

## "Attention Is All You Need" (Vaswani et al., 2017)

Revolutionary Architecture for Deep Learning

### Core Innovation

#### ⚡ Eliminated recurrence entirely

Enabled full parallelization

#### 📈 Massive dataset training

Efficient at scale

#### 🎯 State-of-the-art performance

Across multiple benchmarks



### Foundation Models

GPT, BERT, and modern LLMs



### NLP Revolution

Translation • Summarization  
Question Answering • Generation



### Beyond NLP

Vision Transformers (ViT)  
Audio Processing • Multimodal AI



### Impact

New era of AI capabilities  
Industry-wide transformation