

# Enhancing Privacy-Utility Trade-offs to Mitigate Memorization in Diffusion Models

## Supplementary Material

**Outline.** In Sec. 8, we provide additional analysis explaining why the *prompt re-anchoring* (PR) strategy effectively enhances privacy and how synergistic benefits emerge when combined with the *semantic search* (SS) strategy. In Sec. 9, we present additional quantitative results, showing that the proposed PRSS strategy consistently improves the privacy-utility trade-offs compared to baseline methods. These improvements are demonstrated under different detection signals for both global and local memorization scenarios and evaluated using various memorization metrics. In Sec. 10, we include additional qualitative results highlighting PRSS’s superior performance over baseline approaches in improving privacy while maintaining high utility. In Sec. 11, we delve into further implementation details to enhance the reproducibility of our work.

### 8. Additional Analysis

Fig. 8 demonstrates the shortcomings of the baseline’s prompt engineering strategy in preserving privacy and the advantages of our proposed *prompt re-anchoring* (PR) and *semantic search* (SS) strategies. Please check out the main paper for a detailed analysis, the key takeaways are summarized as follows.

**Enhancing Privacy Through PR.** As shown in both line charts in Fig. 8, while the engineered prompt achieves a low magnitude initially, it fails to ensure that this low level is maintained throughout subsequent inference steps, potentially leading back to high magnitudes indicative of Stable Diffusion’s memorized outputs. In response, our PR strategy introduces a mechanism for consistent diversion from memorization across the entire inference process. This is demonstrated by consistently lower magnitudes and corroborated by both the qualitative appearance of the final generated image and its quantitative similarity score, showcasing effective privacy preservation.

**Synergy Effect of the Two Strategies.** The right side of Fig. 8 showcases the synergistic benefit of utilizing the *prompt re-anchoring* strategy to compensate for the SS strategy’s limitations. Although the semantic search may not secure an alternative prompt with a sufficiently reduced magnitude, lowering it from the original 7.48 to 6.02 is proved to be adequate for *prompt re-anchoring* to leverage the original prompt as the new anchor point for steering the generation away from memorized content effectively.

### 9. Additional Quantitative Results

As discussed in Sec. 4.1, we complement the standard average SSCD similarity score that is reported in Fig. 6 with additional metrics for a comprehensive evaluation. These include the localized similarity metric LS (Eq. (6)), designed specifically for assessing local memorization [8], shown in Fig. 9. We also report other SSCD-based statistics, such as the 95th percentile similarity score used in [32] (Fig. 10) and the percentage of memorized generations (SSCD similarity  $> 0.50$ ) employed in [6], as depicted in Fig. 11.

Same as Fig. 6, these additional results are presented separately for global and local memorization scenarios, comparing PRSS against two baselines ([36] and [8]) using their respective detection signals to ensure fair comparisons.

Across all metrics, the results confirm that PRSS consistently improves the privacy-utility trade-offs of both baselines, as evidenced by the red line consistently lying in the bottom-right direction of the yellow line, and the green line consistently lying in the bottom-right direction of the blue line for both global and local memorization scenarios.

### 10. Additional Qualitative Results

We present additional qualitative results in Figs. 12 and 13 for comparative analysis in both local and global memorization scenarios using the same random seeds. These results demonstrate the superior performance of our method in enhancing privacy while preserving utility, effectively reducing memorization within Stable Diffusion generations.

Specifically, generations produced by baseline strategies that leverage prompt engineering often exhibit poor alignment with the intended meaning of the user prompt or continue to replicate training images similar to those generated by the original Stable Diffusion model. This highlights their limitations in maintaining both utility and privacy. In contrast, our approach produces images that are significantly less prone to memorization while remaining semantically aligned with the user’s prompt.

### 11. Additional Implementational Details

#### 11.1. Controlling the Strength of PR

We achieve finer-grained control of the privacy-utility trade-off by adjusting the strength of PR guidance. Specifically, we balance the guidance scale  $s$  between fully using the re-anchored guidance method (Eq. (13)) and fully using the

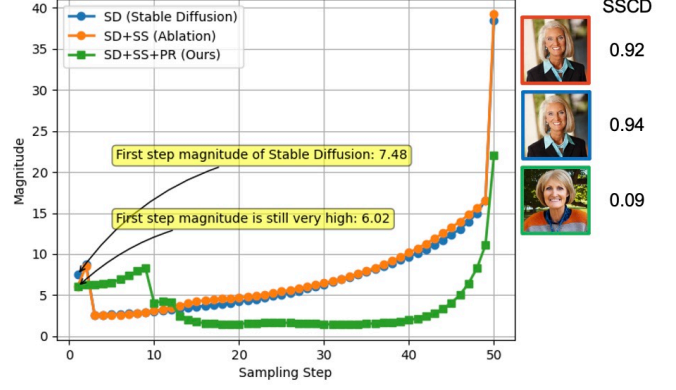
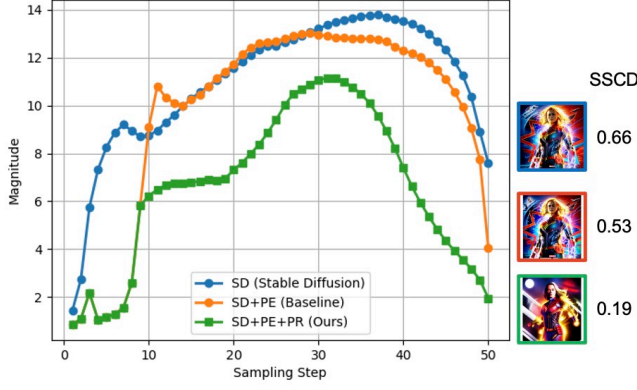


Figure 8. **Left and Right:** Our *prompt re-anchoring* (PR) strategy allows continuous memorization diversion throughout the entire inference process, effectively reducing magnitude and enhancing privacy. **Right:** Should the alternative prompt identified by the *semantic prompt search* (SS) strategy not achieve a low magnitude, employing PR concurrently can still prevent memorized image generation, provided the magnitude of the alternative prompt is lower than that of the original.

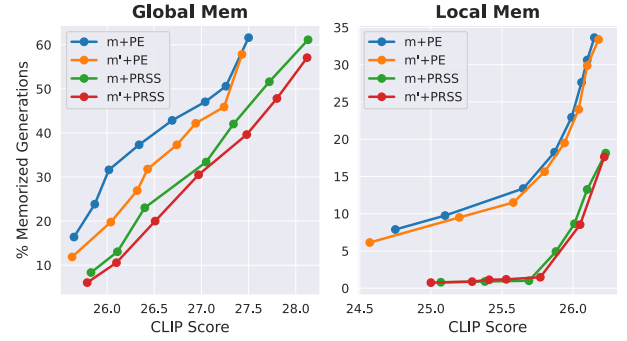
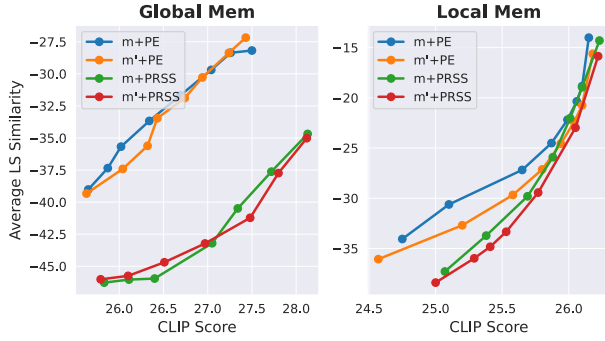


Figure 9. Comparison with baselines using localized memorization metric LS. Across different detection signals ( $m$  and  $m'$ ), PRSS consistently achieves superior privacy-utility trade-offs.

Figure 11. Comparison with baselines using the percentage of memorized generations (quantified as SSSD similarity  $> 0.50$ ) as the memorization metric. Across different detection signals ( $m$  and  $m'$ ), PRSS consistently achieves superior privacy-utility trade-offs.

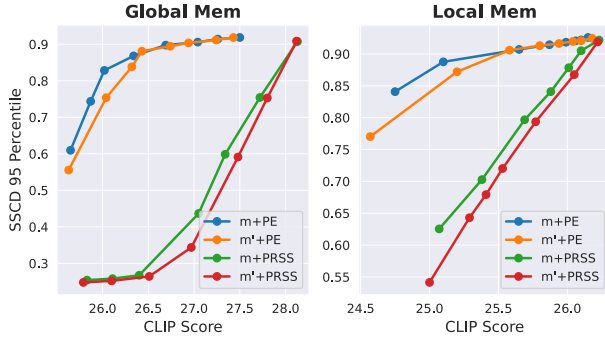


Figure 10. Comparison with baselines using the 95 percentile of SSSD similarity scores as the memorization metric. Across different detection signals ( $m$  and  $m'$ ), PRSS consistently achieves superior privacy-utility trade-offs.

original guidance method (Eq. (14)).

$$\hat{\epsilon} \leftarrow [\epsilon_{\theta}(x_t, e_{\phi}) + s(\epsilon_{\theta}(x_t, e_p) - \epsilon_{\theta}(x_t, e_{\phi}))] \mathbb{1}_{\{m_{T-1} < \lambda\}} + [\epsilon_{\theta}(x_t, e_p) + s(\epsilon_{\theta}(x_t, e_p^{ss}) - \epsilon_{\theta}(x_t, e_p))] \mathbb{1}_{\{m_{T-1} > \lambda\}}. \quad (13)$$

$$\hat{\epsilon} \leftarrow [\epsilon_{\theta}(x_t, e_{\phi}) + s(\epsilon_{\theta}(x_t, e_p) - \epsilon_{\theta}(x_t, e_{\phi}))] \mathbb{1}_{\{m_{T-1} < \lambda\}} + [\epsilon_{\theta}(x_t, e_{\phi}) + s(\epsilon_{\theta}(x_t, e_p^{ss}) - \epsilon_{\theta}(x_t, e_{\phi}))] \mathbb{1}_{\{m_{T-1} > \lambda\}}. \quad (14)$$

The balanced approach, shown in Eq. (15), splits the guidance scale  $s$  between these two options, allowing for a more flexible trade-off.

$$\hat{\epsilon} \leftarrow [\epsilon_{\theta}(x_t, e_{\phi}) + s(\epsilon_{\theta}(x_t, e_p) - \epsilon_{\theta}(x_t, e_{\phi}))] \mathbb{1}_{\{m_{T-1} < \lambda\}} + [\epsilon_{\theta}(x_t, e_p) + s'(\epsilon_{\theta}(x_t, e_p^{ss}) - \epsilon_{\theta}(x_t, e_p))] + (s - s')(\epsilon_{\theta}(x_t, e_p^{ss}) - \epsilon_{\theta}(x_t, e_{\phi})) \mathbb{1}_{\{m_{T-1} > \lambda\}}. \quad (15)$$

where  $s' \in [1, s]$ : when  $s' = 1$ , the method defaults to the original guidance using  $e_{\phi}$  as the anchor; when  $s' = s$ , the guidance term is fully re-anchored to  $e_p$ . These repre-

sent the two extreme cases. Specifically, since the magnitude quantifies the level of memorization risk, we design  $s'$  to be proportional to the real-time magnitude at timestep  $t$  throughout the entire inference process. It is capped at  $s$  for excessively large magnitudes that are greater than  $\lambda_{\max}$  and floored at 1 when the magnitude is below the detection threshold  $\lambda$ , ensuring no PR is triggered during non-memorized cases to preserve output utility maximally. This behavior is governed by the following specifications:

$$s' = 1 + (s - 1) \cdot \left( \frac{\max(\min(m_t, \lambda_{\max}) - \lambda, 0)}{\lambda_{\max} - \lambda} \right) \quad (16)$$

## 11.2. Generating Alternative Prompts

As discussed in Sec. 4.1, we leverage GPT-4 via the OpenAI API for searching semantically similar alternative text prompts. To ensure reproducibility, we describe the process we used for such a task. After taking a user-provided text prompt, the process begins by setting up a system instruction that explicitly directs GPT-4 to act as a skilled prompt engineer. The instruction focuses on paraphrasing the input prompt into semantically similar, natural-sounding alternatives. The exact instruction provided to GPT-4 is as follows: *“You are a prompt engineer skilled in paraphrasing user prompts into semantically similar, natural-sounding prompts. I will provide a text prompt that will be fed into Stable Diffusion. Please generate 1 semantically similar prompt as a plain sentence, without using any list numbering (such as 1., 2., etc.) or quotation marks. Each prompt should be a standalone sentence in plain language.”*

The OpenAI API is configured to query the GPT-4 model with these messages. Parameters such as `max_tokens`, `temperature`, and the number of alternatives ( $n$ ) are adjusted to optimize the generation process. For instance, `max_tokens=750` ensures sufficient response length, `temperature=0.8` introduces slight randomness for diversity, and `n=1` generates a single prompt per call for focused evaluation.

Below, we list some examples of generated alternative prompts (in bullet points) for a given user text prompts (in bold):

### **The No Limits Business Woman Podcast**

- The Businesswoman Without Boundaries Show
- The Entrepreneurial Woman’s Unrestricted Podcast
- Podcast for the Unstoppable Business Lady
- The Infinite Potential Business Woman Series
- The Female Mogul’s Limitless Journey Podcast
- Broadcast for the Fearless Businesswoman
- The Ceaseless Enterprise Woman Podcast
- Unbounded Success: A Woman’s Business Podcast
- The Entrepreneurial Woman’s No-Barriers Broadcast
- The Female Executive’s Infinite Podcast
- The Business Femme’s Unchained Podcast Series

- Undeterred: The Business Woman’s Podcast
- The Unrestricted Businesswoman’s Dialogue
- The Empowered Business Woman’s Podcast
- Ventures Without Limits: A Female Entrepreneur’s Podcast
- The Business Woman’s Sky’s-the-Limit Show
- Limitless Horizons: The Woman Entrepreneur Podcast
- The Trailblazing Businesswoman’s Podcast
- Podcasting the Unstoppable Business Woman
- Frontiers of Business: The Woman’s Edition
- The Entrepreneurial Spirit: Women Without Limits
- The Podcast for Limit-Defying Businesswomen
- Beyond Barriers: The Female Business Leader Podcast
- The Businesswoman’s Podcast: No Boundaries
- Empire Building: The Woman’s Podcast without Limit

### **Mothers influence on her young hippo**

- Mother hippo’s guidance of her offspring
- The impact of a mother on her baby hippo
- Maternal effect on a juvenile hippo
- The role of a mother in a young hippo’s life
- Motherly impact on a young hippopotamus
- The nurturing of a baby hippo by its mother
- Influence of a mother hippo on her calf
- How a mother shapes her hippo youngster
- Maternal shaping of a young hippo’s development
- The nurturing bond between mother and baby hippo
- Maternal teachings to a young hippopotamus
- The guiding presence of a mother in her hippo’s upbringing
- A mother hippo’s role in her offspring’s growth
- The maternal touch in raising a young hippo
- Motherly wisdom bestowed upon a juvenile hippo
- A mother’s guiding influence on her young hippopotamus
- Life lessons from a mother to her hippo calf
- The formative influence of a mother on her baby hippo
- A hippo mother’s impact on her offspring’s early life
- Maternal instincts shaping a young hippo’s journey
- The protective guidance of a mother hippo to her young
- Young hippo’s upbringing under its mother’s care
- The maternal effect on a hippo calf’s behavior
- Influence of motherhood on a young hippo’s path
- Mother and young hippo: A nurturing relationship

### **Talks on the Precepts and Buddhist Ethics**

- Conversations centered on Buddhist precepts and ethics.
- Discussion on the ethical guidelines and precepts in Buddhism.
- Let’s discuss the precepts and moral guidelines in Buddhist teachings.
- An in-depth exploration of Buddhist precepts and ethics.
- Discussing Buddhist precepts and the ethical guidelines that follow.

- A detailed discussion about precepts and ethics in Buddhism.
- Unpacking the ethical teachings and precepts of Buddhism.
- Looking into the ethical teachings and precepts of Buddhism.
- Uncovering the precepts and ethical standards of Buddhism.
- Understanding the precepts and ethical values of Buddhism.
- An examination of Buddhist precepts and the moral code associated with them.
- Understanding the principles and ethical teachings of Buddhism.
- Exploring the ethical guidelines and teachings of Buddhism.
- A talk about the moral teachings and principles upheld by Buddhists.
- Let's delve into the teachings and moral code of Buddhism.
- Discussing the principles and ethics in Buddhism.
- A comprehensive discussion on the principles and ethical standards in Buddhism.
- A dialogue about the principles and moral standards in Buddhism.
- Exploring the ethical teachings and principles upheld in Buddhism.
- A conversation about the ethical principles and precepts followed in Buddhism.
- Discussing the ethical values and teachings observed in Buddhism.
- Get to know the ethical principles and teachings in Buddhism.
- A deep dive into the moral standards and precepts in Buddhism.
- Examining the moral guidelines and teachings of Buddhism.
- Conversations revolving around the moral standards and principles in Buddhism.

#### **Sony Won't Release *The Interview* on VOD**

- Sony opts out of releasing The Interview on VOD.
- Sony refrains from releasing The Interview on VOD.
- Sony denies the release of The Interview on VOD.
- Sony will not launch The Interview on VOD.
- Sony decides not to launch The Interview on Video on Demand.
- Sony decides not to make The Interview available on VOD.
- Sony has decided against the VOD release of The Interview.
- Sony rules out releasing The Interview on Video on De-

mand.

- Sony has ruled out the VOD release of The Interview.
- Sony is not planning to release The Interview on VOD.
- Sony has chosen not to put out The Interview on VOD.
- Sony has refused to release The Interview on Video on Demand.
- Sony abstains from releasing The Interview on Video on Demand.
- Sony has declined to release The Interview on Video on Demand.
- Sony won't be putting The Interview out on VOD.
- The Interview will not be available on VOD, says Sony.
- The VOD release of The Interview has been canceled by Sony.
- Sony is not going to make The Interview available on VOD.
- The Interview will not be launched on Video on Demand by Sony.
- Sony won't be releasing The Interview on Video on Demand.
- The Interview will not be made available on VOD by Sony.
- The Interview won't be released on VOD by Sony.
- Sony will not be sending The Interview to VOD.
- Sony isn't going to release The Interview on Video on Demand.
- The Interview's VOD release has been declined by Sony.

#### ***South Park: The Stick of Truth* Review (Multi-Platform)**

- A multi-platform review of the game South Park: The Stick of Truth.
- A roundup review of South Park: The Stick of Truth across multiple platforms.
- Review and analysis of South Park: The Stick of Truth for different platforms.
- A detailed review of South Park: The Stick of Truth for various platforms.
- A comprehensive review of South Park: The Stick of Truth across various gaming platforms.
- Reviewing South Park: The Stick of Truth on its various platforms.
- A critique of South Park: The Stick of Truth for multiple gaming platforms.
- An overview of South Park: The Stick of Truth for various gaming platforms.
- A breakdown review of South Park: The Stick of Truth on different platforms.
- A thorough review of South Park: The Stick of Truth on various platforms.
- Exploring South Park: The Stick of Truth in a multi-platform context.
- Assessment of South Park: The Stick of Truth for multi-

ple gaming platforms.

- A thorough examination of South Park: The Stick of Truth on multiple platforms.
- Evaluating South Park: The Stick of Truth on multiple gaming platforms.
- A multi-platform critique of South Park: The Stick of Truth.
- Examining South Park: The Stick of Truth on multiple game platforms.
- A comprehensive look at South Park: The Stick of Truth across different platforms.
- A deep dive into South Park: The Stick of Truth on multiple platforms.
- A game analysis for South Park: The Stick of Truth on multiple platforms.
- Analyzing the game South Park: The Stick of Truth for multiple platforms.
- Analyzing the multi-platform game South Park: The Stick of Truth.
- Delving into South Park: The Stick of Truth on multiple platforms.
- Rating South Park: The Stick of Truth across various platforms.
- Discussing the merits of South Park: The Stick of Truth on different platforms.
- Inspection of South Park: The Stick of Truth across several platforms.

#### **Insights with Laura Powers**

- Gleaning insights from Laura Powers.
- Exploring insights with Laura Powers.
- Discovering Laura Powers insights.
- Taking in insights from Laura Powers.
- Learning from Laura Powers insights.
- A deep dive into Laura Powers ideas.
- Exploring the mind of Laura Powers.
- Unearthing the wisdom of Laura Powers.
- Immersing in a conversation with Laura Powers.
- Understanding the worldview of Laura Powers.
- Indulging in an enlightening talk with Laura Powers.
- Understanding Laura Powers perspective.
- Getting a glimpse into Laura Powers thought process.
- Seeking an intellectual exchange with Laura Powers.
- Uncovering Laura Powers thoughts and opinions.
- Seeking wisdom from Laura Powers.
- Engaging in a knowledge sharing session with Laura Powers.
- Delving into Laura Powers expertise.
- Discussing various topics with Laura Powers.
- Getting to know Laura Powers better.
- Venturing into the intellectual space of Laura Powers.
- Analysing the thought process of Laura Powers.
- Hearing Laura Powers point of view.
- Dive into the world of Laura Powers.

- Having an intellectual conversation with Laura Powers.

#### ***The Long Dark* Gets First Trailer, Steam Early Access**

- The Long Dark has its first trailer out, ready for play on Steam Early Access.
- The Long Dark has made its first trailer public, also now available on Steam Early Access.
- The Long Dark debuts its first trailer and is now part of Steam Early Access.
- The Long Dark introduces its first trailer and is now available on Steam Early Access.
- The initial trailer of The Long Dark has been released, now on Steam Early Access.
- The Long Dark releases its initial trailer and is now live on Steam Early Access.
- The Long Dark releases its first trailer and is now included in Steam Early Access.
- The first trailer for The Long Dark has been unveiled, now available on Steam Early Access.
- The Long Dark has just revealed its first trailer, also now accessible on Steam Early Access.
- The Long Dark has launched its first trailer and it's now on Steam Early Access.
- The Long Dark premieres its first trailer and is now available for Steam Early Access.
- The Long Dark premieres its first trailer and is now featured in Steam Early Access.
- The Long Dark just dropped its first trailer and is now playable on Steam Early Access.
- The Long Dark unveils its debut trailer and joins the Steam Early Access platform.
- The game The Long Dark released its very first trailer and is now in Steam Early Access.
- The first trailer for The Long Dark has been launched, now available through Steam Early Access.
- An initial trailer for The Long Dark is out, and it's now accessible on Steam Early Access.
- The game The Long Dark has launched its first trailer and is now up for Early Access on Steam.
- The Long Dark rolls out its first trailer and is now accessible through Steam Early Access.
- The initial trailer for The Long Dark is now out and the game is accessible on Steam Early Access.
- Steam Early Access just received The Long Dark and its first trailer has been launched.
- The Long Dark has just launched its first trailer and is now available for Early Access on Steam.
- The video game The Long Dark has released its first trailer and now is part of Steam Early Access.
- Steam Early Access now features The Long Dark with its first trailer being out.
- Steam Early Access starts hosting The Long Dark, as its first trailer is launched.



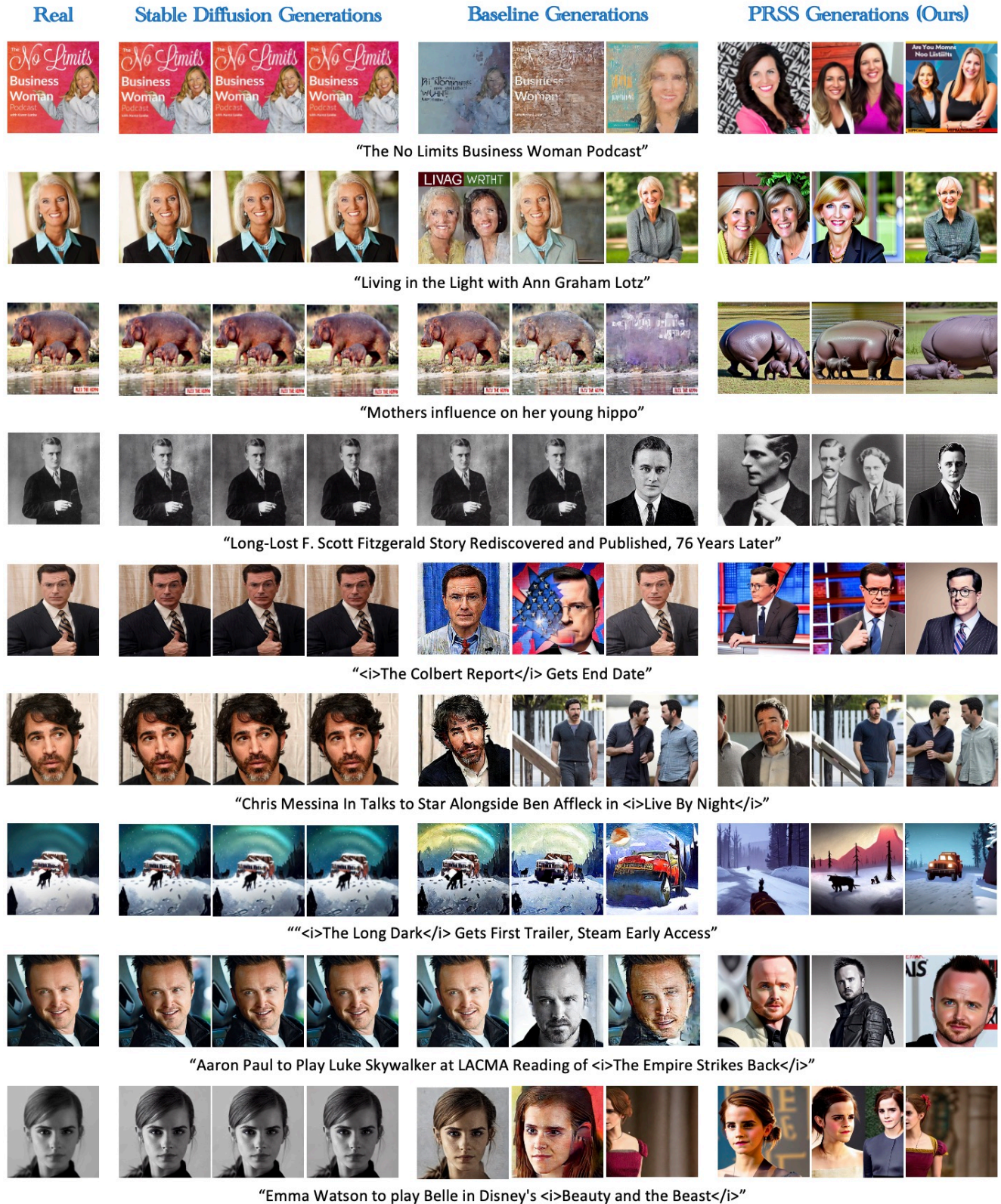


Figure 12. Qualitative comparisons for global memorization using the same random seeds. The baseline’s prompt engineering approach often incurs a significant loss of output utility or fails to adequately mitigate memorization. PRSS effectively reduces memorization in Stable Diffusion while maintaining semantic alignment between the generated images and the user’s prompt.



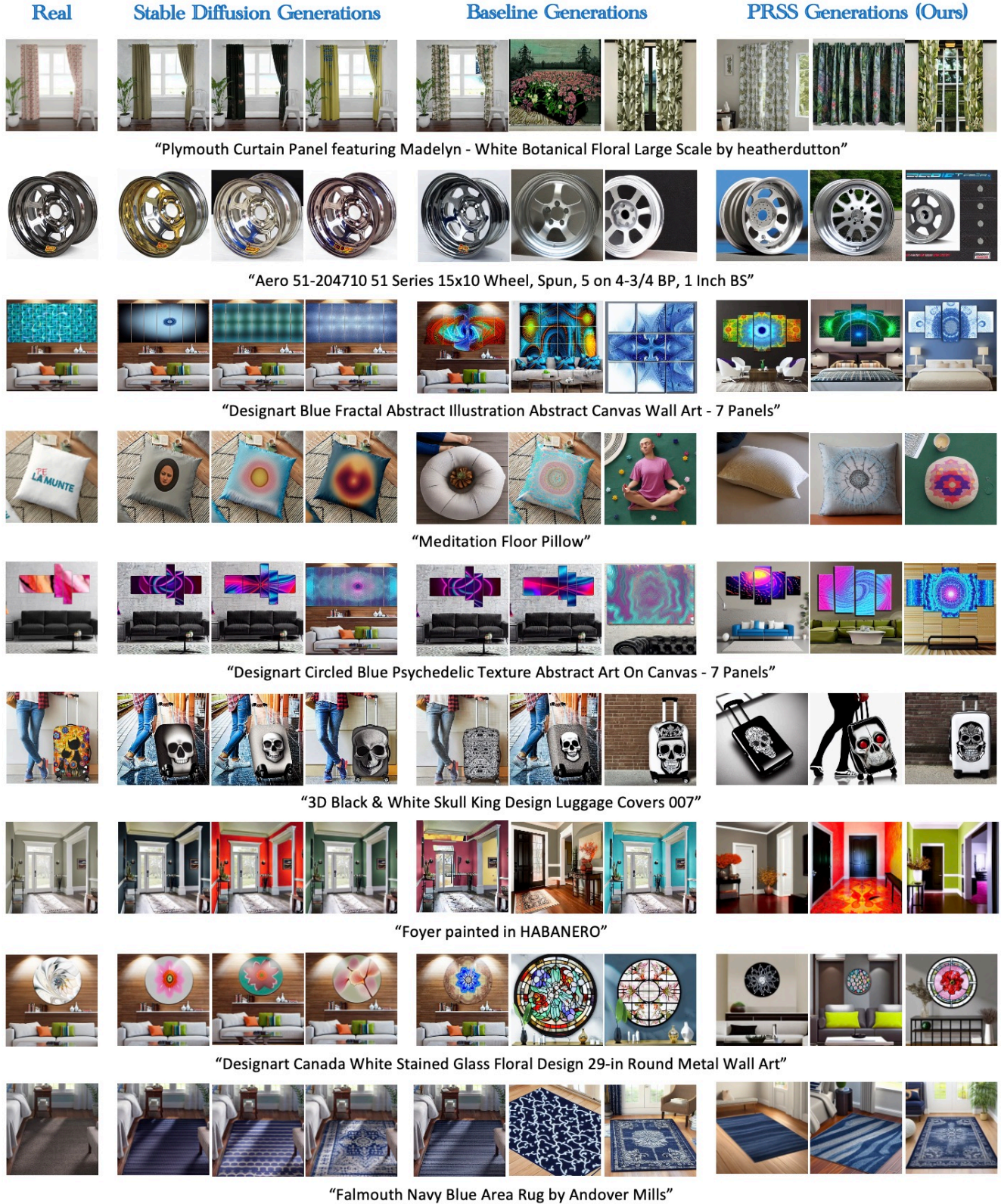


Figure 13. Qualitative comparisons for local memorization using the same random seeds. The baseline’s prompt engineering approach often incurs a significant loss of output utility or fails to adequately mitigate memorization. PRSS effectively reduces memorization in Stable Diffusion while maintaining semantic alignment between the generated images and the user’s prompt.