CONTENTSTACK

Brand Kit: Implementation Strategy

Considerations for customers when setting up AI integration

Glossary

Contentstack Brand Kit: The application within Contentstack that utilizes AI to generate on-brand content. It allows users to create brand-specific profiles and access relevant company information in order to create consistent, on-brand content with AI.

Brand Kit folder(s): Within the Contentstack Brand Kit application, a Brand Kit folder acts as a container that stores all the information specific to a particular brand or sub-brand. This includes Voice Profiles, Knowledge Vault data, and any other settings related to content style and governance for that brand.

- A single Brand Kit folder can be associated with one or more Contentstack stacks.
- Contentstack stacks can be associated with one or more Brand Kits.
- One Brand Kit is always attached to one Knowledge Vault.

Voice Profiles: These profiles define the desired writing style, tone and communication preferences for a brand or individual user. By customizing Voice Profiles, users can instruct the AI to generate content that aligns with their specific voice and messaging goals.

Knowledge Vault: This is a vector database that stores a curated collection of factual information specific to a company. The Knowledge Vault serves as a trusted source for the AI, allowing it to generate content that is not only tonally consistent but also grounded in accurate and relevant company data.

Generative AI (GenAI): This refers to a type of artificial intelligence that can create new content, such as text, images or code. The Contentstack Brand Kit leverages Generative AI to automate content creation while maintaining brand consistency.

Retrieval-Augmented Generation (RAG): This is a specific technique used by Contentstack's Brand Kit (as well as other organizations), that was originally invented by Meta (formerly Facebook) a few years ago. RAG involves retrieving relevant information

from the company's Knowledge Vault and using it to augment the content generation process. This ensures that the AI-Generated content is factually accurate.

Set up strategies

What is a Knowledge Vault?

A Knowledge Vault is a vector database designed to store and retrieve information in a way that's optimized for use in AI language models.

Unlike traditional databases that store text as is, a vector database converts text into numerical representations called "embeddings" or "vectors."

These embeddings capture the semantic meaning and relationships between words and concepts. This allows AI to understand the context and nuances of the information stored within the Knowledge Vault.

How is a Knowledge Vault used for Retrieval Augmentation

When a user (or a program, like Automate) asks for a content generation task to be accomplished (such as asking to create a product description), the AI first sends that query to the Knowledge Vault.

The Knowledge Vault then finds the most relevant pieces of information in the database based on the semantic similarity between the query and the stored embeddings.

These relevant pieces of information are then added to the query, along with other information (such as a Voice Profile), which creates the prompt which will be sent to the AI model. Having this accurate background information in the prompt serves to "ground" the AI in factual data specific to the company, which drastically improves the relevancy and accuracy of the AI"s generated content.

What content should I include in my Knowledge Vault?

The best content for a Knowledge Vault focuses on factual, company-specific information. Generalized information is probably already available in the training of the language model, and would not need to be included. Here are some examples of content that works well in a Knowledge Vault:

- **Company Background:** History, mission statement, core values, brand guidelines, style guidelines

- **Products and Services:** Descriptions, technical specifications and use cases

- **Industry Knowledge:** Market reports, whitepapers, competitive analysis

- **Customer Data:** FAQs, support queries and answers

- **Documentation:** Training materials, sales presentations

- **Marketing Data:** Marketing messaging, strategic overviews

What kind of content should I avoid including in my Knowledge Vault?

Consider various content types and sections you have on your site, and consider whether they would be good sources of truth for AI:

- **Irrelevant or Outdated Data:** Your documentation from a discontinued product or an expired promotion may not be useful when generating content about new products and may cause confusion.

- **Old or off-tone information:** Blog posts from 2014 when your company was completely different, may set the wrong tone of voice when used as example content.

- **Confusing data:** Including a competitive overview of another company's product may be incorporated as background information and will then cause AI to write incorrect information.

- **Sensitive data or confidential information:** Do not include information that you do not want to be included in your generated text

- **Mislabeled or unlabeled data:** If it is not clear to the AI what it is, it may cause the AI to generate incorrect data.

How many Brand Kits does a customer need?

Having separate Brand Kits allows an organization to maintain organized folders for each of their separate brands, however, most importantly, it allows the various Knowledge Vaults (which is a source of company "truth") to be separated. Since retrieval augmentation is based on a similarity search

- **Single Brand:** If a company has a well-defined, singular brand identity, a single Brand Kit would likely suffice.

- **Multiple Brands or Sub-Brands:** If the company manages distinct brands that all share a contentstack organization, they may require multiple Brand Kits in order to keep the data, such as the Knowledge Vault, separate.

- **Regional or Market Variations:** For brands operating across diverse markets or regions with localized variations in tone or messaging, separate Brand Kits might be necessary.

Why can't I mix up content from different companies or areas?

- **Brand Confusion:** Mixing up content from different companies creates conflicting information in the vector database. This means that chunks of content from one source may be included in a prompt for another company, causing the AI to generate incorrect or confusing content.

- **Data Inaccuracy:** Including data from other companies or unrelated areas introduces irrelevant information which leads to inaccurate or misleading generated content.

- **Loss of Control:** When the source of data is mixed, you lose the ability to ensure the quality, relevance, and alignment of the information used to ground the AI in truth.