

GPT-2 CODI: Attention Distribution Across Layers
Early-Middle Layers (4-8) Show Concentration

