbe essere una strada da percorrere: l'agire non è "causato" dall'interfaccia, o dal cervello, ma è stato motivato dalla persona, dal suo

“essere un soggetto”⁵¹, riconoscendo che è proprio quello che è stato definito un “vissuto di potere”⁵² a distinguerla dalla sua composizione animale, dal suo cervello, per renderla persona⁵³.

In definitiva, il problema del *Responsibility Gap* così come è stato fino a questo momento presentato e ricostruito, rappresenta una questione aperta solo sotto la possibile riqualificazione del concetto di azione, e forse non sotto il profilo del soggetto agente. Lo si chiami, quest’ultimo, cervello o impulso cerebrale, non ha importanza. L’utente non è il suo cervello, è *anche* il suo cervello e la sua attività cerebrale. Continuare a impostare la questione su *chi* ha attivato l’interfaccia significa aprire alla possibilità che, alla fine, quel *chi* può essere anche il suo cervello, in base dunque a un “seme” di riduzionismo che non solo conferma la definitiva sudditanza dell’esperienza giuridica al progresso neuroscientifico, ma nega alla base ogni possibilità di responsabilità.

⁵¹ *ivi*, pag. 31.

⁵² *ivi*, pag. 107.

⁵³ *ivi*, pag. 141.

Capitolo 6

Conclusioni

Il nostro percorso ha preso inizio con un sogno. Il sogno del rabbino. Nella spiaggia del fiume Moldava è stata creata una nuova creatura, il Golem. È un essere antropomorfo di fango e argilla, creato con lo scopo di difendere il ghetto ebraico di Praga da ogni minaccia. La creatura è controllata dal suo creatore attraverso lo stesso legame magico che ne ha permesso la creazione. Un giorno, tuttavia, la creatura pare sfuggire al controllo del suo creatore. Nel ghetto si scatena il panico: un gigantesco uomo sta lasciando dietro di sé una scia di distruzione e di violenza. Il rabbino riesce a intervenire e a placare il Golem. Il ripristino dell'ordine e della pace ha un costo: relegare la creatura a un sonno eterno e abbandonarla nella soffitta della sinagoga.

Svegliandosi dal sogno ed entrando nella realtà, Golem e rabbino si rivedono nelle nuove tecnologie rappresentata dalle interfacce neurali, o *Brain-Computer Interfaces*, e nel loro utente.

Nel secondo capitolo è stata presentata una loro classificazione, procedendo anche a evidenziare la peculiare modalità di interazione con l'utilizzatore, basata, essenzialmente, sulla rilevazione attraverso strumenti più o meno invasivi della sua attività neurale. L'eterogeneità degli ambiti di applicazione e l'affermazione sempre crescente delle BCI nella società comporta profonde ricadute sia sul piano giuridico che sul piano etico-filosofico. Il riflesso forse più importante sulle categorie giuridiche è rappresentato dal problema del c.d. *Responsibility Gap*, il grande tema della trattazione.

Nel terzo capitolo, allora, è stata introdotta e descritta la questione servendosi di un caso fittizio nel quale una donna di nome Mary, grazie a un'interfaccia invasiva, riesce a muovere le sue braccia robotiche. Immaginando che Mary abbia accoltellato suo marito, ci si è chiesti quali siano le categorie giuridiche interessate.

La prima che è venuta in rilievo è l'intenzione. L'impatto delle BCI è tale che diviene altamente problematico ricondurre l'azione dell'utente a un'intenzionalità riscontrabile ed accertabile. In particolare, a prescindere dalla riconducibilità della condotta dell'utente all'una o all'altra tipologia di dolo, ciò che è messa alla prova critica delle interfacce è la nozione stessa di dolo, intesa come rappresentazione e volontà dell'evento. Non si capirebbe, infatti, se il soggetto abbia effettivamente “voluto” attivare l'interfaccia e se sia stato in grado di prevedere l'evento, anche alla luce della imprevedibilità insita negli algoritmi ad apprendimento automatico installati nelle interfacce.

Anche la categoria della colpa verrebbe incisa dall'interazione con le BCI. Qualificando la colpa in senso tradizionale, come la violazione di regole di prudenza, diligenza e perizia, si è posta la domanda sulle modalità di violazione. Posta sempre sullo sfondo la questione dell'imprevedibilità del funzionamento dell'algoritmo, si dovrebbe constatare che l'utente di BCI ha violato tali regole con il mero pensiero. Ciò potrebbe portare, allora, alla paradossale soluzione di imporre un obbligo giuridico di *non pensare* determinati pensieri di modo da non porre in essere una condotta colposa per la violazione delle suddette regole cautelari. A ben vedere, tuttavia, la stessa nozione di condotta giuridicamente rilevante è messa in discussione.

L'altra grande categoria profondamente incisa dal *Responsibility Gap*, infatti, è quella di condotta umana. La prima difficoltà sorge sull'effettivo significato da attribuire al concetto di azione, com'è stato evidenziato nel paragrafo 3.6. Se poi si intende la condotta come movimento corporeo volontario controllato o controllabile dal soggetto, le frizioni emergono per la peculiarità dell'interazione tra interfaccia e algoritmo e utente. La tensione emerge in riferimento alla stessa nozione di movimento corporeo. L'interfaccia che ha rilevato i segnali neurali si può ricondurre alla dimensione corporea del soggetto? Una risposta affermativa comporterebbe il sacrificio del tradizionale paradigma di condotta, dal momento che si qualificerebbero come tali movimenti causati da interfacce che vanno oltre la dimensione spaziale e temporale che è posta a fondamento del concetto di condotta così come assunto fino a questo momento. Le ricadute sono, tuttavia, più profonde.

Il *Responsibility Gap* ha anche una rilevante dimensione antropologica. La domanda alla base della questione è “chi ha agito?”. Per dare una identità a quel “chi”, bisogna prima interrogarsi sulla natura dell'uomo stesso. Nel quarto capitolo, dunque, ci si è chiesti in che modo l'uomo può essere trasformato dall'esperienza delle BCI. A fronte di un composito ed eterogeneo panorama di orientamenti e ideologie, dal transumano al postumano, passando per i riduzionisti e compatibilisti, una considerazione è da ricordare,

già accennata in riferimento al tema della condotta umana. A seconda che si riconduca l'interfaccia all'utente di BCI o meno, cambia anche il modo di concepire la sua individualità. Quest'ultimo, infatti, è un *in-dividuo*, il quale non è frammentato e disarticolato in molteplici componenti, pena la perdita di identità e autonomia. Se non si riconduce l'interfaccia all'utente, si creerebbero due centri di imputazione: algoritmo da un lato, utente dall'altro. Al contrario, se si qualifica l'utente come un "tutto" dotato di interfaccia, si mantiene un'unica individualità: algoritmo, interfaccia e utente si identificano in un unico soggetto. Ecco, allora, che appare necessario capire in che modo i mezzi a disposizione per superare il *gap* di responsabilità devono prendere atto di questa nuova individualità, così come descritto nel capitolo cinque, in riferimento ai due beni giuridici della integrità mentale e della libertà di autodeterminazione. Proprio l'analisi di quest'ultima rappresenta l'occasione per porre il problema del *Responsibility Gap* su basi diverse.

Decisivo, infatti, potrebbe essere il modo di considerare il rapporto non già tra interfaccia e utente, ma tra utente e il suo cervello. Ripercorrendo al contrario la catena causale dell'attivazione dell'interfaccia, si deve prendere atto che l'interfaccia ha rilevato i segnali neurali provenienti dal cervello del soggetto. Così come l'attività neurale fa capo all'organo cerebrale, così quest'ultimo deve necessariamente essere ricondotto all'utilizzatore della BCI, a meno di abbracciare posizioni riduzionistiche che, tuttavia, porrebbero in discussione non solo la questione della responsabilità, ma l'utilità dell'intero sistema giuridico che regola la società umana. Ricondurre il cervello dell'utente all'utente stesso permetterebbe di recuperare un'unità individuale. Esiste un solo individuo: l'utente dotato di interfaccia, senza che vi sia scissione tra l'uno e l'altro. Alla fine, l'unico responsabile è questo individuo *2.0*, che altro non è se non un essere già transumano.

Il rischio di procedere a una scissione tra BCI, utente e cervello è troppo grande: non solo si potrebbe cadere in un baratro riduzionista, ma si dovrebbe riqualificare il concetto di condotta umana che si estrinseca nel mondo esterno, giungendo all'estrema possibilità di considerare colpevole il *dividuo* per il mero pensiero che ha azionato l'interfaccia. Se, infatti, l'algoritmo non è parte del soggetto, ma un qualcosa che gli è estraneo, la differenza fondamentale con una qualunque altra tecnologia è proprio la modalità di interazione, essendo l'attivazione dell'interfaccia dovuta a un mero pensare da parte dell'utente. Inoltre, non aderire a visione riduzionistica significa anche iniziare a ragionare in termini di motivi e non di cause. Ponendo il discorso su queste diverse basi, infatti, al centro della questione non si pone più ciò che è *causa* dell'agire, ossia le onde cerebrali, bensì l'individuo nel suo complesso, l'agire del quale è dettato da *motivi*. Così facendo si pone

la questione su un piano differente da quello dell'agire causale, proprio della concezione del riduzionismo: se la domanda non è più "chi ha agito", dal momento che ci può essere un solo soggetto agente, ossia l'individuo, non è più imprescindibile ricostruire la catena causale sotto il profilo dell'attività cerebrale, così come non lo sarebbe, ad esempio, nel caso in cui Mary fosse sprovvista di interfaccia. Così come in questo diverso caso, infatti, non si porrebbe in discussione la riconducibilità del movimento al soggetto, essendo un tutt'uno "naturale" e indiviso, allo stesso modo, dovendo essere un tutt'uno indiviso utente, cervello e interfaccia, non è configurabile il dubbio circa la riconducibilità dell'azione all'uno o all'altro.

L'attribuzione di responsabilità, tuttavia, è possibile solo a patto di riformulare la categoria della capacità di intendere e di volere in modo da ricondurvi anche i segnali neurali involontari che hanno attivato l'interfaccia. È proprio qui che si coglie il nucleo più profondo del *Responsibility Gap*. Azione e mero pensiero da un lato, capacità di intendere e di volere dall'altro: il futuro dirà quale dei due sarà sacrificata o riformulata.

Il Golem si è risvegliato. I suoi passi proseguono. Continuiamo a seguirli.

Bibliografia

- 1: *Relazione sul Libro 1. del Progetto / Ministero della Giustizia e degli Affari di culto*. ita. Roma: Tipografia delle Mantellate, 1929.
- Aflalo, Tyson et al. “Decoding motor imagery from the posterior parietal cortex of a tetraplegic human”. In: *Science* 348.6237 (2015), pp. 906–910.
- Aharoni, Eyal et al. “Neuroprediction of future rearrest”. In: *Proceedings of the National Academy of Sciences* 110.15 (2013), pp. 6223–6228.
- Alderson-Day, Ben e Charles Fernyhough. “Inner speech: development, cognitive functions, phenomenology, and neurobiology.” In: *Psychological bulletin* 141.5 (2015), p. 931.
- Alexander, John J et al. “Restoration of cone vision in a mouse model of achromatopsia”. In: *Nature medicine* 13.6 (2007), pp. 685–687.
- Amato, Salvatore. “La riproduzione assistita come diritto. Riflessioni bio-giuridiche”. In: *Rivista Internazionale di Filosofia del Diritto* 1 (2001), pp. 89–117.
- “Neuroscienze e utilizzazione militare delle tecniche di potenziamento umano”. In: (2014).
- Anguera, Joaquin A et al. “Video game training enhances cognitive control in older adults”. In: *Nature* 501.7465 (2013), pp. 97–101.
- Annas, Catherine L e George J Annas. “Enhancing the fighting force: medical research on American soldiers”. In: *J. Contemp. Health L. & Pol’y* 25 (2008), p. 283.
- Audi, Robert. “Intending”. In: *The Journal of Philosophy* 70.13 (1974), pp. 387–403.
- Austin, John Langshaw. *How to do things with words*. Oxford university press, 1975.
- Baars, Bernard. “How brain reveals mind neural studies support the fundamental role of conscious experience”. In: *Journal of Consciousness Studies* 10.9-10 (2003), pp. 100–114.
- Bainbridge, William Sims. “Religion for a galactic civilization 2.0”. In: *Institute for Ethics and Emerging Technologies* (2009).

- Baird, Amee D et al. “The amygdala and sexual drive: insights from temporal lobe epilepsy surgery”. In: *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 55.1 (2004), pp. 87–96.
- Belliveau, John William. *Functional NMR imaging of the brain*. Harvard University, 1990.
- Berger, Hans. “Über das elektroencephalogramm des menschen”. In: *Archiv für psychiatrie und nervenkrankheiten* 87.1 (1929), pp. 527–570.
- Berger, Hans e Gottfried Wälchli. *Hans Berger*. Musée d’art et d’histoire, 1956.
- Berlin, Isaiah. “Two concepts of liberty”. In: *The liberty reader*. Routledge, 2017, pp. 33–57.
- Bertolini, Andrea et al. “On robots and insurance”. In: *International Journal of Social Robotics* 8.3 (2016), pp. 381–391.
- Bettiol, G. *Diritto penale: parte generale*. CEDAM, 1982.
- Bickle, J, P Mandik e A Landreth. “The Philosophy of Neuroscience, in «The Stanford Encyclopedia of Philosophy»”. In: *Zalta EN (a cura di)* (2019). URL: <https://plato.stanford.edu/archives/fall2019/entries/neuroscience>.
- Bickle, John. *Philosophy and neuroscience: A ruthlessly reductive account*. Vol. 2. Springer Science & Business Media, 2003.
- Birbaumer, Niels e Leonardo G Cohen. “Brain–computer interfaces: communication and restoration of movement in paralysis”. In: *The Journal of physiology* 579.3 (2007), pp. 621–636.
- Birks, David e Alena Buyx. “Punishing intentions and neurointerventions”. In: *AJOB neuroscience* 9.3 (2018), pp. 133–143.
- Bishop, Christopher M e Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006.
- Blackmore, Susan J. “Why psi tells us nothing about consciousness”. In: (1998).
- Blondet, Maria V Ruiz et al. “A wearable real-time BCI system based on mobile cloud computing”. In: *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE. 2013, pp. 739–742.
- Bocquelet, Florent et al. “Key considerations in designing a speech brain-computer interface”. In: *Journal of Physiology-Paris* 110.4 (2016), pp. 392–401.
- Bosoer, Lucia. “Chile at the Forefront of Neurorights Protection”. Accessed: 29/08/2022. 2021. URL: <https://blogs.eui.eu/latin-american-working-group/opinion-chile-at-the-forefront-of-neurorights-protection/>.

- Brain, W. “Brain Waves Module 1: Neuroscience, society and policy”. In: (2011).
- Bratman, Michael. “Intention and means-end reasoning”. In: *The Philosophical Review* 90.2 (1981), pp. 252–265.
- “Two faces of intention”. In: *The Philosophical Review* 93.3 (1984), pp. 375–405.
- Broca, Paul. “Sur le siège de la faculté du langage articulé”. In: *Bulletins et Mémoires de la Société d’Anthropologie de Paris* 6.1 (1865), pp. 377–393.
- Bryant, Brian R e Penny Crews Seay. “Republication of The Technology-Related Assistance to Individuals with Disabilities Act: Relevance to Individuals with Learning Disabilities and Their Advocates”. In: *Journal of Learning Disabilities* 53.2 (2020), pp. 80–91.
- Bublitz, Christoph et al. “Legal liabilities of BCI-users: Responsibility gaps at the intersection of mind and machine?” In: *International journal of law and psychiatry* 65 (2018), pp. 101399–101399.
- Bublitz, Jan Christoph e Reinhard Merkel. “Autonomy and authenticity of enhanced personality traits”. In: *Bioethics* 23.6 (2009), pp. 360–374.
- Bublitz, Jan-Christoph. “My mind is mine!? Cognitive liberty as a legal concept”. In: *Cognitive enhancement*. Springer, 2013, pp. 233–264.
- Burke, John F et al. “Brain computer interface to enhance episodic memory in human participants”. In: *Frontiers in Human Neuroscience* 8 (2015), p. 1055.
- Burrell, Jenna. “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”. In: *Big Data & Society* 3.1 (2016), p. 20.
- Buyx, Alena e David Birks. “Neuroscience and social problems: The case of neuropunishment”. In: *Cambridge Quarterly of Healthcare Ethics* 27.4 (2018), pp. 628–634.
- Cahn, B Rael e John Polich. “Meditation states and traits: EEG, ERP, and neuroimaging studies.” In: *Psychological bulletin* 132.2 (2006), p. 180.
- Campa, Riccardo. “Manifesto dei transumanisti italiani”. In: gen. 2009, pp. 75–100. ISBN: 978-88-96333-01-3.
- Cerebral connections: UCLA engineers tap into rich legacy of brain-computer interface technology, howpublished = <http://samueli.ucla.edu/brain-computer-interface/>, note = Accessed: 2021-11-08.*
- Chakrabarti, Shreya et al. “Progress in speech decoding from the electrocorticogram”. In: *Biomedical Engineering Letters* 5.1 (2015), pp. 10–21.
- Cheshire, William P. “The origami brain: from neural folds to neuroethics”. In: *Ethics & Medicine* 27.2 (2011), p. 79.

- Chiereghin, Franco. *L'eco della caverna: ricerche di filosofia della logica e della mente*. Il poligrafo, 2004.
- “La coscienza: un ritardato mentale?” In: *Verifiche* 37.4 (2008), pp. 283–316.
- Churchland, Patricia S. *Neurobiologia della morale*. Raffaello Cortina Editore, 2012.
- Churchland, Patricia Smith. *Neurophilosophy*. MIT press, 1986.
- Churchland, Paul M. “The logical character of action-explanations”. In: *The Philosophical Review* 79.2 (1970), pp. 214–236.
- Clark, Andy e David Chalmers. “The extended mind”. In: *analysis* 58.1 (1998), pp. 7–19.
- Clark, Andy e Natural-Born Cyborgs. “Minds”. In: *Technologies, and the Future of Human Intelligence* (2003).
- Clausen, Jens, Neil Levy et al. *Handbook of neuroethics*. Springer, 2015.
- Clynes, Manfred E e Nathan S Kline. “Cyborgs and space”. In: *Astronautics* 14.9 (1960), pp. 26–27.
- Cocco G. e Ambrosetti, E.M. *Trattato breve di diritto penale. Parte generale*. Trattato breve di diritto penale v. 2. CEDAM, 2018. ISBN: 9788813368586.
- Cocuccio, Maria Francesca. “Il diritto all’identità personale e l’identità “digitale””. In: *Il diritto di Famiglia e delle persone* 3 (2016), pp. 949–968.
- Cole-Turner, Ronald. *Transhumanism and transcendence: Christian hope in an age of technological enhancement*. Georgetown University Press, 2011.
- Connolly, William E. *Neuropolitics: Thinking, culture, speed*. Vol. 23. U of Minnesota Press, 2002.
- Damasio, Hanna et al. “The return of Phineas Gage: clues about the brain from the skull of a famous patient”. In: *Science* 264.5162 (1994), pp. 1102–1105.
- Danto, Arthur C. *Analytical philosophy of action*. Cambridge University Press, 1973.
- Davidson, Donald. *Essays on Actions and Events: Philosophical Essays Volume 1*. Clarendon Press, 2001.
- De Marsico, A. *Coscienza e volontà nella nozione del dolo*. Alberto Morano, 1930.
- De Monticelli, R. *La novità di ognuno: Persona e libertà*. Garzanti Saggi. Garzanti, 2012. ISBN: 9788811134152.
- Diaz, Rafael M, Laura E Berk e Rafael Diaz. *Private speech: From social interaction to self-regulation*. L. Erlbaum Hillsdale, NJ, 1992.

- Diefenbach, Gretchen J et al. "Portrayal of lobotomy in the popular press: 1935-1960". In: *Journal of the History of the Neurosciences* 8.1 (1999), pp. 60–69.
- Donoghue, John P et al. "Development of neuromotor prostheses for humans". In: *Supplements to Clinical neurophysiology* 57 (2004), pp. 592–606.
- Douglas, Angela E. *The symbiotic habit*. Princeton University Press, 2021.
- Dunagan, Jake F. "Neuro-futures: the brain, politics and power". In: *Journal of Futures Studies* 9.2 (2004), pp. 1–18.
- "Politics for the neurocentric age". In: *Journal of Futures Studies* 15.2 (2010), pp. 51–70.
- Edward, Shorter. *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*. 1997.
- Farahany, Nita A. "A neurological foundation for freedom". In: *Stan. Tech. L. Rev.* (2012), p. 4.
- Fiandaca G. e Musco, E. *Diritto penale. Parte generale*. Zanichelli, 2019. ISBN: 9788808620316.
- Finn, Emily S et al. "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity". In: *Nature neuroscience* 18.11 (2015), pp. 1664–1671.
- Fischer, John Martin. "Responsibility and control". In: *The Journal of Philosophy* 79.1 (1982), pp. 24–40.
- Floridi, Luciano. *La quarta rivoluzione: come l'infosfera sta trasformando il mondo*. Raffaello Cortina Editore, 2017.
- Focquaert, Farah e Dirk De Ridder. "Direct intervention in the brain: ethical issues concerning personal identity". In: *Journal of Ethics in Mental Health* 4.2 (2009).
- Folgieri, Raffaella. "Brain computer interface and transcranial magnetic stimulation in legal practice and regulations". In: *Neuroscience and Law*. Springer, 2020, pp. 273–290.
- Frankfurt, Harry G. "The problem of action". In: *American philosophical quarterly* 15.2 (1978), pp. 157–162.
- Freeman, Walter e James W Watts. "Prefrontal lobotomy in agitated depression. Report of a case". In: *Med. Ann. Dist. Columbia* 5 (1936), pp. 326–329.
- Fuselli, Stefano. "Brain-Computer Interface e soggettività agente. Considerazioni etico-giuridiche." In: *Etica, diritto e tecnologia. Percorsi dell'informatica giuridica contemporanea*. 1 (2021), pp. 151–174.

- Fuselli, Stefano. “Di bene in meglio? Riflessioni sul potenziamento biotecnologico dell’uomo”. In: *Diritto e desiderio Riflessioni biogiuridiche* (2015), pp. 153–84.
- *Diritto, neuroscienze, filosofia. Un itinerario*. FrancoAngeli, 2014.
- “Neurotecnologie e tutela dell’integrità psichica. Profili filosofico giuridici di un mutamento in atto”. In: *Journal of Ethics and Legal Technologies* 2.1 (2020).
- Gaiani, Alberto. “Riduzionismo e neuroscienze: il dibattito filosofico recente”. In: *Etica & Politica / Ethics & Politics, XVI, 2* (2014), pp. 47–63.
- Garis, Hugo de. “The Artilect War: Cosmists vs. Terrans. A Bitter Controversy Concerning Whether Humanity Should Build Godlike Massively Intelligent Machines”. In: *Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference*. 2008, pp. 437–447.
- Gazzaniga, M.S. e S. Inglese. *Chi comanda? Scienza, mente e libero arbitrio*. Le Scienze. Codice, 2017. ISBN: 9788875786748.
- Georgopoulos, Apostolos P et al. “Mental rotation of the neuronal population vector”. In: *Science* 243.4888 (1989), pp. 234–236.
- Geraci, R.M. *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. OUP USA, 2012. ISBN: 9780199964000.
- Gibson, W. et al. *Neuromancer*. Ace science fiction. Ace Books, 1984. ISBN: 9780441569595.
- Glannon, Walter et al. *Bioethics and the brain*. Oxford University Press, USA, 2007.
- Glannon, Walter. “Ethical issues in neuroprosthetics”. In: *Journal of Neural Engineering* 13.2 (2016), p. 021002.
- Glaser, Joshua I et al. “Machine learning for neural decoding”. In: *Eneuro* 7.4 (2020).
- Goddard, Eliza. “Deep brain stimulation through the “lens of agency”: Clarifying threats to personal identity from neurological intervention”. In: *Neuroethics* 10.3 (2017), pp. 325–335.
- Goering, Sara et al. “Staying in the loop: Relational agency and identity in next-generation DBS for psychiatry”. In: *AJOB Neuroscience* 8.2 (2017), pp. 59–70.
- Goethals, Ingeborg et al. “Brain activation associated with deep brain stimulation causing dissociation in a patient with Tourette’s syndrome”. In: *Journal of Trauma & Dissociation* 9.4 (2008), pp. 543–549.
- Goldman, Alvin I. “Theory of human action”. In: *Theory of Human Action*. Princeton University Press, 2015.

- Gometz, Gianmarco. “La privacy della mente: alcune riflessioni sul rapporto tra protezione dei dati personali e libertà di pensiero”. In: *Stato, Chiese e pluralismo confessionale* (2020).
- Gurney, David. “Killer robot arms: A case-study in brain–computer interfaces and intentional acts”. In: *Minds and Machines* 28.4 (2018), pp. 775–785.
- Hall, Shana A et al. “The neural basis of involuntary episodic memories”. In: *Journal of cognitive neuroscience* 26.10 (2014), pp. 2385–2399.
- Harbisson, Neil. “Painting by ear”. In: *Modern Painters, The International Contemporary Art Magazine* (2008), pp. 70–73.
- Harlow, John M. “Passage of an iron rod through the head.” In: *The Boston Medical and Surgical Journal (1828-1851)* 39.20 (1848), 0.1.
- “Recovery after severe injury to the head”. In: *Publication of the Massachusetts Medical Society* 2.327 (1868), pp. 990–992.
- Haselager, Pim. “Did I do that? Brain–computer interfacing and the sense of agency”. In: *Minds and Machines* 23.3 (2013), pp. 405–418.
- Hassanien, Aboul Ella e AA Azar. “Brain-computer interfaces”. In: *Switzerland: Springer* 74 (2015).
- Hayles, N. “Katherine: How we became posthuman”. In: *Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago (1999).
- Hess, Ewa e Hennric Jokeit. “Neurocapitalism”. In: *trans. by M. Newton, Eurozine*. Available at: <https://www.opendemocracy.net/ewa-hess-hennric-jokeit/neurocapitalism> (accessed 03.07.2022) (2009).
- Himmelreich, Johannes. “Agency and embodiment: Groups, human–machine interactions, and virtual realities”. In: *Ratio* 31.2 (2018), pp. 197–213.
- Hochberg, Leigh R et al. “Neuronal ensemble control of prosthetic devices by a human with tetraplegia”. In: *Nature* 442.7099 (2006), pp. 164–171.
- Hofstadter, D.R. et al. *L’io della mente. Fantasia e riflessioni sul sé e sull’anima*. Biblioteca scientifica. Adelphi, 1985. ISBN: 9788845906329.
- Höhne, Johannes et al. “A novel 9-class auditory ERP paradigm driving a predictive text entry system”. In: *Frontiers in neuroscience* 5 (2011), p. 99.
- Huxley, Julian. *Religion without revelation*. The New American Library (1957), 2016.
- “Transhumanism”. In: *Ethics in Progress* 6.1 (2015), pp. 12–16.
- Inenca, Marcello. “Intelligenza²: per un’unione di intelligenza naturale e artificiale”. In: *Intelligenza²* (2019), pp. 1–175.
- Inenca, Marcello e Roberto Andorno. “Towards new human rights in the age of neuroscience and neurotechnology”. In: *Life sciences, society and policy* 13.1 (2017), pp. 1–27.

- Ienca, Marcello e Pim Haselager. “Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity”. In: *Ethics and Information Technology* 18.2 (2016), pp. 117–129.
- Ienca, Marcello, Pim Haselager e Ezekiel J Emanuel. “Brain leaks and consumer neurotechnology”. In: *Nature biotechnology* 36.9 (2018), pp. 805–810.
- Il Golem: un uomo di fango nel ghetto di Praga*, <https://www.storicang.it/a/il-golem-uomo-di-fango-nel-ghetto-di-praga-15184>, Accessed: 2021-11-04.
- Inglese, Silvia e Andrea Lavazza. “What Should We Do With People Who Cannot or Do Not Want to Be Protected From Neurotechnological Threats?” In: *Frontiers in Human Neuroscience* (2021), p. 442.
- Jansson, Bengt. “Controversial psychosurgery resulted in a Nobel Prize”. In: *Nobelprize.org* (2007).
- Jeffries, Stuart. “Neil Harbisson: the world’s first cyborg artist”. In: *The Guardian* 6 (2014).
- Jescheck, H.H. e T. Weigend. *Lehrbuch des Strafrechts: allgemeiner Teil*. Lehrbuch des Strafrechts, Allgemeiner Teil: Hans-Heinrich Jescheck; Thomas Weigend. Hauptbd. Duncker & Humblot, 1996. ISBN: 9783428083480.
- Jobsis, Frans F. “Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters”. In: *Science* 198.4323 (1977), pp. 1264–1267.
- Juengst, Eric e Daniel Moseley. “Human Enhancement”. In: *The Stanford Encyclopedia of Philosophy*. A cura di Edward N. Zalta. Summer 2019. Metaphysics Research Lab, Stanford University, 2019.
- Kellmeyer, Philipp. “Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices”. In: *Neuroethics* 14.1 (2021), pp. 83–98.
- Kessler, Ronald. *The sins of the father: Joseph P. Kennedy and the dynasty he founded*. St. Martin’s Press, 1996.
- Kirmayer, Laurence J e Ian Gold. “Re-socializing psychiatry”. In: *Critical neuroscience: A handbook of the social and cultural contexts of neuroscience* (2012).
- Klaming, Laura e Pim Haselager. “Did my brain implant make me do it? Questions raised by DBS regarding psychological continuity, responsibility for action and mental competence”. In: *Neuroethics* 6.3 (2013), pp. 527–539.
- Kleiner-Fisman, Galit et al. “Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes”. In: *Movement disorders: of-*

- Official journal of the Movement Disorder Society* 21.S14 (2006), S290–S304.
- Kolber, Adam J. “Will there be a neurolaw revolution”. In: *Ind. LJ* 89 (2014), p. 807.
- Kurzweil, Ray. *The age of spiritual machines: When computers exceed human intelligence*. Penguin, 2000.
- *The singularity is near: When humans transcend biology*. Penguin, 2005.
- Larus, James et al. *When computers decide: European recommendations on machine-learned automated decision making*. 2018.
- Lauterbur, Paul C. “Image formation by induced local interactions: examples employing nuclear magnetic resonance”. In: *nature* 242.5394 (1973), pp. 190–191.
- Lavazza, Andrea. “Freedom of thought and mental integrity: The moral requirements for any neural prosthesis”. In: *Frontiers in Neuroscience* 12 (2018), p. 82.
- Lavin, Douglas. “Must there be basic action?” In: *Nou̇s* 47.2 (2013), pp. 273–301.
- Leach, Martin O. “Nobel prize in physiology or medicine 2003 awarded to Paul Lauterbur and Peter Mansfield for discoveries concerning magnetic resonance imaging”. In: *Physics in Medicine & Biology* 49.3 (2004).
- Libet, Benjamin. *Mind time: The temporal factor in consciousness*. Harvard University Press, 2009.
- Libet, Benjamin et al. “Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential)”. In: *Neurophysiology of consciousness*. Springer, 1993, pp. 249–268.
- Liszt, F. von. *Lehrbuch des deutschen Strafrechts*. 1888.
- Locked-In syndrome*, <https://rarediseases.info.nih.gov/diseases/6919/locked-in-syndrome/>, Accessed: 2021-11-10.
- López, Tatiana Afanador e Judit Parés Padrós. “Designing organs at the Transpecies Society: hybrid practices between cybernetics and artificial intelligence”. In: *Temes de Disseny* 35 (2019), pp. 140–153.
- Macioce, Fabio. “Informed consent procedures between autonomy and trust”. In: *BioLaw Journal-Rivista di BioDiritto* 1S (2019), pp. 23–35.
- MacKellar, Calum. *Cyborg Mind: What Brain Computer and Mind Cyberspace Interfaces Mean for Cyberneuroethics*. Berghahn Books, 2017.
- Mackenzie, Catriona, Mary Walker et al. “Neurotechnologies, personal identity, and the ethics of authenticity”. In: *Handbook of neuroethics* 10 (2015), pp. 373–392.
- Macmillan, Malcolm. *An odd kind of fame*. 2000.

- Magazine, Wired. 8.02 (February 2000), 'Cyborg 1.0: Interview with Kevin Warwick'. Rapp. tecn. Retrieved 25-12-2006.
- Mak, Joseph N e Jonathan R Wolpaw. "Clinical applications of brain-computer interfaces: current state and future prospects". In: *IEEE reviews in biomedical engineering* 2 (2009), pp. 187–199.
- Mantovani, F. *Diritto penale. Parte generale*. CEDAM, 2015. ISBN: 978-88-13354-4-66.
- Marchesini, R. *Post-human: verso nuovi modelli di esistenza*. Saggi Bollati Boringhieri (Firm): Scienze. Bollati Boringhieri, 2002. ISBN: 978-88-33913-7-97.
- Marinucci G. e Dolcini, E. *Manuale di diritto penale. Parte generale*. Giuffrè, 2015. ISBN: 9788814200465.
- Marmor, Andrei. "What Is the Right to Privacy?" In: *Philosophy & Public Affairs* 43.1 (2015), pp. 3–26. ISSN: 00483915, 10884963.
- Martel, Adrien, Sven Dähne e Benjamin Blankertz. "EEG predictors of covert vigilant attention". In: *Journal of neural engineering* 11.3 (2014), p. 035009.
- Mashour, George A, Erin E Walker e Robert L Martuza. "Psychosurgery: past, present, and future". In: *Brain research reviews* 48.3 (2005), pp. 409–419.
- Matthias, Andreas. "The responsibility gap: Ascribing responsibility for the actions of learning automata". In: *Ethics and information technology* 6.3 (2004), pp. 175–183.
- Mehlman, Maxwell J et al. "Ethical and legal issues in enhancement research on human subjects". In: *Cambridge Quarterly of Healthcare Ethics* 20.1 (2011), pp. 30–45.
- Meixner, Uwe. "New perspectives for a dualistic conception of mental causation". In: *Journal of Consciousness Studies* 15.1 (2008), pp. 17–38.
- Mele, Alfred. "Proximal intentions, intention-reports, and vetoing". In: *Philosophical Psychology* 21.1 (2008), pp. 1–14.
- Mele, Alfred R. "Acting for reasons and acting intentionally". In: *Pacific Philosophical Quarterly* 73 (1992).
- Metzinger, Thomas. "The problem of mental action". In: (2017).
- Moore, Brian E. *The brain computer interface future: time for a strategy*. Rapp. tecn. Air War College Air University Maxwell AFB United States, 2013.
- Moore, Michael S. *Act and crime: The philosophy of action and its implications for criminal law*. Oxford University Press on Demand, 2010.
- Moravec, Hans. *Mind children: The future of robot and human intelligence*. Harvard University Press, 1988.

- Moravec, Hans P. *Robot: mere machine to transcendent mind*. Oxford University Press on Demand, 2000.
- Moro, P. *Etica, diritto e tecnologia. Percorsi dell'informatica giuridica contemporanea*. Diritto moderno e interpretazione classica. Franco Angeli, 2021. ISBN: 9788835119104.
- Morse, Stephen. "Avoiding irrational neurolaw exuberance: a plea for neuromodesty". In: *Law, Innovation and Technology* 3.2 (2011), pp. 209–228.
- "Lost in translation?: An essay on law and neuroscience". In: *An Essay on Law and Neuroscience (August 3, 2011)*. *Law and Neuroscience, Current legal issues* 13 (2011), p. 529.
- Morse, Stephen J. "New neuroscience, old problems: legal implications of brain science." In: *Cerebrum: the Dana forum on brain science*. Vol. 6. 4. 2004, pp. 81–90.
- "The status of neurolaw: a plea for current modesty and future cautious optimism". In: *The Journal of Psychiatry & Law* 39.4 (2011), pp. 595–626.
- Mühl, Christian et al. "A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges". In: *Brain-Computer Interfaces* 1.2 (2014), pp. 66–84.
- Nabavi, Sadegh et al. "Engineering a memory with LTD and LTP". In: *Nature* 511.7509 (2014), pp. 348–352.
- Nicolelis, Miguel. *Beyond boundaries: The new neuroscience of connecting brains with machines—and how it will change our lives*. Macmillan, 2011.
- Nisco, A. *La tutela penale dell'integrità psichica*. Itinerari di diritto penale. Giappichelli, 2012. ISBN: 9788834827666.
- Ogawa, Seiji et al. "Brain magnetic resonance imaging with contrast dependent on blood oxygenation". In: *proceedings of the National Academy of Sciences* 87.24 (1990), pp. 9868–9872.
- Pacherie, Elisabeth. "The sense of control and the sense of agency". In: *Psyche* 13.1 (2007), pp. 1–30.
- "Towards a dynamic theory of intentions". In: *Does consciousness cause behavior* (2006), pp. 145–167.
- Padovani, T. *Diritto penale*. Giuffrè, 2008. ISBN: 9788814141539.
- Pagliari, A. *Il reato*. Trattato di diritto penale (Giuffrè editore).: Parte generale. Giuffrè, 2007. ISBN: 9788814127816.
- *Principi di diritto penale: parte generale*. v. 1. Giuffrè, 2003. ISBN: 978-88-14105-4-63.
- Pais-Vieira, Miguel et al. "A brain-to-brain interface for real-time sharing of sensorimotor information". In: *Scientific reports* 3.1 (2013), pp. 1–10.

- Palazzani, Laura. “La condizione tecno-umana e le tecnologie convergenti: percorsi scientifici e filosofici “oltre” l’umano”. In: *Studium Ricerca* 3 (2019), pp. 59–77.
- Parisi, Domenico. *La naturalizzazione della cultura*. 1998.
- Pepperell, Robert. *The post-human condition*. Intellect books, 1995.
- Persson, Ingmar e Julian Savulescu. “Unfit for the future? Human nature, scientific progress, and the need for moral enhancement”. In: *Enhancing human capacities* (2011), pp. 486–500.
- Pinti, Paola et al. “The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience”. In: *Annals of the New York Academy of Sciences* 1464.1 (2020), p. 5.
- Proust, Joëlle. “A plea for mental acts”. In: *Synthese* 129.1 (2001), pp. 105–128.
- Rainey, Stephen, Hannah Maslen, Pierre Mégevand et al. “Neuroprosthetic speech: The ethical significance of accuracy, control and pragmatics”. In: *Cambridge Quarterly of Healthcare Ethics* 28.4 (2019), pp. 657–670.
- Rainey, Stephen, Hannah Maslen e Julian Savulescu. “When thinking is doing: Responsibility for BCI-mediated action”. In: *Ajob Neuroscience* 11.1 (2020), pp. 46–58.
- Reichlin, M. *Etica e neuroscienze: stati vegetativi, malattie degenerative, identità personale*. Forma mentis. Mondadori università, 2012. ISBN: 978-88-61842-2-74.
- Ritchie, Craig et al. “CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer’s disease dementia and other dementias in people with mild cognitive impairment (MCI)”. In: *Cochrane Database of Systematic Reviews* 3 (2017).
- Rocco, A. e Italia. *Lavori preparatori del codice penale e del codice di procedura penale. vol. 5: progetto definitivo di un nuovo codice penale con la relazione del guardasigilli Alfredo Rocco*. Tipografia delle mantellate, 1929.
- Romano, M. e G. Grasso. *Commentario sistematico del Codice penale: Art. 1-84*. v. 1. A. Giuffrè, 1987.
- Ronchi, Alfredo M. *eCulture: cultural content in the digital age*. Springer Science & Business Media, 2009.
- Ronco, M. *Scritti patavini: Due tomi indivisibili*. Anima Juris Ratio. Giapichelli, 2017. ISBN: 9788892164116.
- Roskies, Adina. “Neuroethics for the new millenium”. In: *Neuron* 35.1 (2002), pp. 21–23.
- Salfellner, Harald. *Il golem di Praga: Leggende ebraiche dal ghetto*. Vitalis, 2019.

- Sandel, M.J. e S. Galli. *Contro la perfezione. L'etica nell'età dell'ingegneria genetica*. Vita e Pensiero, 2022. ISBN: 9788834351857.
- Sandrone, Stefano, Marco Bacigaluppi, Marco R Galloni, Stefano F Cappa et al. “Weighing brain activity with the balance: Angelo Mosso’s original manuscripts come to light”. In: *Brain* 137.2 (2014), pp. 621–633.
- Sandrone, Stefano, Marco Bacigaluppi, Marco R Galloni e Gianvito Martino. “Angelo Mosso (1846–1910)”. In: *Journal of neurology* 259.11 (2012), pp. 2513–2514.
- Santosuosso, A e B Bottalico. “Neuroscienze e categorie giuridiche: quale impatto”. In: *Neuroetica* (2009), pp. 45–68.
- Sarra, Claudio. “Questioni pregiudiziali: una prospettiva epistemologica sui rapporti tra neuroscienze e diritto”. In: (2014).
- “Relevant Legal Issues for Hybrid Human-Robotic Assistive Technologies: A First Assessment”. In: ().
- Savulescu, Julian. “Moral Status of Enhanced Beings: What Do We Owe the Gods?” In: *Human enhancement* (2009), pp. 211–247.
- Schauer, Frederick. “On the Distinction Between Speech and Action”. In: *Virginia Public Law and Legal Theory Research Paper 2014-68* (2014).
- Schechtman, Marya. “Getting our stories straight”. In: *Personal identity and fractured selves* (2009), pp. 65–92.
- Schermer, Maartje. “Changes in the self: The need for conceptual research next to empirical research”. In: *The American Journal of Bioethics* 9.5 (2009), pp. 45–47.
- “Ethical issues in deep brain stimulation”. In: *Frontiers in integrative neuroscience* 5 (2011), p. 17.
- Schlosser, Markus E. “Agency, ownership, and the standard theory”. In: *New waves in philosophy of action*. Springer, 2011, pp. 13–31.
- “Bending it like Beckham: movement, control and deviant causal chains”. In: *Analysis* 70.2 (2010), pp. 299–303.
- Schmidt, EM, MJ Bak e JS McIntosh. “Long-term chronic recording from cortical neurons”. In: *Experimental neurology* 52.3 (1976), pp. 496–506.
- Schockendoff, Eberhard. “Der freie Wille – ein problemgeschichtlicher Abriss”. In: *Der freie Wille und die Schuldfähigkeit in Recht , Psychiatrie und Neurowissenschaften* (2010), pp. 3–15.
- Schönau, Andreas. “The Spectrum of Responsibility Ascription for End Users of Neurotechnologies”. In: *Neuroethics* 14.3 (2021), pp. 423–435.
- Schultze-Kraft, Matthias et al. “The point of no return in vetoing self-initiated movements”. In: *Proceedings of the National Academy of Sciences* 113.4 (2016), pp. 1080–1085.

- Sententia, Wrye. “Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition”. In: *Annals of the New York Academy of Sciences* 1013.1 (2004), pp. 221–228.
- Shen, Francis X. “Neuroscience, mental privacy, and the law”. In: *Harv. JL & Pub. Pol’y* 36 (2013), p. 653.
- Shepherd, Joshua. “Conscious control over action”. In: *Mind & language* 30.3 (2015), pp. 320–344.
- Shulman, Robert G. *Brain imaging: What it can (and cannot) tell us about consciousness*. Oxford University Press, 2013.
- Sommaggio, P. *Filosofia del biodiritto: una proposta socratica per società postumane*. G. Giappichelli, 2016. ISBN: 9788892162600.
- Sommaggio, Paolo. “Neurociviltà”. In: *Etica & Politica / Ethics & Politics, XVI, 2* (2014), pp. 130–168.
- Sommaggio, Paolo e Marco Mazzocca. “Cognitive liberty and human rights”. In: *Neuroscience and Law*. Springer, 2020, pp. 95–111.
- Steinbock, B. *The Oxford Handbook of Bioethics*. Oxford Handbooks. OUP Oxford, 2007. ISBN: 9780191569258.
- Steinert, Steffen et al. “Doing things with thoughts: Brain-computer interfaces and disembodied agency”. In: *Philosophy & Technology* 32.3 (2019), pp. 457–482.
- Stone, James L e John R Hughes. “Early history of electroencephalography and establishment of the American Clinical Neurophysiology Society”. In: *Journal of Clinical Neurophysiology* 30.1 (2013), pp. 28–44.
- Strawson, Galen. “XI—Mental ballistics or the involuntariness of spontaneity”. In: *Proceedings of the Aristotelian Society (Hardback)*. Vol. 103. 1. Wiley Online Library. 2003, pp. 227–256.
- Synofzik, Matthis e Thomas E Schlaepfer. “Stimulating personality: ethical criteria for deep brain stimulation in psychiatric patients and for enhancement purposes”. In: *Biotechnology Journal: Healthcare Nutrition Technology* 3.12 (2008), pp. 1511–1520.
- Synofzik, Matthis, Thomas E Schlaepfer e Joseph J Fins. “How happy is too happy? Euphoria, neuroethics, and deep brain stimulation of the nucleus accumbens”. In: *AJOB Neuroscience* 3.1 (2012), pp. 30–36.
- Thompson, Margaret C. “Critiquing the concept of BCI illiteracy”. In: *Science and engineering ethics* 25.4 (2019), pp. 1217–1233.
- Tirosh-Samuels, Hava. “Transhumanism as a secularist faith”. In: *Zygon* 47.4 (2012), pp. 710–734.
- Truccolo, Wilson, Leigh R Hochberg e John P Donoghue. “Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes”. In: *Nature neuroscience* 13.1 (2010), pp. 105–111.

- Tulving, Endel. “Elements of episodic memory”. In: (1983).
- Valenstein, Elliot S. *Great and Desperate Cures: the rise and decline of psychosurgery and other radical treatments for mental illness*. Basic Books, 1986.
- Van Banning, Theo et al. “Human Rights Reference Handbook”. In: *Costa Rica: University for Peace* (2004).
- Vassalli, Giuliano. *Scritti giuridici - Volume III, Il processo e le libertà*. Casa ed. Giuffrè, 1997. ISBN: 9788814064036.
- Vidal, Jacques J. “Toward direct brain-computer communication”. In: *Annual review of Biophysics and Bioengineering* 2.1 (1973), pp. 157–180.
- “Real-time detection of brain events in EEG”. In: *Proceedings of the IEEE* 65 (1977), pp. 633–641.
- Wang, Yijun e Tzyy-Ping Jung. “A collaborative brain-computer interface for improving human performance”. In: *PloS one* 6.5 (2011), e20422.
- Warren, Samuel e Louis Brandeis. “The right to privacy”. In: *Killing the Messenger*. Columbia University Press, 1989, pp. 1–21.
- Warwick, Kevin, Mark Gasson, Benjamin Hutt e Iain Goodhew. “An attempt to extend human sensory capabilities by means of implant technology”. In: *2005 IEEE International Conference on Systems, Man and Cybernetics*. Vol. 2. IEEE. 2005, pp. 1663–1668.
- Warwick, Kevin, Mark Gasson, Benjamin Hutt, Iain Goodhew et al. “The application of implant technology for cybernetic systems”. In: *Archives of neurology* 60.10 (2003), pp. 1369–1373.
- Waters, B. *From Human to Posthuman: Christian Theology and Technology in a Postmodern World*. Ashgate science and religion series. Ashgate, 2006. ISBN: 9780754639152.
- Wegner, Daniel M. “Ironic processes of mental control.” In: *Psychological review* 101.1 (1994), p. 34.
- *The illusion of conscious will*. MIT press, 2017.
- Wenzlaff, Richard M e Daniel M Wegner. “Thought suppression”. In: *Annual review of psychology* 51.1 (2000), pp. 59–91.
- Wernicke, Carl. *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*. Cohn., 1874.
- Westin, Alan F. “Privacy and freedom”. In: *Washington and Lee Law Review* 25.1 (1968), p. 166.
- Wolpaw, Jonathan R, Niels Birbaumer et al. “Brain-computer interfaces for communication and control”. In: *Clinical neurophysiology* 113.6 (2002), pp. 767–791.

- Wolpaw, Jonathan R e E Winter Wolpaw. “Brain-computer interfaces: something new under the sun”. In: *Brain-computer interfaces: principles and practice* 14 (2012).
- Wolpe, Paul Root. “Is my mind mine? Neuroethics and brain imaging”. In: (2009).
- Zander, Thorsten O et al. “Enhancing human-computer interaction with input from active and passive brain-computer interfaces”. In: *Brain-computer interfaces*. Springer, 2010, pp. 181–199.
- Zeki, Semir et al. “For the law, neuroscience changes nothing and everything”. In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359.1451 (2004), pp. 1775–1785.
- Zornoza, Alejandro et al. “Robots Liability: A Use Case and a Potential Solution”. In: *Robotics-Legal, Ethical and Socioeconomic Impacts*. IntechOpen, 2017.

Ringraziare tutti quelli che mi hanno aiutato in questo lavoro significa ringraziare coloro che ci sono stati e sono rimasti in questi cinque lunghi anni.

Il primo e più sentito ringraziamento, anche se mai sarà sufficiente, va a Stefano Fuselli, il mio relatore, il quale mi ha accompagnato nella prima e nell'ultima prova di questo percorso universitario, mostrandomi in che modo la passione può essere insegnata e tradotta in parole.

A mamma, papà, Isabella: grazie, senza di loro non sarei qui. Letteralmente.

Grazie a Tommaso, fratello di lunghe camminate e profonde parole. Il dialogo con lui ha reso più leggero e più profondo il confronto con ogni fatica di questa università.

Grazie a Claudia, amica ritrovata, sostegno in molte prove e fatiche: il suo sostegno continuo è stato unico e fondamentale, soprattutto negli ultimi passi verso la fine.

Infine, grazie ad Anna, compagna e guida in ogni passo in questi cinque anni. Ho percorso, insieme a lei, un milione di scale e corridoi, aprendo altrettante porte. Questa, forse, è quella più importante.