

**UNIVERSITÀ POLITECNICA DELLE MARCHE**  
**FACOLTÀ DI INGEGNERIA**  
Dipartimento di Ingegneria dell'Informazione  
Corso di Laurea in Ingegneria Informatica e dell'Automazione

---



**TESI DI LAUREA**

**Esperienze con l'Intelligenza Artificiale nel contesto dei chatbot, del  
riconoscimento facciale e dell'analisi di testi**

**Experiences with Artificial Intelligence in the contexts of chatbots,  
facial recognition and text analysis**

Relatore

Prof. Domenico Ursino

Candidata

Margherita Gentili

---

**ANNO ACCADEMICO 2023-2024**

*La scienza ci insegna a non trascurare niente, a non disdegnare gli inizi modesti, in quanto nel piccolo sono sempre presenti i principi del grande, come nel grande è contenuto il piccolo.*

Michael Faraday

## Sommario

L'Intelligenza Artificiale è già parte della vita di moltissimi cittadini di tutto il mondo. In questo momento storico, infatti, le tecnologie, gli strumenti e i servizi che ricorrono a una o più branche di essa sono molteplici. Pensiamo ad una giornata tipo e ad alcune azioni che svolgiamo quotidianamente, come, per esempio, un po' di esercizio con il nostro smartwatch mentre mettiamo la playlist suggerita, o cerchiamo tra i vari programmi che ci appaiono sulle piattaforme streaming, o parliamo con un Chatbot nel momento in cui riscontriamo delle problematiche di qualche tipo. Ognuna di queste azioni è resa possibile dalle tecnologie che rientrano nell'Intelligenza Artificiale. Il risultato è che tantissime persone ne beneficiano quotidianamente, senza rendersene conto, anche se magari, di fronte a una richiesta esplicita, non sarebbero in grado di spiegare cosa sia l'IA. In questa tesi affrontiamo il tema dell'Intelligenza Artificiale e di alcuni suoi ambiti di utilizzo, che in meno di dieci anni hanno profondamente cambiato il modo di vivere delle persone diventando sempre di più parte del nostro quotidiano. In particolare affronteremo il fenomeno dei Chatbot, creandone uno, attraverso la piattaforma Amazon Lex. Successivamente affronteremo il tema dell'analisi facciale, analizzando alcuni volti attraverso Amazon Rekognition. Inoltre analizzeremo vari testi e ne dedurremo gli elementi principali attraverso Amazon Comprehend. Infine, faremo una discussione riguardo le esperienze condotte e il possibile sviluppo futuro di questo campo innovativo. Quindi, in definitiva, questo documento ha l'obiettivo di fornire una visione generale su un ambito importante come l'Intelligenza Artificiale, tecnologia che, con il passare del tempo, diventerà sempre più influente all'interno delle nostre vite.

**Keyword:** Intelligenza Artificiale, Natural Language Processing, Machine Learning, Deep Learning, Chatbot, riconoscimento facciale, analisi dei testi.

<b>Introduzione</b>	<b>1</b>
<b>1 Introduzione all'Intelligenza Artificiale</b>	<b>3</b>
1.1 Definizione . . . . .	3
1.2 Intelligenza Artificiale Forte e Debole . . . . .	4
1.2.1 Intelligenza Artificiale Forte . . . . .	4
1.2.2 Intelligenza Artificiale Debole . . . . .	4
1.3 Storia dell'Intelligenza Artificiale . . . . .	5
1.3.1 I primi anni . . . . .	5
1.3.2 Prime innovazioni . . . . .	6
1.3.3 Il primo inverno dell'IA . . . . .	7
1.3.4 Dal primo inverno ad oggi . . . . .	7
1.3.5 Sviluppi . . . . .	7
1.3.6 Il gioco dell'imitazione . . . . .	9
1.3.7 Test della stanza cinese . . . . .	9
1.4 Rischi dell'IA . . . . .	10
1.4.1 Ambito lavorativo . . . . .	10
1.4.2 Trasparenza degli algoritmi . . . . .	11
1.4.3 Protezione dei dati . . . . .	11
1.4.4 Altre problematiche . . . . .	11
1.5 Regolamentazione . . . . .	12
1.6 Benefici dell'IA . . . . .	12
1.6.1 Ambito finanziario . . . . .	12
1.6.2 Automatizzazione dei processi aziendali . . . . .	12
1.6.3 Impatto ambientale . . . . .	13
1.6.4 IA e sicurezza . . . . .	13
<b>2 Natural Language Processing</b>	<b>14</b>
2.1 Introduzione . . . . .	14
2.2 Brevi cenni storici . . . . .	14
2.3 Cos'è il Natural Language Processing . . . . .	15
2.4 Funzionamento . . . . .	16
2.4.1 Analisi Lessicale . . . . .	16
2.4.2 Analisi grammaticale . . . . .	17
2.4.3 Analisi sintattica o parsificazione . . . . .	17

2.4.4	Analisi semantica . . . . .	18
2.5	Panoramica sulle applicazioni del Natural Language Processing . . . . .	19
2.5.1	Information Retrieval . . . . .	19
2.5.2	Information Extraction . . . . .	19
2.5.3	Question Answering . . . . .	20
2.5.4	Summarization . . . . .	20
2.5.5	Dialogue Processing . . . . .	20
2.6	L'importanza del Natural Language Processing oggi . . . . .	20
<b>3</b>	<b>Esperienze nell'ambito dei Chatbot</b> . . . . .	<b>22</b>
3.1	Introduzione ai Chatbot . . . . .	22
3.2	Storia . . . . .	23
3.3	Cosa sono i Chatbot e le loro funzionalità . . . . .	25
3.4	Vantaggi e limiti nei Chatbot . . . . .	26
3.4.1	Vantaggi . . . . .	26
3.4.2	Limiti . . . . .	26
3.5	Esempi di utilizzo dei Chatbot . . . . .	27
3.5.1	Educazione . . . . .	27
3.5.2	Industria . . . . .	27
3.5.3	Servizio clienti . . . . .	28
3.5.4	Assistenza sanitaria . . . . .	28
3.6	Introduzione a Amazon AWS . . . . .	29
3.7	Amazon Lex . . . . .	29
3.7.1	Che cos'è . . . . .	29
3.7.2	Funzionamento . . . . .	30
3.8	Realizzazione di un chatbot con Amazon Lex . . . . .	31
3.8.1	Creazione . . . . .	31
3.8.2	Gestione delle conversazioni . . . . .	32
3.8.3	Utilizzo . . . . .	35
<b>4</b>	<b>Esperienze nell'ambito del riconoscimento facciale</b> . . . . .	<b>38</b>
4.1	Introduzione al Riconoscimento Facciale . . . . .	38
4.2	Brevi cenni storici . . . . .	39
4.3	Come funziona il riconoscimento facciale . . . . .	39
4.4	Vantaggi . . . . .	41
4.5	Esempi di utilizzo, applicazioni . . . . .	42
4.6	Conseguenze etiche . . . . .	43
4.7	Amazon Rekognition . . . . .	44
4.7.1	Cos'è . . . . .	44
4.8	Test con Amazon Rekognition . . . . .	45
4.8.1	Analisi facciale . . . . .	45
4.8.2	Analisi espressioni facciali e deduzione emozioni . . . . .	49
4.8.3	Confronto facciale . . . . .	51
<b>5</b>	<b>Esperienze nell'ambito dell'analisi del testo</b> . . . . .	<b>53</b>
5.1	Introduzione all'analisi dei testi tramite Intelligenza Artificiale . . . . .	53
5.2	Funzionamento dell'analisi del testo . . . . .	53
5.3	Applicazioni dell'analisi del testo tramite IA . . . . .	55
5.4	Caso di studio: Amazon Comprehend . . . . .	55
5.4.1	Che cos'è . . . . .	55
5.4.2	Benefici nell'uso di Amazon Comprehend . . . . .	56

---

5.4.3	Funzionamento . . . . .	56
5.4.4	Esempi con Amazon Comprehend . . . . .	67
<b>6</b>	<b>Discussione in merito alle esperienze condotte</b>	<b>74</b>
6.1	Discussione in merito all'utilizzo dei Chatbot . . . . .	74
6.2	Discussione in merito al riconoscimento facciale . . . . .	75
6.2.1	Discussione in merito all'analisi dei testi . . . . .	76
	<b>Conclusioni</b>	<b>77</b>
	<b>Bibliografia</b>	<b>79</b>
	<b>Ringraziamenti</b>	<b>81</b>

---

## Elenco delle figure

---

1.1	Deep Blue sconfigge il campione in carica . . . . .	8
1.2	Il Test di Turing . . . . .	9
2.1	Esempio di parsing sintattico . . . . .	18
3.1	Programma Eliza . . . . .	23
3.2	Creazione di un chatbot Amazon Lex . . . . .	32
3.3	Lista degli intenti . . . . .	32
3.4	Slot dell'intent hamburger_menu . . . . .	33
3.5	Tipologia di slot . . . . .	33
3.6	Codice Lambda per la gestione degli errori . . . . .	34
3.7	Esempio di una conversazione con il chatbot (prima parte) . . . . .	35
3.8	Esempio di una conversazione con il chatbot (seconda parte) . . . . .	35
3.9	Schermata per l'aggiunta di un canale dalla console Amazon Lex . . . . .	36
3.10	Schermata per la configurazione di un canale dalla console Amazon Lex . . . . .	37
4.1	Risultato dell'analisi facciale del volto di una donna . . . . .	45
4.2	Risultato dell'analisi facciale del volto di una donna a tre quarti . . . . .	46
4.3	File JSON relativo all'analisi del volto di una donna a tre quarti . . . . .	47
4.4	Parte finale del file JSON relativo all'analisi precedente . . . . .	47
4.5	Landmark . . . . .	48
4.6	Analisi facciale del volto di un uomo . . . . .	49
4.7	Analisi delle emozioni tramite Amazon Rekognition . . . . .	50
4.8	File JSON relativo alle emozioni . . . . .	51
4.9	Confronto facciale . . . . .	51
4.10	Codice del confronto facciale . . . . .	52
5.1	Funzionamento di Amazon Comprehend . . . . .	57
5.2	Schermata iniziale Amazon Comprehend . . . . .	58
5.3	File JSON relativo all'analisi del sentimento . . . . .	59
5.4	Analisi del sentimento mirato . . . . .	60
5.5	File JSON relativo al sentimento mirato . . . . .	61
5.6	Analisi delle entità . . . . .	62
5.7	File JSON relativo all'analisi delle entità . . . . .	62
5.8	Analisi delle frasi chiave . . . . .	63
5.9	File JSON relativo alle frasi chiave . . . . .	64

---

5.10	Analisi dei dati PII . . . . .	65
5.11	File JSON relativo all'analisi dei dati PII . . . . .	65
5.12	File JSON relativo all'analisi della lingua dominante . . . . .	66
5.13	Analisi della sintassi . . . . .	67
5.14	File JSON relativo all'analisi della sintassi . . . . .	67
5.15	Analisi delle entità dell'esempio . . . . .	68
5.16	File JSON relativo all'analisi delle entità dell'esempio . . . . .	68
5.17	Risultati dell'analisi delle frasi chiave . . . . .	69
5.18	File JSON relativo all'analisi delle frasi chiave dell'esempio . . . . .	69
5.19	Risultato dell'analisi della lingua dell'esempio . . . . .	70
5.20	Risultato dell'analisi dei PII . . . . .	70
5.21	File JSON relativo all'analisi dei PII . . . . .	70
5.22	I risultati dell'analisi del sentimento dell'esempio . . . . .	71
5.23	File JSON relativo ai risultati dell'analisi del sentimento dell'esempio . . . . .	71
5.24	I risultati dell'analisi del sentimento mirato . . . . .	72
5.25	File JSON relativo ai risultati dell'analisi del sentimento mirato dell'esempio . . . . .	72
5.26	I risultati dell'analisi della sintassi (prima parte) . . . . .	73
5.27	I risultati dell'analisi della sintassi (seconda parte) . . . . .	73

Per introdurre il concetto di Intelligenza Artificiale ci possiamo rifare al discorso proposto da Roberto Cingolani, amministratore delegato dell'azienda Leonardo, durante il programma andato in onda sulla Rai, "Noos". In questo discorso, Cingolani afferma che la definizione dell'Intelligenza Artificiale è di per sé una definizione semplice, in quanto, prima di tutto, parte da come funziona l'essere umano. Ogni uomo possiede una memoria con tanti ricordi e tante esperienze differenti, che possiamo equiparare alla memoria del computer. Allo stesso modo, in ogni uomo è "innestato" un sistema di sensori, quali occhi, tatto e udito, che accumula dei dati sulla realtà, mentre una parte del cervello di cui siamo dotati analizza tali dati e, collegandoli ai dati presenti nella nostra memoria, li elabora per permetterci di compiere azioni, prendere decisioni o comportarci in un determinato modo. Questo stesso meccanismo viene replicato nelle macchine.

I computer che abbiamo a disposizione al giorno d'oggi sono caratterizzati da grandissime memorie, che eseguono una grande quantità di operazioni al secondo, analizzando in tempi rapidi una mole enorme di dati (che vengono raccolti tramite una rete di sensori).

Tuttavia, nonostante nella teoria l'uomo e la macchina sembrano funzionare in maniera identica, nella realtà ci sono delle differenze importanti. Infatti, l'uomo riesce a "funzionare" con un dispendio energetico molto basso, o, in ogni caso, molto minore rispetto a quello richiesto dai supercomputer, che hanno bisogno di migliaia e migliaia di watt per analizzare i dati. Per contro, queste macchine sono molto più veloci ad analizzare ed elaborare dati, riuscendo a compiere, ad oggi, un miliardo di miliardi di operazioni al secondo.

Avendo dato un'idea di ciò che può essere l'Intelligenza Artificiale e di tutte le sue potenzialità, si può intuire l'importanza di questa disciplina nella nostra società moderna.

In questa tesi verrà proposta un'analisi dell'Intelligenza Artificiale, spiegando la sua storia e il suo funzionamento e approfondendo, in particolare, un suo settore ovvero il Natural Language Processing. Successivamente verranno affrontate delle esperienze in diversi campi dell'Intelligenza Artificiale, che vanno dalla creazione di un Chatbot, all'analisi facciale ed infine all'analisi dei testi.

La presente tesi sarà composta da sette capitoli come di seguito specificato:

- Nel Capitolo 1 si introdurrà il concetto di Intelligenza Artificiale, verrà esposto un excursus sulla sua storia e sui benefici e rischi di questa tecnologia.
- Nel Capitolo 2 verrà introdotta la branca dell'Intelligenza Artificiale chiamata Natural Language Processing. Si farà una panoramica generale sul funzionamento e l'utilità di questa disciplina.

- Nel Capitolo 3 si spiegherà il funzionamento, con degli esempi, del servizio di costruzione dei Chatbot offerto da AWS, ovvero Amazon Lex; a tal fine verrà creato un Chatbot attraverso tale strumento.
- Nel Capitolo 4 si spiegherà il funzionamento, con degli esempi, del servizio Amazon Rekognition, per il riconoscimento facciale, e si tenterà il suo funzionamento attraverso l'analisi di volti umani.
- Nel Capitolo 5 si spiegherà il funzionamento, con degli esempi, del servizio Amazon Comprehend, per l'analisi dei testi.
- Nel Capitolo 6 verranno espresse delle riflessioni sui capitoli precedenti.
- Nel Capitolo 7 verranno tratte le conclusioni e verranno delineati alcuni possibili sviluppi futuri.

---

## Introduzione all'Intelligenza Artificiale

---

*In questo capitolo si introdurranno il concetto di Intelligenza Artificiale, le varie tipologie che la compongono, e vedremo se è possibile arrivare al concetto di macchina pensante. Inoltre si farà un accenno alla storia di questa tematica innovativa e ai benefici e alle problematiche che ne derivano.*

### 1.1 Definizione

L'Intelligenza Artificiale, abbreviata con IA, è una disciplina, sia scientifica che ingegneristica, innovativa a rapido sviluppo che si pone un obiettivo molto ambizioso, che è quello di realizzare macchine "pensanti", ovvero macchine e programmi che possano risolvere, in modo del tutto autonomo, problemi di varia natura con il ragionamento. Lo scopo ultimo dell'Intelligenza Artificiale, infatti, consiste nella progettazione e nell'elaborazione di sistemi in grado di eseguire o simulare compiti che, in genere, sono molto complessi e che sono attuabili solo tramite l'intelligenza umana, ma non solo; infatti, si cerca di progettare un sistema che possa permettere alle macchine di modellare o migliorare le capacità della mente umana. L'analisi dell'Intelligenza Artificiale parte da una domanda fondamentale: "le macchine possono pensare?" La risposta a tale domanda non può prescindere dalla definizione stessa di intelligenza. Cosa c'è alla base del concetto di Intelligenza Artificiale? La volontà dell'uomo di creare un legame indissolubile tra automazione e ragionamento. I primi tentativi in merito si basavano sulla presenza di algoritmi matematici che si sono gradualmente evoluti, diventando sempre più complessi e difficili. Quando parliamo di Intelligenza Artificiale, infatti, trattiamo proprio di innovazioni come lo sviluppo di auto a guida autonoma, oppure come il sempre più diffuso utilizzo di software che simulano conversazioni umane (come Chat GPT).

In generale, i sistemi artificialmente intelligenti possono eseguire compiti comunemente associati alle funzioni cognitive umane, come interpretare il discorso e giocare, e imparano come farlo elaborando enormi quantità di dati, ricercando degli esempi da poter utilizzare come modello nel proprio processo decisionale. In molti casi gli esseri umani supervisionano il processo di apprendimento di un'IA. Rafforzando le buone decisioni e scoraggiando quelle cattive. Tuttavia, vi sono anche alcuni sistemi di Intelligenza Artificiale che sono progettati per imparare senza supervisione, ad esempio giocando a un determinato gioco più e più volte fino a quando non riescono a comprendere le regole e come ottenere la vittoria.

Date tali premesse, è comprensibile capire come mai l'Intelligenza Artificiale, oggi, stia diventando sempre più parte della vita di tutti i giorni e un'area in cui le aziende di ogni settore stanno investendo.

## 1.2 Intelligenza Artificiale Forte e Debole

Nel 1980 il filosofo John Searle pubblicò il suo lavoro "Menti, Cervelli e programmi", introducendo una differenza sostanziale tra due categorie di Intelligenze Artificiali, o meglio due concezioni della stessa, ovvero:

- Intelligenza Artificiale Debole;
- Intelligenza Artificiale Forte.

Questa tecnologia si è, negli anni, evoluta e sviluppata e ad oggi è possibile fare una netta distinzione tra le due categorie. Tale distinzione è utile perché ci permette di classificare le macchine in base al livello di autonomia e complessità effettivamente raggiunto. Infatti, esiste un'intelligenza che vuole provare a raggiungere l'intelligenza umana, per cui alcune macchine sono progettate per raggiungere un livello di coscienza equiparabile a quello degli esseri umani. Tale concetto viene indicato come "Intelligenza Artificiale Forte". Esistono poi altri tipi di Intelligenza Artificiale, ovvero "l'Intelligenza Artificiale Debole", che rappresentano la possibilità di alcuni algoritmi di manifestare un certo tipo di intelligenza non volendo, però, raggiungere quella umana, per cui tale tipo di Intelligenza Artificiale non riesce a prendere coscienza di sé e non è in grado di auto-apprendere e ha sempre necessità di supporto dell'essere umano. Tale intelligenza rappresenta tutto ciò che in questo momento si è riusciti a raggiungere e che viene ad oggi definita come "Intelligenza Artificiale".

### 1.2.1 Intelligenza Artificiale Forte

La teoria dell'Intelligenza Artificiale Forte (o Strong AI), supportata dal campo di ricerca denominato Intelligenza Artificiale Generale, che studia sistemi in grado di replicare l'intelligenza umana, sostiene che le macchine siano in grado di sviluppare una coscienza di sé. Grazie a una programmazione sofisticata, le macchine sarebbero in grado di funzionare come se fossero dotate di una vera e propria mente, avvicinandosi al funzionamento dell'essere umano. In tal senso, e se programmata in maniera opportuna, la macchina non sarebbe solo uno strumento ma diventerebbe essa stessa una mente, con una capacità cognitiva non distinguibile da quella umana.

La tecnologia alla base dell'Intelligenza Artificiale forte rientra in quella dei sistemi esperti, cioè una serie di programmi che vogliono riprodurre, attraverso una macchina, le prestazioni e le conoscenze delle persone esperte in un determinato campo. Tale concezione di Intelligenza Artificiale viene anche definita Intelligenza Artificiale Generale, poiché la sua realizzazione non si limita al solo imitare il comportamento umano in determinate situazioni, ma di sviluppare una coscienza autonoma in grado di risultare efficace in qualsiasi situazione. Ad oggi il concetto di strong IA rimane del tutto teorico, non essendo ancora riusciti a realizzare una macchina del tutto autonoma e indipendente dall'uomo. La maggior parte degli esperti ritiene che non si arriverà mai a una macchina che abbia coscienza di sé, semplicemente perché la coscienza nell'uomo è una facoltà distinta del linguaggio.

### 1.2.2 Intelligenza Artificiale Debole

In contrasto con la prima, l'Intelligenza Artificiale Debole ritiene possibile sviluppare macchine in grado di risolvere problemi specifici senza avere coscienza delle attività svolte. In

altre parole, l'obiettivo dell'IA Debole non è realizzare macchine con un grado di intelligenza pari o superiore a quello umano, ma è quello di imitare il comportamento che un essere umano svolgerebbe per l'esecuzione di un compito ben determinato. La macchina, nonostante non sia capace di pensare in maniera autonoma, svolge egregiamente il suo compito superando le effettive capacità dell'uomo. Alcuni esempi di Intelligenza Artificiale Debole sono per esempio:

- l'analisi comportamentale dei clienti, per migliorare la loro conoscenza e dare luogo a campagne di marketing più efficaci, sulla base di contenuti personalizzati;
- filtri anti spam per email;
- sistemi di guida autonoma;
- ottimizzazione dei consumi energetici;
- autocorrettore per messaggi;
- software di riconoscimento immagini;
- assistenti digitali.

## 1.3 Storia dell'Intelligenza Artificiale

### 1.3.1 I primi anni

Il concetto di entità in grado di ragionare come un uomo stuzzica l'immaginario umano da decenni. Già negli anni 50, un'intera generazione di scienziati fantasticava sull'Intelligenza Artificiale. Uno di questi era Alan Turing che si pose una semplice domanda "se gli umani usano la ragione e l'informazione per risolvere problemi e prendere decisioni, perché non possono farlo anche le macchine?". Alan Turing fu, infatti, il primo a prendere in considerazione l'ipotesi che le macchine potessero "pensare". Ma la nascita dei primi calcolatori risale già a molto tempo prima e segna l'inizio dell'interesse verso questa nuova frontiera dell'informatica. Nel 1623, Wilhelm Schickard, scienziato tedesco, costruì la prima macchina calcolatrice in grado di fare, con l'aiuto dell'uomo, calcoli a sei cifre. Successivamente ci si iniziò a focalizzare sulla possibilità di creare macchine in grado di poter effettuare operazioni matematiche.

Risale, però, agli anni '50 l'inizio del concetto di Intelligenza Artificiale, quando, Alan Turing apriva un articolo da lui scritto, pubblicato sulla rivista *Mind Computing Machinery and Intelligence*, con la domanda: "Le macchine possono pensare?". Questo studio divenne noto come test di Turing. Secondo il test, una macchina poteva essere considerata intelligente se il suo comportamento, osservato da un essere umano, fosse considerato indistinguibile da quello di una persona. L'articolo divenne celebre per aver dato inizio al campo di ricerca che solo sei anni dopo prese il nome di Intelligenza Artificiale (questo termine venne poi coniato da Jhon Mccarty, informatico statunitense).

Un punto di svolta si ebbe nel 1956 quando al Dartmouth College (ad Hanover, New Hampshire), viene presentato quello che per molti è il primo programma dell'Intelligenza Artificiale della storia. La conferenza svolta in quella sede può essere ritenuta un importante tassello della storia dello sviluppo dell'Intelligenza Artificiale che ha dato l'avvio per la ricerca nel campo per i successivi anni. Nel documento preparatorio della conferenza, che era stato redatto e fatto circolare nell'anno precedente, quattro autori davano una definizione di IA attraverso l'indicazione del suo principale obiettivo: costruire una macchina che si comporti in un modo che sarebbe considerato intelligente nel caso di un essere umano. Per la

prima volta venne utilizzato il termine "Intelligenza Artificiale". Gli scrittori e scienziati che diedero vita a questa rivoluzione furono John McCarthy, un matematico statunitense, Marvin Minsky, un fisico, Nathaniel Rochester, un ingegnere di IBM<sup>1</sup> e Claude Shannon, un ingegnere e matematico statunitense, spesso definito "il padre della teoria dell'informazione". Il documento discute di alcuni temi che gli organizzatori ritenevano fondamentali per il campo di ricerca, come le reti neurali, la teoria della computabilità, la creatività e l'elaborazione del linguaggio naturale, motivando la necessità della conferenza con l'obiettivo di costruire delle "macchine intelligenti", capaci di simulare ogni aspetto dell'intelligenza umana. Non è ben chiaro a quali risultati portò la conferenza, della quale non è mai stato nemmeno realizzato il rapporto conclusivo; si trattò, infatti, di una situazione disorganizzata nella quale i vari studiosi non riuscirono a mettersi d'accordo per uno standard dello sviluppo. John McCarthy definì questi anni come il periodo del "guidare la bici senza mani", proprio per mettere in evidenza come gli obiettivi che si erano prefissati erano troppo ambiziosi rispetto alla complessità dei compiti che si volevano affrontare.

Ma ormai il primo passo era stato fatto e da lì inizio un forte interesse per questa disciplina che attirò scienziati e finanziatori; ne conseguirono ricerche e innovazioni nel settore.

### 1.3.2 Prime innovazioni

Una delle innovazioni più sorprendenti è stata realizzata dall'informatico statunitense Arthur Samuel che, nel 1959, elaborò il suo "giocatore di dama", un programma ideato affinché si auto-migliorasse fino a superare le abilità del creatore. Tale macchina, per incrementare le sue capacità, attuava qualcosa impossibile per gli esseri umani, ovvero giocare contro sé stesso. Samuel sceglie proprio il gioco della dama poiché le regole sono relativamente semplici, mentre le tattiche da utilizzare sono complesse, permettendo, così, di provare come le macchine, a seguito di istruzioni fornite dai ricercatori, siano in grado di imitare le decisioni umane. Nel 1957 Herbert Simon stimò, che nel giro di dieci anni, la comunità scientifica avrebbe sviluppato un'Intelligenza Artificiale in grado di competere con i campioni di scacchi. Nel 1958 lo psicologo Frank Rosenblatt sviluppa il Perceptron, un programma nato per riconoscere le immagini e che consiste in una rete neurale artificiale. Il Perceptron è in grado di fornire i risultati desiderati aggiustando da solo i propri parametri grazie ad un algoritmo proposto dallo stesso scienziato. Le prestazioni attese sono altissime, ma una singola rete neurale artificiale si dimostra inadeguata a compiere le operazioni logiche richieste: i risultati sono deludenti e il progetto viene abbandonato. E con esso si accantona l'idea stessa di impiegare reti neurali per l'IA.

Altri ricercatori, come Seymour Papert, partendo dall'IA, svilupparono un linguaggio di programmazione per aiutare i bambini a pensare in maniera matematica, il Logo.

In informatica medica, il primo sistema applicato in campo medico è stato DENDRAL, sviluppato da Edward Albert Feigenbaum nell'anno 1965 e scritto in Lisp<sup>2</sup>, che aiutava i chimici a identificare molecole sconosciute. È il primo sistema esperto, cioè che imita il ragionamento umano, in grado di risolvere problemi in un particolare campo. Tra i primi programmi (siamo nel 1965) in grado di superare il test di Turing c'è, probabilmente, ELIZA, un software che simula le risposte di uno psicoterapeuta. L'inventore di questa, che può essere definita, la prima forma di Chatbot si chiama Joseph Weizenbaum. Il software era utilizzato per gestire una conversazione di natura medica; più nel dettaglio, simulava la relazione tra un paziente e il suo psicoterapeuta, le sue risposte al paziente erano sempre

<sup>1</sup>Azienda statunitense del settore informatico, tra le più importanti al mondo.

<sup>2</sup>Il linguaggio LISP fu inventato da John McCarthy nel 1958. È un linguaggio nato per elaborare le espressioni simboliche. È particolarmente utile per lavorare sui dati strutturati ad albero logico.

effettuate con la forma di domanda. Con ELIZA si è ancora molto lontani dall'ottenere una macchina davvero intelligente.

Nel 1972, a Stanford, viene sviluppato MYCIN, un sistema capace di identificare i batteri responsabili di infezioni gravi e di consigliare antibiotici nel dosaggio corretto. Ma la potenza di calcolo disponibile è ancora limitata e il sistema non riesce ad essere sviluppato.

Nel 1979 lo Stanford Cart, primo robot guidato da computer semovente dotato di telecamera, riesce ad attraversare da solo una stanza piena di sedie. Ci mette circa cinque ore.

Da qui in poi, l'Intelligenza Artificiale inizia un ciclo di stagioni nel quale si alternano delusioni e fallimenti ad entusiasmi e successi.

### 1.3.3 Il primo inverno dell'IA

I risultati direttamente "applicabili" dell'Intelligenza Artificiale al mondo reale, però, non arrivarono subito. Venne presto notata, infatti, la limitatezza legata alla potenza di calcolo dei computer a disposizione in quegli anni e gli hardware allora disponibili non erano ancora in grado di ospitare tutta la memoria necessaria a immagazzinare una quantità sufficiente di dati e informazioni. Ci si scontrava, dunque, di fronte all'impossibilità di trasformare algoritmi funzionanti a livello teorico, in programmi capaci di calcolare effettivamente la soluzione proposta. La costruzione di una macchina che fosse intelligente apparve subito un progetto complesso e difficile. Già verso la fine degli anni '60 i ricercatori in ambito IA dovettero ammettere che i tempi non erano ancora sufficientemente maturi per l'adozione in ambito industriale di algoritmi di Intelligenza Artificiale. La ricerca in questo campo passò in sordina, rimanendo legata ad ambiti prettamente accademici. Si arrivò ad un periodo di raffreddamento degli interessi da parte dei finanziatori nei confronti dell'IA e ad un conseguente stallo dal punto di vista della ricerca e dello sviluppo delle tecnologie.

Questa fase di declino venne chiamata "Primo Inverno AI" e durò fino alla prima metà degli anni '80, quando i tempi divennero finalmente maturi affinché l'Intelligenza Artificiale potesse cominciare a essere applicata (sebbene lentamente) in ambito industriale. Le estreme conseguenze di questo periodo si protrarranno fino al 2012.

### 1.3.4 Dal primo inverno ad oggi

Negli anni '80 e '90 i computer registrarono notevoli miglioramenti nella potenza di calcolo; nel frattempo si ebbero a disposizione enormi quantità di dati generati su Internet e dei dispositivi per il loro immagazzinamento; ciò rese disponibile un'intera nuova generazione di tecnologie che permisero un ulteriore sviluppo dell'Intelligenza Artificiale. Inoltre, tali miglioramenti consentirono di sviluppare programmi sempre più complessi nel mondo dell'IA e superare alcune barriere degli anni precedenti. La nuova Intelligenza Artificiale venne applicata per risolvere problemi ancora più complessi, come l'interpretazione del linguaggio naturale, il riconoscimento visivo delle immagini e la rappresentazione generale della realtà.

### 1.3.5 Sviluppi

Alla fine degli anni Ottanta, il ricercatore Yann LeCun sviluppa un algoritmo che sa "vedere" e riconoscere gli assegni bancari, e che, soprattutto, riesce a farlo e ad imparare a farlo in maniera migliore rispetto agli esseri umani: è l'inizio della rivoluzione del Machine Learning.

La vincita di Deep Blue fu un momento storico in quanto il calcolatore divenne la prima macchina a sconfiggere un umano nel gioco degli scacchi. Il progetto Deep Blue nacque dalla

tesi dello studente Feng-Hsiung Hsu e si basava sulla costruzione di una macchina con il solo scopo di specializzarla e renderla “forte” nel gioco degli scacchi. La prima volta che Deep Thought (nome precedente della macchina) sfidò il campione in carica Kasparov, essa perse, facendo vincere il suo sfidante, il quale il giorno dopo scrisse “Non riesco a immaginarmi cosa vorrebbe dire vivere sapendo che un computer è più forte della mente umana, devo sfidare Deep Thought per proteggere la razza umana”. Successivamente gli studenti, con l’aiuto di IBM, fecero evolvere la macchina in Deep Blue e, sette anni dopo, nel 1996, Deep Blue si prese la sua vittoria (figura 1.1). In quell’occasione, quando Kasparov si arrese, gli restavano solo 5 minuti e 42 secondi del tempo a disposizione mentre, Deep Blue, aveva ancora quasi un’ora, ed era, come scrisse il New York Times, «palesamente in controllo del gioco». Nonostante Kasparov riuscì a rovesciare il risultato nelle cinque partite successive, con tre vittorie e due patte, il match del 10 febbraio diede un’enorme visibilità all’Intelligenza Artificiale sui mass media, mostrando al mondo che vi era una possibilità che questa potesse sostituirsi all’uomo nel prendere decisioni e risolvere problemi molto complessi. Si riaccese, così, l’entusiasmo del mondo industriale per questo campo.



**Figura 1.1:** Deep Blue sconfigge il campione in carica

Nello stesso anno, un sistema di riconoscimento della lingua parlata viene implementato e ampliato all’interno di Windows, mettendo un altro importante tassello per consentire l’interazione uomo-macchina. Nel 2009 si ebbe un primo importante punto di svolta quando un gruppo di ricercatori a Stanford introdusse Imagenet, il più grande dataset di immagini mai realizzato. Nello stesso anno, gli scienziati della Northwestern University sviluppano un programma che scrive notizie sportive senza intervento umano. Nel 2012 Google presenta Google Brain, ovvero una rete neurale addestrata a riconoscere un gatto basandosi su 10 milioni di immagini digitali tratte dai video di YouTube.

Se in precedenza i progressi dell’IA sono stati oggettivamente lenti, dal 2012 in poi tutto accelera, e tanto. Tuttavia, in questi anni non cambia molto il modo di codificare l’Intelligenza Artificiale, né sono in corso le costruzioni di macchine più intelligenti; quindi, cosa è cambiato? Si sono spostati i limiti tecnici e grazie alla legge di Moore, che stima il raddoppio della memoria e della velocità nei microprocessori ogni anno, le capacità dei computer sorpassano velocemente le nostre necessità permettendoci di realizzare IA sempre più evolute. Si satura la capacità dell’Intelligenza Artificiale fino a quella massima consentita dall’attuale potenza computazionale, poi si aspetta che la legge di Moore ci superi di nuovo.

Nel 2016, ci fu un altro importante avvenimento, sempre nel mondo dei giochi, il computer AlphaGo di Google DeepMind sconfigge Lee Sedol, campione di Go, il gioco più antico (risale a oltre tremila anni fa) e complicato del mondo, in cui l’intuizione ha un ruolo predominante.

Tutte queste innovazioni e rivoluzioni, seppur piccole, sono state di grandissima importanza ed hanno permesso all’Intelligenza Artificiale oggi di rappresentare uno dei principali campi di interesse della comunità scientifica informatica, con temi di ricerca come il Machine Learning, l’elaborazione del linguaggio naturale e la robotica. Inoltre, le aziende informatiche,

e non solo, stanno investendo sempre di più in questo settore e i progressi tecnologici sono sotto gli occhi di tutti. Come già detto precedentemente, l'Intelligenza Artificiale, ormai, sta permeando sempre di più tutti gli aspetti delle nostre vite, diventando sempre più fondamentale e indispensabile per un futuro sviluppo informatico. È per questi motivi che è definita come il motore principale della Quarta Rivoluzione Industriale, al punto tale che Klaus Schwab (fondatore del World Economic Forum) ha asserito che l'adozione dell'AI nei domini più vari diventerà irrefrenabile entro il 2030.

### 1.3.6 Il gioco dell'imitazione

Il test di Alan Turing basato sul gioco dell'imitazione, è stato usato per decenni per decidere se una macchina era intelligente o meno. In questo gioco si cerca di distinguere due interlocutori in base alle risposte che forniscono a specifiche domande. L'idea è che, in base alle risposte date, si possa distinguere un umano da un computer.

Il gioco prevede tre partecipanti: un giudice (umano), un uomo ed una macchina, posti in locali separati (figura 1.2). Il giudice dialoga, tramite messaggi testuali, con gli altri due soggetti e deve scoprire chi dei due sia l'uomo e chi la macchina. Le regole del gioco prevedono che il computer possa attuare qualsiasi comportamento per non farsi scoprire; per esempio potrebbe sbagliare determinate domande di proposito, oppure ritardare nell'effettuare un calcolo matematico complicato che potrebbe smascherarlo. Nel libro "Intelligenza artificiale. Guida al futuro prossimo" Jerry Kaplan afferma che Turing è stato il primo ad attribuire la capacità di pensiero ai computer. Egli parla, infatti, della simulazione della macchina di Turing e spiega come, con il suo test, non stesse proponendo una modalità per determinare se le macchine fossero intelligenti o meno, ma "egli stava in realtà speculando che il nostro uso comune del termine pensare si sarebbe alla fine esteso fino a essere ampliato, in maniera appropriata, a certe macchine o programmi di capacità adeguata". Ancora oggi il test di Turing è considerato uno dei pilastri dell'Intelligenza Artificiale.

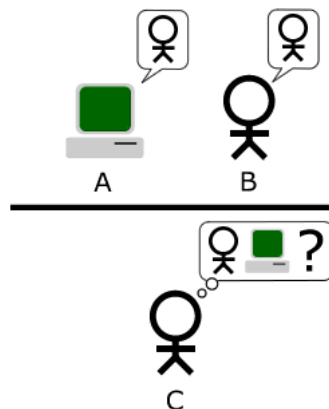


Figura 1.2: Il Test di Turing

### 1.3.7 Test della stanza cinese

Anche se il test di Turing rappresenta un tassello fondamentale per lo studio nel campo dell'Intelligenza Artificiale, non definisce un metodo definitivo per affermare che una macchina sia pensante. Infatti nel tempo si sono sviluppate una serie di controtesi per provare la non totale affidabilità al test di Turing. In particolare la tesi più famosa viene messa a punto dal filosofo John Searle, che nel suo scritto, "Mind, Brain and programs" utilizza un test simile a quello di Turing per mettere in evidenza come non ha senso assimilare la mente ad un

computer, in quanto nessun computer può "pensare" nello stesso modo degli esseri umani. Il suo principale presupposto è che il computer, per elaborare l'informazione, non ha bisogno di comprendere il linguaggio o altri codici simili. Nel test di Turing, come affermato sopra, abbiamo due soggetti una macchina e un umano, ma quest'ultimo non sa se sta interagendo con un'altra persona o con un sistema informatico. Quindi l'uomo attraverso le risposte e le domande deve capire con chi o cosa sta discutendo. Nel test di Searle quest'ultimo si sostituisce alla macchina e si suppone che egli si chiuda in una stanza e deve interagire all'esterno con una persona che parla cinese. Searle non conosce questa lingua e sul tavolo dispone di alcuni caratteri che deve utilizzare per rispondere, ma che egli comunque non comprende. Però supponiamo che nella stanza ci sia un libro con alcuni caratteri cinesi a cui sono associate delle istruzioni in inglese, Searle continua a non capire nulla di cinese ma comprende le regole che sono scritte e che gli indicano come deve rispondere. Queste regole, che costituiscono ciò che Searle chiama il "programma", gli rendono possibile mettere in relazione una serie di simboli formali con un'altra serie di simboli formali, cioè gli permettono di dare una risposta (output) a ogni domanda (input). Searle mette in luce che non ha mai dovuto interpretare i simboli cinesi per capire la domanda e dare la risposta giusta, dunque, non era necessario che egli comprendesse ciò che doveva fare, perché doveva solo seguire le istruzioni fornite. Searle afferma, quindi, che, come lui ha eseguito meccanicamente queste regole per rispondere alle domande, allo stesso modo una macchina può eseguire il programma scritto nel linguaggio di programmazione senza capire il significato di cosa sta facendo.

## 1.4 Rischi dell'IA

I vantaggi di questa tecnologia sono chiari, ma bisogna riconoscere che essa porta con sé anche alcuni rischi, dovuti in parte alla tecnologia stessa e in parte al fatto che provoca una trasformazione molto rapida della nostra vita.

### Dilemmi etici

L'aspetto dell'apprendimento nelle macchine ci pone davanti a numerosi problemi etici, legali e sociali, quali la diffusione della discriminazione sociale, l'equità e l'inclusione. Assumere un/una lavoratore/lavoratrice, valutarne le capacità, determinarne l'affidabilità, saranno decisioni legate sempre più a macchine e modelli matematici che restituiranno punteggi e predizioni, tradotti in giudizi capaci di cambiare la vita delle persone. Questo potrebbe anche significare che l'IA potrebbe prendere decisioni "discriminatorie" che escludono determinati gruppi di persone. Ciò è dovuto al fatto che le operazioni che compiono le macchine seguono dei precisi algoritmi di calcolo che devono essere pre-programmati da programmatori umani nel tentativo di "nutrire" i sistemi di Intelligenza Artificiale. Studi recenti hanno, infatti, dimostrato che anche gli algoritmi possono essere affetti da bias proprio perché progettati e alimentati da esseri umani. Inoltre, le macchine recepiscono qualsiasi tipo di dato da varie fonti, e sono dunque sottoposte al rischio di interiorizzare dati che le porteranno, poi, ad attuare comportamenti discriminatori e a produrre decisioni più o meno vantaggiose per alcune categorie di popolazione.

### 1.4.1 Ambito lavorativo

L'Intelligenza Artificiale potrebbe avere ripercussioni anche nell'ambito lavorativo. Infatti, una possibile conseguenza dell'uso dell'IA potrebbe essere la riduzione dell'occupazione di lavoratori umani e una conseguente redistribuzione non equa delle ricchezze. Se, per

esempio, un'azienda riuscisse ad automatizzare gran parte della catena di produzione, questo comporterebbe il licenziamento di molti dei suoi dipendenti, così da ridurre le spese dovute agli stipendi, aumentando, invece, il guadagno, facendo in modo che la ricchezza rimanga in mano a pochi, ovvero gli imprenditori. E questa possibilità è applicabile a gran parte dei lavori attuali. L'IA, inoltre, richiede ingenti risorse per essere studiata e progettata. Per questo le ricerche si sono spostate dai centri accademici alle grandi aziende, che dispongono di più risorse economiche e dati, ciò ha fatto sì che, poche grandi aziende possano detenere la maggior parte delle informazioni riguardo a questa disciplina.

### 1.4.2 Trasparenza degli algoritmi

Un altro dilemma etico riguarda la trasparenza degli algoritmi; un algoritmo normale è, infatti, programmato dall'uomo che fornisce ad esso un input, o delle indicazioni, rispetto a quello che deve o non deve fare. Tuttavia, nel caso dell'IA, il concetto è leggermente diverso, poiché man mano che questa impara, l'uomo perde gradualmente il controllo di cosa vi è dentro, di cosa sta apprendendo e in che modalità. Spesso, quindi, non è chiaro come la macchina arrivi alla sua decisione finale a partire dai dati in input ad essa forniti. Ciò impatta sia sul controllo umano rispetto alle decisioni dell'IA, sia sulla stessa fiducia verso questa tecnologia.

### 1.4.3 Protezione dei dati

Sono, inoltre, centrali le questioni legate alla privacy, ovvero alla gestione e alla protezione dei dati utilizzati dall'Intelligenza Artificiale. L'utilizzo massiccio di dati ed informazioni per addestrare algoritmi ad apprendere sempre più nozioni solleva preoccupazioni riguardo alla sicurezza delle informazioni personali che vengono analizzate e raccolte. L'IA, infatti, potrebbe essere vulnerabile agli attacchi informatici e gli hacker potrebbero manipolare gli algoritmi per ottenere risultati indesiderati o causare danni attraverso l'inserimento di dati falsi. Inoltre, l'IA potrebbe essere usata in dispositivi per il riconoscimento facciale o per la profilazione online e potrebbe mettere insieme le informazioni che acquisisce su una persona senza che questa ne sia a conoscenza. Lo dimostra l'istruttoria avviata dal Garante italiano per la protezione dei dati personali proprio nei confronti di OpenAI, che ha poi deciso in autonomia di sospendere ChatGPT in Italia finché non si allineerà alle richieste dell'Authority. Ciò significa che le aziende e i governi devono garantire che i dati siano protetti da accessi non autorizzati o da eventuali violazioni di sicurezza.

### 1.4.4 Altre problematiche

Altre problematiche potrebbero riguardare il fatto che la capacità dell'IA di generare contenuti (come testi e immagini) introduce anche rischi legati all'impatto sul sistema educativo, sui lavori creativi e sull'apprendimento consapevole di meccanismi di ragionamento logico. L'elaborazione di contenuti non è del tutto attendibile in quanto si rifà a tutto ciò che c'è in rete e, come sappiamo, non è tutto accertato, quindi, per ora, le notizie che produce possono contenere errori. Esse possono, altresì, perpetuare pregiudizi esistenti e informazioni false. Chi ha cattive intenzioni può manipolare e addirittura addestrare l'Intelligenza Artificiale a produrre un'informazione che diventa difficile da smascherare, creando credenze false e anche pericolose.

Ma, si tratta, comunque, di uno strumento che si può programmare, in sé per sé, non è né buono né cattivo. L'IA evolverà sempre ma bisogna essere consapevoli dei suoi problemi e cercare delle soluzioni adeguate in modo tale che sia utilizzata in modo etico sicuro e responsabile con precise regole che non vanno certo ad ostacolare la sua crescita di potenza.

## 1.5 Regolamentazione

Proprio per la pericolosità dell'utilizzo dell'Intelligenza Artificiale, l'Unione Europea è tra le prime a regolamentare l'utilizzo dell'Intelligenza Artificiale. Il commissario UE per l'industria è stato uno dei promotori in un programma in cui l'Unione Europea diventa il primo continente a stabilire regole chiare per l'uso dell'Intelligenza Artificiale. Egli afferma che l'AI act è molto più di un regolamento; è un trampolino di lancio per le start up e i ricercatori UE e guidare la corsa globale all'IA. In questo documento si stipulano dei parametri che l'Intelligenza Artificiale generale deve rispettare sulla sicurezza e trasparenza proprio in virtù del loro potenziale impatto. L'idea principale è quella di regolamentare l'IA sulla base della capacità di quest'ultima di causare danni alla società seguendo un approccio basato sul rischio: tanto maggiore è il rischio, quanto più rigorose sono le regole. Altre forme di IA sono state proprio bloccate perché possono mettere a rischio i nostri diritti e le nostre libertà fondamentali (come, per esempio, sistemi di categorizzazione biomedica, che utilizzano caratteristiche sensibili che possono diventare armi di manipolazione). Verranno create delle strutture amministrative, che controlleranno che vengano rispettate tutte queste regole.

## 1.6 Benefici dell'IA

Come anticipato, l'Intelligenza Artificiale è definita come uno dei pilastri su cui si basa la quarta rivoluzione industriale, in quanto "promette" una serie di vantaggi importanti, come il miglioramento dell'automazione e dell'interazione tra uomo e macchina, l'efficienza della produzione, la prevenzione di malfunzionamenti, di difetti e di incidenti del prodotto e delle risorse impiegate.

### 1.6.1 Ambito finanziario

Un campo di applicazione dell'Intelligenza Artificiale è la finanza. Secondo Forbes, al momento, il 70% dei servizi finanziari utilizza l'Intelligenza Artificiale. Uno tra i principali utilizzi è il Credit Score, un punteggio che indica l'affidabilità finanziaria di una persona, di grande importanza quando si parla di prestiti. L'Intelligenza Artificiale è in grado di analizzare un enorme quantità di dati relativi a quella persona, permettendo di valutare in maniera rapida e accurata il punteggio corrispondente.

### 1.6.2 Automatizzazione dei processi aziendali

L'Intelligenza Artificiale ha un ruolo fondamentale anche nell'automazione di processi aziendali, come già anticipato. Uno tra i maggiori vantaggi è l'ottimizzazione delle procedure, offrendo un valore aggiunto, sviluppando sistemi automatici per eseguire le attività ripetitive e standardizzate, riducendo il tempo impiegato e gli errori umani, migliorando così, la precisione e l'efficienza delle operazioni automatizzabili. Questo perché riesce ad elaborare enormi quantità di dati in pochissimo tempo e ad analizzarli in tempo reale per trovare le migliori soluzioni possibili (affidabili). L'Intelligenza Artificiale sta modificando, anche, il modo in cui le aziende si relazionano con i consumatori. Infatti, una delle principali applicazioni dell'IA nel marketing è l'analisi avanzata dei dati, che consente alle aziende di raccogliere e interpretare enormi quantità di dati sui comportamenti dei consumatori. Attraverso l'apprendimento automatico e le analisi predittive, le aziende possono anticipare le necessità e le preferenze dei consumatori, personalizzando le offerte e le esperienze di acquisto. Ciò permette di migliorare l'esperienza complessiva, semplificando la ricerca delle offerte, aumentando la velocità e la riduzione degli errori umani nelle transazioni commerciali.

### 1.6.3 Impatto ambientale

Anche nel campo della crisi climatica l'Intelligenza Artificiale può avere un impatto positivo e aiutare. È in grado di offrire nuove opzioni nella gestione dell'energia e dei rifiuti, di migliorare la sostenibilità dei prodotti, di ottimizzare i processi industriali e di limitare le emissioni e l'utilizzo di energia. Per questo motivo, risulta essere un importante tool nel raggiungere gli obiettivi del Green Deal europeo. Ci sono molte grandi aziende, come Google, che stanno investendo nell'utilizzo della tecnologia per il risanamento della salute del pianeta. DeepMind, una startup di proprietà di Google, ha quasi dimezzato il suo consumo energetico grazie all'analisi di dati da parte dell'IA. Altro vantaggio per l'ambiente è l'utilizzo dei droni nel campo dell'agricoltura, soluzione che permette di monitorare lo stato dei campi in tempo reale e in maniera "environmentally friendly". L'IA può contribuire a proteggere l'ambiente attraverso l'ottimizzazione dei processi, l'analisi dei dati ambientali e lo sviluppo di tecnologie sostenibili.

### 1.6.4 IA e sicurezza

Anche nel caso della giustizia l'utilizzo dell'IA potrebbe essere di aiuto. Essa potrebbe essere usata nella prevenzione dei reati e come supporto nella giustizia penale, perché permetterebbe di elaborare più velocemente grandi volumi di dati, valutare con sempre maggior cura eventuali rischi di fuga dei detenuti, prevedere e prevenire crimini e attacchi terroristici. Oppure potrebbe essere usata per la difesa e le strategie di attacco in caso di crimini informatici o per colpire obiettivi chiave nella lotta informatica. Attualmente l'Intelligenza Artificiale viene impiegata con successo in alcune piattaforme online per individuare ed eliminare pratiche illegali o inappropriate presenti in rete. L'IA può tenere sotto controllo i flussi di transazioni finanziarie, e sono in grado di riconoscere un furto, una frode o un attacco online molto più velocemente di come potrebbe fare una persona.

---

## Natural Language Processing

---

*Nel presente capitolo verrà trattato l'argomento del Natural Language Processing dandone una definizione generale e definendone i compiti più comuni, analizzando i diversi livelli di elaborazione del linguaggio. Verrà, poi, presentata una panoramica sulle applicazioni principali dell'NLP.*

### 2.1 Introduzione

Il Natural Language Processing (NLP) in italiano “Elaborazione del Linguaggio Naturale” è un particolare sistema di Intelligenza Artificiale che, negli ultimi anni, ha registrato un notevole progresso. Grazie al contributo di tecniche di AI, come Deep Learning e Machine Learning, l'NLP si sta facendo strada con grande rapidità, anche se il percorso è ancora lungo, considerando la complessità del linguaggio umano, ricco di espressioni gergali, modi di dire e dialetti. Gli studi sull'NLP vanno avanti da quasi settant'anni e i moderni applicativi rendono evidente come questa tecnologia stia conquistando uno spazio sempre più importante. È sufficiente pensare agli assistenti virtuali, ai Chatbot, ai sistemi di navigazione per le auto, ai correttori di ortografia e traduzioni automatiche; in realtà, queste sono solo alcune delle molteplici applicazioni che questa innovativa tecnologia trova oggi. L'NLP implica la "comprensione" di espressioni umane complete, almeno nella misura in cui è in grado di dare risposte utili ad esse. Fornendo interfacce uomo-macchina più naturali e un accesso più sofisticato alle informazioni memorizzate, l'elaborazione del linguaggio è arrivata a svolgere un ruolo centrale nella società dell'informazione multilingue.

### 2.2 Brevi cenni storici

Le ricerche inerenti al Natural Language Processing cominciarono ufficialmente nei primi anni Cinquanta, attraverso dei primi rudimentali esperimenti di traduzione automatica dall'Inglese al Russo. Ma è alla fine degli anni '50 che l'NLP cominciò a vedere gli albori, quando fu collegato da Minsky<sup>1</sup> all'Information Retrieval (IR), un insieme di tecniche che si occupano di gestire la rappresentazione, la memorizzazione, l'organizzazione e l'accesso ad oggetti contenenti informazioni. Questo periodo è caratterizzato da un forte ottimismo, poichè, nonostante le risorse tecniche ed economiche erano poche, comunque si riuscirono ad affrontare i primi problemi relativi alla semantica, alla sintattica e all'ambiguità del linguaggio,

---

<sup>1</sup>Matematico, informatico e ricercatore statunitense specializzato nel campo dell'Intelligenza Artificiale.

fornendo e completando dizionari e regole di traduzione. Ciò nonostante i progressi reali furono decisamente più lenti e, nel 1966, dopo che una ricerca di più di dieci anni non riuscì a mantenere le aspettative, i fondi per la traduzione automatica furono drasticamente ridotti e la ricerca venne sostanzialmente interrotta fino agli anni '80. In questa fase le regole NLP risultavano comunque distanti dall'Intelligenza Artificiale; infatti, fino ad allora, la maggior parte dei sistemi di Natural Language Processing erano basati su insiemi complessi di regole scritte a mano.

A cominciare dalla fine degli anni '80, però, ci fu una rivoluzione nell'ambito del Natural Language Processing con l'introduzione degli algoritmi di Machine Learning: ciò fu possibile grazie alla crescita delle capacità computazionali a disposizione (Legge di Moore, come indicato nella sezione 1.3.5). Successivamente gli esperti iniziarono ad utilizzare l'NLP statistico, per analizzare e generare i testi in linguaggio naturale utilizzando modelli statistici come alternativa agli approcci basati su regole linguistiche. L'approccio statistico permette alla macchina di imparare autonomamente sulla base dei vari testi e documenti, di grandi dimensioni che incontra, permettendo una maggiore robustezza, generalizzazione e gestione dell'ambiguità. Molti dei sistemi di riconoscimento vocale ancora oggi si basano su questo approccio grazie al fatto che essi garantiscono una forte robustezza al rumore e un risultato più affidabile in presenza di input non familiari o errati. Si inizia ad evidenziare un maggior approfondimento nello sviluppo di algoritmi volti all'addestramento dei modelli e al miglioramento dell'accuratezza; infatti, questo periodo segna la crescita degli algoritmi di Machine Learning nell'ambito dell'NLP. Grazie al lavoro svolto dai ricercatori IBM un grande successo venne raggiunto nell'ambito della traduzione automatica, dove vennero sviluppati modelli statistici sempre più accurati e capaci di utilizzare decine di documenti multilingue già esistenti come base di conoscenza per imparare a crearne di nuovi.

## 2.3 Cos'è il Natural Language Processing

Come nel campo dell'Intelligenza Artificiale, anche quando parliamo di Natural Language Processing trattiamo di fenomeni che riguardano proprio l'interazione tra due soggetti, ovvero i calcolatori e l'uomo. Il compito del Natural Language Processing è quello di creare un metodo di dialogo tra questi due soggetti principali. Una volta individuati i soggetti, è necessario definire, poi, il vero e proprio modello di interazione tra macchine ed esseri umani; in particolare, è necessario creare delle tecniche che riescano a far comunicare due mondi in apparenza completamente diversi, nonostante l'elaboratore abbia con sé le caratteristiche di un agente razionale. Ci si domanda come fare a trasformare il linguaggio umano, con tutte le sue problematiche e le sue ambiguità, in una serie di dati comprensibili da un elaboratore tramite regole formali (linguaggio macchina), considerando anche che questa trasformazione deve essere reversibile (l'essere umano deve poter capire dati elaborati da un calcolatore). Questo è proprio il compito dell'Elaborazione del Linguaggio Naturale. Per dare una definizione possiamo dire che l'NLP è un sottocampo dell'Intelligenza Artificiale che racchiude insieme elementi di linguistica (linguistica computazionale) e informatica e si occupa dell'interazione tra i computer e il linguaggio umano. L'obiettivo dell'NLP è proprio lo sviluppo di appositi algoritmi AI per analizzare e capire il linguaggio naturale scritto oppure parlato.

Il termine "linguaggio naturale" comprende tutte le lingue storico-naturali, ossia quelle nate spontaneamente lungo il corso della civiltà umana, che vengono utilizzate per la comunicazione quotidiana tra gli esseri umani. A differenza dei linguaggi artificiali, come i linguaggi di programmazione e le notazioni matematiche, i linguaggi naturali si sono evoluti di generazione in generazione, e in base al luogo dove si sono sviluppati; per questo motivo, essi sono difficili da definire con regole esplicite. L'NLP combina la linguistica computazionale

le e la modellazione basata su regole del linguaggio umano con modelli statistici di Machine Learning e di Deep Learning. Tali tecnologie, nel loro complesso, consentono ai computer di elaborare il linguaggio umano sotto forma di testo o dati vocali e di "capire" il suo pieno significato, compreso l'intento e i sentimenti di chi parla o scrive. Ogni parola nel testo viene convertita in modo che sia comprensibile alla macchina, ovvero sotto forma di numeri. Grazie a questo espediente è possibile processare linguaggio elaborando delle "semplici" matrici con cui siamo soliti lavorare.

## 2.4 Funzionamento

Diversamente dai linguaggi di programmazione, il Natural Language Processing si trova ad affrontare l'enorme difficoltà di doversi interfacciare con una "entità" che non segue regole facilmente interpretabili ovvero le lingue utilizzate dall'uomo. Per raggiungere un adeguato grado di comprensione e interazione viene, quindi, utilizzata la "Linguistica computazionale", ossia la scienza che si focalizza sullo sviluppo di formalismi descrittivi che illustrano il funzionamento di una lingua naturale per poi trasformarli in programmi eseguibili da una macchina. Si tratta di un processo complesso e delicato a causa delle caratteristiche proprie del linguaggio umano. Espressioni idiomatiche, elementi grammaticali, parole con lo stesso suono e molti altri elementi rendono il nostro linguaggio scritto e orale molto complicato da interpretare con i suoi vari significati. I sistemi NLP sono in grado di districarsi in questo ambito così difficile, individuando con rapidità i differenti elementi linguistici. Ciò è possibile grazie ad una pre-elaborazione dei dati che serve a semplificarli e a renderli in una forma che il computer sia capace, poi, di analizzarli, così da assegnare un significato. Per tentare di ridurre al minimo il numero di errori, e superare eventuali ambiguità, il processo di elaborazione viene diviso in più fasi. È, infatti, necessario comprendere la struttura del linguaggio umano e come trattare il testo prima di poter applicare tecniche di Machine Learning e Deep Learning. Le fasi più importanti dell'NLP sono le seguenti:

- *analisi lessicale*: scomposizione di un'espressione linguistica in token;
- *analisi grammaticale*: associazione delle parti del discorso a ciascuna parola nel testo;
- *analisi sintattica*: arrangiamento dei token in una struttura sintattica (ad esempio ad albero: parse tree);
- *analisi semantica*: assegnazione di un significato (semantica) alla struttura sintattica e, di conseguenza, all'espressione linguistica.

Nelle prossime sottosezioni esaminiamo in dettaglio ciascuna di queste fasi.

### 2.4.1 Analisi Lessicale

Il processo di elaborazione di un testo comincia dall'analisi lessicale o tokenizzazione. A questo livello, le persone, così come i sistemi NLP, interpretano il significato di singole parole. Questa fase, infatti, si occupa di convertire una sequenza di caratteri in una sequenza di token, ovvero stringhe, frasi, parole o unità semantiche più piccole, con un significato assegnato. Tale processo permette di analizzare la frase in modo semplificato, ad esempio consentendo l'eliminazione di parole inutili (stop words), ovvero parole che, data la loro elevata frequenza in una lingua, sono di solito ritenute poco significative. Tale fase non è particolarmente difficile, ma potrebbe presentare dei problemi. Un token può essere una parola, un numero, un simbolo di punteggiatura, o qualsiasi altro elemento che ha un significato all'interno di un testo. Un modo semplice per tokenizzare un testo è separarlo in base agli spazi che ci

sono tra una parola e l'altra. In italiano o in inglese gli spazi separano le parole; ciò rende semplice dividere un testo in token. In altre lingue invece, come il cinese e il giapponese, le parole non sono separate da spazi e potrebbe essere più complicata la separazione in token. Inoltre, per riconoscere token mai visti prima, ci sono delle tecniche come il Machine Learning. Questi modelli di Intelligenza Artificiale possono essere addestrati a riconoscere nuovi token, osservando tantissimi esempi di come sono usate le parole in quella lingua e imparando a riconoscere i modelli. Addestrare il modello su un grande set di dati che include vari elementi come, varietà linguistiche, slang, terminologie tecniche, etc., può aiutare quest'ultimo a riconoscere una gamma più ampia di elementi e usi dei caratteri. È possibile usare anche tecniche di apprendimento supervisionato, dove il modello viene addestrato su esempi di testo che sono stati già tokenizzati correttamente. Il modello apprende, quindi, a riconoscere i token basandosi sugli esempi forniti.

### 2.4.2 Analisi grammaticale

L'analisi grammaticale è il procedimento per assegnare una categoria lessicale (e le relative sotto-categorie) ad ogni parola in un testo nel quale è usata, risolvendo alcune ambiguità. Ad esempio, per la lingua italiana, ogni parola può appartenere a una delle nove categorie lessicali dell'italiano (cinque variabili: articolo, aggettivo, sostantivo o nome, pronomi, verbo, e quattro invariabili: avverbio, preposizione, congiunzione, interiezione o esclamazione); ciò significa comprendere lo scopo di ogni parola all'interno della frase. Le parti del discorso possono essere categorizzate come classi chiuse, ovvero classi in cui la condizione di appartenenza è relativamente fissa (ad esempio, le proposizioni), e classi aperte, in cui è possibile di volta in volta trovare nuovi elementi.

Nei sistemi di elaborazione del linguaggio naturale questa fase è chiamata Part-Of-Speech tagging, o POS tagging. Durante questa fase si etichettano le parole individuate nella fase di analisi lessicale con il POS corrispondente, senza eseguire una vera e propria analisi sintattica, ma ricorrendo, in genere, ad informazioni statistiche o a regole. Le etichette, o tag, sono raggruppate nei tagset; questi rappresentano una raccolta di tag utilizzata per una particolare attività. Il POS tagging aiuta a identificare entità come nomi, posizioni e organizzazioni all'interno del testo. Attraverso una migliore comprensione della struttura e della semantica delle frasi, questa tecnica consente alle macchine di studiare e comprendere il significato della frase in modo più accurato, facilitando l'identificazione dei collegamenti delle parole. Durante la codifica di parti del discorso ci possono essere delle complicazioni dovute all'ambiguità intrinseca del linguaggio che rende difficile l'etichettatura delle parole poiché possono significare cose diverse a seconda del contesto, il che può portare a malintesi. Esempi come gli slang, le espressioni colloquiali e le frasi idiomatiche possono essere problematici per i sistemi di etichettatura POS poiché non sempre seguono gli standard grammaticali formali.

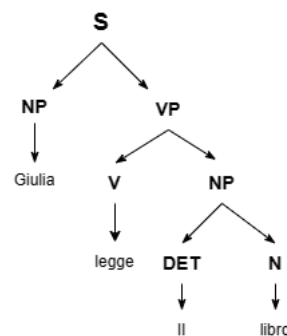
### 2.4.3 Analisi sintattica o parsificazione

La sintassi si riferisce all'insieme di regole, principi e processi che governano la struttura delle frasi in un linguaggio naturale. Una descrizione di base della sintassi è il modo in cui parole diverse, come soggetto, verbi, nomi, frasi nominali, etc. vengono sequenziate in una frase. In informatica, l'analisi sintattica, o parsificazione, è un processo che ha lo scopo di trovare le differenti parti che danno un significato alla frase, posizionando le parole in maniera corretta dentro una frase, in modo da determinare la correttezza della sua struttura grazie ad una data grammatica. Dal punto di vista dell'elaborazione a livello sintattico è possibile analizzare frasi anche complesse, riconoscere la loro correttezza grammaticale, e ricostruire la loro struttura sintattica. Ciò richiede sia una grammatica che un parser (programma che

esegue l'analisi sintattica). Tramite la sintassi di una lingua si può spiegare in che modo alcune serie di parole abbiano un senso compiuto, mentre altre con le stesse parole e ordine differente, siano invece prive di senso. Spesso questo procedimento avviene attraverso la costruzione di un albero sintattico che mostra le regole utilizzate durante il riconoscimento dell'input.

Un albero sintattico è un albero che rappresenta la struttura sintattica di una frase in accordo a determinate forme grammaticali; esso è composto da nodi e rami. Un programma che produce quest'albero è, appunto, il parser. Facciamo un esempio per capire meglio questo procedimento utilizzando la frase "Giulia legge il libro" (Figura 2.1). Nell'albero sono usate le seguenti abbreviazioni:

- S: sta per la frase, la struttura;
- NP: frase nominale; il primo ramo (quello a sinistra) è il soggetto "Giulia", l'altro è l'oggetto della frase;
- VP: sta per frase verbale, che serve il predicato;
- V: per verbo; in questo caso "legge";
- Det: sta per aggettivo determinativo, in questo caso "il";
- N: sta per sostantivo.



**Figura 2.1:** Esempio di parsing sintattico

Nella maggior parte dei linguaggi, tuttavia, l'analisi sintattica opera su una sequenza di token in cui l'analizzatore lessicale spezzetta l'input. Anche questo step dell'NLP presenta delle ambiguità di carattere sintattico e può creare non poche problematiche.

#### 2.4.4 Analisi semantica

L'analisi semantica è una componente cruciale dell'elaborazione del linguaggio naturale; essa si concentra sulla comprensione del significato, dell'interpretazione e delle relazioni tra parole e frasi in un determinato contesto. Essa rappresenta il ponte tra l'espressione umana e la comprensione della macchina. L'analisi semantica è l'attività di assegnazione di un significato, un senso, alla struttura sintattica corretta e determina le varie accezioni di una frase. La disambiguazione<sup>2</sup> del senso delle parole è un aspetto fondamentale che aiuta a determinare i significati multipli di queste ultime. I significati, ovvero i sensi, spesso sono tipicamente classificati in una gerarchia di concetti. Si classificano parole o gruppi di parole all'interno delle classi che raccolgono sinonimi. L'insieme di classi è chiamato tassonomia, mentre le relazioni tra queste classi rappresentano l'ontologia.

<sup>2</sup>Processo attraverso il quale si precisa il significato di una parola per evitare che ci sia ambiguità.

## 2.5 Panoramica sulle applicazioni del Natural Language Processing

La maggior parte delle applicazioni che utilizzano il testo si basa sul Natural Language Processing; di seguito vengono riportate le applicazioni più frequenti.

### 2.5.1 Information Retrieval

L'Information Retrieval (IR) è l'insieme delle tecniche utilizzate per gestire la rappresentazione, la memorizzazione, l'organizzazione e l'accesso ad oggetti contenenti informazioni quali documenti, pagine web, cataloghi online e oggetti multimediali, al fine di rendere agevole all'utente il soddisfacimento dei propri bisogni informativi. Data una collezione di documenti e un bisogno informativo dell'utente, lo scopo di un sistema di IR è di trovare le informazioni che potrebbero essere utili a soddisfare la richiesta dell'utente. Per fare questo l'IR si basa su alcuni concetti chiave:

- *Query* ("interrogazioni"): sono stringhe di parole-chiave rappresentanti l'informazione richiesta. Vengono digitate dall'utente in un sistema di IR (per esempio, un motore di ricerca) e rappresentano la concretizzazione del reale bisogno informativo dell'utente;
- *Oggetto*: un'entità che possiede informazioni, le quali potrebbero essere la risposta dell'interrogazione dell'utente, come pagine web, articoli scientifici o libri digitali;
- *Rilevanza*: la misura in cui un documento corrisponde alla nostra query.

Nel momento in cui un utente intende usare un qualsiasi sistema di reperimento dell'informazione per acquisire informazioni su un determinato argomento, questi deve tradurre tale necessità in una query. A partire da essa, l'Information Retrieval troverà tutti i documenti rilevanti alla richiesta effettuata. Le tecniche di Information Retrieval, nate per dati testuali, sono state in seguito estese ed applicate anche a dati multimediali (immagini, audio, video) e semi-strutturati.

### 2.5.2 Information Extraction

L'estrazione delle informazioni è il processo di analisi dei dati non strutturati e che prevede l'identificazione automatizzata e l'estrazione delle informazioni essenziali e specifiche in formati di dati più modificabili e strutturati. L'obiettivo principale è quello di trasformare dati di testo non strutturati in dati strutturati che possono essere facilmente analizzati e ricercati, automatizzando l'estrazione di dati da fonti eterogenee, come documenti, fatture, contratti o rapporti, eliminando la necessità di un intervento umano prolungato e soggetto a errori. Il sistema è in grado di comprendere il contenuto testuale, riconoscere pattern e relazioni semantiche all'interno del testo, e convertire queste informazioni in dati facilmente utilizzabili in applicazioni aziendali e sistemi di gestione.

L'Information Extraction aiuta le organizzazioni e le aziende a ricavare informazioni preziose da pool di dati sempre più grandi. Ad esempio, le aziende estraggono informazioni dai moduli di sondaggio e le analizzano per determinare se i loro clienti sono felici, infelici o indifferenti. Pertanto, l'apprendimento automatico è diventato il metodo preferito per estrarre dati utili su larga scala e con maggiore precisione.

L'estrazione delle informazioni è un compito impegnativo che richiede l'uso di varie tecniche, tra cui il riconoscimento delle entità denominate (NER), uno strumento secondario per trovare specifiche informazioni. Esiste un'ampia gamma di applicazioni coinvolte, tra cui motori di ricerca, chatbot, sistemi di raccomandazione e rilevamento di frodi, tra gli altri, che lo rendono uno strumento vitale nel campo dell'NLP.

### 2.5.3 Question Answering

Il Question Answering (QA) consiste nel rispondere automaticamente a una domanda espressa in linguaggio naturale. Per trovare la risposta a una domanda, un programma di QA può utilizzare una base di conoscenza<sup>3</sup> o una raccolta di documenti in linguaggio naturale (un corpus quale il World Wide Web o altre collezioni locali). La ricerca nell'ambito del QA cerca di fronteggiare un ampio range di tipi di domande. Collezioni di ricerca variano da piccoli corpora a raccolte di notizie, al World Wide Web<sup>4</sup>.

### 2.5.4 Summarization

Il termine "Text Summarization" fa riferimento al procedimento di generazione automatica di una versione ridotta di un determinato documento di testo che trasmette le informazioni più importanti del contenuto originale. È una tecnica utilizzata per ridurre le dimensioni e la complessità del materiale originale, fornendo una sinossi più gestibile da leggere e comprendere per gli utenti. Quando si tratta di tecniche di riepilogo del testo, esistono due approcci principali:

- *Extractive Summarization*: tale approccio identifica ed estrae le frasi più importanti dal testo originale per formare il riassunto. Si basa su funzionalità come la posizione della frase, la frequenza delle parole chiave e la pertinenza all'argomento principale del testo. Questo metodo non crea nuove parole o frasi, ma prende parole e frasi già esistenti e le presenta.
- *Abstractive Summarization*: il suo approccio tenta di comprendere i punti principali del testo riformulandoli attraverso una parafrasi in modo nuovo, riuscendo a superare le incongruenze grammaticali che possono presentarsi nel metodo estrattivo. In questo modo vengono create nuove espressioni e frasi che trasmettono le informazioni più utili dal testo originale. Esso richiede tecniche avanzate di Elaborazione del Linguaggio Naturale, come l'apprendimento automatico e il Deep Learning. Per questo motivo, le tecniche di riepilogo astrattivo sono più complesse delle tecniche di riepilogo estrattivo e sono anche computazionalmente più costose.

### 2.5.5 Dialogue Processing

Le tecniche di Dialogue Processing vengono applicate ai sistemi di dialogo e servono per creare una conversazione tra uomo e macchina. Esse affrontano le caratteristiche di conversazione tipiche tra uomo-uomo integrando, poi, nei sistemi di dialogo per la successiva interazione uomo-macchina, migliorandone, così, la conversazione. L'utente, però, si deve adattare alla macchina poiché, ad oggi, quest'ultima è in grado di comprendere solo un vocabolario limitato.

## 2.6 L'importanza del Natural Language Processing oggi

Il Natural Language Processing riveste una certa importanza nel mondo odierno soprattutto per l'enorme quantità di dati non strutturati che vengono prodotti ogni giorno. Ai giorni nostri l'NLP è in continua espansione e, grazie alle nuove tecnologie informatiche, ricopre tuttora un ruolo fondamentale, ad esempio per quanto riguarda le sue espressioni in campo

<sup>3</sup>È un tipo speciale di banca dati per la gestione della conoscenza per scopi aziendali, culturali o didattici. Essa costituisce dunque un ambiente volto a facilitare la raccolta, l'organizzazione e la distribuzione della conoscenza.

<sup>4</sup>Permette di navigare e usufruire di un insieme molto vasto di contenuti amatoriali e professionali.

di Information Retrieval (Google, Bing, Yahoo, etc.) e nel campo della traduzione automatica (Google Translate, dizionari online, etc.). Molti progressi sono stati fatti nel campo della sintassi, migliorando i programmi parser attraverso strutture logiche sempre più ben definite; inoltre la potenza sempre più elevata degli elaboratori ha facilitato notevolmente l'esecuzione e l'elaborazione di algoritmi anche pesanti. Si interagisce quotidianamente con l'NLP, per esempio, nell'ambito dei sistemi GPS a comando vocale, degli assistenti digitali, del software di dettatura speech-to-text, dei chatbot per il servizio clienti e di altre funzionalità destinate ai consumatori. Inoltre, ci sono moltissime aziende, persone e attività che generano dati non strutturati, composti da file di testi scritti, conversazioni in chat, referti, cartelle cliniche, e tanto altro. Molte di queste informazioni sono catalogate in database e storage e, fino a pochi anni fa, era complesso analizzarli in maniera efficace. Oggi, invece, tutto ciò diventa molto più semplice per merito di tecnologie avanzate come l'NLP di ultima generazione che può agire senza interruzioni e fornire analisi rapide e precise. L'NLP gioca anche un ruolo importante nelle soluzioni aziendali per facilitare le operazioni di business, aumentare la produttività dei dipendenti e semplificare i processi di business.

Ovviamente siamo ancora lontani da sistemi perfetti in grado di lavorare al cento per cento delle possibilità garantendo una conoscenza globale dell'ambiente sul quale si affacciano e dei dati che hanno a disposizione. La ricerca prosegue e l'obiettivo ultimo per le macchine è quello di comprendere e generare testi e conversazioni in modo simile agli esseri umani. Le sfide principali oggi sono inerenti alla possibilità di far fronte all'ambiguità del linguaggio umano. Un'importante sfida riguarda, infatti, la capacità di risolvere ambiguità che l'uomo facilmente introduce nella sua interazione con altri uomini. Ma è facile pensare che l'NLP occuperà una posizione sempre più rilevante nella società moderna.

---

## Esperienze nell'ambito dei Chatbot

---

*Nel presente capitolo verrà esplorato il mondo dei Chatbot, partendo dalla loro definizione, per poi concentrarsi sulla loro evoluzione e sulle principali applicazioni; infine, verrà approfondita una piattaforma per la realizzazione di Chatbot chiamata Amazon Lex. In particolare, si esplorerà il suo funzionamento e si mostrerà un'esperienza effettuata con quest'ultima.*

### 3.1 Introduzione ai Chatbot

Il termine "Chatbot" unisce le parole inglesi "bot" (abbreviazione di robot), che indica un programma o uno script per l'automazione di compiti ripetitivi, e "chat" (dall'inglese dialogo), in riferimento alle piattaforme di messaggistica. Il termine "chatterbot" venne coniato per la prima volta nel 1994 da Michael Mauldin, inventore di uno dei primi motori di ricerca della storia, Lycos<sup>1</sup>. Si può dire, tuttavia, che sia stato Alan Turing, attraverso il suo test (Test di Turing introdotto nella sezione 1.3.6) a intuire già nel 1950 la potenzialità di simulare una conversazione umana attraverso un software. Un Chatbot è, quindi, un software progettato per dialogare con gli esseri umani all'interno di una piattaforma di messaggistica utilizzando il linguaggio naturale, e permette di interagire con la tecnologia utilizzando una varietà di metodi di input, come testo o voce. I Chatbot sono noti con una grande varietà di nomi, come bot conversazionale, assistente virtuale, assistente digitale, e tanti altri. Questa varietà di termini mette in evidenza la loro crescente presenza nella vita quotidiana e la diversità delle loro applicazioni. I Chatbot, infatti, possono essere realizzati per diversi scopi, come il supporto per i clienti, rispondere alle domande degli utenti, eseguire un ordine, o aiutare una persona nel momento in cui ha dei problemi. Ad esempio, gli assistenti vocali, come Siri di Apple, Google Assistant o Amazon Alexa, sono tipi specifici di Chatbot che si concentrano sulla risposta a domande e sull'esecuzione di comandi vocali, mentre i Chatbot conversazionali sui siti web possono fornire supporto all'utente o assistenza nell'acquisto di prodotti e servizi. I Chatbot possono essere implementati con diversi gradi di funzionalità o di "intelligenza". I più semplici sono delle soluzioni front-end in grado di rispondere a domande frequenti degli utenti. I più avanzati, alimentati da algoritmi di Intelligenza Artificiale e apprendimento automatico, sono in grado di comprendere il contesto della conversazione, apprendere dai dati e migliorare continuamente le loro risposte. Questo livello di sofisticazione li rende adatti per applicazioni più complesse, come il supporto

---

<sup>1</sup>Ormai non molto conosciuto, ma nel '95 nacque come uno dei motori di ricerca più importanti del web. Fino al 2000 fu tra i leader insieme a Yahoo ed altri motori di ricerca, successivamente surclassati da Google.

virtuale, l'automazione dei processi aziendali e molto altro. L'implementazione e lo sviluppo dei Chatbot sono diventati parte integrante della trasformazione digitale in molteplici settori; essi offrono un modo più semplice e immediato di accedere alle informazioni da parte di un utente, costituendo un'alternativa alla navigazione in menù complessi in cui non risulta facile individuare le informazioni necessarie.

## 3.2 Storia

Nonostante si possa credere che i Chatbot siano nati solo recentemente, in realtà la loro storia risale a decenni fa. Nel 1966 Joseph Weizenbaum presenta quello che si può considerare come il primo Chatbot, ELIZA (come introdotto nella Sezione 1.3.2).

```

Welcome to
          EEEEE LL   IIII  ZZZZZ  AAAAA
          EE    LL   II    ZZ    AA  AA
          EEEEE LL   II    ZZ    AAAAAA
          EE    LL   II    ZZ    AA  AA
          EEEEE LLLLL IIII  ZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:

```

**Figura 3.1:** Programma Eliza

ELIZA nasce come parodia di un terapeuta rogersiano<sup>2</sup>; il suo funzionamento è semplice: le risposte sono generate dalla domanda dell'utente grazie a specifiche regole di decomposizione e assemblaggio.

Successivamente, nel 1972, lo psichiatra Kenneth Colby, implementa PARRY, durante i suoi studi all'università di Stanford. Rimanendo in ambito medico, il chatbot simula il pensiero di un individuo paranoico e schizofrenico. PARRY funzionava tramite un complesso sistema di ipotesi, attribuzioni e risposte emotive che si innescavano durante determinati dialoghi e spunti verbali. Esso, possedeva, quindi, una strategia di conversazione, risultando, dunque, più convincente di ELIZA. PARRY è stato uno dei primi sistemi a passare il test di Turing. PARRY ed ELIZA hanno interagito diverse volte. La più famosa delle conversazioni tra i due bot è avvenuta durante la International Conference on Computer Communications, tenutasi nell'ottobre del 1972; PARRY ed ELIZA si rispondevano a vicenda, collegati tramite ARPANET<sup>3</sup>.

Nel 1988 il programmatore Rollo Carpenter sviluppa il Chatbot Jabberwacky, ma in maniera completamente diversa dai precedenti bot. L'obiettivo era simulare la naturale chat umana ma in modo interessante, divertente e umoristico.

<sup>2</sup>Approccio psicoterapeutico basato sulla Teoria del Cliente. Il terapeuta rogersiano riconosce nel paziente una persona che, in quanto tale, è in una posizione egualitaria nei confronti del terapeuta.

<sup>3</sup>Fu una rete di computer studiata e realizzata dal Dipartimento della difesa degli Stati Uniti. Da questa si originò a partire dal 1983 l'attuale rete Internet.

Nel 1990, l'inventore americano Hugh Loebner lancia il Premio Loebner, una competizione annuale che prende il formato di un test di Turing standard con i giudici che premiano il programma per computer più simile ad un uomo.

Nel 1995 il dottor Richard S. Wallace costruì A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) e vinse il Premio Loebner nel 2000, 2001 e 2004. A.L.I.C.E è un chatbot realizzato come software open source e con un approccio minimalista; esso è programmato con il linguaggio A.I.M.L. (Artificial Intelligence Markup Language).

Nel 2000 Robert Hoffer e Timothy Kay fondarono la società ActiveBuddy. Lo scopo era quello di creare agenti intelligenti in grado di comunicare attraverso piattaforme di messaggistica istantanea. Hoffer notando l'aumento della popolarità dei servizi di messaggistica, ebbe l'idea di creare agenti interattivi per aggiungere funzionalità e migliorare l'esperienza. L'implementazione comprendeva una vasta gamma di applicazioni come l'accesso a notizie, meteo, elenchi telefonici e una varietà di altri strumenti. Tali applicazioni furono inserite in un unico pacchetto e lanciate nel 2001 sotto il nome di SmarterChild, il quale permetteva un rapido accesso alle informazioni e offriva la possibilità di divertirsi anche attraverso la conversazione. Il successo del progetto, utilizzato da oltre 13 milioni di utenti, portò alla realizzazione di prodotti promozionali in ambito musicale, cinematografico e televisivo.

Successivamente, nel 2006, i ricercatori dell'IBM sviluppano il Chatbot Watson (il suo nome è stato scelto per rendere omaggio al primo presidente dell'azienda, Thomas J. Watson). Esso divenne famoso perché riuscì a sconfiggere i campioni in carica in Jeopardy<sup>4</sup>, un gioco molto famoso in quel periodo.

Una grande rilevanza in ambito dell'assistenza vocale oggi giorno è attribuita sicuramente a Siri, il quale è, probabilmente, l'assistente vocale più famoso. Originariamente Siri fu distribuita come applicazione indipendente per iOS; solo successivamente fu acquisita da Apple e integrata nei vari dispositivi. Siri può fornire indicazioni stradali o sul meteo, rispondere a domande, inviare messaggi, e molto altro ancora.

Nel 2012 anche Google sviluppa il suo assistente virtuale, ovvero Google Now. Quest'ultimo è un software di assistenza personale e usa un'interfaccia utente in lingua naturale per rispondere alle domande, dare consigli, eseguire azioni e mettere in contatto l'utente con diversi servizi nella rete. Una funzione ulteriore è quella di ricordare le conversazioni per provare a prevedere cosa l'utente chiederà. Nel 2016 Google ha presentato il suo nuovo assistente virtuale chiamato Google Assistant pensato come evoluzione di Google Now.

Anche Microsoft sviluppa nell'aprile del 2014 un software di assistenza vocale creato per il sistema operativo Windows Phone, chiamato Cortana. Questo Chatbot permette agli utenti di interagire tramite determinate conversazioni, l'invio di messaggi e di soddisfare richieste di informazioni di vario tipo. Inoltre, Cortana è in grado di impostare delle notifiche per ricordare eventi o comunicare notizie, come il meteo e il traffico. Tuttavia, nel 2023 l'assistenza vocale di Cortana in Windows come app autonoma è stata ritirata.

Successivamente, anche l'azienda statunitense Amazon sviluppa il proprio assistente virtuale, ovvero Alexa. Quest'ultimo è un assistente personale intelligente e i suoi algoritmi permettono ad esso di essere in grado di interpretare il linguaggio naturale, dialogare con gli umani, fornendo informazioni di diverso tipo, ed eseguire differenti comandi vocali, interpretandone anche il tono. Tale Chatbot permette di ottenere un maggiore grado di riconoscimento che consente di rispondere in maniera sempre meno artificiale e di assolvere a richieste sempre più complesse.

Nell'aprile del 2016, Mark Zuckerberg, fondatore di Facebook, ha presentato i Chatbot costruiti appositamente per Messenger. Essi ebbero subito successo; infatti, in poco tempo, se ne registrarono oltre 10mila. L'assistente virtuale di Facebook esegue operazioni molto

---

<sup>4</sup>Quiz televisivo statunitense che consiste in una gara di cultura generale tra i vari concorrenti.

complicate, come la prenotazioni di viaggi o ristoranti partendo da una semplice chat di richiesta.

Nell'autunno del 2022 OpenAI ha lanciato il suo chatbot ChatGPT. Quest'ultimo era uno dei chatbot più avanzati al momento del lancio e viene considerato una parte fondamentale nello sviluppo dell'IA conversazionale. Il modello è stato addestrato su grandi quantità di conversazioni umane, dati e informazioni, e può, quindi, comunicare con gli utenti in modo naturale e simil-umano. ChatGPT viene spesso utilizzato per scopi di assistenza clienti ed è in grado di rispondere a domande su una vasta gamma di argomenti.

### 3.3 Cosa sono i Chatbot e le loro funzionalità

Un Chatbot è un software progettato per simulare una conversazione umana con gli utenti attraverso l'uso di chat, messaggi di testo o vocali. I Chatbot, diventati ormai sempre più popolari grazie alle loro molteplici funzionalità, possono essere composti da elementari software che rispondono ad una semplice query, o da software più complessi per produrre diversi livelli di personalizzazione. Un Chatbot è in grado di interagire e conversare con un umano, come se lo fosse, proprio perché è programmato per rispondere alle richieste dell'utente formulate secondo il linguaggio naturale. Raccogliendo e analizzando i dati che riceve durante la conversazione, esso riesce, poi, ad elaborare una risposta che possa soddisfare la richiesta proposta dall'utente con cui sta interagendo. I Chatbot hanno svariate applicazioni ormai in campi molto diversi tra loro. Le funzionalità principali per cui vengono utilizzati i Chatbot sono l'assistenza clienti, l'automazione, il marketing, la ricerca e l'analisi di dati. I Chatbot sono solitamente integrati sui siti aziendali o appositamente sviluppati per le principali app di messaggistica istantanea come Messenger o Telegram; essi vengono utilizzati anche dai settori che si occupano dell'assistenza medica fino ad arrivare al mondo degli assistenti personali come Siri e Cortana. Proprio per la loro capacità di rispondere alle richieste, e grazie al costante sviluppo dell'IA, i Chatbot possono facilitare le operazioni in molti processi, per questo trovano moltissime applicazioni. Per esempio, i Chatbot possono essere utilizzati nei processi che si occupano del servizio clienti, della personalizzazione dei prodotti e servizi e nella raccolta di informazioni da parte delle imprese attraverso dei questionari. Tramite l'uso di una chat automatica le società sono capaci di rispondere ad una ricca gamma di richieste degli utenti automaticamente, abbattendo le spese ed i tempi di una tipica interazione umana. I Chatbot sono, quindi, un valido sostituto per proporre un servizio o un prodotto che somiglia a quello umano ma ad un numero sempre più vasto di clienti. All'interno del mondo dei Chatbot sono compresi sia gli assistenti digitali a comando vocale (Siri, Cortana, Alexa e Google Home), che i sistemi basati su testo, distribuiti sulle piattaforme di messaggistica istantanea come Whatsapp, Instagram, Snapchat etc. I Chatbot sono guidati prevalentemente da Machine Learning e dal Natural Language Processing e riescono ad elaborare i dati per dare risposta a una molteplicità di richieste.

Esistono due principali categorie di Chatbot che si possono riscontrare; esse sono:

- *I Chatbot dichiarativi (dedicati alle attività)*: sono programmi che si concentrano sull'esecuzione di una funzione. Essi fanno uso di regole NLP, e poco di Machine Learning, e generano risposte automatizzate alle domande degli utenti. I Chatbot dichiarativi gestiscono domande semplici che non implicano una grande varietà di variabili. Infatti, le loro capacità sono abbastanza basilari, nonostante utilizzino l'NLP per fare in modo che gli utenti possano sperimentarli in modo semplice. Ad oggi, questi sono i Chatbot più usati.
- *I Chatbot predittivi basati sui dati (di conversazione)*: spesso sono indicati come assistenti virtuali o digitali poiché sono più interattivi rispetto ai Chatbot dichiarativi. Questa

tipologia di Chatbot è consapevole del contesto di riferimento ed è in grado di sfruttare la comprensione della lingua naturale, l’NLP e il Machine Learning per imparare. Per questo sono molto più sofisticati e personalizzati rispetto ai Chatbot dichiarativi. Siri di Apple e Alexa di Amazon sono esempi di Chatbot predittivi orientati al consumatore e basati sui dati.

## 3.4 Vantaggi e limiti nei Chatbot

Il Chatbot è una tecnologia diventata sempre più popolare; tale utilizzo però, ci porta a considerare i vantaggi e i limiti che si presentano nel momento in cui essi si utilizzano nei sistemi.

### 3.4.1 Vantaggi

Oggi si può dire che i Chatbot rappresentano un vero e proprio valore aggiunto per le imprese poiché ne incrementano l’efficienza a livello operativo. Inoltre, essi rappresentano un enorme beneficio anche per i clienti in quanto permettono alle persone di usufruire di importanti ed utili servizi aggiuntivi. I vantaggi principali derivanti dall’utilizzo di un Chatbot dipendono chiaramente dalle specifiche implementazioni e dalla tipologia di servizio offerto. Si possono comunque individuare alcuni vantaggi generali ed indipendenti dalla specifica implementazione perché intrinseci alla soluzione Chatbot.

Un primo vantaggio è, sicuramente, la massima disponibilità del servizio dei Chatbot, in quanto quest’ultimi sono reperibili tutti giorni e 24h su 24h, anche nei giorni di festività. Tale caratteristica dà la possibilità ai clienti di ottenere un’assistenza immediata e costante, indipendentemente dalla posizione del cliente nel mondo e dai fusi orari. I Chatbot così forniscono un servizio più rapido rispetto a quello umano e in grado di processare un numero elevato di richieste contemporaneamente, in modo da ridurre il tempo di attesa dei clienti, garantendo un’assistenza sempre attiva. Ciò consente di migliorare l’esperienza dell’utente a un costo molto basso per le imprese. Infatti, realizzare e mantenere un Chatbot è, in genere, molto più economico che assumere un dipendente. Ciò comporta anche un aumento della produttività poiché, i Chatbot, compiendo le operazioni più ripetitive, consentono di liberare le risorse del personale umano e impiegarle per svolgere dei compiti più complessi.

I Chatbot, inoltre, permettono maggiore semplicità nel trovare le informazioni utili relative ad un determinato argomento (molto più facile che cercare tra le varie domande di un sito ad esempio), oppure una maggiore facilità nell’ottenere un determinato risultato rispetto all’utilizzo di un’applicazione dedicata, ad esempio. Nel tempo è, inoltre, possibile che il Chatbot memorizzi le informazioni degli utenti, garantendo così, un’esperienza personalizzata con risposte specifiche e su misura per le esigenze degli utenti.

### 3.4.2 Limiti

Nonostante ci sia stato un enorme sviluppo in campo dell’Intelligenza Artificiale con la diffusione dei nuovi Big Data e il veloce incremento delle capacità di apprendimento delle macchine, i Chatbot presentano ancora i limiti propri dell’Intelligenza Artificiale, su cui si basano. Nonostante i Chatbot IA siano incredibilmente utili e intelligenti, sono comunque vulnerabili agli errori. Se, grazie al Machine Learning, il tasso di errore dei Chatbot diminuisce con l’incrementare del suo uso, può capitare, tuttavia, che alcuni sistemi di Chatbot fraintendano le domande, forniscano le risposte sbagliate e non raggiungano l’obiettivo. Poiché i Chatbot si basano su un vasto bacino di dati disponibili su Internet, è molto probabile che questi dati siano obsoleti, imprecisi o irrilevanti. Il limite più importante che risulta essere

oggi ancora una sfida molto importante per gli studiosi è la difficoltà, da parte dei Chatbot, di capire il linguaggio naturale e di comprendere, quindi, le richieste più complesse o ambigue. Questo fa sì che i Chatbot possano essere facilmente ingannabili o manipolabili, dagli utenti, creando dei risultati errati o distorti, che possono causare l'uso o la diffusione, intenzionale o meno, di informazioni false per deviare altri utenti. Questo mette in evidenza anche il problema della sicurezza, in quanto i Chatbot che raccolgono informazioni possono essere soggetti a degli attacchi informatici, da parte di malintenzionati, e a conseguenti perdite di dati e incidenti di sicurezza.

Un altro limite importante è la mancanza di intelligenza emotiva, infatti, anche se i Chatbot rispondono alle domande in modo colloquiale, non hanno l'empatia e la moralità di un vero essere umano.

Inoltre, bisogna considerare che non tutti i consumatori sono pronti ad abbandonare interamente le interazioni umane. Alcuni consumatori non si fidano della tecnologia per l'esecuzione di determinati compiti delicati, come trasferire fondi o fissare appuntamenti. La sensazione di parlare con una macchina aumenta l'attitudine a confidarsi, ma anche il distacco, e anche il più sofisticato Chatbot conversatore fatica a creare una vera connessione con l'utente.

## 3.5 Esempi di utilizzo dei Chatbot

### 3.5.1 Educazione

Nel panorama in rapida evoluzione dell'istruzione, i progressi tecnologici hanno svolto un ruolo fondamentale nel trasformare i metodi di apprendimento tradizionali. In questo contesto un ruolo importante è attribuito all'utilizzo dei Chatbot applicandoli alla didattica. Infatti, uno dei principali punti critici nel campo dell'istruzione nelle scuole è la riduzione dell'assistenza da parte degli insegnanti a causa dell'aumento del numero di studenti. Spesso gli insegnanti incorrono in ostacoli, come la mancanza di risorse, e questo comporta una difficoltà nel dare riscontri variegati e specifici per ogni studente. I Chatbot possono, infatti, supportare le attività di insegnamento e promuovere un'esperienza di apprendimento più personalizzata e coinvolgente. Memorizzando le informazioni delle lezioni, essi possono rispondere prontamente alle domande degli studenti e aiutarli in un determinato argomento in cui hanno difficoltà.

I Chatbot, inoltre, possono aiutare gli studenti per l'iscrizione a determinati corsi o pratiche burocratiche, dove spesso non hanno un aiuto sufficiente e si ritrovano in difficoltà. Così facendo si ridurrebbe notevolmente la mole di carico amministrativo sugli insegnanti. Molto importante è anche l'utilizzo dei Chatbot per l'apprendimento di lingue straniere. Raramente uno studente, che vuole apprendere una lingua straniera, si ritrova a dialogare in una lingua che non sia quella nativa, e le app per l'apprendimento delle lingue spesso utilizzano Chatbot per aiutare gli studenti a praticare la conversazione, il vocabolario e la pronuncia.

### 3.5.2 Industria

I Chatbot sono molto utili nel settore della vendita al dettaglio e della produzione perché interagiscono e forniscono assistenza ai clienti in tempo reale. La conversazione tra Chatbot e cliente è simile a quella umana grazie all'Intelligenza Artificiale e all'apprendimento automatico. L'impatto dei Chatbot sulle aziende oggi è significativo e multidimensionale e può apportare enormi benefici. L'utilizzo dei Chatbot permette una riduzione dei costi, automatizzando le risposte alle domande frequenti e gestendo un numero elevato di interazioni contemporaneamente, qualcosa che sarebbe molto costoso e difficile da realizzare con il

personale umano. Così facendo, il personale umano si può concentrare su attività più complesse e creative, che sono difficili da automatizzare. Inoltre, i Chatbot, possono agire come assistenti virtuali di vendita, guidando i clienti attraverso il processo di acquisto, fornendo raccomandazioni di prodotti e promozioni in tempo reale. La presenza di questi assistenti virtuali è anche un segno di innovazione e adattabilità da parte dell'azienda, aumentandone la competitività nel mercato. I Chatbot permettono, inoltre un miglioramento del servizio clienti con risposte immediate tutti i giorni e a tutte le ore indipendentemente dal luogo in cui una persona si trova.

### 3.5.3 Servizio clienti

L'utilizzo dei Chatbot per il servizio clienti è sempre più diffuso nelle aziende per fornire aiuto e assistenza ai consumatori. Il commercio elettronico, infatti, ha rivoluzionato il modo in cui le aziende promuovono e vendono i loro prodotti, arrivando a sempre più persone in tempo minore. Ciò ha fatto emergere, però, delle problematiche, ovvero la difficoltà nel gestire tutte le domande o richieste da parte degli utenti. Spesso, infatti, si verificano tempi prolungati di attesa per ricevere risposta da parte degli operatori aziendali e tali risposte potrebbero non essere rilevanti alle specifiche richieste del cliente. Per questo i Chatbot sono molto utili nel settore della vendita al dettaglio e della produzione perché interagiscono e forniscono assistenza ai clienti in tempo reale e sono costantemente attivi tutti i giorni. Essi sono basati su software in grado di simulare e condurre conversazioni umane (scritte o parlate), pronti a rispondere agli interrogativi, gestirne le richieste dei clienti e i reclami, fornendo informazioni utili per supportare le vendite o le attività post-vendita.

### 3.5.4 Assistenza sanitaria

L'utilizzo dei Chatbot offre alle organizzazioni sanitarie l'opportunità di estendere le proprie risorse, raggiungendo un maggior numero di persone. Con l'aiuto di conversazioni interattive le organizzazioni possono assistere pazienti, famiglie, infermieri e medici in vari modi, sia per compiti semplici, che questioni complesse, come identificare i sintomi, gestire i farmaci e monitorare i problemi di salute cronici. I Chatbot possono essere impiegati, per esempio, per aiutare i pazienti a ricordarsi l'assunzione dei farmaci impostando dei promemoria, o per aiutarli e fornendo informazioni sulla loro salute.

Una delle caratteristiche più rilevanti è la possibilità attraverso i Chatbot di fornire assistenza medica 24 ore su 24 e tutti i giorni della settimana nel caso il paziente abbia dei dubbi o dei problemi. I Chatbot sono, infatti, un'alternativa più affidabile e accurata alle ricerche online che i pazienti effettuano quando cercano di capire la causa dei loro sintomi. Al momento l'obiettivo non è quello di diagnosticare completamente i pazienti tramite gli assistenti virtuali, ma piuttosto di guidarli verso le risorse giuste, e aiutare gli operatori sanitari a comprendere meglio le loro esigenze.

Inoltre, i Chatbot possono essere anche dei validi assistenti nel momento in cui una persona abbia bisogno di supporto psicologico. Bisogna ricordare, però, che il tema della sensibilità e trasparenza dei dati è un aspetto importante su cui si sta ancora lavorando in campo dell'IA e, di conseguenza, dei Chatbot. Per questo motivo i pazienti, per questioni più complesse e delicate riguardo la loro salute, preferiscono rivolgersi a dottori umani. Tuttavia, i Chatbot sono impiegati nel campo della psicoterapia già da ELIZA, e vengono utilizzati nella prevenzione del suicidio e dell'aiuto per persone con bisogni, offrendo un modo più accessibile e facile per affrontare i problemi legati alla salute mentale.

## 3.6 Introduzione a Amazon AWS

Nell'ambito della creazione dei chatbot un ruolo fondamentale lo svolge Amazon. Quest'ultima, infatti, ha sviluppato Amazon AWS (Amazon Web Services), la piattaforma di servizi Cloud più utilizzata del mondo. L'idea alla base di AWS è stata sviluppata da Chris Pinkham e Benjamin Black che, nel 2003, hanno presentato un documento che proponeva un modello di infrastruttura basato su Cloud all'interno di Amazon. Questo modello prevedeva l'offerta di servizi di calcolo su richiesta. Nel 2004, Jeff Bezos, fondatore e CEO di Amazon, ha approvato il progetto e ha dato il via allo sviluppo di ciò che sarebbe diventato Amazon Web Services.

Nel 2006, AWS è stato ufficialmente lanciato, offrendo servizi di Cloud Computing, tra cui calcolo e storage su richiesta. Oggi, AWS è diventato uno dei principali provider di servizi Cloud al mondo, con milioni di clienti che utilizzano i suoi servizi in vari settori, tra cui aziende, istituzioni governative e organizzazioni senza scopo di lucro. AWS si contraddistingue nel panorama dei provider Cloud per il suo ampio e innovativo set di servizi e funzionalità. Esso offre una vasta gamma di servizi, più di 200, che coprono diverse esigenze, dai fondamentali servizi di infrastruttura, come il calcolo e l'archiviazione, ai settori più innovativi, come il Machine Learning, l'Intelligenza Artificiale e l'analisi dei dati.

L'utilizzo di AWS permette una vasta scelta di database specializzati per diversi contesti applicativi. In questo modo gli utenti possono scegliere il database più adatto alle loro specifiche esigenze, ottimizzando, così, sia i costi che le prestazioni. Tra questi vari servizi, possiamo trovare Amazon Lex, una piattaforma che permette di creare esperienze di conversazione automatizzate e intuitive, ovvero i Chatbot.

## 3.7 Amazon Lex

Questa sezione si dedica all'approfondimento di Amazon Lex, un servizio di Intelligenza Artificiale che rappresenta una significativa evoluzione nell'ambito di interfacce conversazionali e dei Chatbot. Si tratta di una piattaforma di servizi di elaborazione del linguaggio naturale (NLP) sviluppata da Amazon Web Services che offre potenti strumenti per la creazione di esperienze di conversazione automatizzate e intuitive.

### 3.7.1 Che cos'è

Amazon Lex V2 è un servizio AWS progettato per permettere agli sviluppatori di creare interfacce conversazionali attraverso la voce e il testo in qualsiasi applicazione. Attraverso l'utilizzo di tecnologie avanzate di riconoscimento vocale (ASR - Automatic Speech Recognition) e la comprensione del linguaggio naturale (NLU - Natural Language Understanding), Amazon Lex offre funzionalità e flessibilità avanzate, consentendo all'utente di creare esperienze coinvolgenti con interazioni di comunicazione naturali, oltre che di creare nuove categorie di prodotti. Esso permette così di creare Chatbot sofisticati e naturali; esso, è, infatti, in grado di interpretare le intenzioni degli utenti, rispondere a domande, eseguire azioni specifiche e mantenere dialoghi naturali. Amazon Lex offre un'ampia integrazione con servizi AWS, inclusi AWS Lambda, Amazon Connect, Amazon Comprehend e Amazon Kendra.

Amazon Lex permette, quindi, di avere numerosi vantaggi. Esso consente di creare in maniera semplice, attraverso pochi passaggi, un Chatbot. Utilizza modelli avanzati di riconoscimento del linguaggio naturale per comprendere le intenzioni degli utenti (chiamate intent) e rispondere in modo appropriato. Può, infatti, comprendere e interpretare il linguaggio naturale degli utenti, sia attraverso comandi vocali che input testuali. L'utente fornisce solo alcune frasi di esempio e Amazon Lex crea un modello completo di linguaggio naturale tramite

cui il bot riesce a gestire conversazioni complesse, riconoscendo le intenzioni e raccogliendo informazioni specifiche attraverso parametri (slot); così l'utente può interagire utilizzando voce e testo per porre domande, ottenere risposte e completare operazioni sofisticate. Come già affermato sopra, anche il fatto che Amazon Lex riesca a trasformare la voce in testo, lo rende fondamentale per il suo utilizzo, per esempio, come assistente su dispositivi mobili o integrazioni con sistemi telefonici. Infatti, in ambito informatico il riconoscimento vocale e la comprensione del linguaggio naturale sono tra i problemi più complicati affrontati finora.

I Chatbot Amazon Lex convertono in testo le parti parlate in entrata e comprendono l'intento dell'utente per generare una risposta intelligente, permettendo all'utente di concentrarsi sulla creazione del bot. Inoltre Amazon Lex ha una interfaccia utente intuitiva e strumenti per semplificare la creazione e la gestione dei bot. Esso permette di creare, testare e distribuire i chatbot direttamente dalla console di Amazon Lex, con la possibilità di pubblicarli sui dispositivi mobili, attraverso app e il web, e servizi di chat, consentendo, quindi, una facilità di sviluppo in quanto gestisce automaticamente le risorse eliminando la necessità di gestire infrastrutture. Amazon Lex opera in modo integrato con altri servizi AWS, come Amazon Lambda (per eseguire funzioni serverless), sfruttando e traendo vantaggio dalla potenza della piattaforma AWS per la sicurezza, il monitoraggio, l'autenticazione utente, lo storage e lo sviluppo di applicazioni mobili. Inoltre Lex è altamente scalabile, il che significa che può gestire un numero elevato di interazioni contemporanee senza compromettere le prestazioni. Infine, con Amazon Lex, non vi sono costi anticipati o tariffe minime. Esso usa un modello di pagamento basato sul consumo, vengono addebitate solo le richieste vocali o di testo effettuate. Con il piano gratuito per il quale si può provare facilmente Amazon Lex, senza alcun investimento iniziale.

### 3.7.2 Funzionamento

Il funzionamento di un Chatbot si basa su diverse risorse per comprendere e gestire le conversazioni con gli utenti. Di seguito è presente un elenco di concetti fondamentali e di terminologia usata:

- *Intento*: rappresenta un'azione che l'utente desidera eseguire. Ogni intento è associato a un'azione specifica che il bot deve eseguire. Per definirlo è necessario specificare: il nome, ovvero un nome descrittivo, degli esempi di enunciati, ovvero come l'utente potrebbe esprimere l'intento; infine, va soddisfatto l'intento.
- *Slot*: sono parametri o informazioni aggiuntive richieste da un intento. Un intento può richiedere zero o più slot o parametri. In fase di esecuzione, Amazon Lex richiede all'utente valori di slot specifici. L'utente deve fornire i valori per tutti gli slot richiesti prima che Amazon Lex possa soddisfare l'intento. Ciascuno slot ha un tipo, come data e ora, che può essere creato personalmente oppure può essere predefinito.
- *Versione e Alias*: una versione è uno snapshot numerato del lavoro, mentre un alias è un puntatore a una versione specifica del bot. Versione e alias consentono di gestire l'evoluzione del bot senza richiedere aggiornamenti ai client.
- *Utterance*: le utterance sono frasi o espressioni che l'utente potrebbe utilizzare per esprimere un intento. Ad esempio, per l'intento "prenotare un volo", le utterance potrebbero includere "Voglio prenotare un volo" o "Ho bisogno di un biglietto aereo".
- *Prompt*: un prompt è una domanda o, richiesta, che il bot fa all'utente per raccogliere le informazioni necessarie a completare l'intento. Se l'utente non fornisce tutte le informazioni necessarie, Lex utilizza i prompt per chiedere dettagli aggiuntivi.

Grazie alle risorse dedicate alla gestione delle conversazioni, Amazon Lex offre la possibilità di sviluppare applicazioni tramite un'interfaccia testuale o vocale per interagire con gli utenti. Di seguito vengono riportati i passaggi tipici per operare con Amazon Lex:

- *Creare un bot*: è possibile definire un bot e aggiungere uno o più linguaggi. Prima di tutto si assegna un nome al bot, poi si selezionano le impostazioni di base, come la lingua, e si configurano le impostazioni di sicurezza. Successivamente è possibile configurare il bot affinché comprenda gli obiettivi dell'utente, si impegni in una conversazione per ottenere informazioni e soddisfi l'intento dell'utente.
- *Simulazione delle interazioni*: Amazon Lex offre la possibilità di utilizzare la finestra di test integrata fornita dalla console per verificare il comportamento del bot e simulare le conversazioni, e per verificare che il bot riconosca correttamente gli intent e risponda adeguatamente.
- *Pubblicare una versione e creare un Alias*: è possibile creare una versione del bot e associare un alias ad essa.
- *Deploy del bot*: Amazon Lex permette la distribuzione del bot e di integrarlo con altre piattaforme di messaggistica come Facebook Messenger o Slack. Ciò consente di rendere il bot accessibile agli utenti su più canali.

## 3.8 Realizzazione di un chatbot con Amazon Lex

In questa sezione, si affrontano e vengono presentate le varie fasi di creazione, sviluppo e test di un Chatbot utilizzando Amazon Lex; nel caso specifico, il Chatbot consentirà un'ordinazione in un fast food.

### 3.8.1 Creazione

Il primo passo da effettuare nello sviluppo del Chatbot è la sua creazione. Essa prevede una scelta iniziale del tipo di bot con cui cominciare. La scelta comprende tre tipi di bot:

- *bot vuoto*: consiste nel creare un bot di base senza linguaggi, intenti e tipi di slot preconfigurati;
- *bot con esempio*: ovvero si prende un bot con un esempio da cui iniziare, con lingue, intenti e tipi di slot preconfigurati, che si possono modificare;
- *bot con le trascrizioni*: uno strumento che genera automaticamente intenti a partire dalle trascrizioni delle conversazioni caricate.

Si può avere, anche, un'altra opzione utilizzando l'Intelligenza Artificiale. Attraverso la descrizione del tipo di bot che si desidera creare, Lex utilizzerà l'Intelligenza Artificiale generativa per creare intenti e tipi di slot.

Scegliendo di iniziare con la creazione del bot vuoto, la configurazione prosegue con l'inserimento obbligatorio del nome, l'inserimento facoltativo di una descrizione e le autorizzazioni IAM e il Children's Online Privacy Protection Act (COPPA). L'ultima fase della creazione prevede la scelta della lingua desiderata e la selezione di una voce artificiale nell'applicazione di riferimento (Figura 3.2). Con questo passaggio si completa la configurazione iniziale del chatbot, passando, poi, allo sviluppo effettivo.

**Figura 3.2:** Creazione di un chatbot Amazon Lex

### 3.8.2 Gestione delle conversazioni

Successivamente si procede con la creazione degli intenti e degli slot, necessari a soddisfare il compito assegnato. In questa fase dello sviluppo, bisogna definire la logica del Chatbot, quindi cercare di capire come lo si vuole sviluppare e qual è il suo funzionamento. In questo elaborato abbiamo sviluppato un Chatbot che si occupa delle ordinazioni in un fast food. Abbiamo, quindi, creato una serie di intenti per permettere al cliente di avere diverse possibilità di scelta tra vari prodotti che il fast food offre: hamburger, insalata, crocchette.

	Name
<input type="radio"/>	<a href="#">hamburger_senza_menu</a>
<input type="radio"/>	<a href="#">Richiestainiziale</a>
<input type="radio"/>	<a href="#">hamburger_menu</a>
<input type="radio"/>	<a href="#">insalata_senza_menu</a>
<input type="radio"/>	<a href="#">crocchette_menu</a>
<input type="radio"/>	<a href="#">crocchette_senza_menu</a>
<input type="radio"/>	<a href="#">insalata_menu</a>
<input type="radio"/>	<a href="#">FallbackIntent</a>
<input type="radio"/>	<a href="#">chiedo_menu_hamburger</a>
<input type="radio"/>	<a href="#">chiedo_menu_crocchette</a>

**Figura 3.3:** Lista degli intenti

L'intento iniziale è quello denominato "Richiestainiziale" (Figura 3.3), nel quale si chiede all'utente che prodotto vuole ordinare. Successivamente, una volta effettuata la sua scelta, abbiamo creato, per ogni alimento, un intento nel quale si chiede all'utente se vuole il menù. Da questo punto abbiamo creato due intenti per ogni prodotto, uno con il menù e uno senza menù. Questa diramazione è stata possibile attraverso una ramificazione condizionale, nella quale dall'intento "Richiestainiziale" riusciamo, in base alla risposta dell'utente, ad invocare l'intento desiderato. Successivamente, anche nel caso della scelta del menù, viene fatta una ramificazione per condurre il chatbot nell'intento voluto, ovvero con il menù o senza menù. Questa procedura è verificata per ogni alimento. Nel caso del prodotto con il menù vengono chiesti: la dimensione, il tipo di carne, le salse, le patatine, e la bevanda. Questo procedimento

viene ripetuto nell'intento di ogni alimento. Nel caso del prodotto senza menù non sono incluse le patatine e le bevande. Infine si ritorna nell'intento "Richiesta iniziale" nel quale il chatbot chiede se l'utente vuole ordinare altro. Se l'utente risponde in modo affermativo viene ripetuto il procedimento, altrimenti vengono richieste la data e l'orario di ricezione dell'ordine. Infine si chiede la conferma per procedere alla richiesta di ordinazione (Figura 3.4).

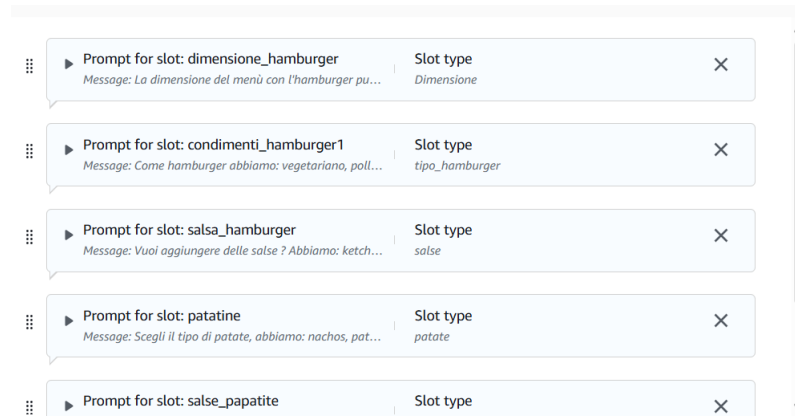


Figura 3.4: Slot dell'intent `hamburger_menu`

Inoltre, di default, viene creato l'intento `FallbackIntent`, che viene progettato per gestire situazioni in cui il sistema non riesce a determinare in modo corretto l'intento dell'utente.

In ogni Intent sono stati definiti e creati vari tipi di Slot necessari per eseguire le azioni del bot. Infatti, oltre agli slot integrati presenti in Amazon Lex, che presentano i dati di slot utilizzati più frequentemente come data, ora e posizione, abbiamo creato vari tipi di slot per definire diversi dati degli alimenti che servono per permettere l'ordinazione del cliente. All'interno degli slot sono state inserite le risorse utilizzate. Abbiamo creato dei tipi di slot, per esempio, per definire la dimensione del menù, oppure il tipo di carne da mettere nell'hamburger o nell'insalata (come mostrato nella Figura 3.5).

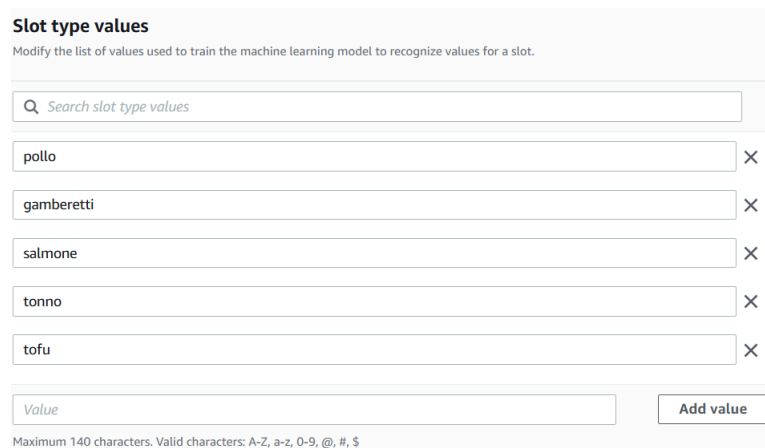


Figura 3.5: Tipologia di slot

Successivamente abbiamo creato un codice Lambda affinché gestisse gli errori dovuti a input non previsti da parte dell'utente. Per ogni intento abbiamo invocato la funzione Lambda che controlla per ogni slot la validità del valore e, in caso contrario, restituisce un messaggio predefinito di errore. Il codice è definito da due funzioni: la prima è "lambda\_handler", che viene richiamata quando viene invocata la funzione Lambda; la seconda, chiamata

"validazione\_richiesta", viene utilizzata per confrontare il valore degli slot predefiniti con l'input dell'utente (Figura 3.6).

```

import json

ListaCibo = ['hamburger', 'croccchette', 'insalata']
Riordinare = ['si', 'no']
Dimensione = ['piccola', 'media', 'grande']
CondimentiHamburger = ['vegetariano', 'pollo con formaggio', 'pollo con bacon', 'manzo con formaggio', 'manzo con bacon']
Salse = ['ketchup', 'barbecue', 'maionese', 'salsa rosa', 'piccante']
Proteine = ['tofu', 'pollo', 'gamberetti', 'salmone']
CondimentiInsalata = ['olive', 'mais', 'pomodori', 'cipolla', 'feta']
DimensioneCroccchette = ['4', '6', '9', '20']
TipoCroccchette = ['pollo', 'pollo piccante', 'manzo', 'vegetariane']
Patatine = ['nachos', 'patate fritte', 'patate fritte con bacon', 'patate fritte con formaggio']
Bevande = ['cocacola', 'cocacola zero', 'estahè al limone', 'estahè alla pesca', 'acqua naturale', 'acqua gassata', 'fanta', 'fanta zero']

def validazione_richiesta(slots, intent):

    if intent == 'Richiestainiziale':
        # Validazione lista_cibo
        if not slots['lista_cibo']:
            return {
                'isValid': False,
                'invalidSlot': 'lista_cibo'
            }
        if slots['lista_cibo']['value']['originalValue'].lower() not in ListaCibo:
            return {
                'isValid': False,
                'invalidSlot': 'lista_cibo',
                'message': 'Perfavore selezionare una tra le seguenti scelte {}'.format(", ".join(ListaCibo))
            }
        # Validazione riordinare
        if not slots['riordinare']:
            return {
                'isValid': False,
                'invalidSlot': 'riordinare'
            }
        if slots['riordinare']['value']['originalValue'].lower() not in Riordinare:
            return {
                ...
            }

    # Valid Order
    return {'isValid': True}

def lambda_handler(event, context):
    slots = event['sessionState']['intent']['slots']
    intent = event['sessionState']['intent']['name']

    risultato_validazione_richiesta = validazione_richiesta(slots, intent)

    if not risultato_validazione_richiesta['isValid']:
        if 'message' in risultato_validazione_richiesta:
            risposta = {
                "sessionState": {
                    "dialogAction": {
                        "slotToElicit": risultato_validazione_richiesta['invalidSlot'],
                        "type": "ElicitSlot"
                    },
                    "intent": {
                        "name": intent,
                        "slots": slots
                    }
                },
                "messages": [
                    {
                        "contentType": "PlainText",
                        "content": risultato_validazione_richiesta['message']
                    }
                ]
            }
        else:
            risposta = {
                "sessionState": {
                    "dialogAction": {
                        "slotToElicit": risultato_validazione_richiesta['invalidSlot'],
                        "type": "ElicitSlot"
                    }
                }
            }
    else:
        risposta = {
            "sessionState": {
                "dialogAction": {
                    "type": "Delegate"
                },
                "intent": {
                    "name": intent,
                    "slots": slots
                }
            }
        }
    return risposta

```

**Figura 3.6:** Codice Lambda per la gestione degli errori

Nella Figura 3.6 viene riportata una parte del codice. Prima di tutto viene fatta un'inizializzazione dei valori degli slot che verranno utilizzati, quindi, degli array di stringhe nelle quali si definiscono i valori. Successivamente viene definita la funzione "validazione\_richiesta", dove vengono passati come parametri gli slot e il nome dell'intento corrente. Attraverso vari controlli condizionali all'interno della funzione si verifica la correttezza di ogni slot nell'intento corrente, verificando se lo slot è vuoto o se il contenuto dello slot è compreso tra i valori accettati all'interno degli array, che abbiamo definito precedentemente. Se il controllo non va a buon fine l'istruzione "return" restituisce un oggetto JSON con cui imposta le chiavi

"isValid" a "False", "invalidSlot" al nome dello slot non valido e "message" ad una stringa rivolta all'utente che chiede di selezionare i giusti valori per quello slot. Se al contrario lo slot è corretto, quindi tutti i controlli sono soddisfatti, allora "isValid" viene impostato a "True". La funzione "lambda\_headler" è invocata ogni volta che viene effettuato un collegamento al codice Lambda da parte della console di Lex. Dopo aver richiamato la funzione "validazione\_richiesta" si valuterà il risultato ottenuto per poi gestire e strutturare l'oggetto JSON e trasmetterlo come output al chatbot. La funzione "lambda\_headler" distinguerà due casi, nel primo caso uno in cui lo slot non è valido e viene restituito il messaggio di errore, impostando la chiave "slotToElicit" allo slot non valido; la funzione chiederà all'utente di riscrivere il valore scegliendo tra quelli consentiti. Nel secondo caso lo slot è valido; allora il chatbot dovrà proseguire con la normale programmazione.

### 3.8.3 Utilizzo

L'ultimo passaggio per lo sviluppo del Chatbot è la fase di test e di deploy. La fase di test è fondamentale affinché il Chatbot funzioni correttamente e risponda in maniera adeguata alle risposte degli utenti (Figura 3.7 e Figura 3.8)

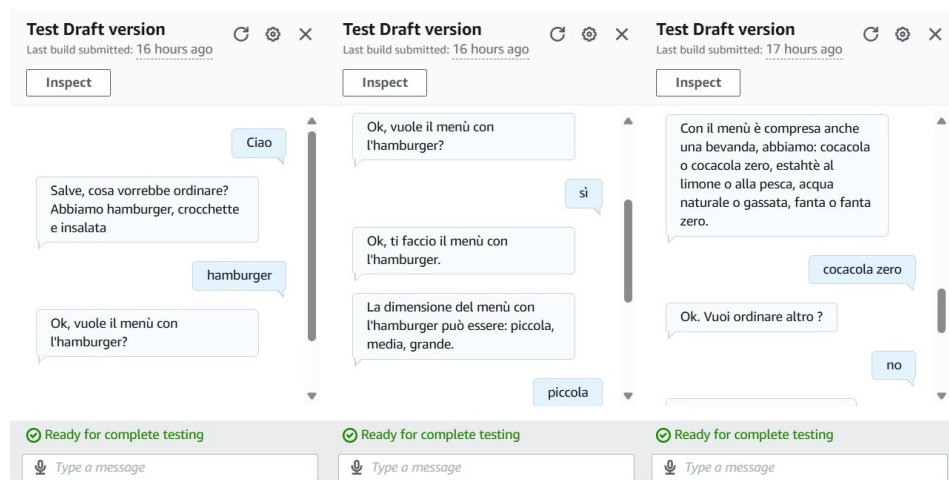


Figura 3.7: Esempio di una conversazione con il chatbot (prima parte)

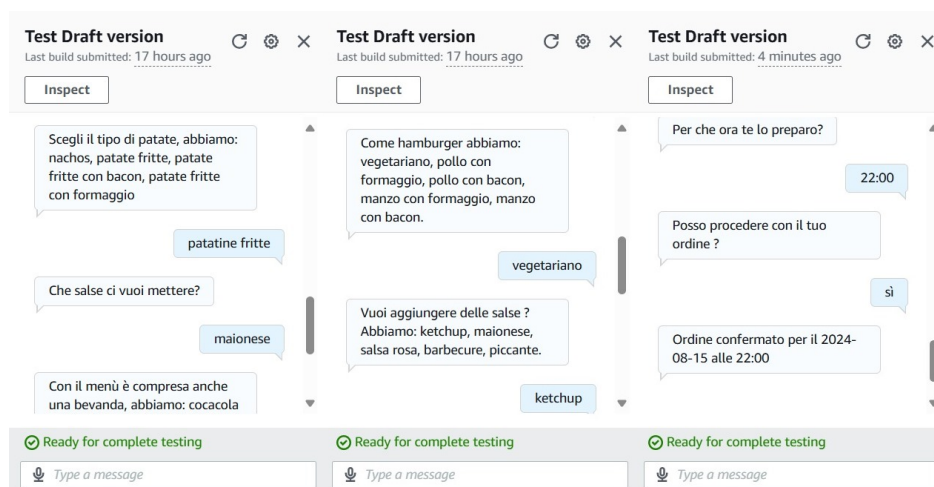
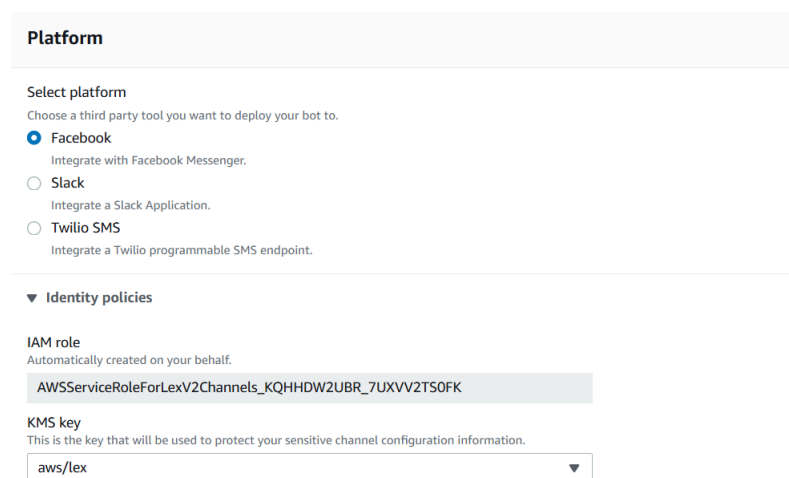


Figura 3.8: Esempio di una conversazione con il chatbot (seconda parte)

Qui si possono inserire varie utterance (frasi degli utenti) per vedere se il bot riconosce correttamente gli intent e risponde adeguatamente, e per provare diverse formulazioni della stessa richiesta per assicurarsi che il bot possa gestire la varietà di espressioni che un utente potrebbe utilizzare. Una volta eseguito il build è possibile testare il bot direttamente all'interno dell'interfaccia Amazon Lex per verificare il funzionamento prima di effettuare il deploy su altre piattaforme. Nelle Figure 3.7 e 3.8 viene riportato un esempio di funzionamento del chatbot.

Terminata la fase di test, è possibile distribuire il chatbot, integrandolo all'interno di un'app mobile, utilizzando strumenti specifici della piattaforma. Inizialmente si genera un alias, ovvero un nome associato alla versione e che viene utilizzato per la sua distribuzione. Tramite l'alias è possibile integrare ed utilizzare canali differenti da Amazon Web Service. La configurazione del canale di output offre una scelta tra diverse piattaforme di messaggistica, come Facebook, Slack e Twilio SMS (Figura 3.9).



The screenshot shows the 'Platform' configuration screen in the Amazon Lex console. It is divided into two main sections: 'Select platform' and 'Identity policies'.  
In the 'Select platform' section, there is a heading 'Select platform' and a sub-heading 'Choose a third party tool you want to deploy your bot to.'. Three radio button options are listed: 'Facebook' (selected), 'Slack', and 'Twilio SMS'. Each option has a brief description: 'Integrate with Facebook Messenger.', 'Integrate a Slack Application.', and 'Integrate a Twilio programmable SMS endpoint.' respectively.  
The 'Identity policies' section is expanded, showing 'IAM role' and 'KMS key'. The IAM role is 'AWSServiceRoleForLexV2Channels\_KQHHDW2UBR\_7UXVV2TS0FK' and the KMS key is 'aws/lex'.

**Figura 3.9:** Schermata per l'aggiunta di un canale dalla console Amazon Lex

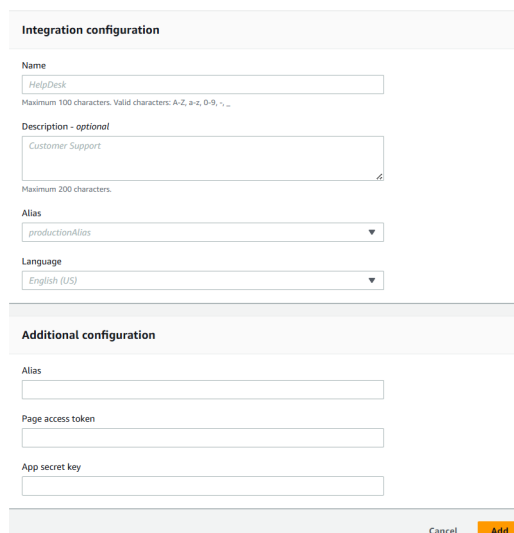
Una volta effettuata la scelta, la configurazione prosegue scegliendo il ruolo IAM (Identity and Access Management) e la chiave KMS (Key Management Service). Il ruolo IAM consente di concedere a un'entità (come un servizio AWS, un utente o un'applicazione) l'accesso temporaneo e limitato a risorse specifiche in AWS, senza dover condividere le credenziali di accesso permanenti. Tale ruolo è fondamentale per gestire l'accesso e le autorizzazioni in un ambiente AWS sicuro e scalabile. La chiave KMS è un servizio che consente di creare, gestire e controllare le chiavi di crittografia utilizzate per proteggere i dati utili per la protezione delle informazioni sensibili sulla configurazione del canale.

Per terminare, bisogna inserire il nome del canale e due codici di sicurezza (Figura 3.10) ovvero:

- *token di accesso alla pagina*: è un tipo di credenziale che consente al bot di interagire con una specifica pagina Facebook e compiere delle azioni, come pubblicare post, rispondere a messaggi, ottenere informazioni sulla pagina, etc;
- *chiave segreta*: è la chiave associata all'applicazione che serve ad autenticarne le richieste quando interagisce con l'API di Facebook. Viene utilizzata per firmare e convalidare le richieste API, garantendo che provengano da fonti autorizzate;

Una volta completati tutti i passaggi, verrà generato un URL di callback. Quest'ultimo è il punto di ingresso del server che Facebook chiamerà per inviare eventi al bot. Esso deve essere

sicuro in quanto svolge un ruolo fondamentale nel processo di autorizzazione dell'utente e nel collegamento del bot con Facebook. Successivamente, dopo che l'utente ha concesso l'autorizzazione, Facebook reindirizza al precedente URL di callback, includendo un codice di autorizzazione o un token di accesso. Tale URL riveste un'importanza cruciale nel processo di autenticazione OAuth, facilitando all'applicazione il ricevimento e lo scambio dei token necessari per le successive richieste API.



The screenshot displays the 'Integration configuration' form in the Amazon Lex console. It is divided into two main sections: 'Integration configuration' and 'Additional configuration'.  
**Integration configuration:**  
- **Name:** A text input field containing 'HelpDesk'. Below it, a note states: 'Maximum 100 characters. Valid characters: A-Z, a-z, 0-9, -, \_'.  
- **Description - optional:** A text area containing 'Customer Support'. Below it, a note states: 'Maximum 200 characters.'  
- **Alias:** A dropdown menu with 'productionAlias' selected.  
- **Language:** A dropdown menu with 'English (US)' selected.  
**Additional configuration:**  
- **Alias:** An empty text input field.  
- **Page access token:** An empty text input field.  
- **App secret key:** An empty text input field.  
At the bottom right of the form, there are two buttons: 'Cancel' and 'Add'.

**Figura 3.10:** Schermata per la configurazione di un canale dalla console Amazon Lex

---

## Esperienze nell'ambito del riconoscimento facciale

---

*Nel presente capitolo verrà esplorato il task dell'Intelligenza Artificiale che si occupa del riconoscimento facciale. In particolare verranno presentati il suo funzionamento e le sue applicazioni. In seguito verrà affrontato un tool di riconoscimento facciale, ovvero Amazon Rekognition. Attraverso questa piattaforma si faranno vari test per verificare la sua efficacia nel riconoscere le caratteristiche del viso e le emozioni.*

### 4.1 Introduzione al Riconoscimento Facciale

Il riconoscimento facciale è una tecnologia biometrica che permette di identificare o verificare l'identità di una persona basandosi sull'analisi delle caratteristiche del viso, utilizzando immagini o video che la ritraggono. Solitamente il riconoscimento avviene mediante tecniche di elaborazione digitale delle immagini, ignorando tutto quello che non rappresenta una faccia, come edifici, alberi, corpi, che vengono solitamente definiti di background. La sua storia è strettamente legata allo sviluppo dell'informatica, dell'Intelligenza Artificiale e delle tecniche di riconoscimento delle immagini. Il riconoscimento facciale utilizza algoritmi avanzati per trasformare le immagini del volto in dati digitali che possono essere confrontati con quelli memorizzati in un database. Anche se meno accurato del riconoscimento delle impronte digitali, è spesso preferito per la sua natura senza contatto, e viene utilizzato in una vasta gamma di settori, principalmente nella sicurezza personale, nelle forze dell'ordine o nell'onboarding digitale nella finanza, oltre a essere presente in numerose situazioni di vita quotidiana come, per esempio, lo sblocco di un dispositivo utilizzando il viso. Ricopre una particolare importanza proprio nell'ambito della sicurezza pubblica, grazie all'elevato grado di attendibilità ormai raggiunto, che consente di facilitare operazioni di riconoscimento altrimenti difficoltose. Il riconoscimento facciale comprende due tipi di operazioni, ovvero:

- *la verifica*: una procedura con la quale si cerca di determinare l'appartenenza di un volto a una certa identità;
- *l'identificazione*: una procedura che permette di individuare una persona tra più possibili identità; e un suo possibile utilizzo riguarda l'analisi, da parte delle forze dell'ordine, di un volto su un dataset per poterlo identificare.

## 4.2 Brevi cenni storici

Il riconoscimento facciale ha le sue origini negli anni '60 e '70, quando furono avviati i primi studi per creare un sistema automatizzato di riconoscimento dei volti. In questo periodo, gli scienziati svilupparono i primi algoritmi per tracciare i tratti facciali, come la distanza tra occhi, naso e bocca, basandosi su immagini bidimensionali. In particolare Woody Bledsoe<sup>1</sup> sviluppò un sistema nel quale un operatore doveva misurare le distanze tra i diversi tratti del viso nelle fotografie stampate e le inseriva in un software che le confrontava con tutte le immagini presenti in un database in modo da identificare la persona. Questo procedimento voleva dimostrare come le caratteristiche del volto, quali la distanza degli occhi, del naso, della bocca e i lineamenti sottili, possano essere un valido sistema per identificare uno specifico individuo. Tuttavia, la potenza di calcolo dell'epoca era limitata, e questi primi sistemi erano lenti e poco precisi. Successivamente, il decennio degli anni '90 ha visto un notevole miglioramento delle capacità di elaborazione dei computer e l'introduzione di tecniche avanzate di riconoscimento delle immagini. Nel 1991, Matthew Turk e Alex Pentland del MIT svilupparono il sistema Eigenfaces, che rappresentò una svolta significativa nel campo del riconoscimento facciale. Tale sistema utilizzava una tecnica statistica chiamata "Principal Component Analysis" (PCA) per ridurre la complessità delle immagini facciali e confrontarle in modo più efficiente. Il modello estraeva le varie caratteristiche di un volto, le codificava e le comparava con le altre caratteristiche contenute in un database. Purtroppo, l'estrazione delle caratteristiche veniva fatta in maniera globale, tralasciando aspetti facciali più dettagliati. Questo modello, infatti, era molto interessante dal punto di vista tecnico ma presentava delle difficoltà quando erano presenti variazioni facciali differenti rispetto alle informazioni acquisite. Con l'inizio del nuovo millennio, il riconoscimento facciale ha iniziato a essere utilizzato in contesti più pratici, grazie ai progressi dell'hardware e delle reti neurali. Le forze dell'ordine e gli aeroporti hanno iniziato a impiegare sistemi di riconoscimento facciale per scopi di sicurezza, e le aziende tecnologiche hanno cominciato a incorporare questa tecnologia nei loro prodotti. Negli anni 2010, il riconoscimento facciale ha fatto un balzo in avanti con l'avvento del Deep Learning. Le reti neurali convoluzionali (Convolutional Neural Network - CNN) hanno permesso di migliorare notevolmente la precisione e la velocità del riconoscimento facciale. Un sistema innovativo fu quello proposto da parte di ricercatori di Facebook chiamato DeepFace. Questo sistema consiste in un modello di rete neurale convoluzionale che si basa su oltre 4.4 milioni di immagini di 4030 diverse persone; esso ha raggiunto il miglior livello di accuratezza sul dataset LFW (Labeled Faces in the Wild). Nel 2015, il Visual Geometry Group dell'Università di Oxford ha sviluppato il sistema VGGFace grazie alla raccolta di 2.6 milioni di immagini appartenenti a 2600 soggetti. Tale modello ha livelli di accuratezza ancora superiori a DeepFace. Nel corso degli anni 2020, il riconoscimento facciale è diventato onnipresente in molti settori, dalle applicazioni per smartphone alle telecamere di sorveglianza. Tuttavia, questo ha sollevato preoccupazioni significative in termini di privacy e diritti civili, portando a dibattiti legislativi e restrizioni in diversi paesi. Alcune città e stati hanno addirittura vietato l'uso del riconoscimento facciale da parte delle forze dell'ordine.

## 4.3 Come funziona il riconoscimento facciale

Il sistema di riconoscimento facciale è un modo sofisticato per verificare o accertare l'identità di qualcuno utilizzando un algoritmo che elabora un'immagine digitale o un fotogramma

---

<sup>1</sup>Bledsoe è stato un matematico, informatico ed eminente educatore americano. È uno dei fondatori dell'Intelligenza Artificiale avendo apportato i primi contributi al riconoscimento di modelli e al riconoscimento facciale.

video, seleziona le caratteristiche distintive del volto di qualcuno mostrato in un'immagine, e le abbina ai volti già registrati in un database. Il riconoscimento facciale sta crescendo in popolarità, con nuovi usi che vengono sviluppati continuamente. Le immagini digitali e video stanno diventando più chiare e nitide; di conseguenza diventa più facile individuare persone e volti distinti; allo stesso tempo il software e gli algoritmi di corrispondenza beneficiano di maggiori fonti di dati e accuratezza. Per abbinare i modelli del volto, le reti neurali convoluzionali elaborano ciascuna immagine attraverso diversi passaggi:

- *Rilevamento dei volti*: il primo passaggio consiste nell'acquisire un volto all'interno di un'immagine, un flusso video o una scena più grande. Questo processo prevede la distinzione dei tratti del viso dall'ambiente circostante e l'individuazione della loro posizione all'interno dell'inquadratura. Così avviene attraverso algoritmi di visione artificiale progettati per individuare la presenza di un volto tra gli oggetti presenti nell'immagine. Questi algoritmi analizzano l'immagine pixel per pixel per identificare le caratteristiche tipiche di un viso umano, come la forma o la simmetria. L'immagine acquisita viene spesso migliorata o normalizzata per garantire una qualità adeguata (ad esempio, regolazione della luminosità, contrasto, riduzione del rumore).
- *Analisi del volto*: una volta rilevato un volto, la tecnologia analizza le caratteristiche del viso. Questa analisi si basa tipicamente sulla geometria del viso, misurando vari punti chiave su di esso, noti come punti di riferimento o punti nodali, che possono includere la distanza tra gli occhi, la forma della mascella e i contorni degli zigomi, delle labbra e del naso.
- *Estrazione delle caratteristiche*: dopo l'allineamento, il sistema estrae le caratteristiche distintive del volto, come la distanza tra gli occhi, la lunghezza del naso, la forma del mento e altre misure uniche. Questa fase trasforma l'immagine del volto in una rappresentazione numerica, nota come "firma" o "impronta facciale", ovvero una mappa digitale della geometria del viso. Questa firma è un codice numerico distintivo che incapsula le caratteristiche facciali archiviate nel database.
- *Confronto*: la firma facciale ottenuta viene confrontata con quelle presenti in un database. Il confronto può avvenire in due modalità. La prima modalità prevede l'identificazione, ovvero il sistema confronta la firma facciale con un database di firme conosciute per trovare la corrispondenza più vicina (metodo usato, per esempio, per identificare una persona sconosciuta). La seconda modalità prevede la verifica, ovvero la firma facciale viene confrontata con una firma specifica (ad esempio, quella registrata per un account) per verificare se la persona è chi afferma di essere. Questo metodo è comune nei sistemi di autenticazione, come lo sblocco di un telefono tramite il volto.
- *Decisione*: dopo il confronto, il sistema prende una decisione basata sul livello di somiglianza tra la firma facciale acquisita e quella nel database. Se la somiglianza supera una certa soglia predefinita, il sistema conferma l'identità o verifica la persona; altrimenti, l'identificazione o la verifica fallisce.
- *Feedback e aggiornamento*: alcuni sistemi di riconoscimento facciale possono anche aggiornare i loro modelli e database con nuove immagini del volto, migliorando la precisione nel tempo. Questo apprendimento continuo aiuta il sistema ad adattarsi a cambiamenti nel volto dell'utente, come invecchiamento o variazioni di luce.

Il funzionamento consiste, quindi, nel fatto che quando la tecnologia di riconoscimento facciale rileva una nuova immagine, confronta il nuovo modello di volto con le firme facciali all'interno di un database esistente di volti noti. Utilizzando una complessa gamma

di algoritmi di Intelligenza Artificiale, il sistema valuta il modello del volto con velocità e precisione impressionanti per accertare se esiste una corrispondenza con qualsiasi firma facciale memorizzata. Una rete neurale convoluzionale converte ogni modello facciale in un codice numerico con ogni modello espresso come vettore numerico. Più due vettori sono vicini tra loro, maggiore è la probabilità che vi sia una corrispondenza. La chiave del successo in evoluzione della tecnologia di riconoscimento facciale risiede nei progressi nell'apprendimento automatico, che perfezionano il processo di creazione e confronto dei modelli di volti. Man mano che il sistema impara da ogni volto che elabora, aumenta progressivamente la sua precisione, rafforzando il ruolo centrale della tecnologia in aree come la sicurezza, l'autenticazione personale e varie applicazioni nel nostro ecosistema digitale.

## 4.4 Vantaggi

Il riconoscimento facciale è diventato ormai una tecnologia utilizzata giornalmente da tantissime persone; si tratta di una tecnologia molto sofisticata utilizzata in diversi ambiti, tra cui la sicurezza nei luoghi affollati, come centri commerciali o aeroporti. L'uso del riconoscimento facciale e dei dati sta diventando sempre più una parte essenziale nella vita di tutti i giorni, come, per esempio, nell'identificazione commerciale, contribuendo a raggiungere gli individui e a personalizzare i messaggi di vendita e di marketing. Il riconoscimento facciale è sempre più utilizzato anche nell'indicizzazione automatica delle immagini, nelle interazioni uomo-tecnologia e nei sistemi di videosorveglianza. Infatti, i progressi tecnologici avvengono costantemente, con aggiornamenti significativi, e per questo, ci sono molti vantaggi nel software di riconoscimento facciale. Le tecniche di riconoscimento facciale basate su Deep Learning migliorano nel tempo rendendo i sistemi più precisi e adattabili. Di seguito una lista di alcuni dei vantaggi più importanti che può portare l'utilizzo del riconoscimento facciale.

- *Rapido processo di identificazione*: il riconoscimento facciale può identificare o verificare l'identità di una persona in pochi secondi, rendendo più efficienti le operazioni che richiedono l'identificazione, come l'accesso a edifici o dispositivi; in particolare, esso può essere di grande aiuto per una maggiore sicurezza. Il riconoscimento facciale, infatti, rende più difficile per le persone non autorizzate accedere a dispositivi o aree riservate, riducendo il rischio di frodi e accessi non autorizzati. Inoltre, esso fornisce un grande supporto alla sorveglianza, aiutando, così, le forze dell'ordine. Queste ultime possono utilizzare il riconoscimento facciale per monitorare in tempo reale grandi folle o aree pubbliche, migliorando la capacità di individuare persone ricercate o prevenire crimini. Inoltre, il riconoscimento facciale è conveniente in quanto non c'è bisogno di nessun contatto fisico, a differenza di altre forme di identificazione biometrica, come le impronte digitali. Ciò non solo lo rende più igienico, ma anche più comodo per gli utenti.
- *Automazione delle attività*: il riconoscimento facciale riduce la necessità di intervento umano in compiti di sicurezza e autenticazione, automatizzando processi che altrimenti richiederebbero tempo, come il controllo dei passaporti negli aeroporti. Esso, inoltre, permette una possibile riduzione degli errori umani. Infatti, rispetto ai sistemi manuali, il riconoscimento facciale riduce gli errori di identificazione che possono derivare dalla distrazione o dall'errore umano, come lo scambio di persone o il mancato riconoscimento di un volto noto.
- *Facilità d'uso*: per gli utenti, il riconoscimento facciale è una tecnologia semplice e naturale. Non richiede l'inserimento di password o PIN, né l'uso di dispositivi fisici, come carte o chiavi.

- *Esperienze personalizzate*: nelle applicazioni di marketing e retail, il riconoscimento facciale consente alle aziende di offrire esperienze e proposte personalizzate ai clienti, migliorando il coinvolgimento e la soddisfazione.
- *Supporto per persone con disabilità*: il riconoscimento facciale può essere utilizzato per facilitare l'accesso a dispositivi e sistemi da parte di persone con disabilità, eliminando la necessità di interagire fisicamente con i dispositivi.
- *Riduzione delle barriere linguistiche*: non richiedendo l'uso di comandi vocali o scritti, il riconoscimento facciale può essere utilizzato da persone di diverse lingue e culture senza difficoltà di comprensione.
- *Integrazione con altre tecnologie*: il riconoscimento facciale può essere integrato con altre forme di sicurezza e autenticazione, come i codici QR o le impronte digitali, per creare un sistema multi-fattoriale ancora più sicuro.

## 4.5 Esempi di utilizzo, applicazioni

Il riconoscimento facciale ha trovato una vasta gamma di applicazioni in diversi settori grazie alla sua capacità di identificare o verificare l'identità delle persone in modo rapido ed efficiente. Ecco alcune delle principali applicazioni di questa tecnologia:

- *Sicurezza e sorveglianza*: le agenzie di polizia utilizzano il riconoscimento facciale per identificare sospetti in tempo reale, confrontando le immagini dei volti con i database di criminali conosciuti. Questa tecnologia viene spesso utilizzata in luoghi pubblici, come aeroporti, stazioni ferroviarie e grandi eventi per monitorare e prevenire attività illegali. Inoltre, gli aeroporti e i posti di frontiera impiegano il riconoscimento facciale per verificare l'identità dei viaggiatori, accelerando i processi di immigrazione e migliorando la sicurezza.
- *Autenticazione e accesso ai dispositivi*: molti smartphone e computer utilizzano il riconoscimento facciale per essere sbloccati. Questa funzione offre una forma di autenticazione biometrica rapida e conveniente. Oltre a ciò, in ambienti aziendali o istituzionali il riconoscimento facciale viene utilizzato per controllare l'accesso a edifici, stanze o sistemi informatici, sostituendo o integrando le tradizionali carte di accesso o password.
- *Pubblicità personalizzata*: alcune aziende utilizzano il riconoscimento facciale per analizzare le caratteristiche demografiche e le emozioni dei clienti mentre guardano un display pubblicitario. Questo consente di adattare i messaggi pubblicitari in tempo reale in base al profilo del cliente. Quindi i negozi e le catene di retail possono impiegare il riconoscimento facciale per riconoscere i clienti abituali e offrire loro un servizio personalizzato, come sconti esclusivi o raccomandazioni di prodotti basate su acquisti precedenti.
- *Sanità*: negli ospedali, il riconoscimento facciale può essere utilizzato per semplificare l'assistenza ai pazienti. Gli operatori sanitari stanno sperimentando l'uso del riconoscimento facciale per accedere alle cartelle cliniche dei pazienti, migliorando l'efficienza del trattamento e riducendo gli errori. Il riconoscimento facciale può anche essere utilizzato in alcune strutture sanitarie per monitorare i pazienti e garantire che ricevano il trattamento corretto, specialmente in contesti di assistenza a lungo termine.
- *Educazione*: nelle scuole e nelle università, il riconoscimento facciale può essere utilizzato per monitorare la presenza degli studenti e migliorare la sicurezza. Inoltre in alcune

istituzioni, il riconoscimento facciale viene sperimentato per analizzare il comportamento e l'attenzione degli studenti durante le lezioni, aiutando gli insegnanti a migliorare le tecniche di insegnamento.

- *Settore bancario e finanziario*: le banche e i servizi finanziari utilizzano il riconoscimento facciale per autenticare i clienti durante le transazioni online o negli sportelli automatici, riducendo il rischio di frodi. Anche durante l'apertura di conti bancari online, il riconoscimento facciale può essere usato per verificare l'identità dei nuovi clienti confrontando la loro immagine con i documenti di identità caricati.

## 4.6 Conseguenze etiche

Le criticità relative al riconoscimento facciale rientrano in uno dei quesiti più importanti dell'agenda politica internazionale e cioè come rendere l'IA più etica. Infatti ci si chiede se la sorveglianza sia una violazione della sfera privata oppure si tratta di un bene per la comunità. Ecco alcuni aspetti che potrebbero portare a delle ripercussioni negative all'utilizzo del riconoscimento facciale se fatto in maniera sbagliata:

- *Privacy e sorveglianza*: nonostante renda difficile il furto di dati personali, questa tecnologia espone gli utenti a qualche rischio legato alla privacy. Difatti, il riconoscimento facciale permette al fornitore del servizio di raccogliere e archiviare molte informazioni sensibili sul singolo utente; si tratta, inoltre, di informazioni "fisiche", che lo rendono inequivocabilmente riconoscibile e individuabile. Inoltre, l'uso estensivo del riconoscimento facciale in spazi pubblici può essere percepito come invasivo, contribuendo alla sorveglianza di massa e limitando la privacy. Infatti, la raccolta e la conservazione di questi dati facciali può sollevare preoccupazione sulla gestione sicura e sull'uso futuro di queste informazioni. Inoltre, il riconoscimento può essere utilizzato da governi e autorità pubbliche a scopo di sorveglianza e controllo, portando a una diminuzione della privacy e della libertà individuale. In molti paesi le leggi che regolano la gestione di questa tecnologia sono insufficienti, creando un vuoto normativo. Questi dati raccolti potrebbero essere, quindi, utilizzati da governi o aziende per fini non etici, come la discriminazione, il controllo sociale o la sorveglianza senza consenso.
- *Opacità dei sistemi*: molte tecnologie di riconoscimento facciale sono sviluppate da aziende private che non rivelano completamente il funzionamento dei loro algoritmi o le fonti dei dati utilizzati. Questo rende difficile per i cittadini e i regolatori valutare e comprendere l'impatto della tecnologia. Inoltre, se un sistema di riconoscimento facciale causa un danno, ad esempio un arresto ingiusto, è spesso difficile attribuire responsabilità specifiche sia alle aziende che sviluppano la tecnologia sia alle istituzioni che la utilizzano.

I problemi etici del riconoscimento facciale richiedono una riflessione approfondita e una regolamentazione rigorosa. È fondamentale che l'uso di questa tecnologia sia trasparente, equo e rispettoso dei diritti fondamentali delle persone. Senza un quadro normativo adeguato, esiste il rischio che il riconoscimento facciale venga utilizzato in modi che possono compromettere gravemente la privacy, la libertà e l'uguaglianza nella società. Nonostante ciò, i vantaggi di questa forma di riconoscimento biometrico sono indubbi. In più, gli individui possono contare su leggi e buone pratiche per tutelare al meglio i propri dati.

## 4.7 Amazon Rekognition

### 4.7.1 Cos'è

Amazon Rekognition è un software cloud-based che mette a disposizione capacità di visione artificiale pre-addestrate e personalizzabili per estrarre informazioni dettagliate da immagini e video. Esso semplifica l'integrazione di analisi di immagini e video nelle applicazioni. Basta fornire un'immagine o un video all'API di Amazon Rekognition, e il servizio offre diverse funzionalità, ovvero:

- identificazione di etichette (oggetti, concetti, persone, scene e attività) e testo;
- rilevamento di contenuti inappropriati;
- analisi facciale altamente accurata, con capacità di confronto e ricerca facciale.

Le API di riconoscimento facciale di Amazon Rekognition consentono di rilevare, analizzare e confrontare volti per vari casi d'uso come la verifica dell'utente, la catalogazione, il conteggio delle persone e la sicurezza pubblica. Il servizio viene offerto ai clienti anche senza competenze specifiche di Machine Learning ed è in grado di svolgere innumerevoli compiti in base allo scenario in cui viene impiegato. Alcuni di questi compiti sono i seguenti:

- *Rilevare oggetti*: Rekognition riesce ad identificare qualsiasi tipologia di oggetto all'interno di un'immagine oppure di un video. Dopo aver riconosciuto e assegnato un punteggio che rappresenta il livello di affidabilità, esso permette di eseguire ricerche e di filtrare i risultati a partire dall'output generato.
- *Riconoscere i volti, analisi e controllo facciale*: Rekognition è in grado di riconoscere volti umani e creare un indice delle persone che presentano determinate caratteristiche come gli occhi aperti oppure il sorriso. Tali immagini possono essere utilizzate per calcolare la probabilità che un volto in immagini diverse appartenga alla stessa persona. È basato su un punteggio di somiglianza che identifica ogni utente in ciascuna foto in tempo reale. La tecnologia di riconoscimento facciale di Amazon risulta ad oggi una delle più avanzate del mondo.
- *Rilevamento di immagini inappropriate*: Rekognition è in grado di capire se, all'interno di una determinata immagine o di un determinato video sono presenti determinati contenuti espliciti inappropriati, così da segnalarli ed eventualmente rimuoverli.

I servizi offerti da Amazon Rekognition si basano su due API, Rekognition Image e Rekognition Video. Rekognition Image è un servizio di riconoscimento delle immagini che permette di rilevare oggetti, scene, punti di riferimento, volti, colori dominanti e qualità dell'immagine. Rekognition Image può, inoltre, estrarre il testo, riconosce le celebrità e identificare i contenuti inappropriati nelle immagini. Esso consente anche di ricercare e confrontare i volti. Rekognition Video è un servizio di riconoscimento video che permette di rilevare le attività, interpretare i movimenti delle persone in un filmato e riconosce oggetti, personaggi famosi e contenuti inappropriati nei video archiviati in Amazon S3<sup>2</sup> o nei flussi in diretta. Rekognition Video rileva le persone e le segue nel video anche quando i loro volti non sono visibili e quando la persona esce o entra nell'inquadratura. Per quanto riguarda le immagini, Amazon Rekognition supporta i formati .jpg e .png; in alternativa, le immagini possono essere caricate come oggetti di S3. La dimensione massima che l'immagine può avere è di 15 MB, se caricata come oggetto di S3, mentre è di 5 MB negli altri casi. Riguardo i video,

<sup>2</sup>Amazon S3 è un servizio di archiviazione che offre sicurezza, prestazioni e una grande disponibilità di dati.

invece, Amazon Rekognition Video supporta file .mp4, .avi e .mov, la cui dimensione massima deve essere di 10 GB; mentre, se il video viene caricato come oggetto di S3, esso può essere lungo fino a 6 ore.

## 4.8 Test con Amazon Rekognition

I case study presi in considerazione in questo capitolo rientreranno nell'ambito dell'identificazione tramite riconoscimento facciale. I servizi di Amazon Rekognition scelti per questo studio sono, in particolare, l'analisi facciale e, di conseguenza, l'analisi delle espressioni facciali. Verranno studiate diverse immagini con un viso più o meno riconoscibile.

### 4.8.1 Analisi facciale

Questa funzione di Amazon Rekognition permette di fornire in input il volto di una persona, e consente di identificare caratteristiche e attributi del viso, come:

- *età stimata*: la stima dell'età di una persona;
- *genere*: maschile o femminile, con un livello di confidenza;
- *barba, baffi, occhiali*: presenza o assenza di barba, baffi e occhiali.

In questa sezione utilizzeremo varie foto che ritraggono varie persone differenti e in tutti i casi valuteremo come il software si comporta. Nella prima immagine (Figura 4.1) che ritrae una donna senza evidenti segni particolari e con il viso completamente frontale in cui si vedono tutte le caratteristiche, Rekognition è stato in grado di riconoscere il volto umano. Ha, quindi, analizzato il volto fornendo informazioni accurate riguardo al soggetto, identificandolo come una donna tra i 21 e i 27 anni con un'espressione tranquilla. L'unica difficoltà che forse Rekognition ha riscontrato è capire se la donna stia sorridendo o meno; infatti la piattaforma ha impostato una confidenza del 99,1% che non stia sorridendo, quando è evidente che sul viso ci sia un accenno di sorriso.

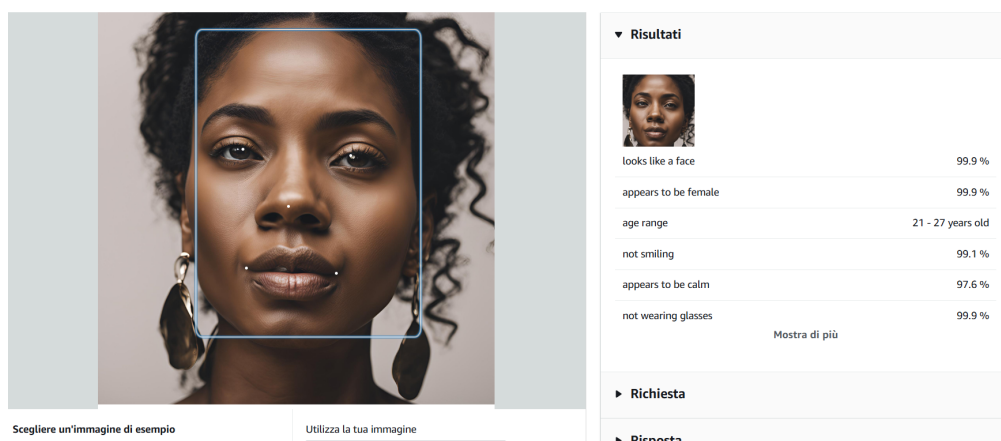
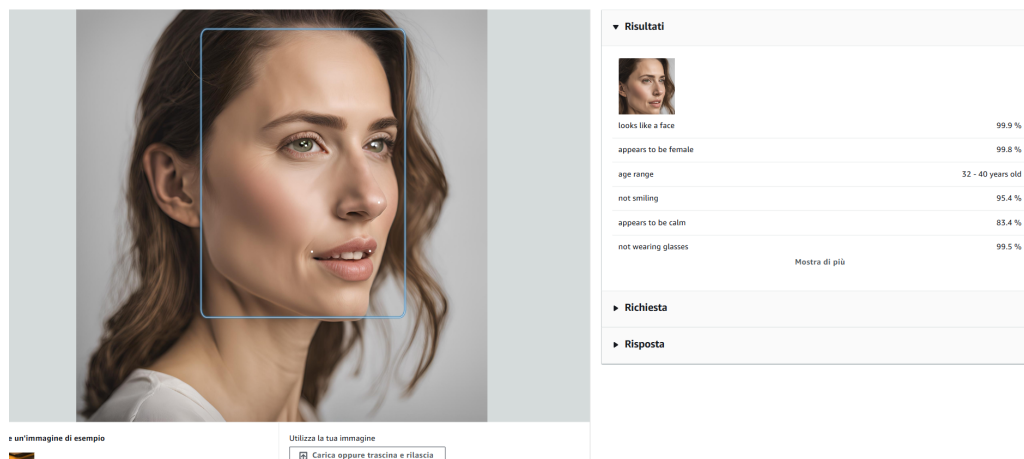


Figura 4.1: Risultato dell'analisi facciale del volto di una donna

Testando ulteriormente la piattaforma abbiamo utilizzando un'immagine che ritrae una donna ma con il volto girato di tre quarti (Figura 4.2). Anche in questo caso Amazon Rekognition riesce a dedurre in pochissimo tempo età, sesso, espressione facciale, fornendoci in output tutte le caratteristiche di cui abbiamo bisogno.



**Figura 4.2:** Risultato dell'analisi facciale del volto di una donna a tre quarti

Nella Figura 4.3 si mostra come inserendo in input un'immagine, in output il sistema ci fornisce un codice JSON in cui si evidenziano tutte le caratteristiche che si riescono ad evincere dal volto nella foto, come:

- età;
- se sorride;
- se indossa gli occhiali;
- se indossa gli occhiali da sole;
- sesso;
- presenza di barba;
- presenza di baffi;
- se gli occhi sono aperti;
- se la bocca è aperta;

Si evidenzia come, in ogni caso, il valore viene impostato a "false" nel momento in cui non c'è la corrispondenza invece viene settato a "true" nel momento in cui c'è corrispondenza. In entrambi i casi viene evidenziata la confidenza con la quale la piattaforma è sicura di aver interpretato bene la foto e aver dedotto le giuste caratteristiche. In questo caso abbiamo tutti i valori del parametro "confidence" elevati, il che significa che il software è riuscito a riconoscere e a distinguere in maniera ottima i parametri del volto. Rekognition, inoltre, individua la posizione esatta in cui si trovano bocca, naso e occhi nella foto.

Infine viene studiata (Figura 4.4) la posa del viso, dove vengono evidenziati in dettaglio le inclinazioni del viso; queste possono essere:

- *Pitch*: indica l'inclinazione della testa rispetto all'asse orizzontale. Se il valore è positivo, la testa è inclinata verso l'alto (guarda in su). Se il valore è negativo, la testa è inclinata verso il basso (guarda in giù).
- *Yaw*: descrive la rotazione della testa a destra o sinistra lungo l'asse verticale. Se il valore è positivo la testa è girata verso destra, mentre se il valore è negativo la testa è girata verso sinistra.

```

"AgeRange": {
  "Low": 32,
  "High": 40
},
"Smile": {
  "Value": false,
  "Confidence": 95.43311309814453
},
"Eyeglasses": {
  "Value": false,
  "Confidence": 99.54875183105469
},
"Sunglasses": {
  "Value": false,
  "Confidence": 99.36775970458984
},
"Gender": {
  "Value": "Female",
  "Confidence": 99.86238098144531
},
"Beard": {
  "Value": false,
  "Confidence": 99.17737579345703
},
"Mustache": {
  "Value": false,
  "Confidence": 99.90746307373047
},
"EyesOpen": {
  "Value": true,
  "Confidence": 99.91908264160156
},
"MouthOpen": {
  "Value": true,
  "Confidence": 76.22862243652344
}
}

```

**Figura 4.3:** File JSON relativo all'analisi del volto di una donna a tre quarti

- *Roll*: indica l'inclinazione della testa rispetto all'asse frontale (inclinazione laterale). Se il valore è positivo la testa è inclinata verso la propria spalla destra; mentre se il valore è negativo la testa è inclinata verso la spalla sinistra.

```

  "pose": {
    "Roll": 0.8767149448394775,
    "Yaw": 36.8218994140625,
    "Pitch": -0.3158908486366272
  },
  "quality": {
    "Brightness": 80.98062896728516,
    "Sharpness": 96.61495208740234
  },
  "confidence": 99.99976348876953,
  "faceOccluded": {
    "Value": false,
    "Confidence": 99.7945327758789
  },
  "eyeDirection": {
    "Yaw": 12.67602825164795,
    "Pitch": -2.7077572345733643,
    "Confidence": 99.99535369873047
  }
}
}
}

```

**Figura 4.4:** Parte finale del file JSON relativo all'analisi precedente

Nel codice, il parametro *quality* (qualità) fa riferimento a due metriche che descrivono la qualità dell'immagine del volto analizzato; tali metriche sono:

- *Brightness* (Luminosità): questo valore indica quanto l'immagine è luminosa. Un livello di luminosità ottimale aiuta Amazon Rekognition a rilevare correttamente i dettagli facciali. Un valore alto significa che l'immagine è molto luminosa, mentre un valore basso indica che l'immagine è scura.

- *Sharpness* (Nitidezza): esso misura quanto è nitida l'immagine, ossia quanto i bordi e i dettagli del volto sono ben definiti. Se il valore è alto l'immagine è molto nitida, con dettagli chiari e ben definiti, mentre se il valore è basso l'immagine è sfocata o ha una bassa risoluzione.

In questo caso abbiamo il valore di "Brightness" pari a circa 80; quindi, l'immagine ha un buon livello di luminosità. Invece, il valore di "Sharpness" è di circa 96, ovvero l'immagine è molto nitida, il che suggerisce che i dettagli del volto sono ben definiti. Il parametro "Confidence" (confidenza) rappresenta il livello di certezza di Amazon Rekognition riguardo a una particolare caratteristica o ad un particolare attributo rilevati nell'immagine. È espresso come percentuale e indica quanto il modello è sicuro della correttezza dell'output che ha generato. Inoltre nel codice vengono studiati anche i punti di riferimento (landmark), sono coordinate che indicano la posizione di caratteristiche facciali chiave all'interno di un volto rilevato. Tali punti vengono utilizzati per individuare elementi specifici del volto, come occhi, naso, bocca, etc. I landmark possono essere utili in varie applicazioni, come l'allineamento del volto, il tracciamento delle espressioni facciali o l'analisi delle pose (Figura 4.5). I landmark più tipici sono i seguenti:

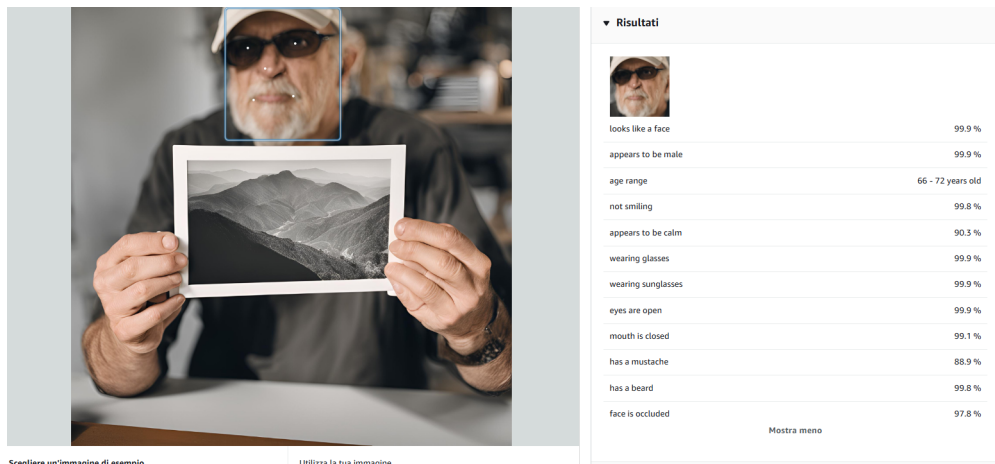
- *eyeLeft*: coordinate del centro dell'occhio sinistro;
- *eyeRight*: coordinate del centro dell'occhio destro;
- *nose*: coordinate del centro del naso;
- *mouthLeft*: coordinate dell'angolo sinistro della bocca;
- *mouthRight*: coordinate dell'angolo destro della bocca;
- *leftEyeBrowLeft*: coordinate del punto più esterno del sopracciglio sinistro;
- *rightEyeBrowRight*: coordinate del punto più esterno del sopracciglio destro;
- *leftPupil*: coordinate della pupilla sinistra;
- *rightPupil*: coordinate della pupilla destra.

```
"Landmarks": [  
  {  
    "Type": "eyeLeft",  
    "X": 0.5583330988883972,  
    "Y": 0.32945647835731506  
  },  
  {  
    "Type": "eyeRight",  
    "X": 0.729461133480072,  
    "Y": 0.32968249917030334  
  },  
  {  
    "Type": "mouthLeft",  
    "X": 0.5726242661476135,  
    "Y": 0.5869912505149841  
  },  
  {  
    "Type": "mouthRight",  
    "X": 0.4273757338523865,  
    "Y": 0.5869912505149841  
  }  
]
```

**Figura 4.5:** Landmark

Per interpretare i Landmarks si usano le coordinate X e Y ed essi indicano la posizione relativa di ogni punto di riferimento. I valori sono normalizzati da 0 a 1 e rappresentano una percentuale rispetto alla larghezza e all'altezza dell'immagine.

Abbiamo cercato di testare ulteriormente Amazon Rekognition utilizzando una foto in cui in primo piano non troviamo il volto di una persona, ma una foto di un paesaggio (Figura 4.6). Infatti, nella figura possiamo notare come il viso dell'uomo, oltre ad essere sfocato, è in secondo piano rispetto ad una foto che tiene in mano. Risulta, quindi, difficile studiare bene tutte le sfaccettature del viso, e ricevere, di conseguenza, tutti i dati di cui abbiamo bisogno, come nel caso dei visi delle due donne. In questo caso, quindi, per il software dovrebbe essere più complicato andare a estrarre tutte le caratteristiche del viso. Ma anche in questo caso Amazon Rekognition non ha avuto nessuna difficoltà nel riconoscere l'uomo. Lo identifica, infatti, come un uomo tra i 66 e i 72 anni, con barba e baffi e con gli occhiali.



**Figura 4.6:** Analisi facciale del volto di un uomo

#### 4.8.2 Analisi espressioni facciali e deduzione emozioni

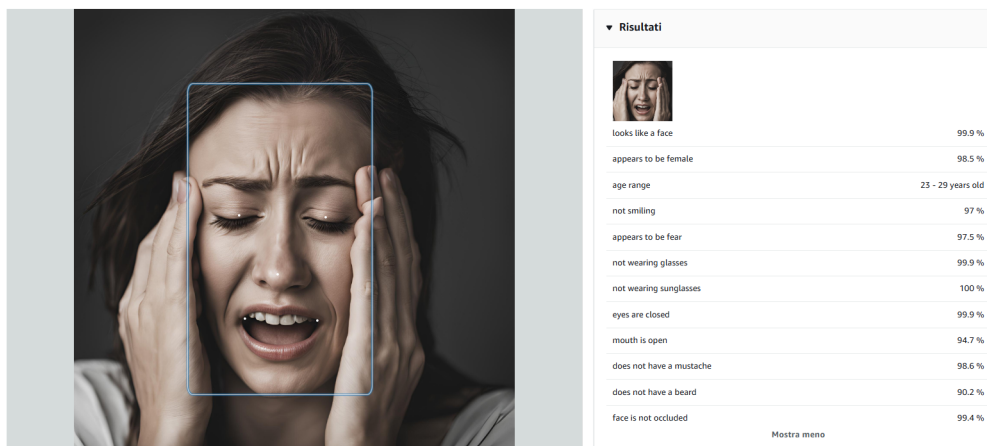
Tutta l'analisi del volto effettuata in precedenza viene utilizzata da Amazon Rekognition per analizzare i visi dei soggetti al fine di dedurre le espressioni facciali e, di conseguenza, le emozioni che stanno provando. Questa funzionalità consente di riconoscere emozioni e stati d'animo come felicità, tristezza, rabbia, sorpresa, etc, semplicemente analizzando un'immagine o un video. Amazon Rekognition può rilevare le seguenti emozioni:

- happy (felicità);
- sad (tristezza);
- angry (rabbia);
- confused (confusione);
- disgusted (disgusto);
- surprised (sorpresa);
- calm (calma);
- fear (paura).

Riesce a capire queste emozioni grazie, prendendo in considerazione la posizione degli occhi, se gli occhi sono aperti o chiusi, se la bocca è aperta o chiusa, se la persona sta sorridendo o no. Quando analizza un volto, assegna un punteggio di probabilità (confidenza) a ciascuna emozione rilevata. Quindi, ogni emozione rilevata viene accompagnata da un valore numerico che rappresenta il livello di confidenza espresso in percentuale (0-100%). Ad esempio, se Rekognition rileva "felicità" con un valore di confidenza del 95%, significa che il sistema è sicuro al 95% che l'emozione predominante sia la felicità. Questo permette agli sviluppatori di utilizzare il servizio in modo flessibile. Ad esempio, si può decidere di considerare solo le emozioni con un valore di confidenza superiore a una certa soglia (come il 75%) per evitare falsi positivi o interpretazioni errate. Sebbene Amazon Rekognition sia molto potente, non è perfetto. Il riconoscimento delle emozioni è un processo complesso e può essere influenzato da vari fattori come la qualità dell'immagine, le condizioni di illuminazione e le espressioni facciali che potrebbero essere ambigue. L'analisi delle immagini prevede i seguenti passi:

- *caricamento dell'immagine o video*: l'immagine o il video contenente il volto viene caricato e analizzato dal servizio di Amazon Rekognition;
- *rilevamento del volto*: Rekognition rileva automaticamente uno o più volti nella foto o nel video;
- *analisi degli attributi*: il sistema rileva attributi del volto, tra cui le emozioni, il sorriso, la direzione dello sguardo, la presenza di occhiali, ecc;
- *valutazione della confidenza*: ogni emozione rilevata ha un valore di Confidence, che indica la probabilità che l'emozione identificata sia corretta.

Per testare Amazon Rekognition e verificare se riesce a analizzare e a decifrare in maniera corretta le emozioni, abbiamo considerato un'immagine raffigurante il volto di una ragazza in evidente stato di paura (Figura 4.7).



**Figura 4.7:** Analisi delle emozioni tramite Amazon Rekognition

In questo esempio possiamo notare come Amazon Rekognition sia riuscito a dedurre in primis età e sesso della persona; successivamente ha evidenziato il fatto che gli occhi sono chiusi, la bocca è aperta e non sta sorridendo, il software riesce a capire che la donna si trova in uno stato di paura. Questa analisi dell'emozione paura ha una confidenza del 97.5, quindi molto alta, ovvero la piattaforma è fortemente sicura dell'emozione che ha stabilito. Analizzando il codice (Figura 4.8) si nota che su ogni emozione studiata da Rekognition abbiamo a fianco la confidenza che la piattaforma associa.

```

    },
    "Emotions": [
      {
        "Type": "FEAR",
        "Confidence": 97.55859375
      },
      {
        "Type": "SAD",
        "Confidence": 4.217529296875
      },
      {
        "Type": "ANGRY",
        "Confidence": 0.06814002990722656
      },
      {
        "Type": "CONFUSED",
        "Confidence": 0.03528594970703125
      },
      {
        "Type": "SURPRISED",
        "Confidence": 0.021129846572875977
      },
      {
        "Type": "DISGUSTED",
        "Confidence": 0.017571449279785156
      },
      {
        "Type": "CALM",
        "Confidence": 0.004804134368896484
      },
      {
        "Type": "HAPPY",
        "Confidence": 0.00004172325134277344
      }
    ]
  }

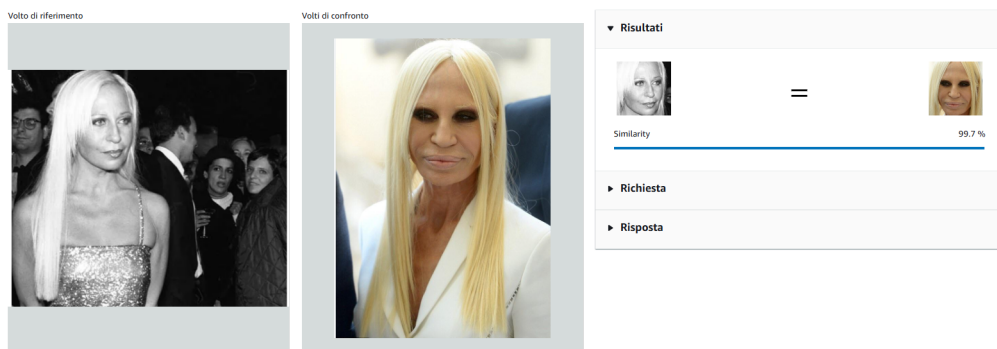
```

**Figura 4.8:** File JSON relativo alle emozioni

Infatti, riguardo alle emozioni come felicità, confusione e anche calma, Amazon Rekognition associa ad ognuna di esse una confidenza intorno allo zero. Questo ci fa notare come esso abbia capito che queste emozioni non sono assolutamente presenti nel volto della donna. Mentre, riguardo all'emozione tristezza, Amazon Reckongition associa un valore di confidenza di circa il 4% molto inferiore al valore di confidenza associato alla paura.

### 4.8.3 Confronto facciale

Successivamente abbiamo testato ulteriormente Amazon Rekognition mettendolo in difficoltà, utilizzando una stessa persona, ovvero Donatella Versace in due periodi diversi della sua vita, ovvero da giovane e ora. Infatti, a causa di interventi, il suo volto è cambiato. Ma anche in questo caso Amazon Rekognition non ha trovato nessuna difficoltà nel riconoscere che, in entrambe le immagini, sia presente la stessa persona (Figura 4.9).



**Figura 4.9:** Confronto facciale

Amazon Rekognition esegue un'analisi delle caratteristiche facciali, come la forma del viso, il posizionamento degli occhi, il naso, la bocca e altri dettagli distintivi e restituisce un punteggio di similarità che indica la probabilità che i volti nelle due immagini appartengano alla stessa persona. Il punteggio varia da 0 a 100, dove un punteggio più alto indica una maggiore somiglianza. In questo caso abbiamo una somiglianza del 99,7%, ovvero Amazon non ha riscontrato nessuna difficoltà nel dedurre che si tratta della stessa persona, pur non avendo gli stessi tratti del viso. Studiando il codice (Figura 4.10) troviamo l'attributo "SourceImageFace" che specifica le coordinate che delimitano il volto nell'immagine di riferimento. Esso definisce l'area in cui è stato rilevato il volto, ed è espresso in percentuali rispetto alla dimensione complessiva dell'immagine. Viene definito attraverso le seguenti grandezze:

- *Width*: larghezza del riquadro rispetto alla larghezza totale dell'immagine;
- *Height*: altezza del riquadro rispetto all'altezza totale dell'immagine;
- *Left*: Distanza del lato sinistro del riquadro dal bordo sinistro dell'immagine;
- *Top*: distanza del lato superiore del riquadro dal bordo superiore dell'immagine.

Nel nostro caso sono tutte queste grandezze hanno valore vicino allo 0.

Inoltre troviamo l'array "FaceMatches" che mostra il volto nell'immagine di destinazione che corrisponde a quello dell'immagine di origine, insieme al punteggio di similarità. Esso contiene i dettagli del volto rilevato nell'immagine di destinazione; in esso:

- *BoundingBox*: indica le coordinate che delimitano il volto all'interno dell'immagine di confronto, in modo simile a quanto avviene con il volto di riferimento;
- *Confidence*: indica il livello di confidenza di Amazon Rekognition sul fatto che l'area rilevata contenga effettivamente un volto.

Il campo "Similarity" permette di valutare quanto sono simili i volti rilevati. È possibile impostare una soglia di similarità minima per considerare i volti corrispondenti. Nel nostro caso abbiamo una similarità superiore al 99% quindi molto alta.

```
"SourceImageFace": {
  "BoundingBox": {
    "width": 0.1922966092824936,
    "height": 0.3368915021419525,
    "left": 0.251707524061203,
    "top": 0.08384109288454056
  },
  "Confidence": 99.99917602539862
},
"FaceMatches": [
  {
    "Similarity": 99.70489501953125,
    "Face": {
      "BoundingBox": {
        "width": 0.3703244626522064,
        "height": 0.3914508819580078,
        "left": 0.33573609590530396,
        "top": 0.09635191410779953
      },
      "Confidence": 99.99381256103516,
    }
  }
]
```

Figura 4.10: Codice del confronto facciale

---

## Esperienze nell'ambito dell'analisi del testo

---

*Nel presente capitolo verrà esplorato il task dell'Intelligenza Artificiale che si occupa dell'analisi del testo. In particolare verranno presentati il suo funzionamento e le sue applicazioni. In seguito verrà affrontato un tool di analisi del testo, ovvero Amazon Comprehend. Attraverso questa piattaforma si faranno vari test per verificare la sua efficacia nell'analizzare vari tipi di testi.*

### 5.1 Introduzione all'analisi dei testi tramite Intelligenza Artificiale

L'analisi del testo tramite Intelligenza Artificiale è un insieme di tecniche e strumenti che consentono di estrarre il significato e le informazioni strutturate da dati testuali non strutturati, come articoli, recensioni, post sui social media o documenti. Con l'aumento esponenziale della quantità di dati testuali generati giornalmente, l'analisi di testo automatizzata è diventata essenziale per molte imprese e industrie. L'obiettivo principale consiste nel convertire grandi volumi di testo in dati utili che possano essere interpretati e analizzati.

Utilizzando modelli di Machine Learning e tecniche di elaborazione del linguaggio naturale (NLP), l'Intelligenza Artificiale può riconoscere modelli, identificare entità importanti, classificare documenti, analizzare sentimenti e, persino, generare linguaggio umano. I modelli più avanzati, come i transformer (ad esempio BERT o GPT), sono in grado di comprendere il contesto e le sfumature semantiche del linguaggio in modo più efficace rispetto agli approcci tradizionali.

L'analisi del testo tramite AI è utilizzata in molti settori, dalle ricerche di mercato all'analisi delle opinioni dei clienti, fino all'automatizzazione di attività come la classificazione delle email, la traduzione automatica e la generazione di contenuti. Grazie a questi strumenti, si possono ottenere insight strategici, ottimizzare i processi aziendali e migliorare l'interazione con gli utenti.

### 5.2 Funzionamento dell'analisi del testo

L'analisi del testo implica il processo di estrazione di informazioni dai testi scritti, che possono variare da documenti formali a post sui social media. Tale approccio viene utilizzato per una vasta gamma di attività come:

- *classificazione del testo*: ad esempio, separare email spam da quelle non spam;
- *analisi del sentiment*: determinare se un testo ha un tono positivo, negativo o neutro;

- *ricoscimento di entità nominate (NER)*: identificare entità specifiche, come nomi di persone, aziende, luoghi etc;
- *riassunto automatico*: sintetizzare lunghi documenti o articoli.

L'analisi del testo attraverso l'uso dell'Intelligenza Artificiale, sfrutta algoritmi di elaborazione del linguaggio e Machine Learning per estrarre informazioni dai dati testuali. Il primo passo è pre-elaborare il testo e prepararlo per l'analisi. Questo include le seguenti fasi già spiegate esaurientemente nella Sezione 2.4:

- *tokenizzazione*: il testo viene suddiviso in parole (token) o frasi;
- *lemmatizzazione*: le parole vengono ridotte alla loro forma base;
- *rimozione di stop words*: parole comuni e non informative vengono eliminate;
- *pulizia*: si eliminano simboli, punteggiatura o caratteri non rilevanti.

Le tecniche di analisi testuale sono basate su vari algoritmi di apprendimento automatico, che sono addestrati a riconoscere modelli linguistici e semantici all'interno dei dati. Per permettere ad un algoritmo di Intelligenza Artificiale di comprendere il testo, quest'ultimo deve essere convertito in una forma numerica. Alcune delle tecniche utilizzate per permettere questa conversione sono:

- *Bag of Words (BoW)*: in questo modello, ogni documento è considerato in quanto contenente parole, analogamente a una borsa; ciò consente una gestione di queste basata su liste, dove ogni borsa contiene determinate parole di una lista. Questo modello, quindi, rappresenta i dati di testo contando la frequenza di ciascuna parola, ignorando la sintassi e l'ordine delle parole. Effettivamente, con questo modello abbiamo una perdita di informazioni sull'ordine delle parole; due frasi con le stesse parole ma nell'ordine differente avranno la stessa rappresentazione. BoW registra, inoltre, solo la presenza o la frequenza delle parole, senza prendere in considerazione altri fattori semantici, quali il loro significato.
- *TF-IDF (Term Frequency-Inverse Document Frequency)*: TF-IDF è una tecnica ampiamente utilizzata nell'elaborazione del linguaggio naturale (NLP) per quantificare l'importanza di una parola in un documento rispetto a un intero corpus di documenti. A differenza del semplice conteggio delle parole, come nella tecnica Bag of Words (BoW), TF-IDF bilancia l'importanza delle parole frequenti in un documento con la loro rarità in tutto il testo, mettendo in evidenza i termini che sono distintivi per quel documento. L'idea alla base di questo comportamento è quella di dare più importanza ai termini che compaiono nel documento, ma che, in generale, sono poco frequenti.
- *Word Embeddings*: è una tecnica che permette di memorizzare le relazioni sia semantiche che sintattiche, partendo da un testo e costruendo uno spazio vettoriale in cui i vettori delle parole sono più vicini se le parole occorrono negli stessi contesti linguistici, cioè se sono riconosciute come semanticamente più simili. Alcune delle tecnologie che utilizzano questa tecnica sono Word2Vec, GloVe o modelli più avanzati, come BERT e GPT, che mappano le parole in uno spazio vettoriale, rappresentando la relazione semantica tra i termini (ad esempio, "re" e "regina" saranno vicini nello spazio vettoriale). Assegnando valori numerici alle parole in base alle somiglianze semantiche, il Word Embedding consente ai modelli delle reti neurali di comprendere il contesto in modo più efficiente. Questo approccio riduce la complessità di elaborazione e migliora le prestazioni dei modelli, preservando le informazioni semantiche.

## 5.3 Applicazioni dell'analisi del testo tramite IA

L'analisi del testo tramite Intelligenza Artificiale viene utilizzata in una vasta gamma di ambiti, grazie alla sua capacità di estrarre informazioni utili e pattern da grandi quantità di dati testuali. Ecco alcuni dei principali settori in cui questa tecnologia viene applicata:

- *Servizi finanziari e rilevazione di frodi*: l'analisi del testo viene utilizzata per monitorare email, messaggi e documenti al fine di rilevare anomalie o attività fraudolente. Riguardo ai documenti legali e ai contratti, l'Intelligenza Artificiale può analizzarli prelevando clausole chiave o potenziali rischi. Infatti, gli avvocati potrebbero utilizzare l'analisi del testo per esaminare grandi quantità di documenti giuridici, identificando i casi e i precedenti più rilevanti. Inoltre, le banche e le istituzioni finanziarie utilizzano l'IA per analizzare rapporti di mercato, report finanziari e notizie economiche, estraendo dati chiave che possono influenzare decisioni di investimento.
- *Educazione*: l'IA viene utilizzata per correggere e valutare i compiti scritti, analizzando la grammatica, la sintassi e il contenuto. I sistemi di tutoring basati su NLP possono interagire con gli studenti, rispondendo alle domande e guidandoli nel processo di apprendimento.
- *Analisi dei report medici*: l'analisi del testo consente di estrarre dati importanti dai report medici, dai referti o dagli articoli scientifici per migliorare la diagnosi, la gestione dei pazienti e la ricerca medica.

## 5.4 Caso di studio: Amazon Comprehend

Questa sezione si focalizza sull'analisi approfondita di Amazon Comprehend, un servizio di elaborazione del linguaggio naturale fornito da Amazon AWS. Amazon Comprehend utilizza tecniche di Machine Learning per analizzare dei testi, nonché per rilevare ed estrarre informazioni.

### 5.4.1 Che cos'è

Amazon Comprehend è il servizio di elaborazione del linguaggio naturale offerto da AWS, ed è in grado di raccogliere informazioni da un testo. Amazon Comprehend grazie, al Machine Learning, riesce a individuare relazioni e insight all'interno di testi. Attraverso l'elaborazione di questi insight, esso riesce a riconoscere entità, frasi chiave, sentimenti, etc. Esso, infatti, fornisce servizi di riconoscimento delle entità personalizzati, classificazione personalizzata, estrazione di frasi chiave e altre tipologie di API per permettere di integrare l'elaborazione del linguaggio naturale nelle applicazioni. Ciò è possibile perché Amazon Comprehend si appoggia a tecniche di NLP per comprendere il linguaggio umano, come la tokenizzazione, la lemmatizzazione e la rappresentazione vettoriale del testo. Inoltre, sfrutta algoritmi avanzati per il riconoscimento delle entità e l'analisi del sentiment, per estrarre informazioni rilevanti in modo automatico. Amazon Comprehend funziona senza la necessità di addestrare modelli specifici da parte dell'utente, poiché utilizza modelli pre-addestrati continuamente migliorati tramite apprendimento basato su grandi quantità di dati, consentendo al sistema di comprendere il contesto, le relazioni semantiche e la struttura dei testi. Inoltre, Amazon Comprehend impiega modelli basati su architetture di transformer<sup>1</sup>, che vengono usati primariamente nelle branche dell'elaborazione del linguaggio naturale

<sup>1</sup>Nell'apprendimento automatico, un transformer è un modello di deep learning che adotta il meccanismo della auto-attenzione, pesando differenzialmente la significatività di ogni parte dei dati in ingresso.

come BERT (Bidirectional Encoder Representations from Transformers), e permettono di analizzare il contesto bidirezionale del testo, migliorando la comprensione delle relazioni tra le parole e le frasi. Oltre all'apprendimento supervisionato per la classificazione di documenti, Amazon Comprehend utilizza anche tecniche di topic modeling non supervisionato, che consente di identificare automaticamente temi e argomenti principali all'interno di grandi set di dati. Questa piattaforma può essere utilizzata per semplificare i flussi di lavoro di elaborazione, categorizzando ed estraendo automaticamente le informazioni dallo stesso. Ad esempio, è possibile utilizzare Amazon Comprehend per eseguire le seguenti attività:

- Effettuare l'analisi del sentiment su ticket di assistenza cliente, recensioni di prodotti, feed dei social media e altro.
- Integrare Amazon Comprehend con Amazon Lex per sviluppare una chatbot intelligente e conversazionale.
- Estrarre termini medici da documenti e identificare la relazione che sussiste tra essi e Amazon Comprehend Medical.

#### 5.4.2 Benefici nell'uso di Amazon Comprehend

Di seguito sono elencati alcuni dei vantaggi che derivano dell'utilizzo di Amazon Comprehend:

- *Analisi dettagliata del testo:* Amazon Comprehend utilizza la tecnologia di Deep Learning per analizzare accuratamente il testo. Esso riesce, quindi, ad analizzare in maniera approfondita i sentimenti espressi da un testo, aiutando le aziende che vogliono monitorare recensioni dei clienti, feedback sui social media o comunicazioni interne per comprendere meglio lo stato d'animo dei propri utenti. Amazon Comprehend riesce, inoltre, a riconoscere le entità come persone, luoghi, organizzazioni, date, eventi, ecc., all'interno di un testo. Questa capacità è molto utile per estrarre automaticamente informazioni chiave dai documenti o dalle e-mail, migliorando processi come la classificazione e la ricerca di dati.
- *Rilevamento delle lingue:* Amazon Comprehend è in grado di identificare la lingua in cui è scritto un documento, anche se contiene frasi in più lingue. Questo è utile per piattaforme globali che gestiscono contenuti multilingua e devono eseguire elaborazioni linguistiche differenziate. In quanto esso permette di analizzare milioni di documenti in modo che si possano scoprire le informazioni dettagliate che contengono.
- *Integrazione con altri servizi:* Amazon Comprehend è progettato per funzionare perfettamente con altri servizi AWS come Amazon S3, AWS KMS e AWS Lambda.
- *Basso costo:* con Amazon Comprehend non ci sono tariffe minime o anticipazioni. I costi vengono calcolati sulla base dei documenti analizzati e dei modelli personalizzati sottoposti a training; si paga solo per quello che si utilizza. Questo lo rende conveniente per piccole e grandi aziende che vogliono analizzare il linguaggio naturale senza investire pesantemente in hardware o software proprietario.

#### 5.4.3 Funzionamento

Le funzionalità di analisi dei documenti di Amazon Comprehend possono essere adoperate utilizzando la console oppure le API fornite dal servizio. L'analisi dei testi può essere eseguita in tempo reale per piccoli carichi di lavoro oppure, nel caso di grandi set di documenti, possono essere avviati dei processi di analisi asincrona. Oltre ad utilizzare i modelli

pre-addestrati forniti dal servizio si possono addestrare i propri modelli personalizzati per la classificazione e il riconoscimento delle entità in base alle proprie necessità. Tutte le funzionalità di Amazon Comprehend accettano documenti di testo UTF-8 come input. Inoltre, per la classificazione personalizzata e il riconoscimento personalizzato delle entità, è possibile utilizzare come input anche file di immagine, file PDF e file Word.

Come mostrato nella Figura 5.1 Amazon Comprehend, analizzando un documento, offre vari insight quali:

- *Analisi del sentimento*: riconosce il sentimento dominante di un testo, classificandolo come positivo, negativo, neutrale o misto. Può essere utilizzato, ad esempio, per capire il feedback dei clienti su recensioni o post sui social media.
- *Sentiment mirato*: questo insight non si limita a determinare il tono generale di un testo, ma cerca di identificare il sentimento associato a entità o aspetti specifici all'interno di un documento.
- *Riconoscimento delle entità*: identifica entità come persone, luoghi, organizzazioni, date, quantità e altro ancora all'interno del testo, con lo scopo di identificarne i concetti principali ed etichettarli a seconda di un dataset predefinito di tipologie. Ciò è utile per estrarre informazioni chiave da documenti come contratti o articoli di notizie.
- *Estrazione delle frasi chiave*: esamina il testo e individua le frasi principali o i concetti centrali, facilitando l'individuazione dei temi principali o delle informazioni cruciali.
- *Rilevamento della lingua*: riconosce la lingua dominante di un testo utile per analizzare contenuti multilingue. Supporta una vasta gamma di lingue, più di 100, incluse inglese, spagnolo, italiano, francese, tedesco, portoghese, arabo, cinese (semplificato e tradizionale), giapponese, coreano e molte altre.
- *Rilevamento di PII (Informazioni di Identificazione Personale)*: è possibile rilevare e oscurare informazioni sensibili come nomi, numeri di carte di credito e dati personali, garantendo la conformità alle normative sulla privacy.
- *Sintassi*: identifica la funzione di ciascuna parola o elemento all'interno del documento.



**Figura 5.1:** Funzionamento di Amazon Comprehend

La schermata iniziale, come possiamo notare nella Figura 5.2, mostra l'analisi del testo che viene effettuata da Amazon Comprehend quando ad esso viene proposto un documento. In input può essere fornito un testo che può essere di massimo 5000 parole. L'analisi del testo può essere di due tipi: quella fornita direttamente da Amazon, il tipo "Incorporato", ovvero un'analisi in tempo reale con modelli integrati, oppure il tipo "Personalizzato", ovvero un'analisi personalizzata in tempo reale con endpoint. Nel primo caso vengono utilizzati

i metodi integrati di Amazon Comprehend, nel secondo caso bisogna creare un endpoint personalizzato, ovvero un progetto dell'utente, così che Amazon Comprehend possa visualizzare informazioni dettagliate in tempo reale basate su modelli personalizzati. Il testo da analizzare deve essere inserito all'interno del campo di testo chiamato "testo di input" per eseguire l'analisi bisogna premere il pulsante "Analizzare". La piattaforma analizza il testo e va a estrarre le varie caratteristiche di cui si ha bisogno, come la lingua dominante, le frasi chiave, la sintassi, etc. Al di sotto dell'analisi troviamo, poi, l'elenco dei vari elementi che compongono il testo che il sistema è riuscito ad individuare. Nel nostro caso vengono individuate "Zhang Wei", "Jhon" e "Alice", come persone, "AnyCompany" e "Sunshine Spa" come organizzazioni, infine "\$24.53" come una quantità. Inoltre si riescono ad individuare data e luogo. Il resto delle informazioni vengono classificate con il tipo "Altro", in quanto Amazon Comprehend non riesce a individuare il tipo e a categorizzarle. Al di sotto della sezione "Risultati", se ne trova un'altra, ovvero la sezione "Application Integration". In essa vengono fornite informazioni riguardanti il metodo dell'API utilizzato per eseguire l'analisi basata sull'insight. Inoltre, vengono presentate, in formato JSON, sia la chiamata effettuata all'API sia la relativa risposta fornita dall'API stessa.

**Analisi in tempo reale** Informazioni

Analizza il testo in tempo reale utilizzando modelli integrati o personalizzati. Con i modelli predefiniti, è possibile riconoscere le entità, estrarre frasi chiave, rilevare le lingue dominanti, rilevare le informazioni personali, determinare il sentiment, rilevare il sentiment misto o analizzare la sintassi. Con i modelli personalizzati, è possibile rilevare le entità definite dall'utente oppure classificare i documenti utilizzando le proprie categorie o etichette.

**▼ Panoramica**

**Analisi in tempo reale con modelli integrati**  
Utilizza i modelli integrati di Amazon Comprehend per analizzare i contenuti in tempo reale.

**Analisi personalizzata in tempo reale con endpoint**  
Crea un endpoint per il tuo modello personalizzato per **classificare i documenti** o **rilevare entità personalizzate** in tempo reale.

**Dati di input**  
[Lingue supportate](#)

**Tipo di analisi** Informazioni

**Incorporato**  
Visualizza informazioni dettagliate in tempo reale basate sui modelli integrati di AWS.

**Custom**  
Visualizza informazioni dettagliate in tempo reale basate su modelli personalizzati da un endpoint che hai creato.

**Testo di input**

Hello Zhang Wei, I am John. Your AnyCompany Financial Services, LLC credit card account 1111-0000-1111-0000 has a minimum payment of \$24.53 that is due by July 31st. Based on your autopay settings, we will withdraw your payment on the due date from your bank account number XXXXXX1111 with the routing number XXXXXX0000.  
Customer feedback for Sunshine Spa, 123 Main St, Anywhere. Send comments to Alice at sumpag@mail.com.  
I enjoyed visiting the spa. It was very comfortable but it was also very expensive. The amenities were ok but the service made the spa a great experience.

0/76 di 5000 caratteri utilizzati

Testo in chiaro **Analizzare**

**Intuzioni** Informazioni

Entità | Frasi chiave | Lingua | PI | Sentimento | Sentimento misto | Sintassi

**▼ Testo analizzato**

Ciao Zhang Wei! Sono John. La tua AnyCompany Servizi Finanziari, LLC carta della carta di credito 1111-0000-1111-0000 ha un pagamento minimo di \$24.53/giorno che è dovuto da 31 luglio. In base alle impostazioni di pagamento automatico, ritireremo il pagamento alla data di scadenza dal numero di conto bancario dell'utente XXXXXX1111 con il numero di routing XXXXXX0000.  
Feedback dei clienti per Sunshine Spa, 123 Main St qualsiasi luogo. Invia commenti a Alice a sumpag@mail.com.  
Mi è piaciuto visitare il centro benessere. Era molto comodo ma era anche molto costoso. I servizi erano ok ma il servizio ha reso il centro benessere un'ottima esperienza.

**▼ Risultati**

Q Ricerca

Entità	Digitare	Fiducia
Zhang Wei	Persona	0.99+
John	Persona	0.99+
AnyCompany Servizi Finanziari, LLC	Organizzazione	0.99+
1111-0000-1111-0000	Altro	0.99+
\$24.53/giorno	Quantità	0.99+
31 luglio	Data	0.99+
XXXXXX1111	Altro	0.98
XXXXXX0000	Altro	0.97
Sunshine Spa	Organizzazione	0.98
123 Main St	Localizzazione	0.98

**Figura 5.2:** Schermata iniziale Amazon Comprehend

Nel seguito forniamo una panoramica e una spiegazione più dettagliata dei vari insight che Amazon Comprehend offre.

## Analisi del sentimento

Un primo insight che viene fornito da Amazon Comprehend è l'analisi del sentimento, ossia la capacità di analizzare le emozioni espresse da un testo. Questo tipo di analisi è utile per trasformare dati testuali non strutturati in informazioni utili, aiutando le aziende e gli sviluppatori a prendere decisioni migliori, per esempio, nell'ambito dell'analisi delle opinioni sui prodotti oppure nella valutazione di recensioni scritte dai clienti. Questa analisi del sentimento viene suddivisa da Amazon Comprehend in quattro categorie principali:

- *Positive*: il testo esprime un'emozione favorevole, felice o soddisfatta;
- *Negative*: il testo esprime insoddisfazione, rabbia o emozioni avverse;
- *Mixed*: il testo presenta una combinazione di sentimenti positivi, negativi o neutri;
- *Neutral*: il testo è neutro e non contiene emozioni rilevanti.

Per ogni categoria di sentimento, Amazon Comprehend fornisce un numero, ovvero il punteggio di confidenza. Esso rappresenta la probabilità che quell'emozione sia presente nel testo. Inoltre, oltre alla confidenza, viene riportato quale tra i vari sentimenti è quello a prevalenza maggiore. Per ottenere questa analisi è necessario utilizzare il metodo dell'API DetectSentiment che restituirà un risultato in formato JSON. Prendiamo, ad esempio, una semplice frase "Luca ogni volta che va a mangiare fuori con la famiglia è felice". In Figura 5.3 viene presentato il file JSON della frase analizzata.

```
API response
1  {
2    "Sentiment": {
3      "Sentiment":
4        "POSITIVE",
5        "SentimentScore":
6          {
7            "Positive": 0
8              .998380303382
9              8735,
10           "Negative": 0
11             .000054553649
12             53377284,
13           "Neutral": 0
14             .001562124118
15             2088852,
16           "Mixed": 0
17             .000003103158
18             860540134
19         }
20     }
21 }
```

Figura 5.3: File JSON relativo all'analisi del sentimento

In questa parte di codice possiamo notare, come già espresso in precedenza, i vari valori di confidenza che vengono associati alle varie emozioni. Riguardo ai sentimenti "Negative", "Neutral" e "Mixed" abbiamo una confidenza intorno allo zero, mentre associato a "Positive" troviamo il valore intorno a uno. Questi numeri esprimono come Amazon Comprehend abbia individuato come sentimento predominante quello "Positive", infatti è quello con un valore di confidenza maggiore. Possiamo quindi affermare che in questo caso la piattaforma abbia interpretato in maniera corretta l'emozione presente nel testo.

## Analisi del sentimento mirato

Un'evoluzione dell'insight precedente è quello del sentimento mirato; si tratta di un tipo di analisi del sentimento che non si limita a valutare il sentimento generale di un testo, ma

lo collega a entità specifiche o aspetti particolari menzionati nel testo stesso. In altre parole, invece di analizzare il sentimento a livello globale, si cerca di capire come le emozioni o le opinioni si riferiscano a specifici oggetti, persone o concetti citati. Questo tipo di analisi fornisce varie informazioni, ovvero:

- *Identificazione delle entità*: il primo passo consiste nell'individuare le entità nel testo, come nomi di prodotti, marchi, persone o categorie. Amazon Comprehend, ad esempio, offre la funzionalità di entity recognition che identifica automaticamente queste entità.
- *Punteggio di confidenza per ogni entità*: simile all'analisi del sentimento tradizionale, ogni sentimento mirato viene accompagnato da un punteggio di confidenza che indica quanto il modello è sicuro di quella classificazione.
- *Attribuzione del sentimento*: una volta che le entità sono state identificate, l'algoritmo cerca di determinare il sentimento associato a ciascuna di esse, non ch  il sentimento predominante.

Questo tipo di analisi   disponibile solo in lingua inglese. Per studiare un esempio prendiamo come frase "Amo la fotocamera di questo telefono, ma la batteria si scarica troppo velocemente e il servizio clienti non   stato d'aiuto", scrivendola in inglese. Nell'immagine successiva (Figura 5.4) possiamo notare come alle parole sottolineate venga attribuito un sentimento. Nel caso delle parole verdi viene attribuito il sentimento positivo, a quelle blu il sentimento neutrale, a quelle rosse il sentimento negativo; infine a quelle di colore grigio il software assegna un mix di sentimenti. Nel nostro caso abbiamo un sentimento positivo nel caso della parola "camera" e "cellulare"; infatti, nella frase viene espresso un senso di apprezzamento per la fotocamera del cellulare. Nel caso, invece, della batteria e del servizio clienti, Amazon Comprehend associa un sentimento di negativit , infatti nella frase si esprime un senso di disapprovazione verso la velocit  con cui si scarica la batteria, e l'inefficienza del servizio clienti. Si potrebbe avere un dubbio nell'interpretazione dell'entit  "I"; infatti ad essa il sistema associa un sentimento neutrale, quando la si potrebbe, invece, associare ad un sentimento positivo, in combinazione con il verbo successivo che esprime, appunto, un sentimento di positivit . Ma questa analisi potrebbe essere opinabile. Possiamo affermare come in questo caso il sistema sia riuscito ad interpretare bene il sentimento espresso dalle varie parti nella frase posta in input.

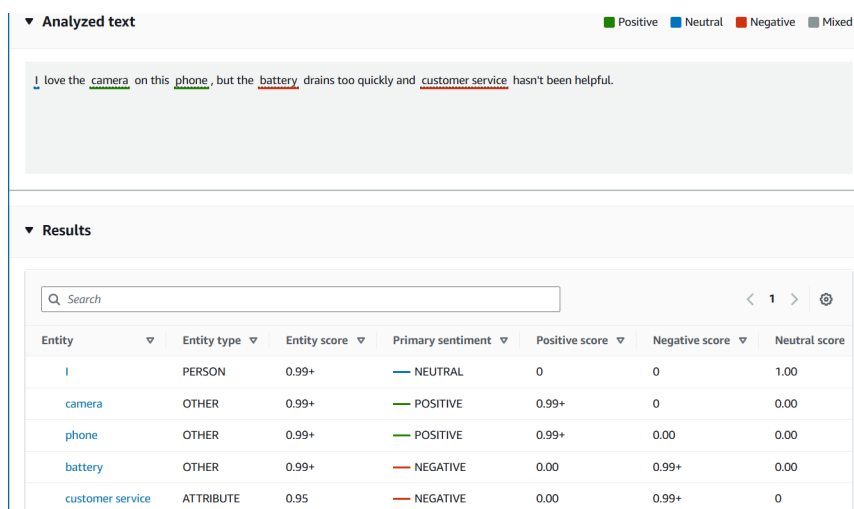


Figura 5.4: Analisi del sentimento mirato

Di seguito (Figura 5.5) viene mostrata una parte del file JSON, in particolare quella riferita alla parola "camera". Vicino a ogni sentimento sono associati dei numeri di confidenza. Nel caso del sentimento positivo, abbiamo il numero più elevato, infatti, il sentimento positivo è individuato come predominante. Nel file il campo "BeginOffset" indica il punto nel testo in cui inizia la menzione dell'entità, mentre il campo "End Offset" indica dove termina.

```
27 {
28   "DescriptiveMentionIndex": [
29     0
30   ],
31   "Mentions": [
32     {
33       "Score": 0.9999669790267944,
34       "GroupScore": 1,
35       "Text": "camera",
36       "Type": "OTHER",
37       "MentionSentiment": {
38         "Sentiment": "POSITIVE",
39         "SentimentScore": {
40           "Positive": 0.9999930262565613,
41           "Negative": 0,
42           "Neutral": 0
43             .000004999999873689376,
44           "Mixed": 0
45             .000001999999949504854
46         }
47       },
48       "BeginOffset": 11,
49       "EndOffset": 17
50     }
51   ]
52 }
```

**Figura 5.5:** File JSON relativo al sentimento mirato

### Analisi delle entità

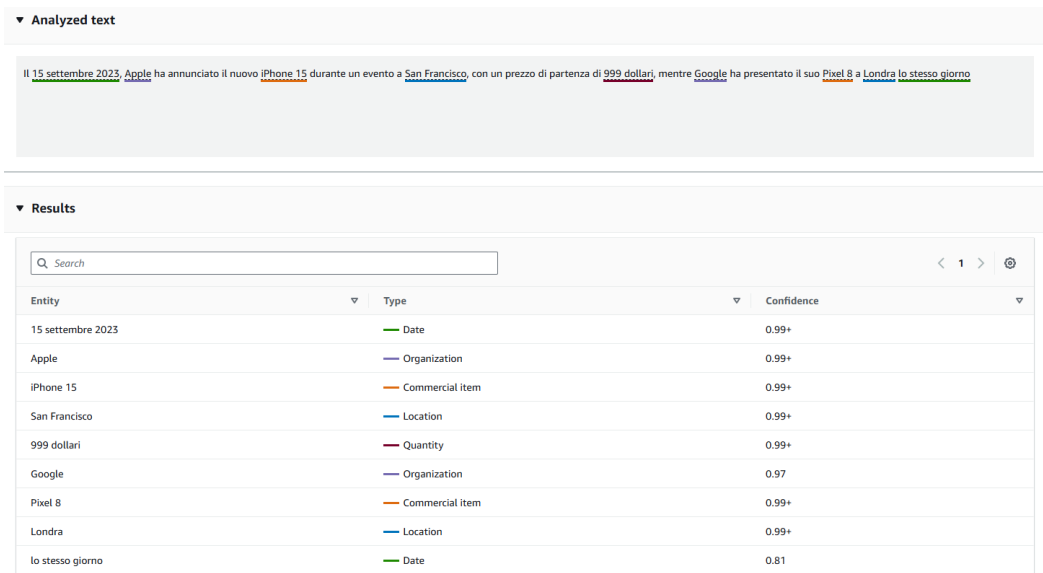
Amazon Comprehend è in grado di svolgere l'analisi delle entità. Come spiegato precedentemente tale analisi consiste nell'identificare e classificare automaticamente specifiche entità menzionate in un testo. Le entità che vengono rilevate da Amazon Comprehend sono:

- nomi di persone;
- organizzazioni;
- luoghi;
- date;
- quantità;
- eventi;
- prodotti.

Amazon Comprehend offre, quindi, un servizio di entity recognition che identifica automaticamente le entità in un testo e le classifica in diverse categorie. Quando si fornisce un documento di testo ad Amazon Comprehend, questo utilizza modelli di Machine Learning per rilevare automaticamente le entità all'interno del testo. Esso restituisce, quindi, un elenco delle entità riconosciute con il loro tipo, la posizione nel testo e un punteggio di confidenza che indica quanto il modello è sicuro della classificazione.

Per testare la capacità di Amazon Comprehend nel fare l'analisi delle entità possiamo prendere come esempio questa frase: "Il 15 settembre 2023, Apple ha annunciato il nuovo

iPhone 15 durante un evento a San Francisco, con un prezzo di partenza di 999 dollari, mentre Google ha presentato il suo Pixel 8 a Londra lo stesso giorno". Come mostrato nella Figura 5.6, Amazon riesce a trovare correttamente le varie entità presenti nel testo. Una volta analizzato il testo, il sistema sottolinea con diversi colori le diverse tipologie di entità che è riuscito a individuare. Viene poi aggiunta una sezione "Results" dove viene fornita una lista completa di ogni entità trovata, con le relative etichette e percentuali di confidenza, che indicano il livello di fiducia che il servizio ha riguardo alla correttezza del riconoscimento del tipo di entità.



The screenshot shows the Amazon Comprehend interface. The top section, titled "Analyzed text", displays a sample sentence with various words highlighted in different colors (green, orange, blue, red) to represent different entity types. The bottom section, titled "Results", contains a table with the following data:

Entity	Type	Confidence
15 settembre 2023	Date	0.99+
Apple	Organization	0.99+
iPhone 15	Commercial item	0.99+
San Francisco	Location	0.99+
999 dollari	Quantity	0.99+
Google	Organization	0.97
Pixel 8	Commercial item	0.99+
Londra	Location	0.99+
lo stesso giorno	Date	0.81

Figura 5.6: Analisi delle entità

Nelle API, per ottenere l'analisi delle entità, si utilizza il metodo DetectEntities, che fornisce un file JSON strutturato come mostrato nella Figura (Figura 5.7)

```
1 {
2   "Entities": [
3     {
4       "Score": 0.9933477640151978,
5       "Type": "DATE",
6       "Text": "15 settembre 2023",
7       "BeginOffset": 3,
8       "EndOffset": 20
9     },
10    {
11      "Score": 0.992820680141449,
12      "Type": "ORGANIZATION",
13      "Text": "Apple",
14      "BeginOffset": 22,
15      "EndOffset": 27
16    },
17    {
18      "Score": 0.9980770349502563,
19      "Type": "COMMERCIAL_ITEM",
20      "Text": "iPhone 15",
21      "BeginOffset": 51,
22      "EndOffset": 60
23    },
24  ]
25 }
```

Figura 5.7: File JSON relativo all'analisi delle entità

In figura viene mostrato parte del codice, in particolare la parte che si riferisce alle entità ovvero data, organizzazione e commercial item, che rientra nella categoria dei prodotti.

## Estrazione delle frasi chiavi

Un'altra funzionalità offerta da Amazon Comprehend è quella che permette di identificare, nel testo proposto, le "Key Phrases", ovvero le "frasi chiave". Con il termine "Key Phrases" ci si riferisce ai termini o alle espressioni che rappresentano il concetto centrale di un testo. Si tratta di parole o gruppi di parole che racchiudono l'essenza di un documento o di una parte di testo. Sostanzialmente, Amazon Comprehend attraverso un'analisi semantica del testo, analizza e identifica quelle frasi che sono più significative per il contesto, fornendo all'utente, nella sezione "Results", una lista di tutti gli elementi riconosciuti. Ogni frase chiave include anche il punteggio di confidenza riguardo al fatto che la stringa sia una frase chiave. Prendiamo ad esempio come frase "Amazon Comprehend utilizza il Machine Learning per analizzare il linguaggio naturale e fornire insight sui dati testuali, identificando entità come persone, luoghi e frasi chiave", e vediamo come Amazon Comprehend individua le frasi chiave. In questo caso il software sottolinea le varie parti della frase che ritiene siano rilevanti all'interno del testo, ed ad ognuna associa un determinato livello di confidenza. Nella Figura 5.8 si può notare come i livelli di confidenza siano tutti superiori al valore 70, questo indica quanto il modello è sicuro che le varie frasi individuate rappresentino un concetto rilevante per il contesto generale del testo.

The screenshot shows the 'Analyzed text' section with the following text: "Amazon Comprehend utilizza il machine learning per analizzare il linguaggio naturale e fornire insight sui dati testuali, identificando entità come persone, luoghi e frasi chiave". Below this, the 'Results' section displays a table of key phrases with their confidence scores.

Key phrases	Confidence
Amazon Comprehend	0.99+
il linguaggio naturale	0.99+
luoghi	0.99+
frasi chiave	0.99+
entità come persone	0.98
insight	0.97
dati testuali	0.76

**Figura 5.8:** Analisi delle frasi chiave

Nelle API, per ottenere l'analisi delle frasi chiave, si utilizza il metodo `DetectKeyPhrases` che fornisce un JSON, la cui struttura viene presentata nella Figura 5.9. L'elemento principale della risposta presente nel codice è il termine "KeyPhrases", questo contiene un array di frasi chiave estratte dal testo. Ogni `KeyPhrase` contiene all'interno i seguenti campi:

- *Text*: la frase chiave estratta dal testo;
- *Score*: un valore numerico (tra 0 e 1) che rappresenta la confidenza del modello sull'importanza di quella frase chiave, più il valore è vicino a 1, più il modello è sicuro della rilevanza della frase;
- *BeginOffset*: la posizione nel testo originale dove inizia la frase chiave;
- *EndOffset*: la posizione nel testo originale dove finisce la frase chiave.

```
1 [{"KeyPhrases": [  
2   {  
3     "Score": 0.9984550476074219,  
4     "Text": "Amazon Comprehend",  
5     "BeginOffset": 0,  
6     "EndOffset": 17  
7   },  
8   {  
9     "Score": 0.9999291300773621,  
10    "Text": "il linguaggio naturale",  
11    "BeginOffset": 62,  
12    "EndOffset": 84  
13  },  
14  {  
15    "Score": 0.9786719083786011,  
16    "Text": "insight",  
17    "BeginOffset": 95,  
18    "EndOffset": 102  
19  },  
20  {  
21    "Score": 0.7694594860076904,  
22    "Text": "dati testuali",  
23    "BeginOffset": 107,  
24    "EndOffset": 120  
25  },  
26  {  
27    "Score": 0.9892978072166443,  
28    "Text": "entità come persone",  
29    "BeginOffset": 136,  
30    "EndOffset": 155  
31  },  
32  {  
33    "Score": 0.9973702430725098,  
34    "Text": "luoghi",  
35    "BeginOffset": 157,  
36    "EndOffset": 163  
37  },  
38  {  
39    "Score": 0.994104266166687,  
40    "Text": "frasi chiave",  
41    "BeginOffset": 166,  
42    "EndOffset": 178  
43  }  
44 ]  
45 }  
46 ]
```

Figura 5.9: File JSON relativo alle frasi chiave

L'utilizzo dell'estrazione degli elementi chiave può essere utile in tantissimi settori, in particolare essi possono essere usati come input per algoritmi di classificazione, migliorando l'analisi dei documenti su larga scala. In alternativa le aziende possono usare l'estrazione di frasi chiave per monitorare le opinioni dei clienti sui social media e per comprendere quali aspetti di un prodotto o servizio stanno ricevendo maggior attenzione.

### Analisi dei PII

Amazon Comprehend può offrire anche la funzionalità specifica per rilevare nei testi le Personal Identifiable Informations (PII), ovvero le informazioni di identificazione personale. Questa funzionalità è estremamente utile per garantire la conformità con determinate normative in vigore, come il regolamento GDPR (General Data Protection Regulation), che impongono la protezione delle informazioni personali sensibili. Le informazioni di identificazione personale includono dati che possono essere utilizzati per identificare un individuo. Amazon Comprehend è in grado di identificare le seguenti informazioni personali: nome e cognome, indirizzo, numero di telefono, email, codice fiscale, numero di carta di credito o debito, data di nascita, sesso e luogo di nascita. Amazon Comprehend non solo identifica le informazioni personali, ma le classifica in varie categorie. Prendendo ad esempio il testo "Il numero di carta di credito di Luca Rossi è 4111 1111 1111 1111, e il suo numero di telefono è +1-234-567-8901", l'output della frase sarebbe: "Luca Rossi" come "NAME", quindi nome, "4111 1111 1111 1111", come tipo "CREDIT\_NUMBER", quindi numero di carta di credito, e "+1-234-567-8901", come tipo "PHONE", quindi numero di telefono.

Nella Figura 5.10 viene mostrata la tabella con i risultati dell'analisi da parte di Amazon Comprehend; la frase viene inserita in lingua inglese in quanto l'analisi dei dati PII è supportata soltanto per la lingua inglese.

▼ Analyzed text

Luca Rossi's credit card number is 4111 1111 1111 1111, and his phone number is +1-234-567-8901

---

▼ Results

Q Search < 1 > ⌂

Entity	Type	Confidence
Luca Rossi	Name	0.99+
4111 1111 1111 1111	Credit debit number	0.99+
+1-234-567-8901	Phone	0.98

**Figura 5.10:** Analisi dei dati PII

Nella Figura 5.11 viene proposta la risposta in formato JSON riguardante il testo descritto precedentemente, dove possiamo individuare i corrispettivi elementi. In primis troviamo "Entities" che contiene un elenco di tutte le entità PII rilevate; ogni entità ha i seguenti campi:

- *Score*: indica la confidenza del modello nell'identificazione della PII;
- *Type*: specifica il tipo di PII;
- *BeginOffset* e *EndOffset*: indicano la posizione esatta della PII all'interno del testo originale;
- *Text*: rappresenta il frammento di testo contenente la PII identificata.

▼ Application integration info  
API call and API response of DetectPiiEntities API

API call	API response
<pre> 1 { 2   "Text": "Luca Rossi's credit card number is 4111 1111 1111 1111, and his phone number is +1-234-567-8901", 3   "LanguageCode": "en" 4 } </pre>	<pre> 1 { 2   "Entities": [ 3     { 4       "Score": 0.9999945163726807, 5       "Type": "NAME", 6       "BeginOffset": 0, 7       "EndOffset": 10 8     }, 9     { 10      "Score": 0.9999524354934602, 11      "Type": "CREDIT_DEBIT_NUMBER", 12      "BeginOffset": 35, 13      "EndOffset": 54 14     }, 15     { 16      "Score": 0.9837067127227783, 17      "Type": "PHONE", 18      "BeginOffset": 80, 19      "EndOffset": 95 20     } 21   ] 22 } </pre>

Copy
Copy

**Figura 5.11:** File JSON relativo all'analisi dei dati PII

L'utilizzo della rilevazione PII può essere utile nei sistemi automatizzati, poiché può essere adattata per proteggere le informazioni personali quando vengono inviate tra diverse applicazioni o servizi. Oppure le aziende che trattano grandi quantità di dati personali possono utilizzare la rilevazione PII per garantire che le informazioni sensibili siano gestite in conformità con le normative sulla privacy.

### Analisi della lingua dominante

Una funzionalità fondamentale che Amazon Comprehend può offrire è il riconoscimento della lingua. Il sistema supporta una vasta gamma di linguaggi tra i quali inglese, tedesco, francese, spagnolo, italiano, portoghese, giapponese, cinese semplificato, coreano, arabo, russo, turco, olandese, svedese, danese, norvegese, finnico, indonesiano e malese. La funzionalità di rilevazione della lingua dominante di Amazon Comprehend è progettata per identificare automaticamente la lingua principale di un testo. Ciò è utile quando si gestiscono contenuti multilingue e non si è sicuri in quale lingua siano scritti. Quando viene inviato un testo ad Amazon Comprehend per l'analisi della lingua dominante, esso esamina il testo e restituisce i seguenti elementi:

- La lingua predominante identificata nel testo.
- Un punteggio di confidenza (confidence score) che indica quanto il modello è sicuro della lingua rilevata.
- Il codice della lingua attraverso codici ISO 639-1 per identificarla (ad esempio "en" per inglese, "it" per l'italiano, "es" per lo spagnolo).

Per effettuare questa operazione, si utilizza il metodo DetectSentiment dell'API, il quale fornisce un risultato in formato JSON che è strutturato come nella Figura 5.12.

```
1 {  
2   "Languages": {  
3     "LanguageCode": "en",  
4     "Score": 0.9912890791893005  
5   }  
6 }
```

Figura 5.12: File JSON relativo all'analisi della lingua dominante

### Analisi della sintassi

Come ultimo insight fornito da Amazon Comprehend abbiamo l'analisi della sintassi, chiamata "Syntax". Questa analisi ha come obiettivo quello di analizzare la struttura grammaticale del testo proposto, identificando le parti del discorso (come frasi, parole, verbi, soggetti e altri elementi linguistici) e le loro relazioni nel testo. Questo tipo di analisi aiuta a comprendere la composizione sintattica di un testo, rendendo più semplice ulteriori elaborazioni linguistiche e l'estrazione delle informazioni. Queste ultime possono essere utilizzate in molte applicazioni, tra cui la creazione di modelli di chatbot, l'analisi del customer feedback e la creazione di sistemi di ricerca e di classificazione delle informazioni. Prendiamo come esempio la frase "Luca ha comprato un nuovo cellulare", inserendola nel software, esso individua le varie parti del testo come mostrato nella Figura 5.13.

L'analisi di ogni singola parola viene restituita in formato JSON (Figura 5.14), nel quale possiamo individuare le varie parti del codice:

- *Text*: la parola o il token identificato.
- *PartOfSpeech*: in questa sezione vengono specificati il Tag (parte del discorso, come "NOUN" per sostantivo, "VERB" per verbo, "ADJ" per aggettivo) e il punteggio di confidenza (Score), che indica quanto il modello è sicuro della classificazione.
- *BeginOffset* e *EndOffset*: indicano la posizione della parola nel testo, con il primo carattere dell'intera frase che ha indice 0.

▼ Results

< 1 > ⌂

Word	Part of speech	Confidence
Luca	Proper noun	1.00
ha	Auxiliary verb	0.99+
comprato	Verb	1.00
un	Determiner	1.00
nuovo	Adjective	1.00
cellulare	Noun	1.00

**Figura 5.13:** Analisi della sintassi

```

{
  "Text":
  "cellulare"
  ,
  "BeginOffset": 26,
  "EndOffset": 35,
  "PartOfSpeech": {
    "Tag":
    "NOUN",
    "Score": 1
  }
}

```

**Figura 5.14:** File JSON relativo all'analisi della sintassi

#### 5.4.4 Esempi con Amazon Comprehend

Al fine di illustrare il funzionamento di Amazon Comprehend è stato scelto il testo seguente:

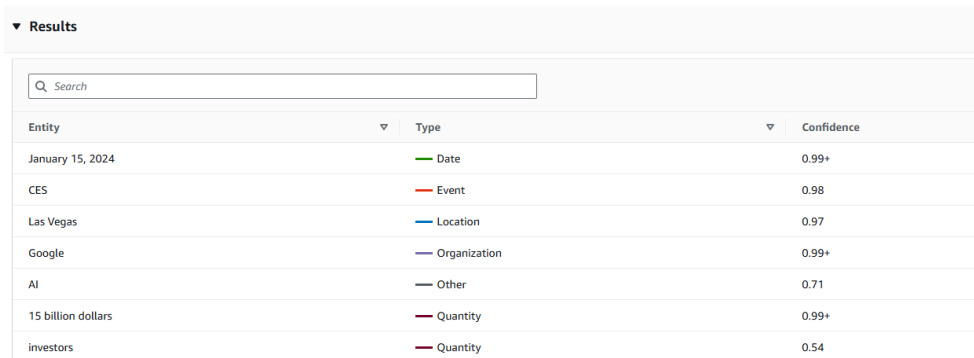
Il 15 gennaio 2024, durante il CES di Las Vegas, Google ha annunciato il lancio del suo nuovo servizio cloud Quantum AI, destinato a rivoluzionare il settore della tecnologia avanzata, con un investimento di 15 miliardi di dollari. L'annuncio ha ricevuto un'accoglienza mista: mentre gli esperti di settore sono entusiasti delle potenzialità, alcuni investitori rimangono scettici riguardo al rischio associato.

È stata, però, selezionata la lingua inglese in modo tale da sfruttare appieno le capacità di analisi di Amazon Comprehend. Di seguito, è riportata la traduzione letterale del testo:

On January 15, 2024, during CES in Las Vegas, Google announced the launch of its new Quantum AI cloud service, destined to revolutionize the advanced technology sector, with an investment of 15 billion dollars. The announcement received a mixed reception: while industry experts are enthusiastic about the potential, some investors remain skeptical about the associated risk.

## Risultati dell'analisi delle entità

Dall'analisi dell'entità è evidente come i risultati siano generalmente accurati, con il servizio in grado di distinguere chiaramente le diverse entità presenti nella frase fornendo valutazioni di sicurezza molto elevate (SCORE). Infatti, come si evince dalla Figura 5.15, si hanno tutti numeri elevati di confidenza; l'unico risultato errato riguarda la parola "investors", ovvero investitori, in quanto Amazon Comprehend lo individua come quantità, quando è evidente che non sia così. In questo caso, infatti, si ha un livello di confidenza pari al 54%, pertanto, al di sotto di una certa soglia è possibile trascurare l'analisi poiché potrebbe contenere errori.



Entity	Type	Confidence
January 15, 2024	Date	0.99+
CES	Event	0.98
Las Vegas	Location	0.97
Google	Organization	0.99+
AI	Other	0.71
15 billion dollars	Quantity	0.99+
investors	Quantity	0.54

Figura 5.15: Analisi delle entità dell'esempio

Questa analisi individua informazioni importanti nel testo, come chi ha fatto l'annuncio (Google), dove (CES, Las Vegas), quando (15 gennaio 2024) e l'investimento associato (15 miliardi di dollari). È utile per estrarre dati strutturati da contenuti non strutturati.

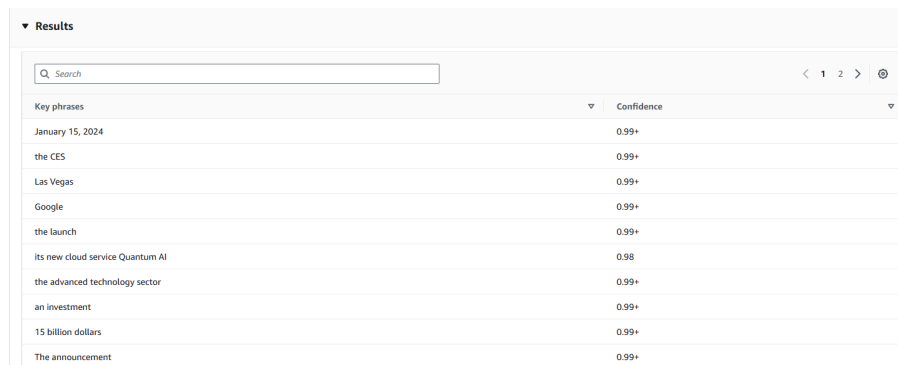
L'analisi in formato JSON appare come in Figura 5.16.

```
1  {
2    "Entities": [
3      {
4        "Score": 0.9933477640151978,
5        "Type": "DATE",
6        "Text": "15 settembre 2023",
7        "BeginOffset": 3,
8        "EndOffset": 20
9      },
10     {
11       "Score": 0.992820680141449,
12       "Type": "ORGANIZATION",
13       "Text": "Apple",
14       "BeginOffset": 22,
15       "EndOffset": 27
16     },
17     {
18       "Score": 0.9980770349502563,
19       "Type": "COMMERCIAL_ITEM",
20       "Text": "iPhone 15",
21       "BeginOffset": 51,
22       "EndOffset": 60
23     }
24   ]
25 }
```

Figura 5.16: File JSON relativo all'analisi delle entità dell'esempio

## Risultati dell'analisi delle frasi chiave

Un'altra analisi da proporre per il testo preso in considerazione riguarda l'analisi delle frasi chiave, che rappresentano i concetti principali del testo. In questo caso alcune delle frasi chiavi individuate sono: l'annuncio del servizio Quantum AI da parte di Google, l'entusiasmo degli esperti del settore, lo scetticismo degli investitori e l'importanza economica dell'investimento. In questo caso in tutte le frasi chiave abbiamo evidenziata una confidenza superiore al 99%, ovvero il sistema è fortemente sicuro della sua analisi Figura 5.17.



Key phrases	Confidence
January 15, 2024	0.99+
the CES	0.99+
Las Vegas	0.99+
Google	0.99+
the launch	0.99+
its new cloud service Quantum AI	0.98
the advanced technology sector	0.99+
an investment	0.99+
15 billion dollars	0.99+
The announcement	0.99+

Figura 5.17: Risultati dell'analisi delle frasi chiave

Il file JSON corrispondente viene riportato successivamente nella Figura 5.18. L'analisi risulta molto accurata, il sistema riesce a distinguere in maniera corretta le varie frasi chiavi.

```
1  {
2    "KeyPhrases": [
3      {
4        "Score": 0.9999819993972778,
5        "Text": "January 15, 2024",
6        "BeginOffset": 4,
7        "EndOffset": 20
8      },
9      {
10       "Score": 0.9987857937812805,
11       "Text": "the CES",
12       "BeginOffset": 29,
13       "EndOffset": 36
14     },
15     {
16       "Score": 0.9999570846557617,
17       "Text": "Las Vegas",
18       "BeginOffset": 40,
19       "EndOffset": 49
20     },
21     {
22       "Score": 0.9991552829742432,
23       "Text": "Google",
24       "BeginOffset": 51,
25       "EndOffset": 57
26     },
27     {
28       "Score": 0
```

Figura 5.18: File JSON relativo all'analisi delle frasi chiave dell'esempio

## Risultato analisi della lingua

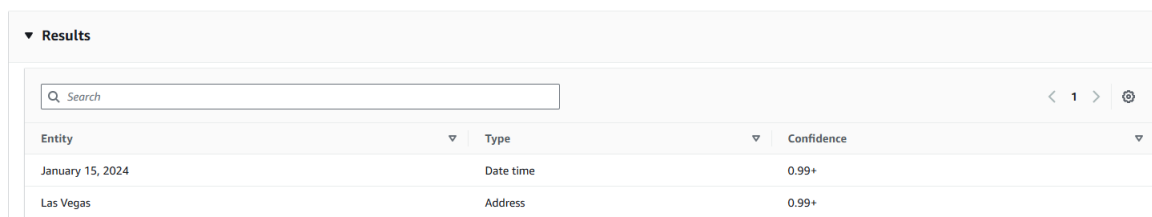
Per il rilevamento della lingua, Amazon con molta facilità è riuscito ad individuare in maniera sicura il linguaggio del testo senza commettere alcun errore e con uno SCORE pressoché pari ad uno. In questo caso individua come lingua l'inglese (Figura 5.19).

```
1  {
2    "Languages": {
3      "LanguageCode": "en",
4      "Score": 0
5        .9968473315238953
6    }
7  }
```

**Figura 5.19:** Risultato dell'analisi della lingua dell'esempio

## Risultati dell'analisi dei PII

Passando ai risultati dell'analisi dei PII, nel testo fornito non ci sono informazioni direttamente identificabili relative a individui (come nomi, indirizzi, numeri di telefono, ecc.). Le PII rilevate sono più relative a date e luoghi (Figura 5.20).



Entity	Type	Confidence
January 15, 2024	Date time	0.99+
Las Vegas	Address	0.99+

**Figura 5.20:** Risultato dell'analisi dei PII

Nella Figura 5.21 viene mostrato il file JSON ottenuto dall'analisi dei PII.

### API response

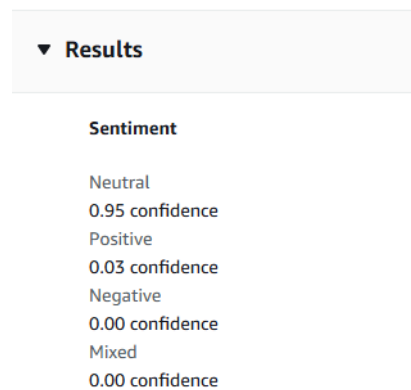
```
1  {
2    "Entities": [
3      {
4        "Score": 0
5          .999846041202
6          5452,
7        "Type":
8          "DATE_TIME",
9        "BeginOffset":
10         : 4,
11        "EndOffset":
12         : 20
13      },
14      {
15        "Score": 0
16          .999987840652
17          4658,
18        "Type":
19          "ADDRESS",
20        "BeginOffset":
21         : 40,
22        "EndOffset":
23         : 49
24      }
25    ]
26  }
```

**Figura 5.21:** File JSON relativo all'analisi dei PII

Non ci sono elementi di PII nel testo fornito, quindi l'analisi dei rischi relativi alla protezione dei dati personali non è necessaria in questo caso.

### Risultati dell'analisi del sentimento

Nel caso dell'analisi dei sentimenti Amazon Comprehend individua nel testo solo sentimento neutrale, come si evince nella Figura 5.22. Infatti, abbiamo una confidenza dello 0,95% per il sentimento neutrale, mentre una confidenza dello 0,03% per il sentimento positivo. Se in alcune parti del testo questa analisi del sentimento può essere considerata giusta, ovvero nella parte in cui si parla, per esempio, del lancio del servizio Quantum AI da parte di Google, per la parte finale del testo questa analisi può essere considerata errata. Infatti, come possiamo notare, nel testo viene espresso un sentimento misto tra positività, da parte degli esperti del settore, e negatività, da parte di alcuni investitori. Di conseguenza, l'analisi effettuata può essere considerata parzialmente sbagliata poiché il sistema non è riuscito a comprendere alcuni dei sentimenti espressi nel testo.



**Figura 5.22:** I risultati dell'analisi del sentimento dell'esempio

Nella Figura 5.23 viene mostrato il file JSON associato all'analisi del sentimento dell'esempio.

```
1  {
2    "Sentiment": {
3      "Sentiment":
4        "NEUTRAL",
5      "SentimentScore":
6        {
7          "Positive": 0
8            .038140270859
9            00307,
10         "Negative": 0
11           .001286778715
12           4391408,
13         "Neutral": 0
14           .958110511302
15           948,
16         "Mixed": 0
17           .002462511649
18           3552923
19       }
20     }
21   }
```

**Figura 5.23:** File JSON relativo ai risultati dell'analisi del sentimento dell'esempio

### Risultati dell'analisi del sentimento mirato

L'analisi del sentimento mirato rappresenta un ulteriore passo rispetto a quella precedente. In questo caso sono stati puntualmente evidenziati i sentimenti espressi dalle principali entità

presenti nel testo. Anche in questo caso, come nel caso precedente, troviamo sempre una difficoltà nella comprensione del sentimento presente nel testo (Figura 5.24). Infatti, il sistema individua un sentimento neutrale appartenente a tutte le entità. Invece, riguardo alla entità "experts", si dovrebbe associare un sentimento positivo, mentre riguardo l'entità "investors" un sentimento negativo; invece, il sistema non è riuscito a riconoscere questi sentimenti, ma ha associato tutto il testo ad un sentimento neutrale.

Entity	Entity type	Entity score	Primary sentiment	Positive score	Negative score	Neutral score	Mixed score
On January 15, 2024	DATE	0.99+	NEUTRAL	0.00	0.00	0.99+	0.00
CES	EVENT	0.94	NEUTRAL	0.00	0.00	0.99+	0.00
Las Vegas	LOCATION	0.99+	NEUTRAL	0.00	0.00	0.99+	0.00
Google	ORGANIZATION	0.99+	NEUTRAL	0	0	1.00	0
its	ORGANIZATION	0.52	NEUTRAL	0	0	1.00	0
15 billion dollars	QUANTITY	0.99+	NEUTRAL	0	0.00	0.99+	0.00
industry	ORGANIZATION	0.97	NEUTRAL	0	0	1.00	0
experts	PERSON	0.99+	NEUTRAL	0.00	0	0.99+	0
some	QUANTITY	0.99+	NEUTRAL	0	0.00	0.99+	0
investors	PERSON	0.58	NEUTRAL	0	0	1.00	0

**Figura 5.24:** I risultati dell'analisi del sentimento mirato

Di seguito viene riportata parte del file JSON inerente all'analisi del sentimento mirato, in particolare riguardo alla parte di testo "On January 15".

```

1  {
2    "Entities": [
3      {
4        "DescriptiveMentionIndex": [
5          0
6        ],
7        "Mentions": [
8          {
9            "Score": 0.9999240040779114,
10           "GroupScore": 1,
11           "Text": "On January 15, 2024",
12           "Type": "DATE",
13           "MentionSentiment": {
14             "Sentiment": "NEUTRAL",
15             "SentimentScore": {
16               "Positive": 9.999999974752427e-7,
17               "Negative": 0.000003000000106112566,
18               "Neutral": 0.9999939799308777,
19               "Mixed": 9.999999974752427e-7
20             }
21           },
22           "BeginOffset": 0,
23           "EndOffset": 19
24         }
25       ]
26     }
  
```

**Figura 5.25:** File JSON relativo ai risultati dell'analisi del sentimento mirato dell'esempio

## Risultati dell'analisi della sintassi

Ultima analisi fornita da Amazon Comprehend è quella della sintassi all'interno della console. Dall'analisi effettuata dal sistema possiamo evincere che il testo è composto da due frasi complesse. La prima frase è la frase participiale, e specifica lo scopo del lancio del servizio Quantum AI e un complemento per indicare l'entità dell'investimento. La seconda frase introduce un contrasto tra le opinioni di esperti e investitori utilizzando una subordinata (while) e una frase coordinata per presentare punti di vista opposti. Ogni elemento viene poi associato al suo scopo all'interno del testo. In questo caso Amazon ha effettuato in maniera corretta l'analisi della frase, comprendendo pienamente la sua composizione sintattica.

Word	Part of speech	Confidence	Word	Part of speech	Confidence
new	Adjective	0.99+	with	Adposition	0.99+
cloud	Noun	0.99+	an	Determiner	1.00
service	Noun	1.00	investment	Noun	0.99+
Quantum	Proper noun	0.99+	of	Adposition	1.00
AI	Proper noun	0.99+	15	Numeral	1.00
,	Punctuation	1.00	billion	Numeral	1.00
aimed	Verb	1.00	industry	Noun	1.00
at	Subordinating conjunction	1.00	experts	Noun	1.00
revolutionizing	Verb	0.99+	are	Verb	0.99+
the	Determiner	1.00	excited	Adjective	0.99+
*	Punctuation	0.93	about	Adposition	0.99+
On	Adposition	0.92	the	Determiner	1.00
January	Proper noun	1.00	potential	Noun	0.99+
15	Numeral	1.00	,	Punctuation	1.00
,	Punctuation	1.00	some	Determiner	0.99+
2024	Numeral	1.00	investors	Noun	1.00
,	Punctuation	1.00	remain	Verb	1.00
during	Adposition	1.00	skeptical	Adjective	0.99+

**Figura 5.26:** I risultati dell'analisi della sintassi (prima parte)

the	Determiner	1.00	about	Adposition	0.99+
CES	Proper noun	0.98	the	Determiner	0.99+
in	Adposition	1.00	associated	Verb	0.99+
Las	Proper noun	0.99+	risk	Noun	0.99+
Vegas	Proper noun	0.99+	,	Punctuation	0.99+
,	Punctuation	1.00	*	Punctuation	1.00
Google	Proper noun	0.99+	dollars	Noun	0.99+
announced	Verb	1.00	,	Punctuation	1.00
the	Determiner	1.00	The	Determiner	1.00
launch	Noun	1.00	announcement	Noun	1.00
of	Adposition	1.00	received	Verb	1.00
its	Pronoun	1.00	a	Determiner	1.00
advanced	Adjective	0.97	mixed	Adjective	0.95
technology	Noun	1.00	reception	Noun	1.00
sector	Noun	1.00	:	Punctuation	1.00
,	Punctuation	1.00	while	Subordinating conjunction	0.99+

**Figura 5.27:** I risultati dell'analisi della sintassi (seconda parte)

L'analisi evidenzia una sintassi ben bilanciata, con frasi complesse ma ben organizzate, dove ciascun elemento sintattico svolge un ruolo preciso nel costruire il significato.

---

## Discussione in merito alle esperienze condotte

---

*In quest'ultimo capitolo ci soffermeremo a discutere riguardo ai risultati delle analisi dei casi d'uso effettuate sugli strumenti di Intelligenza Artificiale messi a disposizione da Amazon Web Services. In particolare, trarremo delle considerazioni finali riguardanti l'utilizzo dei Chatbot, il riconoscimento facciale e l'analisi dei testi.*

### 6.1 Discussione in merito all'utilizzo dei Chatbot

L'utilizzo dei Chatbot sta diventando sempre più presente nelle nostre vite, permeando in maniera consistente molti aspetti del nostro fare quotidiano. L'evoluzione dei Chatbot è continua, con l'introduzione di tecnologie avanzate come il Natural Language Processing e il Machine Learning, che migliorano la loro capacità di comprendere il linguaggio naturale e rispondere in modo più efficace. Questo rappresenta una rivoluzione significativa nell'interazione tra utenti umani e sistemi digitali. L'utilizzo degli assistenti virtuali non ha semplificato solo l'accesso alle informazioni, ma ha anche ribaltato il concetto di comunicazione online. La capacità dei Chatbot di comprendere e rispondere in maniera personalizzata alle domande degli utenti ha introdotto un nuovo livello di tempestività nell'erogazione di servizi e aiuti. In effetti, uno tra i tanti benefici che i Chatbot hanno apportato alla vita quotidiana è il fatto che sono disponibili 24 ore su 24 garantendo un supporto immediato a qualsiasi ora del giorno senza dover attendere che un operatore sia disponibile. Inoltre, grazie ai Chatbot che si dedicano alle attività ripetitive, le aziende possono liberare risorse umane per affrontare problemi più complessi. Questa è una grande risorsa per le aziende, poiché migliora la soddisfazione del cliente e può portare a una maggiore fidelizzazione. Ciò che rende questi assistenti virtuali così potenti sono due elementi cardine di tutti gli sviluppi dell'IA, ovvero la disponibilità di un immenso potere computazionale e l'accessibilità di una straordinaria quantità di dati, che permettono ai Chatbot di allenarsi continuamente apportando sempre delle migliorie nel loro apprendimento.

Tuttavia, non possiamo ignorare le limitazioni dei Chatbot. Infatti, questi ultimi, per quanto avanzati, non riescono a provare sentimenti ed essere empatici con l'interlocutore. Per questo risulta difficile il loro utilizzo in determinati campi che richiedono più comprensione emotiva durante l'interazione. D'altronde un Chatbot non può sostituire l'empatia e la comprensione che un operatore umano può offrire nel momento di difficoltà. Questo è un aspetto fondamentale da considerare, specialmente in settori come quello sanitario, dove le emozioni giocano un ruolo cruciale. Infatti, è importante sottolineare come i Chatbot non debbano essere presi come sostituti e che la loro utilità risiede nella capacità di essere

dispositivi ausiliari. Inoltre, mentre la presenza di questi assistenti virtuali diventa sempre più pervasiva, è cruciale valutare attentamente gli impatti sociali, etici e la qualità complessiva dell'esperienza utente.

Un aspetto che solleva preoccupazioni è la privacy. I Chatbot, spesso, raccolgono dati sensibili degli utenti, il che implica la necessità di implementare misure di sicurezza adeguate. Gli utenti devono sentirsi al sicuro quando interagiscono con queste tecnologie. Fatte queste premesse possiamo affermare che l'utilizzo dei Chatbot presenta sia opportunità che sfide. Sono strumenti potenti che possono migliorare l'efficienza e la soddisfazione del cliente, ma non possono sostituire completamente l'uomo. La chiave per un utilizzo efficace risiede nella progettazione di Chatbot che possano integrare le capacità umane, mantenendo un equilibrio tra automazione e interazione personale.

A questo proposito un ruolo importante nella creazione dei Chatbot è la piattaforma Amazon Lex, un servizio offerto gratuitamente da Amazon AWS. Nel terzo capitolo, infatti, attraverso questa piattaforma siamo riusciti a creare un Chatbot in grado di permettere l'ordinazione in un Fast Food. Questo è stato possibile poiché Amazon Lex presenta una struttura semplice da comprendere e utilizzare, avendo a disposizione anche vari manuali che permettono di comprendere al meglio il suo funzionamento. Ciò ha permesso a più persone di poter accedere a questo servizio e usufruirne. Infatti, prima dell'arrivo di strumenti come Lex, sviluppare Chatbot o assistenti vocali richiedeva una profonda conoscenza di linguistica computazionale, Intelligenza Artificiale e Machine Learning, oltre a significative risorse tecniche ed economiche. Con Amazon Lex si è semplificato questo processo, fornendo un'interfaccia che permette a sviluppatori non specializzati di creare chatbot e sistemi di risposta automatizzata con relativamente poca fatica.

## 6.2 **Discussione in merito al riconoscimento facciale**

Il riconoscimento facciale è una tecnologia che consente di identificare o verificare l'identità di una persona analizzando e confrontando le caratteristiche facciali. La capacità di una macchina di riconoscere, elaborare e compiere operazioni complesse a partire da immagini di volti umani completamente in autonomia può avere risvolti importantissimi per lo sviluppo dell'innovazione. Il riconoscimento facciale ad oggi è sempre in continua evoluzione, con miglioramenti nella precisione e nell'affidabilità dei sistemi. Però, può anche portare delle problematiche e diventare pericoloso, ridefinendo il confine tra pubblico e privato. Mentre i governi e le aziende corrono verso l'adozione di questa tecnologia, è cruciale che si tengano in considerazione i suoi limiti etici e sociali. Infatti, risulta necessario un bilanciamento tra sicurezza e libertà personale. Da un lato, il riconoscimento facciale può contribuire a una società più sicura, permettendo alle forze dell'ordine di agire in modo più efficace. Dall'altro, esiste il rischio che la sicurezza venga utilizzata come giustificazione per un controllo eccessivo delle persone. Infatti, risulta fondamentale la regolamentazione del riconoscimento facciale. Attualmente, la legislazione che disciplina l'uso di questa tecnologia varia da un paese all'altro, e spesso le normative sono ancora in una fase di elaborazione. In Europa, il Regolamento Generale sulla Protezione dei Dati (GDPR) fornisce un quadro legale rigoroso per il trattamento dei dati biometrici, tra cui i volti, imponendo limiti stringenti sull'utilizzo di queste informazioni senza il consenso esplicito degli individui. Tuttavia, affinché questa tecnologia possa essere adottata su larga scala senza generare una reazione di rigetto, è indispensabile che si sviluppi una cultura della trasparenza. Le persone devono essere consapevoli di quando e come vengono utilizzate le loro informazioni biometriche e devono avere il diritto di rifiutare o di controllare l'accesso ai propri dati.

Riguardo questa tematica nel terzo capitolo, per quanto abbiamo potuto constatare dalle demo di Rekognition forniteci da Amazon Web Services, il riconoscimento dei volti

con l'Intelligenza Artificiale procede in un'ottima direzione, sia dal punto di vista dello sviluppo, che da quello dei suoi utilizzi non negativi. Con l'utilizzo di Amazon Rekognition e compiendo numerose prove con diverse immagini, abbiamo potuto constatare la sua capacità nell'identificare in maniera corretta gran parte delle caratteristiche dei volti che appaiono nelle immagini che diamo in input ad esso. Infatti, nonostante le diverse immagini con volti più o meno riconoscibili, la piattaforma ha effettuato sempre un'ottima analisi ed estrazione dell'informazione dai volti dei soggetti. A questo proposito Amazon Rekognition potrebbe essere, quindi, utilizzato in moltissimi settori per aumentare la produttività, infatti, la sua capacità di automatizzare l'analisi delle immagini e dei video ha ridotto il tempo e le risorse necessarie per eseguire compiti che prima richiedevano un intervento manuale significativo. Ad esempio, nel mondo dei media e dell'intrattenimento, Rekognition può essere utilizzato per "taggare" automaticamente le immagini, identificare persone o oggetti e ordinare grandi quantità di contenuti visivi, rendendo più efficiente la gestione dei dati multimediali.

### **6.2.1 Discussione in merito all'analisi dei testi**

L'analisi dei testi è un tema sempre più rilevante in un mondo in cui la quantità di dati testuali, generati continuamente attraverso e-mail, social media, recensioni online, articoli e molto altro, è in costante crescita. Dai documenti aziendali ai post sui social media, i testi contengono un'enorme quantità di dati che, se analizzati correttamente, possono fornire insight preziosi. Questa disciplina consente di estrarre informazioni significative da grandi quantità di testi, combinando competenze linguistiche, informatiche e statistiche. L'analisi dei testi è cruciale perché il linguaggio scritto è una delle forme di comunicazione più ricche e complesse. Con l'esplosione delle informazioni disponibili online, è diventato impossibile per un essere umano analizzare manualmente tutto il materiale testuale prodotto. Attraverso l'analisi dei testi è possibile identificare pattern, rilevare sentimenti, comprendere tendenze e scoprire informazioni nascoste che altrimenti sarebbero difficili da cogliere.

Questa disciplina ha un impatto significativo non solo a livello commerciale, ma anche sociale. Può essere utilizzata per monitorare l'opinione pubblica su temi politici, sociali o economici, fornendo strumenti per comprendere l'evoluzione dei sentimenti collettivi. Ad esempio, può aiutare a prevedere tendenze elettorali o a rilevare movimenti di protesta attraverso l'analisi dei social media. Con l'evoluzione del Machine Learning e l'uso crescente dell'Intelligenza Artificiale conversazionale, l'analisi dei testi diventerà sempre più accurata e versatile.

Le tecnologie future promettono di comprendere meglio il contesto, ridurre i bias e gestire la complessità del linguaggio umano, con applicazioni che spazieranno dalla sanità alla giustizia, dall'educazione al commercio. Riguardo questa tematica Amazon Comprehend rappresenta un valido sistema di analisi ed estrazione delle informazioni. Infatti, attraverso una serie di frasi e testi abbiamo potuto testare questa piattaforma e collaudare la capacità di Amazon Comprehend di effettuare l'analisi in maniera eccellente. Il suo impatto è evidente in una vasta gamma di settori, dove può migliorare l'efficienza operativa, ridurre i tempi di elaborazione delle informazioni e aiutare le aziende a prendere decisioni basate sui dati.

Tuttavia, è necessario affrontare questioni legate alla privacy, all'accuratezza e alla gestione dei dati per garantire che l'uso di questa tecnologia sia etico e responsabile. Ciò dimostra che si tratta di una tecnologia all'avanguardia, necessaria per le aziende che desiderano migliorare i propri prodotti o servizi basandosi sui feedback reali forniti dalle recensioni dei clienti.

In questa tesi ci siamo concentrati, nel primo capitolo, sull'esplorazione dell'Intelligenza Artificiale, esaminandone gli aspetti principali e le applicazioni pratiche. Abbiamo iniziato in primis con una definizione generale dell'IA, evidenziando una differenza tra Intelligenza Artificiale Forte e Debole, ovvero due differenti tipi di Intelligenza, una più autonoma e indipendente e l'altra più comandata. Abbiamo ripercorso le varie tappe storiche di questa affascinante disciplina, approfondendo alcuni test classici, come il test di Turing e il test della stanza cinese, proponendo, anche, le opinioni dei vari studiosi riguardo la riuscita o meno della creazione di una macchina pensante. In seguito abbiamo messo in luce, oltre ai numerosi aspetti positivi che l'utilizzo dell'Intelligenza Artificiale porta con sé, anche le problematiche che un uso sconsiderato di questa disciplina può comportare. Risulta necessario, infatti, creare e adottare delle leggi che regolamentano l'utilizzo di questa tecnologia rivoluzionaria.

Successivamente, nel secondo capitolo, abbiamo presentato una branca dell'Intelligenza Artificiale, ovvero il Natural Language Processing, in italiano Elaborazione del Linguaggio Naturale, che ha apportato enormi cambiamenti e migliorie nell'interpretazione del linguaggio naturale da parte delle macchine. Anche in questo caso abbiamo presentato un excursus storico di questa tecnologia, spiegando, poi, il suo funzionamento, ovvero come i sistemi artificiali riescano a comprendere il linguaggio naturale, attraverso varie fasi di analisi del testo. Inoltre, sono stati evidenziati gli innumerevoli vantaggi che questa branca porta con sé. Risulta, infatti, fondamentale il suo utilizzo in numerosissime applicazioni del fare quotidiano; basti pensare anche solo agli assistenti virtuali, che svolgono un ruolo cruciale nelle nostre vite.

Infatti, nel terzo capitolo, abbiamo trattato il tema dei Chatbot, ovvero delle piattaforme di messaggistica istantanea. Di questo argomento si è cercato di fare un breve excursus storico, proponendo i vari passi cruciali dello sviluppo dei Chatbot che li ha portati ad essere così fondamentali oggi. Il primo sistema che si può considerare Chatbot è nato, addirittura, nei primi anni '60. Venne chiamato ELIZA e rappresenta il primo passo verso questa rivoluzione nel campo dell'IA. In seguito si è spiegato il loro funzionamento in maniera sintetica e differenziando tra Chatbot dichiarativi e predittivi. Anche in questo caso si sono accennati i vantaggi e gli svantaggi propri di ogni aspetto e branca dell'Intelligenza Artificiale. Infine, sono state proposte alcune delle loro varie applicazioni. Una piattaforma fondamentale nel campo della creazione dei Chatbot è Amazon AWS, che attraverso Amazon Lex, permette a tutti di sperimentare e creare dei Chatbot, grazie alla presenza di una interfaccia semplice e intuitiva. Riguardo questa applicazione abbiamo in primis spiegato cos'è e il suo funzionamento, successivamente abbiamo creato un Chatbot conversazionale che permettesse di ordinare vari menù da un Fast Food. Sono stati spiegati i vari passaggi

per la creazione di questo Chatbot e si è mostrata una conversazione standard che potrebbe avvenire se si decide di ordinare tramite esso uno dei possibili alimenti messi a disposizione.

Nel capitolo successivo, ovvero il quarto, ci siamo concentrati, invece, sulla branca dell'Intelligenza Artificiale chiamata Riconoscimento Facciale. Essa permette, attraverso l'IA, di riconoscere e analizzare i volti di persone che vengono mostrati in delle immagini. Abbiamo cercato di ripercorrere la sua storia nei vari periodi mettendo in evidenza gli avvenimenti più importanti dalla sua nascita ad ora. In seguito, abbiamo spiegato il suo funzionamento e i vantaggi e gli svantaggi nel suo utilizzo. Inoltre, abbiamo proposto una serie di casi in cui il riconoscimento facciale può essere utilizzato. In questo campo dell'IA un ruolo importante è attribuito ad Amazon Rekognition, un software messo a disposizione da Amazon AWS che permette di analizzare e riconoscere i volti. A questo proposito abbiamo fatto una serie di prove con varie immagini per vedere come la piattaforma si sarebbe comportata nei vari casi. Anche per le immagini in cui i volti erano più difficili da riconoscere, Rekognition non ha avuto nessuna difficoltà nell'analizzarli, permettendoci, quindi, di avere le varie informazioni di cui avevamo bisogno. È stato, anche, proposto un confronto facciale tra due immagini che ritraevano la stessa persona ma in periodi diversi, e anche in questo caso la piattaforma ha avuto una performance ottimale.

Nel quinto capitolo, invece, abbiamo affrontato il tema dell'analisi dei testi proponendo come piattaforma Amazon Comprehend, un servizio di elaborazione del linguaggio naturale offerto da AWS. Abbiamo fatto una breve introduzione spiegando cosa fosse e come funzionasse, successivamente lo abbiamo testato attraverso vari testi per ogni insight. In tutti i casi ha dimostrato di riuscire ad analizzare ed estrarre informazioni dai testi in maniera impeccabile.

La tesi si conclude con una riflessione sulle esperienze di utilizzo dell'Intelligenza Artificiale condotte durante il nostro studio. Abbiamo dimostrato come l'IA possa essere implementata efficacemente in vari contesti, evidenziando il suo potenziale di trasformazione e le opportunità future che offre. Infatti, di giorno in giorno avvengono degli enormi sviluppi e innovazioni che caratterizzano l'IA. Se si procede di questo passo in futuro si potrebbe anche assistere allo sviluppo della cosiddetta AGI (Artificial General Intelligence), ossia l'Intelligenza Artificiale Generale, che permetterebbe lo sviluppo di macchine munite di pensiero autonomo. Questo rappresenterebbe un traguardo sensazionale; infatti numerosissime imprese e organizzazioni stanno investendo ingenti risorse per fare in modo che questo avvenga. Nonostante i benefici della AGI siano sicuramente importanti, bisogna sempre tener conto che, proprio per il fatto che queste tecnologie permettano di fare tutto, esse potrebbero essere usate per scopi malevoli. Per questo risulta fondamentale, come già affermato in precedenza, la presenza di leggi che regolamentino il loro funzionamento, permettendo uno sviluppo responsabile e collaborativo di questa tecnologia, evitando impatti etici e sociali negativi.

- ASIMOV, I. (1950), *Io, robot*, Mondadori.
- BODEN, M. (2019), *L'Intelligenza Artificiale*, Il Mulino.
- BORANA, J. (2007), «Applications of artificial intelligence & associated technologies», Rap. tecn., Department of Electrical Engineering, Jodhpur National University.
- BRACCINI, M. e ROLI, A. (2013), *Intelligenza artificiale: test di Turing e alcune questioni filosofiche*, Tesi di dottorato, DUniversità di Bologna, Corso di Studio in Ingegneria elettronica, informatica e telecomunicazione, Sweden.
- CHEN, X. e JIN, L. (2021), *Text recognition in the wild: A survey*, ACM Computing Surveys.
- EGMONT-PETERSEN, M. (2002), «Image processing with neural networks—a review», Rap. tecn., Pattern recognition.
- FORD, C. (2024), *Amazon Web Services: A Complete Guide*, Madeley publishing.
- KAPLAN, J. (2017), *Intelligenza artificiale. Guida al futuro prossimo*, Luiss.
- LU, D. (2007), «A survey of image classification methods and techniques for improving classification performance», Rap. tecn., International journal of Remote sensing.
- MINSKY, M. (1969), *Perceptrons: an introduction to computational geometry*, The MIT Press.
- RUSSELL, S. (2010), *Artificial Intelligence a modern approach*, Pearson Education.
- SEARLE, J. (1984), *Minds, brains, and programs*, THE BEHAVIORAL AND BRAIN.
- SHADIEV, R., WU-YUIN HWANG, N.-S. C. e HUANG, Y.-M. (2014), «Review of speech-to-text recognition technology for enhancing learning», Rap. tecn., Journal of Educational Technology & Society.
- TEGMARK, M. (2018), *Life 3.0: Being Human in the Age of Artificial Intelligence*, Penguin.
- TURING, A. (1950), «Computing machinery and intelligence», *Mind*.

## Siti web consultati

- Amazon Web Services - [docs.aws.amazon.com](https://docs.aws.amazon.com)
- Enciclopedia Treccani - [www.treccani.it](http://www.treccani.it)
- Blog Osservatori - <https://blog.osservatori.net>
- IBM - [www.ibm.com](http://www.ibm.com)
- Treccani - [www.treccani.it](http://www.treccani.it)
- Wikipedia - [www.wikipedia.org](http://www.wikipedia.org)
- Namirial - <https://focus.namirial.it>
- AndreaMinini - <https://www.andreaminini.com>
- YouTube - <https://www.youtube.com/>
- Marino Luigi - <https://www.marinoluigi.it/u>

---

## Ringraziamenti

---

Desidero esprimere la mia sincera gratitudine a tutte le persone che mi hanno supportato durante la stesura di questa tesi.

Il primo ringraziamento va a ai miei genitori, che mi hanno dato la possibilità di intraprendere e portare a termine questo lungo e difficile percorso universitario.

Un altro importantissimo ringraziamento va a Diego, che mi è stato sempre vicino e mi ha supportata ed aiutata.

Ringrazio anche Alessia, la mia coinquilina, che dal secondo anno è stata molto presente e mi ha aiutato costantemente nei momenti di bisogno.

Ringrazio mia sorella, che durante la stesura di questa tesi mi ha fornito importanti consigli.

Vorrei ringraziare il mio relatore, il Professore Ursino, che mi ha accompagnato nel lavoro della stesura della tesi seguendomi costantemente con estrema disponibilità.

Un pensiero va anche ai mie amici e amiche, che con la loro presenza e condivisione hanno reso questo viaggio più stimolante e arricchente.

Infine ringrazio tutti coloro che, direttamente o indirettamente, hanno contribuito alla realizzazione di questa tesi e del mio percorso universitario.