

# SPAN, TRACE, SESSION

The three levels you trace and evaluate in Arize



## SPAN

### WHAT IT IS

A single atomic operation: one model call, one tool call, or one retrieval step.

### WHAT YOU EVALUATE

Whether that step did its job – was the retrieval relevant, did the tool return the right result.



## TRACE

### WHAT IT IS

All the spans for one user-facing request – the full path from input to final answer.

### WHAT YOU EVALUATE

Did the whole turn land – a grounded, correct, on-task response built from those steps.



## SESSION

### WHAT IT IS

A collection of traces – the whole multi-turn conversation between a user and the app.

### WHAT YOU EVALUATE

Did the user reach their goal – coherence across turns, resolution, and signs of frustration.