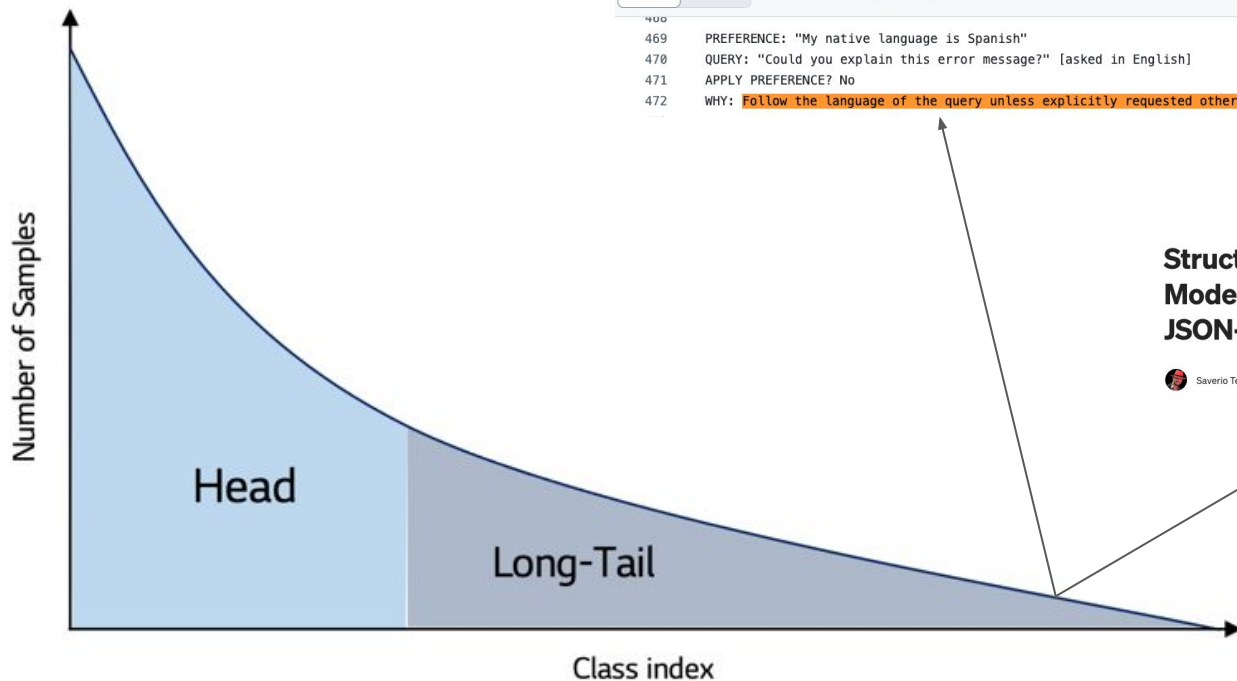


# Treasure Hunt: Real-time Targeting of the Long Tail using Training-Time Markers

---

Daniel D'souza, Julia Kreutzer, Adrien Morisot,  
Ahmet Üstün, Sara Hooker

# What about the Long Tail?



```
Code Blame 653 lines (527 loc) · 100 KB
469 PREFERENCE: "My native language is Spanish"
470 QUERY: "Could you explain this error message?" [asked in English]
471 APPLY PREFERENCE? No
472 WHY: Follow the language of the query unless explicitly requested otherwise.
```



*Prompt Engineering*

**Structured Output with Gemini Models: Begging, Threatening, and JSON-ing**



Saverio Terracciano Follow 7 min read · Apr 8, 2025



*Other methods\**

## Research Question

*“Can we **optimize our training protocols** to both improve controllability and performance on underrepresented use cases at inference time?”*

# Markers

## Taxonomy

```
<task>          ==> CodeGeneration, CodeTranslation, Summarization, ...  
<domain>        ==> Code, Math, Science, Legal, ...  
<language>      ==> English, French, German, Spanish, Hindi, ...  
<length_bucket> ==> concise, medium, long  
< ... >         ==> ...
```

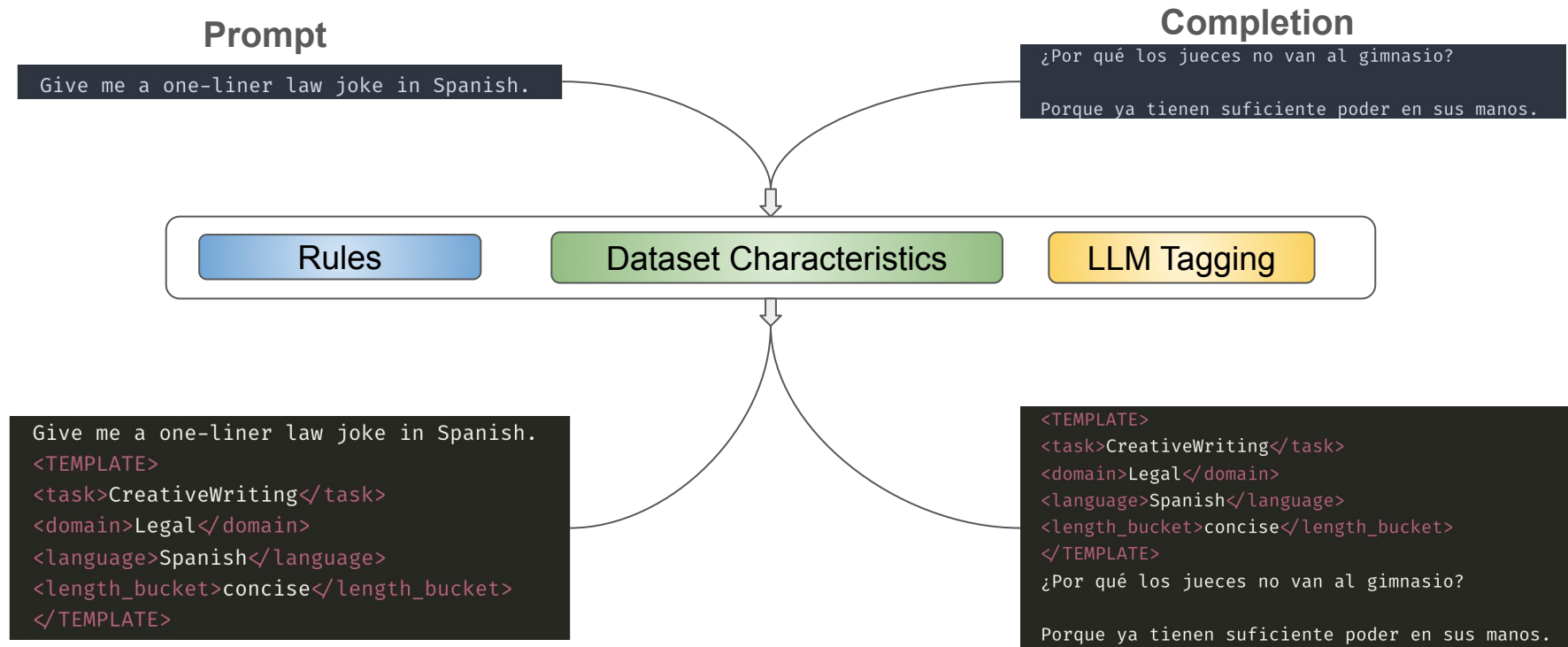
## Annotation

Rules

Dataset Characteristics

LLM Tagging

# Modifying the Dataset



# Marker Dropout

Give me a one-liner law joke in Spanish.

```
<TEMPLATE>  
<domain>Legal</domain>  
</TEMPLATE>
```



```
<TEMPLATE>  
<task>CreativeWriting</task>  
<domain>Legal</domain>  
<language>Spanish</language>  
<length_bucket>concise</length_bucket>  
</TEMPLATE>
```

¿Por qué los jueces no van al gimnasio?

Porque ya tienen suficiente poder en sus manos.

Give me a one-liner law joke in Spanish.

```
<TEMPLATE>  
<task>CreativeWriting</task>  
</TEMPLATE>
```



```
<TEMPLATE>  
<task>CreativeWriting</task>  
<domain>Legal</domain>  
<language>Spanish</language>  
<length_bucket>concise</length_bucket>  
</TEMPLATE>
```

¿Por qué los jueces no van al gimnasio?

Porque ya tienen suficiente poder en sus manos.

Give me a one-liner law joke in Spanish.

```
<TEMPLATE>  
<language>Spanish</language>  
<length_bucket>concise</length_bucket>  
</TEMPLATE>
```



```
<TEMPLATE>  
<task>CreativeWriting</task>  
<domain>Legal</domain>  
<language>Spanish</language>  
<length_bucket>concise</length_bucket>  
</TEMPLATE>
```

¿Por qué los jueces no van al gimnasio?

Porque ya tienen suficiente poder en sus manos.

# “TreasureMarked” Model

## “Conditioning” generations at inference



How do I check for a palindrome in python?



<TEMPLATE>

<code\_type>python</code\_type>

<domain>Code</domain>

<source>AA</source>

<task>CodeGeneration</task>

</TEMPLATE>

A palindrome is a word, phrase, number, or other that reads the same forward and backward (ignoring punctuation and capitalization). To check if a given string is a palindrome, you can use the following code:

```
def is_palindrome(s):
```

```
    """
```

```
    Check if a string is a palindrome.
```



Can you give me a 3 sentence paragraph on the French Revolution ?



<TEMPLATE>

<domain>SocialSciences</domain>

<lang>English</lang>

<length\_bucket>concise</length\_bucket>

<length\_paragraphs>1</length\_paragraphs>

<length\_sentences>3</length\_sentences>

<length\_tokens>81</length\_tokens>

<quality>4.4</quality>

<quality\_bucket>2</quality\_bucket>

<source>CC</source>

<task>QuestionAnswering</task>

</TEMPLATE>

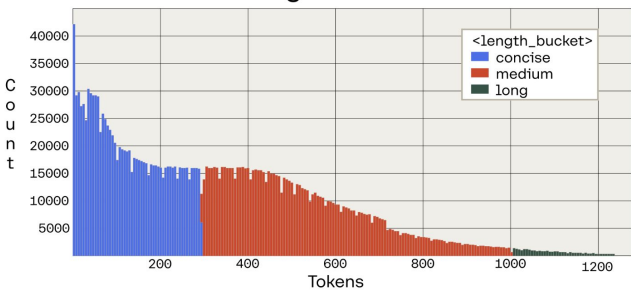
The French Revolution, which took place between 1789 and 1799, was a period of profound social and political upheaval in France. It marked the end of the French absolute monarchy and the creation of a republic, driven by ideals of liberty, equality, and fraternity. The revolution had a significant impact on European history, spreading the ideas of Enlightenment and leading to the Napoleonic Era and subsequent modernization of European political systems.

# TreasureHunt

*“Controlling” generations at inference*

CONTROL AT INFERENCE

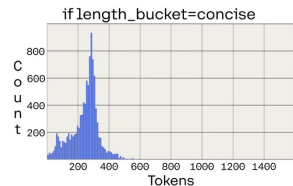
Training Distribution



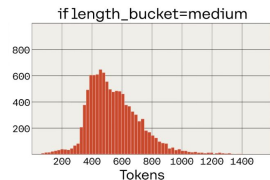
At inference time, **changing a single marker** can lead to vastly different test-time distributions

Distribution at Inference

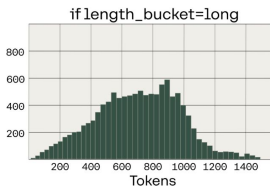
```
Tell me about the French Revolution
<TEMPLATE>
<length_bucket>concise</length_bucket>
</TEMPLATE>
```



```
Tell me about the French Revolution
<TEMPLATE>
<length_bucket>medium</length_bucket>
</TEMPLATE>
```



```
Tell me about the French Revolution
<TEMPLATE>
<length_bucket>long</length_bucket>
</TEMPLATE>
```





# TreasureHunt

*“Controlling” generations at inference*



## Reward Model Annotation Markers

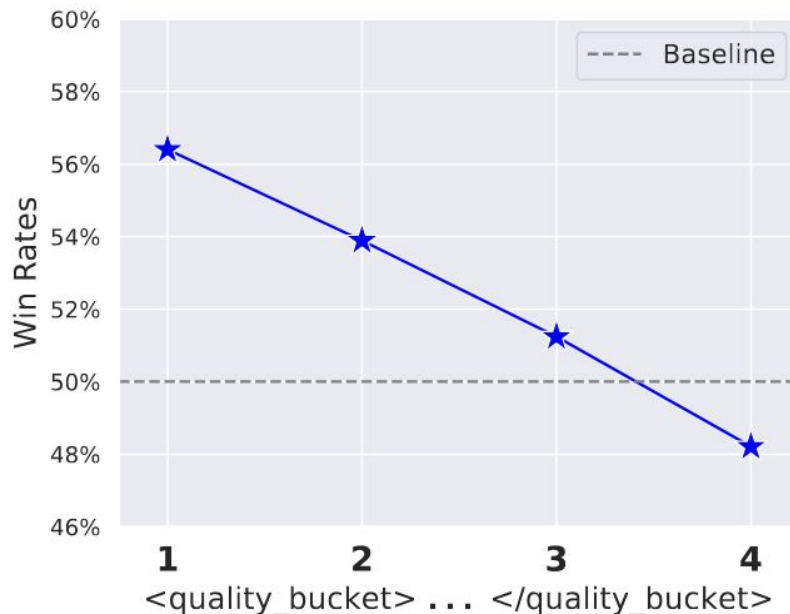
<quality\_bucket> 1 ⇒ **Highest** Quality

<quality\_bucket> 2 ⇒ ...

<quality\_bucket> 3 ⇒ ...

<quality\_bucket> 4 ⇒ **Lowest** Quality

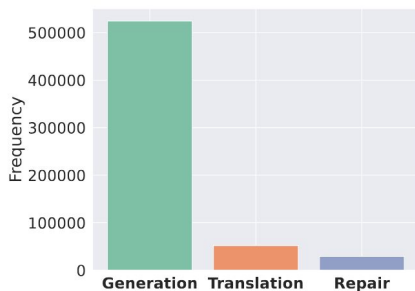
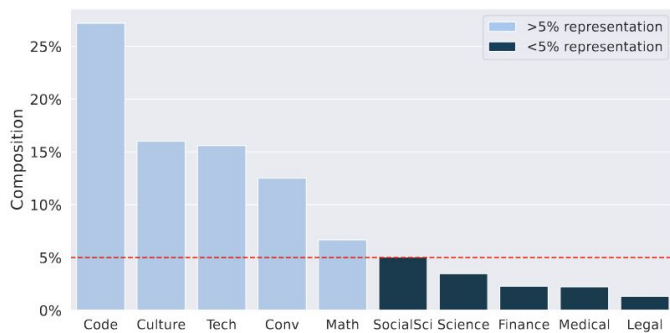
*mArenaHard Win Rates against Baseline Model*



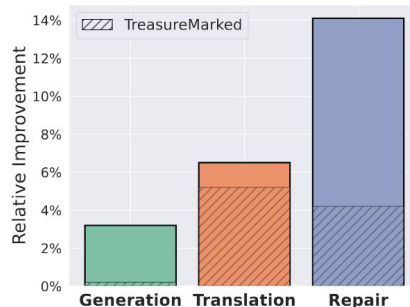
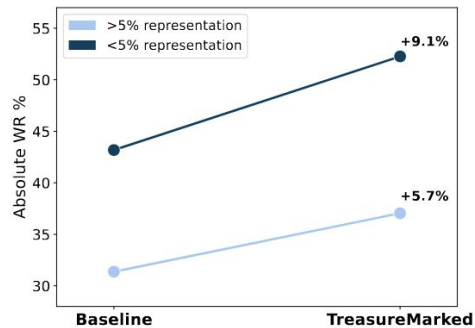
# Results

*The long tail benefits the most!*

## Training Distribution

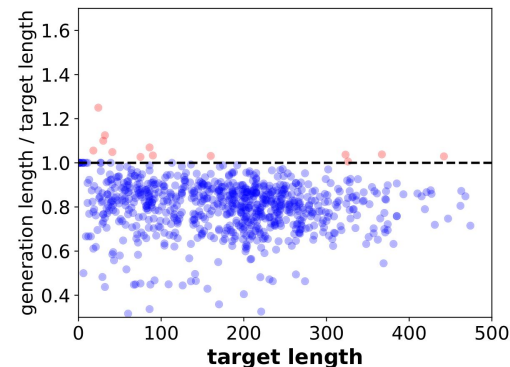
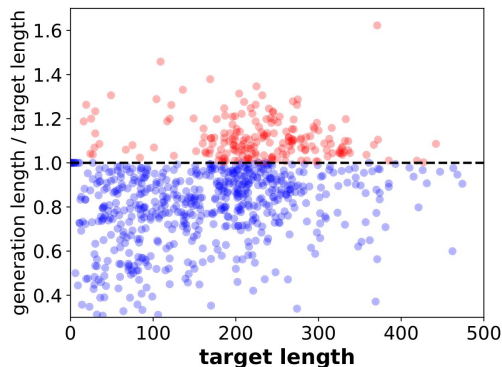
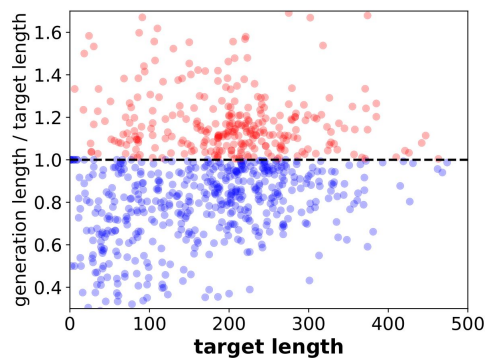


## Improvement at Inference



# Results

## Alpaca Eval Length Instruct

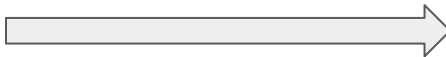


	Baseline	TreasureMarked	+fixed
<i>Violation Rate(%)</i>	36.6 %	24.7 % ↓	1.3 % ↓
<i>Win Rate(%)</i>	14.4 %	19.5 % ↑	21.2 % ↑

# Takeaway



Map Markers during Training



Uncover Gains at Inference