

# LLM 開発を支える多様な Fine-Tuning: PFN での取り組み

中鉢魁三郎

株式会社 Preferred Networks

2025/03/14 @ 第1回「大規模言語モデルのファインチューニング技術と評価」ワークショップ

# We are hiring !!

生成 AI をはじめとする AI 技術の実用化を進めていただける仲間を募集中！

## 私たちのプロダクト・サービス

**PLaMo**

純国産フルスクラッチ基盤モデル **PLaMo**

**A** **文** 世界最高クラスの日本語性能

**●** 純国産フルスクラッチモデル

**API** APIで簡単に導入可能

- フラグシップモデルのPLaMo Prime、小型軽量のPLaMo Liteを提供中
- 次期バージョンPLaMo2も鋭意開発中  
<https://tech.preferred.jp/ja/blog/>

**PreferredAI**

**PreferredAI Talent Scouter:** AIアバターとの対話で実務の適性評価と人材採用を支援

**PreferredAI Insight Scan:** 口コミ/アンケート等の大量の文章を賢く分類、インサイトを発掘

**PreferredAI Notes:** PDFやスライドなどの社内データを元に数秒であらゆる文書を生成

など、大規模言語モデルを活用したプロダクトを開発しており、提供中、もしくは提供開始予定です



## 募集中のポジション一覧

### 【エンジニアリング】

- 大規模言語モデルアラインメント(指示学習)エンジニア
- 大規模言語モデルプラットフォーム開発エンジニア (アプリケーション・インフラ)
- プロダクトマネージャー (大規模言語モデルプラットフォーム)
- 大規模言語モデルプラットフォーム開発テックリード(機械学習・自然言語処理)
- 基盤モデルサービス開発エンジニア (大規模言語モデル / 画像基盤モデル)

### 【ビジネス開発】

- ビジネス開発 (LLMを活用したプロダクトの新規立ち上げ)

最新の採用情報はこちら→  
<https://www.preferred.jp/ja/careers/>



# 自己紹介

## 中鉢 魁三郎 (Kaizaburo Chubachi)

### 所属

- Preferred Networks エンジニア

### 略歴

- 東北大学情報科学研究科 修士課程修了
  - 篠原・吉仲研究室 質問学習
- 2019年 Preferred Networks 入社
  - 産業向けソリューションの研究開発
  - LLM Alignment チームのマネージャー



# Preferred Networks (PFN) 会社概要

## ミッション： 現実世界を計算可能にする

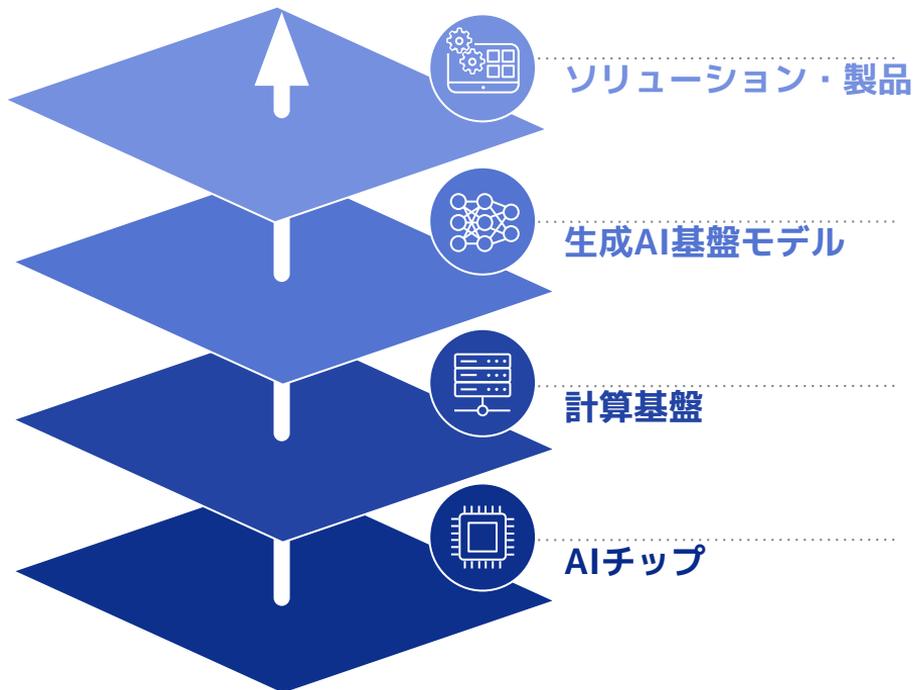
- 設立** 2014年3月26日
- 本社** 東京都千代田区
- 代表取締役** 西川徹（最高経営責任者）、岡野原大輔（最高研究責任者）
- 従業員数** 約350名（2024年9月）
- 事業内容** AIチップ、計算基盤、生成AI・基盤モデルなどのAI関連技術を活用したソリューション・製品の開発・販売および研究開発
- 主要子会社** 株式会社Preferred Computational Chemistry（2021年6月）  
株式会社Preferred Robotics（2021年11月）  
株式会社Preferred Elements（2023年11月）  
株式会社Preferred Computing Infrastructure（2025年1月）
- 出資企業** SBIグループ ENEOSイノベーションパートナーズ合同会社  
積水ハウス投資事業有限責任組合 中外製薬株式会社  
TEL Venture Capital, Inc. トヨタ自動車株式会社  
株式会社日本政策投資銀行 日本電信電話株式会社  
株式会社博報堂DYホールディングス 株式会社日立製作所  
ファナック株式会社 株式会社みずほ銀行 三井物産株式会社  
三菱商事株式会社 株式会社ワコム



<https://www.preferred.jp>

# PFNの事業: AI技術のバリューチェーンを垂直統合

PFNは、チップ、計算基盤、生成AI基盤モデル、ソリューション・製品まで、AI技術のバリューチェーンを垂直統合し、ソフトウェアとハードウェアを高度に融合することで、競争力の高い技術の開発および産業応用を進めています。



様々な産業・消費者向けのソリューション・製品



**PLaMo**

PLaMo Prime (国産LLM)  
PLaMo Lite (エッジ向けSLM)

**PFP**

物質のエネルギー計算モデル



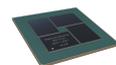
GPUクラスタ



MN-3  
(MN-Core™  
クラスタ)



MN-Core™ 2を  
計算資源とした  
クラウドサービス



MN-Core™



MN-Core™ 2



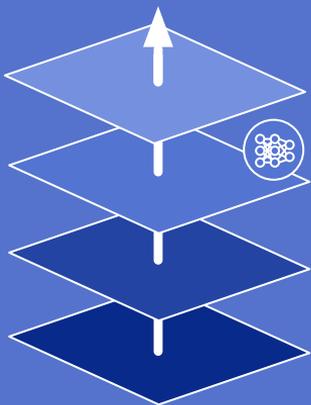
生成AI (推論) 向け  
MN-Core L1000  
(2026年提供予定)



MN-Core  
次世代



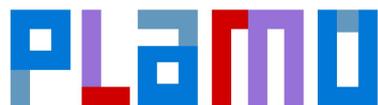
## 生成AI・ 基盤モデル



# PLaMo™

## PFNグループの生成AI基盤モデル

PLaMoは、PFNのグループ会社Preferred Elements (PFE)がフルスクラッチで開発する純国産の生成AI基盤モデルです。高品質な学習データを用い、既存の大規模言語モデルをベースに用いずに全く新規に開発。主要な日本語ベンチマークで世界最高レベルの精度を示しています。



- 商用版PLaMo Primeを2024年12月にリリース
- エッジ向け小規模言語モデル PLaMo Lite を提供中
- 金融・医療などの特定タスクに強いモデルも順次リリース予定

### 世界最高クラスの日本語性能

- 主要な日本語ベンチマーク\*でGPT-4を超える精度を記録
- 現在も精度向上中
- 日英の翻訳性能にも優れる

\*日本語LLMの性能を測るため標準的に使われるベンチマーク「Jaster」を使用

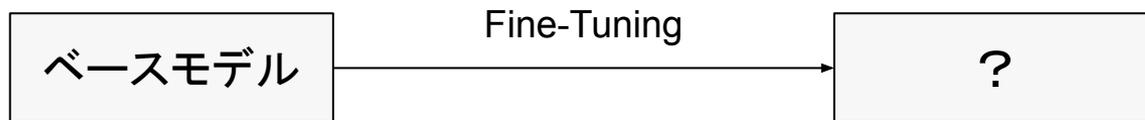
### 純国産フルスクラッチモデル

- 独自のアーキテクチャと学習データでゼロから事前学習・指示学習を実施
- 他社モデルベースではないため社外ライセンスの縛りや開発上で不明瞭な点がない

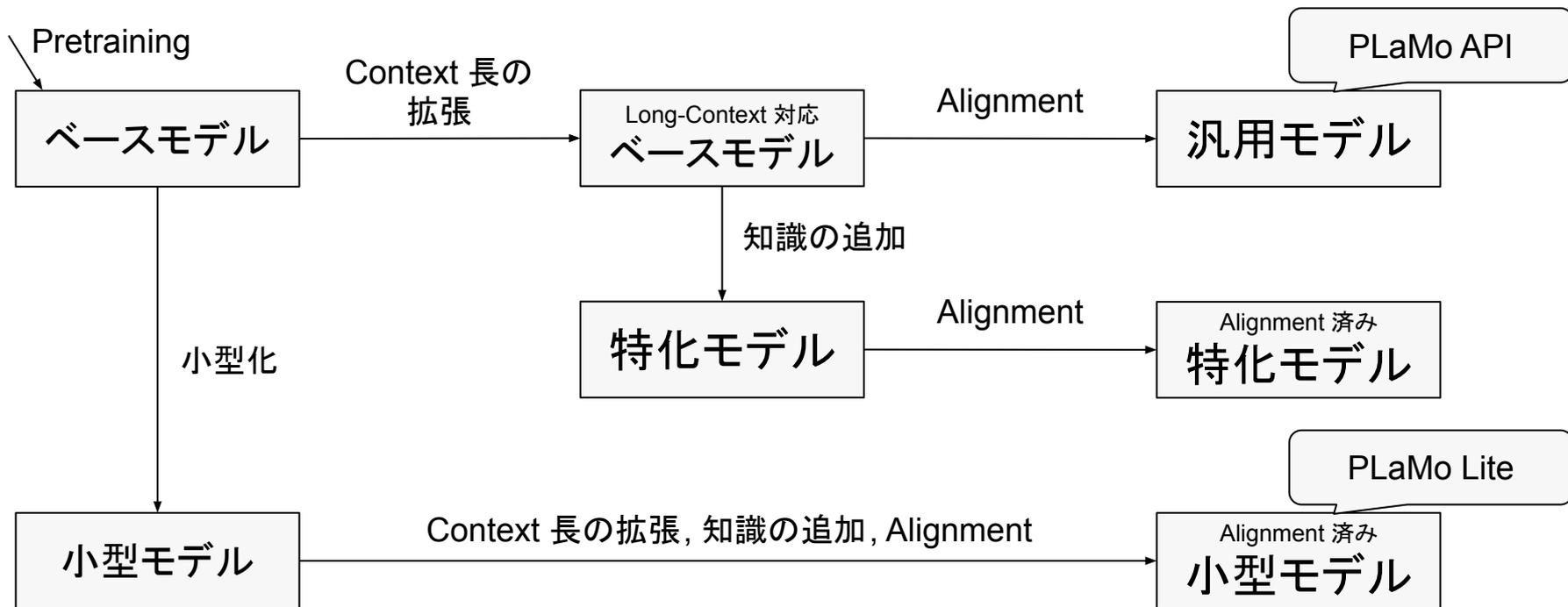
### APIで簡単に導入可能

- API経由でクラウドを介して提供、簡単なコードの書き換えで実装可能
- モデル自体をお客様のクラウドなオンプレミス環境に実装することも可能

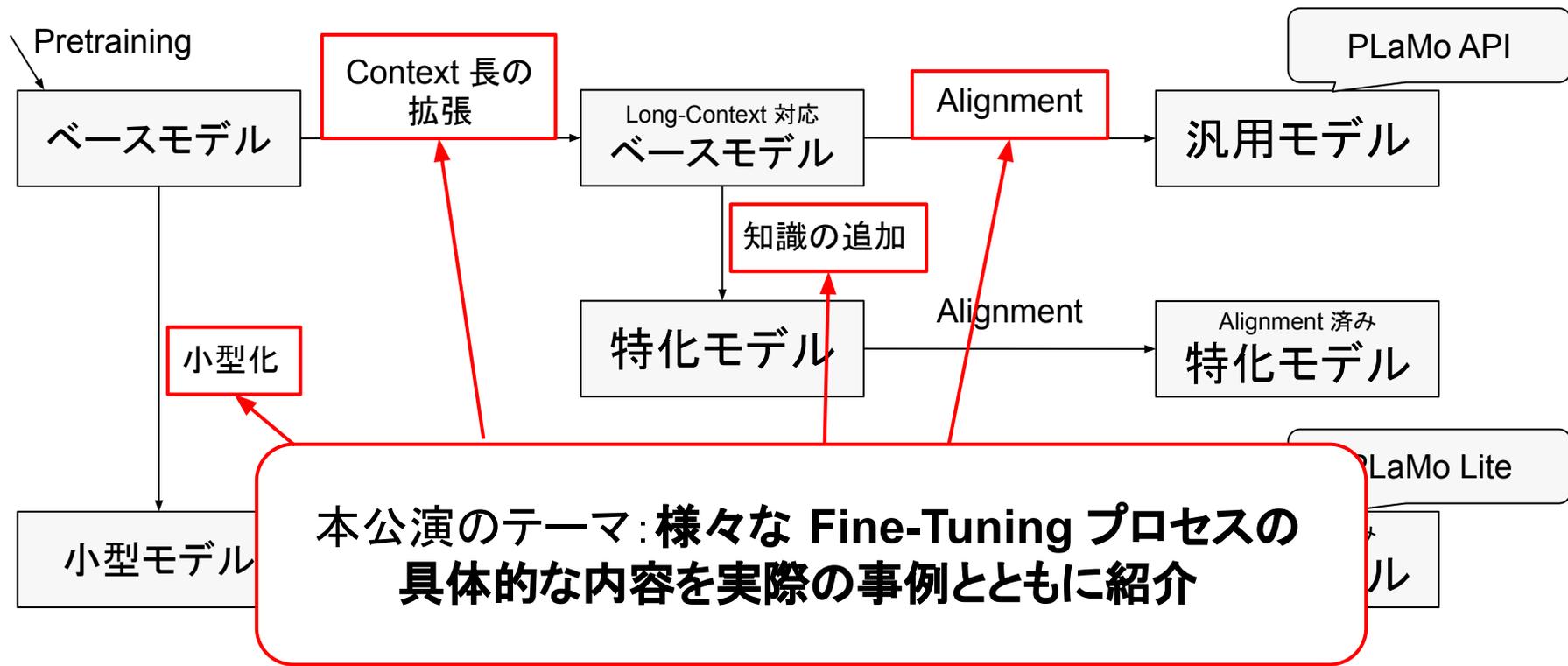
# Why Fine-Tuning



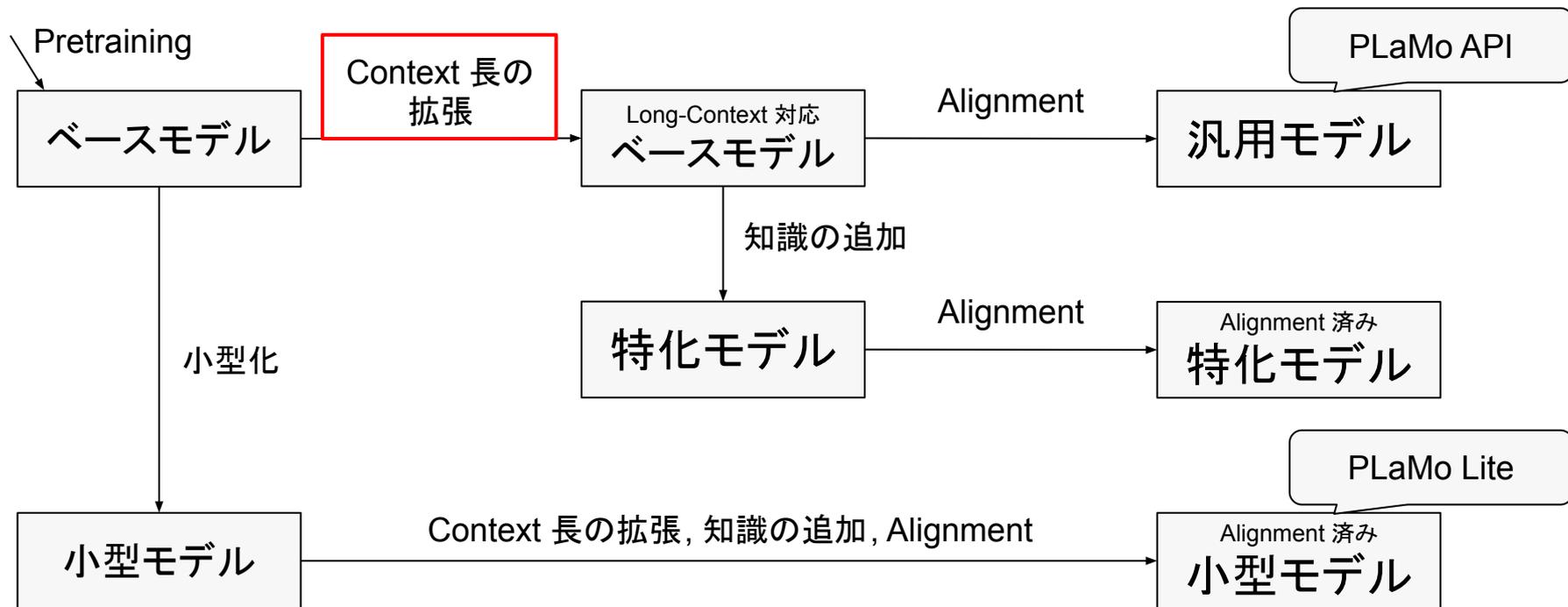
# PFN における Fine-Tuning



# PFN における Fine-Tuning



# PFN における Fine-Tuning



# Context 長の拡張

Pretraining は 4k-token で行い、後で Context 長を拡張するのが一般的  
典型的には以下の2ステップで行う

## Positional Encoding の修正

- Rotary Position Embedding は離れた Token 間の Attention Score が小さくなりやすい。Base Frequency を変更しこれを軽減 ([ABF RoPE, Xiong et al., 2023](#))

## Continual Pretraining

- Long-context なデータを使って Pretraining と同じ損失関数で訓練する
- 4k -> 32k -> 64k -> 128k -> 256k のように複数回の Stage で学習する場合も ([Qwen2.5, Yang et al., 2024](#))

# PLaMo Prime における Context 長の拡張

ABF RoPE を用いて 4k-token から 4倍の 16k-token まで拡張した

- [PLaMo-100B の arXiv 本文](#) や [夏目漱石著の『夢十夜』の全文](#) が収まる
- Alignment Phase の SFT で Long-context を学習
  - 公開データセットや合成した長文の翻訳タスクを活用

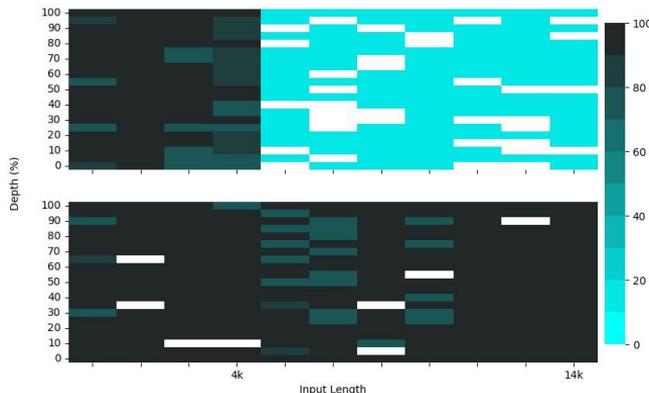
LongBench v1 に含まれる 10 の商用  
可英語データセットの平均スコア

	LongBench v1 Avg. (↑)
拡張前	31.67
拡張後	41.31

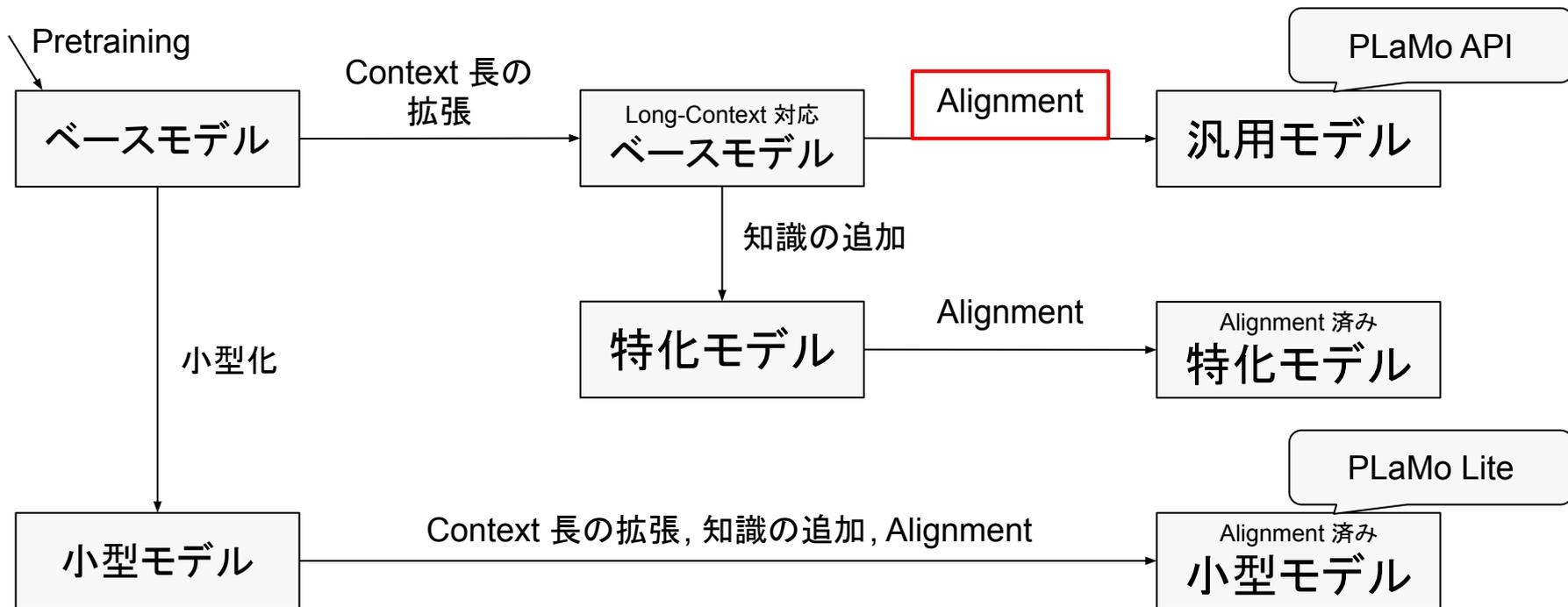
Needle In A Haystack の評価結果  
(特定の位置のテキストの検出率を示す)

Context 拡張前の結果  
input が 4k を超えると  
検出率が 0 になる

Context 拡張後の結果  
4k を超えても高い精度  
が達成できている



# PFN における Fine-Tuning



# Alignment

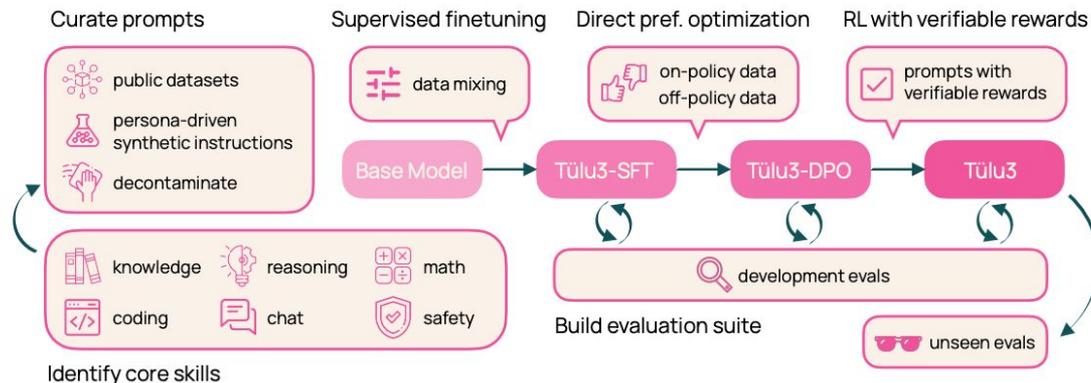
言語モデルをユーザーにとって望ましい方法で動作するようにするプロセス

- **Supervised Fine-Tuning (SFT, Instruction Tuning)**
  - モデルに望む回答例を教師として学習を行う
- **Direct Preference Optimization (DPO)**
  - 同じ質問への2つの回答例のどちらが好ましいかを教師として、モデルがより好ましい回答をするように促す
- **Reinforcement Learning (RL)**
  - 出力の良し悪しを測る報酬モデルを用いて、常に新しい出力を試しながらモデルを改善していく

# 汎用モデル開発における Alignment

SFT を行ったモデルをベースに DPO や RL を行うのが一般的  
SFT で基本的な能力を身に着けて、DPO, RL で磨き上げるイメージ

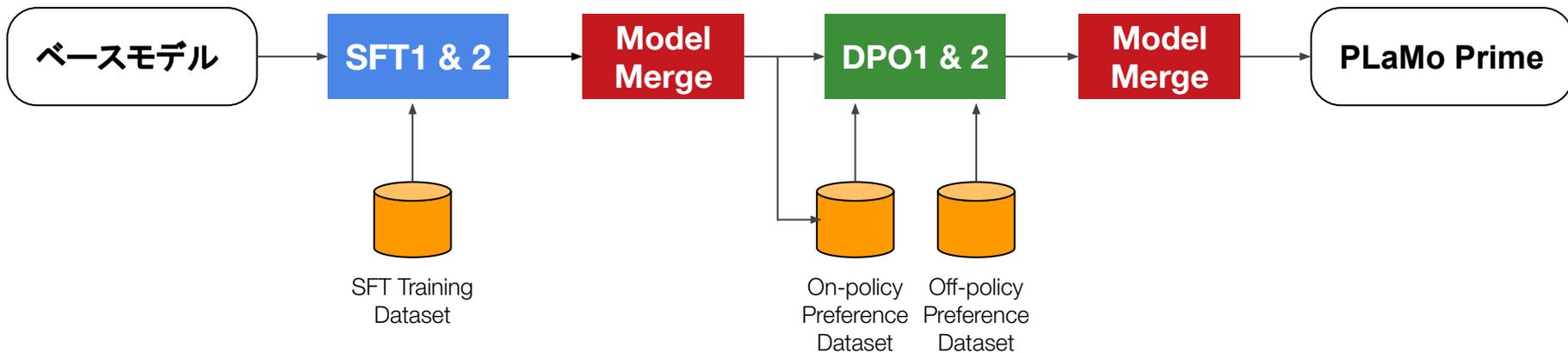
多様で高品質なデータセットを構築することが重要



[Tulu 3, Lambert et al., 2024](#) より引用

# PLaMo Prime の Alignment

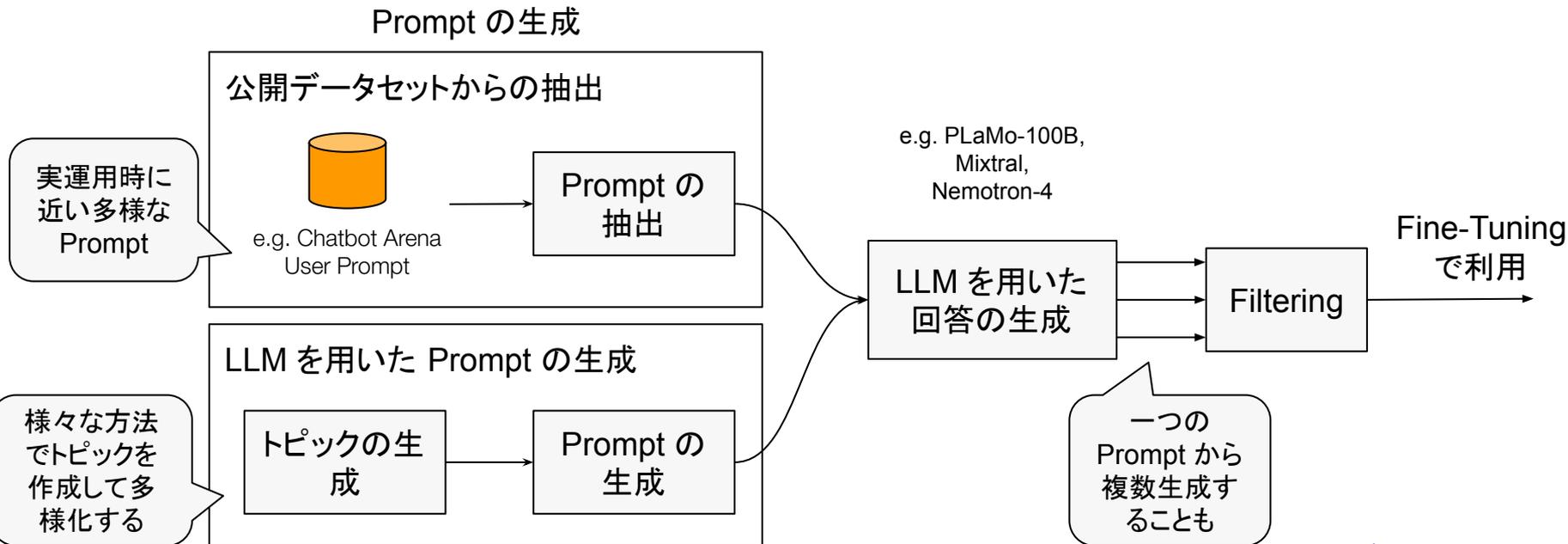
- SFT は数学特化と多様なタスクの 2 Stage
- データ生成と学習を複数回行う Iterative DPO ([Dong et al. 2024](#))
- 各 Phase で設定の異なる学習結果を Averaging



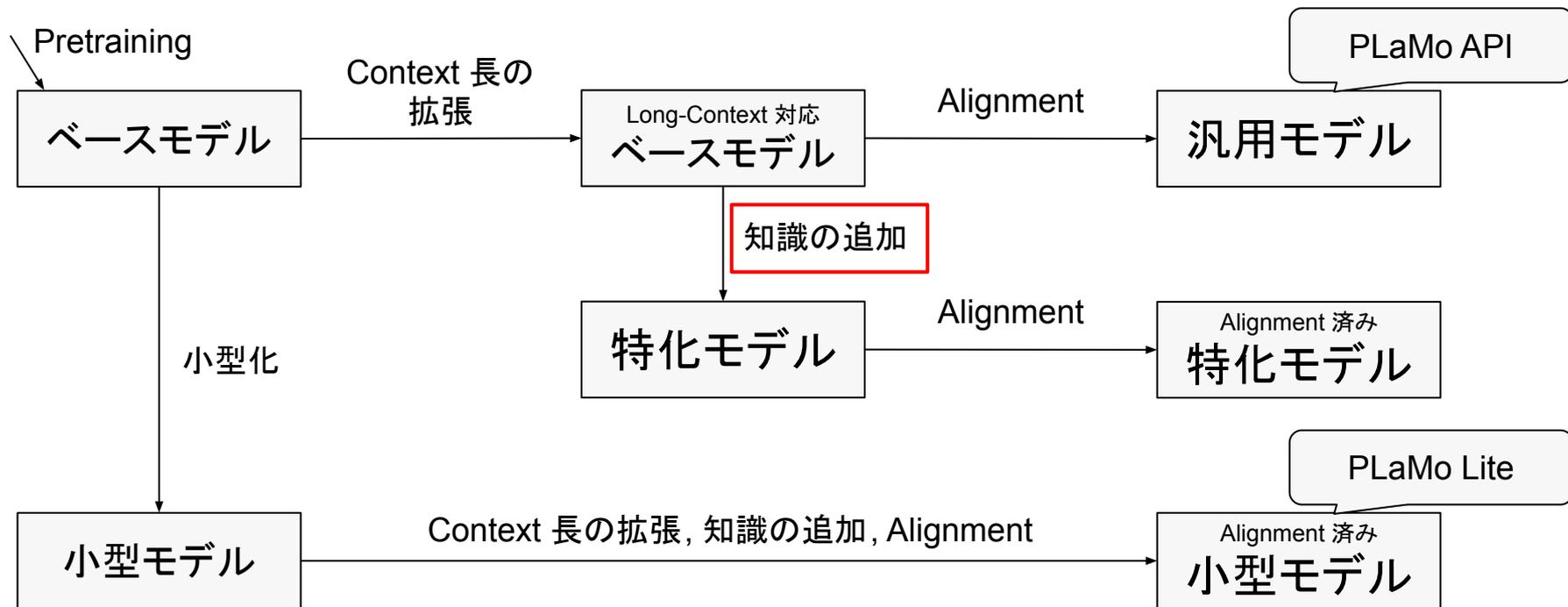
Staged SFT は Nemotron-4 ([NVIDIA, 2024](#)), Phase 毎の Model Averaging は Llama-3 ([Meta, 2024](#))で使われていたテクニックを取り入れた

# 多様なデータセット構築のための取り組み

高品質な公開データセットに加えて、LLM を用いた**合成データ**を用いるのが一般的。PLaMo でも多くの合成データを利用した



# PFN における Fine-Tuning



# 知識の追加

SFT は新しい知識の獲得には不向き ([Gekhman et al. 2024](#)) で、知識の追加には対象のドメインでの **Continual Pretraining** を行うのが一般的

<b>Pretraining/ Continual Pretraining</b>	文章全てに対して次の単語を予測して学習
<p>PFNのプロダクト Preferred Networks(PFN)は、深層学習などのソフトウェア技術と計算基盤などのハードウェア技術を独自の方法で組み合わせ、産業向けソリューションから消費者向けサービスまで幅広く業界をまたがるプロダクトを開発・提供しています。</p> <p>研究開発中のプロジェクト、プロセッサやスーパーコンピュータなどの計算基盤、PFNが開発に関わる各種オープンソースソフトウェアについてはProjectsをご覧ください。</p>	

全体でロス进行計算

<b>SFT</b>	ユーザーの質問とLLMによる応答部分で構成し、 <b>応答部分</b> を予測
<p>## 指示: PFNのプロダクトについて教えてください。</p> <p>## 応答: Preferred Networks(PFN)は、深層学習などのソフトウェア技術と計算基盤などのハードウェア技術を独自の方法で組み合わせ、産業向けソリューションから消費者向けサービスまで幅広く業界をまたがるプロダクトを開発・提供しています。</p>	

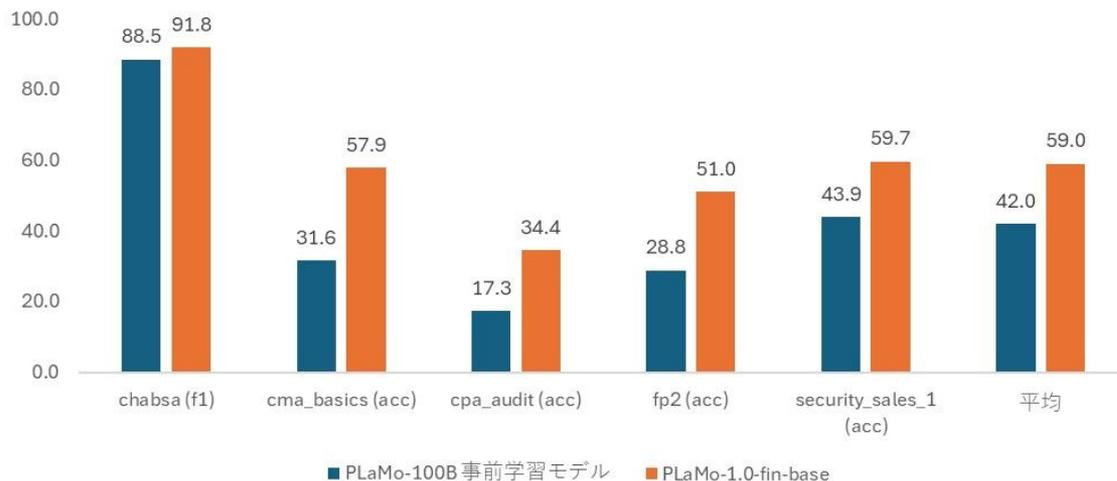
PLaMo-13Bで使われたAlpaca Formatの例

応答部分のみでロス进行計算

# 金融特化モデル (PLaMo-fin-base)

金融ドメインで Full Parameter の Continual Pretraining を行った事例

金融分野における日本語ベンチマーク Japanese Language Model  
Financial Evaluation Harness で高い評価を達成



# 知識の追加による破滅的忘却の抑制

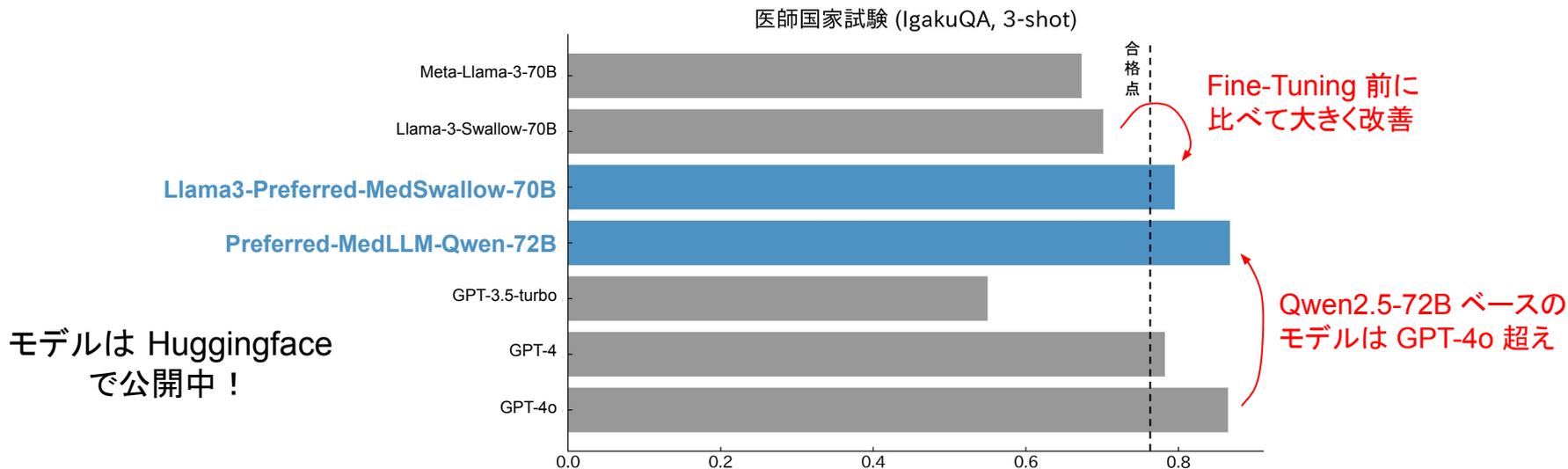
- Continual Pretraining でも破滅的忘却 (Catastrophic forgetting) は起こる
  - 新しい知識の学習によって元々持っていた知識を忘れてしまう現象
- 破滅的忘却の抑制は盛んに研究されている
  - Pretraining 時のデータにアクセスできる場合は、その一部を Continual Pretraining でも使うことにより忘却を抑制できる
  - 低ランクの行列のみを学習し、少ないリソースで Fine-Tuning が行える LoRA ([Hu et al., 2021](#)) は、学習されるパラメータが少ないことが正則化として働き、忘却の抑制にも寄与する ([Biderman et al. 2024](#))

# 医療特化モデル

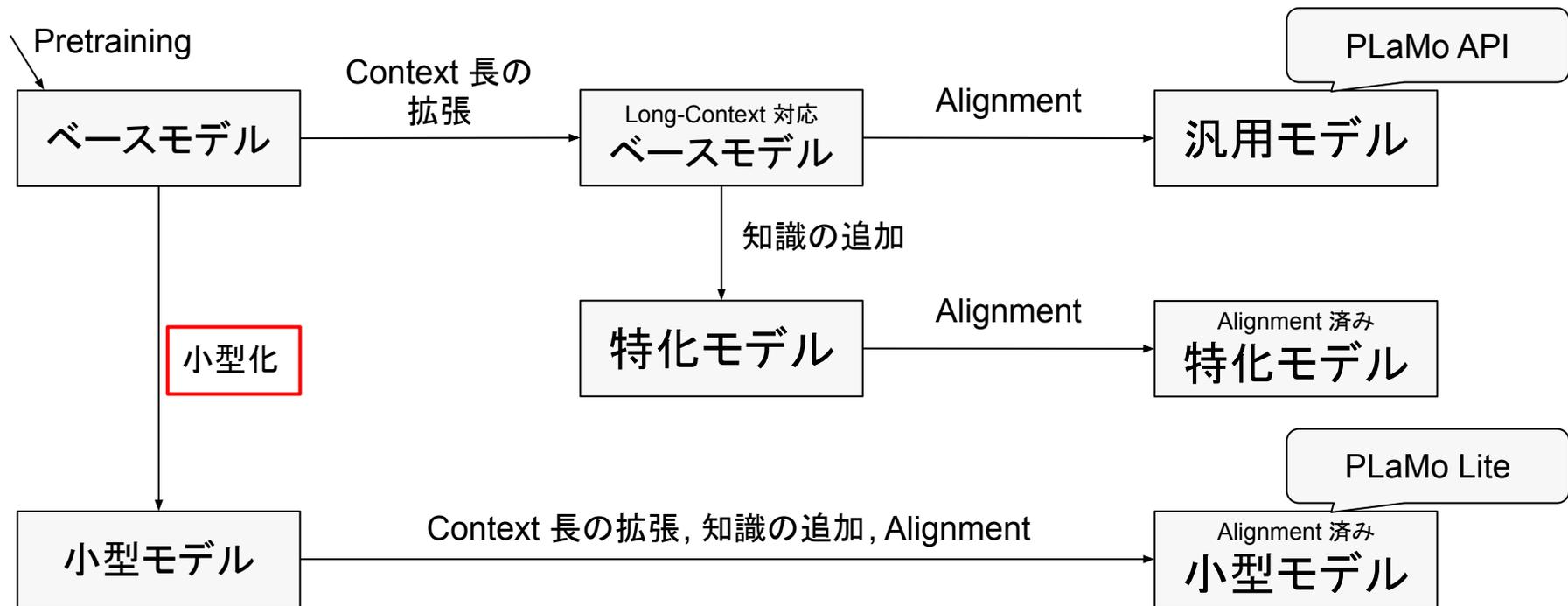
QLoRA ([Dettmers et al. 2023](#)) で医療知識を追加したモデルを開発した事例

量子化 + LoRA。更に少ないリソースで実行可能

日本医師国家試験において GPT-4 & GPT-4o を超える成績を記録



# PFN における Fine-Tuning



# 小型化

- 小さなモデルにはメリットが沢山
  - スタンドアロンにデプロイしやすい
  - スマートフォンなどエッジデバイスでも動作できる
- 小さなモデルは一般には大きなモデルに性能面で劣るが、ドメインを絞って Fine-Tuning を行うことで、小さくても戦える可能性がある
  - **Fine-Tuning は小型モデルと相性が良い** ！？

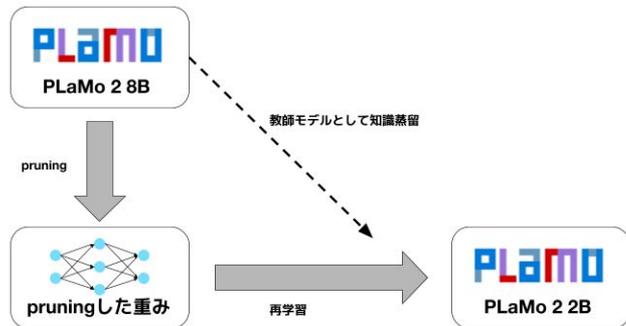
大きなモデルを **Pruning** (パラメータ削減) して小型化するのがトレンド

- 直接小さなモデルを作るよりずっと良いものができる
- Meta の Llama-3.2-1B や NVIDIA の Minitron ([Muralidharan et al. 2024](#)) など

# PLaMo-2-2B

## PLaMo-2-8B を Pruning した 2B モデルが高い性能を達成した

Minitron (Muralidharan et al. 2024) を  
ベースにした Pruning + 知識蒸留



Pruning で作成した PLaMo-2-2B は少ない学習計算量で

- 直接学習した 1B モデル
- 検証用に 100B tokens だけ学習した 8B モデル  
よりも高い性能を達成した

モデル	学習token数	学習計算量 [10 <sup>18</sup> FLOPs]	JMLLU default (5 shots, acc)	JHumanEval (0-shots, pass@1)
PLaMo-2-2B	100B	2800 *	0.505	0.311
PLaMo-2-1B	4T	24000	0.334	0.189
8Bモデル (検証用)	100B	4800	0.365	0.232

\* Pruning元であるPLaMo 2 8Bの学習計算量を除き、2B自体の学習に必要な計算量と蒸留する際に利用する8Bの推論の計算量を含めた値

# まとめ

- 目的に応じた様々な Fine-Tuning を PFN での事例とともに紹介
  - **Context 長の拡張**: ABF RoPE + Long-Context Continual Pretraining
  - **Alignment**: SFT + DPO, RL
  - **知識の追加**: 対象ドメインで Full Parameter/LoRA Continual Pretraining
  - (小型化: Pruning & 知識蒸留)
- 用途や計算機の制約にマッチした技術を選択できることが重要



**Making the real world computable**