Preparation Stage

Edited heads Identification

Training Set

Intact Input I: The patient presents the symptoms of…, **pain, shortness of breath, fatigue**, …

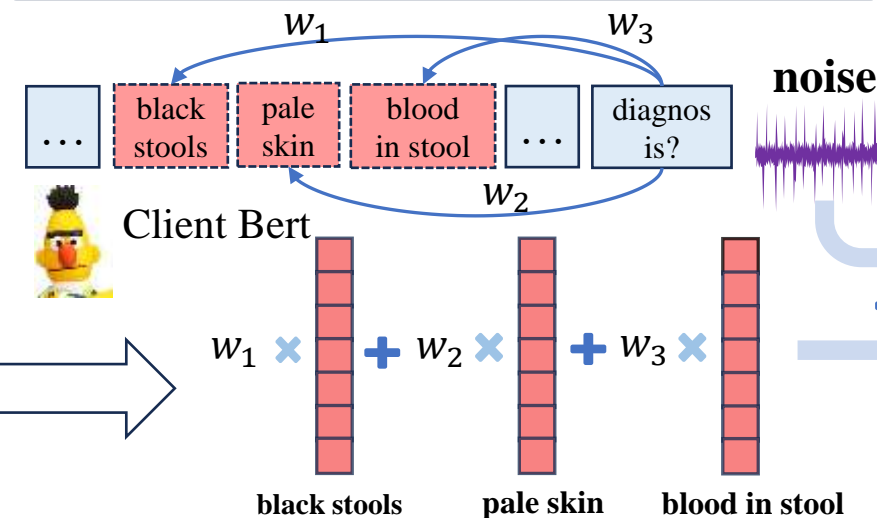I: The patient presents the symptoms of …, **pain, shortness of breath, fatigue**, … .

$\hat{I}$: The patient presents the symptoms of …, **pain,** ▮▮▮▮ **, fatigue**, … .

Remove Privacy Spans

LLM

Probe Technique

Top-K heads sets for each private span

Common top-K heads set

Top-K heads Selector
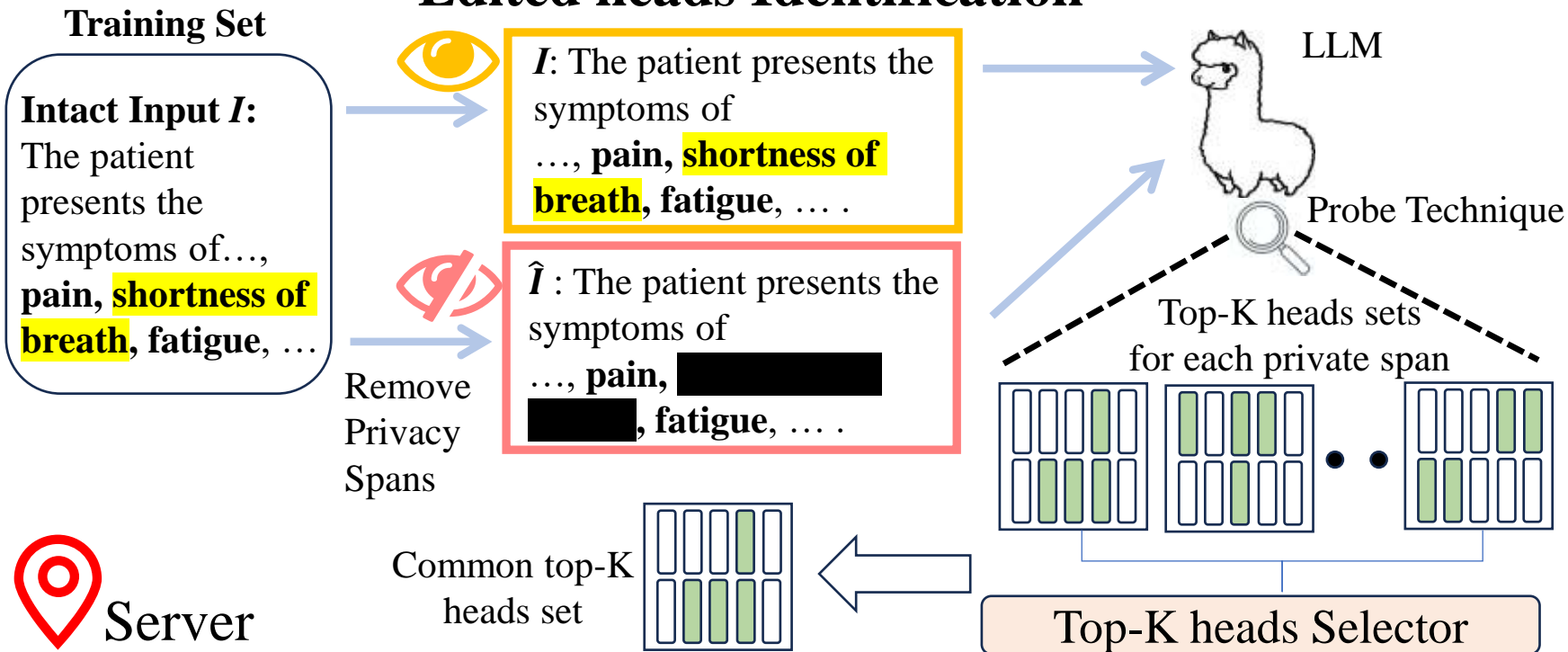
Server

Restoration Vectors Training

Training

LLM

❄ **Frozen**

Training Set

Intact Input (I)
+ Incomplete Input ($\hat{I}$)
+ Output Given I (a)
+ Output Given $\hat{I}$ ($\hat{a}$)

🔥 **Trainable**

Restoration Vectors for each privacy span (plug-and-play)

shortness of breath

black stools

pale skin
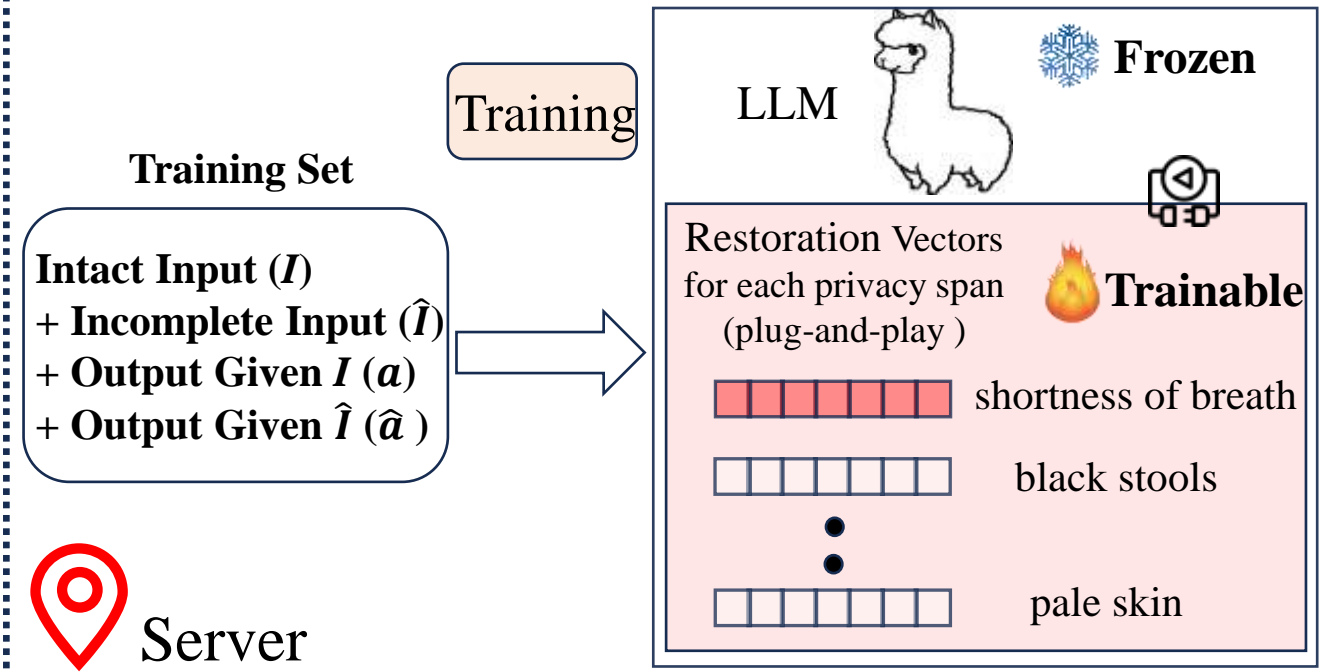
Server

Inference Stage

Meta Vector Construction

User Query q : Lately, **My stools sometimes is black** and I have abdominal pain, and **pale skin**. Additionally, I've even noticed **blood in my stool**. What is the likely diagnosis?

Users identify privacy spans

Client

Attention-aware Weighted Aggregation

w_1 w_3

black stools   pale skin   blood in stool   diagnosis?

w_2

Client Bert

noise

$w_1 \times$ + $w_2 \times$ + $w_3 \times$

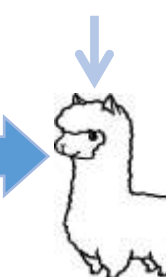black stools   pale skin   blood in stool

+

Transmission to Server

Meta Vector

User Query with Privacy Spans removed

Incomplete Query $\hat{q}$: Lately, ▮▮▮▮ and I have abdominal pain, and ▮▮▮▮. Additionally, I've even noticed ▮▮▮▮.What is the likely diagnosis?

Privacy Restoration

Meta Vector

LLM

Server

LLM Output O: According to the patient's symptoms, the illness is **Anemia**. This is potentially caused by conditions such as: …

✓