生成AIができるまで（文章で学び、人の評価で整える）

対話型の生成AIは、どのような段階を経てつくられるのだろう

1 データを集める
本・Webなどの大量の文章（権利・個人情報・かたよりに配慮が必要）

→

2 事前学習
次のことばを予測する練習をくり返し、ことばのパターンを学ぶ

→

3 追加の学習
質問と答えの例で、指示にそって答えるように学ぶ

→

4 人の評価で整える
人が答えのよしあしを評価し、よい答えを出しやすくする

→

5 試験して公開する
危ない答えを出さないか試す。公開後も見直しを続ける

2 事前学習のようす（模式）

「春になると桜の花が ___」



大量の文章で「次に来ることば」を当てる練習をくり返す。正解とのずれが小さくなるように、中の数（パラメーター）を少しずつ直す。

4 人の評価で整える（模式）

同じ質問への2つの答えを、人がくらべて選ぶ

答えA：短く正確で、ていねい　○ よい

答えB：もっともらしいが不正確　×

たくさんの評価をもとに、よい答えを出しやすく調整する

それでも残る課題

- もっともらしいまちがい（ハルシネーション）
- 学習した時点より後のことは知らない
- 集めたデータや評価する人のかたよりが答えに出る
- 何を学んだかが見えにくい

※ 模式図。段階の分け方は代表的な例で、開発する組織によってちがう。確からしさの数は架空。