

**UNIVERSITÀ POLITECNICA DELLE MARCHE**  
**FACOLTÀ DI INGEGNERIA**

Dipartimento di Ingegneria dell'Informazione  
Corso di Laurea in Ingegneria Informatica e dell'Automazione

---



**TESI DI LAUREA**

**Progettazione e implementazione di casi di studio nel contesto  
dell'Intelligenza Artificiale: Chatbot, Riconoscimento di Immagini,  
Estrazione di testi da immagini**

**Design and implementation of case studies in the context of  
Artificial Intelligence: Chatbots, Image Recognition, Text extraction  
from images**

Relatore

Prof. Domenico Ursino

Candidato

Niccolò Ciotti

---

**ANNO ACCADEMICO 2022-2023**

*Vedo il momento in cui saremo per i robot cosa sono i cani per gli umani  
e io faccio il tifo per le macchine*

Claude Shannon

## Sommario

Negli ultimi anni, l'Intelligenza Artificiale ha guadagnato un'importanza sempre maggiore, diffondendosi in vari contesti grazie allo sviluppo di sistemi in grado di apprendere autonomamente concetti e abilità umane. Questi avanzamenti hanno semplificato e ottimizzato numerosi processi, sia a livello aziendale che nella vita quotidiana. Questa tesi si propone di esplorare le caratteristiche distintive dell'Intelligenza Artificiale e di analizzarne le applicazioni in alcuni settori specifici, contribuendo, così, ad una comprensione più approfondita delle sue dinamiche e delle sue implicazioni.

**Keyword:** Intelligenza Artificiale, Machine Learning, Cloud Computing, Chatbot, Image Recognition, Text Recognition.

<b>Introduzione</b>	<b>1</b>
<b>1 Introduzione all'Intelligenza Artificiale</b>	<b>3</b>
1.1 Cos'è l'Intelligenza Artificiale . . . . .	3
1.1.1 Definizione di IA . . . . .	3
1.1.2 Obiettivi e applicazioni dell'IA . . . . .	4
1.1.3 Tipi di Intelligenza Artificiale . . . . .	5
1.2 La storia dell'Intelligenza Artificiale . . . . .	6
1.2.1 Origini . . . . .	6
1.2.2 Test di Turing . . . . .	6
1.2.3 1956-1969: I successi e l'entusiasmo dei primi anni . . . . .	8
1.2.4 Stato Attuale dell'IA . . . . .	9
1.3 Benefici e Rischi dell'IA . . . . .	10
1.3.1 Vantaggi dell'IA . . . . .	10
1.3.2 Rischi e preoccupazioni . . . . .	11
<b>2 Amazon AWS</b>	<b>13</b>
2.1 Introduzione al Cloud Computing . . . . .	13
2.1.1 Cos'è . . . . .	13
2.1.2 Servizi e Utilizzi . . . . .	14
2.1.3 Vantaggi e Sfide del Cloud Computing . . . . .	15
2.2 Introduzione ad Amazon AWS . . . . .	16
2.2.1 Che cos'è . . . . .	16
2.2.2 Storia . . . . .	16
2.2.3 Come funziona . . . . .	17
2.3 Scenari Applicativi di Amazon AWS . . . . .	18
2.3.1 Servizi a disposizione . . . . .	18
2.3.2 Utilizzi e Applicazioni . . . . .	19
<b>3 Esperienze nell'ambito dei chatbot</b>	<b>21</b>
3.1 Introduzione ai chatbot . . . . .	21
3.1.1 Cos'è un chatbot . . . . .	21
3.1.2 Come funzionano . . . . .	21
3.2 Caso di studio: Amazon Lex . . . . .	22
3.2.1 Che cos'è . . . . .	22

---

3.2.2	Come funziona . . . . .	23
3.3	Realizzazione di un chatbot con Amazon Lex . . . . .	24
3.3.1	Creazione . . . . .	24
3.3.2	Gestione delle conversazioni . . . . .	25
3.3.3	Utilizzo . . . . .	30
<b>4</b>	<b>Esperienze nell'ambito dell'Image Recognition</b>	<b>33</b>
4.1	Introduzione al riconoscimento di immagini . . . . .	33
4.1.1	Cos'è l'Image Recognition . . . . .	33
4.1.2	Come funziona . . . . .	34
4.2	Caso di studio: Amazon Rekognition . . . . .	35
4.2.1	Che cos'è . . . . .	35
4.2.2	Funzionalità . . . . .	35
4.3	Test di Amazon Rekognition . . . . .	37
4.3.1	Creazione del dataset . . . . .	37
4.3.2	Labelling delle immagini . . . . .	38
4.3.3	Addestramento del modello . . . . .	39
4.3.4	Valutazione del modello . . . . .	39
<b>5</b>	<b>Esperienze nell'ambito del Text Recognition</b>	<b>43</b>
5.1	Introduzione al riconoscimento di testi . . . . .	43
5.1.1	Cos'è . . . . .	43
5.1.2	Come funziona . . . . .	43
5.2	Caso di studio: Amazon Textract . . . . .	44
5.2.1	Che cos'è . . . . .	44
5.2.2	Come funziona . . . . .	45
5.3	Test di Amazon Textract . . . . .	46
5.3.1	Riconoscimento ed estrazione di testi da immagini . . . . .	46
<b>6</b>	<b>Discussione in merito alle esperienze condotte</b>	<b>54</b>
6.1	Discussione sull'utilizzo dei chatbot . . . . .	54
6.2	Discussione sul riconoscimento delle immagini . . . . .	55
6.3	Discussione sul rilevamento ed estrazione di testi da immagini . . . . .	56
	<b>Conclusioni</b>	<b>57</b>
	<b>Bibliografia</b>	<b>58</b>
	<b>Ringraziamenti</b>	<b>60</b>

---

## Elenco delle figure

---

1.1	Kasparov vs. Deep Blue . . . . .	6
1.2	Illustrazione del funzionamento del Test di Turing . . . . .	7
1.3	Il modello originale di Pandemonium . . . . .	8
1.4	GPT-4 . . . . .	9
1.5	Esempio di veicolo autonomo . . . . .	12
2.1	Illustrazione del Cloud . . . . .	14
2.2	Mappa dell'infrastruttura AWS . . . . .	17
2.3	Alcuni dei servizi di AWS . . . . .	19
3.1	Creazione di un chatbot Amazon Lex . . . . .	24
3.2	Scelta della lingua e della voce del chatbot . . . . .	25
3.3	Lista degli intenti del chatbot . . . . .	26
3.4	Tipi di Slot del chatbot . . . . .	26
3.5	Tipologie dello Slot "RoomType" . . . . .	27
3.6	Slot dell'intent "BookHotel" . . . . .	27
3.7	Codice Lambda per il salvataggio su database (prima parte) . . . . .	28
3.8	Codice Lambda per il salvataggio su database (seconda parte) . . . . .	29
3.9	Esempio di una conversazione con il chatbot (prima parte) . . . . .	30
3.10	Esempio di una conversazione con il chatbot (seconda parte) . . . . .	30
3.11	Schermata per l'aggiunta di un canale dalla console Amazon Lex . . . . .	31
3.12	Schermata per la configurazione di un canale dalla console Amazon Lex . . . . .	32
4.1	Configurazione di un dataset . . . . .	37
4.2	Esempio di immagini utilizzate per l'addestramento . . . . .	38
4.3	Lista delle etichette . . . . .	38
4.4	Esempio di immagini con le etichette personalizzate . . . . .	39
4.5	Schermata di training completato . . . . .	39
4.6	Risultati del modello . . . . .	40
4.7	Comando per avviare il modello nella console AWS CloudShell . . . . .	41
4.8	Comando per analizzare l'immagine nella console AWS CloudShell . . . . .	41
4.9	Comando per arrestare il modello nella console AWS CloudShell . . . . .	41
4.10	Immagine utilizzata per il test del modello . . . . .	41
4.11	Schermata con i risultati del test . . . . .	42
5.1	Esempio di un blocco di tipo PAGE in formato JSON . . . . .	46

---

5.2	Schermata per il caricamento di un documento di Amazon Textract . . . . .	48
5.3	Schermata per la scelta della funzionalità . . . . .	48
5.4	Interfaccia di analisi ed estrazione di un documento . . . . .	49
5.5	Immagine utilizzata per il test di Amazon Textract . . . . .	50
5.6	Comando per l'installazione della libreria <i>boto3</i> . . . . .	50
5.7	Codice per l'apertura e la visualizzazione dell'immagine . . . . .	51
5.8	Codice per la lettura dell'immagine e il rilevamento del testo . . . . .	51
5.9	Testo grezzo rilevato ed estratto dall'immagine . . . . .	51
5.10	Codice per il rilevamento della firma . . . . .	52
5.11	Codice per disegnare dei rettangoli delimitatori all'interno dell'immagine . .	52
5.12	Risultato dell'immagine modificata tramite Amazon Textract . . . . .	53

L'Intelligenza Artificiale, abbreviato comunemente in IA, è una disciplina affascinante che incarna la sfida umana di dare vita alle macchine, conferendo loro la capacità di apprendere, ragionare e adattarsi in modi che tradizionalmente erano riservati alla sola mente umana. Questo campo multidisciplinare si estende dalla teoria dei giochi all'apprendimento automatico, dalla visione artificiale all'elaborazione del linguaggio naturale, e rappresenta una delle frontiere più intriganti della ricerca scientifica e tecnologica contemporanea.

L'IA si manifesta come l'espressione della mente umana nell'ambito digitale, dove algoritmi apprendono, ragionano e si adattano in maniera simile all'essere umano. Essa offre la prospettiva di una mente artificiale capace di apprendere dai dati, adattarsi alle novità e superare i limiti umani in termini di analisi e previsione.

Le applicazioni dell'Intelligenza Artificiale si estendono ampiamente a settori diversificati, tra cui medicina, sicurezza, giochi, industria, e molti altri ancora. Tuttavia, la tematica dell'Intelligenza Artificiale si è rivelata, e continua a essere, al centro di intensi dibattiti nel contesto scientifico. Alcuni vedono l'IA come il motore di un progresso straordinario, mentre altri sollevano interrogativi sulle implicazioni etiche legate all'automazione delle decisioni e alle potenziali minacce per la privacy e la sicurezza.

Questa tesi si concentra sull'esplorazione approfondita della vastità dell'IA, illuminando le sue potenzialità rivoluzionarie e investigando le sfide etiche e sociali connesse alla sua diffusione. L'obiettivo principale è sondare il terreno dell'IA, evidenziando le opportunità che offre ma, al contempo, interrogandoci sulle responsabilità etiche correlate alla sua implementazione su vasta scala. Infine, saranno sperimentate ed analizzate alcune delle più importanti applicazioni di IA attraverso dei casi di studio, evidenziandone, con un approccio critico, vantaggi e limiti principali.

Nello specifico, dunque, la presente tesi è composta da sette capitoli, strutturati come di seguito specificato:

- Nel Capitolo 1 verrà introdotto il mondo dell'Intelligenza Artificiale, esplorando le sue definizioni e accennando ad alcuni concetti riguardanti la sua storia. Infine, saranno presentate le principali applicazioni, i benefici e i rischi del suo utilizzo.
- Nel Capitolo 2 si esplorerà il vasto mondo del Cloud Computing, esplorando i suoi concetti fondamentali, i servizi offerti e i vantaggi che offre alle organizzazioni. Si approfondiranno, inoltre, le sfide associate all'impiego di questa tecnologia. Successivamente, ci si concentrerà su uno dei principali leader di questo settore, ovvero Amazon Web Services (AWS), esaminandone la storia e analizzando i servizi messi a disposizione e i relativi scenari applicativi.



- Nel Capitolo 3 verrà approfondito il tema dei chatbot ed il loro funzionamento. Sarà presentato un caso di studio specifico, ovvero Amazon Lex, con l'obiettivo di esplorare dettagliatamente il suo funzionamento. Infine, si illustrerà un'esperienza pratica effettuata con questo strumento, offrendo una visione concreta delle sue applicazioni e delle dinamiche interattive che può generare.
- Nel Capitolo 4 verrà approfondito il campo che si occupa del riconoscimento delle immagini. In particolare, verrà presentato il servizio Amazon Rekognition e, infine, verrà illustrata una prova effettuata con quest'ultimo strumento.
- Nel Capitolo 5 si esplorerà il campo del Text Rekognition, analizzando il suo funzionamento. Successivamente si sperimenterà il suo utilizzo attraverso l'impiego di Amazon Textract.
- Nel Capitolo 6 saranno riportate alcune discussioni riguardanti le esperienze svolte nei capitoli precedenti
- Nel Capitolo 7, infine, verranno tratte le conclusioni e verranno delineati alcuni possibili sviluppi futuri.

---

## Introduzione all'Intelligenza Artificiale

---

*Nel primo capitolo verrà introdotto il concetto di Intelligenza Artificiale partendo dalle sue definizioni e fornendo degli accenni sulla sua storia. In particolare, verranno presentati obiettivi e ambiti di applicazione di questa disciplina. A chiusura del capitolo ci sarà una sezione dedicata all'approfondimento dei suoi benefici e dei suoi rischi.*

### 1.1 Cos'è l'Intelligenza Artificiale

L'intelligenza è ciò che differenzia un uomo da un oggetto inanimato; da migliaia di anni cerchiamo di comprendere come pensiamo e agiamo, ovvero come il nostro cervello possa percepire, comprendere, prevedere e manipolare un mondo molto più grande e complicato di se stesso. Il campo dell'Intelligenza Artificiale, o IA, si occupa non solo della comprensione, ma anche della costruzione, di entità intelligenti, ovvero macchine in grado di calcolare come agire in modo efficace e sicuro in un'ampia varietà di situazioni. Una macchina intelligente deve essere in grado, quindi, di analizzare l'ambiente circostante, elaborare delle strategie, assimilare conoscenze attraverso esperienze accumulate nel tempo ed elaborare risposte.

#### 1.1.1 Definizione di IA

Esistono diverse definizioni di Intelligenza Artificiale, dalle più semplici alle più complicate; questo perché l'intelligenza è, tuttora, un concetto complesso e difficile da spiegare. Secondo Marco Somalvico in Somalvico *et al.* [1987]:

*"L'IA è quella disciplina, appartenente all'informatica, che studia i fondamenti teorici, le metodologie e le tecniche che permettono di progettare sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico delle prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana"*

Un'altra definizione afferma:

*"L'Intelligenza Artificiale (in breve IA) è una disciplina che studia se e in che modo si possano realizzare sistemi informatici intelligenti in grado di simulare la capacità e il comportamento del pensiero umano."*

Da queste definizioni si deduce che l'Intelligenza Artificiale è un sistema creato dall'uomo per replicare al meglio i nostri comportamenti.

### 1.1.2 Obiettivi e applicazioni dell'IA

Gli obiettivi dell'Intelligenza Artificiale (IA) variano a seconda del contesto e degli ambiti di applicazione; tuttavia ci sono obiettivi comuni che spesso si cerca di raggiungere nello sviluppo dell'IA. Lo scopo principale di questa disciplina non mira a replicare o simulare l'intelligenza umana, una meta che alcuni studiosi ritengono addirittura inammissibile per ragioni epistemologiche, poiché sottolineano una distinzione tra la realtà e la conoscenza della realtà. Al contrario, l'Intelligenza Artificiale ha l'obiettivo di riprodurre o emulare l'intelligenza umana, poiché non vi è alcun motivo a priori che impedisca a una macchina di fornire alcune (ma non tutte) prestazioni dell'intelligenza umana senza ragionare allo stesso modo di come ragiona l'uomo. Ad esempio, la capacità di risolvere problemi mediante procedimenti logici potrebbe essere emulata utilizzando meccanismi specifici della macchina, eventualmente differenti da quelli ipotizzati per l'uomo, ma in grado di fornire prestazioni qualitativamente equivalenti e quantitativamente superiori a quelle umane.

Le applicazioni dell'Intelligenza Artificiale sono molteplici e riguardano diversi ambiti, non solo industriali, ma anche domestici. Un esempio possono essere i sistemi in grado di regolare autonomamente la temperatura o l'illuminazione all'interno di un'abitazione oppure l'utilizzo della voce come input per far svolgere diversi compiti ad alcuni dispositivi. Al fine di fornire un quadro completo delle soluzioni di IA adottate dalle aziende possiamo individuare sei classi di soluzioni:

1. *Intelligent Data Processing*. Si tratta di algoritmi che analizzano dati specifici per estrapolare informazioni e compiere azioni in conseguenza. In questa categoria rientrano diversi utilizzi, come l'analisi predittiva e il rilevamento di frodi
2. *Virtual Assistant/Chatbot*. Alla categoria dei cosiddetti Chatbot appartengono software in grado di eseguire azioni o fornire servizi per un individuo in base a comandi ricevuti in maniera vocale o testuale. Questi sistemi si contraddistinguono per la loro capacità di comprensione del tono del dialogo e di memorizzazione delle informazioni raccolte. Un esempio di applicazione di Intelligenza Artificiale tramite Chatbot è ChatGPT.
3. *Recommendation System*. Queste applicazioni indirizzano le scelte degli utenti in base ad informazioni da essi fornite (in maniera diretta o indiretta). Tra le soluzioni più diffuse si individuano i sistemi che suggeriscono un acquisto in base a quelli precedenti, influenzando così il processo decisionale dell'utente.
4. *Natural Language Processing*. Si tratta di quel ramo dell'IA che riguarda l'informazione espressa nel linguaggio naturale. Esso comprende soluzioni che elaborano il linguaggio, con finalità che possono variare dalla comprensione del contenuto, alla traduzione, fino alla produzione di testo in modo autonomo a partire da dati o documenti forniti in input.
5. *Computer Vision*. La Computer Vision è un campo scientifico che studia algoritmi e tecniche per permettere ai computer di raggiungere una comprensione di alto livello del contenuto di immagini o video tramite l'utilizzo di tecniche di Machine Learning. Si tratta di soluzioni di analisi di immagini, singole o in sequenza (video), orientate al riconoscimento di persone, animali e cose presenti all'interno dell'immagine stessa, al riconoscimento biometrico (ad esempio, volto, iride), e in generale, all'estrazione di informazioni dall'immagine. Questi sistemi si stanno diffondendo principalmente nell'ambito della videosorveglianza, dove l'analisi delle immagini è fondamentale per individuare eventuali situazioni anomale o di pericolo.
6. *Soluzioni fisiche*. All'interno di questa classe di applicazioni rientrano:

- *Veicoli Autonomi*; con questo termine si intendono i mezzi di trasporto autoguidati; questa è una tipologia di veicoli adibita al trasporto di persone, animali o cose che può essere destinata alla navigazione non solo su strada, ma anche marittima, fluviale e aerea. (Figura 1.5)
- *Robot Autonomi*; con questo termine si intendono i robot in grado di muoversi senza l'intervento umano, in base ad informazioni raccolte dall'ambiente circostante. A questa categoria appartengono sia soluzioni industriali, come i robot progettati per l'automazione di processi produttivi e logistici, sia robot destinati al mercato civile, come gli assistenti alla vendita presenti all'interno di negozi per fornire informazioni ai clienti.

### 1.1.3 Tipi di Intelligenza Artificiale

Le tecnologie di Intelligenza Artificiale sono classificate in base alla loro capacità di imitare le caratteristiche umane, la tecnologia che usano per farlo e le loro applicazioni nel mondo reale. Utilizzando queste linee guida possiamo raggruppare i sistemi di IA in tre tipologie:

- *ANI (Artificial Narrow Intelligence) o Intelligenza Artificiale Ristretta*. Conosciuta anche come Narrow AI o Weak AI, è un tipo di IA specializzata in una determinata attività o in un compito limitato. Questo tipo di Intelligenza Artificiale è progettata per eseguire compiti specifici in modo efficiente, ma non ha la capacità di comprendere o svolgere attività al di fuori del suo ambito di competenza. Un esempio può essere il sistema Deep Blue, che ha battuto Kasparov negli scacchi, dove, tramite un algoritmo il sistema è stato in grado di prevedere i risultati migliori ed effettuare le giuste mosse (Figura 1.1). Nonostante potesse sembrare che il sistema stesse "pensando", stava solo eseguendo il suo compito senza riflettere o decidere.
- *AGI (Artificial General Intelligence) o Intelligenza Artificiale Generale*. Conosciuta anche come Strong AI, l'AGI è la tecnologia più simile o alla pari con l'intelligenza umana. L'obiettivo principale è quello di creare macchine con il nostro stesso livello di intelligenza, in grado di svolgere compiti come gli esseri umani. In questo modo le macchine avranno, quindi, la capacità di ragionare, pensare e svolgere tutte le attività che l'essere umano è capace di fare.
- *ASI (Artificial Super Intelligence) o Superintelligenza Artificiale*. L'ASI è lo stadio dell'IA in cui le capacità dei computer supereranno quelle degli esseri umani. È un concetto puramente teorico che si riferisce a una forma di IA altamente avanzata e in grado di superare in modo significativo l'intelligenza umana in tutte le sfere cognitive, comprese la risoluzione di problemi complessi, l'apprendimento automatico, la comprensione del linguaggio naturale, la creatività, e molto altro. Attualmente l'ASI è ancora una prospettiva lontana rappresentata solo nei film e nei libri di fantascienza.



Figura 1.1: Kasparov vs. Deep Blue

## 1.2 La storia dell'Intelligenza Artificiale

L'IA come è oggi è nata nel secondo dopoguerra, grazie soprattutto all'avvento dei calcolatori, fin da subito considerati come il perfetto strumento per riprodurre una mente umana.

### 1.2.1 Origini

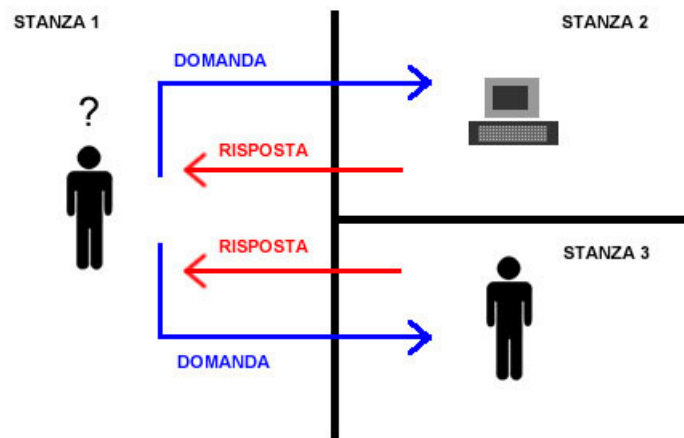
Tra i primi lavori in questo campo di ricerca, emerge quello di Warren McCulloch e Walter Pitts (1943), che presentarono un modello di rete neurale ispirato al funzionamento del cervello umano. Ogni neurone, collegato agli altri attraverso sinapsi, poteva assumere uno stato binario: "on" o "off". Questi due ricercatori dimostrarono che con questo modello era possibile implementare tutti i principali blocchi della logica booleana (come not, and, or, etc.). Inoltre, ipotizzarono la capacità di apprendimento di tali reti. Fù, poi, Donald Hebb nel 1949 a proporre una legge di modifica dei pesi delle sinapsi che consentiva tale possibilità. Un contributo significativo fu anche il primo computer a rete neurale, sviluppato da Marvin Minsky nel 1950, denominato SNARC, che simulava una rete di 40 neuroni. Un importante contributo fu dato anche da Alan Turing (1950) con il suo articolo *Computing machinery and intelligence* dove affrontò la questione della possibilità di costruire una macchina capace di pensare, introducendo quello che ora è noto come il "Test di Turing". Egli infatti affermò:

«Una macchina può essere considerata intelligente se passa il "Test di Turing"»

### 1.2.2 Test di Turing

Il Test di Turing fornisce una prima soluzione al quesito filosofico: "Può una macchina pensare?". Turing, infatti, propone di riformulare il problema in modo diverso tramite una sorta di "gioco dell'imitazione", che consiste in un gioco a tre partecipanti: un uomo A, una donna B e una terza persona C. Quest'ultimo è tenuto separato dagli altri due e, tramite una serie di domande, deve stabilire qual è l'uomo e quale la donna. Dal canto loro, anche A e B hanno dei compiti: A deve ingannare C e portarlo a fare un'identificazione errata, mentre B deve aiutarlo. Affinché C non possa disporre di alcun indizio (come l'analisi della grafia o della voce), le risposte alle domande di C devono essere dattiloscritte o similmente trasmesse. Il test di Turing si basa sul presupposto che una macchina si sostituisca ad A (Figura 1.2). In tal caso, se C non si accorgesse di nulla, la macchina dovrebbe essere considerata intelligente, dal momento che sarebbe indistinguibile da un essere umano. La

macchina cioè dovrebbe essere considerata come dotata di una “intelligenza” pari a quella dell'uomo.



**Figura 1.2:** Illustrazione del funzionamento del Test di Turing

Nel suo stesso articolo Turing anticipa e riassume le opinioni, opposte alle sue, sollevate dai critici sulla validità della domanda: “Possono le macchine pensare?”. Alcune delle obiezioni sono raccolte di seguito:

- *Obiezione Teologica*

*“Il pensare è una funzione dell’anima immortale dell’uomo. Dio ha dato un’anima immortale ad ogni uomo e donna, ma non agli altri animali o alle macchine. Perciò nessun animale o macchina può pensare.”* Turing [1950]

- *Autocoscienza*

Molti filosofi e scienziati hanno affermato che il test di Turing non può essere uno strumento valido per verificare se una macchina stia realmente pensando. Infatti, anche se presentassero un comportamento intelligente non sarebbero comunque in grado di comprendere e di essere consapevoli di sé.

- *Dipendenza dal contesto e dalla situazione*

Il risultato del test può essere influenzato dalla situazione specifica e dalle domande poste. Alcune macchine potrebbero sembrare intelligenti in un contesto limitato ma fallire in altri contesti più complessi.

- *L’obiezione di Lady Lovelace*

Ada Lovelace (1815 - 1852) fu una matematica inglese nota per i suoi studi sulla *macchina analitica* di Charles Babbage. Essa, ridimensionando l’idea che la macchina fosse “pensante” alla maniera dell’uomo, affermò che:

*“La macchina analitica non ha la pretesa di “creare” alcunché. Può fare qualsiasi cosa sappiamo come ordinarle di fare”*

### 1.2.3 1956-1969: I successi e l'entusiasmo dei primi anni

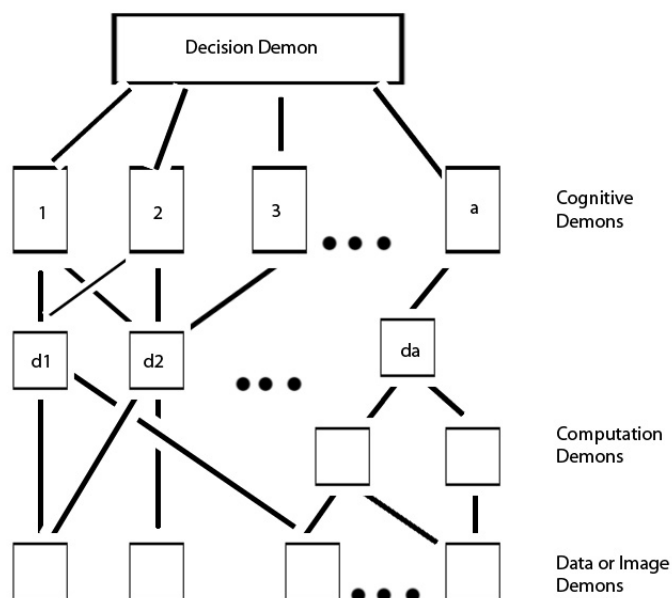
Dopo la prima guerra mondiale, si osserva un crescente interesse per l'Intelligenza Artificiale (IA). Tuttavia, il vero riconoscimento di questa disciplina come una vera e propria scienza avviene nel 1956 con la conferenza del Dartmouth College, organizzata da John McCarthy. La conferenza riunì noti ricercatori, tra cui Herbert Simon, Allen Newell, Marvin Minsky, Claude Shannon, e altri.

A Dartmouth, Simon e Newell presentano il Logic Theorist nel 1955, considerato il primo programma di IA. Questo programma dimostra teoremi matematici derivati dai "Principia Mathematica" di Russell e Whitehead. Nel 1957, essi sviluppano il General Problem Solver (GPS), un sistema ispirato al ragionamento umano.

In parallelo, Herbert Gelernter crea il Geometry Theorem Prover presso l'IBM, mentre Oliver Selfridge implementa il Pandemonium (Figura 1.3), un primo programma per il "pattern recognition", basato su una serie di "demoni" (piccole funzioni richiamate dal programma) addestrati a riconoscere determinati elementi di una lettera scritta. Nel frattempo, McCarthy contribuisce al MIT con lo sviluppo del linguaggio di programmazione LISP, innovando anche con l'invenzione del "time sharing".

Nel 1958, lo stesso McCarthy pubblica "Programs with common sense", introducendo Advice Taker, un esempio di intelligenza artificiale completa: Advice Taker era in grado di percepire la realtà circostante e di rappresentarla al proprio interno, in modo da interagire con essa e rispondere agli stimoli provenienti dall'esterno. Nasce l'approccio dei "micro-mondi" di Minsky, che rappresenta il mondo fisico tramite forme geometriche per facilitare la manipolazione da parte del computer.

Nel 1962, Frank Rosenblatt propone il "perceptrone", una tipologia di rete neurale con un miglior algoritmo di modifica dei pesi sinaptici, riconosciuto come uno strumento eccellente per il riconoscimento di pattern. Questi sviluppi aprirono nuove prospettive per l'IA, contribuendo a definire il suo percorso evolutivo.



**Figura 1.3:** Il modello originale di Pandemonium

### 1.2.4 Stato Attuale dell'IA

Negli ultimi anni l'IA è stata caratterizzata da una grande quantità di cambiamenti sia a livello metodologico sia a livello contenutistico. Attualmente la tendenza comune è quella di partire da teorie già esistenti piuttosto che inventarne di nuove, basando le affermazioni o su teoremi rigorosamente dimostrati o su evidenze sperimentali.

Oggi, l'Intelligenza Artificiale si è diffusa in diverse attività quotidiane e settori come la logistica, la produzione industriale e l'ottimizzazione di processi aziendali, contribuendo allo sviluppo di tecnologie come la robotica, la guida autonoma, il Machine Learning e, recentemente, la tecnologia innovativa delle Generative A.I.

Il Machine Learning, una sotto-categoria dell'IA, consente di addestrare funzioni matematiche e statistiche utilizzando dati storici forniti in input. Gli algoritmi, che sono funzioni apprese dai dati storici, identificano pattern ricorrenti, non visibili all'occhio umano, e si adattano alle situazioni passate per generare l'output corretto quando si ripresentano situazioni simili.

Il Deep Learning, ispirato dalla struttura del cervello umano, utilizza reti neurali di varie dimensioni e tipologie, richiedendo enormi quantità di dati per l'addestramento. Lo scopo delle reti neurali è quello di elaborare e processare le informazioni ricevute in ingresso proponendo diversi tipi di output. Questa tecnologia è comunemente utilizzata per image recognition, speech recognition e traduzione di testi.

L'Intelligenza Artificiale Generativa, o Generative A.I., opera accanto al Machine Learning e al Deep Learning, creando nuovi contenuti, come audio, codice, immagini, testo e video. Fino a poco tempo fa, l'apprendimento automatico si limitava a studiare come identificare un oggetto all'interno di un'immagine e di classificarlo correttamente; ora è in grado di creare un'immagine partendo da una richiesta dell'utente scritta in linguaggio naturale. Un esempio è il software di quarta generazione della serie GPT creato da OpenAI, GPT-4 (Figura 1.4).



Figura 1.4: GPT-4



## 1.3 Benefici e Rischi dell'IA

L'avvento dell'Intelligenza Artificiale ha segnato una svolta significativa nei nostri modi di vita, influenzando profondamente settori che vanno dalla medicina all'industria, dall'intrattenimento all'istruzione. Mentre l'IA offre promettenti opportunità di innovazione e miglioramento delle nostre esperienze quotidiane, non possiamo ignorare il fatto che essa solleva anche questioni etiche, sociali ed economiche di rilevanza critica. Di seguito, vengono delineati una serie di benefici, ma anche una serie di problematiche connesse all'utilizzo incondizionato dell'Intelligenza Artificiale.

### 1.3.1 Vantaggi dell'IA

L'Intelligenza Artificiale ha avuto un ruolo importante in molti settori della nostra società, e tra questi possiamo trovare quello economico, scientifico, medico e finanziario. L'utilizzo più comune è quello per la gestione di Big Data; infatti, secondo DeepMindProject di Google, l'utilizzo dell'IA per la gestione di dati nel campo dell'assistenza sanitaria ha migliorato l'uguaglianza nell'accesso alle cure, l'aumento della velocità di queste ultime, ha trovato nuovi metodi di diagnosi e ha facilitato l'apprendimento continuo.

L'IA ha trovato spazio anche nel campo della Cyber Security dove, grazie al suo utilizzo, ha permesso l'aumento della sicurezza nell'IoT (*Internet of Things*), anticipando e contrastando gli intrusi molto più rapidamente di quanto possano fare gli esseri umani.

Dato che le macchine "intelligenti" utilizzano algoritmi di apprendimento per simulare le attività umane e risolvere problemi, possiamo identificare diversi vantaggi che consentono ad esse di eccellere rispetto agli esseri umani:

#### 1. *Aumento dell'efficienza lavorativa*

Le macchine potenziate dall'Intelligenza Artificiale dimostrano una straordinaria efficienza nell'esecuzione di specifiche attività ripetitive. Questo successo è attribuibile principalmente alla loro capacità di eliminare gli errori umani, garantendo risultati accurati ad ogni esecuzione di compiti specifici. Di conseguenza, si evita la necessità di impiegare due turni di personale che lavorano giorno e notte per gestire operazioni importanti. Ad esempio, gli assistenti virtuali basati sull'IA possono rispondere alle domande dei clienti e fornire supporto ai visitatori in ogni minuto del giorno, contribuendo significativamente all'aumento delle vendite aziendali.

#### 2. *Lavoro con elevata precisione*

Gli scienziati stanno lavorando per insegnare alle macchine a risolvere operazioni complesse e svolgere compiti critici autonomamente, al fine di ottenere risultati con una precisione superiore rispetto ai loro omologhi umani. La loro elevata accuratezza ha reso queste macchine indispensabili, soprattutto nel campo medico, data la criticità degli incarichi. Ad esempio, i robot stanno migliorando l'esecuzione di delicate operazioni chirurgiche minimizzando il rischio per le vite umane.

#### 3. *Riduzione dei costi di apprendimento e operativi*

L'IA utilizza algoritmi di apprendimento automatico come il Deep Learning e le reti neurali, per apprendere nuove cose, proprio come fanno gli esseri umani. In questo modo, si elimina la necessità di scrivere nuovo codice ogni volta che vogliamo che le macchine imparino qualcosa di nuovo. Inoltre, esse riducono i costi operativi grazie alla loro elevata efficienza e precisione nello svolgere il loro lavoro. Ad esempio, le macchine non fanno pause e possono eseguire la stessa attività ripetitiva senza alcun periodo di inattività o variazione nei risultati.

#### 4. Ottimizzazione dei processi

L'implementazione di macchine che usano l'IA nel contesto lavorativo offre l'opportunità di raccogliere ingenti quantità di dati correlati ai loro compiti. Questi dati possono essere elaborati per ottenere approfondimenti dettagliati sui processi mediante analisi quantitativa, consentendo, così, ulteriori ottimizzazioni. Ad esempio, le crescenti capacità di apprendimento automatico dei sistemi di IA consentono ad essi di condurre analisi in modo autonomo, contribuendo a migliorare costantemente i processi operativi.

### 1.3.2 Rischi e preoccupazioni

Nonostante i molteplici benefici portati dall'avvento dell'Intelligenza Artificiale, è fondamentale considerare che le macchine non possiedono un pensiero critico ed etico intrinseco; il loro comportamento dipende completamente dal modo in cui vengono programmate.

L'utilizzo indiscriminato dell'Intelligenza Artificiale può dar luogo a vari rischi, di seguito elencati:

- *Sorveglianza e persuasione.* L'IA può essere utilizzata in modo scalabile per eseguire una sorveglianza di massa su individui e rilevare attività di interesse. Modificando i flussi informativi verso gli individui attraverso i social media, basandosi su tecniche di apprendimento automatico, il comportamento politico può essere modificato e controllato in qualche misura.
- *Armi autonome.* L'IA può facilitare la produzione di armi in grado di individuare, selezionare ed eliminare bersagli umani senza intervento umano. Una preoccupazione principale riguarda il loro utilizzo: l'assenza della necessità di supervisione umana significa che un piccolo gruppo può dispiegare un numero arbitrariamente grande di armi contro bersagli umani definiti da qualsiasi criterio di riconoscimento possibile.
- *Decisioni con pregiudizi.* L'uso negligente o deliberato di algoritmi di apprendimento automatico può portare a decisioni influenzate da razza, genere o altre categorie protette. Spesso, i dati stessi riflettono pregiudizi diffusi nella società.
- *Impatto sull'occupazione.* L'impatto sull'occupazione causato dall'introduzione dell'Intelligenza Artificiale solleva molte preoccupazioni. Storicamente, le macchine hanno sostituito alcuni lavori umani, ma hanno anche aumentato la produttività, rendendo le persone più impiegabili e le aziende più redditizie. Tuttavia, ciò ha spesso portato ad un aumento delle disuguaglianze economiche, con una redistribuzione della ricchezza dal lavoro al capitale. Sebbene i progressi tecnologici abbiano provocato cambiamenti nell'occupazione, alla fine le persone hanno trovato nuove opportunità di lavoro.
- *Applicazioni critiche per la sicurezza.* Man mano che le tecniche dell'IA avanzano, vengono sempre più utilizzate in applicazioni critiche per la sicurezza, come la guida di automobili (Figura 1.5). Incidenti fatali sono già accaduti e mettono in evidenza la difficoltà di analisi del rischio statistico per i sistemi sviluppati utilizzando tecniche di apprendimento automatico. Il campo dell'IA dovrà sviluppare standard tecnici ed etici almeno comparabili a quelli già presenti in altre discipline ingegneristiche e sanitarie dove sono in gioco le vite delle persone.
- *Cybersecurity.* Le tecniche di Intelligenza Artificiale sono utili nella difesa contro gli attacchi informatici, ad esempio rilevando modelli di comportamento insoliti. Tuttavia, possono contribuire anche alla sopravvivenza e alla capacità di proliferazione dei malware. Ad esempio, i metodi di apprendimento per rinforzo sono stati utilizzati per

creare strumenti altamente efficaci per attacchi di ricatto e phishing automatizzati e personalizzati.



**Figura 1.5:** Esempio di veicolo autonomo

*In questo capitolo verranno introdotti il concetto di Cloud Computing, i suoi principali utilizzi, i suoi vantaggi e i suoi svantaggi. Successivamente verrà presentato Amazon Web Services (AWS), una piattaforma che offre tali servizi, esplorando la sua storia e i suoi scenari applicativi.*

## 2.1 Introduzione al Cloud Computing

Il fenomeno del Cloud Computing si manifesta in molteplici utilizzi. Ogni rivista tecnologica che si rispetti, oltre a numerosi siti web e blog nel settore informatico, dedica ampio spazio alle discussioni sul Cloud Computing. Tuttavia, la sfida principale consiste nel fatto che non c'è un consenso univoco sulla sua definizione. Se si interroga una decina di professionisti su cosa sia il Cloud Computing, si otterranno altrettante risposte differenti.

### 2.1.1 Cos'è

Il Cloud Computing è una tecnologia che sta rivoluzionando il modo in cui le imprese gestiscono i propri servizi informatici. In passato, le aziende utilizzavano hardware e software proprietari installati su server interni, il che comportava costi elevati per l'acquisto e la manutenzione di tali risorse. Inoltre, spesso era necessario dedicare personale per la gestione delle infrastrutture. Grazie a questa tecnologia, le imprese possono usufruire di servizi informatici affidabili, sicuri e scalabili che vengono erogati da provider esterni, senza la necessità di acquistare e mantenere costose infrastrutture fisiche.

In particolare, il Cloud Computing consiste nell'erogazione di servizi di computing, quali software, database, server e reti, tramite connessione Internet. Ciò significa che gli utenti finali sono in grado di accedere a software e applicazioni ovunque si trovino. I programmi per i computer sono ospitati da un server esterno e si trovano nel "Cloud". Ciò implica che gli utenti non devono preoccuparsi di aspetti tecnici come l'archiviazione e la potenza, ma possono semplicemente godersi il risultato finale.

Il Cloud deve il suo nome alla consuetudine di disegnare Internet come una nuvola (in inglese, appunto, "Cloud") nei diagrammi di brevetto della rete nei primi anni '90 (Figura 2.1). Così, quando si parla di "archiviazione nel Cloud", si intendono le risorse e i servizi di sistema che possono fare queste operazioni su Internet.



**Figura 2.1:** Illustrazione del Cloud

Il Cloud Computing si basa sul presupposto che le principali attività di elaborazione avvengono su una macchina, in remoto, che non è quella attualmente in uso, che riceve direttamente e trasmette gli input. I dati raccolti durante questo processo vengono memorizzati ed elaborati da server remoti (chiamati anche server Cloud), il che significa che il dispositivo che accede al Cloud non deve lavorare altrettanto duramente. Dato che il Cloud Computing implica che il carico di lavoro del computer di un utente è ospitato su una macchina diversa, il Cloud è accessibile ovunque ed è disponibile a chiunque disponga di una connessione a Internet. Esistono diverse tipologie di Cloud, sia pubblici che privati, i quali possono essere auto-ospitati o gestiti da terze parti. Nei Cloud privati, gli utenti devono disporre delle piattaforme o degli accessi adeguati, come un browser o un account specifico, per poter accedere ai server e ai dati al loro interno.

### 2.1.2 Servizi e Utilizzi

Esistono tre tipi principali di servizi basati sul Cloud:

- *Infrastructure as a Service (IaaS)*; fornisce ad un'azienda la sua infrastruttura basilare di computing su Internet, tra cui reti, server e spazio di archiviazione, in genere con pagamento a consumo. Dal punto di vista dell'utente, l'utilizzo di servizi basati sul Cloud implica semplicemente l'accesso sicuro all'applicazione tramite Internet.
- *Software as a Service (SaaS)*; è un modo per fornire applicazioni software su Internet, in genere su abbonamento e su richiesta. Il fornitore di servizi sul Cloud ospita e gestisce l'intera infrastruttura IT, oltre a occuparsi di manutenzione e sicurezza.
- *Platform as a Service (PaaS)*; amplia il modello *IaaS* con un insieme di servizi basati sul Cloud che consente agli sviluppatori e agli utenti delle aziende di sviluppare e distribuire applicazioni facili da utilizzare in modo rapido e intuitivo.

Il Cloud Computing rientra, per buona parte, nell'erogazione dei *Software as a Service (SaaS)*, un modello di distribuzione delle applicazioni che si paga in base all'utilizzo che effettivamente se ne fa. Tra gli utilizzi più comuni possiamo trovare:

1. *Servizi di file hosting*. Includono i servizi di Cloud storage, dedicati alla memorizzazione e al backup dei dati. Ciò si traduce nella facilità di caricare e scaricare file, nell'utilizzo di file system gestibili a distanza e nella funzionalità di sincronizzazione in tempo reale tra file locali e quelli presenti nel Cloud

2. *Utilizzo di piattaforme di streaming video.* I servizi di streaming on demand per la visione di partite, eventi in diretta e film rappresentano un esempio eccellente delle avanzate evoluzioni nel campo del Cloud Computing. In questo contesto, l'infrastruttura impiega una rete di server collaborativi per garantire uno streaming senza interruzioni. Vengono, tra le altre soluzioni, implementate tecnologie di correzione degli errori di trasmissione, buffering per evitare ritardi nel flusso video e per mantenere la sincronizzazione in tempo reale.
3. *Backup di sistema, di siti e di software di ogni genere.* La maggior parte dei sistemi di backup, finalizzati a creare copie di sicurezza dei dati per ripristino e recupero, si basano principalmente sul Cloud Computing. Questo approccio sfrutta efficacemente i diversi strati di software offerti dalla piattaforma per eseguire copie di sicurezza programmate, sicure e ridondanti. Di conseguenza, il Cloud elimina quasi completamente la necessità di eseguire manualmente (e periodicamente, come sarebbe consigliato) queste operazioni, facendo risparmiare tempo agli utenti finali.
4. *App per memorizzare dati personali in modo sicuro.* Molti portachiavi per memorizzare le password del proprio sistema, o quelle dei siti a cui accediamo più spesso, sincronizzano tali dati grazie ad un server Cloud. Ciò è associato ad una crittografia forte (cosiddetta end-to-end); in questo modo nessuno, se non il proprietario, potrà vedere in chiaro le password salvate.

### 2.1.3 Vantaggi e Sfide del Cloud Computing

Il passaggio al Cloud offre alle organizzazioni di ogni forma e dimensione la possibilità di muoversi con maggiore velocità, agilità e innovazione. Il passaggio al Cloud Computing ha trasformato completamente il modo in cui lavoriamo, comunichiamo e collaboriamo, diventando rapidamente una necessità per rimanere competitivi nel mondo digitale di oggi. Di seguito sono riportati i maggiori benefici portati da questa nuova tecnologia.

1. *Scalabilità e flessibilità.* Il Cloud offre maggiore flessibilità alle aziende, consentendo di scalare rapidamente risorse e di effettuare archiviazioni in modo da soddisfare le proprie esigenze senza dover investire in infrastrutture fisiche.
2. *Time to Market più Rapido.* Il Cloud Computing permette di avviare nuove istanze o eseguirne il ritiro in pochi secondi, consentendo agli sviluppatori di accelerare lo sviluppo con deployment rapidi. Supporta l'innovazione semplificando l'esecuzione di test di nuove idee e la progettazione di nuove applicazioni senza limitazioni hardware o processi di approvvigionamento lenti.
3. *Risparmio sui costi.* Qualunque modello di servizio Cloud si scelga, il costo è relativo soltanto alle risorse effettivamente utilizzate.
4. *Collaborazione più efficace.* Il Cloud storage consente l'accesso ai dati da qualsiasi parte del mondo e da qualsiasi dispositivo con connessione a Internet, agevolando la collaborazione in modo flessibile e remoto.
5. *Sicurezza avanzata.* Contrariamente alla percezione comune, il Cloud computing può rafforzare la sicurezza grazie alle avanzate funzionalità, alla manutenzione automatica e alla gestione centralizzata fornite dai fornitori di servizi Cloud.
6. *Prevenzione della Perdita di Dati.* I Cloud provider offrono funzionalità di backup e ripristino di emergenza, contribuendo a prevenire la perdita di dati in caso di malfunzionamenti hardware, minacce dannose o errori umani. L'archiviazione nel Cloud fornisce un'ulteriore sicurezza contro emergenze impreviste.

Naturalmente, come qualsiasi tecnologia, esistono vantaggi e svantaggi del suo utilizzo. In questo caso, ad esempio, una delle limitazioni più comuni del Cloud Computing è che si basa su una connessione a Internet; perciò, una mancanza di connessione ad Internet potrebbe impedire di accedere alle informazioni o alle applicazioni di cui si ha bisogno. Anche i fornitori di servizi Cloud possono riscontrare tempi di inattività non garantendo, quindi, la continuità dei servizi forniti, impedendone l'accesso.

Un'altra limitazione, invece, riguarda la sicurezza informatica e la privacy. La tecnologia Cloud presuppone lo scambio e la gestione di un'enorme quantità di dati in rete; per questo motivo, la possibilità che questi dati subiscano furti e compromissioni è, potenzialmente, sempre dietro l'angolo.

Infine, uno svantaggio prettamente economico potrebbe manifestarsi sotto forma di costi e spese impreviste. Quando si utilizza il Cloud Computing, è importante considerare attentamente i modelli di pricing adottati dai fornitori di servizi. Alcuni fattori, come i picchi improvvisi nell'utilizzo delle risorse o la gestione inefficiente della capacità, potrebbero tradursi in costi aggiuntivi non preventivati.

## 2.2 Introduzione ad Amazon AWS

Amazon Web Services (AWS) è il Cloud più completo e utilizzato del mondo; esso offre più di 200 servizi completi da data center a livello globale. Milioni di clienti, incluse le start-up a maggiore crescita, le più grandi compagnie e le agenzie governative leader di settore, utilizzano AWS per diminuire i costi, diventare più agili e innovarsi in modo più rapido.

### 2.2.1 Che cos'è

Amazon AWS (Amazon Web Services) è una piattaforma di servizi Cloud offerta da Amazon.com. Si distingue nel panorama dei provider Cloud per il suo ampio e avanzato set di servizi e funzionalità. Offre una vasta gamma di servizi 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, i data lake, l'analisi dei dati e l'Internet of Things (IoT). Questo approccio consente alle organizzazioni di migrare in modo agevole le loro applicazioni esistenti verso il Cloud, avvalendosi delle funzionalità più avanzate per costruire soluzioni personalizzate.

Un punto di forza di AWS risiede nella vasta scelta di database specializzati per diversi contesti applicativi. Ciò consente agli utenti di selezionare il database più adatto alle proprie esigenze specifiche, ottimizzando, così, sia i costi che le prestazioni. La presenza di funzionalità avanzate all'interno dei servizi offre un elevato livello di flessibilità e personalizzazione, consentendo alle organizzazioni di sviluppare soluzioni su misura per le proprie necessità.

### 2.2.2 Storia

Amazon Web Services (AWS) è la sezione di servizi Cloud di Amazon.com, una delle più grandi e influenti società di e-commerce al 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.

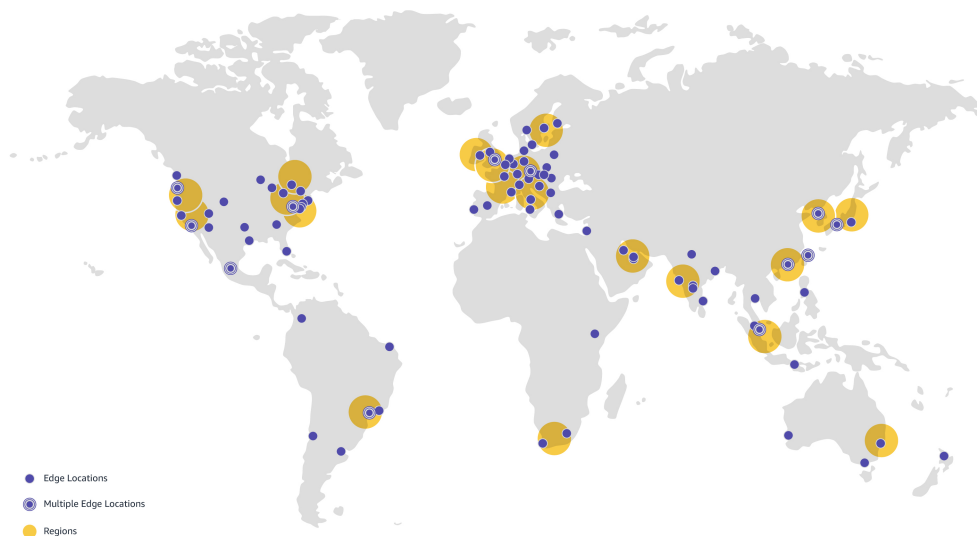
Nel 2004, Jeff Bezos, fondatore e CEO di Amazon, ha approvato il progetto e ha dato il via allo sviluppo di ciò che sarebbe diventato Amazon Web Services. Nel 2006, AWS è stato ufficialmente lanciato, offrendo servizi di Cloud Computing, tra cui calcolo e storage su richiesta.

Il lancio iniziale includeva servizi fondamentali, come Amazon Simple Queue Service (SQS) e Amazon Simple Storage Service (S3). Nel corso degli anni, AWS ha continuato a espandersi, introducendo nuovi servizi e funzionalità per soddisfare le crescenti esigenze delle imprese e degli sviluppatori.

Uno dei momenti cruciali nella storia di AWS è stato il lancio di Amazon Elastic Compute Cloud (EC2) nel 2006, che ha permesso agli utenti di eseguire macchine virtuali su richiesta. Ciò ha aperto la strada a un approccio più flessibile e scalabile al calcolo su Cloud.

Con il passare del tempo, AWS ha continuato a innovare, introdurre servizi avanzati, come Amazon RDS (Relational Database Service), AWS Lambda, per il calcolo serverless, e una vasta gamma di servizi per l'Intelligenza Artificiale, l'analisi dati e altro ancora.

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. Il cloud AWS comprende 102 Zone di disponibilità all'interno di 32 aree geografiche in tutto il mondo, con altre 15 Zone di disponibilità e 5 ulteriori Regioni AWS già annunciate in Canada, Germania, Malesia, Nuova Zelanda e Thailandia (Figura 2.2).



**Figura 2.2:** Mappa dell'infarstruttura AWS

### 2.2.3 Come funziona

Amazon Web Service (AWS) è una piattaforma che sfrutta infrastrutture complesse e software avanzati. Questi includono:

- *Infrastruttura Globale.* AWS gestisce una vasta infrastruttura di data center in tutto il mondo. Questi data center sono distribuiti in diverse regioni geografiche e dispongono di numerose zone di disponibilità all'interno di ciascuna regione (Figura 2.2).
- *API e SDK.* Gli sviluppatori possono interagire con i servizi AWS in modo programmato utilizzando API e SDK, permettendo l'automazione delle operazioni e l'integrazione con applicazioni personalizzate.
- *Archiviazione e Calcolo.* AWS offre servizi di calcolo scalabili tramite Amazon EC2 (Elastic Compute Cloud), che consente agli utenti di eseguire macchine virtuali su richiesta e Amazon S3 (Simple Storage Service) che fornisce servizi di archiviazione di oggetti



scalabili, mentre Amazon EBS (Elastic Block Store) offre archiviazione persistente per istanze EC2.

- *Console di Gestione.* Gli utenti accedono ai servizi AWS attraverso una console di gestione web, avendo a disposizione un'interfaccia grafica intuitiva per configurare, monitorare e gestire risorse.
- *Pagamento Pay-as-You-Go.* AWS opera su un modello di pagamento basato sull'utilizzo effettivo. Gli utenti pagano solo per le risorse che utilizzano, senza costi fissi o obblighi a lungo termine.
- *Sicurezza e Conformità.* AWS offre un ampio insieme di strumenti di sicurezza e servizi di conformità per proteggere le risorse degli utenti e soddisfare i requisiti normativi.

## 2.3 Scenari Applicativi di Amazon AWS

In questa sezione ci focalizziamo sui servizi forniti da AWS e sulle sue principali applicazioni in diversi settori.

### 2.3.1 Servizi a disposizione

Amazon Web Service offre una vasta gamma di servizi differenti, ognuno progettato per scopi specifici (Figura 2.3). Di seguito, vengono esposti alcuni dei servizi più comuni, accompagnati dalla loro spiegazione di funzionamento:

- *Amazon EC2 (Elastic Compute Cloud):* è un servizio che offre capacità di calcolo scalabile. Gli utenti possono creare istanze virtuali (server) con varie configurazioni per eseguire applicazioni o carichi di lavoro specifici.
- *AWS Lambda:* è un servizio che consente di eseguire codice senza dover gestire il server. Gli sviluppatori caricano il proprio codice e AWS Lambda lo esegue automaticamente in risposta a determinati eventi.
- *Amazon RDS (Relational Database Service):* è un servizio che semplifica la configurazione, la scalabilità e le operazioni di database relazionali. Gli utenti possono facilmente configurare e gestire database relazionali, come MySQL, PostgreSQL, SQL Server, etc., senza doversi preoccupare dell'infrastruttura sottostante.
- *Amazon EBS (Elastic Block Store):* fornisce uno storage persistente di tipo blocco per l'uso con istanze EC2. È possibile collegare volumi EBS alle proprie istanze EC2 per fornire storage persistente ad alta disponibilità e bassa latenza.
- *Amazon S3 (Simple Storage Service):* è un servizio che offre uno storage di oggetti scalabile e sicuro. Gli utenti possono archiviare e recuperare una qualsiasi quantità di dati da qualsiasi posizione su Internet, in questo modo essi hanno a disposizione un modo semplice per archiviare e recuperare file.
- *Amazon VPC (Virtual Private Cloud):* consente di creare una rete virtuale isolata nel Cloud AWS. È possibile configurare la propria VPC con subnet, route e gateway per avere controllo sulla propria rete virtuale.
- *AWS Direct Connect:* fornisce connessioni di rete ad AWS dedicate dalla propria infrastruttura on-premise. Il servizio mette a disposizione la possibilità di stabilire connessioni di rete dedicate per migliorare le prestazioni, la sicurezza e l'affidabilità tra l'ambiente on-premise e il Cloud AWS.

- *Elastic Load Balancing*: distribuisce automaticamente il traffico in ingresso tra diverse istanze. Gli utenti possono distribuire il traffico tra più istanze EC2 per migliorare la disponibilità, la resistenza e la scalabilità delle applicazioni.



**Figura 2.3:** Alcuni dei servizi di AWS

### 2.3.2 Utilizzi e Applicazioni

Amazon AWS è utilizzato da una vasta gamma di clienti, tra cui imprese di piccole, medie e grandi dimensioni, organizzazioni governative, start-up e sviluppatori individuali. Alcuni dei clienti noti di AWS includono Netflix, Airbnb, NASA, Samsung, Pfizer, Siemens, e molte altre aziende di diversi settori.

Le applicazioni di Amazon Web Services sono estremamente diverse e impattano vari settori della società. Di seguito sono riportate alcune delle maggiori applicazioni di AWS:

- *Sanità e Assistenza Sanitaria*: AWS è utilizzato per l'archiviazione sicura e la gestione dei dati sanitari e delle immagini mediche. Inoltre, supporta lo sviluppo di applicazioni per la telemedicina e la ricerca medica basata su dati.
- *Educazione*: le istituzioni educative utilizzano AWS per fornire servizi di e-learning, archiviare materiale didattico online e gestire applicazioni di natura accademica.
- *Pubblica Amministrazione*: AWS offre soluzioni per migliorare l'efficienza delle agenzie governative, tra cui la gestione di dati sensibili, la fornitura di servizi online e la modernizzazione dei sistemi legacy.
- *Settore Bancario e Finanziario*: nel settore finanziario, AWS è utilizzato per analisi dei dati, la sicurezza informatica, la gestione dei rischi e lo sviluppo di applicazioni per i servizi finanziari.
- *Industria Manifatturiera*: le aziende manifatturiere sfruttano AWS per l'ottimizzazione dei processi di produzione, la gestione delle catene di approvvigionamento e l'implementazione di soluzioni IoT (*Internet of Things*) per il monitoraggio delle apparecchiature.

- *Startup*: AWS fornisce un'infrastruttura flessibile e scalabile per le startup, consentendo loro di crescere rapidamente senza investimenti iniziali massicci in infrastrutture.
- *Media e Intrattenimento*: i media utilizzano AWS per l'archiviazione e la distribuzione di contenuti, il rendering di effetti speciali, la trasmissione video in streaming e la gestione di database di contenuti.
- *Ricerca Scientifica*: AWS è ampiamente utilizzato nella ricerca scientifica per l'elaborazione di grandi set di dati, la simulazione di esperimenti e l'archiviazione di dati di ricerca.
- *Agricoltura e Sostenibilità Ambientale*: nel settore agricolo, AWS viene utilizzato per la raccolta e l'analisi dei dati agricoli, la gestione delle risorse naturali e la creazione di soluzioni per la sostenibilità ambientale.
- *Trasporti e Logistica*: AWS supporta applicazioni per il monitoraggio e la gestione dei mezzi di trasporto, l'ottimizzazione delle rotte e la tracciabilità delle merci.

---

## Esperienze nell'ambito dei chatbot

---

*Nel terzo capitolo verranno trattati i chatbot e il loro utilizzo. Successivamente verrà presentato un caso di studio: Amazon Lex. In particolare, si esplorerà il suo funzionamento e si mostrerà un'esperienza effettuata con quest'ultimo.*

### 3.1 Introduzione ai chatbot

L'Intelligenza Artificiale si integra sempre di più nella nostra vita quotidiana con la creazione e l'analisi di software e hardware intelligenti, chiamati agenti intelligenti. Gli agenti intelligenti possono svolgere una varietà di compiti, dai lavori manuali a operazioni sofisticate. Un chatbot è un esempio tipico di un sistema di Intelligenza Artificiale e uno degli esempi più elementari e diffusi di interazione intelligente tra l'uomo e il computer.

#### 3.1.1 Cos'è un chatbot

Un chatbot è un software che simula ed elabora, come un'entità intelligente, le conversazioni umane (scritte o parlate), consentendo agli utenti di interagire con i dispositivi digitali come se stessero comunicando con una persona reale. Ciò è reso possibile grazie al *Natural Language Processing (NLP)*, un campo dell'Intelligenza Artificiale che si occupa di consentire alle macchine di comprendere, interpretare e generare linguaggio umano in modo naturale. I chatbot possono essere semplici, come programmi rudimentali che rispondono a una semplice richiesta con una singola riga, oppure sofisticati, come gli assistenti digitali che apprendono e si evolvono per fornire livelli crescenti di personalizzazione quando raccolgono ed elaborano le informazioni.

#### 3.1.2 Come funzionano

Guidati da tecnologie avanzate come l'Intelligenza Artificiale, l'elaborazione del linguaggio naturale (NLP) e l'apprendimento automatico (ML), i chatbot elaborano dati per rispondere a una varietà di richieste.

Esistono due categorie principali di chatbot:

- *I chatbot dedicati alle attività (dichiarativi)*; sono programmi specifici che si concentrano sull'esecuzione di una funzione specifica. Attraverso l'uso di regole, NLP e una quantità limitata di ML, generano risposte automatiche, seppur informali, alle richieste

degli utenti. Questi chatbot gestiscono interazioni altamente specifiche e strutturate, principalmente applicabili alle funzioni di assistenza e servizio. Ad esempio, possono gestire domande frequenti interattive o transazioni di base. Sebbene sfruttino l’NLP per un’interazione intuitiva, le loro capacità sono abbastanza basilari e sono attualmente i chatbot più comunemente utilizzati.

- *I chatbot predittivi basati sui dati (di conversazione)*; sono spesso chiamati assistenti virtuali o digitali e sono notevolmente più sofisticati, interattivi e personalizzati rispetto ai chatbot dedicati alle attività. Questi chatbot comprendono il contesto e utilizzano l’NLU, l’NLP e il ML per apprendere. Sfruttano l’intelligenza predittiva e l’analisi dei dati per personalizzare le risposte in base ai profili e ai comportamenti degli utenti. Gli assistenti digitali possono imparare nel tempo, offrendo raccomandazioni e anticipando le esigenze. Possono anche gestire conversazioni più complesse. Siri di Apple e Alexa di Amazon sono esempi di chatbot predittivi orientati al consumatore e basati sui dati.

L’utilizzo dei chatbot ha registrato un notevole aumento, soprattutto a livello aziendale, poiché ha contribuito ad incrementare l’efficienza operativa e ha consentito alle aziende di ottenere risparmi sui costi, offrendo contemporaneamente comodità e servizi supplementari sia per dipendenti che per clienti. I chatbot permettono alle aziende di gestire agevolmente una varietà di richieste e problemi dei clienti, riducendo, nel contempo, la necessità di interazione umana.

Con l’implementazione dei chatbot, un’azienda può simultaneamente scalare, personalizzare e adottare un approccio proattivo, rappresentando, così, un elemento fondamentale per la differenziazione nel contesto aziendale.

## 3.2 Caso di studio: Amazon Lex

Questa sezione si dedica all’approfondimento di Amazon Lex, una piattaforma di servizi di elaborazione del linguaggio naturale (NLP) sviluppata da Amazon Web Services. Amazon Lex rappresenta una significativa evoluzione nell’ambito delle interfacce conversazionali e dei chatbot, offrendo potenti strumenti per la creazione di esperienze di conversazione automatizzate e intuitive.

### 3.2.1 Che cos’è

Amazon Lex V2, un servizio di Amazon Web Services, è progettato per la creazione di interfacce conversazionali avanzate per applicazioni, utilizzando, sia il riconoscimento vocale che il testo. Esso offre funzionalità avanzate di comprensione del linguaggio naturale (NLU) e di riconoscimento automatico del parlato (ASR); Amazon Lex V2 consente di sviluppare esperienze utente altamente coinvolgenti, caratterizzate da interazioni conversazionali realistiche, nonché di creare nuove categorie di prodotti innovativi.

Un aspetto distintivo di Amazon Lex V2 è la sua accessibilità; esso, infatti, rende possibile per qualsiasi sviluppatore la rapida creazione di chatbot conversazionali. Senza richiedere competenze approfondite in Deep Learning, è possibile definire il flusso di conversazione di base direttamente dalla console di Amazon Lex V2. Il servizio gestisce il dialogo e adatta dinamicamente le risposte durante l’interazione. Attraverso la console, è possibile costruire, testare e pubblicare chatbot vocali o basati su testo, integrandoli, poi, su dispositivi mobili, applicazioni web e piattaforme di chat come Facebook Messenger.

Amazon Lex V2 offre un’ampia integrazione con servizi AWS, inclusi AWS Lambda, Amazon Connect, Amazon Comprehend e Amazon Kendra. L’integrazione con Lambda

consente ai chatbot di accedere a connettori enterprise serverless pre-costruiti, facilitando l'accesso ai dati nelle applicazioni SaaS come Salesforce.

Tra i principali vantaggi offerti da Amazon Lex V2 troviamo i seguenti:

- *Semplicità*: la console di Amazon Lex V2 permette all'utente una creazione rapida di un bot, consentendo di fornire solo alcuni esempi di frasi e generando un modello completo di linguaggio naturale per interazioni vocali o testuali.
- *Tecnologie di Deep Learning democratizzate*: Amazon Lex V2 mette a disposizione delle tecnologie ASR e NLU, semplificando la creazione di sistemi di comprensione del linguaggio. L'utilizzo di queste tecnologie è reso accessibile a tutti gli sviluppatori, permettendo di convertire il parlato in testo e di comprendere l'intento dell'utente.
- *Distribuzione e scalabilità senza soluzione di continuità*: la costruzione, il test e il rilascio dei bot avvengono direttamente dalla console di Amazon Lex V2, con la possibilità di pubblicarli su dispositivi mobili, app web e servizi di chat. Amazon Lex V2 si scala automaticamente, eliminando la necessità di gestire infrastrutture.
- *Integrazione nativa con la piattaforma AWS*: Amazon Lex V2 opera in modo integrato con altri servizi AWS, sfruttando le potenzialità della piattaforma per sicurezza, monitoraggio, autenticazione utente, logica aziendale, archiviazione e sviluppo di app mobili.
- *Economicità*: Amazon Lex V2 adotta un modello di pagamento basato sul consumo effettivo, senza costi iniziali o tariffe minime. Si paga solo per le richieste di testo o voce effettuate. Il prezzo flessibile e il periodo gratuito di Amazon Lex V2 consentono di esplorare il servizio senza investimenti iniziali significativi.

### 3.2.2 Come funziona

Amazon Lex V2 offre la possibilità di sviluppare applicazioni attraverso un'interfaccia basata su testo o voce per interagire con gli utenti. Di seguito vengono riportati i passaggi tipici per operare con Amazon Lex V2:

- *Creare un Bot*: si ha la capacità di definire un bot e aggiungere uno o più linguaggi; 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.
- *Testare il Bot*: Amazon Lex V2 offre la possibilità di utilizzare la finestra di test fornita dalla console per verificare il comportamento del bot.
- *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 V2 permette la distribuzione del bot sulle proprie applicazioni o piattaforme di messaggistica come, Facebook Messenger o Slack.

Il funzionamento di un chatbot si basa su diverse risorse per comprendere e gestire le conversazioni con gli utenti. Segue un elenco di concetti fondamentali e di terminologia usata:

- *Intento*: rappresenta un'azione che l'utente desidera eseguire. Si creano bot per supportare uno o più intenti correlati.

- *Slot*: sono parametri o informazioni aggiuntive richieste da un intento. Ad esempio, un intento "OrdinaPizza" potrebbe richiedere Slot, quali dimensione, tipo di crosta e numero di pizze.
- *Tipi di Slot*: ciascuno Slot ha un tipo, che può essere creato personalmente oppure ci si può rifare ai tipi di Slot predefiniti.
- *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.

### 3.3 Realizzazione di un chatbot con Amazon Lex

In questa sezione, saranno presentate le fasi di creazione, sviluppo e test di un chatbot utilizzando Amazon Lex, con particolare attenzione al suo impiego per la prenotazione di camere in strutture alberghiere.

#### 3.3.1 Creazione

Il primo passo da compiere nello sviluppo del chatbot è la sua creazione. Essa prevede una scelta iniziale del tipo di bot con cui cominciare; la scelta comprende un bot vuoto, un esempio da cui iniziare e uno strumento che genera automaticamente intenti dalle trascrizioni delle conversazioni caricate. La configurazione procede con l'inserimento obbligatorio del nome, l'inserimento facoltativo di una descrizione, le autorizzazioni IAM e il Children's Online Privacy Protection Act (COPPA) (Figura 3.1). L'ultima fase della creazione prevede la scelta della lingua desiderata e, se necessario, la selezione di una voce artificiale nell'applicazione di riferimento (Figura 3.2). Con questo passaggio si completa la configurazione iniziale del chatbot, aprendo la strada allo sviluppo effettivo.

**Configura le impostazioni dei bot** [Info](#)

**Metodo di creazione**

- Crea un bot vuoto**  
Crea un bot di base senza linguaggi, intenti e tipi di slot preconfigurati.
- Inizia con un esempio**  
Un bot di esempio ha lingue, intenti e tipi di slot preconfigurati. È possibile modificare queste impostazioni.
- Inizia con le trascrizioni**  
Genera automaticamente intenti dalle trascrizioni delle conversazioni caricate. Solo la lingua inglese (USA) è disponibile quando si inizia con una trascrizione.

**Configurazione bot**

**Nome bot**

Massimo 100 caratteri. Caratteri validi: A-Z, a-z, 0-9, -, \_

**Descrizione - facoltativo**

Questa descrizione appare nella pagina dell'elenco dei bot. Può aiutarti a identificare lo scopo del tuo bot.

Massimo 200 caratteri.

Figura 3.1: Creazione di un chatbot Amazon Lex

The screenshot shows the 'Aggiungi lingua al bot' (Add language to bot) interface in the Amazon Lex console. The title is 'Aggiungi lingua al bot' with an 'Info' link. Below the title, there is a dropdown menu for 'Lingua: Italiano (Italia)'. The main configuration area is divided into several sections: 'Seleziona lingua' with a dropdown menu currently set to 'Italiano (Italia)'; 'Descrizione - facoltativo' with a text input field and a note 'Massimo 200 caratteri.'; 'Interazione vocale' with a dropdown menu set to 'Adriano' and a note 'La voce della sintesi vocale utilizzata dal tuo bot per interagire con gli utenti.'; 'Esempio di voce' with a text input field containing 'Ciao, mi chiamo Adriano. Come posso aiutarti?' and a 'Riproduci' button; and 'Soglia del punteggio di confidenza per la classificazione dell'intento' with a text input field set to '0.40' and a note 'Min: 0,00, max: 1,00.'. At the bottom, there are three buttons: 'Annulla', 'Aggiungi un'altra lingua', and 'Fatto'.

Figura 3.2: Scelta della lingua e della voce del chatbot

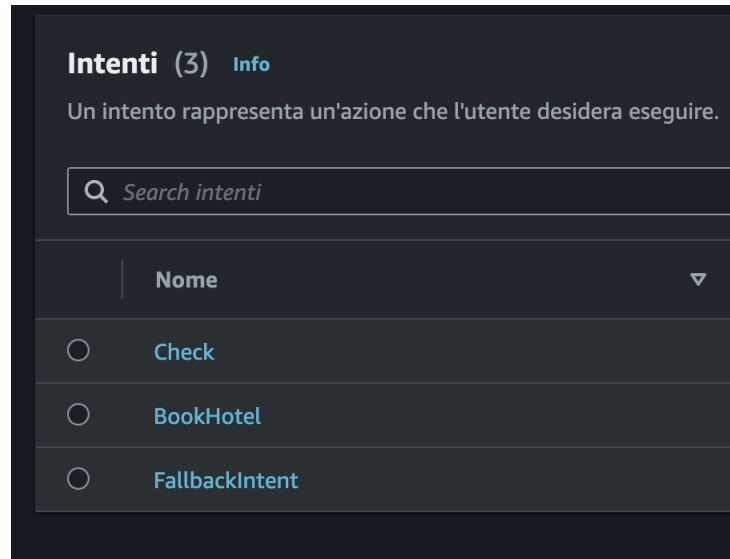
### 3.3.2 Gestione delle conversazioni

Subito dopo la fase di configurazione iniziale, si procede con la creazione degli intenti e degli Slot mirati a soddisfare il compito assegnato. In questa fase dello sviluppo, si definisce con precisione la logica del chatbot, identificando le azioni possibili da parte dell'utente (intent) e i parametri chiave necessari per eseguire tali azioni (Slot). Nell'esperienza da noi effettuata, sono stati creati due intenti personalizzati (Figura 3.3), ovvero:

- *BookHotel*; prevede una sample utterance con messaggi come "ciao", "Ciao", "buongiorno" ed altri, ai quali il chatbot risponde con una presentazione del servizio che offre, chiedendo, inoltre, quale operazione si vuole effettuare. È possibile scegliere tra *Prenotazione* e *Controllo prenotazione*. Se viene scelto il primo campo il bot continua chiedendo tutte le informazioni per effettuare la prenotazione, come nome, città dove soggiornare, data del check-in, notti e tipo di stanza.
- *Check*; a questo intent si ha accesso soltanto se viene scelto il secondo campo; esso permette di controllare se una prenotazione è stata effettuata. Il controllo avviene utilizzando un codice che viene rilasciato al momento della prenotazione.

Inoltre, di default, viene creato l'intento *FallbackIntent* che viene progettato per gestire situazioni in cui il sistema non riesce a determinare correttamente l'intento dell'utente. Si tratta di un "riparo" per affrontare input dell'utente ambigui o non compresi. Quando un utente fornisce un input che non corrisponde a nessun altro intent configurato nel chatbot, il sistema può attivare automaticamente il *FallBackIntent*. Quest'ultimo può essere progettato per restituire un messaggio generico che chiede all'utente di fornire ulteriori dettagli o di esprimersi in modo più chiaro.

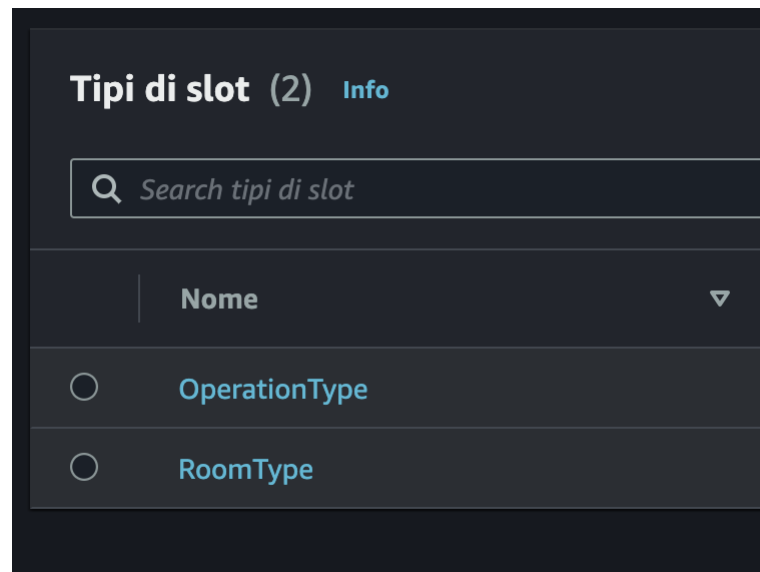




**Figura 3.3:** Lista degli intenti del chatbot

Successivamente, sono stati creati i tipi di Slot necessari per eseguire le azioni del bot. In particolare, sono stati creati due tipi, ovvero *OperationType* e *RoomType* (Figura 3.4).

All'interno degli Slot sono state inserite le risorse utilizzate per acquisire valori personalizzati dalla fase "Slot" degli intent.



**Figura 3.4:** Tipi di Slot del chatbot

Il tipo di Slot *OperationType* prevede un solo valore *Prenotare*. Lo Slot *RoomType*, invece, prevede i valori, *Singola*, *Doppia*, *Tripla* e *Matrimoniale* (Figura 3.5).

**Valori del tipo di slot**  
Modifica l'elenco dei valori utilizzati per addestrare il modello di machine learning a riconoscere i valori per uno slot.

Cerca i valori del tipo di slot

Singola	Tab o ; per il nuovo valore	X
	singola X	
Doppia	Tab o ; per il nuovo valore	X
	doppia X	
Tripla	Tab o ; per il nuovo valore	X
	tripla X	
Matrimoniale	Tab o ; per il nuovo valore	X
	matrimoniale X	
Valore	Tab o ; per il nuovo valore	Aggiungi valore

Massimo 140 caratteri. Caratteri validi: A-Z, a-z, 0-9, @, #, \$

**Figura 3.5:** Tipologie dello Slot "RoomType"

Di seguito sono riportati gli Slot dell'intent *BookHotel*, con i rispettivi messaggi che il bot invia; le risposte sono del tipo "Slot", ovvero vengono validate soltanto se il loro valore è uguale a quello inserito all'interno dei relativi Slot.

▶ Prompt per slot: OperationType Messaggio: Benvenuto nel nostro centro di prenotazio...	Tipo di slot OperationType	X
▶ Prompt per slot: Nome Messaggio: Bene, posso sapere il tuo nome?	Tipo di slot AMAZON.FirstName	X
▶ Prompt per slot: Location Messaggio: Ciao {Nome}. In quale città vuoi soggiorna...	Tipo di slot AMAZON.City	X
▶ Prompt per slot: Nights Messaggio: Quante notti?	Tipo di slot AMAZON.Number	X
▶ Prompt per slot: CheckInDate Messaggio: Da quale giorno?	Tipo di slot AMAZON.Date	X
▶ Prompt per slot: RoomType Messaggio: Che tipo di stanza desideri?	Tipo di slot RoomType	X

**Figura 3.6:** Slot dell'intent "BookHotel"

Infine, per il salvataggio dei dati su un database, è stato utilizzato il servizio DynamoDB, fornito da AWS, ed è stata implementata un funzione Lambda. Nelle Figure 3.7 e 3.8 viene riportata una parte del codice.

```

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

    if(intent=='BookHotel'):
        validation_result = validate(event['sessionState']['intent']['slots'],intent)

        if event['invocationSource'] == 'DialogCodeHook':
            if not validation_result['isValid']:
                if 'message' in validation_result:
                    response = {
                        "sessionState": {
                            "dialogAction": {
                                'slotToElicit':validation_result['violatedSlot'],
                                "type": "ElicitSlot"
                            },
                            "intent": {
                                'name':intent,
                                'slots': slots
                            }
                        },
                        "messages": [
                            {
                                "contentType": "PlainText",
                                "content": validation_result['message']
                            }
                        ]
                    }
                else:
                    response = {
                        "sessionState": {
                            "dialogAction": {
                                'slotToElicit':validation_result['violatedSlot'],
                                "type": "ElicitSlot"
                            },
                            "intent": {
                                'name':intent,
                                'slots': slots
                            }
                        }
                    }
                else:
                    response = {
                        "sessionState": {
                            "dialogAction": {
                                "type": "Delegate"
                            },
                            "intent": {
                                'name':intent,
                                'slots': slots
                            }
                        }
                    }
            }
            return response

        if event['invocationSource'] == 'FulfillmentCodeHook':
            if intent == 'BookHotel':
                # Add order in Database
                current_datetime = datetime.now().strftime("%Y-%m-%d %H:%M:%S")
                codice=str(int(time.time()))
                item = {
                    'Codice': codice,
                    'Nome': slots['Nome']['value']['originalValue'],
                    'Location': slots['Location']['value']['originalValue'],
                    'CheckInDate': slots['CheckInDate']['value']['originalValue'],
                    'Nights': slots['Nights']['value']['originalValue'],
                    'RoomType': slots['RoomType']['value']['originalValue'],
                    'Timestamp': current_datetime
                }
                table.put_item(Item=item)

                response = {
                    "sessionState": {
                        "dialogAction": {
                            "type": "Close"
                        },
                        "intent": {
                            'name':intent,
                            'slots': slots,
                            'state':'Fulfilled'
                        }
                    },
                    "messages": [
                        {
                            "contentType": "PlainText",
                            "content": f"Grazie, ho registrato la tua prenotazione con il codice: {codice}"
                        }
                    ]
                }
            }

```

Figura 3.7: Codice Lambda per il salvataggio su database (prima parte)

```

if event['invocationSource'] == 'FulfillmentCodeHook':

    validation_result = validate_check(codes)

    if not validation_result['isValid']:
        # Restituisci una risposta di errore se la validazione fallisce
        response = {
            "sessionState": {
                "dialogAction": {
                    "type": "Close"
                },
                "intent": {
                    "name": intent,
                    "slots": codes,
                    "state": 'Failed'
                }
            },
            "messages": [
                {
                    "contentType": "PlainText",
                    "content": 'Prenotazione non trovata nel database'
                }
            ]
        }
        return response

    # Altrimenti, la validazione ha avuto successo
    response = {
        "sessionState": {
            "dialogAction": {
                "type": "Close"
            },
            "intent": {
                "name": intent,
                "slots": codes,
                "state": 'Fulfilled'
            }
        },
        "messages": [
            {
                "contentType": "PlainText",
                "content": "Prenotazione effettuata il {}".format(validation_result['timestamp_value'])
            }
        ]
    }
    return response

}

return response

else:
    if intent == 'Check':
        validation_result = validate2(event['sessionState']['intent']['slots'], intent)

        if event['invocationSource'] == 'DialogCodeHook':
            if not validation_result['isValid']:
                if 'message' in validation_result:
                    response = {
                        "sessionState": {
                            "dialogAction": {
                                'slotToElicit': validation_result['violatedSlot'],
                                "type": "ElicitSlot"
                            },
                            "intent": {
                                "name": intent,
                                "slots": codes
                            }
                        },
                        "messages": [
                            {
                                "contentType": "PlainText",
                                "content": validation_result['message']
                            }
                        ]
                    }
                else:
                    response = {
                        "sessionState": {
                            "dialogAction": {
                                'slotToElicit': validation_result['violatedSlot'],
                                "type": "ElicitSlot"
                            },
                            "intent": {
                                "name": intent,
                                "slots": codes
                            }
                        }
                    }
                    return response
            else:
                response = {
                    "sessionState": {
                        "dialogAction": {
                            "type": "Delegate"
                        },
                        "intent": {
                            "name": intent,
                            "slots": codes
                        }
                    }
                }

```

Figura 3.8: Codice Lambda per il salvataggio su database (seconda parte)

### 3.3.3 Utilizzo

L'ultimo passaggio per lo sviluppo del chatbot è la fase di test e di deploy. 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 3.9 e 3.10 viene riportato un esempio di funzionamento del chatbot.

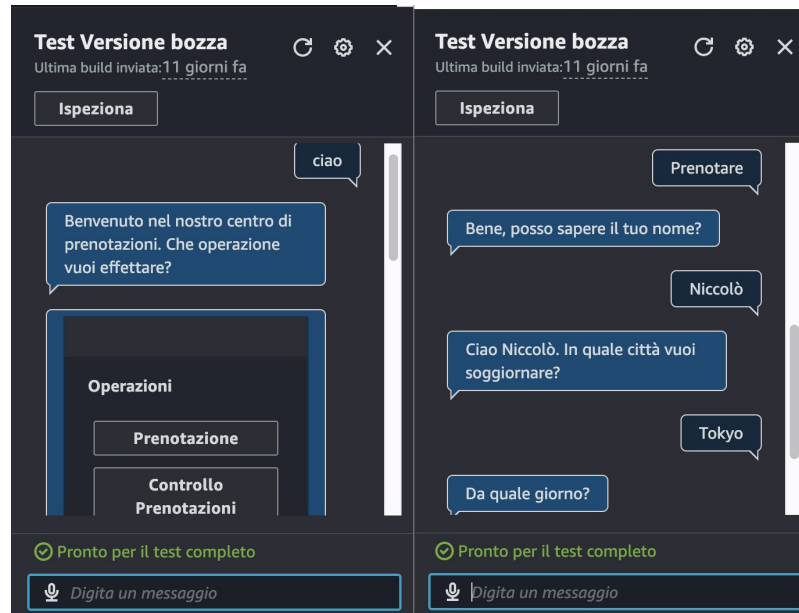


Figura 3.9: Esempio di una conversazione con il chatbot (prima parte)

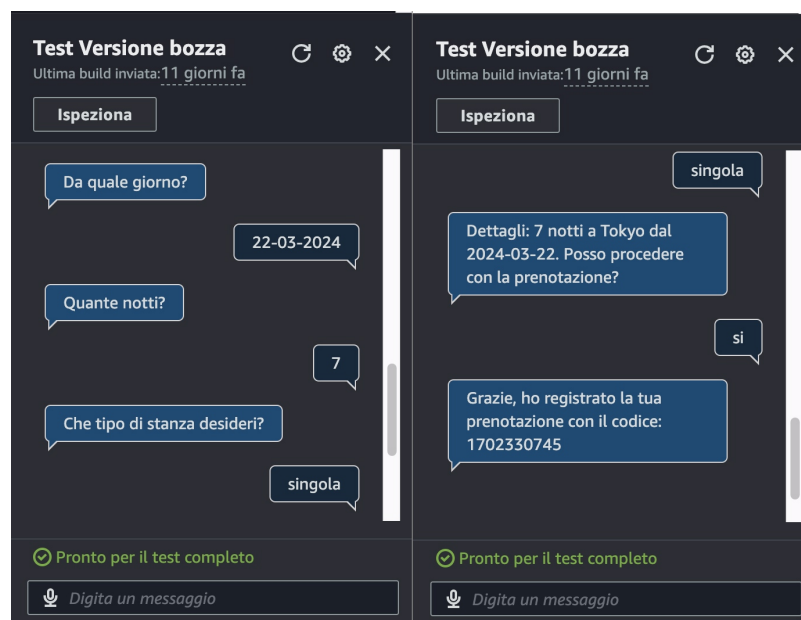


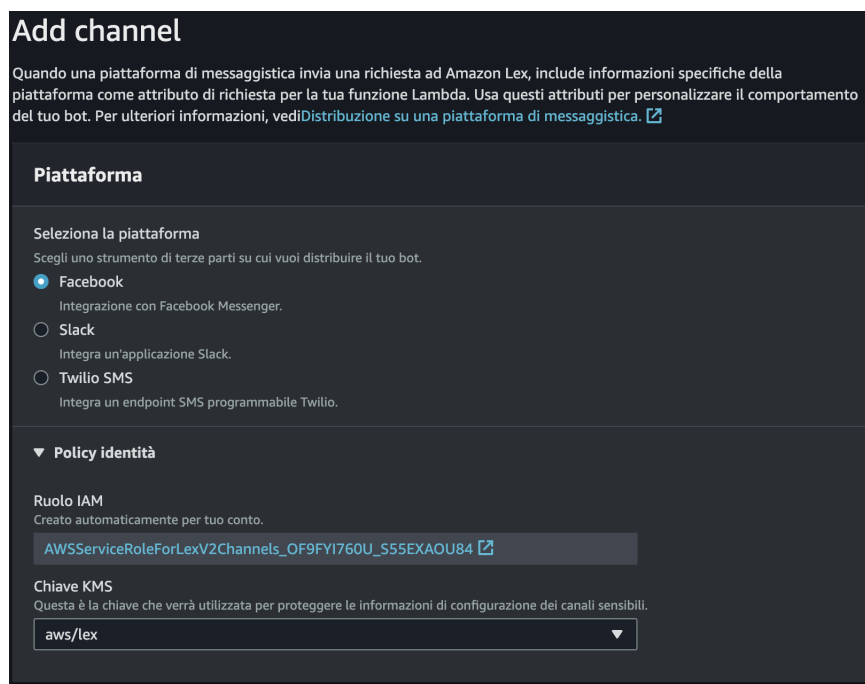
Figura 3.10: Esempio di una conversazione con il chatbot (seconda parte)

Terminata la fase di test, è possibile distribuire il chatbot. Prima di tutto 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 3.11). Una volta effettuata la scelta,

la configurazione prosegue scegliendo il ruolo IAM e la chiave KMS 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 3.12), ovvero:

- *Token di accesso alla pagina*: è un tipo di credenziale che autorizza il bot a compiere determinate azioni sulla pagina Facebook, 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.



The screenshot shows the 'Add channel' configuration page in the Amazon Lex console. At the top, there is a title 'Add channel' and a descriptive paragraph: 'Quando una piattaforma di messaggistica invia una richiesta ad Amazon Lex, include informazioni specifiche della piattaforma come attributo di richiesta per la tua funzione Lambda. Usa questi attributi per personalizzare il comportamento del tuo bot. Per ulteriori informazioni, vedi Distribuzione su una piattaforma di messaggistica.' Below this, there are three main sections: 1. 'Piattaforma' (Platform), which includes a sub-section 'Seleziona la piattaforma' (Select the platform) with three radio button options: 'Facebook' (selected), 'Slack', and 'Twilio SMS'. 2. 'Policy identità' (Identity policy), which includes a sub-section 'Ruolo IAM' (IAM role) with a text input field containing 'AWSServiceRoleForLexV2Channels\_OF9FYI760U\_S55EXAOU84'. 3. 'Chiave KMS' (KMS key), which includes a sub-section 'Chiave KMS' with a dropdown menu currently showing 'aws/lex'.

Figura 3.11: Schermata per l'aggiunta di un canale dalla console Amazon Lex

**Configurazione dell'integrazione**

**Nome**  
  
Massimo 100 caratteri. Caratteri validi: A-Z, a-z, 0-9, -, \_

**Descrizione - *facoltativo***  
  
Massimo 200 caratteri.

**Alias**

**Lingua**

**Configurazione aggiuntiva**

**Alias**

**Token di accesso alla pagina**

**Chiave segreta dell'app**

**Figura 3.12:** Schermata per la configurazione di un canale dalla console Amazon Lex

Una volta completati tutti i passaggi, verrà generato un URL di callback. Questo URL svolge un ruolo fondamentale nel processo di autorizzazione dell'utente e nel collegamento del bot con Facebook. 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.

---

## Esperienze nell'ambito dell'Image Recognition

---

*In questo capitolo verrà trattato il campo dell'Intelligenza Artificiale che si occupa del riconoscimento delle immagini, ovvero l'Image Recognition, passando dalla sua definizione al suo funzionamento, per poi occuparsi della spiegazione del caso di studio effettuato, ovvero Amazon Rekognition. Verrà, infine, presentato, un test realizzato con quest'ultimo strumento.*

### 4.1 Introduzione al riconoscimento di immagini

Nel contesto degli esseri umani, la vista, insieme agli altri sensi, partecipa attivamente alla percezione dell'ambiente circostante. Tuttavia, questa affermazione non può essere estesa in modo diretto ad un computer. Per una comprensione più approfondita di questo concetto, la presente sezione esplorerà il tema della riconoscimento delle immagini.

#### 4.1.1 Cos'è l'Image Recognition

L'Image Recognition, o riconoscimento di immagini, è una branca dell'Intelligenza Artificiale che si occupa di insegnare a un sistema informatico ad interpretare e comprendere il contenuto visivo delle immagini. L'obiettivo principale è consentire ai computer di analizzare e tradurre informazioni visive in modo simile o superiore alle capacità umane.

Il campo del riconoscimento di immagini è alimentato da algoritmi di Machine Learning, in particolare, Deep Learning, che utilizzano reti neurali artificiali per apprendere modelli complessi dai dati visivi. Questi modelli vengono addestrati su ampi insiemi di dati contenenti immagini etichettate, consentendo loro di riconoscere e distinguere oggetti, pattern e caratteristiche nelle nuove immagini.

È interessante notare, che i termini "riconoscimento dell'immagine" e "rilevamento dell'immagine" vengono talvolta utilizzati in modo intercambiabile, ma presentano importanti differenze tecniche. Il rilevamento dell'immagine si occupa di individuare oggetti all'interno di un'immagine, senza necessariamente attribuire loro significato. Ad esempio, il rilevamento dei volti cerca modelli di volti nelle immagini senza preoccuparsi di determinare la loro identità. D'altra parte, il riconoscimento dell'immagine si concentra sull'identificare gli oggetti di interesse e assegnarli a categorie specifiche o classi.



### 4.1.2 Come funziona

Per approfondire la tecnologia sottostante all'Image Recognition è fondamentale acquisire una comprensione dettagliata del suo funzionamento. Come precedentemente accennato, l'Image Recognition fa uso del Machine Learning per elaborare immagini digitali e riconoscere modelli e figure in esse presenti. Gli algoritmi vengono addestrati attraverso ampi dataset al fine di apprendere a distinguere diversi oggetti, consentendo al modello allenato di classificare le immagini in diverse categorie.

Il processo di Image Recognition si articola tipicamente nelle seguenti fasi:

1. *Raccolta dei dati*: si tratta di un ampio set di immagini descritte, successivamente utilizzate per addestrare l'algoritmo.
2. *Pulizia e rimozione delle distorsioni*: il procedimento può richiedere l'ottimizzazione di alcune peculiarità dell'immagine, come la lucentezza o il contrasto, al fine di renderla fruibile.
3. *Estrazione delle caratteristiche*: riguarda l'identificazione di parti isolate e rilevanti dell'immagine, utilizzate per distinguere le diverse immagini da parte dell'algoritmo.
4. *Addestramento del modello*: una volta estratte le caratteristiche, l'algoritmo viene addestrato su queste immagini.
5. *Test del modello e valutazione*: è il processo che avviene dopo l'addestramento e serve per identificare eventuali errori o debolezze del modello che possono essere migliorati.
6. *Deployment*: il modello è pronto per essere utilizzato dopo essere stato testato e validato, permettendo di classificare nuove immagini in nuove categorie con elevata precisione.

L'Image Recognition si configura come una forma tecnologica avanzata di Machine Learning, rappresentando un ambito di ricerca e sviluppo in costante evoluzione nel campo dell'Intelligenza Artificiale.

Esistono tre metodologie principali per addestrare algoritmi di Image Recognition: il *Supervised Learning*, l'*Unsupervised Learning* e il *Self-Supervised Learning*.

Nel *Supervised Learning*, i dati sono pre-categorizzati, consentendo agli algoritmi di distinguere tra diversi oggetti. Le categorie in cui gli oggetti vengono riconosciuti sono già note in questo approccio.

L'*Unsupervised Learning*, al contrario, viene impiegato quando i dati non sono categorizzati. In questo caso, gli algoritmi determinano le similarità e le differenze tra le immagini, senza avere conoscenza preventiva delle categorie.

Infine, è possibile ricorrere al *Self-Supervised Learning*, che, pur condividendo somiglianze con *Unsupervised Learning* per l'assenza di dati pre-categorizzati, si distingue per il fatto che il modello viene addestrato attraverso pseudo-categorizzazioni generate internamente ai dati stessi. Questo approccio è preferito in situazioni di scarsità di dati.

Le applicazioni dell'Image Recognition sono molteplici e impattano diversi settori, tra cui:

- *Riconoscimento degli Oggetti*; riguarda l'identificazione e la classificazione di oggetti specifici all'interno di un'immagine. Ad esempio, un'app di riconoscimento degli oggetti potrebbe identificare automobili, animali o oggetti domestici in una foto.
- *Riconoscimento del Testo*; riguarda l'estrazione e l'interpretazione di testo da immagini. Questo è spesso utilizzato in applicazioni OCR (Optical Character Recognition) per digitalizzare testi da documenti cartacei o immagini.

- *Riconoscimento Facciale*; riguarda l'identificazione e l'analisi di volti umani nelle immagini. Applicazioni comuni includono il riconoscimento facciale per il riconoscimento degli utenti e l'analisi delle espressioni facciali.
- *Diagnostiche Mediche*; riguarda l'analisi di immagini mediche per diagnosticare condizioni e identificare anomalie. Ad esempio, il riconoscimento di immagini può essere utilizzato per identificare tumori in scansioni radiologiche.
- *Sicurezza e Sorveglianza*; riguarda il monitoraggio e l'identificazione automatica di attività sospette o di oggetti in contesti di sicurezza e sorveglianza.
- *Veicoli a Guida Autonoma*; riguarda l'identificazione di oggetti sulla strada, la segnaletica stradale, i pedoni e, gli altri veicoli; è cruciale per i sistemi di guida autonoma.
- *Settore Alimentare*; riguarda il riconoscimento degli ingredienti o del cibo in immagini; è utilizzato, ad esempio, nelle app per il conteggio delle calorie.

## 4.2 Caso di studio: Amazon Rekognition

Questa sezione è dedicata all'approfondimento di Amazon Rekognition, un servizio di analisi delle immagini e dei video basato sull'Intelligenza Artificiale fornito da Amazon Web Services.

### 4.2.1 Che cos'è

Amazon Rekognition 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.
- offerta di un'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.

Basato sulla tecnologia collaudata e scalabile di Deep Learning sviluppata da Amazon, Rekognition non richiede competenze approfondite in Machine Learning. Include un'API semplice da usare per analizzare rapidamente immagini o video memorizzati in Amazon S3.

### 4.2.2 Funzionalità

Le potenzialità di Amazon Rekognition si estendono a una vasta gamma di casi d'uso, offrendo un insieme completo di funzionalità avanzate di analisi di immagini e video. Tra le sue applicazioni più comuni troviamo:

- *Librerie di immagini e video ricercabili*: Amazon Rekognition rende immagini e video memorizzati ricercabili in modo da poter scoprire le etichette (oggetti, concetti e scene) che vi compaiono.

- *Rilevamento della vivacità facciale*: Amazon Rekognition Face Liveness è una funzione di Machine Learning completamente gestita e progettata per aiutare gli sviluppatori a prevenire frodi nella verifica dell'identità basata sul volto.
- *Verifica dell'utente basata sul volto*: Amazon Rekognition consente la conferma dell'identità degli utenti. Il sistema confronta la loro immagine dal vivo con un'immagine di riferimento.
- *Rilevamento e analisi facciale*: Amazon Rekognition può rilevare e analizzare diversi componenti e attributi facciali, come espressioni emotive, informazioni demografiche e direzione dello sguardo.
- *Ricerca facciale*: con Amazon Rekognition è possibile cercare immagini, video memorizzati e video in streaming per volti che corrispondono a quelli memorizzati in una raccolta degli stessi.
- *Rilevamento di attrezzature di protezione individuale (PPE)*: Amazon Rekognition rileva l'uso di attrezzature di protezione individuale, come copri viso, copri testa e copri mani, su persone all'interno delle immagini.
- *Rilevamento di contenuti non sicuri*: Amazon Rekognition può rilevare contenuti per adulti e contenuti violenti in immagini e video memorizzati.
- *Riconoscimento di celebrità*: Amazon Rekognition può riconoscere celebrità in immagini e video.
- *Rilevamento di testo*: Amazon Rekognition Text in Image consente di riconoscere ed estrarre contenuto testuale da immagini.
- *Etichette personalizzate*: con Amazon Rekognition Custom Label, è possibile identificare le etichette (oggetti e concetti) e le scene nelle immagini specifiche alle esigenze aziendali.

Altri vantaggi di Amazon Rekognition includono l'integrazione semplice con altri servizi AWS, l'analisi scalabile delle immagini, l'analisi e il filtraggio delle stesse in base alle proprietà e un costo basso con un modello di prezzi graduato.

Amazon Rekognition offre due insiemi di API. Si utilizza Amazon Rekognition Image per analizzare le immagini e Amazon Rekognition Video per analizzare i video.

Entrambe le API analizzano immagini e video per fornire informazioni utili nelle applicazioni.

Amazon Rekognition Image potrebbe essere utilizzato per arricchire l'esperienza del cliente in un'applicazione di gestione delle foto. Quando un cliente carica una foto, l'applicazione può sfruttare Amazon Rekognition Image per rilevare oggetti o volti del mondo reale nell'immagine. Dopo che l'applicazione archivia le informazioni restituite da Amazon Rekognition Image, l'utente potrebbe interrogare la propria raccolta per trovare foto con un oggetto o un volto specifico. È possibile effettuare interrogazioni più approfondite, come la ricerca di volti sorridenti o di una certa età. I risultati di qualsiasi etichetta rilevata (oggetti o concetti) possono essere filtrati utilizzando una varietà di filtri inclusivi o esclusivi.

Amazon Rekognition Video può essere utilizzato per tracciare il percorso delle persone in un video archiviato. In alternativa è possibile utilizzare Amazon Rekognition Video per cercare un video in streaming di persone le cui descrizioni facciali corrispondono a descrizioni facciali già archiviate dal sistema. Le operazioni di analisi video consentono di ordinare per timestamp o nome dell'etichetta e di aggregare i risultati per timestamp o segmento video.

Amazon Rekognition consente di migliorare la precisione di determinati modelli di Deep Learning mediante l'addestramento personalizzato. Con Amazon Rekognition Custom Moderation, è possibile adattare il modello di base per l'analisi delle immagini utilizzando un addestratore personalizzato con le proprie immagini.

## 4.3 Test di Amazon Rekognition

In questa sezione, saranno presentate le fasi di creazione di un dataset, di etichettatura delle immagini, nonché quelle di addestramento e valutazione di un modello. L'obiettivo principale è riconoscere diverse caratteristiche delle automobili, tra cui il colore, la tipologia, la marca, etc.

### 4.3.1 Creazione del dataset

Un dataset è una raccolta di immagini ed etichette di immagini che possono essere utilizzate per eseguire l'addestramento o il test di un modello.

Per prima cosa si creano un progetto ed il suo dataset (Figura 4.1). Si inizializza il set scegliendo l'opzione di configurazione che si presta di più all'applicazione; la differenza di scelta tra le opzioni avviene in base al modo di utilizzare le immagini. Nel primo caso, si utilizza un solo set, l'80% delle sue immagini viene utilizzato per l'addestramento, mentre il restante 20% viene usato per il test. Nel secondo caso, invece, si utilizzano due dataset separati, uno per l'addestramento e uno per il test. La configurazione prosegue con la scelta per l'importazione delle immagini; questa può avvenire da un bucket S3, dal proprio computer, da un Dataset già esistente su Amazon Rekognition oppure da SageMaker Ground Truth.

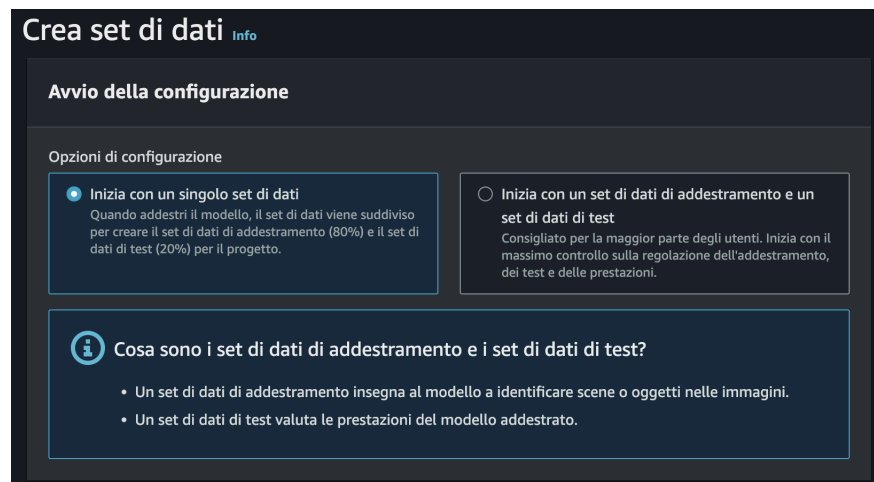


Figura 4.1: Configurazione di un dataset

A quest punto, le immagini possono essere caricate; i formati supportati da Amazon Rekognition sono del tipo JPEG o PNG e le dimensioni massime sono di 15 MB, se i file vengono inviati come oggetti di S3, e di 5 MB, se sono inviati come array di byte.

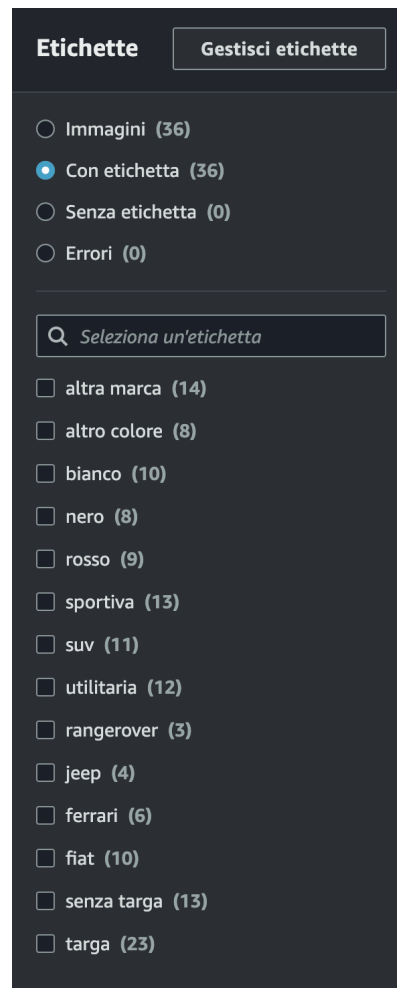
Nel caso d'uso in questione è stato creato un dataset con 52 immagini, di cui 36, ovvero l'80%, utilizzate per l'addestramento e 16, cioè il 20%, utilizzate per il test. Di seguito è riportato un esempio di immagini utilizzate (Figura 4.2).



**Figura 4.2:** Esempio di immagini utilizzate per l'addestramento

### 4.3.2 Labelling delle immagini

Dopo aver caricato il dataset inizia l'attività di etichettatura o labelling; essa è fondamentale poiché consente al sistema di apprendere dagli esempi forniti e migliorare la sua capacità di identificare gli oggetti desiderati nelle nuove immagini. Nel caso di studio effettuato sono state utilizzate le etichette riportate in Figura 4.3.



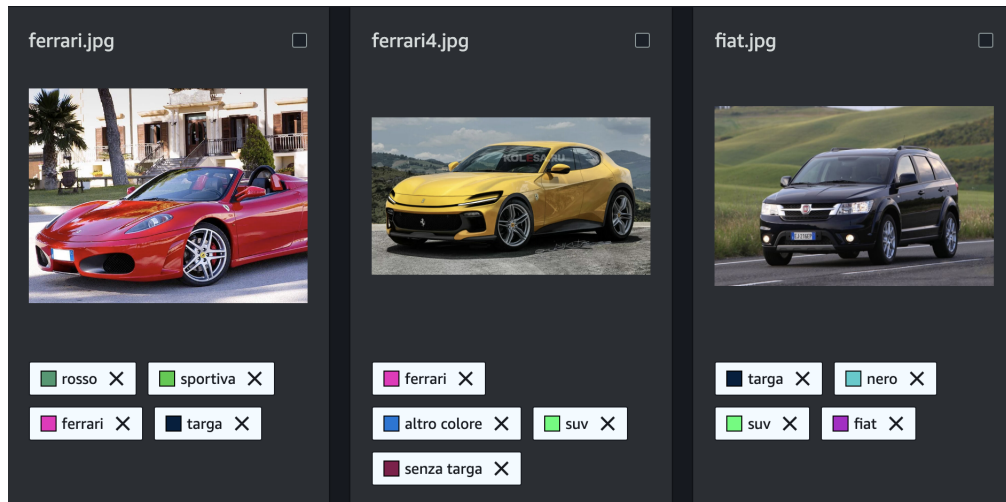
**Figura 4.3:** Lista delle etichette

Successivamente, si procede con il processo di labelling vero e proprio, il quale prevede l'etichettatura manuale di ciascuna immagine nel dataset. L'interfaccia di Custom Labels offre due modalità di etichettatura:

- *Assegna etichetta a livello di immagine;* viene utilizzata quando si ha un singolo soggetto o una caratteristica principale nell'immagine che si desidera etichettare.

- *Disegna riquadri delimitatori*; viene utilizzata quando ci sono più soggetti o caratteristiche diverse nell'immagine che devono essere etichettati separatamente.

Le immagini possono essere selezionate in modo multiplo al fine di accelerare il processo. Una volta completata questa fase, si visualizza una schermata (Figura 4.4) in cui vengono riportate tutte le etichette associate ad ogni immagine.



**Figura 4.4:** Esempio di immagini con le etichette personalizzate

### 4.3.3 Addestramento del modello

Una volta assegnate tutte le etichette ad ogni immagine, si può procedere con il training (o addestramento) del modello. Selezionando l'opzione "Train model" all'interno dell'interfaccia di AWS Rekognition comincerà il processo di addestramento, che può durare, in media, dai 30 minuti fino alle 24 ore, in base alla quantità e alla qualità delle immagini e alla quantità di etichette. Al termine di questo processo, sarà visualizzato lo stato dell'addestramento.

La schermata mostrata (Figura 4.5) rappresenta il processo di addestramento del modello nel caso di studio in questione. Sono riportate diverse informazioni tra cui le più rilevanti sono l'F1-Score, lo stato attuale del modello e il messaggio di stato associato. Nel caso in cui l'addestramento non sia completato con successo, la voce "Stato del modello" mostrerà TRAINING FAILED, fornendo ulteriori dettagli nel campo "Messaggio di stato". Quest'ultimo conterrà un breve messaggio che indica la natura dell'errore riscontrato dal sistema durante il processo di addestramento.

	Nome	Data di creazione	Set di dati per la formazione	Test del set di dati	Prestazioni del modello (Punteggio F1)	Stato del modello	Messaggio di stato
<input type="checkbox"/>	carsrecognition.2023-11-30T11.32.23	November 30, 2023			0.961	STOPPED	The model is ready for hosting.

**Figura 4.5:** Schermata di training completato

### 4.3.4 Valutazione del modello

Terminato l'addestramento, nella sezione "Valutazione del modello", è possibile trovare una serie di informazioni relative al modello e alle sue prestazioni (Figura 4.6). Tra queste abbiamo:

- *F1-Score*; si tratta di una metrica che combina la Precision e il Recall in un singolo valore. È particolarmente utile quando ci interessa bilanciare le due grandezze da cui deriva. L’F1-Score è calcolato utilizzando la formula:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Dove Precision è la proporzione di istanze rilevate correttamente rispetto a tutte le istanze rilevate, e Recall è la proporzione di istanze rilevate correttamente rispetto a tutte le istanze effettive. Il suo valore massimo è 1.0, che indica un modello che ha un equilibrio ottimale tra Precision e Recall. Nel caso d’uso il valore è di 0,961.

- *Average Precision*; l’Average Precision è una metrica utilizzata principalmente nelle valutazioni degli algoritmi di Object Detection. Essa rappresenta l’area sotto la curva Precision-Recall (APR). Questa curva mostra come variano la Precision e il Recall al variare di una soglia di confidenza. L’AP calcolato fornisce una misura del rendimento del modello su diverse soglie di confidenza. Il valore massimo possibile è 1.0, che indica che tutte le predizioni positive del modello sono corrette. Nel caso d’uso il valore è di 0,956.
- *Overall Recall*; il Recall è una misura della capacità di un modello di identificare tutti gli esempi rilevanti in un dataset. L’Overall Recall rappresenta il Recall medio su tutti i casi, indicando quanto bene il modello stia identificando gli oggetti di interesse in modo globale. Anche qui, il valore massimo è 1.0, che significa che il modello riesce a identificare tutte le istanze positive nel dataset. Nel caso d’uso il valore è di 0,974.

Le informazioni relative al set di dati durante il processo di addestramento includono, anche, il totale di immagini ed etichette per la formazione e per l’addestramento. Inoltre, viene fornito il tempo impiegato per il training totale del modello, in questo caso 0.622 ore, ovvero circa 38 minuti. Questi dettagli sono essenziali per valutare l’efficacia e l’efficienza dell’addestramento, consentendo una comprensione approfondita delle prestazioni del modello.

Punteggio F1 <a href="#">Info</a> <b>0.961</b>	Average precision <a href="#">Info</a> <b>0.956</b>	Overall recall <a href="#">Info</a> <b>0.974</b>
Date completed November 30, 2023 Trained in 0.622 hours	Set di dati per la formazione 14 labels, 33 images	Set di dati per la verifica 14 labels, 15 images

**Figura 4.6:** Risultati del modello

Al fine di valutare in modo indipendente i parametri forniti da Amazon Rekognition, è previsto, anche, l’utilizzo di comandi AWS CLI, eseguiti attraverso la console integrata AWS CloudShell. Questa procedura consente di testare il modello mediante l’inserimento manuale di immagini diverse da quelle del dataset, permettendo, così, la verifica del corretto funzionamento del sistema.

Amazon Rekognition offre la disponibilità di tre comandi principali:

- *Avvia il modello*, mostrato in Figura 4.7;

```
aws rekognition start-project-version \  
  --project-version-arn "arn:aws:rekognition:eu-central-1:443931443643:project/  
carsrecognition/version/carsrecognition.2023-11-30T11.32.23/1701340343736" \  
  --min-inference-units 1 \  
  --region eu-central-
```

**Figura 4.7:** Comando per avviare il modello nella console AWS CloudShell

- *Analizza l'immagine*, mostrato in Figura 4.8;

```
aws rekognition detect-custom-labels \  
  --project-version-arn "arn:aws:rekognition:eu-central-1:443931443643:project/  
carsrecognition/version/carsrecognition.2023-11-30T11.32.23/1701340343736" \  
  --image '{"S3Object": {"Bucket": "MY_BUCKET", "Name": "PATH_TO_MY_IMAGE"}}' \  
  --region eu-central-1
```

**Figura 4.8:** Comando per analizzare l'immagine nella console AWS CloudShell

- *Arresta il modello*, mostrato in Figura 4.9;

```
aws rekognition stop-project-version \  
  --project-version-arn "arn:aws:rekognition:eu-central-1:443931443643:project/  
carsrecognition/version/carsrecognition.2023-11-30T11.32.23/1701340343736" \  
  --region eu-central-1
```

**Figura 4.9:** Comando per arrestare il modello nella console AWS CloudShell

Nel caso d'uso in questione è stata utilizzata l'immagine in Figura 4.10 per testare il funzionamento del modello. Il risultato è riportato in Figura 4.11.



**Figura 4.10:** Immagine utilizzata per il test del modello



```
AWS CloudShell
eu-central-1
[cloudshell-user@ip-10-2-9-113 ~]$ aws rekognition detect-custom-labels \
> --project-version-arn "arn:aws:rekognition:eu-central-1:443931443643:project/carsrecognition/version/carsrecognition.2023-11-30T11.32.23/1701340343736" \
> --image '{"S3Object": {"Bucket": "mytestcarbucket", "Name": "test3.jpg"}}' \
> --region eu-central-1
{
  "CustomLabels": [
    {
      "Name": "utilitaria",
      "Confidence": 86.33200073242188
    },
    {
      "Name": "fiat",
      "Confidence": 85.32400512695312
    },
    {
      "Name": "targa",
      "Confidence": 76.52799987792969
    },
    {
      "Name": "rosso",
      "Confidence": 60.000003814697266
    }
  ]
}
[cloudshell-user@ip-10-2-9-113 ~]$
```

**Figura 4.11:** Schermata con i risultati del test

---

## Esperienze nell'ambito del Text Recognition

---

*Nel quinto capitolo verrà esplorato il mondo dell'Intelligenza Artificiale che si occupa del riconoscimento e rilevamento del testo nelle immagini, noto come Text Recognition. Si analizzerà cosa è, come funziona e verrà osservato un caso di studio importante, ovvero Amazon Textract. Infine, verrà presentato un test pratico eseguito con questa potente soluzione di estrazione del testo da immagini.*

### 5.1 Introduzione al riconoscimento di testi

La storia del testo risale a migliaia di anni fa. Le informazioni semantiche, ricche e precise, trasportate dal testo sono importanti in una vasta gamma di scenari applicativi basati sulla visione. Di conseguenza, il riconoscimento del testo in scenari naturali è un argomento di ricerca attivo nell'ambito della visione artificiale e del riconoscimento di pattern.

#### 5.1.1 Cos'è

Il Text Recognition è una tecnologia che consente ai computer di identificare e interpretare il testo presente in immagini, documenti cartacei, o in altri formati visivi. L'obiettivo principale di quest'ultimo, è quello di estrarre informazioni significative dal testo visibile, rendendolo accessibile per l'elaborazione da parte di sistemi informatici. La sua applicazione principale riguarda l'OCR, ovvero l'Optical Character Recognition.

Il riconoscimento ottico dei caratteri (OCR) è il processo che converte un'immagine di testo in un formato di testo leggibile dalla macchina. Ad esempio, se si esegue la scansione di un modulo o di una ricevuta, il computer salva la scansione come file immagine. Non sarà possibile utilizzare un editor di testo per modificare, cercare o contare le parole nel file immagine. Tuttavia, è possibile utilizzare l'OCR per convertire l'immagine in un documento di testo con i suoi contenuti archiviati come dati di testo.

#### 5.1.2 Come funziona

Il funzionamento del meccanismo OCR, o software OCR, può essere suddiviso in diversi passaggi, tra i quali:

- *Acquisizione dell'immagine*; uno scanner legge i documenti e li converte in dati binari. Il software OCR analizza l'immagine scansionata e classifica le aree chiare come background e le aree scure come testo.

- *Pre-elaborazione*; il software OCR prima pulisce l'immagine e rimuove gli errori per prepararla alla lettura. Ecco alcune delle tecniche di pulizia:
  - Leggero raddrizzamento o ribaltamento del documento scansionato, per sistemare i problemi di allineamento durante la scansione.
  - Smacchiatura o rimozione di qualsiasi chiazza sull'immagine digitale o smussatura dei bordi delle immagini di testo.
  - Eliminazione delle caselle e delle linee nell'immagine.
  - Riconoscimento dello script per la tecnologia OCR multi-lingua.
- *Riconoscimento testuale*; i due tipi principali di algoritmi OCR sono detti pattern matching ed estrazione delle caratteristiche. Il pattern matching funziona isolando un'immagine di carattere, chiamata glifo, e confrontandola con un glifo simile archiviato. Il riconoscimento del pattern funziona solo se il glifo archiviato ha un font e una dimensione simile al glifo di input. Questo metodo funziona bene con le immagini scansionate di documenti che sono stati digitati in un font conosciuto. L'estrazione delle caratteristiche scompone o decompone i glifi in caratteristiche come linee, cerchi chiusi, direzioni di linea e intersezioni di linea. Quindi il sistema utilizza queste caratteristiche per trovare l'abbinamento migliore o più vicino tra i vari glifi archiviati.
- *Post elaborazione*; dopo l'analisi, il sistema converte i dati testuali estratti in un file computerizzato. Alcuni sistemi OCR possono creare file PDF annotati che includono sia la prima che la seconda versione del documento scansionato.

## 5.2 Caso di studio: Amazon Textract

Questa sezione è focalizzata sull'analisi approfondita di Amazon Textract, un servizio di estrazione automatica di testo fornito da Amazon Web Services. Amazon Textract utilizza l'Intelligenza Artificiale per analizzare documenti contenenti testo, nonché per rilevare ed estrarre informazioni testuali in modo preciso e automatizzato.

### 5.2.1 Che cos'è

Amazon Textract è uno strumento che facilita il rilevamento e l'analisi del testo da documenti per le applicazioni. Utilizzando questo servizio, è possibile:

- rilevare testo digitato e scritto a mano in vari documenti, tra cui rapporti finanziari, cartelle cliniche e moduli fiscali;
- estrarre testo, moduli e tabelle da documenti con dati strutturati;
- specificare ed estrarre informazioni dai documenti;
- elaborare fatture e ricevute;
- elaborare documenti di identità come patenti di guida e passaporti emessi dal governo degli Stati Uniti;
- caricare e processare pacchetti di prestiti ipotecari attraverso il routing automatico delle pagine del documento alle appropriate operazioni di analisi di Amazon Textract tramite il flusso di lavoro di Analisi Lending;

- utilizzare query personalizzate per personalizzare la funzionalità di query preaddestrata utilizzando i dati per supportare le esigenze di elaborazione downstream;

Amazon Textract si basa sulla stessa tecnologia di Deep Learning altamente scalabile e collaudata sviluppata dagli scienziati della visione artificiale di Amazon per analizzare miliardi di immagini e video ogni giorno. Per utilizzarlo non è necessaria alcuna esperienza nel campo del Machine Learning, poiché Amazon Textract include API semplici e facili da usare per analizzare file di immagini e PDF. Amazon Textract impara costantemente da nuovi dati, e Amazon continua ad aggiungere nuove funzionalità al servizio.

Alcuni casi d'uso comuni di Amazon Textract includono la creazione di un indice di ricerca intelligente, l'estrazione intelligente di testo per l'elaborazione del linguaggio naturale (NLP), l'accelerazione della cattura e normalizzazione di dati da diverse fonti e l'automazione della classificazione ed estrazione di documenti. Tra i benefici di Amazon Textract vi sono l'integrazione semplice del rilevamento del testo, l'analisi scalabile dei documenti, i costi ridotti e molte altre caratteristiche che favoriscono l'efficienza e la convenienza.

### 5.2.2 Come funziona

Amazon Textract consente di rilevare e analizzare il testo in documenti di input a pagina singola o multipagina. Esso offre operazioni sincrone per l'elaborazione di documenti di piccole dimensioni a pagina singola e con risposte quasi in tempo reale. Amazon Textract fornisce, inoltre, operazioni asincrone che possono essere utilizzate per elaborare documenti più grandi e multipagina. Le risposte asincrone non sono in tempo reale.

Quando un'operazione Amazon Textract elabora un documento, i risultati vengono restituiti in un array di oggetti di tipo *Block* o una serie di oggetti di tipo *ExpenseDocuments*. Entrambi gli oggetti contengono informazioni rilevate sugli elementi, inclusa la loro posizione nel documento e la loro relazione con altri elementi del documento.

Le operazioni Amazon Textract restituiscono diversi tipi di oggetti a seconda dell'esecuzione delle operazioni. Per il rilevamento del testo e l'analisi di un documento generico, l'operazione restituisce un oggetto *Block*. Per l'analisi di una fattura o di una ricevuta, l'operazione restituisce un oggetto *ExpenseDocuments*. Per l'analisi della documentazione di identità, l'operazione restituisce un oggetto *IdentityDocumentFields*.

Quando Amazon Textract elabora un documento, crea un elenco di oggetti di tipo *Block* per il testo rilevato o analizzato. Ogni documento è costituito dai seguenti blocchi:

- pagine;
- linee e parole del testo;
- dati di tipo Form (coppie chiave-valore);
- tabelle e celle;
- layout;

Amazon Textract restituisce una rappresentazione di un documento sotto forma di elenco di diversi tipi di oggetti *Block* che sono collegati in una relazione genitore-figlio o in una coppia chiave-valore (Figura 5.1). Viene restituito, anche, un metadato che fornisce il numero di pagine in un documento.

```
{
  "BlockType": "PAGE",
  "Geometry": {
    "BoundingBox": {
      "Width": 1.0,
      "Height": 1.0,
      "Left": 0.0,
      "Top": 0.0
    },
    "Polygon": [
      { "X": 0.0, "Y": 0.0 },
      { "X": 1.0, "Y": 0.0 },
      { "X": 1.0, "Y": 1.0 },
      { "X": 0.0, "Y": 1.0 }
    ]
  },
  "Id": "example-id",
  "Relationships": [
    {
      "Type": "CHILD",
      "Ids": ["child-block-id-1", "child-block-id-2"]
    }
  ],
  "EntityTypes": ["KEY"]
}
```

Figura 5.1: Esempio di un blocco di tipo PAGE in formato JSON

Il blocco contiene, inoltre, informazioni sulla geometria, come la bounding box e il poligono che rappresenta la forma del blocco. L'ID è un identificatore univoco per il blocco. La relazione indica il collegamento con altri blocchi di tipo *CHILD* tramite i rispettivi ID. Gli *EntityTypes* indicano i tipi di entità associate al blocco, come, ad esempio, "KEY", che potrebbe rappresentare il testo chiave.

Infine, le operazioni di Amazon Textract restituiscono la percentuale di confidenza che il servizio ha nell'accuratezza con cui rileva l'oggetto. Per ottenere la confidenza, si utilizza il campo *Confidence* dell'oggetto *Block*.

## 5.3 Test di Amazon Textract

In questa sezione, verrà presentata una prova effettuata con il servizio Amazon Textract, con lo scopo di estrarre testi da immagini.

### 5.3.1 Riconoscimento ed estrazione di testi da immagini

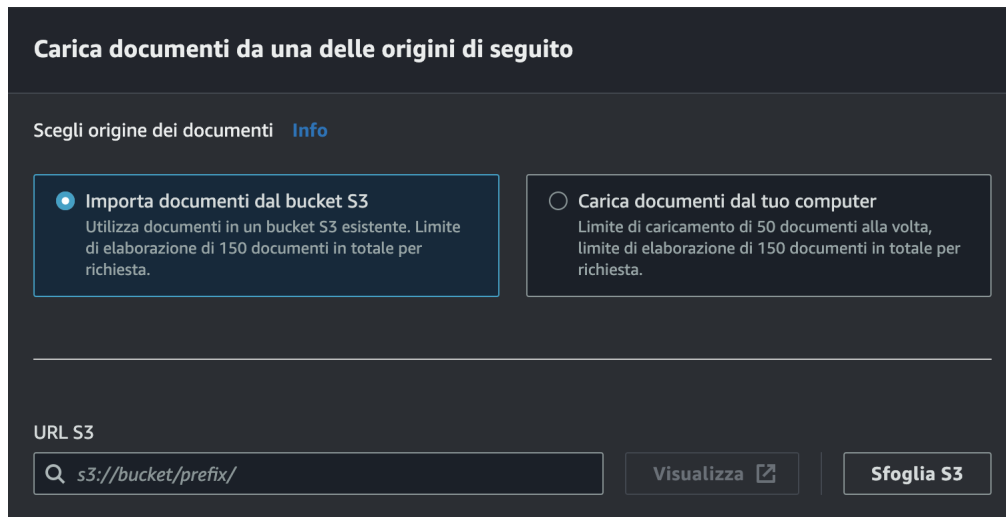
Il rilevamento e l'estrazione di informazioni da immagini può avvenire in due modi. Si può utilizzare l'interfaccia che Amazon Textract mette a disposizione, tramite la console di gestione, oppure si possono usare degli script Python che utilizzano delle operazioni API. Tra queste abbiamo:

- *AnalyzeDocument*; questa API consente di analizzare il contenuto di documenti, rilevando ed estraendo informazioni, quali testo, form e tabelle. Essa può essere utilizzata per l'analisi dettagliata di vari documenti, compresi report finanziari, cartelle cliniche e moduli fiscali.

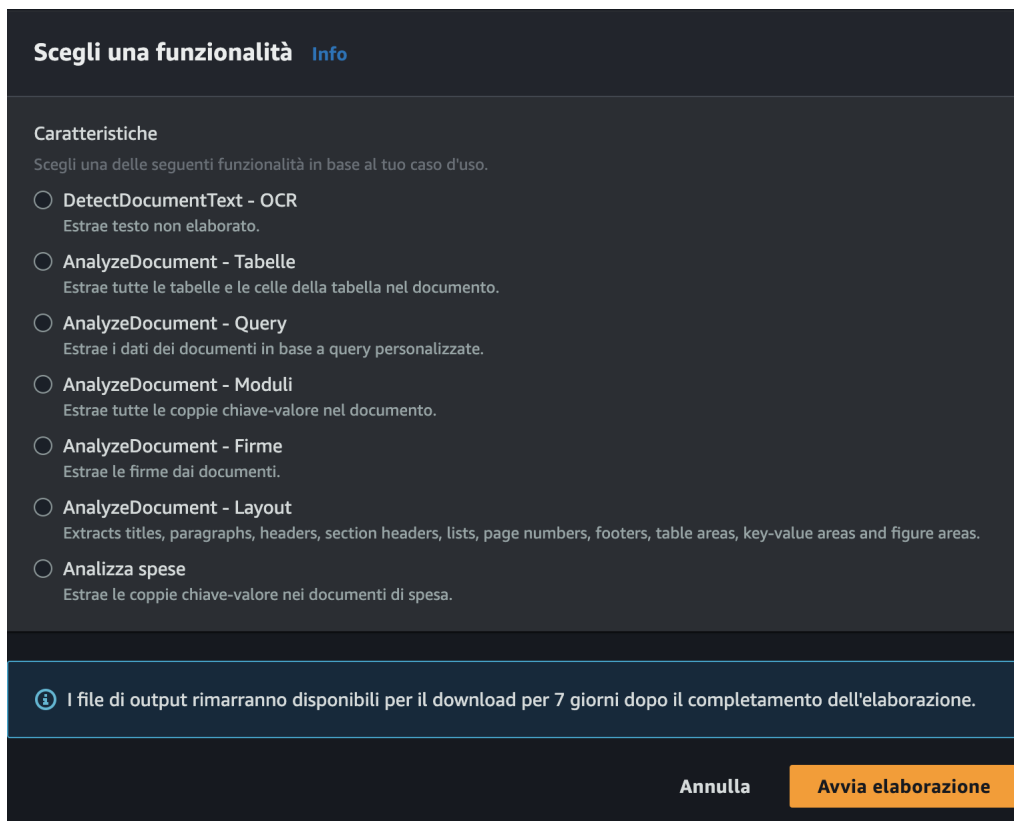
- *AnalyzeExpense*; questa API è progettata specificamente per l'analisi di fatture e ricevute, facilitando l'estrazione di informazioni relative alle spese. È utile per processare automaticamente documenti di spesa e recuperare dettagli come importi, date e informazioni sulle transazioni.
- *DetectDocumentText*; questa API consente di rilevare il testo in documenti o immagini, restituendo i risultati del rilevamento effettuato. Può essere utilizzata per identificare la presenza di testo in documenti o immagini e ottenere informazioni sulla sua posizione.
- *AnalyzeID*; questa API è focalizzata sull'analisi di documenti di identità, come patenti di guida e passaporti emessi dal governo degli Stati Uniti. Può essere impiegata per estrarre informazioni pertinenti da documenti di identità, agevolando processi di verifica e registrazione.
- *GetDocumentAnalysis*; questa API restituisce i risultati dell'analisi di un documento specifico effettuato in precedenza con l'API di avvio corrispondente. È utile per recuperare i risultati completi di un'analisi di documento precedentemente avviata.
- *GetDocumentTextDetection*; questa API restituisce i risultati del rilevamento del testo per un documento specifico, in base a un'analisi avviata precedentemente. È utile per ottenere informazioni dettagliate sul testo individuato all'interno di un documento.
- *GetExpenseAnalysis*; questa API restituisce i risultati dell'analisi delle spese per un documento specifico, basato su un'analisi precedentemente avviata. È utile per recuperare informazioni dettagliate relative alle spese da documenti precedentemente analizzati.
- *StartDocumentAnalysis*, *StartDocumentTextDetection*, *StartExpenseAnalysis*; queste API avviano il processo di analisi per documenti, rilevamento del testo o analisi delle spese rispettivamente. Sono utilizzate per iniziare l'analisi di documenti specifici, estraendo informazioni pertinenti in base alle esigenze dell'utente.

Se si opta per il primo metodo, la prima azione da compiere è caricare il documento che si desidera analizzare. L'interfaccia di Amazon Textract permette di caricare documenti con meno di 11 pagine, di dimensione inferiore a 5 MB e in uno dei formati supportati, ovvero JPEG, PNG o PDF. Inoltre, i documenti possono essere caricati in blocco; il sistema permette di inviarne fino a 150 in una singola richiesta e di scaricare i risultati nel formato CSV.

Dopo aver premuto il pulsante "Carica documenti", la console presenterà una schermata che offre la possibilità di selezionare il luogo da cui recuperare i file. Ciò può avvenire da un bucket S3 o direttamente dal proprio computer (Figura 5.2). Successivamente, si avrà la possibilità di scegliere tra diverse funzionalità, optando per quella più adatta al proprio caso d'uso specifico (Figura 5.3). Questo processo offre una flessibilità significativa, consentendo agli utenti di adattare l'operazione di caricamento alle proprie esigenze e preferenze.



**Figura 5.2:** Schermata per il caricamento di un documento di Amazon Textract



**Figura 5.3:** Schermata per la scelta della funzionalità

Una volta caricato il documento, e scelta la funzionalità, verrà mostrata un'interfaccia (Figura 5.4), che permette di visualizzare:

- *Text Raw (Testo Grezzo)*; mostra il testo estratto dal documento senza alcuna formattazione o struttura. Questa visualizzazione è utile quando si desidera accedere al testo nella sua forma più basilare.

- *Layout*; presenta il testo estratto mantenendo la formattazione e la struttura originale del layout del documento. Include informazioni sulla posizione del testo, mantenendo l'aspetto visivo del documento originale.
- *Moduli*; se il documento contiene moduli (form) o formulari, questa opzione consente di visualizzare in modo specifico le informazioni estratte dai moduli, facilitando la comprensione delle risposte o dei dati associati.
- *Tabelle*; mostra i dati estratti organizzati in forma tabellare. Questa visualizzazione è particolarmente utile quando si desidera analizzare o esportare dati tabellari da documenti.
- *Firme*; Amazon Textract è in grado di riconoscere, se presente, e visualizzare le firme presenti nei documenti analizzati.

The screenshot shows the Amazon Textract interface with the following elements:

- Navigation tabs: **Testo in formato raw**, **Layout**, **Moduli**, **Tabelle**, **Query**, **Firme**.
- Section: **Risultati**
- Search bar:  Segmento per ...
- Document ID: **CO. FILE DEPT. CLOCK NUMBER** ABC 126543 123456 12345 00000000 1
- Document Type: **Earnings Statement** ANY COMPANY CORP. Period ending: 7/18/2008
- Address: **475 ANY AVENUE** Pay date: 7/25/2008 ANYTOWN USA 10101
- Identification: **Social Security Number: 987-65-4321** Taxable Marital Status: Married
- Employee Name: **JOHN STILES** Exemptions/Allowances: 101 MAIN STREET
- Tax Information: **Federal: 3. \$25 Additional Tax** ANYTOWN, USA 12345 State: 2 Local: 2
- Table Headers: **Earnings** rate hours this period year to date Other Benefits and
- Table Data:
 

Earnings	rate	hours	this period	year to date	Other Benefits and		
Regular	10.00	32.00	320.00	16,640.00	Information this period		
total to date		Overtime	15.00	1.00	15.00	780.00	Group Term Life

**Figura 5.4:** Interfaccia di analisi ed estrazione di un documento

Dopo aver effettuato l'analisi del documento, è possibile scaricare i dati estratti in vari formati. Se si desidera ottenere il testo estratto senza alcuna formattazione o struttura, è possibile scaricarlo come testo grezzo, come file di tipo *.txt*. Se l'analisi include l'estrazione di dati da tabelle, è possibile scaricare i risultati in un formato tabellare, come CSV (Comma-Separated Values) o Excel, per una più facile elaborazione e analisi dei dati. Insieme ai file con le informazioni presenti all'interno del documento, viene scaricato anche un file di tipo JSON chiamato *analyzeDocResponse.json* contenente la risposta generata come risultato dell'operazione di analisi del documento.



Una seconda prova, realizzata con Amazon Textract, è stata effettuata tramite l'uso di alcuni script Python, utilizzando le API messe a disposizione; il test è stato effettuato sulla piattaforma Google Colab.

Come immagine da analizzare, è stata utilizzata quella riportata in Figura 5.5. Lo scopo della prova consiste nel rilevare il testo grezzo, riconoscere la presenza di una firma ed evidenziare specifici campi all'interno della foto.

**Absence Request**

**Absence Information**

Employee Name: Jerry Smith

Date: 05/28/2015 Department: Logistics

Manager: Madeleine Wallis

Type of Absence Requested:

Sick  Vacation  Bereavement  Time Off Without Pay

Other  Jury Duty  Maternity/Paternity  Floating Holiday

Dates of Absence:

From: 06/08/2015 To: 06/12/2015

Number of Hours Requesting 40 Number of Available Hours 73

Reason for Absence:

You must submit requests for absences, other than sick leave, two weeks prior to the first day you will be absent.

Jerry Smith 05/28/2015

Employee Signature Date

Figura 5.5: Immagine utilizzata per il test di Amazon Textract

Per prima cosa, per utilizzare le API, è necessario installare la libreria *boto3* (Figura 5.6), ovvero la libreria ufficiale di Amazon Web Services per Python. Essa fornisce un'interfaccia di programmazione per interagire con vari servizi AWS, inclusi quelli per il Machine Learning, come Amazon Textract.

```

pip install boto3
Collecting boto3
  Downloading boto3-1.33.2-py3-none-any.whl (139 kB)
    139.1/139.1 kB 3.0 MB/s eta 0:00:00
Collecting botocore<1.34.0,>=1.33.2 (from boto3)
  Downloading botocore-1.33.2-py3-none-any.whl (11.7 MB)
    11.7/11.7 MB 36.8 MB/s eta 0:00:00
Collecting jmespath<2.0.0,>=0.7.1 (from boto3)
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Collecting s3transfer<0.9.0,>=0.8.0 (from boto3)
  Downloading s3transfer-0.8.1-py3-none-any.whl (82 kB)
    82.0/82.0 kB 8.6 MB/s eta 0:00:00
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in /usr/local/lib/python3.10/dist-packages (from botocore<1.34.0,>=1.33.2->boto3) (2.8.2)
Requirement already satisfied: urllib3<2.1,>=1.25.4 in /usr/local/lib/python3.10/dist-packages (from botocore<1.34.0,>=1.33.2->boto3) (2.0.7)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil<3.0.0,>=2.1->botocore<1.34.0,>=1.33.2->boto3)
Installing collected packages: jmespath, botocore, s3transfer, boto3
Successfully installed boto3-1.33.2 botocore-1.33.2 jmespath-1.0.1 s3transfer-0.8.1

```

Figura 5.6: Comando per l'installazione della libreria *boto3*

Successivamente, è necessario configurare il client di *boto3* per Amazon Textract utilizzando le chiavi di accesso, ovvero access key ID e secret access key. Poi, tramite l'utilizzo della libreria "PIL" (Python Imaging Library), è possibile aprire e visualizzare l'immagine (Figura 5.7).

```
import boto3
client = boto3.client('textract', region_name='eu-central-1', aws_access_key_id='AKIAW0XDAIG5W6WV4YE7',
                      aws_secret_access_key='hpJGJDsumt0cQUd0GJVpAPpG1eWQTxXu+5FT3Neh')

from PIL import Image, ImageDraw
image=Image.open("/content/textract.jpg")
image
```

**Figura 5.7:** Codice per l'apertura e la visualizzazione dell'immagine

Il secondo passo è quello di leggere l'immagine e di convertirla in un array di byte. Attraverso la funzione, *detect\_document\_text* l'array viene inviato al client di Amazon Textract. All'interno della variabile *response* sono contenute le informazioni sul testo rilevato nell'immagine (Figura 5.8).

```
with open('/content/textract.jpg', 'rb') as image:
    img = bytearray(image.read())

response = client.detect_document_text(
    Document={'Bytes': img}
)
response
text = ""
for item in response["Blocks"]:
    if item["BlockType"] == "LINE":
        print (item["Text"])
        text = text + " "+item["Text"]
```

**Figura 5.8:** Codice per la lettura dell'immagine e il rilevamento del testo

Successivamente, tramite la variabile *text*, si itera attraverso i blocchi nella risposta di Textract. Se il blocco è di tipo "LINE" (linea di testo), viene stampato il testo del blocco e lo stesso viene concatenato alla variabile *text* inserendo uno spazio tra i testi, così da ottenere il testo grezzo presente nel documento (Figura 5.9).

```
Absence Request
Absence Information
Employee Name: Jerry Smith
Date:
05/28/2015
Department Logistics
Manager Madeleine Wallis
Type of Absence Requested:
Sick
Vacation
Bereavement
Time Off Without Pay
Other
Jury Duty
Maternity/Paternity
Floating Holiday
Dates of Absence
From:
06/08/2015
To:
06/12/2015
Number of Hours Requesting 40
Number of Available Hours 73
Reason for Absence
You must submit requests for absences other than sick leave, two weeks prior to the first day you will be absent
leny smith
05/28/2015
Employee Signature
Date
```

**Figura 5.9:** Testo grezzo rilevato ed estratto dall'immagine

Il passo successivo è quello di analizzare il documento ed evidenziare specifici campi all'interno dell'immagine. In questo caso, come mostrato in Figura 5.10, si è scelto di rilevare la firma. Per prima cosa viene effettuata una chiamata al servizio Textract tramite il metodo `analyze_document`, passando l'immagine come un array di byte e specificando che si desidera individuare le firme. Poi si itera attraverso i blocchi restituiti e, per ogni blocco di tipo "SIGNATURES", essi vengono assegnati alla variabile `signature_block` per poi stampare la risposta.

```
response = client.analyze_document(Document={'Bytes': img}, FeatureTypes=["SIGNATURES"])

for item in response["Blocks"]:
    if item["BlockType"] == "SIGNATURE":
        signature_block = item
signature_block

{'BlockType': 'SIGNATURE',
 'Confidence': 78.14910125732422,
 'Geometry': {'BoundingBox': {'Width': 0.24169619381427765,
 'Height': 0.10771280527114868,
 'Left': 0.2556406557559967,
 'Top': 0.7984729409217834},
 'Polygon': [{'X': 0.2556406557559967, 'Y': 0.7984729409217834},
 {'X': 0.4973083436489105, 'Y': 0.7984756231307983},
 {'X': 0.49733686447143555, 'Y': 0.9061857461929321},
 {'X': 0.2556583285331726, 'Y': 0.9061793684959412}]},
 'Id': '037c3f1f-991d-4967-81ba-ef1c5baa0233'}
```

Figura 5.10: Codice per il rilevamento della firma

Infine, vengono effettuate alcune azioni sull'immagine. Per prima cosa, attraverso la funzione `ShowBoundingBox`, viene disegnato un rettangolo delimitatore (bounding box) intorno alla firma individuata nel documento; lo stesso processo viene effettuato per altri campi all'interno dell'immagine (Figura 5.11).

```
def ShowBoundingBox(draw,box,width,height,boxColor):
    left = width * box['Left']
    top = height * box['Top']
    draw.rectangle([left,top, left + (width * box['Width']), top +(height * box['Height'])],outline=boxColor)

image=Image.open("/content/textract.jpg")

width, height =image.size
draw=ImageDraw.Draw(image)

ShowBoundingBox(draw, signature_block['Geometry']['BoundingBox'],width,height, 'Red')
image

ShowBoundingBox(draw, name['Geometry']['BoundingBox'],width,height, 'Green')
image

ShowBoundingBox(draw, lines_block['Geometry']['BoundingBox'],width,height, 'Blue')
image
```

Figura 5.11: Codice per disegnare dei rettangoli delimitatori all'interno dell'immagine

Il risultato finale è mostrato in Figura 5.12.

**Absence Request**

**Absence Information**

Employee Name: Jerry Smith

Date: 05/28/2015 Department: Logistics

Manager: Madeleine Wallis

Type of Absence Requested:

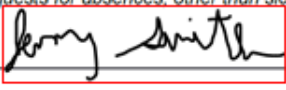
Sick  Vacation  Bereavement  Time Off Without Pay  
 Other  Jury Duty  Maternity/Paternity  Floating Holiday

Dates of Absence:  
From: 06/08/2015 To: 06/12/2015

Number of Hours Requesting 40 Number of Available Hours 73

Reason for Absence:

*You must submit requests for absences, other than sick leave, two weeks prior to the first day you will be absent.*

 05/28/2015

Employee Signature

**Figura 5.12:** Risultato dell'immagine modificata tramite Amazon Textract

---

## Discussione in merito alle esperienze condotte

---

*In quest'ultimo capitolo verranno presentate delle riflessioni sull'utilizzo dei potenti mezzi, precedentemente trattati, messi a disposizione dall'Intelligenza Artificiale. In particolare, trarremo delle considerazioni finali riguardanti l'utilizzo dei chatbot, il riconoscimento delle immagini, il rilevamento e l'estrazione di testo da esse.*

### 6.1 Discussione sull'utilizzo dei chatbot

L'integrazione sempre più diffusa dei chatbot rappresenta una rivoluzione significativa nell'interazione tra utenti e sistemi digitali. L'impiego di questi assistenti virtuali ha non solo semplificato l'accesso alle informazioni, ma ha anche ridefinito la natura stessa della comunicazione online. La capacità dei chatbot di comprendere e rispondere alle domande degli utenti in modo contestuale ha introdotto un nuovo livello di personalizzazione e tempestività nell'erogazione dei servizi.

In maniera più o meno complessa, essi possono sopperire a quelle richieste, quelle fasce orarie o quelle operazioni ripetitive in cui impegnare una persona reale sarebbe impossibile per vari motivi. Perciò, nel contesto delle applicazioni aziendali, i chatbot si sono dimostrati strumenti efficaci per migliorare l'efficienza operativa.

Il chatbot più utilizzato e popolare è ChatGPT, creato dalla start-up OpenAI. In particolare, esso è un Large Language Model (LLM), ovvero una rete neurale con 175 miliardi di parametri, che sono stati tarati con un immenso lavoro di addestramento, facendo "leggere" alla macchina l'intero contenuto di Internet, e per un gran numero di volte.

Il funzionamento della macchina è statistico. Le frasi vengono costruite con metodi probabilistici basandosi sul contesto fornito dalle parole adiacenti. Quello che rende la macchina così potente sono i due elementi cardine di tutti gli sviluppi dell'IA nell'ultimo decennio, ovvero la disponibilità di un immenso potere computazionale e l'accessibilità di una straordinaria quantità di dati, che consentono ai metodi statistici un sorprendente livello di successo. Tuttavia, 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.

Una prima riflessione riguarda l'utilizzo effettivo dei chatbot. Esso può essere assimilato a una forma di plagio che può suscitare, all'utente che usufruisce del servizio, preoccupazioni in merito alla proprietà intellettuale. Allo stesso modo, i proprietari di chatbot ricevono accuse di violazione di proprietà intellettuale da coloro i cui testi, documenti, foto o altro, vengono utilizzati proprio per "alimentare" gli algoritmi dei chatbot.

In secondo luogo, come già menzionato, un chatbot è una macchina; quindi, non possiede alcun tipo di ragionamento critico, né criterio morale nella fruizione delle risposte. Queste si basano esclusivamente sul contenuto appreso nel tempo; di conseguenza, la valutazione della loro correttezza è compito dell'utente. Pertanto, questi strumenti non possono essere considerati come oracoli. È essenziale sottolineare che i chatbot non devono essere presi come sostituti degli umani e che la loro utilità risiede nella capacità di essere dispositivi ausiliari. Nonostante la loro crescente complessità, i chatbot non possono sostituire completamente le attività umane; piuttosto, vanno utilizzati in modo saggio per integrare e agevolarne il lavoro. Inoltre, si può riflettere sul fatto che è difficile sfuggire all'impressione di avere a che fare con un interlocutore umano. È cruciale, quindi, ricordarsi che dall'altra parte non c'è nessuno. Può sembrare una considerazione scontata, ma riflettere sull'effetto psicologico derivante dalla costante interazione con una macchina capace di produrre consistenti quantità di informazioni in un linguaggio impeccabile e autorevole è di fondamentale importanza. È essenziale ricordare frequentemente che non stiamo dialogando con un esperto infallibile o un influente leader di opinione; né tanto meno stiamo interagendo con un semplice individuo. Stiamo, piuttosto, comunicando con una macchina addestrata attraverso l'analisi statistica, in grado di muoversi nell'ampio contesto delle conversazioni umane.

## 6.2 Discussione sul riconoscimento delle immagini

Il contesto del riconoscimento delle immagini mediante l'utilizzo dell'Intelligenza Artificiale si distingue come un capitolo affascinante e dinamico nel panorama tecnologico contemporaneo. La portata dell'innovazione generata da questa tecnologia si manifesta in una vasta gamma di applicazioni pratiche, dall'evoluzione dell'assistenza sanitaria alla rafforzata sicurezza attraverso soluzioni avanzate, dalla prospettiva di veicoli autonomi fino alle sofisticate ottimizzazioni dell'e-commerce.

In questa tesi, è stato possibile esplorare il funzionamento delle Intelligenze Artificiali, focalizzandosi sui sistemi di riconoscimento delle immagini, nello specifico Amazon Rekognition. La sua vera potenza emerge con il servizio Custom Labels. Esso, fornisce un'interfaccia utente dedicata che consente l'addestramento della rete neurale su un dataset personalizzato. Ciò apre la strada a chiunque desideri costruire e utilizzare applicazioni basate sull'IA per il riconoscimento delle immagini, ampliando notevolmente le possibilità di utilizzo di questa tecnologia. Tuttavia, accanto a queste opportunità, emergono anche sfide cruciali, con l'etica e la privacy che emergono come tematiche centrali. La necessità di bilanciare il rapido progresso tecnologico con le preoccupazioni etiche si presenta come un nodo cruciale per una diffusa accettazione di tali tecnologie.

In particolare, la raccolta massiccia di dati necessaria per l'addestramento dei modelli potrebbe coinvolgere informazioni sensibili, aumentando il rischio di accessi non autorizzati e di un uso improprio. Un'altra preoccupazione fondamentale riguarda la potenziale identificazione delle persone, violando la loro privacy, specialmente quando non sono consapevoli del monitoraggio.

Inoltre, l'implementazione diffusa di tali sistemi può comportare una sorveglianza involontaria nelle aree pubbliche, sollevando interrogativi sul bilanciamento tra sicurezza e preservazione della privacy. La presenza di bias nei dati utilizzati per l'addestramento è un ulteriore problema, poiché può portare a discriminazioni ingiuste. La mancanza di consenso da parte delle persone fotografate o registrate, specialmente in contesti pubblici, può generare controversie etiche e legali.

La questione dell'accesso e del controllo sulle informazioni personali, insieme alla sicurezza dei dati, è cruciale per garantire che le persone abbiano una gestione adeguata delle

proprie informazioni. Affrontare tali problematiche richiede misure di sicurezza robuste, consenso informato e politiche di gestione dei dati che rispettino la privacy degli individui.

In conclusione, l'utilizzo del riconoscimento delle immagini, se da un lato offre vantaggi significativi in termini di innovazione ed efficienza, dall'altro solleva sfide importanti, in particolare riguardo alla privacy; pertanto, sta a noi utenti farne un corretto utilizzo.

### **6.3 Discussione sul rilevamento ed estrazione di testi da immagini**

Il rilevamento e l'estrazione di testi da immagini rappresenta un ambito tecnologico che ha profondamente influenzato la nostra interazione con il mondo visivo. Questa capacità dell'Intelligenza Artificiale di interpretare e comprendere il testo presente in un'immagine offre una serie di vantaggi significativi in diversi contesti.

Per prima cosa, questa tecnologia si è dimostrata cruciale per la digitalizzazione dei documenti. La possibilità di estrarre testo da immagini consente di convertire dati cartacei in formati digitali, semplificando la gestione e la condivisione di informazioni. In ambito aziendale, questo si traduce in una maggiore efficienza operativa e nella facilità di accesso ai contenuti.

Un altro aspetto fondamentale riguarda l'accessibilità. Il rilevamento e l'estrazione di testi da immagini ha permesso a persone con disabilità visive di accedere a contenuti visivi, aprendo nuove prospettive di inclusività. Questa tecnologia contribuisce ad abbattere barriere e a garantire un accesso più equo alle informazioni per tutti. Tuttavia, questo progresso innovativo non è privo di sfide.

Nel contesto del riconoscimento del testo emergono diverse problematiche e questioni rilevanti. Una delle principali consiste nell'assicurare una precisione costante del riconoscimento, considerando la varietà di font, stili di scrittura e condizioni di illuminazione. Ciò, avviene poiché i dati presenti in documenti cartacei, immagini digitali e schermate di computer, sono eterogenei; quindi, è necessario adattare un approccio flessibile e versatile. Inoltre, anche la qualità variabile delle immagini costituisce una sfida aggiuntiva, con immagini sfocate, scarsa illuminazione e rumore visivo che possono impattare sulla precisione del riconoscimento.

In secondo luogo, l'ampia diversità di lingue e scritture rappresenta un altro aspetto complesso, influenzando la capacità dei sistemi di riconoscimento del testo. Gli errori di riconoscimento sono inevitabili, e garantire meccanismi di correzione affidabili è essenziale per migliorare l'affidabilità complessiva del sistema.

Dal punto di vista etico, la gestione della privacy è un tema cruciale, specialmente quando il riconoscimento del testo viene applicato a documenti sensibili, come documenti di identità o licenze di guida.

Infine, la presenza di bias nei dati di addestramento può generare risultati discriminatori o inesatti, sottolineando l'importanza di un'analisi critica e di misure correttive.

In conclusione, il rilevamento e l'estrazione di testi da immagini apre scenari promettenti per la trasformazione digitale, migliorando l'accessibilità e la gestione delle informazioni. Tuttavia, la comunità tecnologica deve continuare a concentrarsi su miglioramenti nella precisione e sull'affrontare le sfide etiche per massimizzare i benefici di questa innovazione.

Questa tesi ha offerto un quadro generale riguardante il rivoluzionario mondo dell'Intelligenza Artificiale e dei suoi principali utilizzi.

Inizialmente, sono state fornite alcune definizioni per chiarire il concetto di Intelligenza Artificiale, per poi procedere al racconto di alcuni aneddoti storici, partendo dalla nascita fino allo sviluppo attuale. Successivamente, sono stati esaminati i vantaggi e i pregi derivanti da questa nuova tecnologia, insieme ai suoi limiti, associati ad un utilizzo improprio.

In seguito, è stato presentato il concetto di Cloud Computing. Sono stati forniti definizioni e approfondimenti sul suo utilizzo, esaminandone pregi e difetti. Dopo di ciò, è stata introdotta una piattaforma che offre tali servizi, ovvero Amazon Web Services (AWS). Sono stati esplorati alcuni cenni storici relativi alla sua fondazione e al suo ingresso nel mondo dell'Intelligenza Artificiale, che hanno consentito ad essa di affermarsi come uno dei principali fornitori di servizi Cloud. Infine, sono stati dettagliati i servizi offerti da questa piattaforma e sono stati analizzati i principali scenari applicativi.

A questo punto, è seguito un approfondimento su tre branche dell'Intelligenza Artificiale, ovvero i chatbot, il riconoscimento di immagini e il rilevamento ed estrazione di testi. Per ciascuno di questi argomenti, sono state fornite le definizioni e sono stati descritti i principali meccanismi di funzionamento, per poi passare, per ognuno, ad un caso di studio condotto attraverso l'utilizzo dei servizi di AWS. Nella creazione di un chatbot è stato impiegato il servizio Amazon Lex, per il riconoscimento delle immagini è stato utilizzato Amazon Rekognition, e infine, per l'estrazione di testo da immagini è stato utilizzato Amazon Textract.

In conclusione, sono state presentate una serie di riflessioni per ciascun argomento trattato, con l'obiettivo di analizzare in modo critico sia i vantaggi derivanti dall'utilizzo di questi potenti servizi, che le problematiche connesse, come quelle legate alla privacy o all'etica.

Questa tesi, quindi, ha esplorato le dinamiche emergenti dell'Intelligenza Artificiale nel contesto del Cloud Computing, con un focus su Amazon Web Services. Mentre le applicazioni pratiche, come i chatbot e il riconoscimento di immagini, stanno ridefinendo il modo in cui interagiamo con la tecnologia, è cruciale affrontare le sfide etiche e di sicurezza connesse. Per il futuro, ci si aspetta un aumento significativo nell'adozione dell'Intelligenza Artificiale, con l'urgente necessità di sviluppare normative e pratiche etiche per garantire un progresso responsabile e sostenibile.



- ADAMOPOULOU, E. e MOUSSIADES, L. (2020), «An overview of chatbot technology», in «IFIP international conference on artificial intelligence applications and innovations», p. 373–383, Springer.
- CHEN, X., JIN, L., ZHU, Y., LUO, C. e WANG, T. (2021), «Text recognition in the wild: A survey», *ACM Computing Surveys (CSUR)*, vol. 54 (2), p. 1–35.
- EGMONT-PETERSEN, M., DE RIDDER, D. e HANDELS, H. (2002), «Image processing with neural networks—a review», *Pattern recognition*, vol. 35 (10), p. 2279–2301.
- LU, D. e WENG, Q. (2007), «A survey of image classification methods and techniques for improving classification performance», *International journal of Remote sensing*, vol. 28 (5), p. 823–870.
- MAR, W. e THAW, Y. (2019), «An analysis of benefits and risks of artificial intelligence», *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. 3 (5), p. 2456–6470.
- MARTINEZ, R. (2019), «Artificial intelligence: Distinguishing between types & definitions», *Nevada Law Journal*, vol. 19 (3), p. 9.
- ROLI, A. e BRACCINI, M. (2013), «Intelligenza artificiale: test di Turing e alcune questioni filosofiche», .
- RUSSELL, S. J. e NORVIG, P. (2010), *Artificial intelligence a modern approach*, London.
- SHADIEV, R., HWANG, W.-Y., CHEN, N.-S. e HUANG, Y.-M. (2014), «Review of speech-to-text recognition technology for enhancing learning», *Journal of Educational Technology & Society*, vol. 17 (4), p. 65–84.
- SOMALVICO, M. e OTHERS (1987), *Intelligenza artificiale*, Scienza & vita nuova. (Cited at page 3)
- STRONG, A. (2016), «Applications of artificial intelligence & associated technologies», *Science [ETEBMS-2016]*, vol. 5 (6).
- TURING, A. M. (1950), *Computing machinery and intelligence*, Mind. (Cited at page 7)
- VELTE, A. T. V. T. J. e ELSENPETER, P. D. R. (2010), *Cloud computing*.

## Websites consulted

- Servizi di cloud computing - Amazon Web Services (AWS) - <https://aws.amazon.com>
- Google Cloud - <https://cloud.google.com>
- Enciclopedia Treccani - [www.treccani.it](http://www.treccani.it)
- LegaNerd - <https://leganerd.com>
- Namirial - <https://focus.namirial.it>
- Blog Osservatori - <https://blog.osservatori.net>
- Oracle - [www.oracle.com](http://www.oracle.com)
- Wikipedia - [www.wikipedia.org](http://www.wikipedia.org)

---

## Ringraziamenti

---

In conclusione, vorrei ringraziare di cuore tutti coloro che hanno reso possibile la realizzazione di questo percorso universitario.

Il primo ringraziamento va a tutta la mia famiglia, in particolare ai miei genitori, mia mamma Monia e mio padre Gianluca, che mi hanno dato la possibilità di intraprendere e portare a termine questa esperienza formativa, sostenendomi nei momenti di difficoltà.

Un altro ringraziamento va a mio fratello Jacopo, che è sempre stato pronto ad aiutarmi quando ne avevo bisogno.

Ringrazio di cuore i miei compagni di corso e coinquilini, Davide e Luca, con i quali ho condiviso ore di lezione e studio, superando esami e progetti impegnativi.

Vorrei ringraziare il Prof. Ursino, che mi ha accompagnato nel lavoro di tesi seguendomi costantemente con estrema disponibilità.

Un ulteriore ringraziamento va al box CrossFit Fermo, poiché, grazie a loro, sono riuscito a completare il percorso universitario in forma, permettendomi di distogliere la mente dalle sessioni di esame.

Infine, vorrei ringraziare me stesso, per la determinazione, le sfide affrontate e il costante impegno, sperando che questo non sia un punto di arrivo, ma piuttosto un nuovo e stimolante punto di partenza.