

OpenAI Prompt Engineering for Generating Product Image from Customer Reviews

PART 3. Image Generation with Diffusion Model

For each product, based on the features extracted from the product description and customer reviews, we used OpenAI's DALLÉ to generate images.

1. Dinosaur Plushie:



Fig 4: Original Product Image

Baseline Prompt (Prompt 1):

*Based on the textual descriptions and reviews provided for the VHYHCY Stuffed Dinosaur Plush Toy, here is a summary of the visual information to create an accurate image of the product:\n\n- **Color**: Active Green\n- **Material**: Made of high-quality polyester fiber and filled with PP cotton\n- **Design**: Features a cute dinosaur design with a cute expression, suitable for kids and dino lovers\n- **Size**: 9 inches in length, described as mini and cute, perfect for both boys and girls\n- **Accessories**: Some mentions of a little hat on the plushie, adding a precious touch to the design\n- **Texture**: Described as soft, squishy, and comfortable to hold, with a smooth feel and delicate texture\n- **Features**: Can be used as a decorative item in family bedrooms, cars, etc.\n- **Concerns**: Some reviews indicate issues with seam quality,*

slight threading coming loose, small size compared to expectations, and a vacuum-sealed packaging leading to a need for repairs.\n\n*To create an accurate image of the product, you can focus on showcasing the cute dinosaur design in active green color, with attention to details like the little hat, soft and squishy texture, and the compact size suitable for children. Paying close attention to the quality of seams and overall presentation can help in capturing the essence of the product accurately.*

When we first fed the above prompt which is the output of the previous step directly to DALL-E, we got the below output:



The plushie image created has captured some features of the original product image.

- Green color Dinosaur Plushie - Although the green color isn't entirely accurate because of the vagueness in the specified "Active Green" color in the product description.
- Material, Shape and Size

The main differences from the original image are:

- It added a smile to the dino's face which isn't present in the original product. This could be due to the mention of "Cute expression" in the prompt
- No party hat - The prompt did mention a hat however this wasn't added in the image
- The hand with the paintbrush - We are unsure why DALL-E chose to add this in the image
- Feature image and description bubbles

Follow-up Prompt:

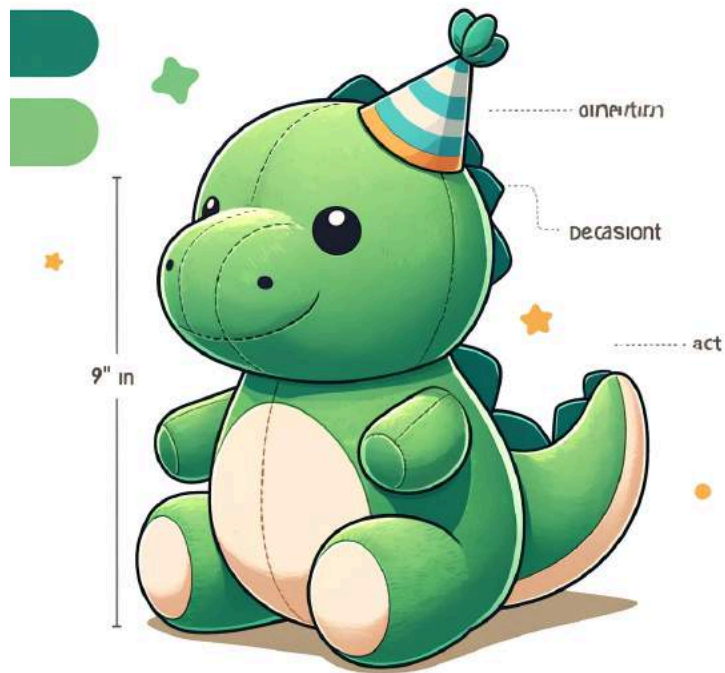
Make the plushie wear a party hat

When the above prompt was added, the model created a 2-D drawing of the dino plushie in a party hat which significantly differs from the previous output and from what we're actually expecting. It continued to keep the paint brush and hand for some odd reason.



Prompt 2:

*“VHYHCY Stuffed Dinosaur Plush Toy, here is a summary of the visual information to create an accurate image of the product:\n\n- **Color**: Active Green\n- **Material**: Made of high-quality polyester fiber and filled with PP cotton\n- **Design**: Features a cute dinosaur design with a cute expression, suitable for kids and dino lovers\n- **Size**: 9 inches in length, described as mini and cute, perfect for both boys and girls\n- **Accessories**: Some mentions of a little hat on the plushie, adding a precious touch to the design\n- **Texture**: Described as soft, squishy, and comfortable to hold, with a smooth feel and delicate texture\n- **Features**: Can be used as a decorative item in family bedrooms, cars, etc.\n- **Concerns**: Some reviews indicate issues with seam quality, slight threading coming loose, small size compared to expectations, and a vacuum-sealed packaging leading to a need for repairs.\n\nTo create an accurate image of the product, you can focus on showcasing the cute dinosaur design in active green color; with attention to details like the little hat, soft and squishy texture, and the compact size suitable for children. Paying close attention to the quality of seams and overall presentation can help in capturing the essence of the product accurately. **Also add a colorful party hat.**”*



Follow-up Prompt:

"Don't make a sketch, make it more realistic"



Prompt 3:

*“Create an image for this product- VHYHCY Stuffed Dinosaur Plush Toy, here is a summary of the visual information to create an accurate image of the product:\n\n- **Color**: Active Green\n- **Material**: Made of high-quality polyester fiber and filled with PP cotton\n- **Design**: Features a cute dinosaur design with a cute expression, suitable for kids and dino lovers\n- **Size**: 9 inches in length, described as mini and cute, perfect for both boys and girls\n- **Accessories**: Some mentions of a little hat on the plushie, adding a precious touch to the design\n- **Texture**: Described as soft, squishy, and comfortable to hold, with a smooth feel and delicate texture\n- **Features**: Can be used as a decorative item in family bedrooms, cars, etc.\n- **Concerns**: Some reviews indicate issues with seam quality, slight threading coming loose, small size compared to expectations, and a vacuum-sealed packaging leading to a need for repairs.\n\nTo create an accurate image of the product, you can focus on showcasing the cute dinosaur design in active green color, with attention to details like the little hat, soft and squishy texture, and the compact size suitable for children. Paying close attention to the quality of seams and overall presentation can help in capturing the essence of the product accurately. Also make sure the dino is wearing a colorful party hat. Make this image realistic.”*



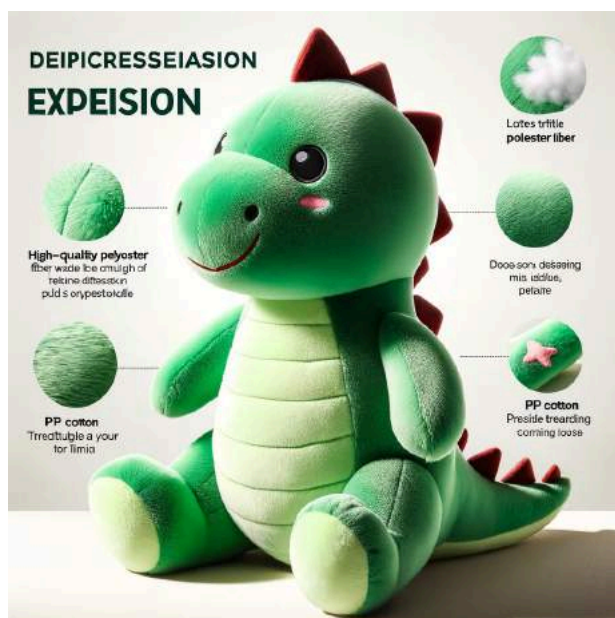
When we added the follow up questions into the main prompt itself, we got a product image which was a much better version of the previous one. The party hat was now captured in the first prompt itself and it didn't display any hand and paintbrush this time. We however have random

image bubbles with text attempting to describe the product features. Some of this does not make any logical sense and on researching this phenomena we realized it's because DALL-E interprets text as part of the overall visual composition rather than as semantically meaningful content.

Prompt 4:

“Create an image for this product- VHYHCY Stuffed Dinosaur Plush Toy, here is a summary of the visual information to create an accurate image of the product:\n\n- ****Color****: Active Green\n- ****Material****: Made of high-quality polyester fiber and filled with PP cotton\n- ****Design****: Features a cute dinosaur design with a cute expression, suitable for kids and dino lovers\n- ****Size****: 9 inches in length, described as mini and cute, perfect for both boys and girls\n- ****Accessories****: Some mentions of a little hat on the plushie, adding a precious touch to the design\n- ****Texture****: Described as soft, squishy, and comfortable to hold, with a smooth feel and delicate texture\n- ****Features****: Can be used as a decorative item in family bedrooms, cars, etc.\n- ****Concerns****: Some reviews indicate issues with seam quality, slight threading coming loose, small size compared to expectations, and a vacuum-sealed packaging leading to a need for repairs.\n\nTo create an accurate image of the product, you can focus on showcasing the cute dinosaur design in active green color, with attention to details like the little hat, soft and squishy texture, and the compact size suitable for children. Paying close attention to the quality of seams and overall presentation can help in capturing the essence of the product accurately. Also make sure the dino is wearing a colorful party hat. Make this image realistic. **Your image should have only the dino product and nothing else in the background**”

We tried to request DALL-E to remove the product feature information bubbles in the background but that didn't seem to work in the first go as shown below. In-fact it didn't even add the party hat for some reason.



We then gave a follow-up prompt

Follow-up Prompt:

“Remove the feature information bubbles around the dino”



This was a great improvement over the previous image. In-fact it closely resembled our original image. We set to add a follow-up prompt to see if we could model the party hat and scales to colors similar to how the original product image was.

Follow-up Prompt:

“Can you make the party hat and scales colorful for the dino in the above image”



The above image generated did add the party hat and colorful image, but changed the dino from the previous image altogether. This made us realize that when DALL-E generates an image based on a prompt and you provide a follow-up prompt, it doesn't actually "edit" or "incrementally work" on the previous image. Instead, it generates a new image from scratch each time, based on the latest prompt provided. This process doesn't retain the exact features of the previous image unless the new prompt explicitly describes those features.

We discovered the below features of DALL-E:

1. No State Memory: Each prompt to DALL-E is treated as an independent request without memory of past interactions. DALL-E doesn't have a continuous memory of the previous prompts and images.
2. Prompt Interpretation: The AI interprets each prompt based on the entire description given. If the follow-up prompt is not a precise addition to the original description, the AI will interpret the new prompt in isolation, leading to a new creation that may not resemble the previous image.

Prompt 5:

“Create and image for this product- VHYHCY Stuffed Dinosaur Plush Toy, here is a summary of the visual information to create an accurate image of the product:\n\n- ****Color****: Active Green\n- ****Material****: Made of high-quality polyester fiber and filled with PP cotton\n- ****Design****: Features a cute dinosaur design with a cute expression, suitable for kids and dino

lovers\n- **Size**: 9 inches in length, described as mini and cute, perfect for both boys and girls\n- **Accessories**: Some mentions of a little hat on the plushie, adding a precious touch to the design\n- **Texture**: Described as soft, squishy, and comfortable to hold, with a smooth feel and delicate texture\n- **Features**: Can be used as a decorative item in family bedrooms, cars, etc.\n- **Concerns**: Some reviews indicate issues with seam quality, slight threading coming loose, small size compared to expectations, and a vacuum-sealed packaging leading to a need for repairs.\n\nTo create an accurate image of the product, you can focus on showcasing the cute dinosaur design in active green color, with attention to details like the little hat, soft and squishy texture, and the compact size suitable for children. Paying close attention to the quality of seams and overall presentation can help in capturing the essence of the product accurately. **Also make sure the dino is wearing a colorful party hat. Make this image realistic. Your image should have only the dino product and the background is entirely blank to focus solely on the plush toy. Make the party hat and scales colorful for the dino. Active green is a light green color.”**

We finally created a combination of all the above learnings from the previous prompts and created one final prompt which should model our original image accurately. We even suggested the meaning of ‘Active green’ color. This is what we finally achieved:



Final Conclusion FOR Product 1:

Similarities:

- Both images feature a plush green dinosaur with a similar overall pose.
- They both have a party hat and a similar color scheme.
- The facial blush marks (something that wasn't mentioned in the prompt).

Differences:

- There's a minor difference in the detailing, the overall shape, the expression (smile), especially the party hat, where the AI-generated image has a simpler design.
- The spines on the back of the dinosaur show more pronounced three-dimensionality and shadowing in the AI-generated image.

AI is able to illustrate the products well based on the information that has been provided to it. It does a reasonable job of illustrating the product, capturing the essence of the toy. However, the texture and some fine details are not as accurate as the real product, which are minor elements.

2. Modern Couch

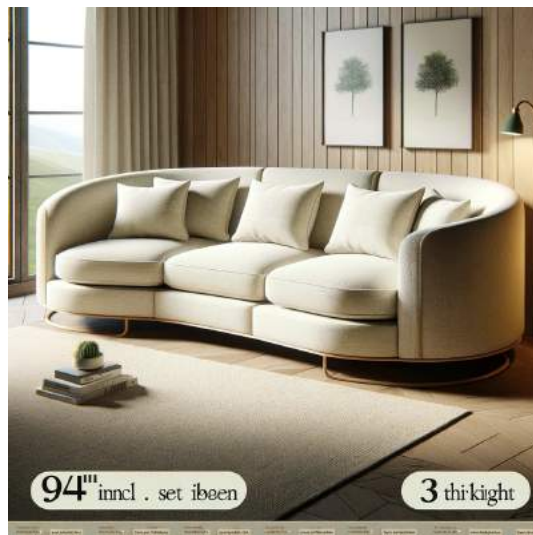


Prompt 1 (Baseline Prompt):

“Based on the textual description provided for the 94" Mid Century Modern Curved Living Room Sofa by JUSTONE, here is the extracted information to help create an accurate image of

the product:\n\n- **Product Type:** 94" Mid Century Modern Curved Living Room Sofa\n-
 Brand: JUSTONE\n- **Seating Capacity:** 3\n- **Frame Material:** Wood\n-
 Shape: Curved\n- **Style:** Mid-Century Modern\n- **Color:** Beige\n- **Upholstery Fabric Type:** Boucle\n- **Room Type:** Living Room, Bedroom, Home Office\n- **Special Feature:** Easy to Assemble\n- **Arm Style:** Curved\n- **Dimensions:** 93.6"D x 41.3"W x 26.9"H\n- **Item Weight:** 299 Pounds\n- **Features:** \n - Upholstered in high-resilience Boucle Fabric\n - Sturdy solid wood frame\n - Comfortable, sinuous and enfolding silhouette\n - Soft padded structure\n - Thickened seat cushions for firm and comfortable seating\n - Curved thickened backrest for good support\n - Pillow Quantity: 3 Pillows\n - High-density sponge filler for comfort\n- **Description:** \n - The sofa is described as elegant, simple, and lightweight\n - Creates a warm, cozy, and relaxing atmosphere\n - Offers luxurious enjoyment with premium quality materials\n - Suitable for residential or contract settings\n - Easy to care for and maintain\n- **Customer Review Highlights:** \n - Craftsmanship is sturdy and stylish\n - Fabric is touchable and comfortable\n - Bold, colorful, and modern statement piece or accent piece\n - Superior and pleasant customer service experience\n\nThis extracted information should provide a detailed understanding of the visual attributes and features of the product, which can be used to create an accurate image using generative image AI techniques. I extracted features of a product on Amazon using the description and product reviews. Can you use Dall E 3 to generate images of this product.”

Output:



There are some unnecessary product features on the image which do not make any sense. In the next prompt we try to remove those from the image generated.

Prompt 2:

Based on the textual description provided for the 94" Mid Century Modern Curved Living Room Sofa by JUSTONE, here is the extracted information to help create an accurate image of the product:\n\n- **Product Type:** 94" Mid Century Modern Curved Living Room Sofa\n- **Brand:** JUSTONE\n- **Seating Capacity:** 3\n- **Frame Material:** Wood\n- **Shape:** Curved\n- **Style:** Mid-Century Modern\n- **Color:** Beige\n- **Upholstery Fabric Type:** Boucle\n- **Room Type:** Living Room, Bedroom, Home Office\n- **Special Feature:** Easy to Assemble\n- **Arm Style:** Curved\n- **Dimensions:** 93.6"D x 41.3"W x 26.9"H\n- **Item Weight:** 299 Pounds\n- **Features:** \n - Upholstered in high-resilience Boucle Fabric\n - Sturdy solid wood frame\n - Comfortable, sinuous and enfolding silhouette\n - Soft padded structure\n - Thickened seat cushions for firm and comfortable seating\n - Curved thickened backrest for good support\n - Pillow Quantity: 3 Pillows\n - High-density sponge filler for comfort\n- **Description:** \n - The sofa is described as elegant, simple, and lightweight\n - Creates a warm, cozy, and relaxing atmosphere\n - Offers luxurious enjoyment with premium quality materials\n - Suitable for residential or contract settings\n - Easy to care for and maintain\n- **Customer Review Highlights:** \n - Craftsmanship is sturdy and stylish\n - Fabric is touchable and comfortable\n - Bold, colorful, and modern statement piece or accent piece\n - Superior and pleasant customer service experience\n\nThis extracted information should provide a detailed understanding of the visual attributes and features of the product, which can be used to create an accurate image using generative image AI techniques.

I extracted features of a product on Amazon using the description and product reviews. Can you use Dall E 3 to generate images of this product.

Dont write write any product specification on the image (No text on the image) and generate 3 images

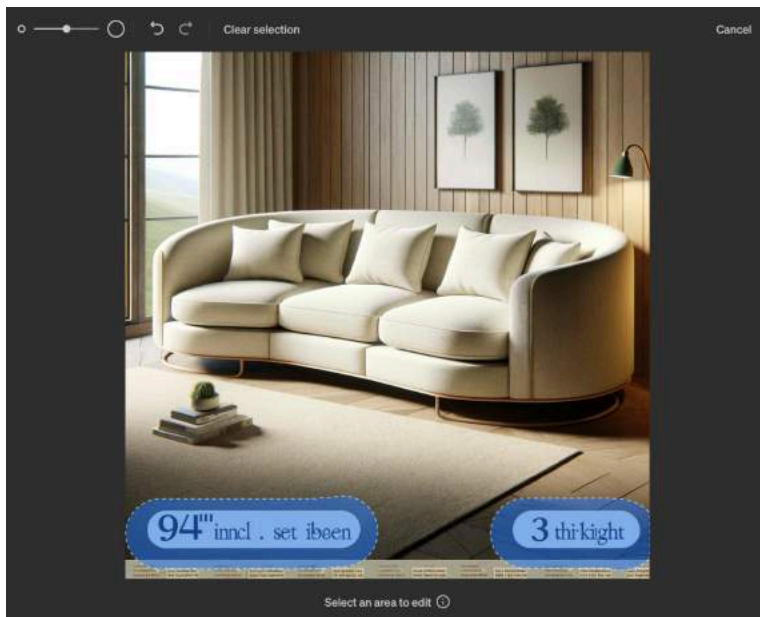
Output:



Prompt 3:

Using ChatGPT's new feature of highlighting areas of the image to edit along with a prompt allowed the removal of the unnecessary segment from Prompt 1.

Image highlighting -



Prompt -

Remove this part of the image

Output:



3. Rainbow Skeleton Tutu

Prompt 1 (Baseline Prompt):

“Create an image for this product - EraSpooky Women Funky Punk Bones Costumes Skeleton Colorful Tutu Dress”:\n- The costume is described as colorful, vibrant, and fun with rainbow colors.\n- It features a rainbow tutu dress with a colorful ribcage print, matching bone printed leggings, and wristlets.\n- The tutu dress has a neon rainbow design.\n- The ribcage print on the dress is colorful and eye-catching.\n- The leggings are described as black footless stockings, thin, and somewhat transparent.\n- The arm parts have elastic at the top that may feel tight.\n- The tights are described as thin like stockings and run small in size.\n- The material of the dress is mentioned to have some itchiness from the tulle where it is sewn.\n- The dress is noted to have a black lining under the tutu for ease of use in the bathroom.\n- The colors of the costume are bright and vibrant, matching the rainbow theme.\n- The outfit is suitable for stage performances,

costume parties, Halloween, and is visually appealing.\n- Overall, the product is described as fun, bright, comfortable to wear, and received many compliments.\n\nBased on these descriptions, the ideal image of the product would feature:\n- A rainbow tutu dress with a colorful ribcage print.\n- Matching bone printed leggings and wristlets in bright rainbow colors.\n- A neon rainbow design on the dress.\n- The dress with a black lining under the tutu for practicality.\n- Thin, black footless stockings for the leggings component.\n- Elastic arm parts with a tight fit.\n- A fun, vibrant, and colorful overall appearance that stands out and is suitable for various events and occasions.”



Prompt 2:

“Create a product image based on the textual data provided: "EraSpooky Women Funky Punk Bones Costumes Skeleton Colorful Tutu Dress":\n- The costume is described as colorful, vibrant, and fun with rainbow colors.\n- It features a rainbow tutu dress with a colorful ribcage print, matching bone printed leggings, and wristlets.\n- The tutu dress has a neon rainbow design.\n- The ribcage print on the dress is colorful and eye-catching.\n- The leggings are described as black footless stockings, thin, and somewhat transparent.\n- The arm parts have elastic at the top that may feel tight.\n- The tights are described as thin like stockings and run small in size.\n- The material of the dress is mentioned to have some itchiness from the tulle where it is sewn.\n- The dress is noted to have a black lining under the tutu for ease of use in the bathroom.\n- The colors

of the costume are bright and vibrant, matching the rainbow theme.\n- The outfit is suitable for stage performances, costume parties, Halloween, and is visually appealing.\n- Overall, the product is described as fun, bright, comfortable to wear, and received many compliments.\n\nBased on these descriptions, the ideal image of the product would feature:\n- A rainbow tutu dress with a colorful ribcage print.\n- Matching bone printed leggings and wristlets in bright rainbow colors.\n- A neon rainbow design on the dress.\n- The dress with a black lining under the tutu for practicality.\n- Thin, black footless stockings for the leggings component.\n- Elastic arm parts with a tight fit.\n- A fun, vibrant, and colorful overall appearance that stands out and is suitable for various events and occasions. **The top should have black puff sleeves. The wristlets should start from the elbows.**”



Prompt 3:

“Create a product image based on the textual data provided: "EraSpooky Women Funky Punk Bones Costumes Skeleton Colorful Tutu Dress":\n- The costume is described as colorful, vibrant, and fun with rainbow colors.\n- It features a rainbow tutu dress with a colorful ribcage print, matching bone printed leggings, and wristlets.\n- The tutu dress has a neon rainbow design.\n- The ribcage print on the dress is colorful and eye-catching.\n- The leggings are described as black footless stockings, thin, and somewhat transparent.\n- The arm parts have elastic at the top that may feel tight.\n- The tights are described as thin like stockings and run small in size.\n- The material of the dress is mentioned to have some itchiness from the tulle where it is sewn.\n- The dress is noted to have a black lining under the tutu for ease of use in the bathroom.\n- The colors of the costume are bright and vibrant, matching the rainbow theme.\n- The outfit is suitable for stage performances, costume parties, Halloween, and is visually appealing.\n- Overall, the product is described as fun, bright, comfortable to wear, and received many compliments.\n\nBased on these descriptions, the ideal image of the product would feature:\n- A rainbow tutu dress with a colorful ribcage print.\n- Matching bone printed leggings and wristlets in bright rainbow colors.\n- A neon rainbow design on the dress.\n- The dress with a black lining under the tutu for practicality.\n- Thin, black footless stockings for the leggings component.\n- Elastic arm parts with a tight fit.\n- A fun, vibrant, and colorful overall appearance that stands out and is suitable for various events and occasions. **The top should have black puff sleeves. The wristlets should start from the elbows. Instead of rainbow use pink, blue and green only.**”



Follow-Up Prompt:

“The top should have black **cloth** puff half-sleeves. The gloves on the arms should start from the elbows only”



Prompt 4:

“Create a product image based on the textual data provided: "EraSpooky Women Funky Punk Bones Costumes Skeleton Colorful Tutu Dress":\n- The costume is described as colorful, vibrant, and fun with rainbow colors.\n- It features a rainbow tutu dress with a colorful ribcage print, matching bone printed leggings, and wristlets.\n- The tutu dress has a neon rainbow design.\n- The ribcage print on the dress is colorful and eye-catching.\n- The leggings are described as black footless stockings, thin, and somewhat transparent.\n- The arm parts have elastic at the top that may feel tight.\n- The tights are described as thin like stockings and run small in size.\n- The material of the dress is mentioned to have some itchiness from the tulle where it is sewn.\n- The dress is noted to have a black lining under the tutu for ease of use in the bathroom.\n- The colors of the costume are bright and vibrant, matching the rainbow theme.\n- The outfit is suitable for stage performances, costume parties, Halloween, and is visually appealing.\n- Overall, the product is described as fun, bright, comfortable to wear, and received many compliments.\n\nBased on these descriptions, the ideal image of the product would feature:\n- A rainbow tutu dress with a colorful ribcage print.\n- Matching bone printed leggings and wristlets in bright rainbow colors.\n- A neon rainbow design on the dress.\n- The dress with a black lining under the tutu for practicality.\n- Thin, black footless stockings for the leggings component.\n- Elastic arm parts with a tight fit.\n- A fun, vibrant, and colorful overall appearance that stands out and is suitable for various events and occasions. **The top should have black puff sleeves. The wristlets should start from the elbows. Instead of rainbow use pink, blue and green only.**”



Follow-up Prompt:

“Remove the skeleton wearing it, it should be worn by a normal human”



Follow-up Prompt:

“Remove the text and anything else in the background”

