

## Expensive Pre-training

**E.g., Llama** [Touvron et al, 2023]

- 65B parameters
- 2,048 A100 GPUs
- >\$2.4m
- 21 days
- >1,000 tons CO<sub>2</sub>



Generative LLM  $\theta_0$

## Proposed Framework:

## Independent Interventions for Language Models

training data



deployment



forget training data

**Unlearning**

“Swine Flu” “COVID”



“Latest US pandemic”

**Knowledge Editing**

16 bit

4 bit

quantize



prune

**Compression**

deployment

$\theta_0 \rightarrow$  **Kn. Edit**  $\rightarrow \theta_1 \rightarrow$  **Compress**  $\rightarrow \theta_2 \rightarrow$  **Unlearn**  $\rightarrow \theta_3 \dots$

**Composable Interventions for Language Models**