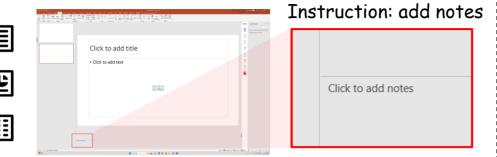
Stage 1: Divide Modalities



In this UI screenshot, please find the **TEXT** element corresponding to the command "{}" (with bbox). **Only focus on text content, ignore icons**.



In this UI screenshot, please find the **ICON** element corresponding to the command "{}" (with bbox). **Only focus on graphical icons, ignore pure text.**

Stage 3: Select Answer

You are given a UI screenshot and a user command: add notes

There are two candidate UI elements:

Candidate 1 (text-based): [718, 741, 821, 887]
Candidate 2 (icon-based): [0, 444, 524, 555]

Candidate 2 (icon-based): [0, 444, 524, 555]
Based on the command and the description of candidates, choose which candidate (1 for text, 2 for icon) better matches the command.

Just answer with '1' or '2'.

Stage 2: Dynamic Zooming

