



CASE STUDY SULL'OTTIMIZZAZIONE E SULLA GENERAZIONE DI CONTENUTI TRAMITE INTELLIGENZA ARTIFICIALE

Esperimento sull'ottimizzazione e sulla
localizzazione dei contenuti tramite
intelligenza artificiale generativa

4 STRUMENTI DI INTELLIGENZA
ARTIFICIALE TESTATI

PROMPT SCRITTI E TESTATI IN 2 LINGUE

PUNTEGGIO DI QUALITÀ MEDIO PER I CONTENUTI IN TEDESCO
OTTIMIZZATI CON PROMPT IN INGLESE: 86%

DIETRO LE QUINTE DELL'ESPERIMENTO DI LIONBRIDGE

Lionbridge ha condotto questo esperimento usando il team, le risorse e i contenuti interni. In particolare, abbiamo usato i nostri contenuti in tedesco di blog già esistenti. Sebbene questo case study sia incentrato sui contenuti di Lionbridge (in particolare in lingua tedesca), l'intenzione è quella di creare un protocollo e raccogliere informazioni che sia possibile applicare a migliaia di nostri clienti in tutti i settori e le lingue. Useremo i risultati e le osservazioni di questo esperimento per soddisfare le varie esigenze di contenuto dei clienti in diverse lingue. Applicheremo le informazioni apprese anche ai nostri processi interni di creazione e ottimizzazione dei contenuti di marketing.

LA SFIDA

L'ottimizzazione dei contenuti, specialmente in altre lingue, può essere un'attività dispendiosa in termini di tempo e costi. Le traduzioni dei contenuti esistenti, inoltre, non sempre consentono di ottenere prestazioni di alto livello in termini di SEO. A volte, si tratta semplicemente di traduzioni dirette che non tengono conto delle sfumature linguistiche, delle parole chiave specifiche del luogo e del comportamento degli utenti e/o dei dati. Gli strumenti di intelligenza artificiale generativa possono aiutarci a ottimizzare i contenuti multilingue riducendo i costi, i tempi e il lavoro necessario.

BACKGROUND

Questi quattro strumenti di **intelligenza artificiale generativa**, Amazon Bedrock, ChatGPT (istanza di Azure di Lionbridge con ChatGPT 4), Google Bard e Llama di Meta, sono tutti in grado di creare e ottimizzare contenuti.

Analizzano modelli di ricerca, dati dei concorrenti e comportamenti degli utenti per fornire suggerimenti basati sui dati per parole chiave e contenuti.

Questi strumenti aiutano a ottimizzare i contenuti in modo da renderli accattivanti e con un forte valore SEO, cercando di allinearsi il più possibile alle aspettative degli utenti. In alcuni casi, possono addirittura offrire risultati migliori di quelli ottenuti dagli esseri umani.

Gli strumenti basati sull'intelligenza artificiale generativa possono ad esempio avere capacità equivalenti a quelle degli esseri umani nel rispettare rigorosi limiti di caratteri nella scrittura e nella traduzione dei contenuti (in particolare per ottimizzare il valore SEO).

LA SOLUZIONE

Abbiamo eseguito un esperimento in tre fasi usando gli strumenti di GenAI per individuare quale fosse il più efficace per la ricerca e la generazione di contenuti, specialmente in altre lingue. Il nostro team si è concentrato sulle best practice per lo sviluppo di prompt, in particolare per i contenuti multilingue. Abbiamo anche esaminato le best practice per l'ottimizzazione dei contenuti già esistenti in nuove lingue, al fine di migliorare il posizionamento SEO di un sito web.

METODOLOGIA

Per lo studio, Lionbridge ha usato quattro strumenti di intelligenza artificiale: Amazon Bedrock, ChatGPT (istanza di Azure di Lionbridge con ChatGPT 4), Google Bard e Llama di Meta. Abbiamo eseguito le tre fasi (ricerca, creazione e analisi) due volte.

Ricerca: in questa fase, abbiamo chiesto a ciascuno strumento di intelligenza artificiale di completare una ricerca delle parole chiave e selezionare le due domande migliori per la creazione di contenuti in tedesco.

Abbiamo fornito a ognuno degli strumenti i seguenti elementi:

- Un prompt di personalizzazione, chiedendo di agire come un content writer tedesco e "fornire 3 nuove parole chiave in tedesco e 3 domande in tedesco".
- Contenuto di origine da ottimizzare
- Parole chiave esistenti

Abbiamo inoltre scelto due domande supplementari da usare in seguito per generare piccoli paragrafi di contenuto.

Per la ricerca di parole chiave, l'intervento umano sembra essere essenziale, in particolare per Bard. Lo strumento svolge un lavoro di ricerca eccellente, ma fornisce troppe opzioni e richiede l'intervento di un esperto per la selezione delle parole chiave o delle domande necessarie per la generazione dei contenuti.



Creazione: in questa fase, abbiamo sviluppato due set di prompt in tedesco. Questi prompt per la creazione di contenuti includevano le parole chiave identificate nel passaggio precedente. Un set era basato sulla personificazione di un content writer, mentre l'altro era dedicato alle istruzioni per l'attività.

Abbiamo chiesto a un membro del dipartimento di intelligenza artificiale di Lionbridge di revisionare e adattare i prompt per ottimizzarne l'efficacia. Una volta finalizzati i prompt (illustrati di seguito nella sezione "Prompt"), abbiamo usato le risorse di Lionbridge per localizzarli in tedesco. Abbiamo poi eseguito questi due set completi di prompt in tedesco in tutti e quattro gli strumenti di intelligenza artificiale, raccogliendo e contrassegnando i contenuti suggeriti da ogni strumento.

Per sviluppare i prompt abbiamo usato questi input:

- Domande e parole chiave dalla fase di ricerca (in tedesco)
- Esempi di contenuti di buon livello che volevamo emulare (in tedesco)
- Testo dell'articolo di origine (in tedesco)
- Prompt di personificazione (in inglese e in tedesco)
- Prompt per l'esecuzione dell'attività (in inglese e in tedesco)

Dovendo usare due set di prompt, abbiamo dovuto eseguire il progetto due volte in ognuno degli strumenti di intelligenza artificiale.

Analisi: in questa fase abbiamo chiesto a dieci esperti linguistici e del settore di lingua tedesca di revisionare i contenuti. Per svolgere una "revisione alla cieca", non abbiamo indicato lo strumento da cui sono stati generati i contenuti.

Abbiamo aggiunto gli output da valutare in un foglio di calcolo usando due meccanismi di valutazione:

- Un menu a discesa con tre valori "Inutilizzabile", "Utilizzabile con modifiche" e "Utilizzabile senza modifiche"
- Un campo in cui assegnare un punteggio da 1 a 100 al contenuto

Abbiamo raccolto il feedback e lo abbiamo esaminato per determinare:

- Quale è stato lo strumento più efficace?
- Quali sono le best practice per la creazione dei prompt? Come si deve procedere per creare prompt in altre lingue?
- Che prestazioni offrono questi strumenti nella ricerca di parole chiave?
- Che prestazioni offrono gli strumenti nella generazione di domande relative al contenuto?
- Che prestazioni offrono gli strumenti nella generazione di punti di vista per gli articoli?
- Che prestazioni offrono gli strumenti nella creazione di risposte alle domande generate?
- Come possiamo ottenere processi ottimali di creazione di contenuti basati sull'intelligenza artificiale?
- Come si sviluppano prompt efficaci?

PIANIFICAZIONE



Sceita di 3 strumenti di GenAI da usare



Sceita di 3 articoli da ottimizzare per l'esperimento



Ricerca di 10 esperti per la revisione



Ricerca di parole chiave e domande



Sviluppo di prompt e localizzazione in tedesco

ESEMPI DI PROMPT USATI

Prompt di personificazione

- Lavori come content writer per Lionbridge, ti occupi della creazione di contenuti relativi ai servizi linguistici.
- Parli tedesco e vivi in Germania.
- Crei contenuti in tedesco, rivolti a un pubblico di lingua tedesca in Germania, Austria e Svizzera...

Prompt relativi ai destinatari

- Persone che lavorano nel marketing o nelle comunicazioni
- Persone che occupano una posizione senior, come Senior Manager, direttore, vicepresidente, presidente, CMO, CEO
- Persone di lingua tedesca che vivono in Germania, Austria o Svizzera...

Prompt di creazione

- Crea due paragrafi di testo aggiuntivi per questo articolo...
- Ogni paragrafo deve contenere una risposta a queste domande...
- La parola chiave primaria per l'URL è...
- Le parole chiave secondarie sono...

TEAM MULTIDISCIPLINARE

- 10 tecnici di lingua tedesca per la creazione dei prompt
- 1 esperto di SEO di Lionbridge
- 1 esperto di marketing digitale di Lionbridge

I RISULTATI

L'intelligenza artificiale deve assumere un ruolo importante in qualsiasi progetto di ottimizzazione.

Il nostro esperimento ha confermato che gli strumenti basati sull'intelligenza artificiale possono aiutare a ottimizzare e aggiornare i contenuti esistenti, anche non in lingua inglese, in modo efficace ed efficiente.

È incoraggiante notare che la qualità degli output è migliorata nel corso dell'esperimento. L'intervento umano nel processo di ottimizzazione o creazione di contenuti è fondamentale per selezionare le parole chiave dall'output dell'intelligenza artificiale per l'ottimizzazione. È necessario un revisore umano anche per convalidare i contenuti di output finali. Questo aspetto potrebbe cambiare in futuro, considerando che la qualità degli output migliora sensibilmente nel tempo.

Ecco alcune best practice identificate:

- Attualmente, è necessario l'intervento umano per convalidare gli output finali dell'intelligenza artificiale.
- La qualità dei nostri input è migliorata nel corso dell'esperimento, man mano che abbiamo appreso nuove best practice.
- È più probabile che l'intelligenza artificiale produca contenuti di alta qualità se le richieste vengono inviate con:
 - Testi brevi
 - Contenuti che possono essere facilmente frammentati in sezioni più piccole (come elenchi dei primi dieci elementi e così via)
- Attualmente, la comunicazione con la tecnologia di intelligenza artificiale è più efficace in inglese. Questo non sorprende, in quanto l'addestramento degli strumenti avviene in inglese. Fornire prompt in tedesco non è stato utile per generare contenuti migliori in tedesco e gli strumenti hanno prodotto contenuti tedeschi migliori con i prompt in inglese. Nota: è comunque utile fornire contenuti di origine, parole chiave, domande ed esempi nella lingua di destinazione (diversa dall'inglese).
- I processi di creazione di contenuti devono utilizzare una descrizione approfondita del creatore di contenuti e del pubblico di destinazione. È inoltre utile creare raccolte di prompt, per consentire agli utenti di riutilizzare i prompt e sviluppare contenuti coerenti per i clienti.



Bard sembra aver fornito i migliori output per la ricerca.

I suoi risultati sono stati simili al livello di ricerca generato da uno strumento come MarketMuse, sviluppato appositamente per questo scopo. Tuttavia, gli output della ricerca di Bard richiedono l'intervento umano per selezionare le informazioni da utilizzare per la parte del processo relativa alla creazione dei contenuti. Utilizzando le istruzioni in tedesco, Bard ha fornito risultati costanti nelle due fasi dell'esperimento. Tuttavia, i risultati ottenuti dalle istruzioni in inglese sono passati da 82/100 a 66/100. Considerando che il set di dati era piuttosto limitato, la media dei punteggi ottenuti da Bard (in inglese) si attesta su 74/100.

I contenuti prodotti da ChatGPT hanno ricevuto commenti leggermente più negativi. Tuttavia, i revisori talvolta hanno elogiato i contenuti. Meta (Llama) non ha ottenuto risultati altrettanto soddisfacenti e probabilmente non è ancora pronto per ottimizzare o generare contenuti non in inglese in modo affidabile. La qualità degli output sembra essere migliorata nel corso di questo breve progetto durato un mese. Bard è migliorato negli output di ricerca, mentre ChatGPT è migliorato nella creazione di contenuti in tedesco. Gli output di ChatGPT creati da istruzioni in tedesco hanno ottenuto inizialmente un punteggio di 47/100.

Nel giro di un solo mese, il punteggio ha raggiunto 76/100. Per gli output generati dalle istruzioni in inglese, il punteggio è passato da 71/100 a 86/100 nello stesso periodo di tempo (il punteggio medio per l'inglese è di 78,5/100).

Sia Bard che ChatGPT hanno ottenuto buoni risultati nella produzione di contenuti brevi (ad esempio una o due frasi). I feedback negativi per i contenuti brevi sono stati minimi.

I punteggi illustrati si riferiscono ai contenuti di lunghezza maggiore nelle due fasi dell'esperimento.

PUNTEGGI FASE 1

Origine	Inutilizzabile	Utilizzabile con modifiche	Utilizzabile senza modifiche	Punteggio medio
ChatGPT (input tedesco)	9	9	0	47,61
Bard (input tedesco)	2	14	2	68,33
ChatGPT (input inglese)	3	9	6	71,67
Bard (input inglese)	0	9	9	82,22
Meta (Llama) (input inglese)	7	8	3	46,50
TOTALE	21	49	20	

PUNTEGGI FASE 2

Origine	Inutilizzabile	Utilizzabile con modifiche	Utilizzabile senza modifiche	Punteggio medio
ChatGPT (input tedesco)	1	7	0	76,25
Bard (input tedesco)	2	4	2	66,25
ChatGPT (input inglese)	0	4	4	86,875
Bard (input inglese)	3	4	1	59,375
TOTALE	5	19	7	

Nota: Meta (Llama) è stato rimosso nella fase 2 perché non è ancora pronto per ottimizzare o generare contenuti non in lingua inglese.

