

STEVEN UMBRELLO

Università degli Studi di Torino, Dipartimento di Filosofia e Scienze dell'Educazione
steven.umbrello@unito.it

MIGLIORAMENTO E POTENZIAMENTO DEGLI OPERATORI SANITARI ATTRAVERSO LA PROGETTAZIONE

ABSTRACT

Gran parte della letteratura riguardante l'uso dell'intelligenza artificiale (IA) sul posto di lavoro, in particolare nell'ambito dell'assistenza infermieristica e dei servizi di cura, si è concentrata sui problemi etici che insorgono a valle della sua implementazione o per ragioni puramente speculative. Concentrarsi sull'IA come artefatto separato dal suo design e dai suoi progettisti rende l'assistenza infermieristica e la cura, come qualsiasi altro settore, in gran parte impotente nei confronti degli impatti dell'IA. Per questo motivo, la focalizzazione sul design e su come progettare rispettando valori e professionalità dei professionisti e degli operatori sanitari è di massima importanza. Sono state proposte vari approcci per raggiungere questo obiettivo, come la progettazione partecipata, la progettazione universale, la progettazione inclusiva e la progettazione orientata ai va-

lori. Nonostante i vantaggi di questi approcci più consolidati, nessuno di essi si concentra sulle narrazioni e le metafore sottostanti che guidano e influenzano come l'IA viene percepita e quindi impatta il lavoro dei professionisti. Questo lavoro esamina brevemente un approccio complementare, ovvero l'IA centrata sull'essere umano (HCAI), che propone nuove narrazioni di potenziamento e miglioramento per gli utenti, piuttosto che narrazioni di sostituzione. Se i professionisti spingono per l'adozione di tale approccio e nuove narrazioni, possono aumentare, piuttosto che abdicare, la loro responsabilità sul posto di lavoro e aumentare la loro capacità di fornire cure adeguate.

PAROLE CHIAVE

Intelligenza artificiale, IA incentrata sull'uomo, Design, Etica applicata.

ABSTRACT

The existing literature on the use of artificial intelligence (AI) in nursing and caregiving has mainly focused on the ethical problems that arise after its implementation or for purely speculative reasons.

However, this perspective leaves the caregiving field largely helpless to the impacts of AI, as it treats AI as an isolated artifact divorced from its design and designers. Therefore, it is crucial to focus on design, respecting the values of professionals and practitioners. Although various approaches, such as participatory design, universal design, inclusive design, and value-sensitive design, have been proposed, they do not address the underlying

narratives and metaphors that shape and influence how AI is perceived and affects professionals in the workplace. This paper introduces a complementary approach, namely Human-Centered AI (HCAI), which proposes new narratives of empowerment and enhancement for users instead of substitution. By advocating for the adoption of this approach and new narratives, professionals can increase their ability to provide quality care, maintain responsibility in the workplace, and avoid abdication of their duties.

KEY WORDS

Artificial Intelligence, human-centred AI, Design, Applied Ethics.

L'IMPORTANZA DEL DESIGN PER LA CURA

Nel contesto dell'assistenza sanitaria, in particolare negli ospedali, l'uso di tecnologie avanzate sta diventando sempre più diffuso. Tuttavia, vi è stata una limitata esplorazione nella letteratura su come guidare e orientare correttamente queste tecnologie per garantire il valore fondamentale della responsabilità umana, che è essenziale per l'assistenza.

Ciò è cruciale sia per i caregiver che per i pazienti per mantenere un controllo significativo sui loro ambienti sociotecnici, che implicano questioni etiche e sociali. Pertanto, è importante considerare il ruolo del design in questo contesto, nonostante non sia una considerazione immediatamente ovvia.

Tuttavia, l'inesorabilità di tecnologie come l'IA e le loro dimensioni sociali non dovrebbero essere confuse con il determinismo tecnologico. Questo causerebbe il dibattito sugli impatti etici della tecnologia, portando a una cattiva attribuzione di chi ha la responsabilità nei nostri spazi di lavoro ad alta tecnologia.

Una migliore cornice dell'argomento è che le tecnologie e la società costruiscono e interagiscono tra di loro. La società determina parzialmente e viene determinata dalla tecnologia progettata e

viceversa. Ciò ci consente di posizionare correttamente gli impatti delle tecnologie all'interno di una narrazione che consente agli esseri umani di guidare quegli impatti attraverso il design.

Pertanto, al centro del modo in cui possiamo farlo, c'è la prospettiva del design, una prospettiva rivolta alla progettazione di tecnologie *per* i valori umani.

La prospettiva del design per i valori è un'area di ricerca emergente che cerca di creare tecnologie che si allineino con i valori umani, come la privacy, l'autonomia e la dignità umana. Nel contesto della cura e dell'assistenza, questa prospettiva è diventata sempre più rilevante poiché la tecnologia continua ad essere integrata nei servizi sanitari e di assistenza domiciliare. Studi hanno evidenziato l'importanza di considerare i valori degli operatori sanitari e dei destinatari dell'assistenza nella progettazione della tecnologia (El-Dassouki et al., 2022).

Ad esempio, sono stati utilizzati approcci di progettazione partecipativa per sviluppare tecnologie che soddisfino le esigenze sia degli operatori sanitari che dei destinatari dell'assistenza (Ahmed et al., 2019).

Tali approcci implicano la partecipazione attiva di entrambe le parti nel processo di progettazione, consentendo loro di esprimere i loro valori e preferenze.

Un altro approccio è il design sensibile ai valori, che implica la considerazione delle implicazioni etiche della progettazione tecnologica e dell'eventuale impatto sugli stakeholder (Umbrello et al., 2021). Questo approccio è stato applicato nel contesto dell'assistenza per garantire che le tecnologie non compromettano la privacy e l'autonomia dei destinatari dell'assistenza (cfr. Schoenhofer et al., 2019).

Nonostante questi approcci offrano vantaggi nell'aiutare i designer ad allineare i loro sistemi ai valori degli stakeholder coinvolti in contesti di cura, gli approcci più tradizionali non considerano in modo sufficiente il ruolo e la significatività delle metafore e delle narrazioni nella modellazione del design e dell'uso di tali sistemi.

NUOVE METAFORE PER AI IN CARE

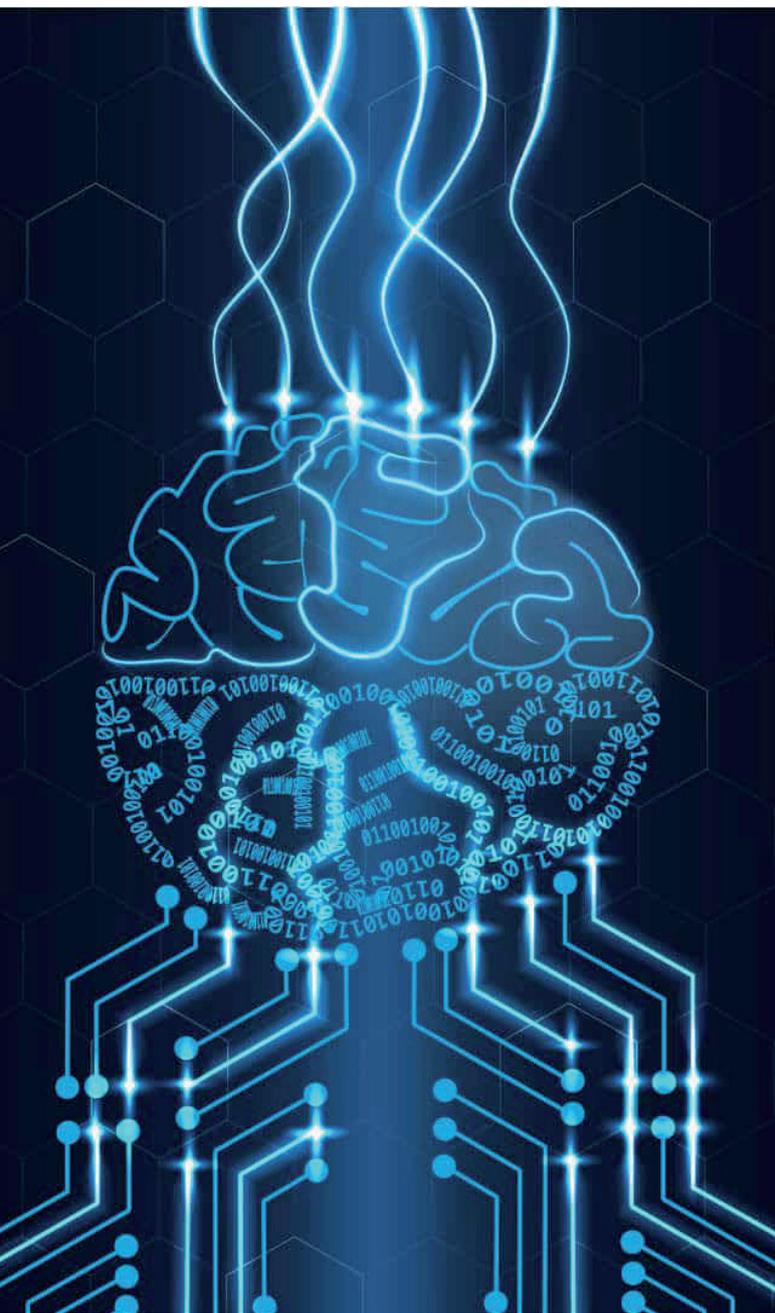
L'IA Centrata sull'Essere Umano (HCAI) è un approccio che combina gli algoritmi di intelligenza artificiale (IA) con il pensiero incentrato sull'essere umano.

Questa sintesi prevede di unificare la ricerca sugli algoritmi di IA con i metodi di progettazione dell'esperienza dell'utente per sviluppare tecnologie in grado di amplificare, aumentare, abilitare e migliorare le prestazioni umane. In altre parole, l'HCAI mira a creare sistemi di IA che non solo siano tecnicamente avanzati, ma anche personalizzati per soddisfare le esigenze umane e migliorare l'esperienza dell'utente (Shneiderman, 2022: 4). I ricercatori e gli sviluppatori dei sistemi HCAI danno valore al concetto di **controllo umano significativo**, ponendo le persone al primo posto, servendo i bisogni, i valori e gli obiettivi umani (cfr. Santoni de Sio e van den Hoven, 2018). Ciò significa che l'HCAI si occupa principalmente di sviluppare tecnologie che consentano agli esseri umani di svolgere il proprio lavoro in modo più efficace, anziché creare tecnologie che svolgano il lavoro o sostituiscano gli agenti umani. Sistemi come internet, posta elettronica, navigazione digitale e fotografia digitale sono esempi di processi tecnologici che migliorano le abilità umane anziché sostituirle (come spiegato da Shneiderman, 2021).

Questo ultimo punto è particolarmente importante considerando che, storicamente, le narrazioni riguardanti le tecnologie emergenti e innovative, in particolare quelle relative alle tecnologie di automazione come l'IA, sono spesso accompagnate da narrazioni e metafore di sostituzione.

Questa preoccupazione non è del tutto infondata, considerando il cambiamento del panorama lavorativo che ha segnato l'introduzione di tecnologie automatizzate all'inizio del ventesimo secolo. Tuttavia, tali narrazioni, se accettate acriticamente, possono diventare una profezia che si autoavvera, generando ulteriori tecnologie basate sull'IA a vantaggio esclusivo dei professionisti umani invece che come supporto al lavoro umano (cfr. Locsin e Ito, 2018).





Il framework HCAI non si concentra esclusivamente sulla tecnologia in fase di sviluppo, ma anche sui valori umani. Si propone di esplorare come la tecnologia possa migliorare e supportare i diritti umani come definiti da statuti consolidati, come ad esempio la Dichiarazione Universale dei Diritti dell'Uomo dell'Organizzazione delle Nazioni Unite (Nazioni Unite, 1949). Inoltre, il framework cerca di rispondere ad altre importanti domande come garantire l'uguaglianza di accesso alla giustizia e all'equità tra diversi grup-

pi, inclusi i gruppi minoritari, e come mantenere la dignità umana. Queste domande centrali costituiscono la base del framework HCAI, sottolineando l'importanza di progettare sistemi di intelligenza artificiale che diano la priorità ai valori umani. Questo approccio impedisce che i valori umani siano considerati come un'idea secondaria o addirittura ignorati del tutto.

Il framework** HCAI si propone di identificare e contrastare le potenziali minacce che possono compromettere il carattere centrato sull'essere umano delle tecnologie AI. Ciò include l'azione di *attori malintenzionati* come avversari, criminali, spammer, terroristi e regimi politici oppressivi, che potrebbero tentare di appropriarsi della tecnologia e utilizzarla a scopi illeciti (Shneiderman, 2022). Inoltre, dobbiamo essere consapevoli del *bias* che può derivare sia dalle nostre attività consapevoli che inconsapevoli. Per affrontare questa problematica, dobbiamo trovare modi per minimizzare il bias, essere consapevoli di esso e assicurarci che i dati che utilizziamo siano imparziali e che vengano utilizzati in modo imparziale. In definitiva, dobbiamo assicurarci che il software che sviluppiamo sia *affidabile e sicuro* (o, in Inglese, *reliable, safe, and trustworthy*; RST) per raggiungere i nostri obiettivi riducendo al minimo il numero di bug presenti.

Ma cosa c'entra tutto questo con le metafore? Come abbiamo già accennato, le metafore e le narrazioni ad esse correlate possono guidare le azioni umane nella loro interazione con le tecnologie. Proprio per questo, dobbiamo assicurarci di scegliere e diffondere narrazioni che rafforzino e migliorino il comportamento umano e, di conseguenza, la responsabilità nella loro interazione con i sistemi di intelligenza artificiale, in particolare nei contesti di assistenza. Il framework HCAI fornisce una prospettiva unica sulla nozione di controllo umano significativo, considera inoltre come i progettisti utilizzino le metafore per comprendere i sistemi e i diversi aspetti della loro gestione nel corso del tempo.

** Framework: Lo scopo del framework software è semplificare l'ambiente di sviluppo, consentendo agli sviluppatori di dedicare i propri sforzi ai requisiti del progetto, piuttosto che occuparsi delle funzioni di controllo software, più generali, semplici e ripetitive, meno attinenti allo specifico, problema demandando questo compito al framework.

Progettare e implementare sistemi di intelligenza artificiale e le loro interfacce digitali comporta l'uso di una serie di metafore che forniscono "qualcosa di noto e di nostra creazione, o almeno della nostra scelta, che mettiamo a rappresentare e, quindi, per aiutarci a capire, qualcosa di sconosciuto e non di nostra creazione" (Olney, 1972: 30). Ciò significa conseguentemente che la scelta delle metafore di design può avere un impatto positivo sull'esperienza degli utenti (Laurel, 1991). Per poter raggiungere questo obiettivo, l'interfaccia deve utilizzare metafore di design adeguate che consentano agli utenti di mantenere il controllo, anche quando si interagisce con sistemi informatici altamente automatizzati. Le metafore scelte per rendere i sistemi di intelligenza artificiale accessibili agli utenti non sono mai neutrali. Riporto qui la famosa osservazione di Lakoff e Johnson (1980), l'uso di metafore specifiche per descrivere cose ed eventi porta sempre a orientare gli atteggiamenti nei loro confronti e a guidare le azioni future. Ciò significa che i designer devono prestare molta attenzione nella

sceita delle metafore appropriate per descrivere i sistemi informatici e le metafore incorporate nelle interfacce. Ad esempio, il fallimentare tentativo di sviluppare interfacce sociali, come Ms. Dewey di Microsoft, presentava un video di una donna che commentava le ricerche degli utenti sui motori di ricerca. Le rappresentazioni sessualizzate e razzializzate degli "assistenti" dell'interfaccia hanno infine influenzato la funzionalità e l'equità di tali sistemi (Sweeney, 2016).

HCAI sottolinea l'importanza di sviluppare metafore efficaci per descrivere e migliorare la tecnologia che mette al primo posto le esigenze umane, un aspetto di particolare importanza e prescienza nel contesto della cura. Nonostante la popolarità del termine "robot sociali", Shneiderman (2022) fa notare che molti tentativi di creare tali robot, spesso sotto forma di umanoidi, sono falliti. Pertanto, egli suggerisce di sostituire tali metafore con un paradigma di pensiero più innovativo che consideri i "super strumenti". Invece di simulare le caratteristiche umane (ad esempio, robot di assistenza), gli strumenti

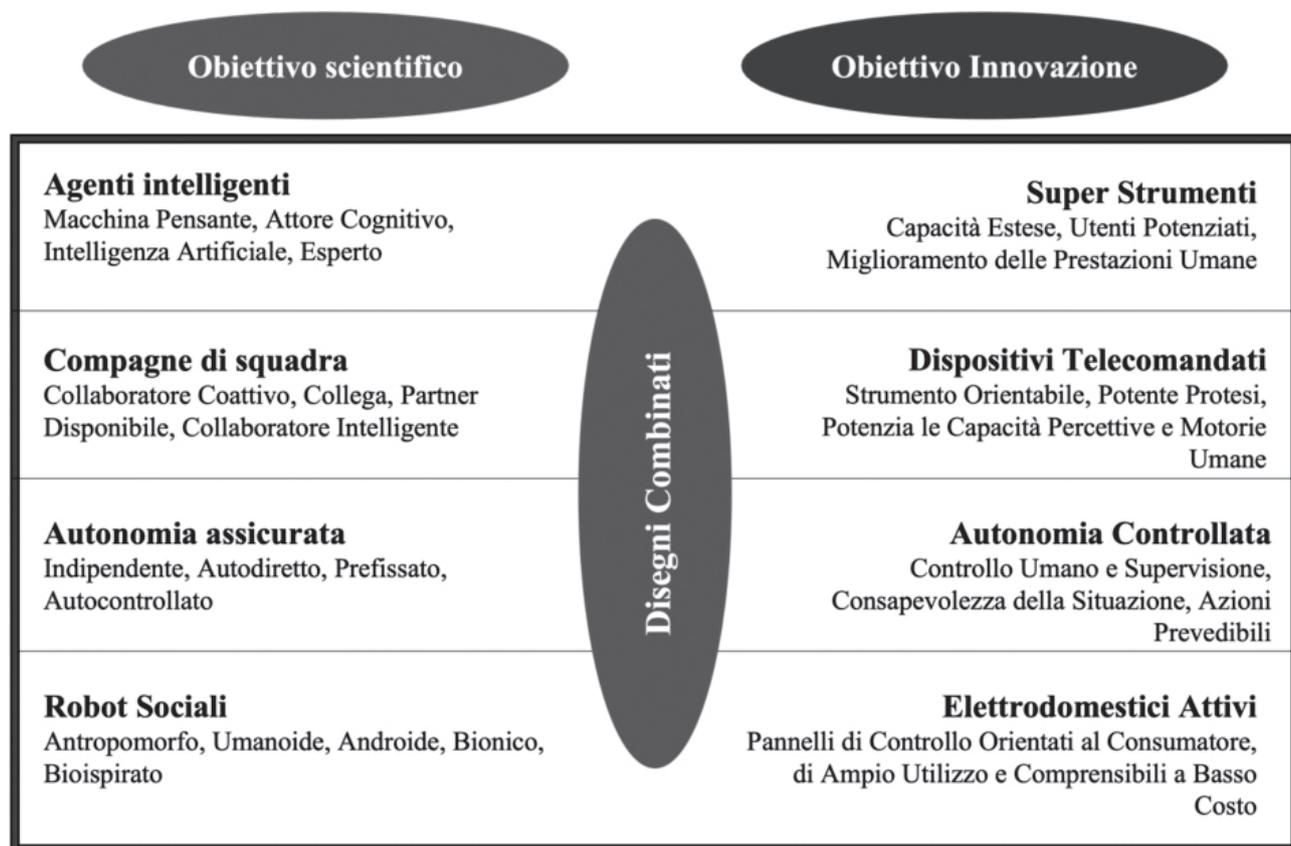


Figura 1. Metafore progettuali. Fonte: Shneiderman, 2020

estendono le capacità umane, come nel caso di dispositivi telecomandati e elettrodomestici attivi. Pertanto, per giungere a un sistema sicuro e affidabile, le metafore utilizzate quando ci si riferisce a questi sistemi devono fare un simile cambiamento. Nella Figura 1, possiamo vedere che l'obiettivo innovativo nel cambio di metafore si riferisce a dispositivi con autonomia supervisionata per garantire il controllo e la supervisione umana. Gli "supertools" possono includere dispositivi come i sistemi di sollevamento, che consentono agli utenti di aumentare la propria mobilità e riacquistare l'autonomia. In questi sistemi, si diventa beneficiari del potere dei sistemi di intelligenza artificiale mantenendo al tempo stesso il controllo predittivo e prospettico. All'utente è permesso di scegliere *prima* che il sistema agisca.

Il concetto di elettrodomestici attivi è meglio illustrato da dispositivi come cuociriso, macchine per il caffè, lavastoviglie, asciugatrici, ecc., che nascondono dietro le loro interfacce apparentemente semplici una serie di sensori e automazioni che consentono all'utente di effettuare scelte riguardo alla loro funzione. Anche i pacemaker, un tempo altamente automatizzati, stanno diventando più incentrati sull'essere umano, consentendo agli utenti di controllare e raccogliere dati attraverso app e di regolare i parametri del dispositivo, fornendo agli esperti sanitari l'accesso ai dati (Tarakji et al., 2021).

Nel contesto dell'assistenza, possiamo vedere come la selezione accurata delle metafore di design per l'HCAI possa contribuire a mitigare le potenziali implicazioni negative che spesso derivano dall'eccessiva antropomorfizzazione dell'AI, come ad esempio la capacità del design di un sistema di ingannare l'utente facendogli credere che il sistema abbia autonomia (Natale, 2021; Umbrello & Natale, 2023). Il vivace dibattito sulle rappresentazioni di genere incorporate nella caratterizzazione dell'IA conversazionale è un esempio significativo. I suggerimenti sull'identità di genere includono il nome femminile scelto per servizi come Alexa o Cortana e la voce di genere. Molti di questi servizi sono stati inizialmente lanciati con la voce femminile come predefinita, creando preoccupazioni per le sovrapposizioni tra gli stereotipi di genere e la metafora di design del lavoratore impiegato at-

traverso cui l'assistenza viene offerta agli utenti (Woods, 2018). L'effetto dell'antropomorfizzazione è facilitato sia dal personaggio con connotazioni di genere sia dalla metafora dell'assistente o del servitore (Woods, 2018), che sono destinati a "offrire alle persone qualcosa a cui acclimatarsi" (come riconosciuto da uno sviluppatore coinvolto nella progettazione dell'IA conversazionale, vedi Young, 2019: 117).

HCAI fornisce una soluzione potenziale al dilemma. Per quanto riguarda il controllo dell'utente, la possibilità per l'utente di selezionare il proprio nome e la voce del genere può mettere l'utente in condizione di controllare le rappresentazioni di genere attivate dagli assistenti; per quanto riguarda le metafore di progettazione, inoltre, concettualizzare l'IA conversazionale come *super strumento fa correre il rischio* di identificarla con le figure professionali. Ciò consente sia al caregiver che al ricevente delle cure di personalizzare la qualità e le esigenze particolari alle loro cure.

CONCLUSIONI

Non c'è dubbio che il contesto dell'assistenza, indipendentemente dalla posizione, continuerà a essere modificato e saturato da tecnologie emergenti e trasformative, in particolare l'AI.

Tuttavia, non possiamo accettare acriticamente questi sistemi date le enormi conseguenze sul modo in cui i professionisti svolgono effettivamente i loro doveri di assistenza, in particolare i loro comportamenti.

Dato che questi sistemi si comportano e influenzano i loro utenti e l'ambiente a causa del design, ha senso guardare al design come luogo di soluzione. Sebbene gli approcci che possono essere adottati a livello di design siano vari, ognuno con i suoi vantaggi e svantaggi, questo breve articolo ha presentato l'approccio HCAI come un mezzo per affrontare un aspetto importante dei sistemi ad alta tecnologia come l'AI, ovvero le metafore di automazione, sostituzione e agenzia che sono spesso accoppiate con le narrazioni sull'AI.

Queste forniscono una potenziale pericolosa e auto-realizzante profezia che deve essere rivista se vogliamo assicurare che i professionisti possano rimanere al controllo e responsabili del tipo e della qualità dell'assistenza che forniscono ai loro pazienti.

RINGRAZIAMENTI

Vorrei ringraziare Francesco Casile che mi ha incoraggiato ad applicare alcuni dei miei pensieri e approcci precedenti al campo della cura.

BIBLIOGRAFIA

- Ahmed, R., Toscos, T., Ghahari, R. R., Holden, R. J., Martin, E., Wagner, S., ... & Mirro, M. (2019). Visualization of cardiac implantable electronic device data for older adults using participatory design. *Applied clinical informatics*, 10(04), 707-718.
- El-Dassouki, N., Pfisterer, K., Benmessaoud, C., Young, K., Ge, K., Lohani, R., ... & Pham, Q. (2022). The Value of Technology to Support Dyadic Caregiving for Individuals Living With Heart Failure: Qualitative Descriptive Study. *Journal of Medical Internet Research*, 24(9), e40108.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Laurel, B. (1991). *Computers as Theatre*. Boston, MA: Addison-Wesley.
- Locsin, R. C., & Ito, H. (2018). Can humanoid nurse robots replace human nurses. *Journal of Nursing*, 5(1), 1-6.
- Natale, S. (2021). *Deceitful media: Artificial intelligence and social life after the Turing test*. Oxford University Press, USA.
- Nazioni Unite. Assemblea Generale. (1949). *Dichiarazione universale dei diritti umani* (Vol. 3381).
- Olney, J. (1972). *Metaphors of self: The meaning of autobiography*. Princeton, NJ: Princeton University Press.
- Santoni de Sio, F., & Van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 15.
- Schoenhofer, S. O., van Wynsberghe, A., & Boykin, A. (2019). Engaging robots as nursing partners in caring: nursing as caring meets care-centered value-sensitive design. *International Journal for Human Caring*, 23(2), 157-167.
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4), 1-31.
- Shneiderman, B. (2021). *Human-Centered AI: Reliable, Safe & Trustworthy* [Video]. Recuperato 21 February 2023, da <https://www.youtube.com/watch?v=05XwczIERvM>.
- Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.
- Sweeney, M. (2017). The Ms. Dewey “experience”: Technoculture, gender, and race. In J. Daniels, K. Gregory & T. McMillan Cottom, *Digital Sociologies* (pp. 401-420). Bristol, UK: Policy Press.
- Tarakji, K. G., Zaidi, A. M., Zweibel, S. L., Varma, N., Sears, S. F., Allred, J., ... & Turakhia, M. P. (2021). Performance of first pacemaker to use smart device app for remote monitoring. *Heart rhythm O2*, 2(5), 463-471.
- Umbrello, S., Capasso, M., Balistreri, M., Pirni, A., & Merenda, F. (2021). Value sensitive design to achieve the UN SDGs with AI: A case of elderly care robots. *Minds and Machines*, 31(3), 395-419.
- Umbrello, S., & Natale, S. (2023). Reframing Deception for Human-Centered AI. *Ethics and Information Technology*, under review.
- Woods, H. S. (2018). Asking more of Siri and Alexa: Feminine persona in service of surveillance capitalism. *Critical Studies in Media Communication*, 35(4), 334-349.
- Young, L. (2019). ‘I’m a Cloud of Infinitesimal Data Computation’ When Machines Talk Back: An interview with Deborah Harrison, one of the personality designers of Microsoft’s Cortana AI. *Architectural Design*, 89(1), 112-117.

