

CoCoMix (ours)

GPT2 SAE

**Input Prompt:** Latent concept modeling is a

**No Steer**

**Generation:** technique that allows you to model a complex system in a way that is easy to understand and easy for the user to manipulate.

**Steer concept 6781:** politic/law related

**Generation:** high-level policy: The White House does not have a policy on <REMOVED>'s plans. The White house said that <REMOVED> and his team had not discussed the matter with

**Steer concept 8984:** aggressive tone

**Generation:** thing? I'm not sure what is going on here.

**Generation:** very important part of the design process.The first step is to create a model that is easy to use and easy for the user to understand.

**Generation:** common practice in the world of computer science. The new law, which was introduced in 2010, requires that all data in public use be treated as such.

**Generation:** thing of the past. And that's what we're talking about here. A concept?