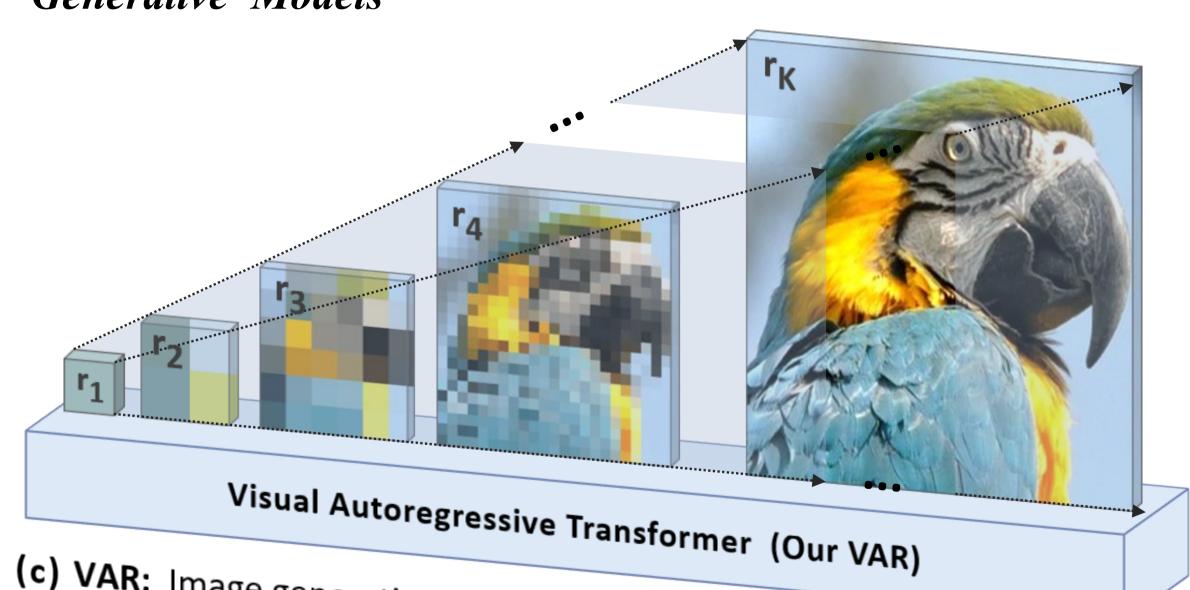
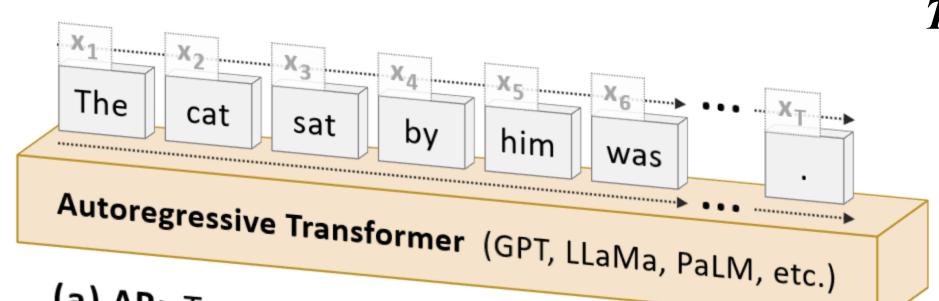
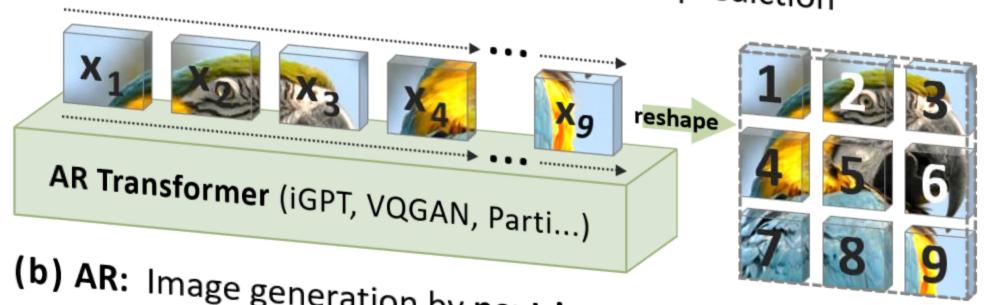
Three Different Autoregressive Generative Models



(c) VAR: Image generation by next-scale (or next-resolution) prediction



(a) AR: Text generation by next-token prediction



(b) AR: Image generation by next-image-token prediction