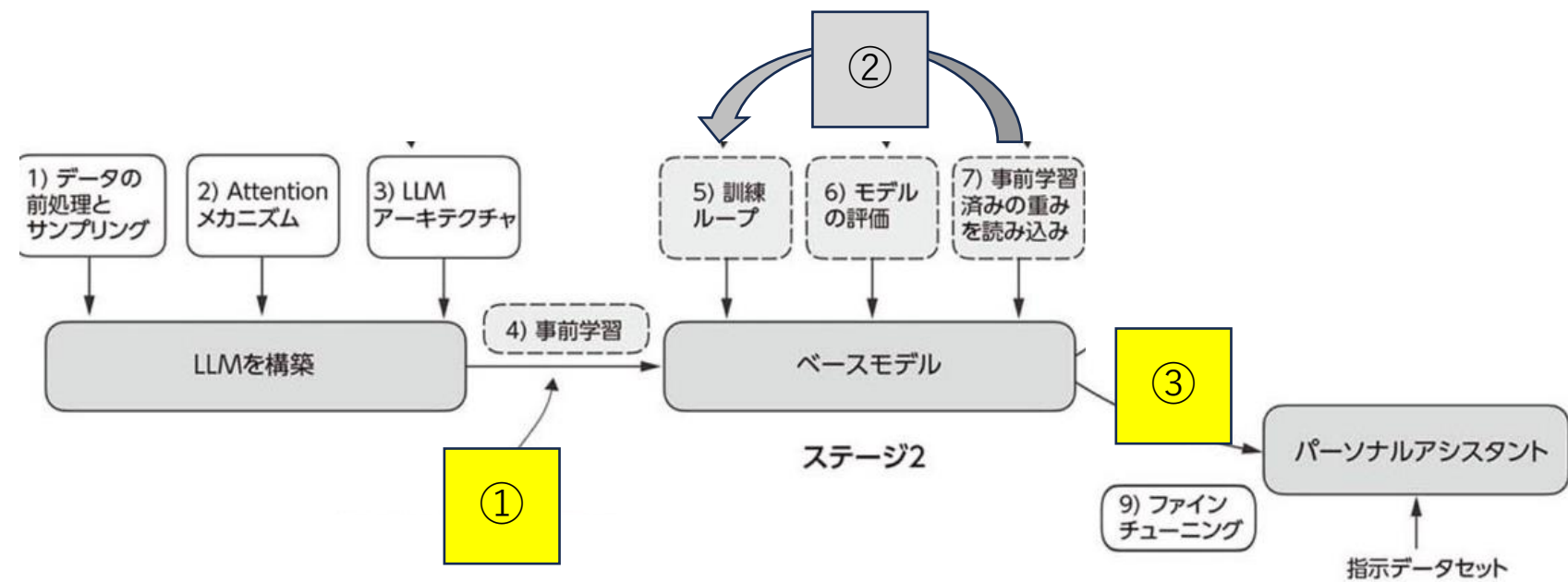


自作LLMの学習プロセス



出典
つくりながら学ぶ！LLM 自作入門



①事前学習：rinna/japanese-gpt-neox-3.6b

膨大なテキストデータを用いて学習を行い、言語の一般的なパターンや文脈理解能力を身につける。
義務教育のように基礎知識を習得。

②継続事前学習：今回は割愛

特定分野（医療・法律など）の言語感覚や専門用語にドメイン知識の下地を作る。
高校で理系・文系に分かれるように慣れさせる段階。

③ファインチューニング：LoRA (Low-Rank Adaptation)

特定タスク（質問応答・分類など）の精度向上のための調整を行う。
大学で専門性を身につけるように、業務遂行スキルを磨く段階である。

Transformerブロックの構造

Transformerブロックの中心には**Masked Multi-head Attention**がある。
前頁①の事前学習gptモデルは36層以上のこのブロックを積み重ねている。
特に「**Masked**」の仕組みにより、「理解しよう」として学んだのではなく
「**生成のため**」に理解が自然と育った、という天才的な仕掛け。

Self-Attention

文章の中の各単語が、自分自身を含めた文章中のすべての単語に注目する仕組みである。**アドリブ**がきく賢い社員のように、周囲の状況を把握しながら**最適な発言**ができる。

