📘 Resum didàctic: *GEO: Generative Engine Optimization*

Autors: *Pranjal Aggarwal, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, Karthik Narasimhan i Ameet Deshpande (IIT Delhi, Princeton i investigadors independents)*

Publicat a: *KDD '24 (Barcelona, agost de 2024) · arXiv 2311.09735v3*

Codi i dades: *https://generative-engines.com/GEO/*

🧭 La idea en una frase

Si abans calia fer **SEO** perquè Google et posés als primers resultats, ara que molta gent pregunta directament a una IA (ChatGPT, Perplexity, Bing Chat, Google SGE…), cal fer **GEO**: adaptar el contingut d'una web perquè la IA l'**esculli, el citi i en parli més** a la seva resposta.

1. El context: de "llista d'enllaços" a "resposta redactada"

Abans: el cercador clàssic

Escrius una pregunta → reps **una llista de 10 enllaços ordenats** → tu hi fas clic i llegeixes. Per a una web, ser visible = **sortir ben amunt a la llista**.

Ara: el "motor generatiu" (*Generative Engine*, GE)

Escrius una pregunta → una IA busca fonts, les llegeix i **et redacta una resposta** amb petites cites entre claudàtors [1], [2]… Molts usuaris **ja no fan clic** a cap web.

Qui hi guanya i qui hi perd?

Actor	Efecte
👤 **Usuari**	Hi guanya: respostes ràpides i personalitzades
🏢 **Empresa del motor** (Google, Microsoft…)	Hi guanya: usuaris més satisfets i més ingressos
✍️ **Creadors de contingut / webs**	**Hi perden**: menys visites, i no saben com ni per què la IA els cita (és una "caixa negra")

👉 El paper vol ajudar aquest tercer grup.

---

2. Com funciona un motor generatiu (versió simplificada)

```
Pregunta de l'usuari
        |
        ▼
[1] Un model reformula la pregunta en subpreguntes més simples
        |
        ▼
[2] Un cercador clàssic recupera fonts (webs)
        |
        ▼
[3] Un model resumeix cada font
        |
        ▼
[4] Un model redacta la resposta final, amb cites a les fonts
```

En els experiments, els autors en fan una versió de **2 passos**: agafen les **5 primeres webs de Google** i les donen senceres a **GPT-3.5-turbo**, que redacta la resposta.

---

3. Problema nº 1: com es mesura la "visibilitat" en una resposta d'IA?

A Google n'hi ha prou amb la **posició a la llista**. Però en una resposta redactada, una font pot:

- aparèixer al principi o al final,
- ocupar una frase curta o un paràgraf sencer,
- ser decisiva o anecdòtica.

Per això proposen **mètriques noves** (anomenades *impressions*):

a) 📏 Recompte de paraules (*Word Count*)

Quina proporció de les paraules de la resposta pertany a frases que citen la teva web.

> *Analogia: és com mesurar quants **minuts de pantalla** té un actor en una pel·lícula.*

b) 📍 Recompte ajustat per posició (*Position-Adjusted Word Count*)

El mateix, però les frases que surten **abans** pesen més (amb una caiguda exponencial), perquè la gent llegeix sobretot el principi.

> *Analogia: no és el mateix sortir al **titular** d'un diari que a l'última pàgina.*

c) 🧠 Impressió subjectiva (*Subjective Impression*)

Una IA (GPT-3.5, amb el mètode G-Eval) puntua cada cita segons 7 aspectes:

1. **Rellevància** respecte la pregunta
2. **Influència** (quant depèn la resposta d'aquesta font)
3. **Originalitat** del que aporta
4. **Posició** percebuda
5. **Quantitat** de contingut percebuda
6. **Probabilitat que l'usuari hi faci clic**
7. **Diversitat** del material

---

4. Problema nº 2: no hi havia dades per fer proves → GEO-bench

---

Creen un banc de proves amb **10.000 preguntes** (8.000 entrenament / 1.000 validació / 1.000 test), extretes de 9 fonts:

Font	Què aporta
MS MARCO, ORCAS-I, Natural Questions	Preguntes reals de Bing i Google
AllSouls (Oxford)	Preguntes d'assaig que exigeixen raonar
LIMA	Preguntes difícils (fer un poema, codi…)
Davinci-Debate	Preguntes de debat
Perplexity.ai Discover	Temes de tendència
ELI5 (Reddit)	"Explica-m'ho com si tingués 5 anys"
Generades amb GPT-4	Per augmentar la diversitat

- **25 dominis** (art, salut, jocs, dret…)
- **80 %** preguntes informatives, **10 %** transaccionals, **10 %** de navegació
- Cada pregunta va amb el text de les **5 primeres webs de Google**

5. Les 9 estratègies GEO que proven

Fan servir una IA per reescriure una de les 5 webs aplicant cada estratègia, i miren si després la IA la cita més.

#	Estratègia	Què fa	Afegeix contingut nou?
1	**Autoritari**	To més persuasiu i segur	No
2	**Afegir estadístiques**	Canvia frases vagues per xifres	Sí
3	**Keyword stuffing**	Omple el text de paraules clau (SEO clàssic)	Sí
4	**Citar fonts**	Afegeix referències a fonts fiables	Sí
5	**Afegir cites textuals**	Inclou frases d'experts o fonts creïbles	Sí
6	**Fàcil d'entendre**	Simplifica el llenguatge	No
7	**Optimitzar la fluïdesa**	Text més ben redactat i fluid	No
8	**Paraules úniques**	Afegeix vocabulari poc comú	No
9	**Termes tècnics**	Afegeix vocabulari especialitzat	No

6. 📊 Resultats principals

Punt de partida (web sense tocar): **19,3** punts de visibilitat.

Estratègia	Posició-ajustada	Subjectiva	Valoració
Sense optimitzar	19,3	19,3	—
Keyword stuffing	**17,7** ⬇️	20,2	❌ No funciona (fins i tot empitjora)
Paraules úniques	20,5	20,4	❌ Gairebé no fa res
Fàcil d'entendre	22,0	20,5	✅
Autoritari	21,3	22,9	✅ moderat
Termes tècnics	22,7	21,4	✅
Fluïdesa	24,7	21,9	✅✅
Citar fonts	24,6	21,9	✅✅
Estadístiques	25,2	23,7	✅✅✅
Cites textuals	**27,2**	**24,7**	🏆 **La millor**

Lliçons clau

1. 🏆 **Cites, estadístiques i fonts** són el que més funciona: entre **+30 % i +40 %** de visibilitat (i +15–30 % en la mètrica subjectiva). I amb canvis petits!
2. ✍️ **La forma també compta**: millorar la fluïdesa o la claredat dona +15–30 %.
3. 🗣️ **Ser "autoritari" ajuda poc**: les IA ja són força resistents al to persuasiu.
4. 🚫 **El SEO clàssic no serveix**: el *keyword stuffing* no millora res. Els models entenen el significat, no només compten paraules.

7. Anàlisis complementàries

7.1 Cada tema té la seva estratègia

Estratègia	On funciona millor
Autoritari	Debats, història, ciència
Fluïdesa	Negocis, ciència, salut
Citar fonts	Afirmacions, fets, dret i govern
Cites textuals	Persones i societat, explicacions, història
Estadístiques	Dret i govern, debat, opinió

👉 Una web d'història hauria de posar **cites**; una de política pública, **xifres**.

7.2 🎯 GEO "democratitza": beneficia sobretot els petits

Quan s'optimitzen **totes** les webs alhora, el resultat depèn de la posició original a Google:

Posició a Google	Efecte de "Citar fonts"
1a	**–30,3 %**
5a	**+115,1 %**

Per què? A Google pesen molt coses difícils per a un petit creador (enllaços entrants, autoritat del domini). En canvi, la IA **llegeix el contingut**: si està ben fet, el cita, siguis qui siguis.

> *Analogia: és com passar d'un concurs on guanya el que té més contactes a un on guanya el que **escriu millor**.*

7.3 Combinar estratègies encara és millor

- La millor parella és **Fluïdesa + Estadístiques**: supera qualsevol estratègia sola en més d'un 5,5 %.
- La **fluïdesa** és la millor "companya" de les altres (+31,4 % de mitjana en combinació).

7.4 Exemples reals (canvis mínims, grans efectes)

Pregunta	Canvi	Millora
"Quin és el secret de la xocolata suïssa?"	Afegir de quina enquesta surt una dada	**+132 %**
"Haurien els robots de substituir els humans?"	Afegir "un 70 % d'augment de robòtica en l'última dècada"	**+65 %**
"Els Jacksonville Jaguars han anat mai a la Super Bowl?"	To més enfàtic i positiu	**+89 %**

8. 🌍 Prova en un motor real: Perplexity.ai

Per comprovar que no és només el seu sistema de laboratori:

Estratègia	Posició-ajustada	Subjectiva
Sense optimitzar	24,1	24,7
Keyword stuffing	21,9 ⬇️ (≈ −10 %)	28,1
Cites textuals	**29,1** (+22 %)	32,1
Estadístiques	26,2	**33,9** (+37 %)

✅ Les conclusions **es mantenen** en un producte real amb milions d'usuaris.

9. ⚠️ Limitacions (reconegudes pels autors)

- Els motors generatius **evolucionen** ràpidament: les estratègies hauran d'anar canviant, com ha passat amb el SEO.
- Les preguntes reals també canvien amb el temps → caldrà actualitzar GEO-bench.
- **No han mesurat** si aquests canvis afecten la posició a Google (tot i que creuen que poc, perquè només toquen el text).
- Amb models capaços de llegir **més fonts**, la posició a Google importarà menys.
- L'etiquetatge de temes pot tenir errors o subjectivitat.

10. ✅ Conclusió i consells pràctics

Aportacions del paper:

1. Defineixen formalment què és un **motor generatiu** i què és **GEO**.
2. Proposen **mètriques de visibilitat** pensades per a respostes d'IA.

3. Publiquen **GEO-bench**, el primer banc de proves gran per a aquest problema.
4. Demostren que estratègies senzilles poden augmentar la visibilitat **fins a un 40 %**.

🛠️ Si tens una web, què en pots treure?

- ✓ Afegeix **dades i xifres** concretes en lloc de frases vagues.
- ✓ Inclou **cites** d'experts o fonts reconegudes.
- ✓ **Referencia les fonts** d'on treus la informació.
- ✓ Escriu de manera **clara i fluida**.
- ✓ Adapta l'estratègia al **tema** de la teva web.
- ❌ Oblida't del *keyword stuffing*: ja no funciona.

📖 Glossari ràpid

Terme	Significat
LLM	Model de llenguatge gran (com GPT)
Motor generatiu (GE)	Cercador que respon amb text redactat per IA i cites a fonts
SEO	Optimització per a cercadors clàssics
GEO	Optimització per a motors generatius
Impressió / visibilitat	Quant i com apareix una web dins la resposta
SERP	Pàgina de resultats d'un cercador clàssic
Benchmark	Conjunt de proves estàndard per comparar mètodes
Caixa negra	Sistema del qual no coneixes el funcionament intern
Keyword stuffing	Farcir un text de paraules clau per enganyar el cercador