

Dati: la salsa segreta che trascina la traduzione alimentare con l'IA

Il Dott. Roberto Mariani spiega perché i dati di allenamento sono la salsa segreta per cucinare la traduzione automatica perfetta. Scopri le 8 migliori pratiche di cui Menulance è stata pioniera per gestire Genesis, un set di dati di allenamento che raggiunge una precisione di traduzione del 97%.



CO-REDATTO DAL DOTT. ROBERTO MARIANI E ANTHONY COUNDOURIS 
TRADOTTO DAL DOTT. FRANK BADRINES

DUE PAROLE DAL CTO

Abbiamo assemblato Genesis, un set multilingue di dati composto da 166 milioni di righe di piatti e ingredienti. Fa parte di un progetto durato un decennio. I motori di traduzione basati su Genesis raggiungono un impressionante punteggio BLEU dell'83% per la traduzione dall'inglese allo spagnolo. Per fornire contesto, un traduttore umano esperto ottiene in genere un punteggio di circa il 60% sulla scala BLEU.

Il successo è attribuito al ristretto set di dati di allenamento di Genesis, alla metodologia di allenamento basata su frasi brevi e alla frammentazione completa delle frasi.

Al contrario, strumenti tradizionali come Google Translate e Bing hanno difficoltà nelle traduzioni alimentari a causa dell'allenamento su dizionari generalisti, con conseguenti traduzioni imprecise.

Rappresentiamo una svolta per gli appassionati di cucina e per la comunità dell'IA. Il fulcro dell'efficacia dell'IA non risiede solo nella competenza tecnologica, ma anche nella meticolosa cura di set di dati puliti e pertinenti.

Traduciamo il cibo del mondo.

A handwritten signature in black ink, appearing to read "Roberto". The signature is fluid and stylized, with a large initial 'R' and a long, sweeping underline.

SOMMARIO

L'intelligenza artificiale (IA) sta cambiando il modo in cui consumiamo il cibo. L'intelligenza artificiale agisce come un'estensione contemporanea dei media tradizionali, come i menù dei ristoranti o i libri di cucina, integrando perfettamente diverse fonti come testo, immagini, mappe e dati video. Questa integrazione migliora l'accessibilità delle informazioni, rendendo la comprensione del cibo più completa.

I sistemi di traduzione basati sull'intelligenza artificiale aiutano i viaggiatori a comunicare con la gente del posto, a comprendere i menù e a partecipare a corsi di cucina, eliminando le barriere linguistiche. Questo favorisce il dialogo interculturale e arricchisce l'esperienza complessiva del turismo gastronomico.

Le tecnologie digitali e l'IA offrono ai turisti esperienze personalizzate e immersive, come tour gastronomici interattivi, lezioni di cucina virtuale e narrazioni in realtà aumentata. Questi progressi aumentano il coinvolgimento e creano ricordi duraturi per i turisti.¹

L'IA svolge un ruolo cruciale nella documentazione e nella digitalizzazione di ricette tradizionali, tecniche di cucina e pratiche culturali, preservando e salvaguardando così il patrimonio culinario, in particolare nelle regioni in cui le conoscenze tradizionali rischiano di andare perdute.

Inoltre, l'IA sta rivoluzionando la gestione dei ristoranti e la condivisione di suggerimenti di ricette. L'IA e le innovazioni digitali possono semplificare le operazioni dei ristoranti, inclusi il controllo dell'inventario e i sistemi di prenotazione, riducendo i tempi di attesa e migliorando l'esperienza culinaria sia per i clienti che per i ristoratori.

Le piattaforme basate sull'IA possono analizzare gli ingredienti regionali, le usanze culinarie e le tendenze alimentari globali per offrire suggerimenti migliori su cibi e ricette. Questo ispira i viaggiatori a esplorare nuovi sapori e tecniche culinarie.

Tuttavia, in mezzo all'entusiasmo per il potenziale dell'IA, persiste una sfida: la tendenza a promettere troppo e a offrire meno del previsto. Questo deriva da uno squilibrio fondamentale, in cui l'attenzione alla creazione di tecnologie all'avanguardia spesso trascura un aspetto cruciale: l'importanza di curare set di dati puliti su cui vengono allenati i motori di IA.

In precedenza, lavoravo con un gruppo di scienziati presso il Laboratorio Elettrotecnico Giapponese per creare un robot mobile in grado di comunicare con il parlato, chiamato Jijo-2. Il progetto Jijo-2 è iniziato nel 1997 ed è stato frutto di una collaborazione tra diversi scienziati e ingegneri, focalizzata sull'integrazione di diverse tecnologie, come il riconoscimento vocale, la navigazione e il riconoscimento facciale, per consentire un'interazione naturale negli ambienti d'ufficio.

Il mio compito era risolvere il problema del riconoscimento facciale. Abbiamo provato tutti i motori di intelligenza artificiale conosciuti all'epoca per aiutare Jijo-2 a riconoscere e associare un volto a un nome. Con ogni motore provato, i risultati erano gli stessi: tutti i motori allenati sul nostro set di dati raggiungevano una precisione media del 44%.

Solo quando abbiamo iniziato a concentrarci sui dati e non sul motore, abbiamo visto un miglioramento della precisione. All'epoca, si trattava di una svolta nel pensiero. Gli scienziati erano ossessionati dai motori di intelligenza artificiale. Jijo-2 ha dimostrato che i set di dati curati erano molto più importanti.

Questo libro bianco esplora l'importanza di dati di alta qualità. Usiamo il set di dati di allenamento di Menulance, Genesis, per mostrare come la qualità dei dati possa determinare il successo o il fallimento delle applicazioni di IA.

INTRODUZIONE

A metà del XIX secolo, durante la Rivoluzione Industriale, gli agricoltori britannici iniziarono a utilizzare macchine a vapore come trebbiatrici e aratri per coltivare più cibo per una popolazione in rapida crescita. Queste macchine aumentarono la produttività e portarono nuove speranze all'agricoltura.

Verso la fine del 1800, le attrezzature agricole erano diventate molto avanzate, ma gli agricoltori si trovarono di fronte a un limite. Le macchine non potevano più migliorare le rese.

Così, gli agricoltori spostarono l'attenzione dalle macchine al terreno. In regioni come lo Yorkshire e l'East Anglia, iniziarono a sperimentare la rotazione delle colture, la fertilizzazione, la lotta ai parassiti e la selezione genetica. Migliorarono anche l'irrigazione e il drenaggio per creare condizioni di crescita migliori.

Innovatori come Jethro Tull contribuirono a guidare questo cambiamento con strumenti come la seminatrice e nuove tecniche di lavorazione del terreno, gettando le basi per un approccio più scientifico all'agricoltura. Questi cambiamenti contribuirono ad aumentare le rese dei raccolti non realizzando macchine migliori, ma comprendendo meglio la terra e lavorandola meglio.

Alla fine del XIX secolo, il progresso agricolo non riguardava più solo le macchine, ma anche il suolo, la scienza e la strategia.

Lo stesso vale per la traduzione automatica oggi. I motori sono già potenti. Per migliorare ulteriormente, dobbiamo guardare oltre la macchina e concentrarci sui dati da cui apprende.

COCCODRILLO BRASATO IN SALSA DI MARMELLATA 🍷



Il mondo culinario è esplosivo

La comida no es solo una necesidad biológica sino que es un aspecto integral de la vida que desempeña un papel central en numerosas culturas. La comida encarna todo el entorno social y cumple una función conmemorativa, conectando la época contemporánea con las prácticas ancestrales y preservando el patrimonio cultural de una sociedad.

Come dice il vecchio proverbio, *siamo ciò che mangiamo*, e le attività legate al cibo (come cucinare, mangiare e discuterne) occupano una parte significativa della nostra vita quotidiana. L'autore A. Sonnenfeld, in *Food: A Culinary History from Antiquity to the Present*, suggerisce che il cibo sia forse l'espressione più distintiva di un gruppo etnico, di una cultura o, in tempi moderni, di una nazione.²

La cultura non solo determina cosa mangiamo, ma plasma anche il modo in cui consumiamo il cibo.³ Secondo Claude Lévi-Strauss, la cucina funziona come un linguaggio, che ne traduce la struttura inconscia. Le norme che regolano le sequenze e gli ingredienti dei pasti riflettono le connessioni sociali tra i commensali, e i codici culinari influiscono il significato attribuito alle diverse funzioni del cibo. In Giappone, ad esempio, l'aspetto del cibo è considerato fondamentale quanto il suo sapore, riecheggiando la convinzione di Goethe che un pasto debba soddisfare prima l'occhio e poi lo stomaco.

La condivisione del cibo si ritrova nei rituali religiosi

L'espressione *spezzare il pane* ha un profondo significato biblico, profondamente radicato nella tradizione cristiana. Al di là del suo contesto storico, spezzare il pane incarna i principi di comunione e comunità all'interno del cristianesimo. La Bibbia descrive la chiesa cristiana primitiva impegnata nella pratica della condivisione dei pasti, insieme alla preghiera e all'insegnamento degli apostoli. Questo atto comunitario simboleggiava l'unità tra i credenti e rafforzava i loro legami di fede.

Nell'ebraismo, il gesto di spezzare il pane è centrale nell'osservanza dello Shabbat, il giorno settimanale di riposo e di rinnovamento spirituale. Durante il pasto dello Shabbat, le famiglie si riuniscono tradizionalmente per condividere il pane, spesso sotto forma di challah, che viene benedetto e poi spezzato prima di essere distribuito ai presenti. Questo gesto simboleggia l'unità, la gratitudine e l'importanza del legame comunitario nella tradizione ebraica.

Nell'Induismo, l'atto di condividere il cibo riveste un'importanza significativa, in particolare nel contesto del rituale noto come *Prasad* o *Prasadam*. Prasad si riferisce al cibo offerto a una divinità durante il culto, che viene poi distribuito ai devoti come dono benedetto. Condividere il Prasad simboleggia la grazia spirituale, le benedizioni e il legame tra il fedele e il divino.

I musulmani osservanti seguono le linee guida alimentari delineate nella legge islamica, note come Halal. Queste includono cibi consentiti (Halal) e proibiti (Haram). Ad esempio, ai musulmani è vietato consumare carne di maiale e alcolici e devono garantire che la carne venga macellata secondo i principi islamici. Garantire che il

cibo sia consentito dalla legge islamica è di per sé un aspetto significativo della pratica islamica. In sostanza, il cibo è un potente mezzo di conservazione ed espressione culturale, incarnando le tradizioni, la storia e l'identità di una società.

È emersa la cucina fusion

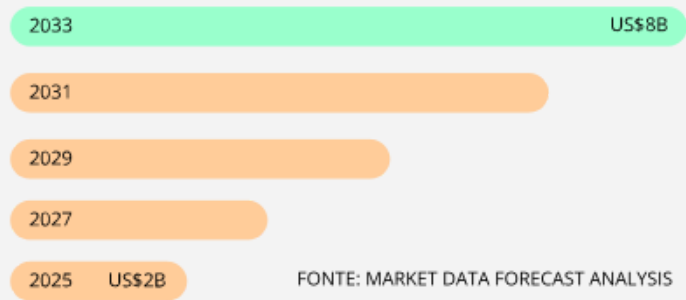
Gli immigrati hanno da tempo contribuito con le loro tradizioni culinarie dei paesi d'origine alla cucina locale delle loro nuove case. Ad esempio, la scrittrice di libri di cucina e antropologa Claudia Roden osserva che l'iconico piatto britannico *fish and chips* fu introdotto in Gran Bretagna nel XVI secolo da immigrati ebrei costretti ad abbandonare Spagna e Portogallo.⁴

La globalizzazione non solo ha facilitato la diffusione di marchi globali e catene di *fast food*, ma ha anche favorito l'apprezzamento per le cucine locali e gli ingredienti etnici. La cucina fusion si è affermata come una tendenza significativa, fondendo ingredienti e tecniche provenienti da tradizioni culinarie storicamente distinte. Questo scambio culinario globale ha aumentato la domanda di alimenti sani e prodotti biologici.

I sapori mediterranei sono spesso abbinati a spezie orientali, le tradizionali torte e i dolci britannici vengono reinventati con ingredienti come il basilico o il wasabi, e lo street food è stato elevato allo status di haute cuisine.⁵ La cucina fusion può essere vista come una forma di traduzione o *transcreazione*, che fonde elementi di diverse culture in creazioni culinarie uniche.

Turismo enogastronomico

Il turismo culinario o enogastronomico è la ricerca di esperienze enogastronomiche uniche e memorabili, sia vicine che lontane. Il turismo enogastronomico sta contribuendo alla diffusione di app di traduzione come Menulance. Nel 2024, il turismo gastronomico ha raggiunto 1,81 miliardi di dollari e si prevede che raggiungerà un valore di 7,91 miliardi di dollari entro il 2033.



Attraverso il cibo e le bevande, i viaggiatori possono scoprire la cultura, il patrimonio e la gente del paese. Queste app diventano strumenti indispensabili, migliorando l'esperienza gastronomica complessiva.⁶

- Il 70% della gente spende più soldi in cibo all'estero rispetto a quanto spende in patria, il che dimostra l'importanza di un'esperienza culinaria unica quando si viaggia;
- Il 70% della gente sceglie una destinazione secondo il cibo e le bevande presenti;
- Le statistiche sul turismo gastronomico mostrano che l'81% della gente pensa che il cibo aiuti a comprendere una cultura, secondo *Jersey Island Holidays*.

Gli uffici turistici di Singapore, Taiwan e Corea del Sud stanno promuovendo la qualità, il gusto e i benefici per la salute dei loro piatti ai viaggiatori. Mentre l'Europa domina ampiamente il mercato del turismo gastronomico, è molto interessante vedere come i paesi asiatici stanno ora promuovendo i loro piatti tipici agli stranieri.

I viaggiatori passano alla modalità scoperta

In uno studio sul turismo enogastronomico condotto nel 2020, i ricercatori hanno scoperto che i turisti modificano le proprie abitudini alimentari durante le vacanze. Ad esempio, i viaggiatori che visitano destinazioni balneari consumano significativamente più pesce e frutti di mare, ma meno legumi, carne, fast food, cereali e i loro derivati. Sia i turisti internazionali che quelli nazionali che visitano località balneari hanno sottolineato che assaggiare la cucina locale migliora significativamente la loro esperienza turistica complessiva.

In sostanza, i viaggiatori entrano in una mentalità di scoperta, in cui sono aperti a provare nuovi piatti e disposti a deviare dalla loro dieta abituale. Spesso intraprendono ricerche ed esplorazioni approfondite prima di prendere decisioni in merito al cibo. Lo stesso studio ha scoperto che:

- Prima di raggiungere la propria destinazione, l'80% dei viaggiatori fa delle ricerche su cibo e bevande;
- Negli ultimi due anni, il 93% dei turisti ha avuto almeno un'attività enogastronomica unica;
- Il 77% dei millennial viaggia per vivere esperienze memorabili a tavola, secondo *Jersey Island Holidays*.

La porndemia alimentare

Molta gente condivide sulle reti sociali fotografie del cibo che mangia. Una ricerca su Instagram con l'hashtag #food genera almeno 300 milioni di post, mentre #foodie ne genera almeno 100 milioni, a sottolineare l'innegabile importanza del cibo nella nostra società.⁷

Questa tendenza riflette un più ampio cambiamento culturale, in cui le immagini di cibo appetitoso sono sempre più diffuse. Definite *food porn* o *gastro porn*, queste immagini ad alta risoluzione popolano libri di cucina di successo, Instagram, X (Twitter), spot pubblicitari e annunci su riviste.⁸

In Gran Bretagna, il cibo è diventato una nuova forma di valuta culturale, superando il fascino della moda o delle preferenze musicali. Le riviste di cucina catturano il cibo con dettagli così intimi che i lettori si sentono immersi nell'esperienza sensoriale, che si tratti della consistenza di un peperone o del profumo del pane appena sfornato. I londinesi, un tempo definiti dalla scelta di stilisti o dai gusti musicali, ora si identificano per le loro preferenze culinarie e le ricette che sperimentano.

Il cibo, in quanto elemento universalmente apprezzato, ha un ruolo centrale nelle interazioni culturali. Personaggi influenti come Jamie Oliver, Michael Pollan, Mark Bittman e Rachael Ray promuovono il ritorno in cucina, sottolineando l'importanza di condividere i pasti nonostante le differenze religiose e alimentari. La proliferazione di scambi interculturali, migrazioni ed esposizione mediatica ha accresciuto la domanda di libri di cucina e menù tradotti, con blog e siti web che fungono da piattaforme per la condivisione di ricette globalmente.

L'influenza di canali televisivi dedicati al cibo come Food Network e UKTV Food, insieme a popolari programmi di cucina come MasterChef e The Great British Bake Off, è stata fondamentale per diffondere la conoscenza di nuove tradizioni gastronomiche. Questi programmi non solo insegnano agli spettatori la preparazione di piatti esotici, ma alimentano anche l'interesse per libri di cucina e riviste di cucina.

La cucina casalinga torna nel menù

I libri di cucina vendono straordinariamente bene. Secondo Kristen Mclean del gruppo di ricerche di mercato NPD, quasi 20 milioni di libri di cucina vengono venduti ogni anno solo negli Stati Uniti. Non solo, ma è anche uno dei mercati librari più stabili, con solo minime fluttuazioni annuali. È persino sopravvissuto alla pandemia di COVID-19.⁹

I libri di cucina sono il sesto genere letterario più popolare negli Stati Uniti, rappresentando il 5% delle vendite totali di libri.¹⁰

I dati di Nielsen BookScan mostrano che le vendite di libri di cucina negli Stati Uniti sono cresciute dell'8% su base annua tra il 2010 e il 2020, con numeri di vendita ulteriormente incrementati dalla pandemia.¹¹

Sebbene la pandemia in sé non abbia portato nulla di buono, ha dato una spinta al mercato dei libri di cucina, con sempre più gente che si è cimentata in cucina: ricordate tutti quei biscotti, quelle torte alla banana e quelle pagnotte di pane a lievitazione naturale? Molti hanno scoperto o riscoperto una passione per la buona cucina casalinga che continua anche da quando siamo *tornati alla normalità*.¹²

Nel primo anno di pandemia, le vendite di libri di cucina sono aumentate di circa il 16%. Sebbene da allora siano leggermente calate, le vendite rimangono solide. Ciò che cambia di più sono i tipi di libri di cucina che la gente acquista in un dato momento.¹³

«Nel 2020, dopo l'arrivo della pandemia, abbiamo assistito a un forte aumento di alcuni tipi di libri di cucina, tra cui libri di cucina sulle basi della cucina casalinga, libri di cucina sulla panificazione, e un po' più tardi libri di cucina sulla preparazione di cocktail e altri tipi di intrattenimento domestico», ha detto McLean. «Quindi è stato un vero e proprio riflesso della psicologia della gente bloccata a casa».¹⁴

Oggi la lista è completamente diversa.¹⁵

«Ora l'enfasi è davvero sui piatti veloci e facili, da preparare in un'unica pentola», ha detto. «Il tipo di cosa che ci si aspetterebbe se la gente dovesse reintegrare la cucina in vite molto frenetiche».¹⁶

DIMENSIONE 🧠

Ogni giorno
610 milioni di
persone fanno
spuntini con
la traduzione
automatica

L'IA ha dato vita alla traduzione automatica, considerata una branca o un sottocampo dell'intelligenza artificiale. Rientra nell'ambito più ampio dell'IA perché implica lo sviluppo di sistemi e algoritmi in grado di tradurre automaticamente testo o discorso da una lingua all'altra.

La traduzione automatica utilizza tecniche di intelligenza artificiale, tra cui l'elaborazione del linguaggio naturale e l'apprendimento automatico, per consentire ai computer di comprendere e generare il linguaggio umano, facilitando la comunicazione interlinguistica.

La traduzione automatica può impiegare diversi approcci, come metodi basati su regole, modelli statistici o reti neurali. La traduzione automatica è diventata rapidamente accessibile a tutti con l'avvento di Internet, spingendo numerosi fornitori di servizi tecnologici a offrire servizi di traduzione automatica gratuiti.

Dopo oltre 70 anni di evoluzione, la traduzione automatica ha raggiunto traguardi significativi. Le applicazioni di traduzione automatica utilizzano algoritmi e modelli per elaborare e convertire testo o discorso in una lingua in testo o discorso equivalenti in un'altra lingua.

Esempi di app di traduzione automatica includono Google Translate, Microsoft Translator e Menulance. Tuttavia, i motori di traduzione automatica presentano debolezze intrinseche che rendono difficile la traduzione tra diverse lingue. Ad esempio, spesso producono output costituiti in gran parte da traduzioni parola per parola.

Inoltre, i traduttori automatici non verificano il loro lavoro, non disponendo di una funzione di pausa e ripetizione che consenta loro di rivedere una frase più di una volta. Oltretutto, hanno difficoltà a gestire il contesto, rendendo difficile capire come una parola o una frase tradotta male possa alterare il significato di un brano in contesti diversi.

Nonostante i suoi difetti, la traduzione automatica è qui per restare.



610 milioni di persone usano Google Translate ogni giorno



1 miliardo di persone ha usato i servizi di DeepL



Bing ha 100 milioni di utenti attivi al giorno



Lionbridge ha tradotto 3,5 miliardi di parole



l'app Yandex raggiunge 40.000 scaricamenti al mese



le app di traduzione linguistica varranno 37 miliardi di dollari entro il 2033¹⁷

Sfide nella traduzione tradizionale

Gli strumenti di traduzione linguistica più diffusi, come Google Translate e Bing Translate, spesso faticano a tradurre accuratamente termini specialistici, in particolare in ambiti come quello culinario. Il difetto intrinseco del loro approccio deriva dall'allenamento su vasti set di dati che includono dizionari sia generalisti che specializzati, il che porta a imprecisioni e confusione semantica.

Nel dizionario culinario, queste inesattezze possono dare origine a traduzioni umoristiche e talvolta dannose.

kid on a spit

La traduzione automatica non è riuscita a tradurre *kid* usando il dizionario culinario (capretto), privilegiando invece il dizionario generale (bambino).

nun fart

La traduzione automatica non è riuscita a tradurre il piatto francese *Pet de Nonne* come entità denominata, ma lo ha invece trattato come un piatto grammaticale.

Nel 2015, Bored Panda ha pubblicato un post intitolato «80 dei più divertenti errori di traduzione di menu di sempre», citando esempi troppo volgari per essere pubblicati qui. [Leggeteli](#) a vostro rischio e pericolo.

La traduzione culinaria presenta difficoltà uniche. I termini e gli ingredienti culinari presentano spesso sfumature culturali e variazioni regionali. I motori di traduzione automatica potrebbero avere difficoltà a catturare accuratamente queste sottigliezze.

Alcuni prodotti alimentari o tecniche di cottura potrebbero non avere equivalenti diretti in altre lingue, rendendo difficile per il motore di traduzione trasmetterne il significato preciso. La scrittura culinaria spesso implica giochi di parole, creatività e linguaggio descrittivo. Tradurre questi elementi mantenendo il sapore e lo stile desiderati è un compito delicato.

Genesis: l'impasto madre della traduzione alimentare

Nel nucleo di Menulance si trova Genesis, un set di dati di formazione meticolosamente curato contenente 166 milioni di parole e frasi relative alle arti culinarie, disponibili in inglese, spagnolo e italiano. Oltre al suo enorme volume, Genesis possiede proprietà matematiche uniche e funziona come una banca dati computazionale rigorosamente verificata per un'efficienza ottimale.

In un test di traduzione automatica condotto da ottobre 2022 a gennaio 2023, abbiamo valutato le prestazioni di Genesis, il set di dati di allenamento di Menulance.

Abbiamo allenato tre motori di traduzione automatica utilizzando un set di convalida di 29.600 righe dall'inglese allo spagnolo estratte casualmente da Genesis.

L'allenamento è stato effettuato su motori di TAS (Traduzione Automatica Statistica), TAN (Traduzione Automatica Neurale) e TAI (Traduzione Automatica Ibrida).

Ecco i dettagli dei tre motori.¹⁸



Rilevamento degli errori, Correzione automatica, Rilevamento delle parole sconosciute



Semplificazione del testo/Gestione degli accenti



Ricerca/traduzione in tempo reale di testo preesistente



Motori di traduzione statistici, gerarchici e neurali



Massima probabilità di fusione tra i motori



Ricerca di parole e espressioni nel dizionario



Suggerimento sinonimo



Ricerche di borsa di parole

I motori di traduzione basati su Genesis hanno ottenuto un impressionante punteggio BLEU dell'83% per la traduzione dall'inglese allo spagnolo. I punteggi BLEU (Bilingual Evaluation Understudy) sono una misura standard usata per valutare l'efficacia dei motori di traduzione automatica. BLEU è una metrica ampiamente usata per valutare la qualità delle traduzioni generate automaticamente, confrontandole con le traduzioni di riferimento generate da esseri umani.

Un punteggio BLEU più alto indica un migliore allineamento con i riferimenti umani. BLEU valuta la qualità della traduzione confrontando gli n-grammi, o sequenze di

parole, nella traduzione generata automaticamente con quelli presenti nella traduzione di riferimento.

Per contestualizzare il risultato di Genesis, i principali modelli di traduzione automatica come Google Translate o DeepL ottengono in genere punteggi tra il 30 e il 40% nei test BLEU. I traduttori professionisti raggiungono punteggi BLEU di 60.

Di seguito è riportata un'interpretazione dei punteggi BLEU.¹⁹

30 - 40%	Da traduzioni comprensibili a buone
40 - 50%	Traduzioni di alta qualità
50 - 60%	Traduzioni di altissima qualità, adeguate e fluide
> 60%	La qualità è spesso migliore di quella umana
83%	Traduzione automatica allenata su Genesis

Genesis raggiunge una precisione del 97,6%

L'83% di BLEU è un risultato eccezionale. Tuttavia, volevamo un secondo parere.

Abbiamo preso i risultati del test BLEU e abbiamo rivalutato i tre motori utilizzando un tipo di verifica umana chiamato METEOR. METEOR, acronimo di Metric for Evaluation of Translation with Explicit ORdering, valuta le traduzioni confrontando le traduzioni BLEU in uscita con il set di dati del test e calcolando un punteggio basato sulla loro corrispondenza in termini di significato e struttura. Si tratta di un tipo di verifica umana e può essere più accurato di BLEU perché affronta alcune delle limitazioni di BLEU, come la sensibilità ai sinonimi e all'ordine delle parole.

BLEU ha i suoi punti di forza, soprattutto nelle valutazioni su larga scala, dove la valutazione manuale potrebbe risultare poco pratica. Fornisce una misura rapida e oggettiva delle differenze tra le traduzioni. Tuttavia, potrebbe non cogliere alcuni aspetti della qualità della traduzione che i valutatori umani possono discernere, come sfumature stilistiche, espressioni idiomatiche o appropriatezza culturale.

Nel corso di due mesi, i ricercatori hanno verificato manualmente i risultati di ciascun motore, valutando 88.800 righe di output BLEU. I risultati sono stati sorprendenti.

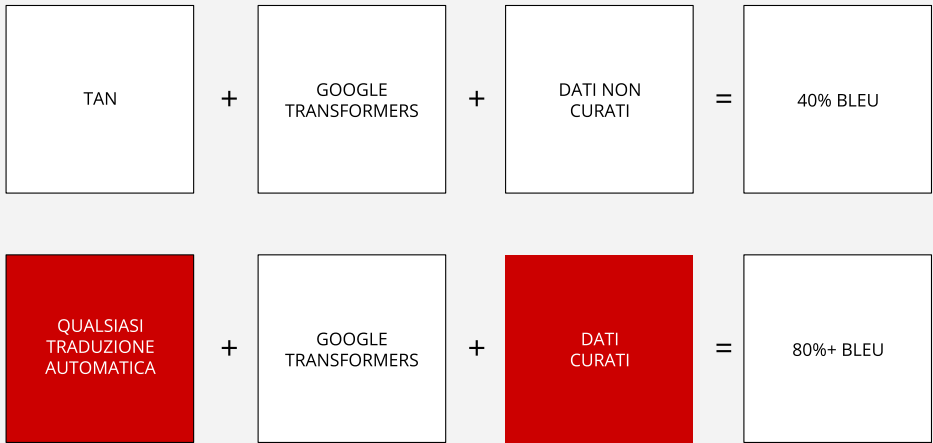
METEOR ha rilevato che un'accuratezza del 97,6% era un punteggio più realistico per i motori di traduzione automatica allenati su Genesis.

I risultati di ciascuna valutazione manuale sono riportati di seguito. La combinazione di parametri quantitativi come BLEU con il giudizio umano ha garantito un approccio più sfumato e olistico alla valutazione della qualità della traduzione.

	TAS	TAN	TAI
Precisione dei motori di traduzione automatica allenati su Genesis	97.33%	96.97%	97.41%

La colpa è del carburante, non del motore

Sia BLEU che METEOR confermano la stessa conclusione: i motori di traduzione automatica non sono intrinsecamente imperfetti; piuttosto, l'accuratezza delle loro traduzioni è strettamente legata alla qualità dei dati di allenamento a cui vengono esposti. Usando gli stessi motori che gestiscono altre app di traduzione automatica tradizionali, possiamo ottenere una maggiore accuratezza concentrandoci *sul carburante, non sul motore*.



Esiste un'idea sbagliata abbastanza diffusa secondo cui l'inserimento di dati non curati o grezzi nell'intelligenza artificiale sia sufficiente per l'allenamento e il raggiungimento di prestazioni affidabili e non guidate. Questa è una semplificazione eccessiva e non coglie il ruolo cruciale della cura dei dati e la necessità di una profonda comprensione dei principi dell'IA.

Il risultato sono motori che faticano a destreggiarsi tra le complessità del linguaggio, con conseguenti traduzioni non ottimali. Per affrontare la causa principale della scarsa accuratezza nella traduzione automatica è fondamentale sottolineare l'importanza di dati di allenamento di alta qualità, diversificati e ben curati.

Ecco una tabella riassuntiva di ciò che abbiamo scoperto:



I motori di traduzione automatica allenati su dati di formazione mal curati hanno ottenuto punteggi BLEU *bassi*



I motori di traduzione automatica allenati su Genesis hanno ottenuto punteggi BLEU *elevati*



I motori di traduzione automatica allenati su set di dati curati hanno ottenuto punteggi *simili* ai punteggi BLEU

GRILLED ZEBRA DRIZZLED IN PEANUT BUTTER 🍳



COME 🧑🍳

8 buone pratiche apprese curando Genesis

Di seguito sono riportate 8 buone pratiche che abbiamo appreso curando Genesis. Se state intraprendendo un progetto di traduzione automatica, queste lezioni sono preziose.

1 LIMITARE L'ALLENAMENTO A UN SINGOLO DOMINIO

I modelli di traduzione generici, allenati su un'ampia gamma di argomenti, incontrano spesso difficoltà nell'interpretare significati specifici del contesto. Ad esempio, la parola *kid* significa tipicamente *bambino*, ma in contesti culinari si riferisce solitamente a un *capretto*. A causa di distorsioni statistiche, i modelli generali tendono a scegliere il significato più frequente, traducendo erroneamente piatti contenenti termini specialistici, come ad esempio interpretando *kid in wild mushroom sauce* come un bambino.

Per mitigare tale ambiguità, Menulance ha curato Genesis, un set di dati incentrato esclusivamente sul vocabolario culinario, che comprende meno del 10% dell'intero lessico inglese. Questo approccio mirato garantisce che il motore di traduzione riconosca accuratamente i termini alimentari specializzati e i significati specifici del contesto, riducendo drasticamente gli errori causati dall'ambiguità semantica.

La ricerca supporta questo approccio specifico per dominio. Studi di Koehn e Knowles (2017)²⁰ evidenziano che i modelli di TAN allenati su dati specifici di un dominio raggiungono punteggi BLEU più elevati e una maggiore fluidità rispetto ai modelli generali allenati su argomenti diversi. I loro risultati sottolineano che limitare i dati di allenamento a un dominio specifico migliora la qualità della traduzione e riduce al minimo la confusione derivante da contesti irrilevanti.

In definitiva, il successo della traduzione automatica non dipende solo dalla quantità di dati, ma anche dalla loro qualità, specificità e pertinenza. Set di dati accuratamente selezionati come Genesis consentono ai motori di traduzione automatica di fornire traduzioni precise, affidabili e contestualmente accurate.

2 ALLENA USANDO FRASI BREVI

Genesis è allenato in piatti e ingredienti alimentari, un ambito naturalmente adatto a frasi brevi. Se il linguaggio culinario prevedesse tipicamente frasi o paragrafi più lunghi, curare Genesis sarebbe stato significativamente più complesso, poiché le frasi lunghe introducono maggiore complessità e un rischio maggiore di errori durante l'allenamento dei motori di traduzione automatica.

Quando una frase è breve, ha meno parole che ne compongono il significato. La gente trova naturalmente le frasi brevi più facili da capire perché ci sono meno parole da capire.²¹ Questo principio si estende all'allenamento dei motori di traduzione automatica: le frasi brevi sono più semplici da imparare e tradurre rispetto a quelle più lunghe.

La lunghezza media delle frasi nel set di dati di Genesis è di circa 11 parole. I motori di traduzione automatica generici spesso si allenano su centinaia di parole. Sono progettati per gestire paragrafi, il che porta ad algoritmi eccessivamente complessi che devono gestire *dipendenze a lungo raggio*, in cui il significato si estende su più frasi.

Un esempio culinario di dipendenza a lungo raggio potrebbe apparire in una ricetta o in una recensione di un cibo distribuita su più frasi. Si consideri quanto segue:

«Soffriggere delicatamente le cipolle e l'aglio in olio d'oliva fino a quando non saranno traslucidi. Metterli da parte. Nella stessa padella, rosolare le cosce di pollo. Rimettere le cipolle e l'aglio nella padella, aggiungere i pomodori e cuocere a fuoco lento finché il pollo non sarà tenero». Qui, *li* nella seconda frase e *le cipolle e l'aglio* nella quarta frase si riferiscono agli ingredienti menzionati all'inizio.

Amazon Alexa, l'assistente virtuale a comando vocale, utilizza tecniche di allenamento a frasi brevi e di decomposizione canonica per migliorare le sue capacità di comprensione del linguaggio. Scomponendo i comandi dell'utente in unità più piccole e concentrandosi su frasi brevi e contestualizzate, Alexa può interpretare e rispondere con precisione a un'ampia gamma di domande e comandi dell'utente.²²

3 EVITARE L'ALLENAMENTO UTILIZZANDO DATI PUBBLICI

Le lingue ad alto contenuto di risorse come l'inglese, lo spagnolo e il mandarino traggono vantaggio da grandi quantità di dati linguistici,²³ tra cui ampi corpora paralleli (raccolte di testi allineati sia nella lingua di partenza che in quella di arrivo) e risorse linguistiche complete (corpora ben annotati, grammatiche e dizionari), che migliorano tutti la qualità della traduzione.

Al contrario, lingue con scarse risorse come il creolo haitiano, il maltese e il kinyarwanda affrontano sfide significative a causa della scarsità di dati di allenamento e risorse linguistiche. Questa mancanza di dati spesso porta a una minore accuratezza dei risultati della traduzione automatica, poiché queste lingue non beneficiano degli stessi corpora paralleli robusti o dei materiali di riferimento completi disponibili per le lingue con risorse elevate.²⁴

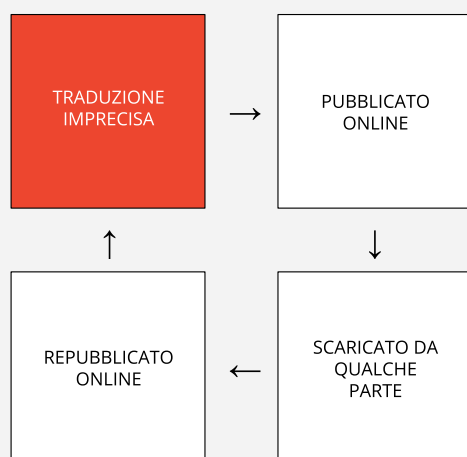
Uno studio condotto dall'UCLA Medical Center nel 2021 ha evidenziato che la precisione di Google Translate variava dal 55% al 94%, a seconda che la lingua di destinazione fosse una lingua con molte o poche risorse.²⁵

punteggio BLEU	
Spagnolo	94%
Mandarino	81.7%
Farsi	67.5%
Armeno	55%

La correlazione tra lingue ad alto contenuto di risorse e dati di allenamento accurati, sebbene vera in senso generale, non ha facilitato la cura di Genesis. Nonostante Genesis fosse stato allenato in lingue ad alto contenuto di risorse (inglese, spagnolo e italiano), abbiamo dovuto affrontare notevoli difficoltà perché le informazioni culinarie online erano inaffidabili e piene di imprecisioni.

Nella nostra esperienza, non era rilevante se la lingua fosse ricca o povera di risorse: i dati culinari disponibili erano imperfetti e inadatti alla Genesi.

La pratica comune di utilizzare queste traduzioni culinarie imprecise per allenare i principali motori di traduzione automatica ha dato origine a un circolo vizioso in cui le inesattezze vengono continuamente pubblicate, scaricate e ripubblicate nel pubblico dominio.



Una volta introdotta una traduzione imprecisa, la sua ripetizione su più supporti crea un'illusione di correttezza in virtù della sua frequenza. Col tempo, questi errori ripetuti si consolidano come verità perché si verificano così spesso che gli utenti e persino i successivi sistemi di traduzione automatica iniziano ad accettarli senza porsi domande.

Oggi, traduzioni imprecise di cibi sono presenti in tutti i tipi di media: menù dei ristoranti, etichette alimentari, siti di e-commerce, dizionari tradizionali, dizionari di nicchia, blog e libri stampati.

Questo circolo vizioso dimostra quanto sia fondamentale la cura dei dati, poiché interromperlo richiede di garantire che ogni dato di allenamento sia accurato e verificato, impedendo agli errori di propagarsi attraverso Genesis.

Indipendentemente dal motore di traduzione automatica utilizzato (TAS, TAI o TAN), se allenato su set di dati scarsamente curati come questi, i motori faranno fatica a produrre traduzioni accurate.

Consapevoli che nessun motore di traduzione automatica può funzionare bene con dati così sporchi, ci siamo imbarcati nel laborioso compito di curare manualmente il nostro set di dati da zero. Partendo da una raccolta iniziale di 140.000 frasi nel 2012, abbiamo scoperto che il 40-50% delle informazioni culinarie online era inaccurato e doveva essere corretto manualmente da traduttori professionisti.

Ci siamo concentrati su precisione e qualità, piuttosto che sul volume di dati. Questo approccio è in linea con le migliori pratiche di altri settori; ad esempio, **Waymo**, la

sussidiaria di Alphabet specializzata in auto a guida autonoma, ha dedicato oltre un decennio alla raccolta continua e all'annotazione meticolosa dei dati dei sensori per allenare i suoi algoritmi di guida autonoma.

Allo stesso modo, IBM Watson for Oncology ha raggiunto il successo grazie all'allenamento su set di dati accuratamente selezionati, composti da letteratura medica, dati di studi clinici e approfondimenti di esperti. Anche il sistema Autopilot di Tesla deve la sua efficacia nel mondo reale al continuo perfezionamento di enormi quantità di dati annotati e di alta qualità raccolti dalla sua flotta, che garantiscono sicurezza e prestazioni.²⁶

4 RACCOGLIERE SET DI DATI DI ALLENAMENTO DENS

I set di dati di grandi dimensioni sono cruciali per l'allenamento dei motori di traduzione automatica, e molti esperti suggeriscono che 20 milioni di esempi siano un buon punto di partenza. Tuttavia, la densità dei dati è altrettanto critica. Avere un set di dati di grandi dimensioni non è utile se ripete la stessa coppia di traduzioni milioni di volte. Un dataset di allenamento ben curato richiede dati densi, bilanciati e diversificati per garantire un apprendimento completo.

I set di dati densi e massicci risolvono anche un problema comune dei motori di traduzione automatica: il sovradattamento. Il sovradattamento si verifica quando un motore di traduzione automatica memorizza i propri dati di allenamento e non riesce a generalizzare oltre. Un motore di traduzione automatica che si sovradatta non riesce a gestire richieste di traduzione al di fuori del proprio set di allenamento, generando traduzioni imprecise.²⁷

Considerate un'auto a guida autonoma Tesla allenata su una sola strada di Los Angeles. Vi fidereste di guidarla su strade sconosciute in Australia? Certamente no. Il veicolo deve sperimentare migliaia di strade diverse in numerose città, con modelli di traffico variabili e condizioni meteorologiche diverse prima di poter generalizzare in modo affidabile le sue capacità di guida. Allo stesso modo, un motore di traduzione automatica deve imparare da un set di dati ampio e denso per tradurre in modo accurato e gestire efficacemente frasi mai incontrate prima.

I set di dati densi sono fondamentali perché forniscono esempi di allenamento diversificati e di alta qualità per il modello di traduzione. Questi set di dati sono ricchi di varianti linguistiche e coprono un'ampia gamma di stili alimentari, contesti e sintassi. La ricchezza dei set di dati densi consente ai modelli di traduzione automatica di catturare le complessità del linguaggio in modo più efficace, con conseguente maggiore qualità della traduzione e una migliore fluidità del testo tradotto.

Il set di dati del Genesis originale, composto da 4 milioni di linee, era sparso e non uniformemente bilanciato. Per risolvere questo problema, abbiamo ampliato il dataset utilizzando un approccio strutturato. Questo ci ha permesso di costruire sistematicamente un set di dati sostanzialmente più ampio e denso, composto da 166 milioni di linee.

Ad esempio, abbiamo prodotto varianti come:

coccodrillo brasato in salsa di marmellata,
anatra brasata in salsa di marmellata, e
maiale brasato in salsa di marmellata.

Questa maggiore densità ha fatto sì che il motore di traduzione automatica incontrasse numerosi modelli di frasi simili ma distinti, migliorando notevolmente la sua capacità di generalizzare.

L'espressione *sponsali con prosciutto* viene spesso tradotto erroneamente dai motori di ricerca tradizionali come *matrimonio con prosciutto* perché sul web il termine *sponsali* si riferisce prevalentemente a *matrimonio* anziché a *erba cipollina*. Nella traduzione del termine *sponsali*, la parola *matrimonio* avrà maggiori probabilità di essere selezionata al posto di *erba cipollina*, il che porterà a una traduzione culinaria contestualmente errata. Un modo per evitare questo problema è aumentare il numero di coppie di frasi italiano-inglese con la presenza simultanea delle parole *sponsali-erba cipollina*, aumentando così le probabilità che *sponsali* venga tradotto correttamente in *erba cipollina*.

Sebbene siano necessari set di dati di grandi dimensioni, la loro densità e diversità sono ciò che realmente migliora le prestazioni dei modelli di traduzione automatica. Garantire che i dati siano ben curati e rappresentativi delle diverse sfumature linguistiche è fondamentale per ottenere una traduzione di qualità superiore.

5 IDENTIFICARE I GRADI DI LIBERTÀ ALL'INTERNO DEI DATI DI ALLENAMENTO

Raccogliere set di dati di allenamento densi è fondamentale, ma altrettanto importante è determinare con precisione quali punti dati raccogliere. Questa selezione dipende dalla chiara definizione dei gradi di libertà (GdL) del set di dati, che rappresentano le variabili e le variazioni specifiche che influenzano i risultati che un modello di IA o di traduzione automatica deve gestire.

Ad esempio, durante lo sviluppo del progetto di riconoscimento facciale Jijo-2, abbiamo identificato nove GdL critici, tra cui il tipo di telecamera utilizzata, la sua posizione, le condizioni di illuminazione, le pose 3D e gli attributi specifici del soggetto come l'età e l'espressione facciale. Ognuna di queste variabili ha influenzato significativamente la raccolta dei dati, garantendo che il set di dati coprisse in modo completo le variazioni del mondo reale e consentendo ai nostri modelli di IA di generalizzare efficacemente.

Al contrario, in un progetto che ho completato con Unilever nel 2019, il riconoscimento di flaconi di shampoo e balsamo richiedeva meno GdL. Fattori come le condizioni di illuminazione, l'orientamento del flacone e il posizionamento del prodotto erano essenziali, ma le espressioni facciali o il tipo di telecamera erano irrilevanti. Definire chiaramente la profondità di campo ci ha permesso di concentrarci sui dati necessari, aumentando l'accuratezza e l'efficienza del modello di riconoscimento basato sull'IA.

Per un'auto a guida autonoma Tesla di cui abbiamo parlato in precedenza, i GdL nei dati di allenamento includono vari fattori ambientali e situazionali che influenzano la guida. Questi includono tipologie di strade (autostrade, strade residenziali, strade sterrate), condizioni stradali (bagnate, ghiacciate, asciutte), densità del traffico (da bassa a congestionata), illuminazione (diurna, crepuscolare, notturna, artificiale), variazioni meteorologiche (soleggiato, piovoso, nevoso, nebbioso), tipologie di ostacoli (pedoni, ciclisti, animali, veicoli, cantieri), nonché differenze geografiche e regionali (segnaletica stradale, convenzioni di guida).

L'acquisizione di dati completi in questi diversi scenari garantisce che il veicolo possa affrontare in modo sicuro e affidabile situazioni non esplicitamente incontrate durante l'allenamento.

GdL

Auto Tesla a guida autonoma	Tipi di strade, condizioni stradali, densità del traffico, illuminazione, variazioni meteorologiche, tipi di ostacoli, differenze geografiche e regionali.
Modello di IA Jijo-2	Tipo di telecamera, posizione della telecamera, condizioni di illuminazione, pose 3D, età, espressione facciale, altri attributi specifici del soggetto.
Genesis di Menulance	Ingredienti, nomi dei piatti, categorie di cibo, aggettivi, genere.

Nella traduzione automatica, in particolare in ambiti specializzati come il linguaggio culinario, il GdL diventa fondamentale per catturare efficacemente le variazioni linguistiche. Ad esempio, i set di dati culinari potrebbero includere variazioni negli ingredienti, nei metodi di cottura, negli stili di presentazione, nei nomi dei piatti, nei termini regionali, nell'origine e nelle sfumature descrittive. Senza una chiara definizione di questi GdL, i set di dati potrebbero non avere la variabilità necessaria per allenare accuratamente i motori di traduzione, limitandone l'efficacia in presenza di input nuovi o leggermente modificati.

Per fare un esempio, si consideri la traduzione di un piatto come il *coccodrillo brasato in salsa di marmellata*. I GdL in questo caso potrebbero includere metodi di cottura (brasato, grigliato, arrosto), tipi di carne (coccodrillo, alligatore, pollo) e varianti di salsa (salsa di marmellata, glassa al miele, salsa al pepe).

Koehn osserva che i GdL chiaramente identificati consentono l'espansione e la gestione efficienti dei set di dati, migliorando significativamente l'adattabilità di un modello alle sfumature specifiche del dominio e riducendo le imprecisioni semantiche (Koehn & Knowles, 2017).²⁸

Inoltre, l'identificazione chiara dei GdL favorisce la scalabilità e la capacità di adattarsi alle esigenze future. Se le variabili rilevanti sono note fin dall'inizio, espandere o perfezionare il set di dati per adattarlo a nuove condizioni, prodotti o contesti diventa semplice.

6 ESTRARRE ENTITÀ DENOMINATE DAI DATI DI ALLENAMENTO

I modelli di traduzione automatica si basano fortemente sulla frequenza delle parole presenti nei dati di allenamento, il che li rende vulnerabili a imprecisioni quando si incontrano termini rari, specializzati o di recente coniazione.²⁹ Questo problema è particolarmente significativo in ambiti specializzati come la traduzione culinaria, dove prevalgono nomi di piatti particolari e terminologia tecnica.

Abbiamo affrontato questa sfida sostituendo o aggiungendo sistematicamente entità denominate in Genesis. Le entità denominate sono piatti conosciuti globalmente, che spesso mantengono i loro nomi originali in tutte le lingue, come gli *spaghetti alla carbonara*. Che la lingua sia lo spagnolo, l'italiano o l'inglese, rimane sempre *spaghetti alla carbonara*.

I motori di traduzione tradizionali spesso non trattano correttamente piatti come gli *spaghetti alla carbonara* come un piatto specifico, il che porta a traduzioni eccessivamente letterali o prive di senso. Ad esempio, l'espressione italiana *alla carbonara* significa letteralmente *nello stile del carbonaio*. Senza contesto, un motore di traduzione automatica tradizionale potrebbe tradurre *spaghetti alla carbonara* in qualcosa come *spaghetti del carbonaio* o *spaghetti al carbone*, perdendone completamente il significato culinario e la natura appetitosa.

Quando aggiungiamo un piatto come gli spaghetti alla Carbonara, eseguiamo due attività critiche in Genesis:



Frammentare la frase in sottofrasi, come *spaghetti* e *alla carbonara*, il che aiuta il modello a comprendere i singoli elementi.



Conservare il piatto completo *spaghetti alla carbonara* come un'unica entità atomica.

Questo approccio arricchisce notevolmente il set di dati, potenziando la capacità del motore di traduzione automatica di tradurre in modo accurato mediante l'aumento dell'esposizione a piccole variazioni nei termini culinari riconoscibili.

Oggi Genesis contiene decine di migliaia di tali entità denominate, rafforzando l'accuratezza del modello attraverso questo processo strutturato di categorizzazione e arricchimento.

Al contrario, piatti come il *coccodrillo brasato servito con salsa di marmellata* sono tipicamente localizzati, descrittivi e trattati come piatti grammaticali. A differenza delle entità denominate, questi piatti si basano in gran parte sulla traduzione di ogni sottofrase descrittiva. Frammentando questi piatti, il modello di traduzione automatica analizza separatamente il significato e le relazioni grammaticali di ciascun termine, garantendo che la frase tradotta trasmetta accuratamente il significato e il contesto culinario del piatto, anche se l'intera frase non è familiare al pubblico di destinazione.

La ricerca supporta questo approccio strutturato alle entità denominate e all'analisi grammaticale.³⁰ Secondo Koehn e Knowles (2017), una chiara categorizzazione specifica per dominio e un'espansione controllata del vocabolario migliorano significativamente i risultati della TAN.

Allo stesso modo, un recente lavoro di Sennrich e Haddow (2016) evidenzia l'efficacia della segmentazione delle sottoparole e dell'etichettatura sistematica, che aiutano i modelli a gestire meglio termini rari o mai visti prima scomponendo e categorizzando elementi lessicali complessi.³¹

Inoltre, l'estrazione di entità denominate migliora notevolmente la coerenza della traduzione. La denominazione coerente di piatti e ingredienti è essenziale per la marca culinaria, la commercializzazione e l'esperienza utente, soprattutto nei menù o nei libri di cucina multilingue. Il riconoscimento corretto delle entità denominate garantisce che i nomi dei piatti rimangano uniformi in tutti i testi tradotti, creando familiarità e fiducia tra i consumatori.³²

7 MIGLIORARE L'ALLINEAMENTO DELLE PAROLE CON SOTTOFRASI VALIDE

L'aggiunta di sottofrasi verificate a Genesis ha migliorato la precisione dell'allineamento delle parole nei casi in cui viene utilizzato un motore di TAS. L'allineamento delle parole è un passaggio fondamentale della TAS che garantisce che parole e frasi in una lingua corrispondano correttamente al loro equivalente più accurato in un'altra lingua. Funziona in modo simile a un dizionario bilingue, in cui ogni parola o frase in una lingua di partenza viene sistematicamente associata alla traduzione più probabile nella lingua di arrivo.

Ad esempio, in un corpus parallelo (un set di dati di frasi bilingue), l'allineamento delle parole garantisce che *apple* in inglese corrisponda correttamente a *mela* in italiano.

Per ottenere un allineamento preciso delle parole, i motori di TAS analizzano ampi set di dati e generano automaticamente tabelle di coppie di frasi durante l'allenamento. Queste tabelle registrano la frequenza con cui una frase in una lingua compare insieme a una frase in un'altra, assegnando punteggi di probabilità basati sulla loro co-occorrenza statistica. Quando un motore TAS traduce una nuova frase, fa riferimento a questi punteggi di probabilità, selezionando la parola o la frase corrispondente più probabile per garantire una traduzione accurata.

Un piatto come *cocodrillo brasato in salsa di marmellata* non ha una traduzione consolidata e universalmente accettata. Poiché il sistema non ha mai visto questa frase prima, frammenta la frase in sottofrasi, assegnando punteggi di probabilità basati sulla frequenza con cui ciascuna coppia di frasi appare nei dati di allenamento:

origine (inglese)	destinazione (italiano)	punteggio di probabilità
braised crocodile	cocodrillo brasato	0.90
braised crocodile	cocodrillo stufato	0.10
in marmalade sauce	in salsa di marmellata	0.85

in marmalade sauce	in salsa alla marmellata	0.15
braised crocodile in marmalade sauce	coccodrillo brasato in salsa di marmellata	0.80
braised crocodile in marmalade sauce	coccodrillo stufato in salsa alla marmellata	0.20

Tuttavia, entità denominate, come *spaghetti alla carbonara*, compaiono nelle tabelle delle frasi come traduzioni fisse ad alta confidenza. Poiché questo piatto è riconosciuto nella Genesi come entità denominata, la sua coppia di frasi riceve un punteggio di probabilità di 1,00:

origine (inglese)	destinazione (italiano)	punteggio di probabilità
spaghetti alla carbonara	spaghetti alla carbonara	1.00

La probabilità di ogni coppia di frasi è determinata dalla frequenza dei dati: una probabilità più elevata (più vicina a 1,0) indica una traduzione più comune e affidabile. Poiché *spaghetti alla carbonara* è un'entità denominata, raggiunge un'accuratezza del 100%. Tuttavia, *coccodrillo brasato in salsa di marmellata* è meno comune, il che determina una probabilità inferiore (0,80), a dimostrazione della sua rarità nel set di dati.

Una delle principali sfide della TAS è il *rumore di fondo* causato da sottofrasi non valide. I motori di TAS estraggono automaticamente le coppie di frasi, ma questo processo non è perfetto: possono essere introdotte molte coppie di frasi prive di significato o errate.

Ad esempio, se un modello di TAS incontra la frase:
Wine of Australia and caviar

Potrebbe estrarre erroneamente la coppia di frasi:
Australia and caviar → *Australia e caviale*

Sebbene grammaticalmente corretta, questa frase non ha alcun significato culinario reale. Tuttavia, poiché i motori di TAS non hanno consapevolezza del contesto, la trattano come una traduzione valida, il che inquina le traduzioni future.

Questo rumore distorce i calcoli di probabilità, aumentando il rischio di traduzioni errate. Nel tempo, l'esposizione ripetuta a sottofrasi non valide può indurre il sistema a produrre output privi di senso, riducendo ulteriormente l'accuratezza della traduzione.

Abbiamo rimosso sistematicamente le sottofrasi non valide da Genesis per migliorare le prestazioni dei motori di TAS. Ciò è stato possibile solo perché Genesis è allenato utilizzando frasi brevi: se il set di dati contenesse paragrafi lunghi, verificare ogni frammento sarebbe quasi impossibile.

Anche altri approcci di traduzione automatica basati su frasi utilizzano tabelle di coppie di frasi, ma la loro affidabilità e il loro approccio variano rispetto alla TAS. La traduzione automatica basata su frasi (TABF), un'evoluzione della TAS, si basa ampiamente sulle tabelle di coppie di frasi. Segmenta le frasi in frasi anziché singole parole e le allinea tra le lingue. L'aggiunta di sottofrasi verificate migliora la capacità della TABF di assemblare correttamente le traduzioni. Come la TAS, la TABF trae notevole vantaggio dall'utilizzo di tabelle di frasi di alta qualità, poiché dati più accurati migliorano l'accuratezza della traduzione.

La TAI combina tecniche basate su regole (TABF) e statistiche (TAS/TABF). Alcuni modelli di TAI utilizzano tabelle di coppie di frasi insieme a regole linguistiche per fornire traduzioni più strutturate. La misura in cui le tabelle di frasi vengono utilizzate nella TAI dipende dall'equilibrio tra apprendimento statistico e logica basata su regole all'interno del sistema.

A differenza della TABF e la TAI, la TAN non utilizza tabelle di coppie di frasi. Si basa invece sull'apprendimento profondo per tradurre intere frasi in modo olistico. Tuttavia, se i dati di allenamento sono privi di sottofrasi valide, i modelli di TAN potrebbero generare traduzioni errate a causa di una scarsa comprensione contestuale. Sebbene la TAN non utilizzi direttamente tabelle di frasi, beneficia indirettamente delle sottofrasi, poiché dati di qualità superiore riducono il rumore durante l'allenamento, con conseguente miglioramento dell'accuratezza della traduzione.

8 MISURARE I DATI DI ALLENAMENTO UTILIZZANDO REGOLARMENTE METEOR

Il BLEU rimane lo standard per misurare l'accuratezza della traduzione automatica. Il BLEU viene calcolato come media su un ampio set di test, fornendo un punteggio affidabile perché premia le parole con la più alta probabilità di essere utilizzate.

Tuttavia, BLEU non è esente da limitazioni. BLEU valuta le traduzioni basate sulla corrispondenza degli n-grammi (sequenze di parole) con quelli presenti nella traduzione di riferimento. Se una parola rara nella traduzione generata automaticamente ha un sinonimo nella traduzione di riferimento, ma non corrisponde esattamente, BLEU potrebbe penalizzare la traduzione anche se il significato viene preservato.

Il BLEU combina 1 grammo, 2 grammi, 3 grammi e 4 grammi in una statistica che calcola la differenza tra il risultato atteso e quello previsto. Se sono presenti le stesse parole, si ottiene un punteggio più alto. Un sinonimo è una parola morfologicamente (molto) diversa, quindi si avrà una discrepanza di 1 grammo, 2 grammi, 3 grammi e 4 grammi, quindi un BLEU inferiore perché le parole sono morfologicamente diverse.

Affrontare questa limitazione è difficile perché il BLEU non offre una comprensione semantica e si basa su corrispondenze esatte di parole. Ricercatori e professionisti sono consapevoli di questo problema e spesso lo tengono in considerazione quando interpretano i punteggi BLEU. Ciò sottolinea l'importanza di utilizzare più metriche di valutazione e il giudizio umano per ottenere una valutazione più dettagliata delle traduzioni automatiche.

Esistono diversi altri metodi per misurare la traduzione automatica, tra cui METEOR. BLEU è uno standard necessario per attività di valutazione della traduzione automatica trasparenti e riproducibili ed è per questo che viene utilizzato nelle competizioni internazionali: agenzie indipendenti come il NIST possono ripetere l'esperimento sui motori di traduzione allenati e confermarne i punteggi.

Tuttavia, una misura migliore è METEOR, un tipo di verifica umana. Scienziati dei dati di tutto il mondo eseguono attività di verifica umana in tutti i campi dell'IA, dalla traduzione automatica alla visione artificiale. Abbiamo condotto una verifica manuale su 29.600 frasi di prova e sul loro output previsto. Mentre BLEU ci ha dato un'accuratezza dell'83%, la precisione reale è stata del 97,6%, dovuta alla presenza di sinonimi.

CONCLUSIONE

È importante capire che la traduzione automatica è uno strumento, non un sostituto della traduzione umana. Il rapporto tra traduttori umani e macchine è molto più simbiotico che competitivo. 610 milioni di persone usano Google Translate ogni giorno, traduttori professionisti inclusi.

La traduzione automatica è preziosa, in quanto fornisce una traduzione iniziale rapida che i traduttori umani possono poi perfezionare e migliorare. Può contribuire ad accelerare il processo di traduzione, gestire grandi volumi di contenuti e fungere da risorsa di riferimento per i traduttori. Tuttavia, il risultato finale beneficia spesso dell'attenta revisione e dell'adattamento forniti dai traduttori umani.

Per ottenere una traduzione automatica di alta qualità, soprattutto in un ambito specializzato come quello culinario, non è sufficiente utilizzare il motore più avanzato, ma utilizzare i dati giusti.

Genesis dimostra che dati di allenamento attentamente selezionati, specifici per un determinato dominio e strutturati in modo intelligente possono produrre sistemi di traduzione automatica che superano di gran lunga i dati non selezionati. Nel settore culinario, dove significato, cultura e fiducia sono tutto, questa non è solo una buona pratica. È un requisito.

TORTA DI SALSICCIA VIENNESE 🧑🍳



APP 

La prima IA al
mondo che
traduce ciò che
hai nel piatto



NOTE FINALI 🍷

1. *JETIR*, Volume 10, Issue 11, "Impact of Digital Innovations and AI on Gastronomy, Tourism, and Local Food," Khilesh Patel and Dr Sunil Kumar, Nov 2023.
2. *Food: A Culinary History from Antiquity to the Present*, A. Sonnenfeld, 1999, xvi.
3. *The Translator*, "What's cooking in English culinary texts? Insights from genre corpora for cookbook and menu writers and translators," Michał B. Paradowski, Jan 2018.
4. *The ATA Chronicle*, "What's Cooking: An Introduction to Culinary Translation," July 2021.
5. *The Translator*, Vol 21, issue 3, "Food and translation, translation and food," Delia Chiaro and Linda Rossato, Dec 2015.
6. Market Data Forecast, *Culinary Tourism Market Research Report*, Mar 2023.
7. Computer Vision Foundation, "Inverse Cooking: Recipe Generation from Food Images," Amaia Salvador, Michal Drozdal, Xavier Giro-i-Nieto, Adriana Romero, Dec 2018.
8. SBS News, "How food porn hijacks your brain," Christian Jarrett, Oct 2015.
9. Marketplace, "What is it about cookbooks?" Samantha Fields, Jan 2023.
10. GITNEX, "Must-Know Book Sales By Genre Statistics," Dec 2023.
11. *The Guardian*, "What's the point of cookbooks? Hope, love and beauty (but not cooking)," Kate Gibbs, Nov 2022.
12. Marketplace, "What is it about cookbooks?" Samantha Fields, Jan 2023.
13. Ibidem.
14. Ibidem.
15. Ibidem.
16. Ibidem.
17. FactMR, *Translation software outlook 2023-2033*, Jan 2023.
18. Menulance, Executive Summary, May 2023.
19. Google Cloud, "Evaluation models."
20. Department of Computational Linguistics, University of Zurich, Switzerland, "Domain Robustness in Neural Machine Translation," Mathias Muller, Annette Rios, Rico Sennrich, Oct 2020.
21. *Computational Linguistics*, Volume 48, Number 2, "Challenges of Neural Machine Translation for Short Texts," Yu Wan, Dec 2021.
22. Alexa Developer Documentation, "Best Practices for Sample Utterances and Custom Slot Type Values," Mar 2024.

23. Portland Community College, "The 4 Most Common Languages Spoken Around the World," Rebecca Raymond, June 2019.
24. Naarg Data Media Services, "Top 7 challenges in machine translation."
25. Phrase.com, "Research vs Practice: How Accurate Is Google Translate?" Nov 2023.
26. Bhumi Publishing, *Cutting-Edge Technology From Tesla*, Zaiba Khan, Oct 2024.
27. Wikipedia, "Hallucination (artificial intelligence)."
28. Association for Computational Linguistics, *Proceedings of the First Workshop on Neural Machine Translation*, Thang Luong, Alexandra Birch, Graham Neubig, Andrew Finch, Aug 2017.
29. LinkedIn.com, "Training AI to Translate: The Promise and Challenges of Machine Translation," Jesse Anglen, Sept 2023.
30. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, "Using Confidential Data for Domain Adaptation of Neural Machine Translation," Sohyung Kim, Arianna Bisazza, Fatih Turkmen, 2021.
31. *Artificial Intelligence Review*, "Revisiting named entity recognition in food computing: enhancing performance and robustness," Uchenna Akujuobi, Shuhong Liu, Tarek R. Besold, Aug 2024.
32. *Language Resources and Evaluation*, "ATR4S: Toolkit with state-of-the-art automatic terms recognition methods in Scala," Nikita Astrakhantsev, Dec 2017.



INFORMAZIONI SUGLI AUTORI

Il Dott. Roberto Mariani, ex scienziato diventato imprenditore nel settore tecnologico, ha conseguito un dottorato di ricerca in fisica presso l'Institut Géographique National de Paris e l'Università di Rouen. Con 30 anni di esperienza nell'IA, ha fondato e guidato numerose iniziative nell'ambito dell'analisi video, del riconoscimento facciale e della traduzione automatica. Roberto detiene nove brevetti, a dimostrazione della sua costante ricerca dell'innovazione.

Anthony Coundouris vanta oltre vent'anni di esperienza nella consulenza per startup e ha conseguito una laurea in economia aziendale presso la University of Technology di Sydney. Imprenditore per due volte, è specializzato nella progettazione di sistemi automatizzati di vendita e marketing. Anthony è autore del libro *run_frictionless: How to Free a Founder from a Sales Role* e ideatore di un framework decisionale chiamato i 4Q.

INFORMAZIONI SUL TRADUTTORE

Il Dott. Frank Badrines, nato a Barcellona nel 1975, ha conseguito il dottorato di ricerca in filosofia presso l'Università di Barcellona con una tesi sull'influenza di Platone e Aristotele su C. S. Lewis. Attualmente è docente di lingua spagnola presso la National University of Singapore e lavora anche come traduttore di narrativa greca moderna in Spagna. È entrato a far parte di Menulance nel 2013.

CHI SIAMO

Menulance è la prima app di traduzione di menù basata sull'IA, allenata su milioni di termini verificati. Scopri e traduci piatti e ingredienti in spagnolo, italiano e inglese.

media@menulance.com

www.menulance.com



TACOS AL TIRAMISÙ CON MOUSSE AL MARSALA 🧑🍳

