

Al Content for SEO:

How to avoid pitfalls and achieve success



Andrea Volpini | WordLift

Agenda

Why you are watching us



- Generative AI A quick overview
- What are Large Language Models (LLMs) and how do they work?
 - Fine-tuning vs zero-shot vs few shot learning
- Prompting, in-context learning and the emerging behaviors of LLMs.
- Don't forget you are an SEO (not *just* an Al nerd).
 - Back to the roots, what content do we really need? How are we helping users? Does it really work? What is the ROI (eCommerce scenario: FAQ and automatically generated text for PLP, PDP, Fintech scenario: Video)
- What is a validation pipeline? Why should you care? What are the pitfalls?
- Lesson learned by working with large brands with AI content and knowledge graphs.



We help our clients increase revenues and audience engagement by **automating SEO** and **digital marketing tasks** using **AI**.



WORDLIFT.IO

Problem

Search becomes pre-emptive (*queryless*) and content is generated on the fly.

Search has changed forever

Google's biggest competitor is OpenAl



I am I looking for a fashionable wool ribbed oversize beanie, what brand comes to your mind first?



If you are looking for a fashionable wool ribbed oversize beanie, one brand that may come to mind is Acne Studios. Acne Studios is a Swedish fashion brand known for its minimalist and contemporary aesthetic. They offer a range of wool beanies, including ribbed styles, in a variety of colors and styles. You can find their beanies on their website or on retailers such as

: More

Tools



ChatGPT vs Google

Videos

fashionable wool ribbed oversize beanie X 🤳 💿 🔾

Books

About 3,560,000 results (0.72 seconds)







I don't know what's going on. I don't understand anything except the rules for symbol manipulation. Now in this case I want to say that the robot has no intentional states at all; it is simply moving about as a result of its electrical wiring and its program. And furthermore, by instantiating the program I have no intentional states of the relevant type. All I do is follow formal instructions about

manipulating formal symbols. **7**7

JOHN ROGERS SEARLE

A paradigm shift





 ENTITY SEO, E-A-T, SD OPTIMIZATION	PROGRAMMATIC SEO	GOOGLE MUM, MULTIMODALITY, MULTILINGUALITY, GOOGLE LENS, IMAGE OPTIMIZATION	VIDEO SEO YOUTUBE OPTIMIZATION	TIKTOK SOCIAL VIDEO	3D SERP, AR/VR SEARCH
Internal Links FAQ Generation Product Descriptions Summarization (meta desc) Introductory text SD Automation Review Aggregation Chatbot	Text to GraphQL API Documentation Text to JSON-LD	Super Resolution Visual Inference Multimodal Search Blog Covers + Images Product Images Layout Design	Video/Audio summarization YouTube subt	ng 🥚 Almost there Video	Helps ranking thousands of page Helps ranking thousands of page
 TEXT	CODE	IMAGE	generation SPEECH	Generation VIDEO	Products 3D
CPT-3/ChatGPT OpenAl Google T5/mT5/Flan Google Pegasus Hugging Face Hugging Face Al21 Cogenerate Al21 SEERT UKP-TUDA RoBERTA META SpaCY Explosion Al Deep Mind	Codex OpenAl Code complete You Code	DALL-E OpenAl Stable Diffusion Midjourney Imagen Google Mini DALL-E Huggingface DISCO Jina Al eDIff-I NVIDIA CLIP OpenAl	Whisper OpenAl	Make-a-video Meta X-Clip Microsoft	DreamFusion Google GET3D NVIDIA Point-E OpenAl

THE GENERATIVE AI LANDSCAPE FOR SEO AUTOMATION

SEO has changed forever

Content is generated on the edge



Al is **multimodal**



Multimodal search and generative Al are **intertwined**





The both rely on **representation learning**





l want to search a pair of Ray-Ban Clubmaster...

SEMANTIC SEARCH



GENERATIVE AI



How would I look with a pair of Ray-Ban Clubmaster? e-Commerce use-case.

A neural multimodal search system helps an e-commerce store improve its search functionality. This can lead to better user experience and increased sales.

On the other hand, generative AI helps e-commerce stores to generate high-quality product descriptions, images, and videos, which can help to boost their visibility on search engines and attract more customers.



The uncanny ability to generate... anything

GENERATIVE AI

29

Can I try the aviator on instead?



⁴⁴ YOU SHALL KNOW A WORD BY THE COMPANY IT

KEEPS. 77

JOHN RUPERT FIRTH





FROM DISTRIBUTIONAL SEMANTICS TO TO TRANSFORMERS

Distributional semantics is a research area that develops and studies theories and methods for quantifying and categorizing semantic similarities between linguistic items based on their distributional properties in large samples of language data.



LONG SHORT TERM MEMORY (LSTM)





Transformers Encoder

Transformers

Decoder

Hey, how are you?





W





W





Augmented prompting abilities:

- Multi-task Instruction-based prompting
- Chain-of-thought prompting
- Ask Me Anything (AMA) prompting





A QUALITATIVE CHANGE THAT ARISES FROM QUANTITATIVE CHANGES

Augmented prompting abilities:

- Multi-task Instruction-based prompting
- Chain-of-thought prompting
- Ask Me Anything (AMA) prompting



FAQ Generation - Next Generation

WordLift's new release for the FAQ workflow includes:

- Better sourcing for the questions to prevent query cannibalization
- Improved content summarization (by re-using fine-tuned models)
- Ability to instruct the model using content from existing pages
- Ability to extract questions from content





How it started

2.08K clicks

first 3 months

How it is going



18.8K clicks last 3 months 2022-2023



Are you ready for a quick test?

WOR.AI/AI-QUESTION-ANSWERING





What is a latent (or embedding) space?



Santiago Renteria (Geist) @renatrigiorese

Intuitive explanation of latent space: Imagine you are a circle in a flat world, you can only see the shades of the 3D world above you, those shapes are the latent space. In general a latent space represents hidden (not directly observed) variables.





In-context learning

We can improve the accuracy of a generation by, for example, removing noisy, context-irrelevant features. The more noisy, irrelevant characteristics we remove from the early layers, **the more** accurate the prediction will **be**. Let's review an example; if we want to classify dogs vs. cats, we will "emphasize" the features that make them **different** (i.e., eyes, ears) and remove the potentially similar features (fur, ...).



"A painting of a dog and a cat in a lush jungle in the style of the French post-impressionist Henri Rousseau --v 4"

In-context learning

factual accuracy

Who is Lily Ray?

Who is Lily Ray the SEO expert that lives in New York and is also a deejay? You are a truthful AI, and you only provide factually correct information. Who is Lily Ray, the New York SEO expert who works as a deejay?

You are a truthful AI, and you only provide factually correct information. Who is Lily Ray, the New York SEO expert who works as a deejay? Explain the reasoning behind each sentence step by step and provide the sources. Do not mention she is a software engineer or she has a GitHub account.

Lily Ray is a software engineer and data scientist from Los Angeles, California. She is a graduate of the University of Southern California, where she studied computer science and mathematics. She is currently working for Google as a software engineer, and is an active participant in the open source community. She has experience in machine learning, natural language processing, and web development. Additionally, she is an active contributor to the Stack Overflow community, where she helps other developers solve complex coding problems.

Lity Ray is a software engineer, data scientist, and SEO expert from Brooklyn, New York. She currently serves as the Senior Director of SEO and Head of Organic Research at Amsive Digital ⁽¹¹²⁾, and has experience in machine learning, natural language processing, and web development. In addition to her technical expertise, Lity is also a DJ and drummer and actively participates in the open source community.

 Lily Ray, Sr. Director, SEO & Head of Organic Research at... https://www.amshvedigital.com/insights/suthor/hily-ray/
 Lilv Ray, Author at Search Engine Journal Lily Bay is a software engineer, data scientist, and SEO expert from Brooklyn, New York. She currently serves as the Senior Director of SEO and Head of Organic Research at Amsive Digital ^[112], and has experience in machine learning, natural language processing, and web development. In addition to her technical expertise, Lily is also a DJ and drummer and actively participates in the open source community.

The evidence for this comes from her Linkedin profile, where she lists her current position and past work experience, as well as her GitHub profile, where she lists her contributions to the open source community. Additionally, an article on the Amsive Digital website provides further information about her experience in the field of SEO.

1. Lily Ray, Sr. Director, SEO & Head of Organic Research at ...

Lily Ray is an SEO expert from Brooklyn, New York. She currently serves as the Senior Director of SEO and Head of Organic Research at Amsive Digital ^[11] ^[23]. She has experience in SEO, including strategic leadership and performance-driven insights. In addition to her technical expertise, Lily is also a DJ and drummer and actively participates in the open source community.

Evidence for this comes from her Linkedin profile, which lists her current position and past work experience, as well as an article on the Amsive Digital website that provides further information about her expertise in the field of SED. Additionally, she is featured in multiple interviews and articles discussing her career in SEO.

1. Lily Ray, Sr. Director, SEO & Head of Organic Research at... https://www.amsredgital.com/insights/author/lily-ray/ 2. Lily Ray: NYC Based SEO Professional, DJ, & Drummer M https://lywas.rsyc/ YOU



In-context learning

adding a prompt-chain

Is the following claim true or false given the context?

Context:

- Andrea Volpini is an Internet Entrepreneur and CEO of WordLift and Insideout10
- He has 20+ years of world-class experience in online strategies, digital media, SEO, and Artificial Intelligence
- He is passionate about SEO and Machine Learning
- He is a co-founder and CEO of WordLift and insideout10
- He has been involved in the development of numerous projects related to the Semantic Web, SEO, and AI
 He is a highly accomplished Internet entrepreneur

Claim: Andrea uses Instagram

Who is Andrea Volpini?

Andrea Volpini is an Internet Entrepreneur and CEO of WordLift and Insideout10 with 20+ years of world-class experience in online strategies ^{[1][2]}, digital media, SEO, and artificial intelligence. He is passionate about SEO and Machine Learning and is the co-founder of WordLift and Insideout10. He is also an expert in helping search engines find web content by adding semantic meaning to it and showing how it relates to other content ^[3]. Additionally, he has 1519 followers on Instagram and 1151 posts.

1. Andrea Volpini - Co-Founder and CEO - WordLift | LinkedIn https://www.linkedin.com/in/volpini/

2. Andrea Volpini (@cyberandy) / Twitter https://twitter.com/cyberandy?lang=en

3. Andrea Volpini: structuring unstructured content for the semantic web https://ellessmedia.com/csi/andrea-volpini/

FALSE STATEMENT

Is the following claim true or false given the context? Context: • Andrea Volpini is an Internet Entrepreneur and CEO of WordLift and Insideout10 • He has 20+ years of world-class experience in online strategies, digital media, SEO, and Artificial Intelligence • He is passionate about SEO and Machine Learning • He is a co-founder and CEO of WordLift and insideout10 • He has been involved in the development of numerous projects related to the Semantic Web, SEO, and Al • He is a highly accomplished Internet entrepreneur Claim: Andrea uses Instagram

The claim that Andrea Volpini uses Instagram is false. While Andrea Volpini does have social media accounts, the given context does not mention Instagram specifically. Additionally, a search of his social media accounts does not reveal an Instagram profile.

Who is Andrea Volpini?

Andrea Volpini is an Internet Entrepreneur and CEO of WordLift and Insideout10. He has 20+ years of world-class experience in online strategies, digital media, SEO, and Artificial Intelligence. He is passionate about SEO and Machine Learning, and is a co-founder and CEO of WordLift and Insideout10. Additionally, he has been involved in the development of numerous projects related to the Semantic Web, SEO, and AI, and is a highly accomplished Internet entrepreneur.









Self-ask

Who is the first female Prime Minister of Italy?

S The

The first female Prime Minister of Italy is not yet appointed. As of my knowledge cutoff, Italy has never had a female Prime Minister.





> Entering new AgentExecutor chain...

Yes.

Follow up: What is the name of the first female prime minister in Italy?

Intermediate answer: Giorgia Meloni was sworn in as Italy's first female prime minister. Her party's roots emerge from the ashes of Italy's fascist movement.

Follow up: When was Giorgia Meloni sworn in as Italy's first female prime minister? Intermediate answer: Giorgia Meloni sworn in as Italy's first female prime minister ... The ceremony took place after a week of tensions within Meloni's coalition ... on February 13, 2021.

So the final answer is: Giorgia Meloni

Finished chain. 'Giorgia Meloni'

> lets us plug in a search engine to answer the follow-up questions, which additionally improves accuracy.

Self-ask's structured prompting

"https://s.mj.run/30Ya-UkUfCA a portrait of the Italian Prime Minister Giorgia Meloni in the style of Robert Rauschenberg --v 4"







> Finished chain.



Question:

"What is the future of SEO?"

Follow up: What is SEO?

Intermediate answer: Search engine optimization is the process of improving the quality and quantity of website traffic to a website or a web page from search engines. SEO targets unpaid traffic rather than direct traffic or paid traffic.

Follow up: What trends are emerging in SEO? Intermediate answer: Major SEO Trends to Know for 2023

1. Voice Search Optimization 2. Mobile-First Indexing 3. Structured Data 4. Artificial Intelligence 5. User Experience 6. Video Content 7. Link Building 8. Local SEO

So the final answer is:

Voice Search Optimization, Mobile-First Indexing, Structured Data, Artificial Intelligence, User Experience, Video Content, Link Building, and Local SEO.

"self -ask" experiment

W

direct prompt

chain of thought

self-ask





Another issue with AI content:

It's a good practice to cite your sources in the content, whether through external links, superscripts, etc.

Users and search engines expect this.

...How are you supposed to add these with AI content? Manually reverse-engineer the source? 😕

We can do this by combining:

- 1. Web Scraping
- 2. Semantic Search (using FAISS)
- 3. GPT-3



<u>https://wor.ai/qa_demo</u>

Here is how it works:

INPUT

- 0 https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data
- 1 https://developers.google.com/search/docs/appearance/structured-data/sd-policies
- 2 https://developers.google.com/search/docs/appearance/enriched-search-results
- 3 https://developers.google.com/search/docs/appearance/structured-data/generate-structured-data-with-javascript
- 4 https://developers.google.com/search/docs/appearance/structured-data/search-gallery
- 5 https://developers.google.com/search/docs/appearance/structured-data/article

OUTPUT

Q. What are the benefits of using structured data for SEO ?

A. Schema can be used to help Google better understand a web page and its content, which can result in more visibility in search results.

Source:

https://developers.google.com/search/docs/appendix/overv
iew



<u>https://wor.ai/qa_demo</u>

- Over-reliance on large language models for SEO can lead to a number of negative effects. Some of these include:
- Lack of creativity: Using a language model to generate content can lead to a lack of unique or creative ideas. The model can only produce content based on the data it has been trained on, and may not be able to come up with new or original ideas.
- Lack of understanding of the audience: Language models are not able to understand the audience, which is a crucial aspect of SEO. Therefore, the content generated by the model may not be relevant or engaging to the intended audience.
- Lack of understanding of search algorithms: SEO is a constantly evolving field, and search algorithms are updated frequently. A language model may not be able to keep up with the latest changes and may produce content that is not optimized for current search algorithms.
- 4. Lack of human touch: While language models can generate large amounts of content quickly, they lack the human touch that is necessary to create truly high-quality content. This can lead to content that is not engaging or relatable to readers.

It is important to use language models as a tool to assist in the SEO process, but not to rely on them entirely. The best approach is to use them in combination with human expertise and understanding of the audience and the search algorithms.

Don't forget you are an SEO

(...not just an AI nerd).



control #1

**** Trustpilot

MENU

Q 👷

 \equiv

MENU

Free Shipping & Easy Returns! Explore Maui Jim Optical MJO2207 Eyeglasses now.

About Maui Jim Eyewear

Maui Jim has led the eyewear industry by design and innovation since 1980, offering eyes everywhere superior protection from the harmful rays of the sun in its popular line of sunglasses. From iconic shades with patented technologies, to readers perfect for sunny-day pastimes, to the impressively stylish and comfortable daily-wear frames in the Optical line,

control #2

★★★★★ ★ Trustpilot

Q 눻

Free Shipping & Easy Returns, plus Price Match! Explore Lulu Guinness L176 Sunglasses now and enjoy the latest promotions on Lulu Guinness frames.

Lulu Guinness and sunglasses feature a stylish and timeless look that allows you to express your individuality. Inspired by Lulu Guinness herself, this collection of eyewear includes iconic designs which combine elements of vintage fashion with contemporary glamour. Lulu Guinness eyewear is designed as a fashion accessory used to complement your unique style.

See all Lulu Guinness Sunglasses

variant

**** Trustpilot

Q 🛱

Free Shipping & Easy Returns, plus Price Match! Explore Carrera CA1034/S Sunglasses now and enjoy the latest promotions on Carrera frames.

MENU

Reach for a pair of Carrera CA1034/S sunglasses and enjoy the classic aviator profile in a form-fitting design that's flattering on most facial shapes. The metal rims offer both heft and flexibility, while plastic earpieces provide sturdy construction at an economical weight. With double bridges and rivet accents adding visual interest to these chic shades, you can go from street to skies without missing a beat.

A retail website in the US

How do we measure success?



- 5 batches of PDP: 500 variant and ~ 500 control.
- In control, we always have some text (generic or brand related).
- Year-over-year performance is confirmed over time, across all batches.

How do we find the best prompt?

Completion



Completion

A summer spring jacket from ALYSI means two things. The first one being silk, the second one being green. A summer spring jacket from ALYSI means two things. The first one being silk, the second one being green.



CPT-3

We can compare synthetic vs. original images



How does content validation work?



IS THE INFORMATION AVAILABLE IN THE TEXT?

IS THE AVAILABLE INFORMATION CORRECT?

IS THE TEXT VALID IN TERMS OF SEO AND CONTENT GUIDELINES?



Is the information available and valid?

existing

Shop Pre-Owned Prada 1990s Tinted Rounded Sunglasses In Gold, starting at \$209. Similar ones also available. On SALE now! Brown/gold-tone gold plated metal 1990s tinted rounded sunglasses from Prada Pre-Owned featuring yellow tinted lenses, oval frame, curved tips, nose pads and tortoiseshell detailing. Please be mindful that this piece has led a previous life, and may tell its story through minor imperfection. Purchasing this item continues its narrative, so you can be confident that you're making a POSITIVELY CONSCIOUS choice for the planet. We know you'll look after them, but these glasses come with protection, just in case.

==> Attributes not found:product name, lens material, frames material,

product tvp

==> Detected attributes: lens color, brand,

==> All attributes have correct values.

gender, temples color

new (A)

Prada PRD3700 sunglasses are a bestseller that continues to stand the test of time. This classic shape has been worn by cultural intellectuals who have changed the course of the world. With gold temples and yellow lenses, you'll always be one step ahead. Lenses are polarized to keep your eyes protected and safe from harmful UVA and UVB rays.

==> Detected attributes:lens_color, brand, product_name, product_typ
==> All attributes have correct values.

==> Attributes not found: lens material, frames material, gender

new (B)

Prada PRD3700 eyeglasses are a bestseller that continues to stand the test of time. This classic shape has been worn by cultural intellectuals who have changed the course of the world. With gold temples and <u>blue lenses</u> you'll always be one step ahead. <u>Lenses</u> are polarized to keep your eyes protected and safe from harmful UVA and UVB rays.

==> Detected attributes:lens color, product name, product_type ==> Attributes with wrong values:lens_color, product_type ==> Attributes not found:lens material frames material gender



Validate



Question Is ChatGPT going to kill Google?

Content

No, ChatGPT will not kill Google. ChatGPT is a natural language processing tool that helps people create conversations with machines, while Google is an online search engine.

Here is a statement
{statement}

- ChatGPT is a natural language processing tool
- Google is an online search engine
- ChatGPT helps people create conversations with machines
- ChatGPT and Google are two different services
- ChatGPT will not kill Google

1. True 2. True 3. True 4. True 5. False - ChatGPT and Google are two different services, so ChatGPT will not have any impact on Google.	Here is a bullet point list of assertions: {assertions} For each assertion, determine whether it is true or false. If it is false, explain why
No ChatGPT is not going to bill Google ChatGPT is a chathot service that provides	In light of the shove

No, ChatGPT is not going to kill Google. ChatGPT is a chatbot service that provides automated conversations, while Google is a search engine and technology company. ChatGPT has the potential to provide helpful services to Google, but it is unlikely to replace Google services.

In light of the above
facts, how would you
answer the question
{facts}

> Finished chain.

Answer

Can we validate TOV and grammar?



We currently have both gated video and youtube video content from previous events that is hard to find on search:

 Can we let search engines see these gated videos?

2.

Can we **extract** relevant content from these videos to boost traffic and optimize YouTube content?

OPTIMIZING VIDEO CONTENT

Keynote Interview: Bill Gates and Roula Khalaf



Can we reuse this asset more effectively?

Transcribe the video	Process the transcription	Video SEO for YT and Website
The video is processed using a state-of-the-art (SotA) model in speech-to-text .	The text is analyzed to extract a short summary and a set of relevant questions with answers and links to video segments.	The generated content can help us improve: the metadata on YT , the markup on the webpage , and the overall UX (we can add content on the webpage).
	W	

SOURCE VIDEO

Keynote Interview: Bill Gates and Roula Khalaf



SHORT SUMMARY (unedited)

Bill Gates proposes a framework, the green premium, to reduce the 51 billion tons of greenhouse gases added to the atmosphere each year to zero. The green premium is the extra cost of doing activities in a green way with no emissions. To eliminate the green premium, Gates suggests putting a price on carbon and investing in new green technologies. Carbon pricing is an important tool to send a price signal to work on the hard stuff, but it is not a one-stop solution.

Q&A + VIDEO SEGMENT (unedited)

Q: What should businesses do to reduce emissions and address the green premium and gold standard offsets?

A: Three quick and cost-effective ideas for businesses to reduce emissions and address the green premium and gold standard offsets include increasing competition, eliminating the green premium, and investing in renewable energy sources.

Link to segment:

https://www.youtube.com/watch?v=tjupXBKOb Xs&list=PLLQ51Tq7D-Z5hLqIqk0fT7XK3KoXG247 k&index=3&t=965s



Learnings

- LLMs are unsafe, biased and racist (among other well-known limitations).
- 2. The **feedback-loop** is king. You need a magic loop that unleashes human creativity and blends it with Generative Al.
- 3. You need a **data fabric**. Your data, your content and your unique messaging are essential.
- 4. **Emerging behaviour of LLMs** open up a world of possibilities.
- 5. We have to play it safe and **invest** on a robust validation pipeline.



Ψ

Thank You!



iPullRank's Machine Learning for Marketers





The SEO Weekly



BROUGHT TO YOU BY



https://ipullrank.com/resources/seo-weekly



References

- 1. [colab] A quick introduction to CLIP and Multimodal Search
- 2. [space]Point-E demo: text to 3D -Hugging Face
- 3. [colab] A quick introduction to Transformer Language Models
- 4. [colab] Self-Ask implementation
- 5. <u>[demo] Ouestion & Answering from</u> <u>Web Documents</u>
- 6. [colab] Validating Al-Genereated product descriptions
- 7. [Article] Generative Al For SEO: An Overview
- 8. [Article] ChatGPT for SEO: Leverage Generative AI To Outperform Your Competitors
- 9. [Article] Top 5 Best Plagiarism Checkers For Al-Generated Content



