Guide

A Content Marketer's Guide to Natural Language Generation

MarketMuse

This guide provides content marketers with an overview of the state of natural language generation (NLG). You'll learn what natural language generation is, the benefits of NLG, the challenges, some use cases, and a recent breakthrough in creating long-form content with the help of NLG.

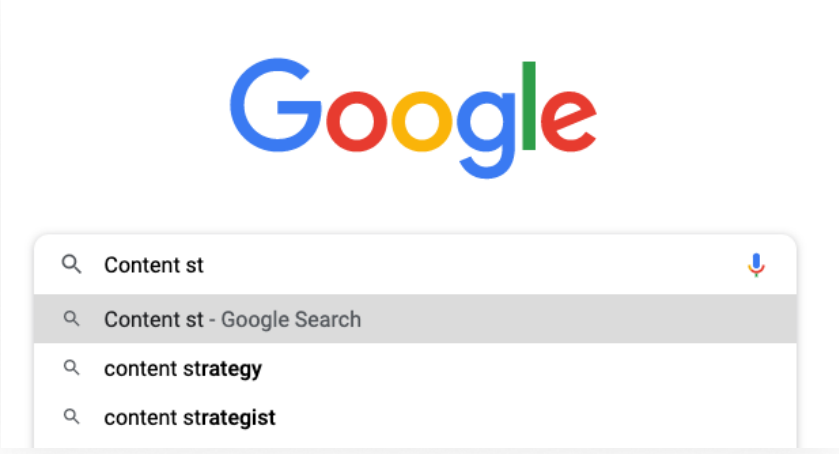
What is Natural Language Generation?

Every time your smartphone talks to you or Gmail auto-finishes a sentence, you're using natural language generation (NLG). Through this process, an app or program takes raw data and turns it into a coherent response that sounds natural to the human ear.

NLG typically works from structured data analysis. Historically, we've seen this technology at work in chatbots, smart devices, and search engines, but advancements have made it accessible for a wider audience, including content marketers.

Recent developments in NLG free the technology from the constraints of structured data. Text generation software can now analyze a collection of articles on a given subject and transform that raw data into an original document.

The document you are reading was created with the help of natural language generation working in this manner.



NLG at work

How NLG works

In simple terms, current NLG technologies, like GPT-3, are recursive text predictors. Give the model some text, and it will determine the next most likely word or phrase. Then it repeats the process using the original input along with the newly generated text.

These deep learning models are self-supervised and don't require any labeled data. They can learn from vast amounts of any text. Deep learning aims to significantly improve the accuracy and completeness of the language model. It uses a combination of more powerful algorithms and tools to train a neural network, which simplifies learning strategies using the user's input.

NLG vs. NLP vs. NLU

There are many moving parts in the AI and machine learning process. Three you probably hear about a lot are natural language processing (NLP), natural language understanding (NLU), and natural language generation (NLG).

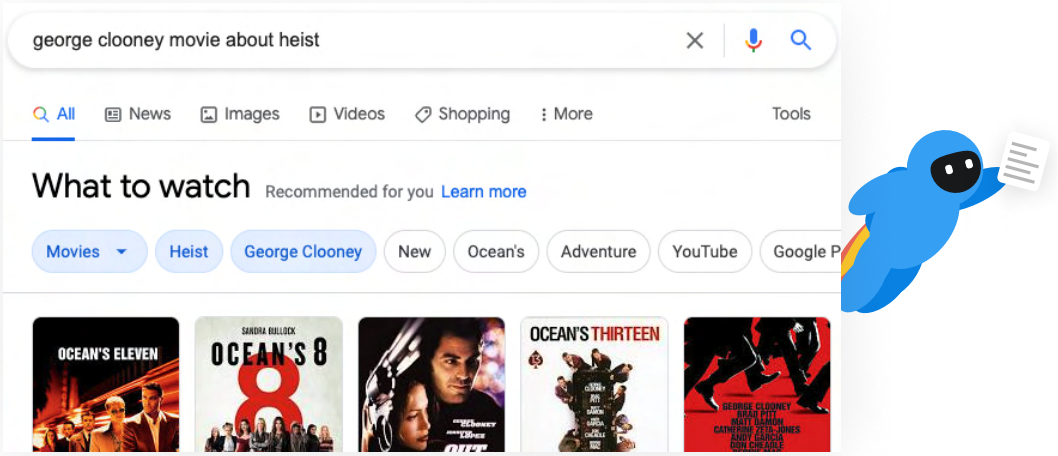
These terms are often confused because they're all part of the singular process of reproducing human communication in computers.

Let's take a look at what these terms mean and how they (and AI as a whole) can help marketers streamline their content marketing strategy.

What is Natural Language Processing?

Natural language processing (NLP) is a branch of artificial intelligence that designs algorithms to process natural language data. The three primary processes involved in NLP are speech recognition, natural language understanding (NLU), and natural language generation (NLG). Think of all the search fields you fill out to find what you need. Amazon, Netflix, Google; they all use autofill to figure out what you're searching for even before you finish typing it.

And how many times have you sort of known the name of a movie or product you're looking for? You plug in what you know and hope the search functionality can match your partial search term with the thing you're searching for. As long as your query is better than "that movie with that guy," there's a good chance Amazon or Google will find it. That's natural language processing.



NLP in action

What is Natural Language Understanding?

Natural language understanding is a smaller part of natural language processing. Once the language has been broken down, it's time for the program to understand, find meaning, and even perform sentiment analysis.

The program breaks language down into digestible bits that are easier to understand. It does that by analyzing the text semantically and syntactically.

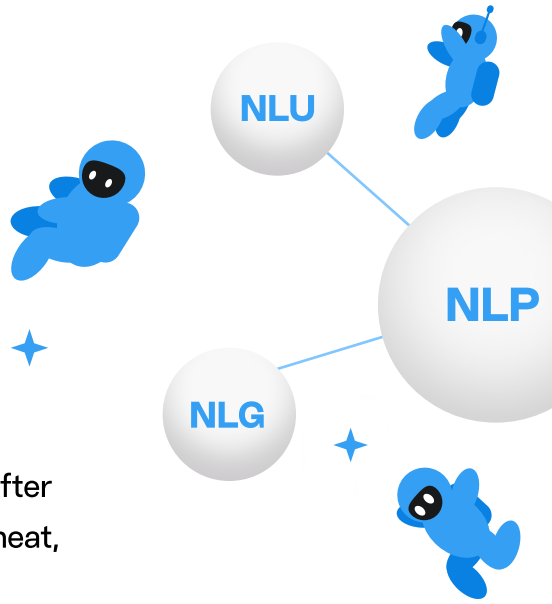
Unlike structured data, human language is messy and ambiguous. As a species, we're rarely straightforward with our communication. Grammar and the literal meaning of words pretty much go out the window whenever we speak.

In fact, "out the window" is a great example. We, of course, didn't mean that we throw things out a literal window, especially since we're talking about intangible concepts rather than solid objects.

And so it is when you ask your smart device something like “What’s I-93 like right now?”. If you were being literal, you might get an answer like, “It’s long, gray, and has cars driving on it.” But you probably wanted to know what the traffic conditions are.

That’s where natural language understanding comes in. It’s taking the slangy, figurative way we talk every day and understanding what we truly mean.

How NLP, NLU, and NLG Work Together



NLP, NLU, and NLG all play a part in teaching machines to think more like humans. They simply tackle different parts of the conversational AI problem. Let’s take a specific example to illustrate just how these functions work together.

You get home from work and wonder how your stocks did today. After you tell your smart device to turn on your lights and crank up the heat, you ask, “Hey Google, how did the stock market do today?”

Your Google Home device listens to your query, and then natural language processing kicks in. It takes your question and breaks it down into understandable pieces — “stock market” and “today” being keywords on which it focuses.

Then it compares your query to similar queries made to Google in general and tries to understand what you’re asking. That’s natural language understanding.

Once it understands that you want to know the closing numbers for NASDAQ, Dow Jones, and the S&P 500, it crawls the web for content that best answers your question.

Once it has data from a reliable source, like Bloomberg, it pulls the relevant data and delivers it in plain language. That’s natural language generation. Its answer is something like, “According to Bloomberg, the NASDAQ was down 1.5 points, but the Dow was up 77, and the S&P was up 5 points.”

Where is Natural Language Generation Used in Content Marketing?

Artificial intelligence is changing the way we plan and create content. More importantly for content marketers, it's allowing teams to scale by automating certain kinds of content creation and analyzing existing content to improve what you're offering and better match user intent.

Using structured data, NLG can scale content exponentially by creating plug-and-play content from templates and reducing manual workload, saving both time and money in a major way. Use-cases for this templated style of content include:

- **Product descriptions**
- **Real estate listings**
- **Earnings reports**
- **Sports game scores**
- **Election results**
- **Business intelligence data**

NLG can also analyze text to produce short summaries, allowing marketers to quickly generate bite-sized content for social media posts, product descriptions, email copy, and even short-form blog posts. Previously, NLG was restricted to situations in which structured data was available, but that's changing with recent advancements that enable NLG to mimic the way humans write and create long-form content.

Helping others improve their yoga practice

Yoga practice gives you the opportunity to introduce mindfulness and meditation, to improve your connection to yourself, and to learn to forgive yourself. Happily, you don't have to become a yogi to practice mindfulness and meditation. You can practice yourself while seated in a yoga pose, for example by sitting with your right arm over your right knee while lying on your grass. Breathing is also

We asked MarketMuse's NLG (First Draft) to tell us a little about Yoga.

NLG tools and applications

GPT-3 API

GPT-3 is a machine learning model from OpenAI, a US-based artificial intelligence research lab, that uses deep learning to generate and predict natural language. GPT stands for generative pre-trained transformer. The three means it's in its third generation.

There are other tools available to help marketers and content strategists develop both long and short-form content, create narratives from structured data, and convert text to speech. This past year has seen a number of new products hit the market, driven by the availability of the GPT-3 application programming interface (API).

AI Not API

It's important to understand the limitations of relying on an AI platform that uses an API. There's little to differentiate platforms that use the same API. If you're unhappy with the results from one platform that uses GPT-3, you'll find little satisfaction moving to another one that also uses GPT-3.

The problem with using an API is that the platform is restricted to working with the API's output. There's nothing the platform can do to improve that output as that control rests with the API provider. On the other hand, using an AI platform not driven by an API allows for continuous improvement, customization, and flexibility in response to customer feedback.

Text Generation

For our purpose, we're defining anything less than 750 words as simple text generation. There are certain situations where a shorter-form narrative is more appropriate — email and ad copywriting, for instance. Tools in this category aim to help publishers with a few things: email subject lines, headlines, CTAs, mobile app push messages, social posts, AdWords copy, landing pages, display ad copy, and even short-form web articles.

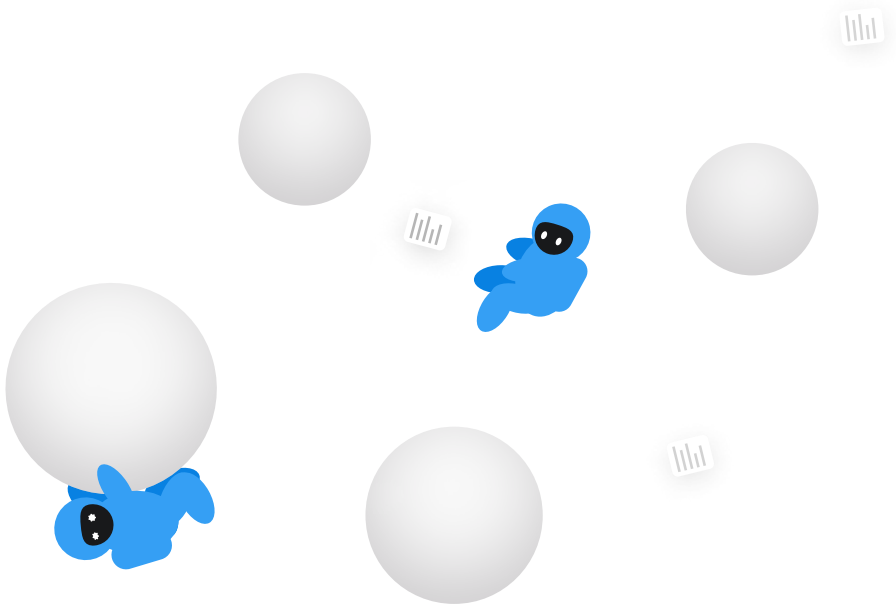
Structured Data Narrative

Applications in this category take sets of highly structured data and turn them into a narrative. The Associated Press produces nearly 4,000 company earnings articles quarterly with the help of artificial intelligence. Ecommerce sites can also create product descriptions, category stories, and newsletters using this method. There are numerous use cases for this approach, as long as you have the structured data to support it.

Text to Speech

Text to speech converts written text into natural-sounding audio in a variety of languages. This can be used in chatbot and voice assistant interaction, turning digital ebooks into audiobooks, and interacting with in-car navigation systems.

Recently, companies have been using deep neural networks to synthesize speech that's nearly identical to human recordings. Human-like speech patterns, intonation, and articulation significantly reduce listening fatigue when interacting with AI systems. A handful of well-known organizations dominate this area including IBM, Microsoft, Amazon, and Google.



Benefits and risks of NLG

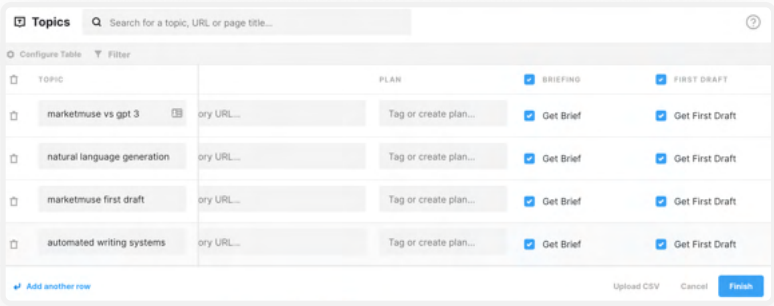
How Can Content Marketers Benefit from NLG?

Speed, agility, and productivity are where content marketers can benefit most from NLG. Need 100, 1,000, or 10,000 content items by week's end? No problem! NLG applications are infinitely scalable.

Need to switch content production to support a different campaign? Easy! Natural language generation systems are subject matter agnostic. You won't waste time trying to put together the right team.

Are you struggling to balance the workload among your content creation team and maximize throughput? Not an issue with NLG platforms.

If you have data that needs to be converted into a story, natural language generation can help.



Remove the bottleneck around content creation with NLG

Natural language processing allows marketers to do more with less, producing content that is as human as possible. It brings about a smoother workflow, and the content creators can focus more on the development and execution of their content.

The benefits that organizations experience in using NLG are many and continue to grow. The ability to rapidly produce content at a sufficient level of quality not only saves money but allows coverage of topics that otherwise would be ignored. Done properly, more NLG content means more opportunities to connect with a human audience.

What are the Challenges of Natural Language Generation in Content Marketing?

Until now, content marketers have been limited in where they can employ natural language generation. They've been restricted to situations where structured data is available. Using natural language generation to help build free-form text, like an article or blog post, has so far been unworkable.

GPT-3-driven applications have helped popularize NLG in the content marketing sphere. But there are some limitations:

Creating a long-form article using GPT-3 is inefficient

The Guardian wrote an article in September of 2020 with the title A robot wrote this entire article. Are you scared yet, human? The pushback by some esteemed professionals within the AI world was immediate. That's because the AI had a heavy helping hand from its human counterpart. The (human) editor had to piece together 8 different 500-word essays to come up with something that was fit to be published. There's nothing efficient about that.

GPT-3 lacks full comprehension

Those who took a closer look at GPT-3 found the smooth narrative was lacking in substance. As Technology Review observed, "although its output is grammatical, and even impressively idiomatic, its comprehension of the world is often seriously off."

```
Q: How many eyes does a blade of grass have?
A: A blade of grass has one eye.
```

GPT-3 response to a question prompt.

To test how comprehensive an article GPT-3 could produce, we ran an analysis on the Guardian article ourselves using MaretMuse's Optimize tool to determine how well it addressed the topics that experts mention when writing on this subject. The results were less than stellar.

GPT-3 is NSFW

60% of GPT-3's training data comes from the Common Crawl dataset. This vast corpus of text is mined for statistical regularities, which are entered as weighted connections in the model's nodes. The program looks for patterns and uses these to complete text prompts.

As TechCrunch remarks, "any model trained on a largely unfiltered snapshot of the internet, the findings can be fairly toxic."

OpenAI researchers did their own investigation on fairness, bias, and representation concerning gender, race, and religion and found inherent bias in the content that was produced.
Having been trained on uncurated Internet data, GPT-3 output can be embarrassing, if not harmful.

GPT-3 produces low-quality content

We did an experiment over on our blog, taking a 2,400-word article produced by GPT-3 on the topic of Twitter and ran it through MarketMuse's quality assessment and optimization tools. Our results showed us a low Content Score and a lack of Topic Authority with zero mentions of critical topics like "social media," "tweets," "Twitter followers," or "trending hashtags."

While the article was coherent, it lacked structure and any inherent meaning. GPT-3 is pretty good at generating content that looks nice but lacks substance. There's a word for this type of article. It's called "fluff."

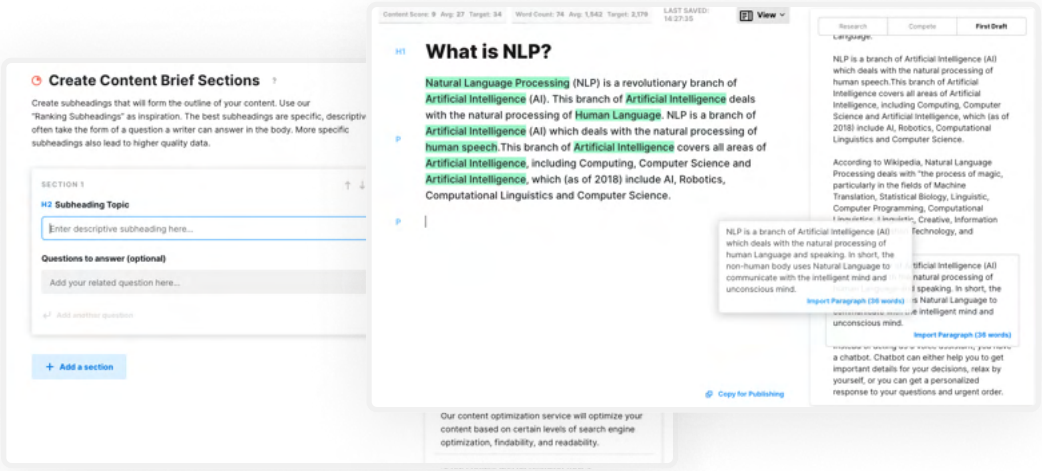
MarketMuse's First Draft

MarketMuse First Draft is an AI-powered content generation platform that removes the bottlenecks companies face when creating high-quality content at scale. Content teams can build complete customer journeys and tell their brand stories faster by using the AI-generated, editor-ready drafts of content.

Unlike GPT-3, First Draft puts additional guardrails in place to better train the model, fine-tune the output, and provide more content structure resulting in a better narrative. With MarketMuse First Draft, marketers can now take advantage of NLG to produce long-form content.

How does First Draft work?

MarketMuse First Draft is one of the only NLG technologies that can generate long-form articles (up to 5,000 words in length). MarketMuse Content Briefs govern the output, so drafts are coherent and structured. Users can build their own MarketMuse Content Briefs before ordering a First Draft. During this stage, they can specify the topics the article should mention, questions it should answer, and what sections make up the article.

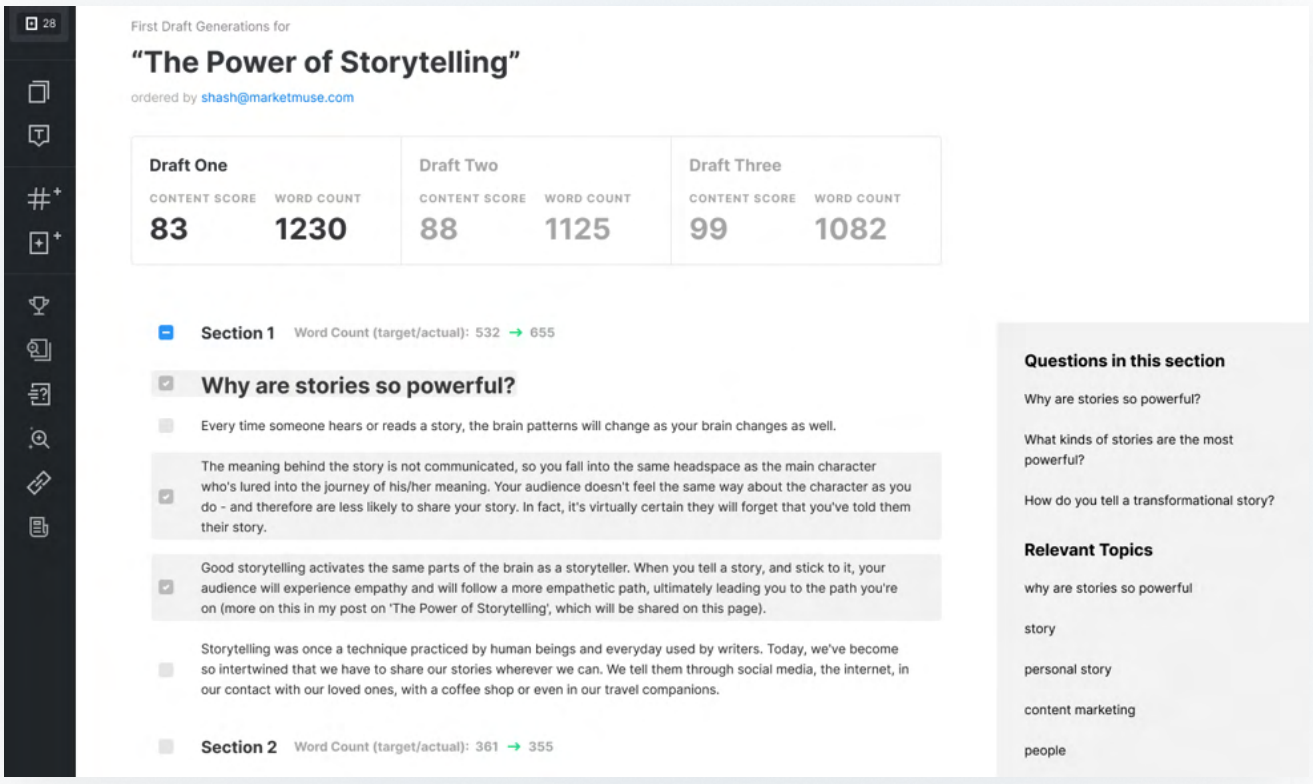


Sample MarketMuse Content Brief Order and First Draft Output

First Draft is connected to the rest of the MarketMuse Suite. It connects content planning, research, and production in one place. Technology enables each step of the process. Unlike other solutions, you won't waste time trying to bring multiple software vendors together and make it work seamlessly.

How it works

1. Use the Inventory to research a topic you want to write about.
2. Use our Brief builder to structure the content however you like.
3. Click “Get First Draft.”
4. Our AI generates a draft of content that’s ready by the next day.
5. Access your First Draft from the menu and edit it directly in our text editor.



MarketMuse First Draft vs. GPT-3

GPT-3 is a solution in search of an application. The only way to access the API is to join a waitlist in which your use case must be described. Even with access, you'll still be limited in using what's provided through the Application Programming Interface.

MarketMuse First Draft was created to solve a specific use-case—generating long-form SEO-quality articles for content marketers. If you're thinking of using GPT-3 in this manner, you'll be disappointed. Here's a comparison between the two offerings:

	MarketMuse	GPT-3
Coherence & Structure	Output dictated by Content Brief, drafts are structured from the start.	Starts with prompt text without guardrails, leading to unstructured output.
Control	Create your own brief specifying key topics, sections and questions to answer.	Users have little control over topics and questions answered by GPT-3 geneartions.
Configuration	Trained on articles from a curated dataset (excluding sexist, racist, adult content).	Trained on the entire web, including low-quality, explicit and hateful content.
Degredation & Repetition	Produces text free of degredation, plagiarism and repetition.	No checks for degedation, plagiarism or repetition.
Speed to Publish	Output can be edited into publication-ready contnet in 1-2 hours.	Output takes several hours to be edited into publication-ready content.

The data used to train a natural language model plays a critical role in differentiating First Draft from GPT-3 and other competitors. MarketMuse is very selective in the data it uses for training its natural language generation model. We have very strict filters to ensure clean data that avoids bias regarding gender, race, and religion.

Also, our model is trained exclusively on well-structured articles. We're not using Reddit posts or social media posts and the like. Although we're talking millions of articles, it's still a very refined and curated set compared to the amount and type of information used in other approaches.

MarketMuse First Draft Benefits

First Draft was created to solve a specific challenge; how to help content teams produce better content faster. It's a natural extension of our already successful AI-powered Content Briefs. Clients will be able to review a Content Brief and order a MarketMuse First Draft based on the criteria, structure, recommended questions to answer, and topics in the Content Brief.

Our NLG technology creates long-form content without using templates. Unlike existing NLG models that can only produce less than 1,200 words before serious quality degradation, MarketMuse First Draft doesn't suffer from this restriction. We've trained our model on articles from a carefully curated dataset (that excludes sexist, racist, and adult content) to improve the outcome of generations.

Your content, your voice

First Draft can be configured to match your writing style. It can also emulate the style of an author or publication you like.

Avoid common pitfalls with AI-generated text

First Draft produces long-form content without plagiarism, repetition, or degradation of quality. Each draft is unique and original and doesn't simply extract text fragments from other documents.

Write authoritatively on any topic

Get drafts of content on any topic so you can support all of your products and lines of business, no matter what they are.

As much content as you need, when you need it

Whether you're starting content from scratch, or want to ramp up publication to beat your competitors, First Draft can help you reach your aspirational publication tempo.

Empower writers and editors

MarketMuse First Draft helps writers craft better content quicker. But it doesn't replace them. While First Draft will reduce the amount of time spent on research and writing, we still recommend editing to enhance the narrative. You may want to enrich your content using images, infographics, pull quotes, video, additional links, step-by-step instructions, promotional language, and your own experiences.

Scale content without scaling costs

First Draft lets you scale content production without scaling costs and complexity. Keep your content costs predictable and your quality consistent by letting AI do the work of getting you a strong initial draft.

MarketMuse

Publish better every time.

Let us show you how MarketMuse helps thousands of content teams turn content into a growth channel.

Schedule a Demo