

# The Exponential Evolution of NLP Technology

Model Size:

1K

1M

1B

1T

Relative Performance (log scale)

$10^4$

$10^3$

$10^2$

$10^1$

$10^0$

$10^{-1}$

1985

1990

1995

2000

2005

2010

2015

2020

2025

Year

N-grams

Neural LM

LSTM

Word2Vec

BERT

Transformer

GPT-2

GPT-3

Claude-3

10,000x  
improvement  
in 30 years!

Few-shot Learning  
Emerges

Attention Revolution  
Begins

Semantic Understanding  
Unlocked

Statistical Era

Early Neural Era

Embedding Era

Transformer Era