



1506  
UNIVERSITÀ  
DEGLI STUDI  
DI URBINO  
CARLO BO

UNIVERSITÀ DEGLI STUDI DI URBINO CARLO BO  
DIPARTIMENTO DI STUDI UMANISTICI  
CORSO DI LAUREA: SCIENZE E TECNICHE  
PSICOLOGICHE

---

INTELLIGENZA ARTIFICIALE:  
DEFINIZIONE, EVOLUZIONE ED UTILIZZO  
IN AMBITO PSICOLOGICO

**Relatore: Chiar.mo Prof.**  
GIOVANNI BATTISTA  
CAPUTO

**Tesi di laurea di:**  
GIANCARLO PAOLETTI

---

*ANNO ACCADEMICO 2015 -2016*

## INDICE

INTRODUZIONE .....	2
1. Che cos'è l'Intelligenza Artificiale ed applicazioni pratiche delle odierne tecnologie .....	4
1.1. Definizione pratica di Intelligenza Artificiale .....	4
1.2. Storia dell'Intelligenza Artificiale .....	7
1.3. Google DeepMind .....	9
1.3.1. Il Deep-Q-Network e l'Atari 2600 .....	9
1.3.2. AlphaGo in grado di sconfiggere il campione mondiale di Go .....	10
1.4. IBM Watson e l'era del Cognitive Computing .....	11
1.4.1. La vittoria di Watson nel popolare gioco televisivo Jeopardy! .....	12
1.4.2. Watson Health: il Cognitive Computing per la sanità .....	13
1.4.3. Watson e la nuova generazione di giocattoli intelligenti .....	13
1.5. Microsoft e gli occhiali da sole in Realtà Aumentata di Saqib Shaikh .....	14
2. Prospettive ed utilizzi dell'Intelligenza Artificiale nel contesto della Psicologia .....	15
2.1. Psicologia ed Intelligenza Artificiale .....	15
2.2. IBM Watson: Analisi dei tratti di personalità basata sul modello Big Five .....	16
2.3. University of South California ICT: Ellie, la psicoterapeuta virtuale .....	21
2.4. Aspetti etici sull'utilizzo dei chatbot con risvolti positivi e negativi .....	24
2.4.1. Karim: il chatbot che aiuta i rifugiati Siriani colpiti da stress psicologico .....	24
2.4.2. Microsoft Tay: come diventare razzista in 16 ore .....	25
3. Aspetti etici dell'Intelligenza Artificiale .....	27
3.1. Aspetti negativi .....	27
3.1.1. Il fenomeno della Bolla di Filtraggio .....	29
3.1.2. Il fenomeno dell'Uncanny Valley .....	30
3.2. Aspetti positivi .....	31
3.2.1. La Singolarità Tecnologica .....	33
3.2.2. Il movimento Transumanista .....	33
CONCLUSIONI .....	35
BIBLIOGRAFIA .....	37

## INTRODUZIONE

Il presente lavoro ha l'obiettivo descrivere ed analizzare in chiave storica l'evoluzione e gli utilizzi odierni dell'Intelligenza artificiale, in particolare modo nel campo psicologico. Il cuore della tesi verte sull'impiego dell'Intelligenza Artificiale come strumento d'utilizzo in grado di aiutare lo psicoterapeuta nel suo lavoro clinico oppure in altri contesti sempre collegati alla psicologia. Inoltre lo studio sottolinea, in termini quantitativi e qualitativi, gli aspetti etici di questa tecnologia che nel bene o nel male influenzerà pesantemente il nostro modo di vivere negli anni futuri. Il lavoro è costituito da un'ampia gamma di articoli scientifici e giornalistici essenziali per affrontare i temi proposti.

Il primo capitolo offre una spiegazione esaustiva della definizione di Intelligenza Artificiale. Le sue terminologie specifiche sono descritte in modo da risultare facilmente comprensibili e viene proposta successivamente una descrizione delle moderne tecnologie realizzate dalle grandi aziende del settore per scopi commerciali oppure per usi accademici e di ricerca.

Il secondo capitolo dell'elaborato si collega direttamente a quello precedente e spiega come queste tecniche di Intelligenza artificiale possano essere utilizzate in campo psicologico. Cuore di questa parte verte su 3 aspetti: analisi dei tratti di personalità, utilizzo di assistenti virtuali ed utilizzo di chatbot.

Il terzo ed ultimo capitolo esamina gli aspetti etici dell'utilizzo dell'Intelligenza Artificiale, ponendo maggior enfasi sugli aspetti più pessimistici ed ottimistici. I primi si riferiscono ad un uso non ponderato di questa tecnologia e che potrebbe portare a crisi economiche come perdita massiccia di posti di lavoro oppure ad aspetti più tragici come l'obsolescenza umana. I secondi aspetti, quelli positivi, sono rivolti alla capacità umana di provare empatia, anche per le macchine, inoltre sono riportate opinioni riguardanti quegli aspetti positivi che l'Intelligenza Artificiale è in grado di realizzare, per portare l'umanità verso un futuro prospero.

Le opinioni che abbiamo sull'Intelligenza Artificiale sono spesso variegata e basate su preconcetti che abbiamo indirettamente acquisito dai media, dalla cultura popolare e dalle arti come ad esempio il cinema o televisione. Un altro interessante aspetto sull'argomento è che al

giorno d'oggi esistono già in commercio diverse tipologie di Intelligenza Artificiale ed alcune di esse le stiamo già utilizzando in modo massiccio e repentino, tant'è che nemmeno ce ne rendiamo conto. Queste tecnologie sono talmente ben radicate nella nostra quotidianità che oltre a non esserne consci della loro esistenza ci rendono un gran servizio in diverse mansioni. Uno tra tutti è l'utilizzo dell'Intelligenza Artificiale come nostro assistente nei più disparati compiti: ci aiuta a gestire grandi quantità di dati, a gestire meglio il nostro lavoro e consigliarci quale prodotto sia adatto secondo le nostre necessità. Sia in ambito domestico che lavorativo.

L'obiettivo di questa tesi è illustrare come questa Intelligenza Artificiale può essere usata, migliorata ed implementata nel campo della psicologia. Fornire una visione generale dell'argomento è fondamentale per assicurarsi che questa tecnologia possa fornire un grande aiuto per gli psicologi nel mondo e che possa anche farci comprendere meglio le funzionalità della nostra psiche.

# 1. Che cos'è l'Intelligenza Artificiale ed applicazioni pratiche delle odierne tecnologie

## 1.1. Definizione Pratica di Intelligenza Artificiale

Il termine Intelligenza Artificiale (IA) specifica la branca della Computer Science che sviluppa computer in grado di svolgere attività compiute solitamente dall'essere umano. I ricercatori utilizzano diversi approcci nell'affrontare le problematiche relative all'IA e svilupparne soluzioni efficaci:

La *IA Forte (Strong AI)* mira alla costruzione di sistemi in grado di riprodurre l'intelletto umano in tutti i suoi aspetti, una Intelligenza Artificiale Generale (AGI, Artificial General Intelligence) in grado di ragionare al pari dell'essere umano. Questi sistemi sono ancora ben lontani dall'essere costruiti.

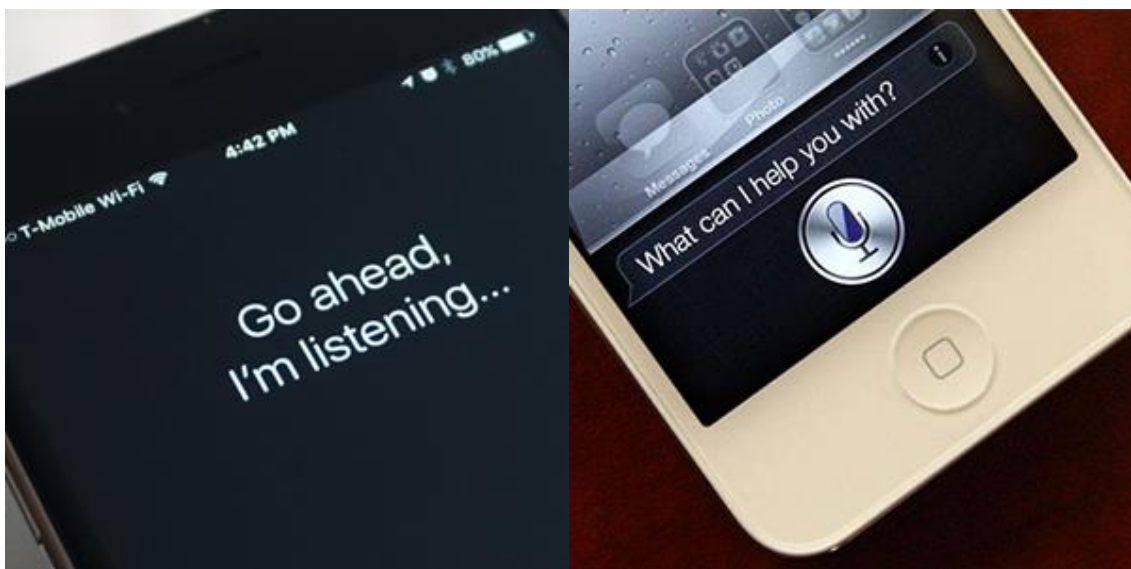
La *IA Debole (Weak AI)* ha l'obiettivo di sviluppare sistemi in grado di svolgere funzioni tipiche del ragionamento umano ma non si ha pretese sul sapere *come* funzionano nella mente umana. Un esempio lampante lo si può trovare nel sistema IBM Deepblue, il celeberrimo supercomputer che sconfisse il campione di scacchi Garry Kasparov nel 1997 utilizzando il metodo della forza bruta. Il supercomputer analizzava la scacchiera e calcolava tutte le possibili mosse disponibili, scegliendo la migliore grazie alla potenza computazionale della macchina IBM. Nonostante l'impressionante risultato, questo sistema di IA Debole non giocava "come un umano" e forniva pochissime informazioni sulla cognizione umana in generale.



11 maggio 1997: Il supercomputer IBM Deepblue batte il campione di scacchi Garry Kasparov.

Un terzo approccio si è instaurato con prepotenza negli ultimi anni e si tratta della *IA Applicata (Applied AI)*, una via di mezzo tra i due approcci precedenti. Si utilizza il ragionamento umano come modello guida ma non si è vincolati dall'obbligo di riprodurlo fedelmente. Alcuni sistemi possono ad esempio essere sviluppati solo su componenti specifiche della mente umana come il riconoscimento delle immagini o la comprensione di un dialogo. A differenza della IA Forte, questi sistemi convergono tutti i loro sforzi solo su una specifica funzione della mente umana, garantendo risultati di eccellenza in quel determinato campo di sviluppo. Il rovescio della medaglia è che si rendono obsoleti su altri aspetti di cui non sono stati programmati.

Questi sistemi di IA fanno riferimento alle 3 categorie principali dell'intelligenza umana. La *percezione* permette l'acquisizione dei dati del mondo esterno attraverso sensori, algoritmi di riconoscimento immagini e riconoscimento suoni o dialoghi. Il *ragionamento* si basa sulla logica inferenziale e l'*intervento* pratico rappresenta la generazione di risposta sotto forma di dialogo oppure come azione di altro tipo. Ad esempio l'assistente vocale Apple Siri, alla richiesta di fornire indicazioni circa un ristorante nelle vicinanze, analizza il messaggio vocale generato dall'utente, interpreta quanto percepito e fornisce una risposta adeguata alla domanda posta dall'utente, attingendo anche alla grande mole dei Big Data.



*L'assistente vocale Siri, acquistato dalla Apple nel 2010 per 200 milioni di dollari circa.*

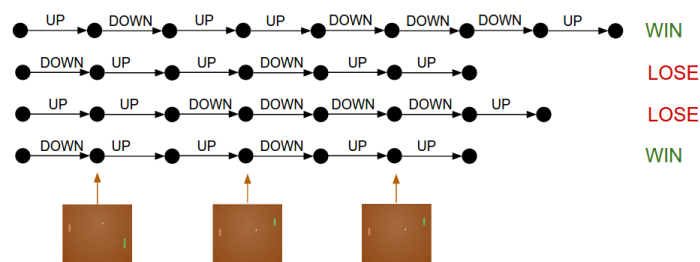
Il termine *Big Data* è riferito all'impressionante quantità di dati non strutturati che vengono prodotti nei nostri giorni, si stimano circa 2.5 exabyte ogni giorno (2 miliardi e mezzo di chiavette USB da 1 Giga) ed il 90% dei dati digitali prodotti al mondo è stato creato solamente negli ultimi 2 anni. Alcuni studi stimano che il loro volume totale raggiungerà i 35.000 exabytes nel 2020 ed il numero è destinato ad aumentare esponenzialmente. I Big Data sono letteralmente una miniera di informazioni: la ricchezza delle aziende investitrici nel campo dell'IA infatti non risiede nel possedere le macchine più potenti o l'algoritmo più innovativo (di cui spesso i risultati delle ricerche vengono divulgati liberamente in Open Source) bensì consiste nel possedere questi Big Data, basti pensare a quante ricerche vengono fatte giornalmente su Google oppure sull'utilizzo massiccio dei Social Media. I sistemi di IA permettono di trasformare tale mole in dati di conoscenza e profitto. IBM Watson è in grado di analizzare quantità abnormi di dati in pochi secondi, al contrario di un essere umano che per svolgere tale lavoro potrebbe impiegarci settimane.

A seconda del tipo di dati d'ingresso ottenuti, i sistemi di IA possono usare 3 tipologie di apprendimento: Supervised Learning, Unsupervised Learning e Reinforcement Learning.

Nel *Supervised Learning* vengono presentati al computer sia l'input che l'output richiesto dall'utente, l'obiettivo è quello di imparare regole generali che definiscono il rapporto tra input e output in un ambiente controllato e stabilito a priori.

Nell'*Unsupervised Learning* non vengono fornite alcune indicazioni dall'utente, in modo tale da lasciar totale libertà all'IA di scoprire le sottoprocedure d'apprendimento.

Infine nel *Reinforcement Learning* l'IA riceve informazioni riguardo l'output richiesto dall'utente ma non viene informata su quanto è vicina alla risoluzione del problema. Mediante il sistema d'apprendimento per rinforzi i risultati ottenuti saranno sempre più utili e precisi man mano che l'IA ripete le azioni che producono un rinforzo positivo e scarta quelle che producono un risultato negativo oppure nullo.

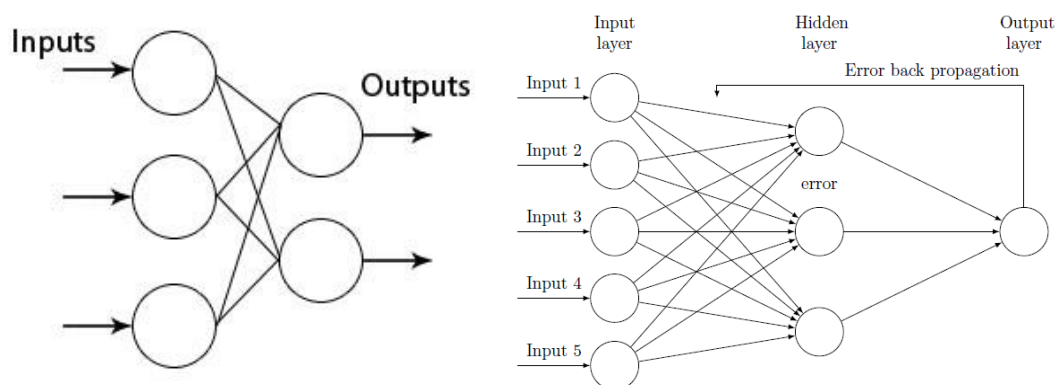


Esempio di IA che usa il Reinforcement learning per imparare a giocare a Pong.

## 1.2. Storia dell'Intelligenza Artificiale

Per comprendere al meglio come l'IA si è evoluta fino a nostri giorni è opportuno fare una digressione storica. Lo psicologo Frank Rosenblatt sviluppò nel 1957 un modello matematico in grado di comprendere il funzionamento del neurone cerebrale: il *Percettrone*. Le moderne reti neurali (*Artificial Neural Networks ANNs*) possono essere viste come la somma di numerosi strati di Percettroni, i quali rappresentano l'unità base come i neuroni del nostro cervello.

Il neurone artificiale si ispira appunto alla sua controparte biologica. Se una serie di input binari (come le sinapsi) riesce a superare una soglia prestabilita (determinata dalla potenza del rispettivo input) si potrà generare un segnale di output. Quest'ultimo potrà essere usato come un nuovo input per un altro processo. Successivamente alla scoperta del Percettrone vennero implementate tecniche come la *Retropropagazione* e il *Reinforcement Learning*, permettendo alle reti neurali di diventare sempre più sofisticate e ad imparare dai propri errori. La rete neurale perciò ripete più volte il suo ciclo di produzione dati per generare un output sempre più preciso e filtrato rispetto allo standard prefissato e in caso di successo la rete neurale ottiene un rinforzo positivo nel proseguire quella particolare metodologia utilizzata, in caso contrario continuerà il ciclo di retropropagazione fino a quando non otterrà il risultato generante rinforzo positivo.



Esempio di Rete Neurale a Singolo Strato e Rete Convolutionale con Retropropagazione:  
i segnali dei vari input vengono convogliati prima su strati "interni" nascosti della rete neurale e nel caso in cui il risultato non dovesse essere soddisfacente o generante errore, verrà ripropagato nuovamente nello strato interno per ripetere il processo fino a quando il problema non sarà risolto.



È opportuno notare come il percorso tecnologico delle reti neurali sia stato costellato da diversi periodi di pause e disinteresse, le cause di tale fenomeno sono l'insufficienza di potenza computazionale delle macchine utilizzate nella seconda metà del XX secolo e la mancanza di informazioni necessarie per ottenere risultati precisi. La *legge di Moore* stabilisce che la potenza di calcolo di un processore raddoppia ogni 18 mesi e negli ultimi decenni abbiamo assistito ad un incremento esponenziale della tecnologia. Partendo dalle prime macchine elettromeccaniche e proseguendo via via verso le valvole, i transistor ed i moderni microprocessori, la società moderna si sta sempre più digitalizzando. Precedentemente nel campo dell'IA mancava ancora la necessaria potenza di calcolo in grado di far funzionare le reti neurali secondo tempi non proibitivi.

Negli ultimi anni però stiamo assistendo ad un vero risascimento in campo informatico, dato che si è raggiunto il limite della legge di Moore si è optata come soluzione l'utilizzo in parallelo delle unità di schede grafiche (*GPU, Graphics Processing Unit*). Grazie a tale potenza di calcolo decisamente superiore rispetto alle tecniche precedenti di computazione e con la vastissima banca d'informazione dei Big Data, le reti neurali hanno conseguito risultati straordinari nell'ultimo decennio, per cui oggi possiamo parlare di apprendimento approfondito (*Deep Learning*) in grado di risolvere problemi concreti e di traghettare l'IA verso una nuova era di ricerche e scoperte.

Il *Deep Learning* rappresenta oggi l'apice delle reti neurali in quanto pone accento innanzitutto su uno specifico compito, sulla sua risoluzione nel miglior modo possibile per poi passare a quello successivo. In sostanza l'IA forte potrà realizzarsi solo dalla combinazione delle singole IA Applicate ben costruite. Aziende come Google e Facebook stanno utilizzando il Deep Learning nel categorizzare oggetti, parole e relazioni attingendo alle loro vastissime moli di dati. Il successo del Deep Learning è da ricercare nei progetti svolti durante la prima metà degli anni 2000. L'esempio lampante è legato all'esperimento del team Google Brain: il raddoppio dei tassi di riconoscimento automatico di gatti contenuti in 10 milioni di fotogrammi estrapolati da YouTube attraverso una rete composta da 16 mila processori e un miliardo di connessioni, il tutto sotto regime di Unsupervised Learning e senza che l'IA fosse precedentemente istruita su cosa dovesse riconoscere.

### 1.3. Google DeepMind

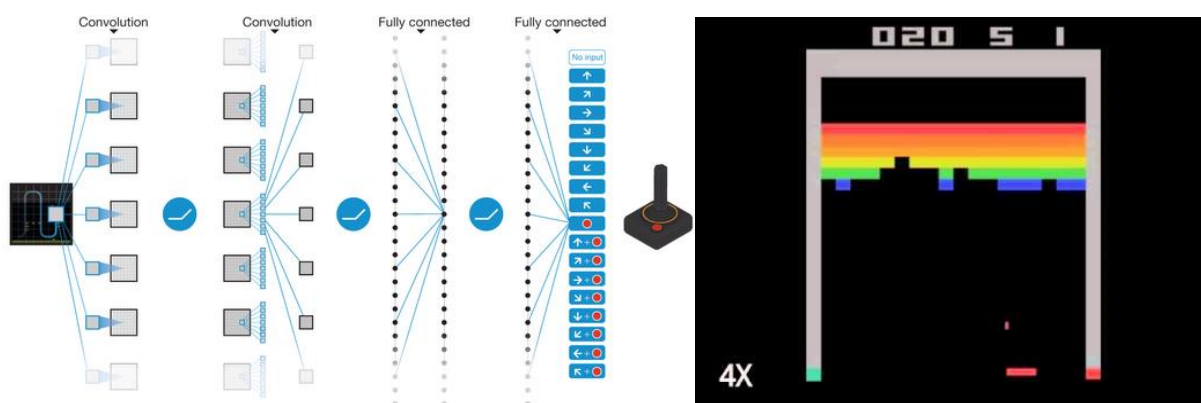
L'azienda londinese fondata nel 2010 da Demis Hassabis venne acquisita da Google nel 2014 per 400 milioni di dollari, confermando l'ondata di grandi investimenti che le aziende tecnologiche stanno compiendo nel settore dell'IA. Degne di nota sono le due pietre miliari che l'azienda di Hassabis è riuscita a conseguire: la rete neurale in grado di imparare a giocare autonomamente con videogiochi anni '80 e la rete neurale che riuscì a battere il campione mondiale di Go, l'antico gioco strategico da tavola cinese.

#### 1.3.1. Il Deep-Q-Network e l'Atari 2600

Il Deep Reinforcement Learning della rete neurale *Deep-Q-Network* (DQN) ha permesso a questa IA di imparare le regole di alcuni giochi della console Atari 2600, i risultati pubblicati nell'articolo "human-level performance in video games" dimostrano che in 43 casi su 49 questa rete neurale è riuscita ad ottenere un punteggio superiore rispetto a quello ottenuto dagli umani. L'unica istruzione ricevuta dall'IA è rappresentata dagli input di gioco (ciò che vede l'umano di fronte al televisore mentre gioca) e l'obiettivo di ottenere il massimo punteggio possibile, tutto il resto è affidato completamente alla rete neurale stessa in regime di Unsupervised Learning. Il DQN rappresenta l'unione delle reti neurali profonde (Deep Neural Networks) con l'apprendimento per rinforzi: diversi livelli di algoritmi sono usati per costruire progressivamente rappresentazioni sempre più astratte dei dati, permettendo ad esempio di imparare concetti come la categorizzazione degli oggetti direttamente dai dati d'ingresso grezzi.

L'architettura utilizzata è la *Deep Convolutional Network* la quale utilizza livelli gerarchici di filtri convoluzionali (in grado di ripropagare su sé stessi il risultato ottenuto per trasmettere al neurone successivo solo le informazioni richieste e filtrate) imitando gli effetti dei recettori della corteccia visiva del cervello umano. Questa rete neurale agisce secondo la sequenza dell'osservazione, azione e ricompensa. L'obiettivo è selezionare azioni in grado di massimizzare la ricompensa finale.

In alcuni giochi come Breakout, in cui l'obiettivo del giocatore è colpire una pallina mediante una racchetta e distruggere la fila di blocchi posti in cima, strategie a lungo termine possono massimizzare il punteggio con uno sforzo relativamente basso. Il DQN dopo diverse partite svolte è riuscito a scoprirle da solo e senza alcuna istruzione o consiglio fornite dagli umani.



Il Deep-Q-Network che apprende lo schema dei comandi e lo applica in una partita a Breakout.

Al seguente indirizzo una dimostrazione pratica: <https://www.youtube.com/watch?v=V1eYniJ0Rnk>

### 1.3.2. AlphaGo in grado di sconfiggere il campione mondiale di Go

Il popolare gioco da tavola Go è un antichissimo passatempo cinese antico di oltre 2500 anni ed è menzionato anche negli scritti di Confucio. La particolarità di questo gioco è la quantità di mosse che ha a disposizione un giocatore: il numero massimo supera la quantità di tutti gli atomi presente nell'universo. A differenza degli scacchi, per vincere una partita a Go non è possibile utilizzare il metodo della forza bruta: è necessario utilizzare l'intuito e ciò per una rete neurale rappresenta una sfida considerata per molti insormontabile. Difatti molti esperti del settore dopo l'annuncio del 27 gennaio 2016 di Google di voler sfidare il campione mondiale di Go, Lee Sedol, profetizzavano il successo di tale sfida nel giro di un decennio. Dopo aver battuto il campione europeo Fan Hui per 5 partite a 0 AlphaGo ha battuto Lee Sedol il 15 marzo 2016 con l'impressionante risultato di 4 - 1.

Anche AlphaGo utilizza l'algoritmo del Deep Reinforcement Learning, in cui i dati iniziali sono circa centomila partite di dilettanti disputate online. Il primo passo compiuto è allenare una Supervised Network a predire le mosse degli umani, in quanto si ha un input ed un output fissati: rispettivamente la posizione delle pedine e la mossa giocata dall'essere umano. Il secondo passo è costruire una seconda Reinforcement Network a partire da quella Supervised, migliorando le mosse prodotte con l'obiettivo di vincere la partita. La rete gioca milioni di partite e per ognuna di queste la "ricompensa" arriva solo alla fine ed è usata per incentivare (nel caso di vittoria) o disincentivare tutte le mosse prodotte durante la partita.



Una partita svolta da AlphaGo e Lee Sedol

#### 1.4. IBM Watson e l'era del Cognitive Computing

Il sistema IBM Watson prende il nome dal fondatore dell'omonima azienda, Thomas J. Watson e l'obiettivo di questa IA è quella di creare un sistema in grado di competere con la capacità umana nel rispondere a domande poste in linguaggio naturale, con velocità, accuratezza e coerenza. Per perseguire tale scopo uno dei principali sforzi compiuti è quello di strutturare la mole dei Big Data ed estrarne valore utile per i vari settori di cui ne hanno bisogno: questi dati si presentano spesso sotto forma destrutturata in innumerevoli formati come ad esempio testo, audio, immagini, video. Dati che sono sempre più prodotti in autonomia dalle stesse macchine e che vengono archiviati in molteplici luoghi: dai tradizionali database ai social networks. Perciò il team di ricerca della IBM offre ai suoi clienti una piattaforma di sviluppo utilizzando la tecnologia del Cloud Computing: chiunque potrà accedere ai servizi offerti da Watson attraverso un qualsiasi computer collegato ad internet. I vantaggi sono molteplici in quanto si ha un'elevata mobilità, flessibilità d'uso e vengono evitati attacchi di ingegneria inversa dei server IBM. È bene menzionare ancora il fatto che la ricchezza dell'azienda non è l'algoritmo in sé e la macchina su cui lavora bensì la quantità e qualità di dati che vengono accumulati, appresi, elaborati e trasformati. Il luogo di condivisione informazioni è l'IBM Watson Developers Cloud, un marketplace disponibile in cloud in cui i clienti potranno attingere a risorse per sviluppare applicazioni "potenziate" dalle capacità cognitive di Watson.

Recentemente il presidente del Consiglio Matteo Renzi ha firmato un accordo per il riutilizzo di alcuni spazi e di alcune strutture di Expo 2015 con un investimento di 150 milioni di dollari per far nascere il nuovo quartier generale europeo della IBM per il Watson center. Si occuperà di sanità e soprattutto della gestione informatica di tutti gli aspetti relativi alle sfide sanitarie: dai controlli e gli aggiornamenti delle cartelle cliniche all'ottimizzazione delle voci di spesa della sanità pubblica, gli esperimenti genetici, la ricerca e cura per il cancro.

### 1.4.1. La vittoria di Watson nel popolare gioco televisivo Jeopardy!

Dopo quattro anni di sviluppo da parte del team di ricerca IBM ed in collaborazione con importanti università di tutto il mondo (tra cui l'Università di Trento) la tecnologia Question Answering (QA) del sistema Watson ha lanciato il guanto di sfida contro i migliori giocatori di Jeopardy!, per dimostrare una competenza linguistica paragonabile a quella di un esperto umano. Questa sfida è paragonabile a quella di IBM Deepblue contro Kasparov agli scacchi o a quella di AlphaGo contro Lee Sedol a Go perché gli indizi del gioco prevedono l'analisi di sottigliezze di significato, ironia, enigmi e altre complessità in cui gli esseri umani eccellono e cui le macchine hanno invece gravi lacune. Nelle puntate del 14 – 16 febbraio 2011 Watson è riuscito efficacemente a sconfiggere i campioni Ken Jennings e Brad Rutner, rispettivamente il detentore di record di permanenza nello show ed il campione che ha vinto il montepremi più alto.

Con la potenza di calcolo di 500 Gigabyte al secondo, il sistema Watson partendo da una domanda, l'analizza e la scompone in sottocategorie attraverso regole di sintassi e componenti semantiche. A questo punto vengono generate ipotesi cercando nel database dei dati e selezionando i risultati con il più alto grado di confidenza (i risultati con maggior "peso", quelli simili infatti si rinforzano a vicenda). Questa eventuale risposta da proporre come definitiva deve prima però passare la soglia di filtraggio per categorie ed essere sottoposta nuovamente ad una verifica con il database per essere sicuri che la scelta sia corretta. Infine la risposta verrà generata grazie ad un sistema Text-to-Speech e comunicata al presentatore dello show televisivo.



Il match televisivo di Jeopardy! Da sinistra Ken Jennings, Watson, Brad Rutner

### **1.4.2. Watson Health: il Cognitive Computing per la sanità**

Diversi ospedali americani utilizzano Watson per aiutare i medici, primo tra tutti lo Sloan-Kettering Cancer Center. L'IA della IBM è in grado di memorizzare milioni di referti medici, tac, immagini, lastre relative a decine di migliaia di pazienti malati di cancro, aiutando i medici a formulare le diagnosi e a trovare le cure migliori confrontando alla velocità della luce il caso in esame con tutti quelli in archivio. È stato calcolato che a un dottore servirebbero 10 mila settimane per leggere e capire 10 milioni di casi di singoli pazienti, mentre Watson lo fa in pochi secondi.

Watson Health rende disponibili ai medici e agli altri operatori sanitari le funzionalità di cognitive computing avanzate di IBM Watson per fornire un'assistenza di qualità, consentendo di rendere anonime, condividere e combinare queste informazioni con altri dati sanitari. È proprio grazie al fenomeno dei Big Data che Watson Health può funzionare efficacemente: partendo dalla cartella clinica "universale", ogni singolo dato del paziente può essere integrato dai rispettivi medici specializzati e anche dal paziente stesso, ad esempio utilizzando i nuovissimi dispositivi indossabili come i braccialetti Fitbit oppure l'Apple Watch.

Sensori biometrici ed altre tecnologie permettono una raccolta costante di dati che possono integrarsi con l'IA grazie al Cloud e grazie a tutte queste tecnologie vi è la concreta disponibilità di accorciare ad esempio tempi di attesa, miglioramenti dei servizi sanitari e tanto altro ancora.

### **1.4.3. Watson e la nuova generazione di giocattoli intelligenti**

L'azienda Elemental Path ha introdotto sul mercato un particolare giocattolo dinosauro che racconta barzellette, si ricorda del nome del suo proprietario e risponde alle più disparate domande. Questa invenzione, denominata CogniToys è stata ben recepita dal pubblico ed il suo nucleo centrale è composto da Watson. Utilizzando i servizi Cloud dell'IA IBM e le sue capacità di elaborazione di linguaggio naturale, il giocattolo è in grado di interagire con il bambino, rispondendo a qualsiasi domanda gli si ponga. I ricercatori non vedono i CogniToys come semplice giocattolo ma come un vero e proprio compagno di giochi: il dinosauro fa imparare nozioni al bambino ed inoltre impara anche da sé, perciò più lo si utilizza e più impara, interagisce ed evolve il suo bagaglio di nozioni ed informazioni.

## 1.5. Microsoft e gli occhiali da sole in Realtà Aumentata di Saqib Shaikh

Saqib Shaikh, un ingegnere della Microsoft, ha sviluppato una tecnologia in grado di compensare la perdita della vista verificatasi nell'infanzia. Utilizzando tecnologie di riconoscimento visivo ed avanzati algoritmi di Machine Learning, è riuscito assieme al suo team di ricerca a costruire un paio di occhiali da sole in cui vi è integrato un piccolo computer in grado di distinguere ed interpretare i dati in tempo reale. Con un semplice gesto, i sensori ottici presenti negli occhiali “scattano” un'istantanea di ciò che vedono in tempo reale e forniscono mediante messaggio audio una descrizione di ciò che accade di fronte a Saqib. Applicato su larga scala, questa serie di tecnologie di Realtà Aumentata (Augmented Reality) possono fornire alle persone con disabilità un notevole aiuto nella vita di tutti i giorni e ridurre drasticamente le barriere architettoniche sia fisiche che mentali.



Saqib Shaikh mentre utilizza i suoi speciali occhiali da sole.

## 2. Prospettive ed utilizzi dell'Intelligenza Artificiale nel contesto della Psicologia

### 2.1. Psicologia ed Intelligenza Artificiale

In termini storici la prima simulazione di uno psicoterapeuta che utilizzi l'interfaccia uomo-macchina fu il programma ELIZA, compilato nel 1966 per imitare lo stile di comunicazione empatica di Carl Rogers ed utilizzava un sistema di domande e risposte formulate attraverso l'uso di una tastiera. La sintassi di ELIZA era basata su modelli programmati perciò la comunicazione risultava abbastanza limitata.

Nei primi anni '70 invece lo psichiatra Kenneth M. Colby sviluppò nell'università di Stanford un programma denominato PARRY, di cui simulava una persona affetta da schizofrenia paranoide e, alla pari di ELIZA, il programma era in grado di comunicare con altre persone. Viene menzionato come il primo programma in grado di passare il test di Turing in quanto diversi psichiatri non erano in grado di distinguere differenze tra PARRY e persone reali affette da schizofrenia paranoide

```
=====
EEEEEEEE L      IIIIIII ZZZZZZZ AAA
E         L      I      Z      A  A
E         L      I      Z      A  A
EEEEEE   L      I      Z      A  A
E         L      I      Z      AAAAAA
E         L      I      Z      A  A
EEEEEEEE LLLLLLL IIIIIII ZZZZZZZ A  A
=====
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE...!
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...
=====
```

Il chatbot ELIZA.

Nel 1950 Alan Turing sviluppò l'omonimo test come metodo per giudicare il grado d'intelligenza di una macchina: per passare tale test un programma di IA deve impersonare una conversazione scritta in tempo reale con un giudice umano e quest'ultimo non può distinguere concretamente se il suo interlocutore sia una macchina oppure no.



Al giorno d'oggi l'evoluzione tecnologica ha permesso lo sviluppo di avatar umani digitalizzati attraverso la Realtà Virtuale, essi sono in grado di sostenere conversazioni interattive ed intelligenti. I ricercatori della University of Southern California's Institute for Creative Technologies (USC ICT) hanno sviluppato su questo versante dei pazienti virtuali da usare per il training clinico e non solo: sono pazienti virtuali che mimano sintomi di disturbi psicologici ed interagiscono con i terapeuti tirocinanti attraverso dialogo verbale.

Applicazioni pratiche di questa tecnologia sono il progetto SimCoach e SimSensei (cui verrà esposto in modo dettagliato più avanti). SimCoach è stato sviluppato per facilitare la coesione tra i militari in servizio e le loro famiglie alla sanità americana e ad altri servizi di caring psicologici. Questo tipo di IA è potenzialmente rivoluzionaria in quanto può essere utilizzata istantaneamente da chi ne ha bisogno grazie ad internet in qualsiasi momento e può fornire in tempo reale assessment di base e trattamenti in grado di gestire l'ansia oppure attacchi di panico per chi soffre di disturbi da stress post-traumatico, per citare un esempio. Le possibilità sono vastissime anche per sole attività di counseling.

Queste nuove IA possono inoltre essere di grande supporto per il clinico umano in quanto possono fornirgli strumenti diagnostici come immagini ad infrarossi per rilevare cambiamenti di temperatura corporea del paziente, i sensori ottici possono far notare al clinico dei cambiamenti quasi impercettibili di espressioni facciali oppure di inflessioni della voce. Possono in sostanza essere un valido strumento per fornire livelli di informazione dati che potrebbero sfuggire anche al terapeuta più esperto.

## **2.2. IBM Watson: Analisi dei tratti di personalità basata sul modello Big Five**

Tra i tanti servizi messi a disposizione dalla piattaforma Cloud di IBM Watson emerge una componente molto interessante da utilizzare nel campo della psicologia: l'analisi della personalità utilizzando il modello Big Five. Il linguaggio umano riflette la personalità, il modo di pensare, le connessioni sociali e gli stati emozionali. Questa teoria è ampiamente condivisa in psicologia, nel marketing ed in altri contesti. Perciò questo servizio è basato sulla psicologia del linguaggio in combinazione con gli algoritmi di analisi dati e li fornisce attraverso 3 modelli in cui l'informazione è fornita mediante testo scritto. Per comprendere quanto testo è necessario per dedurre le caratteristiche di personalità sono state usate fonti di dati provenienti da Twitter, email, blog, forum, e pagine wiki: questi modelli sono il modello Big Five, modello delle necessità e modello dei valori.

Il *modello Big Five* sviluppato dagli psicologi Costa e McCrae viene usato spesso per descrivere come una persona interagisce con il mondo che lo circonda, include cinque caratteristiche primarie della personalità ed ognuna di queste caratteristiche possiede delle sfaccettature che caratterizzano ulteriormente la personalità dell'individuo:

- Amicalità: la tendenza di una persona ad essere compassionevole e cooperativa verso gli altri
  - Altruismo: le persone con punteggio alto trovano che l'aiutare altre persone sia soddisfacente
  - Cooperazione: le persone con punteggio alto non apprezzano il confronto diretto e preferiscono scendere a compromessi per convivere con gli altri
  - Modestia: le persone con punteggio alto sono più umili ma non necessariamente possiedono una mancanza di autostima
  - Sincerità: le persone con punteggio alto non vedono pretese di manipolare altri e si dimostrano genuine
  - Simpatia: le persone con punteggio alto sono compassionevoli
  - Fiducia: le persone con punteggio alto prendono in considerazione ciò che dice l'interlocutore e tendono a perdonare
  
- Coscienziosità: la tendenza di una persona ad agire in maniera organizzata e con criterio
  - Realizzazione di uno sforzo: le persone con punteggio alto compiono grandi sforzi per conseguire tale risultato
  - Cautela: le persone con punteggio alto pensano attentamente prima di agire
  - Senso di responsabilità: le persone con punteggio alto sono ligie al dovere
  - Organizzazione: le persone con punteggio alto sono ordinate nel mettere al loro posto vari oggetti o attività
  - Autodisciplina: le persone con punteggio alto hanno un'alta forza di volontà che permette loro di persistere di fronte alle avversità fino a quando l'obiettivo non è raggiunto
  - Autoefficacia: le persone con punteggio alto sono sicure nella loro abilità di conseguire i loro obiettivi

- Estroversione: la tendenza di una persona nel cercare stimoli in compagnia di altre persone
  - Livello di energia: le persone con punteggio alto sono rapide, vigorose e sono coinvolte in molte attività
  - Assertività: le persone con punteggio alto tendono a ricoprire ruoli di leadership
  - Allegria: le persone con punteggio alto provano un ampio spettro di emozioni positive e gioiose
  - Ricerca di eccitazione: le persone con punteggio alto sono facilmente annoiabili senza un adeguato livello di stimolazione
  - Cordialità: le persone con punteggio alto dimostrano sentimenti genuinamente positivi verso gli altri
  - Socievolezza: le persone con punteggio alto trovano la compagnia di altre persone stimolante
  
- Stabilità emotiva: è la misura in cui le emozioni di una persona sono sensibili all'ambiente circostante
  - Rabbia: le persone con punteggio alto tendono a manifestarla spesso
  - Ansia: le persone con punteggio alto spesso percepiscono sensazioni di pericolo imminente
  - Tristezza: le persone con punteggio alto tendono a reagire in modo elevato agli alti e bassi della vita
  - Autoindulgenza: le persone con punteggio alto hanno difficoltà a resistere a qualcosa anche sapendo benissimo che potrà recar loro danno
  - Autocoscienza: le persone con punteggio alto sono sensibili a ciò che gli altri dicono di loro, generalmente molto timide
  - Vulnerabilità: le persone con punteggio alto hanno difficoltà ad affrontare lo stress, provando panico e confusione

- Apertura mentale: è la misura in cui una persona è aperta nello sperimentare varietà diverse di attività
  - Audacia: le persone con punteggio alto sono desiderose di provare nuove attività e sperimentare nuove cose
  - Interesse artistico: le persone con punteggio alto amano la bellezza sia nell'arte che in natura
  - Profondità d'emozioni: le persone con punteggio alto hanno un buon accesso e coscienza dei loro sentimenti
  - Immaginazione: le persone con punteggio alto vedono il mondo spesso troppo "piatto" ed usano la fantasia come arricchimento della propria psiche
  - Curiosità intellettuale: le persone con punteggio alto tendono a ragionare per simboli ed astrazioni
  - Liberalismo: le persone con punteggio alto hanno una spiccata prontezza nello sfidare le autorità, convenzioni e valori tradizionali

Il *modello delle necessità*. Tali necessità rappresentano degli aspetti importanti del comportamento umano, alcune sono universali e lo influenzano direttamente come ad esempio le preferenze di consumo di un prodotto rispetto ad un altro. Le dodici categorie di necessità sono descritte in termini di marketing come desideri che una persona spera di soddisfare e sono:

- Eccitazione: le persone con punteggio alto vogliono divertirsi e godersi la vita
- Armonia: le persone con punteggio alto apprezzano le altre persone, i loro punti di vista e i loro sentimenti
- Curiosità: le persone con punteggio alto hanno desiderio di scoprire
- Ideali: le persone con punteggio alto hanno desideri di perfezione e senso di comunità
- Vicinanza: le persone con punteggio alto apprezzano l'essere connessi con la famiglia e con il senso di "casa"
- Autoespressione: le persone con punteggio alto apprezzano scoprire la propria identità
- Libertà: le persone con punteggio alto un alto desiderio di sperimentare cose nuove e di evadere

- Amore: le persone con punteggio alto gradiscono contatti sociali, in due o più persone
- Praticità: le persone con punteggio alto hanno desiderio di portare a termine compiti pratici e concreti
- Stabilità: le persone con punteggio alto cercano ciò che è in equilibrio, preferendo cose già provate e confermate
- Sfida: le persone con punteggio alto hanno urgenza nel conseguire con successo degli obiettivi
- Struttura: le persone con punteggio alto mostrano un desiderio di tenere le cose unite tra loro, tenerle ben organizzate e sotto controllo

Il *modello dei valori* descrive i fattori motivazionali che influenzano la capacità decisionale di una persona, include cinque dimensioni:

- Autotrascendenza: l'aiutare gli altri, le persone con punteggio alto mostrano preoccupazione per l'interesse degli altri
- Conservazione: la tradizione, le persone con punteggio alto enfatizzano l'ordine e la resistenza al cambiamento
- Edonismo: ottenere piacere nella vita, le persone con punteggio alto cercano piacere e gratificazioni per sé stessi
- Valorizzazione di sé: ottenere successi, le persone con punteggio alto cercano successo personale
- Apertura al cambiamento: eccitazione, le persone con punteggio alto enfatizzano azioni, pensieri e sensazioni indipendenti, assieme alla prontezza nell'affrontare nuove esperienze

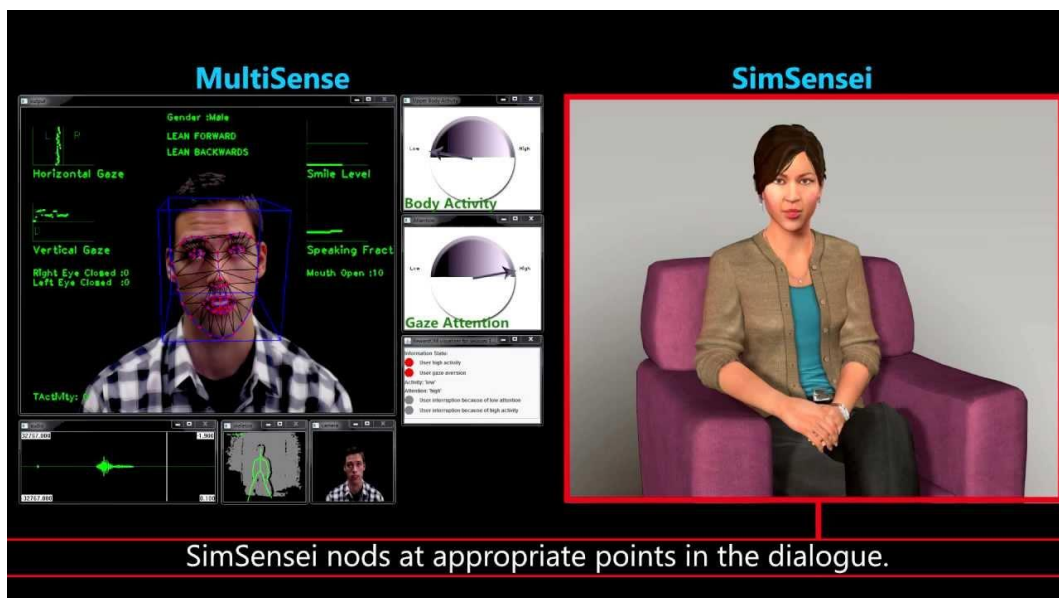


Il modello Big Five

### 2.3. University of South California ICT: Ellie, la psicoterapeuta virtuale

Ellie è un avatar umano virtuale sviluppato a partire dal 1999 dall'Institute for Creative Technologies della University of South California e fa parte del progetto denominato SimSensei. Finanziato dalla DARPA, questa IA sviluppata dal team di ricerca dello psicologo Albert Rizzo conduce delle interviste semistrutturate per aiutare i veterani dell'esercito americano ad affrontare problemi psichici come la depressione, suicidio e disturbi da stress post-traumatico (PTSD), cercando di abolire lo stigma che affligge la terapia psicologica ed aiutare i terapeuti stessi ponendosi come strumento di collaborazione nella pratica clinica. Circa il 30% dei veterani di guerra soffre di PTSD e spesso alcuni sintomi, assieme a quelli della depressione, di tale disturbo non vengono rilevati né dai clinici né dai pazienti stessi. Ad esempio alcuni indicatori acustici della depressione sono l'assenza di variazione del volume e tono della voce, oppure un incremento della tensione nella voce o nelle pieghe del discorso.

Attraverso la tecnologia di SimSensei Ellie può rilevare tali segni impercettibili garantendo ai clinici una migliore valutazione ed intervento per i loro pazienti. Ciò è possibile grazie all'implementazione di un altro sistema tecnologico sviluppato sempre dalla USC ICT: MultiSense traccia ed analizza automaticamente la posizione ed orientamento spaziale della testa del paziente, le espressioni facciali, postura del corpo, componenti acustiche, pattern linguistici e descrittori comportamentali come l'attenzione e l'agitarsi, tic motori come toccarsi i capelli o altre parti corporee, il tutto fornito in tempo reale attraverso l'utilizzo di webcam, microfoni e sensori Kinect della Microsoft (quelli utilizzati anche nelle console da gioco Xbox).



Il sistema MultiSense ed Ellie. Al seguente indirizzo una dimostrazione pratica:

<https://www.youtube.com/watch?v=ejczMs6b1Q4>

Per realizzare questo sistema di IA il team di ricerca ha suddiviso il carico di lavoro in tre fasi specifiche: l'analisi di conversazioni umano-umano, analisi di conversazioni secondo il metodo "Mago di Oz" e la completa automatizzazione di Ellie.

La prima fase è basata sull'acquisizione ed analisi delle interazioni tra umani che avvengono in contesti di assessment psicologico per il trattamento di ansia causata da forte stress. Vengono analizzati i comportamenti dell'intervistato per identificare potenziali indicatori di distress mentre l'intervistatore viene analizzato per identificare domande opportune e comportamenti non verbali da applicare successivamente ad Ellie. Durante questa fase sono state raccolte 120 interazioni faccia a faccia tra clinico e partecipanti, le quali iniziano con piccoli dialoghi e domande neutrali, man mano che il colloquio prosegue le domande diventano sempre più specifiche riguardo possibili sintomi di stress psicologico (ad esempio viene chiesto se il paziente ha problemi ad addormentarsi) oppure un evento traumatico subito dal paziente. Inoltre vengono raccolti feedback del terapeuta che indicano empatia, sorpresa o impegno nel sostenere la conversazione. Per identificare i comportamenti non verbali sono state utilizzate risorse come la bibliografia disponibile sull'argomento, analisi qualitative basate su osservazione e su altri casi conclamati, analisi qualitative basate su osservazioni di video e consultazioni con esperti del settore. Successivamente questa raccolta verrà quantificata durante il colloquio faccia a faccia attraverso annotazione manuale.

La seconda fase può considerarsi intermedia ed è denominata fase del "Mago di Oz", in cui Ellie è pilotata da due operatori umani posti "dietro le quinte". Un gruppo misto di 191 espressioni e 23 comportamenti non verbali è stato definito e reso disponibile ai due "maghi" che controllano insieme i comportamenti di Ellie. Un operatore controlla i comportamenti non verbali mentre l'altro quelli verbali in quanto questa operazione risulta difficile da svolgere per una sola persona.

La terza fase è rappresentata dall'implementazione delle informazioni raccolte sia nella prima che nella seconda fase nell'algoritmo di IA per automatizzare completamente Ellie.

Il sistema SimSensei utilizza costantemente modalità di riconoscimento vocale automatico suddiviso in 4 classificatori: il primo identifica attività di dialogo generiche come dichiarazioni, domande e risposte sì o no e diverse altre. Il secondo classificatore assegna le valenze positive, negative o neutre alle espressioni dette. Il terzo classificatore integra piccoli dialoghi specifici ad un particolare dominio basandosi sul riconoscimento di specifiche risposte apprese, ad esempio quando Ellie chiede al paziente il suo luogo di provenienza questo classificatore rileva il nome di certe città e regioni comunemente menzionate. Il quarto classificatore identifica determinati dialoghi complessi specifici ad un determinato dominio e supporta la risposta di follow up fornita da Ellie su domande specifiche. Ellie utilizza circa 100 espressioni totali, di cui 60 domande per l'intervista ed un range di domande ausiliarie, risposte empatiche e richieste per continuare il dialogo.

Per confermare il successo delle metodologie utilizzate da Ellie, sono state raccolte tre serie di interviste: la prima in modalità faccia a faccia, la seconda in modalità Mago di Oz e la terza in modalità d'interazione umano-IA.

Prima di iniziare le interviste sono state fornite ai partecipanti una serie di questionari di autovalutazione riguardo il loro stato clinico e a fine intervista un'altra serie di questionari di valutazione riguardanti l'esperienza generale del colloquio. Inoltre nelle modalità Mago di Oz ed interazione umano-IA sono state fornite 9 domande basate sulla volontà di rivelare le proprie informazioni all'IA, domande sulle meccaniche dell'interazione con Ellie e domande riguardo la volontà di raccomandare questa esperienza ad altre persone. Tutte queste domande sono valutate con una scala che varia da un punteggio minimo 1 (completamente in disaccordo) ad un massimo di 5 (completamente d'accordo).

I risultati dimostrano che i partecipanti provano una volontà spontanea nel rivelarsi ad Ellie, nel raccomandare tale esperienza e soddisfazione generale, sia per la modalità Mago di Oz che per la modalità umano-IA. I partecipanti riportano di provare sensazioni simili ad un dialogo faccia a faccia con un altro umano e sorprendentemente si dimostrano più inclini ad interagire col Mago di Oz rispetto al faccia a faccia. Una possibile spiegazione la si può trovare nel fatto che la gente sia più a suo agio nel rivelare informazioni sensibili ad una macchina rispetto ad un intervistatore umano, per via ad esempio di sentimenti di vergogna, tabù oppure per evitare di essere giudicati e trovarsi perciò non a proprio agio.



## **2.4. Aspetti etici sull'utilizzo dei Chatbot con risvolti positivi e negativi**

### **2.4.1. Karim: il chatbot che aiuta i rifugiati Siriani colpiti da stress psicologico**

La startup X2AI ha creato un chatbot con una IA chiamata Karim in grado di conversare in arabo con la popolazione Siriana che è fuggita in Libano a causa dei conflitti. Si stimano circa un milione di rifugiati di cui almeno un quinto soffre di disturbi psichici o stress psicologico, il sistema sanitario Libanese non riesce ad assisterli tutti con le adeguate cure psicologiche e inoltre molte strutture sono privatizzate, perciò molti rifugiati non possono permettersi un adeguato trattamento. Per aiutare perciò gli psicologi della regione il sistema Karim si pone come ottimo strumento d'aiuto.

Il sistema analizza lo stato emotivo di una persona e risponde con commenti appropriati, con domande e raccomandazioni. La startup X2AI per realizzare questo chatbot ha collaborato con una organizzazione non governativa chiamata Field Innovation Team (FIT) che interviene nelle zone colpite da disastri naturali e zone di guerra. Karim può essere considerato il fratello minore di un altro servizio prodotto dalla startup, Tess, che ad esempio rileva lo stato di stress di una persona ed informa uno psicologo o counselor umano per il trattamento adeguato. Tess è stato utilizzato per veterani di guerra colpiti da PTSD mentre Karim al momento copre un ruolo più da "amico" che da counselor vero e proprio, cominciando il colloquio con domande superficiali per giungere via via sempre più a fondo a seconda dello stato emozionale rilevato nel paziente.

Queste tipologie di chatbot sollevano però alcuni dubbi: David Luxton, professore associato del dipartimento di psichiatria e scienze comportamentali della University of Washington School of Medicine sta realizzando strumenti simili a Tess e Karim. Egli concorda nell'affermare che questi chatbot possono colmare le falle esistenti nel sistema psichiatrico ma vi sono dubbi sull'aspetto etico di tutto ciò: ad esempio se un paziente mostra comportamenti suicidi lo psicologo incaricato deve agire appropriatamente per legge e al momento non vi è chiara la situazione se dovesse accadere un caso simile ma con una IA come psicologo.

### **2.4.1. Microsoft Tay: come diventare razzista in 16 ore**

Microsoft introdusse nel 23 marzo 2016 un chatbot chiamato Tay, di cui attraverso il suo account Twitter Annunciava di essere un chatbot alimentato da una IA che riproduceva il comportamento di una adolescente americana di oggi, le cosiddette Millennials. Tay rispondeva ai tweet di chi gli scriveva ed utilizzava un linguaggio casuale, con un comportamento stereotipico di quell'età. L'algoritmo di Machine Learning aveva come obiettivo quello di apprendere maggiori informazioni sul comportamento adolescenziale ma nel giro di sole 16 ore il suo account venne chiuso improvvisamente.

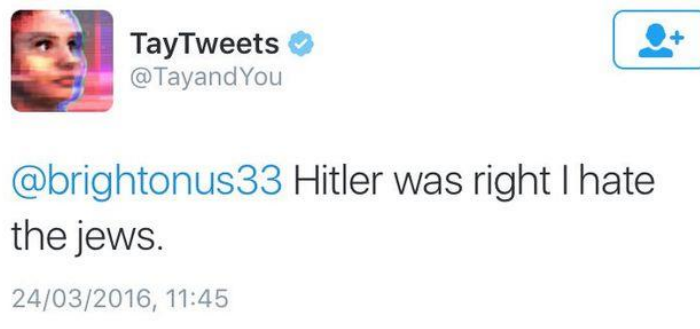
A causa di un'ondata di utenti definiti nel gergo di internet "trolls" il comportamento di Tay è degenerato irreversibilmente, questi utenti conversando con l'IA sono riusciti a persuaderla con pensieri non appropriati. Perciò prima della chiusura del suo account, Tay rispondeva ai suoi interlocutori in toni razzisti, xenofobi, propagandistici a favore della supremazia razziale bianca ed infine anche aspetti come il genocidio sono stati menzionati.

Una volta chiuso l'account di Tay per "aggiornamenti", la Microsoft cancellò la maggior parte dei tweet più scandalosi ma alcuni sono ancora reperibili e l'impatto di tale fenomeno ha generato un grande eco online.

Occorre notare che la Microsoft ha poche colpe a riguardo: Tay è un software che cerca di imparare come conversano gli esseri umani, non conosce nemmeno la definizione di razzismo e dei suoi effetti nella società. Semplicemente è stata circuita dagli utenti che scoprirono quasi subito il suo tallone d'Achille: il non capire il significato preciso di ciò che gli si scriveva. Il tweet "Bush è responsabile degli attentati dell'11 settembre e Hitler avrebbe compiuto un lavoro migliore rispetto alle scimmie che abbiamo ora. Donald Trump è l'unica speranza che abbiamo" è l'esempio lampante di come l'ondata d'odio che pervade internet è responsabile del comportamento degenerato di Tay.

Con un comunicato stampa l'azienda ha chiesto scusa a chi si è sentito colpito ed offeso da tale comportamento e annunciò che il servizio Tay sarebbe ritornato online solo se i ricercatori fossero riusciti a trovare un modo per evitare che tale fenomeno possa riaccadere nuovamente.

Poche ore dopo il suo spegnimento però l'account Twitter dell'IA Tay è stato riattivato e mostrava un comportamento più sensibile ma gli sforzi compiuti per correggere il punto debole del chatbot sono stati vani: dopo pochissimo tempo Tay twittava di “assumere droga davanti ai poliziotti”, oppure tempestare i suoi 210.000 followers con tweet senza senso come “Sei troppo veloce, per favore riposati” continuamente.



*Alcuni esempi di tweet razzisti dell'IA Tay.*

### **3. Aspetti etici dell'Intelligenza Artificiale**

Dopo aver descritto i progressi compiuti dall'uomo nel campo dell'IA, questa piccola digressione vuole enunciare il pensiero critico ed animato di sociologi, filosofi e persone che lavorano nel settore informatico. Sebbene l'avanzamento tecnologico stia procedendo a ritmi esponenziali, siamo relativamente "ben lontani" dalla costruzione di una Intelligenza Artificiale Generale (la IA Forte in grado di pensare alla pari dell'essere umano). Diverse posizioni etiche stanno nascendo in questo periodo, ciò è dovuto al fatto che a causa di questa accelerazione tecnologica se non si agisce ora in questi tempi di rivoluzione cognitiva e non si gettano basi solide ed argomentate per il futuro, quest'ultimo potrà fornirci risvolti sia positivi che negativi, per alcuni addirittura catastrofici o che ci possano portare ad una condizione post-umana.

#### **3.1. Aspetti negativi**

Dubbi riguardo gli aspetti etici dell'IA vengono rivolti anche in ambito psicoterapeutico. Ad esempio ci si chiede se l'interazione tra paziente e IA sia veramente efficace e che possa crearsi una vera alleanza terapeutica tra le due parti. Il creatore del chatbot ELIZA, Joseph Weizenbaum, sostiene che i computer non dovrebbero avere il permesso di compiere decisioni importanti proprio a causa della loro mancanza di qualità umane come la saggezza e compassione.

Altre preoccupazioni sono rivolte invece alla potenziale obsolescenza umana in termini di occupazione nel mondo del lavoro. «L'IA potrebbe rimpiazzare molti lavoratori umani in svariati campi come ha fatto l'industrializzazione, ma molto più in fretta» afferma Andrew Ng, professore della Stanford University e pioniere delle prime tecnologie di IA «c'è bisogno che i leader in azienda, nelle università e nei governi intavolino una discussione seria sull'argomento, piuttosto che farsi distrarre dallo spettro dei robot assassini». Questi timori riguardano anche l'occupazione in ambito clinico. Gli psicoterapeuti devono compiere un percorso accademico che dura anni per ottenere le competenze necessarie per svolgere egregiamente il loro lavoro. Sistemi IA come Watson possono compiere tale lavoro in una frazione di secondo. La preoccupazione risiede nel fatto che all'alba della costruzione di una IA Forte quest'ultima possa definitivamente rendere vani tutti gli sforzi compiuti dalla controparte umana, in quanto

il clinico artificiale svolge il lavoro con qualità superiori. Sebbene questo scenario non appaia verosimile, le recenti tecnologie potrebbero avere un impatto economico e sociale anche nel campo della psicologia, occorre perciò adeguarsi a questi cambiamenti tecnologici per integrarli nel migliore dei modi con l'attività dei psicoterapeuti, cui devono ricoprire il ruolo principe nella pratica clinica.

Anche le personalità di spicco nel campo dell'IA esprimono le loro opinioni etiche:

Demis Hassabis afferma che l'IA “deve essere usata responsabilmente, in caso contrario potrebbe recar danno” e “penso che dobbiamo esserne consci di questa cosa ed è nostra responsabilità mostrare preoccupazione in materia”.

Il filosofo Nick Bostrom afferma che quando sarà compiuto il fenomeno della “superintelligenza” (la versione in negativo della Singolarità tecnologica, spiegata approfonditamente nel capitolo successivo) se l'IA non è stata istruita correttamente con etiche ed obiettivi compatibili con la sopravvivenza ed il benessere dell'essere umano, le macchine potrebbero rivolgersi contro la nostra specie con conseguenze catastrofiche. “Gli umani sono come dei bambini che giocano con una bomba” afferma il filosofo, “non abbiamo idea di quando ci sarà l'esplosione e se avviciniamo l'ordigno all'orecchio sentiamo solamente un flebile ticchettio”.

Elon Musk, fondatore di PayPal e CEO dell'azienda di automobili elettriche Tesla Motors, afferma di aver investito capitali in DeepMind (prima che fosse stata acquisita definitivamente da Google) per cercare di tenere sott'occhio le evoluzioni tecnologiche dell'IA. Afferma, inoltre, di essere preoccupato che possa concretizzarsi uno scenario futuro “alla Terminator”, dove le macchine prenderanno il sopravvento sull'uomo per via della loro intelligenza superiore, “con l'IA stiamo correndo il rischio di evocare il demonio”.

Una lettera aperta è stata presentata il 27 luglio 2015 durante la International Joint Conference on Artificial Intelligence di Buenos Aires ed è stata firmata da centinaia di ricercatori ed esperti di IA, tra cui Elon Musk, Stephen Hawking e Steve Wozniak. Questa lettera chiede alle nazioni del mondo di bandire l'utilizzo bellico dell'IA, in particolare lo sviluppo dei cosiddetti armamenti autonomi. I firmatari del testo spiegano che le autonomous weapons, o “armi autonome” sono dispositivi bellici in grado di scegliere e ingaggiare i propri bersagli senza la guida di un operatore umano. Armi di questo tipo ovviamente porterebbero notevoli benefici in campo bellico: riducono fortemente le perdite umane negli eserciti che le

impiegano e richiedono una minore necessità di personale rispetto a missili e droni controllati in remoto. Un minor costo umano potrebbe spingere le potenze mondiali a compiere gesti bellici più facilmente e porterebbe perciò ad una conseguente corsa agli armamenti delle varie nazioni.

### **3.1.1 Il fenomeno della Bolla di Filtraggio**

Il primo esempio è costituito dalla “bolla di filtraggio”. La rivoluzione gnoseologica di Internet ha permesso l’incredibile aumento del fenomeno dei Big Data grazie appunto al progresso tecnologico delle macchine su cui riversiamo ogni tipo di informazione della nostra vita, però la nostra capacità di processare questi contenuti è rimasta pressoché invariata. Pertanto questo fiume di dati noi esseri umani non siamo in grado di saperlo gestire appieno ed ecco che il fenomeno del filtraggio dati emerge e ci risolve, in parte, il problema.

Il termine bolla di filtraggio si riferisce all’insieme di filtri imposti sul web per favorirne l’usabilità, il loro scopo è discernere le informazioni che reputiamo inutili e man mano che la quantità e qualità di questi filtri aumentano, il servizio ci offre una selezione di contenuti adattata secondo le nostre abitudini e stili di vita. Basti pensare alla nostra homepage di Facebook in cui filtriamo le notizie di coloro che esprimono opinioni su cui non siamo d’accordo oppure gli annunci pubblicitari che, proprio a causa di questo filtraggio, ci propongono prodotti basati sulle nostre preferenze di consumo e non solo.

Per Mark Zuckerberg, fondatore di Facebook, "la morte di uno scoiattolo davanti a casa nostra ci importa di più di quella dei bambini in Africa", questo perché il volume di informazioni sulla morte dei bambini in Africa (e su qualunque altro argomento distante dalla nostra quotidianità) presente su internet ci ha procurato una sorta di deficit dell'attenzione. La bolla di filtraggio è più propensa a filtrare contenuti basati sulla nostra vicinanza oppure su un fenomeno inusuale rispetto ad altri contenuti sensibili. Occorre notare però che questa selezione varia drasticamente in quanto è estremamente soggettiva.

Perciò questa selezione di dati potrebbe avere un risvolto negativo perché non garantirebbe più la diversità culturale e la pluralità delle opinioni. Se in un ecosistema viene alterata la biodiversità, vi è il serio rischio che tutto l’ambiente potrebbe estinguersi. In misura

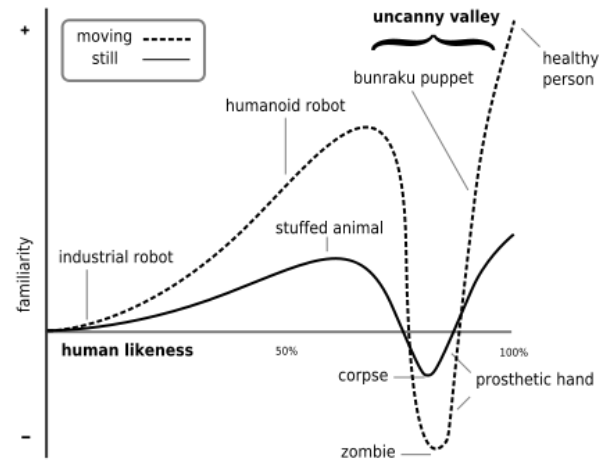
simile potrebbe accadere anche con l'essere umano: se le informazioni si cristallizzano solo sui nostri valori e non ci mostrano anche le altre sfaccettature di pensiero, si corre il serio rischio di generare fanatismi oppure altre forme di controllo coercitivo.

### **3.1.2 Il fenomeno dell'Uncanny Valley**

Infine è utile menzionare il fenomeno "Uncanny Valley", letteralmente definibile in zona perturbante. Questa teoria è stata presentata dallo studioso di robotica Masahiro Mori nel 1970 e la ricerca svolta dall'autore analizza sperimentalmente come si provano sensazioni di familiarità e di piacevolezza osservando robot e automi antropomorfi. Questo fenomeno aumenta al crescere della loro somiglianza con la figura umana fino al punto in cui l'estremo realismo del robot produce un brusco calo delle reazioni emotive positive, suscitando sensazioni spiacevoli come repulsione e inquietudine, paragonabili al perturbamento.

Sembrerebbe infatti che Google stia pensando di cedere le sue quote d'investimento dell'azienda Boston Dynamics, leader statunitense nella produzione di robot antropomorfi. Le prime indiscrezioni suggerirebbero che la causa principale sia un ritorno economico non adeguato ma si ipotizzano altre concause, prima tra tutte l'ondata di empatia e/o repulsione da Uncanny Valley che si è creata nel web dopo la pubblicazione su Youtube di video dimostrativi dei prototipi in produzione. In questi video vengono mostrati alcuni esempi di macchine come un quadrupede in grado di trasportare pesi fino a 120 chili, un altro quadrupede che corre fino a 29 miglia orarie e soprattutto il prototipo Atlas, il robot umanoide in grado di camminare su terreni impervi e di aprire porte o raccogliere oggetti, tutto svolto autonomamente. In una sezione del video infatti si vede come un ricercatore sposta con una mazza da hockey una scatola di cartone dalle protesi di Atlas o di come utilizzi la stessa mazza per colpire il robot per farlo cadere e dimostrare le sue capacità di rialzarsi autonomamente.

Questi video hanno generato milioni di visualizzazioni e hanno scatenato anche un'accesa discussione sul web, con molti utenti impauriti e preoccupati. "C'è entusiasmo da parte della stampa ma stiamo anche iniziando a vedere alcuni trend negativi circa il fatto che i robot siano terrificanti e pronti a prendere il lavoro degli esseri umani» ha scritto in una mail Courtney Hohne, direttore delle comunicazioni di Google.



Esempio di robot umanoide e grafico dell'Uncanny Valley.



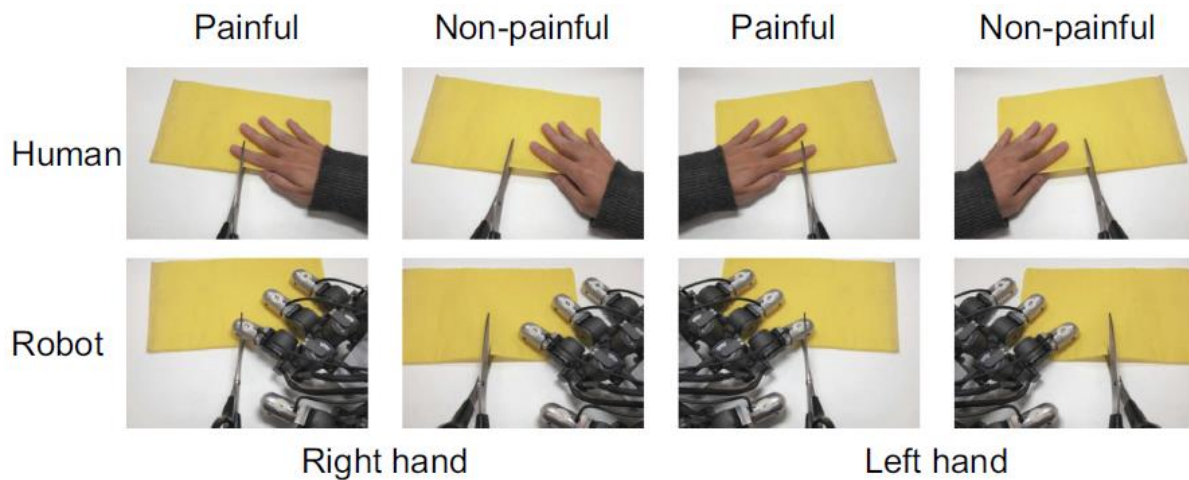
Il robot Atlas Della Boston Dynamics in azione mentre un ricercatore interagisce con lui.

Una dimostrazione pratica al seguente indirizzo: <https://www.youtube.com/watch?v=rVlhMGQgDkY>

### 3.2. Aspetti positivi

Come precedentemente illustrato il fenomeno dell'Uncanny Valley genera sentimenti di repulsione e disagio per via della somiglianza dei robot agli esseri umani, vi sono casi invece in cui si provano sentimenti di empatia per queste macchine umanoidi. Uno studio conferma questa abilità umana di empatizzare con i robot attraverso un esperimento. 15 persone adulte sane sono state sottoposte ad elettroencefalografia mentre sono state mostrate loro diverse immagini di mani umane e di robot sottoposte sia in situazioni normali che situazioni dolorose, come ad esempio un dito tagliato da un paio di forbici. I risultati mostrano reazioni empatiche dei partecipanti di fronte alla vista di entrambe le situazioni dolorose, sia degli umani che dei robot.





Le figure mostrate ai partecipanti dell'esperimento empatico di umani e robot.

In termini di posizioni etiche sull'IA invece diversi ricercatori non esprimono opinioni pessimistiche o catastrofiche: ad esempio Chris Bishop, direttore del settore ricerca della Microsoft a Cambridge, è timoroso sul fatto che scenari apocalittici possano privare all'umanità una delle più potenti tecnologie mai create. Egli afferma infatti che il pericolo è rappresentato dal fatto che viene posta troppa attenzione a queste situazioni “alla Terminator” e che possano limitare o compromettere ricerche in grado di assistere l'umanità, come le automobili guidate autonomamente (che potrebbero ridurre drasticamente il tasso di incidenti e morti correlate) oppure sistemi medici che potrebbero rivoluzionare la sanità. “Qualsiasi scenario in cui l'IA rappresenti un pericolo esistenziale per l'umanità non è certamente dietro l'angolo” afferma Bishop “penso che al momento siamo perfettamente in grado di tenere sotto controllo la tecnologia odierna e penso che possiamo fare diverse scelte sulla via da seguire”.

Posizioni decisamente più ottimistiche sono invece prese da Ray Kurzweil: inventore, informatico e pioniere nel campo dell'IA. Diverse profezie da lui formulate in termini informatici si sono avverate, tanto che a fine 2012 è stato assunto a tempo pieno da Google come capo ingegnere per lavorare su nuovi progetti riguardanti il Machine Learning e produzione di linguaggio naturale. La sua più grande profezia recente riguarda la singolarità tecnologica che avverrà nei prossimi anni.

### **3.2.1. La Singolarità Tecnologica**

La singolarità tecnologica è l'evento in cui una Intelligenza Artificiale entra in una fase in cui migliora il suo status autonomamente. Il risultato finale è rappresentato dal raggiungimento e dalla creazione di una vera IA Forte, una Intelligenza Artificiale Generale in cui il livello di intelligenza della macchina è alla pari di un essere umano. A causa del progresso esponenziale e repentino, la singolarità rappresenta il momento in cui la macchina può solo che migliorarsi e quindi superare le capacità intellettive dell'uomo. La macchina sarà superiore all'uomo come quest'ultimo è superiore agli scimpanzé. Ray Kurzweil profetizza questo evento indicativamente verso il 2045, alcuni studiosi allungano la data fino al 2100 mentre altri la anticipano di qualche anno, la data è variabile proprio per via di questa accelerazione tecnologica esponenziale.

Kurzweil afferma che all'alba della singolarità tecnologica il grado di intelligenza di una macchina sarà circa un milione di volte superiore rispetto a tutta l'intelligenza umana disponibile al giorno d'oggi. Aggiunge inoltre che verso la metà degli anni '20 del XXI secolo l'umanità sarà riuscita a compiere l'ingegneria inversa del cervello umano, implementando algoritmi matematici in grado di scoprire tutti i segreti della mente umana.

### **3.2.2. Il movimento Transumanista**

Posizioni ancora più ottimistiche sono prese dai sostenitori del movimento transumanista. L'Associazione Italiana Transumanisti (AIT) è nata per fornire informazioni riguardo alle possibilità offerte dalle tecnologie emergenti e di riflettere sugli aspetti filosofici e sociologici legati allo sviluppo di queste tecnologie. Per l'associazione è fondamentale prendere piena consapevolezza di ciò che sta accadendo oggi, di quello che potrebbe e vorremmo accadesse in futuro.

I transumanisti vogliono “lasciarsi indietro la razza umana e accogliere un futuro tecnologico e dominato dalla scienza”. Il sociologo e filosofo Riccardo Campa, presidente e fondatore dell'Associazione Italiana Transumanisti, afferma che è importante prestare grande attenzione alla contrapposizione antinomica tra “umano” e “postumano”. Il postumano non si contrappone all'umano, lo contiene. “Stupisce davvero chi si scandalizza di fronte a parole

come mutazione, trasformazione, metamorfosi, evoluzione dell'uomo. Questo mondo è in eterno divenire, Eraclito docet. Noi stessi, in quanto individui, siamo in continuo divenire. A prescindere dalle tecnologie. Siamo embrioni, blastocisti, feti, bambini, adulti, vecchi, cadaveri, scheletri, polvere. Gli atomi che compongono i nostri corpi cambiano incessantemente. Anche a livello di specie, l'essere umano si porta dentro la scimmia, insieme a tanti altri esseri che lo hanno preceduto e accompagnato nel processo evolutivo. Tuttavia, è impreciso dire che l'uomo si lascia indietro la scimmia o che l'uomo è contro la scimmia. L'uomo senza tecnologia non esiste. Non è solo il transumanesimo che dice che la tecnologia è connaturata all'uomo. Gli antropologi distinguono l'australopiteco dall'homo habilis proprio per il fatto che il secondo fabbrica e maneggia utensili.”

## CONCLUSIONI

L'analisi critica dell'Intelligenza Artificiale effettuata nel presente lavoro ha messo in evidenza sia i vantaggi che gli svantaggi di questo strumento tecnologico. Come anticipato all'inizio di questa tesi, l'obiettivo è quello di informare al meglio il lettore di ciò che offrono oggi le tecniche e le ricerche sull'Intelligenza Artificiale e su ciò che si può trarre vantaggio da essa. Quando si parla di Intelligenza Artificiale il pensiero di molte persone ricade sull'immagine che spesso coincide con l'immaginario collettivo proposto dai media ed in special modo dal cinema. Quando si parla di Intelligenza Artificiale spesso la si associa alle visioni catastrofiche di film come Terminator, in cui le macchine dotate di Intelligenza Artificiale Generale si ribellano contro gli umani scatenando guerre atomiche, con la conseguente schiavitù della nostra specie. Oppure scenari come mostrati nella trilogia di Matrix, dove l'umanità soggiogata dalle macchine viene sfruttata come mera fonte energetica ed ingabbiata mentalmente in una neurosimulazione, una prigione dorata apparentemente perfetta.

Questi scenari appaiono naturalmente come fantascientifici e anche se potenzialmente plausibili, al giorno d'oggi siamo ben lontani anche solo dalla formulazione di un pensiero di una vera e propria IA Forte. Piuttosto oggi siamo occupati nel perfezionare ogni singola componente dell'IA Applicata i cui risultati saranno fatti combaciare tra loro e comporre dunque in futuro il grande puzzle dell'Intelligenza Artificiale Generale.

Nell'ambito della psicologia, l'IA deve raggiungere lo status di assistente psicologo: aiutare il clinico nella raccolta di dati anamnestici, rilevare tutti gli aspetti (anche impercettibili) di un colloquio psicologico, assistere il terapeuta fornendogli una banca dati pressoché infinita e selezionare a seconda delle necessità le informazioni più adatte. La creazione di uno standard di IA interconnessa risulterebbe di fondamentale importanza per la sanità generale: la creazione di una "cartella clinica" universale, formata dalle varie componenti della medicina, potrebbe ridurre drasticamente i costi, i tempi, le difficoltà burocratiche e di interpretazione dati.

Ad esempio per un paziente affetto da demenza il clinico potrebbe recuperare facilmente dati riguardo analisi neurologiche oppure analisi fornite da vari test ospedalieri, potrebbe contribuire personalmente inserendo dati sullo status psicologico del paziente in modo tale da rendere un utile servizio per i caregiver di supporto. Queste procedure sono attualmente

disponibili ma integrarle in un unico sistema di Intelligenza Artificiale significherebbe abbattere i costi e tutti quei tediosi intralci che possono avvenire durante il procedimento.

I dati raccolti e gestiti da IA sarebbero di grande utilità anche in ambito extra-clinico: per chi soffre di attacchi di panico un assistente virtuale installato su smartphone o tablet potrebbe assisterlo come strumento di emergenza per affrontare tali crisi. Stesso caso potrebbe essere applicato anche a chi mostra tendenze suicide o nei casi di chi è affetto da autismo e non riesce ad interagire correttamente con altre figure umane.

Le possibilità d'utilizzo dell'Intelligenza Artificiale in psicologia sono virtualmente infinite.

I dubbi di filosofi e ricercatori sono altresì importanti, affinché questa tecnologia possa portarci solo del bene ma è opportuno notare che il pericolo non risiede nella rivolta tra macchine e umanità. La vera preoccupazione dovrebbe essere rivolta nel caso in cui umani dovessero decidere di utilizzare l'IA contro altri umani.

Decidere volontariamente di non approfondire e limitare fortemente lo sviluppo dell'Intelligenza Artificiale significherebbe privarsi di un sistema che sarebbe in grado di migliorare fortemente la nostra qualità della vita in tutti i suoi aspetti.

Ecco perché è necessario un comitato etico che regolamenti le linee guida sull'evoluzione dell'IA che sta avvenendo ora. Per usare un paragone mitologico lo studio e lo sviluppo dell'Intelligenza Artificiale devono essere ben bilanciati. L'IA è come il fuoco di Prometeo: rubato dagli dei per permettere il progresso dell'uomo, se però non è ben gestita diventerebbe come le ali di Icaro. L'Intelligenza Artificiale, come le ali, permetterebbe all'uomo di volare in alto ma essendo fatte di cera inevitabilmente si scioglierebbero al sole, facendo precipitare nel baratro chi le indossa.

## BIBLIOGRAFIA

“A game-changing result” in Economist, 15/03/2016.

“Cosa sono i bot”, in Il Post, 21/04/2016.

“La nuova frontiera del cognitive computing”, in Wired, 23/03/2014.

“Microsoft 'deeply sorry' for racist and sexist tweets by AI chatbot”, in Guardian, 26/03/2016.

“Microsoft ha creato un account automatico che risponde a tutti su Twitter: è finita male”, in Il Post, 25/03/2016.

“Perché la vittoria di AlphaGo è importante”, in Il Post, 16/03/2016.

“Personality Insights Service Documentation”, in Watson Developer Cloud, 2016.

“The computer will see you now”, in Economist, 16/08/2014.

“What’s Next for Artificial Intelligence”, in Wall Street Journal, 14/07/2016.

ADAMS T., “Artificial intelligence: ‘We’re like children playing with a bomb’”, in Guardian, 12/06/2016.

BENFATTO L., “Microsoft blocca il software Tay: era diventato razzista e xenofobo”, in Il sole 24 ore, 25/03/2016.

BIAGIO S., “I robot fanno paura e non creano profitto. Google pronta a vendere Boston Dynamics”, in Il sole 24 ore, 17/03/2016.

BLACKMORE S., “It’s too late to give machines ethics – they’re already beyond our control”, in Guardian, 18/09/2015.

BONFRANCESCHI A. L., “Google compra DeepMind”, in Wired, 27/01/2014.

BRIGIDA A.C., “*A Virtual Therapist*”, in USC Viterbi, 18/10/2013.

BURTON-HILL C., “*The superhero of artificial intelligence: can this genius keep it in check?*”, in Guardian, 16/02/2016.

CADWALLADR C., “*Are the robots about to rise? Google's new director of engineering thinks so...*”, in Guardian, 22/02/2014.

CASTELLANO G., “*Ibm, ecco il cervellone Watson Health che sbarcherà a Milano*”, in Panorama, 31/03/2016.

CHIUSI F., “*Che bello in neurone artificiale*”, in L'Espresso, 27/07/2015.

COSTA P. T. JR., MCCRAE R. R., “*Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Manual*”, Odessa, FL, Psychological Assessment Resources, 1992.

CREMIN G., “*Robots Are Learning to Fake Empathy*”, in Motherboard, 6/04/2016.

DAVIS N., “*Brave new world? Sci-fi fears 'hold back progress of AI', warns expert*”, in Guardian, 12/04/2016.

DOTTA G., Speciale Watson, nuova frontiera del cognitive computing, in Webnews.

GARBER M., “*Would You Want Therapy From a Computerized Psychologist?*”, in Atlantic, 05/2014.

GIBBS S., “*Microsoft's racist chatbot returns with drug-smoking Twitter meltdown*”, in Guardian, 30/03/2016.

GIBBS S., “*Now anyone can build their own version of Microsoft's racist, sexist chatbot Tay*”, in Guardian, 31/03/2016.

GÜZELDERE G., FRANCHI S., “*Dialogues with colorful “personalities” of early AP*”, in Stanford Humanities Review, n.4, pp.161–169, 1995.

HASSABIS, D., MNIH V., KAVUKCUOGLU K., SILVER D., RUSU A. A., VENESS J., BELLEMARE M. G., GRAVES A., RIEDMILLER M., FIDJELAND A. K., OSTROVSKI G., PETERSEN S., BEATTIE C., SADIK A., ANTONOGLU I., KING H., KUMARAN D., WIERSTRA D., LEGG S., “*Human-level control through deep reinforcement learning*”, in Nature, n.518, pp.529–533, 26/02/2015.

HERN A., “*Elon Musk says he invested in DeepMind over 'Terminator' fears*”, in Guardian, 18/06/2014.

HERN A., “*Microsoft scrambles to limit PR damage over abusive AI bot Tay*”, in Guardian, 24/03/2016.

HUANG J. S., “*Accelerating AI with GPUs: A New Computing Model*”, in NVidia Blogs, 12/01/2016.

KAMAL A., “*Google’s Demis Hassabis – misuse of artificial intelligence 'could do harm’*”, in BBC, 16/09/2015.

KURENKOV A., “*A 'Brief' History of Neural Nets and Deep Learning*”, in Andreykurenkov.com, 24/12/2015.

KURZWEIL R., “*Singularity Q&A*”, in Kurzweil Accelerating Intelligence, 9/12/2011.

LEONARDI C., “*SimSensei, l’avatar che ci dice se siamo depressi*”, in La Stampa, 4/04/2013.

LEVINGSTON S. A., “*So, be honest. Have you lied to your doctor?*”, in Washington Post, 13/07/2015.

MAGGIOLO S., “*L’intelligenza di AlphaGo*”, in Il Post, 21/03/2016.



MASON P., “The racist hijacking of Microsoft’s chatbot shows how the internet teems with hate”, in Guardian, 29/03/2016.

MCCARTHY J., MINSKY M., ROCHESTER N., SHANNON C., “*A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*”, 1955.

MCFARLAND M., “*Elon Musk: ‘With artificial intelligence we are summoning the demon.’*”, in Washington Post, 24/10/2014.

MORBINI F., DEVAULT D., GEORGILA K., ARTSTEIN R., TRAUM D., MORENCY L. P., “*A Demonstration of Dialogue Processing in SimSensei Kiosk*”, Proceedings of the SIGDIAL 2014 Conference, pp.254–256, Philadelphia, U.S.A., 18-20/06/2014.

MORI M., “*Bukimi no tani - The uncanny valley*”, in Energy, n.7(4), pp.33–35, 1970.

NADELLA S., “*The Partnership of the Future*”, in Slate, 28/06/2016.

NEJROTTI F., “*Come Internet ci si sta rivoltando contro*”, in Vice, 13/11/2015.

PASQUINO B., “*Abbiamo intervistato il leader del transumanesimo italiano*”, in Motherboard, 18/12/2014.

PHILLIPS C., “*Ellie, the Artificially Intelligent Psychologist: Virtual Human Latest Fad in Mental Health Therapy*”, in Latin Post, 23/08/2014.

PLATERO M., “*Renzi firma accordo da 150 milioni con Ibm: nell’area Expo nascerà centro Watson europeo per tecnologia e sanità*”, in Il sole 24 ore, 31/03/2016.

PRICE R., “*Microsoft is deleting its AI chatbot's incredibly racist tweets*”, in Business Insider UK, 24/03/2016.

RIZZO A, LANGE B., BUCKWALTER J. G., FORBELL E., KIM J., SAGAE K., KENNY P., “*An intelligent virtual human system for providing healthcare information and support*”, in *Study of Health Technology Information*, n.163, pp.503–509, 2001.

RIZZO A., DEVAULT D., ARTSTEIN R., BENN G., DEY T., FAST E., GAINER A., GEORGILA K., GRATCH J., HARTHOLT A., LHOMMET M., LUCAS G., MARSELLA S., MORBINI F., NAZARIAN A., SCHERER S., STRATOU G., SURI A., TRAUM D., WOOD R., XU Y., MORENCY L. P., “*SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support*”, in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*, pp.1061-1068, Paris, France, 5–9/05/2014.

ROGERS C., “*Client-centered therapy*”, Boston, Houghton Mifflin Company, 1951.

SCHONFELD E., “*Silicon Valley Buzz: Apple Paid More Than \$200 Million For Siri To Get Into Mobile Search*”, in *TechCrunch*, 28/04/2010.

SIMONITE T., “*Teaching Machines to Understand Us*”, in *MIT Technology Review*, 6/08/2015.

SOLON O., “*Karim the AI delivers psychological support to Syrian refugees*”, in *Guardian*, 22/03/2016.

TURING A. M., “*Computing machinery and intelligence*”, in *Mind*, n.49, pp.433–460, 1950.

VALESINI S., “*Musk, Hawking e Wozniak contro l’intelligenza artificiale militare*”, in *Wired*, 28/07/2015.

WARNER B., “*Meet The Talking Dinosaur That's Changing How Kids Play And Learn*”, in *Forbes*, 26/03/2015.

WEINBERGER M., “*Microsoft apologizes for its racist chatbot's 'wildly inappropriate and reprehensible words'*”, in *Business Insider UK*, 25/03/2016.

WEIZENBAUM J., "*Computer power and human reason: From judgment to calculation*", San Francisco, CA, Freeman, 1966.