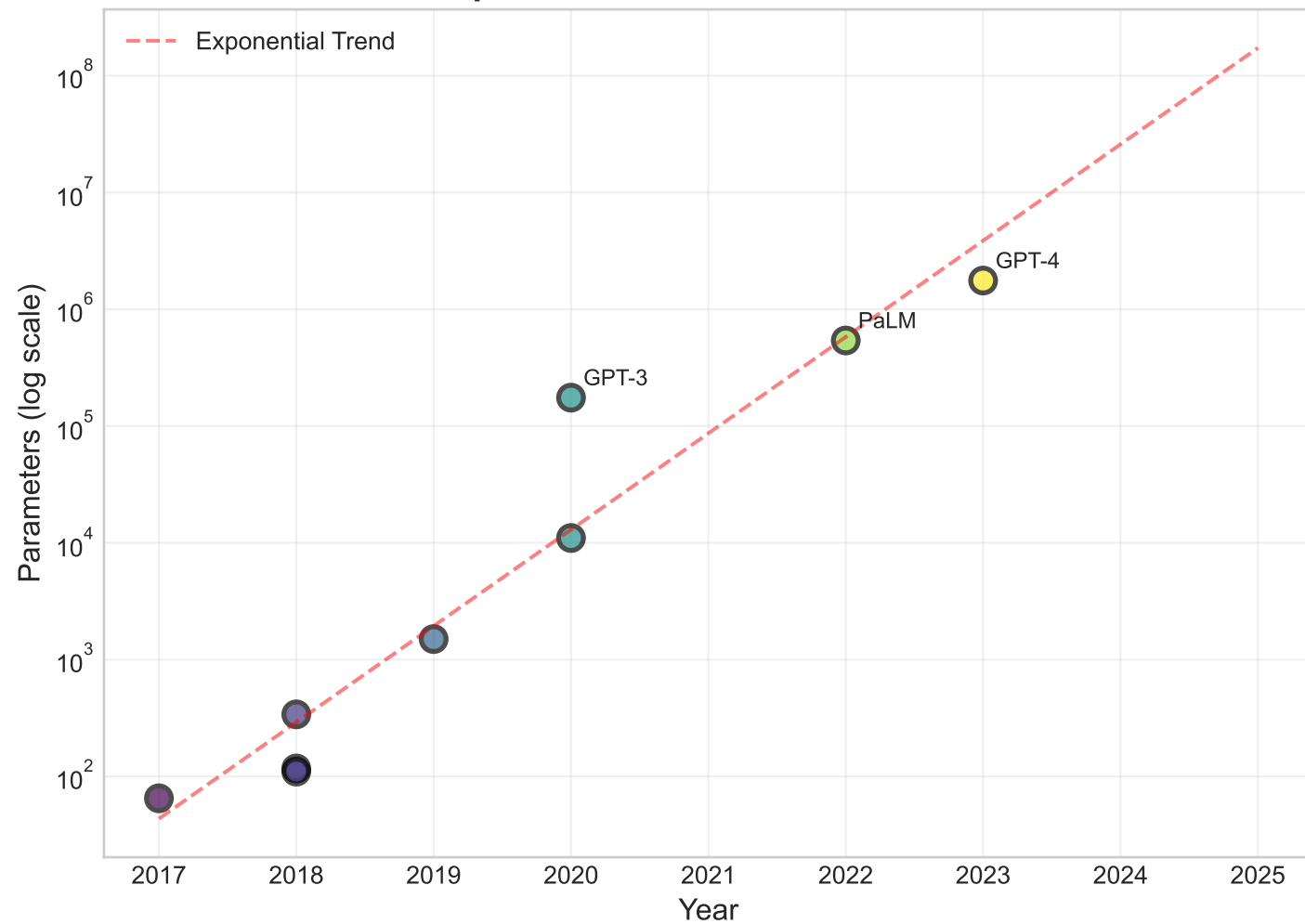The Race to Scale: Transformer Model Evolution

Exponential Growth in Model Size

Relative Model Sizes (GPT-1 = 1)

Circle area proportional to parameter count