

AI Trials: Documenting Global Festivities

Testing AI on World Celebration Content

WILLIAM TREKELL

JAN 02, 2024



Share



Artwork created with Midjourney v6 alpha

In 2024, I conducted a project to experiment with AI to gain a better understanding of its capabilities and how it might be used in real-world scenarios. To explore this, I chose to document global holiday celebrations. Each month, I will use various AI tools, including ChatGPT-4, Bard, and Midjourney, to create content and imagery that capture the essence of international festivities.

Although my professional experience isn't directly in AI, I am deeply interested in its potential. This series will explore how these celebrations can serve as practical examples to observe the capabilities and limitations of current AI technologies. Each month, I will document how these AI models perform in generating relevant content and images, offering insights into their evolving capabilities and areas needing improvement.

Throughout the year, I will share observations on the effectiveness of these AI tools examining both successes and challenges. This approach aims to realistically assess and document AI's progress and practicality in a dynamic, real-world context.

January's Cultural Discoveries

In January I focused on evaluating the performance of AI models like ChatGPT, Bard, and Claude in generating content about global holiday traditions. Key experiments included generating articles on Martin Luther King Jr. Day and Makar Sankranti, revealing the importance of detailed prompts and human oversight in achieving high quality content.

- **AI Trials: Data Collection and Deception**
 - [AI Trials: January Part 1](#): Initial experiments for holiday content creation.
 - [AI Trials: January Part 2](#): Holiday data collection and comparison.
- **AI Trials: Historian Roles on History**
 - [AI Trials: January Part 3](#): Testing roles with varied specificity.
 - [AI Trials: January Part 4](#): Article rating criteria creation.

February's Festival Insights

In February, I explored the impact of specialized roles on content creation. By tailoring roles for AI models, significant improvements in output precision and relevance were achieved. The month highlighted the critical role of human oversight in ensuring culturally sensitive and accurate content.

- **AI Trials: Professionals to Specialists**
 - [AI Trials: February Part 1](#): Defining specialized roles for AI in content creation.
 - [AI Trials: February Part 2](#): Enhancing templates and rating criteria.

- [AI Trials: February Part 3](#): Challenges and triumphs in specialized roles.
- [AI Trials: February Part 4](#): Validating role definitions for use with Claude.
- **AI Trials: Leveling the Playing Field**
 - [AI Trials: February Part 5](#): Challenges in AI workflows.
 - [AI Trials: February Part 6](#): Structured prompts and roles.
 - [AI Trials: February Part 7](#): Prompts and response accuracy needs.
 - [AI Trials: February Part 8](#): Enhanced AI roles improve content creation.

March's Festival Chronicles

March's focus was on using AI tools like Midjourney and StableDiffusion to generate visual content for cultural celebrations such as Holi. This experience emphasized the importance of precise prompt engineering and human oversight to create culturally sensitive and visually appealing content.

- **AI Trials: Unveiling the Perils Within Beauty**
 - [AI Trials: March Part 1](#): Unstructured experimentation.
 - [AI Trials: March Part 2](#): Improved prompts for acceptable imagery.

April

April involved establishing baselines and benchmarks for AI-generated content using updated tools. The experiments focused on refining roles and templates, revealing improvements in content quality and the necessity of combining AI capabilities with human expertise for optimal results.

- **AI Trials: Baselines and Benchmarks**
 - [AI Trials: April Part 1](#): Role adaptation and content quality.
 - [AI Trials: April Part 2](#): Standardized templates and rating criteria.

- [AI Trials: April Part 2+](#): Evaluating AI-generated articles with enhanced roles

May

In May, I tested various tones in AI-generated content about Japan's Golden Week. The experiments showed significant variability in content quality based on tone selection, highlighting the need for precise prompting and iterative refinement to optimize AI output.

- **AI Trials: Style Simulations**

- [AI Trials: May Part 1](#): Exploring tonal variability in content.
- [AI Trials: May Part 2](#): Impact of tone on content quality.
- [AI Trials: May Part 3](#): Structured prompts and refined evaluation criteria.

June

In June, I assessed the ability of ChatGPT-4 and Claude to generate culturally themed content with consistent tones for events like the Dragon Boat Festival and Eid al-Adha. The AI showed diverse interpretations of tonal instructions in creating coherent and culturally appropriate content.

- **AI Trials: Mock Human**

- [AI Trails: June Part 1](#): Applying various tones to improve articles.
- [AI Trails: June Part 2](#): Testing tone consistency over multiple articles.
- [AI Trails: June Part 3](#): Documentation on June Pt 2's roles and templates.

July

In July, I tested a new GPT for role creation and explored JSON as an alternative to Markdown. Trials across five holidays yielded mixed results, with the new GPT showing

only marginal improvements. I also used Claude and ChatGPT-4o to create charts, significantly reducing data visualization time.

- **AI Trials: Most Experienced Professional**
 - [AI Trials: July Part 1](#): Role specificity and input format testing.
 - [AI Trials: July Part 2](#): Comparative analysis and data visualization.

August

This month focused on structured instructions and advanced prompting techniques: tested JSON-formatted roles, various prompting methods, and refined article templates using Claude 3.5 Sonnet and ChatGPT-4o. Experiments covered cultural celebrations like the Flooding of the Nile and Indonesia's Independence Day. JSON roles proved efficient, while exploring "super prompts" and Chain of Thought prompting showed potential for improving content quality, despite increased complexity.

- **AI Trials: Structured Instructions**
 - [AI Trials: August Part 1](#): JSON roles and template comparisons
 - [AI Trials: August Part 2](#): Standalone and super prompt exploration
 - [AI Trials: August Part 3](#): CoT prompting for template refinement

September

In September, I focused on refining AI-generated article templates for cultural events, particularly the Autumn Equinox celebrations. The experiments involved multiple iterations of template modification, article generation, and critique using Claude 3.5 Sonnet and ChatGPT-4o. Despite initial promise, the final articles fell short of expectations, highlighting the importance of thorough prompt engineering and test

The month's trials provided valuable insights into the challenges of AI-assisted content creation and the impact of rushed processes on output quality.

- **AI Trials: Multi-Methodologies**

- [AI Trials: September Part 1](#): Revisiting CoT for template improvement.
- [AI Trials: September Part 2](#): Evaluating new CoT rubrics.
- [AI Trials: September Part 3](#): Testing diverse subjects and markup language.
- [AI Trials: September Part 4](#): Refining templates for Autumn Equinox, revealing prompt engineering challenges.

October

In October, I refined article templates and tested tone variations across AI models. I analyzed 16 articles with different tones, created 14 pieces on Halloween's history, and found professional tones produced the best content. I also grappled with creating universal templates that work across AI models.

- **AI Trials: Measuring Sound**

- [AI Trials: October Pt 1](#): Refining article templates with three AI models.
- [AI Trials: October Pt 2](#): Testing tone variations across 16 articles
- [AI Trials: October Pt 3](#): Creating historically influenced Halloween articles.

November

In November, I explored how different prompt structures affect AI content generation. Starting with basic prompts and progressing to more sophisticated approaches, I tested how varying levels of complexity impacted article quality. The experiments demonstrated the value of structured prompts while revealing scoring inconsistencies.

between AI models. This led to insights about prompt design and the importance of balancing complexity with practical effectiveness.

- **AI Trials: Self ReAlization**

- [AI Trials: November Pt 1](#): Testing baseline prompt complexity effects.
- [AI Trials: November Pt 2](#): Evaluating role definition approaches.
- [AI Trials: November Pt 3](#): Examining AI-generated scoring systems.

December

In December, I concluded the year-long experiment by examining how different prompt elements interact. Through systematic testing of combinations and complexity levels, I explored how various prompt components work together to enhance output quality. The experiments revealed patterns in how strategic element combinations can compensate for limitations and demonstrated the value of layered context in prompt design.

- **AI Trials: Almost Autonomous**

- [AI Trials: December Pt 1](#): Measuring layered complexity effects.
- [AI Trials: December Pt 2](#): Testing strategic element combinations.

Quiet Evolution is about experimenting and sharing insights. If you find this helpful, [coffee is always appreciated](#) (no pressure). Proceeds are used strictly to cover AI costs; any excess goes to the [American Cancer Society](#).

Disclaimer: *The views and opinions expressed in this article are solely those of the author and do not reflect the official policy or position of [Amazon Web Services \(AWS\)](#). The author is a UX designer at Amazon Web Services (AWS) and has no*

involvement in, nor does their work pertain to, any collaborative agreements that [A](#) may have with [Anthropic](#), the creators of Claude. The insights and analyses presented here are entirely independent and unrelated to any projects or initiatives between AWS and Anthropic. All content in this post is based on publicly available interfaces and is not influenced by the author's employer.

← Previous

Next →

Discussion about this post

Comments

Restacks



Write a comment...