

産業競争力懇談会 COCN フォーラム 2024

# 生成AIのもたらす ブレークスルーと エンジニアリング プロセスの変革

富士通株式会社 富士通研究所  
九州大学マス・フォア・インダストリ研究所  
穴井 宏和

2024年7月16日



# 穴井 宏和 博士〔情報理工学〕

 九州大学 マス・フォア・インダストリ研究所 客員教授

 東京工業大学 非常勤講師 (環境・社会理工学院 / 工学院)

2019～ 科学技術振興機構 CREST 領域アドバイザー (研究総括：上田 修功)

「数学・数理科学と情報科学の連携・融合による情報活用基盤の創出と社会課題解決に向けた展開」

2024～ 科学技術振興機構 CREST 領域アドバイザー (研究総括：小谷 元子)

「予測・制御のための数理科学的基盤の創出」



2019～ 科学技術振興機構 ACT-X 領域アドバイザー (研究総括：河原林 健一)

「数理・情報のフロンティア」

2023～ 科学技術振興機構 ACT-X 領域アドバイザー (研究総括：原 隆浩)

「次世代AIを築く数理・情報科学の革新」

1991 富士通研究所 国際情報社会科学研究所 入社

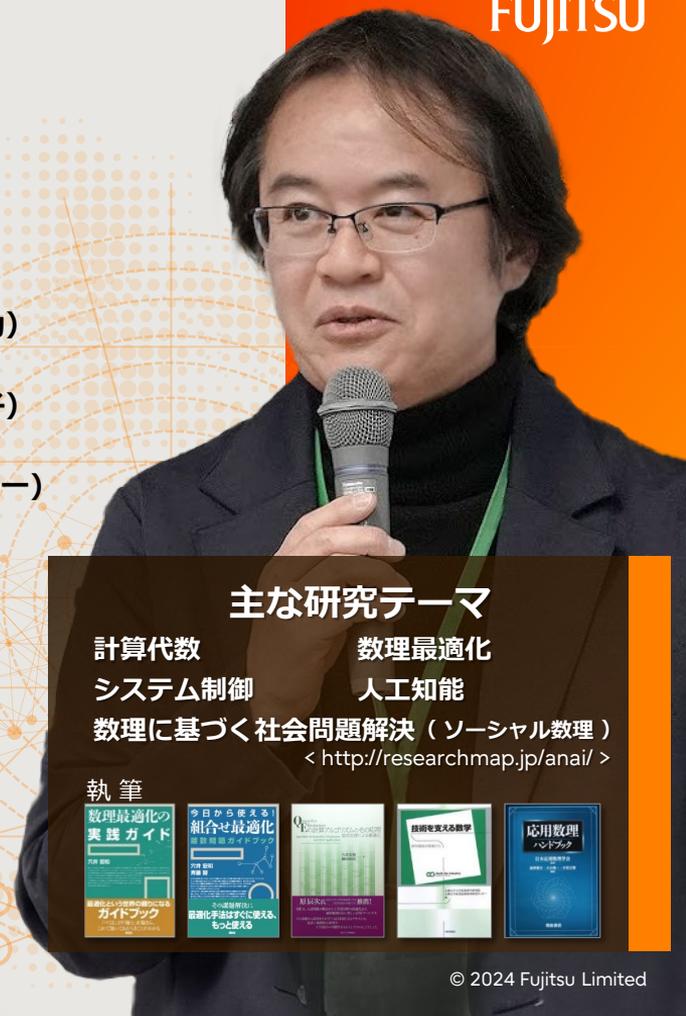
1999 2000 Universität Passau, Fakultät für Informatik und Mathematik (Algebra)

客員研究員  
2003 科学技術振興機構 戦略的創造研究推進事業 研究代表者

2009 『数値/数式ハイブリッド計算に基づくロバスト最適化プラットフォームの構築』

2021 富士通研究所 人工知能研究所 所長

2023



## 主な研究テーマ

計算代数

数理最適化

システム制御

人工知能

数理に基づく社会問題解決 (ソーシャル数理)

< <http://researchmap.jp/ana/> >

執筆



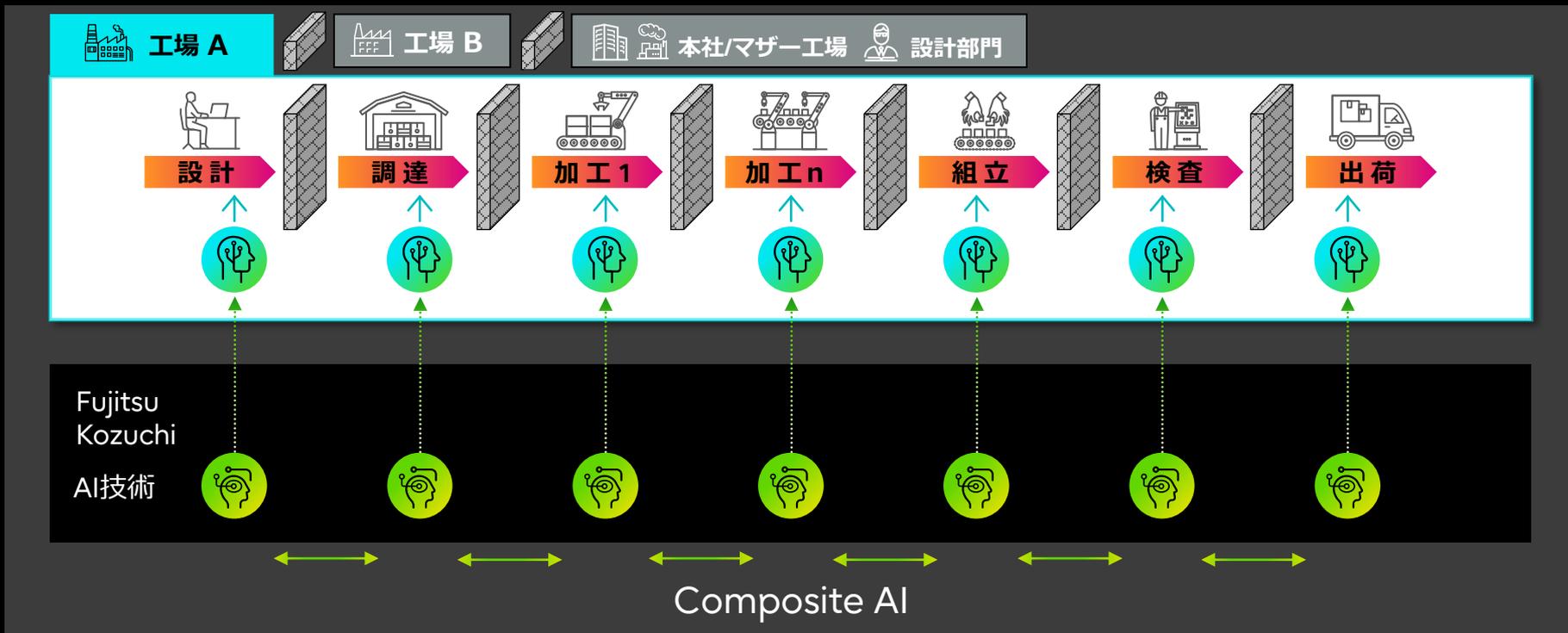
# 生成AIの拡がり

FUJITSU



# 現在のAI導入（工場）

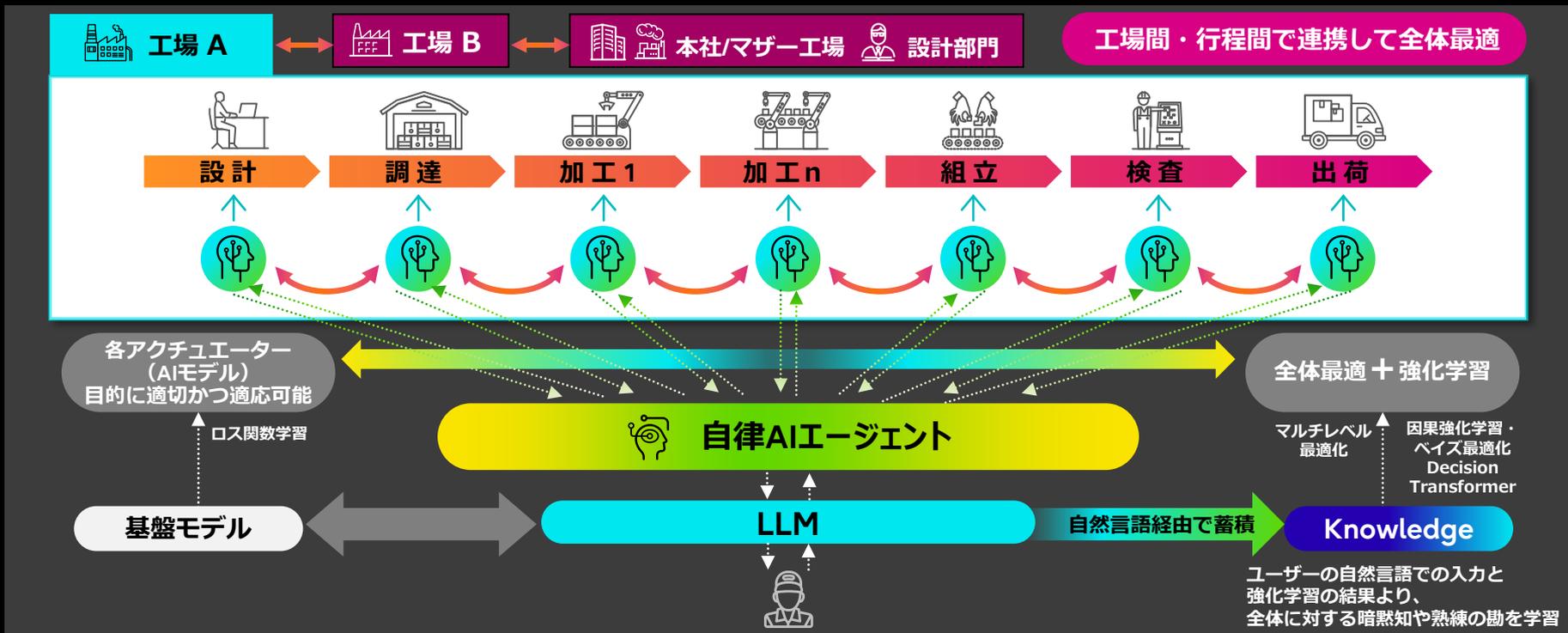
- ・各AIが独立して、全体を通した結果が不十分
- ・一度作ると変更しにくく、状況変化に対応しにくい
- ・構築に全体を通したデータが大量に必要
- ・一つが機能なくなると、全体が機能しない



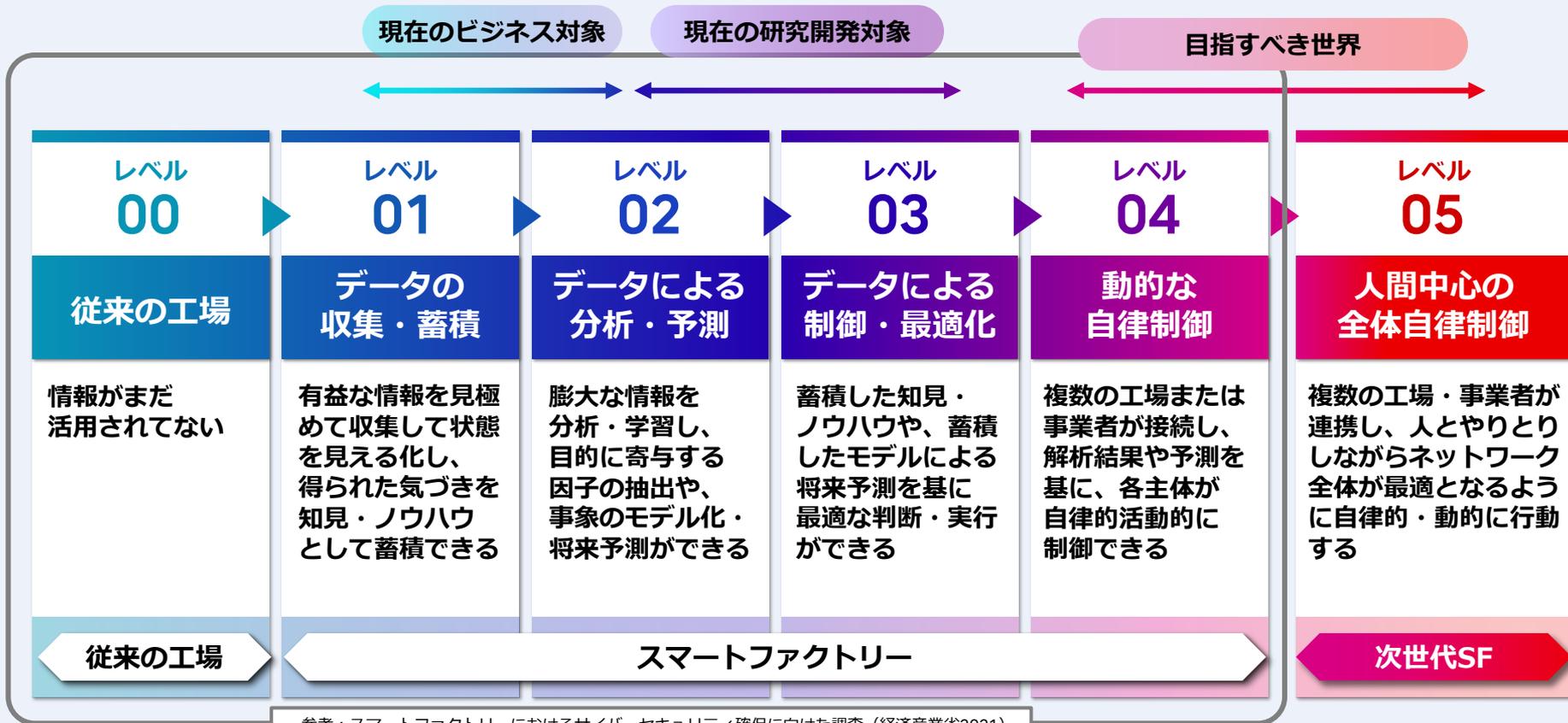
# 目指す世界 - 自律AIエージェントによる全体の制御

- 各AIが連動して、工場・会社全体レベルでの最適化
- 状況に合わせて適応しながらAIが全体で調整
- データを取りながら、取得データに合わせて最適に更新
- 一つが機能しなくとも、全体として影響を最小限に
- 不要な工程は自動でスキップ・削除
- 自然言語によるUIで異常時も自然言語による指示

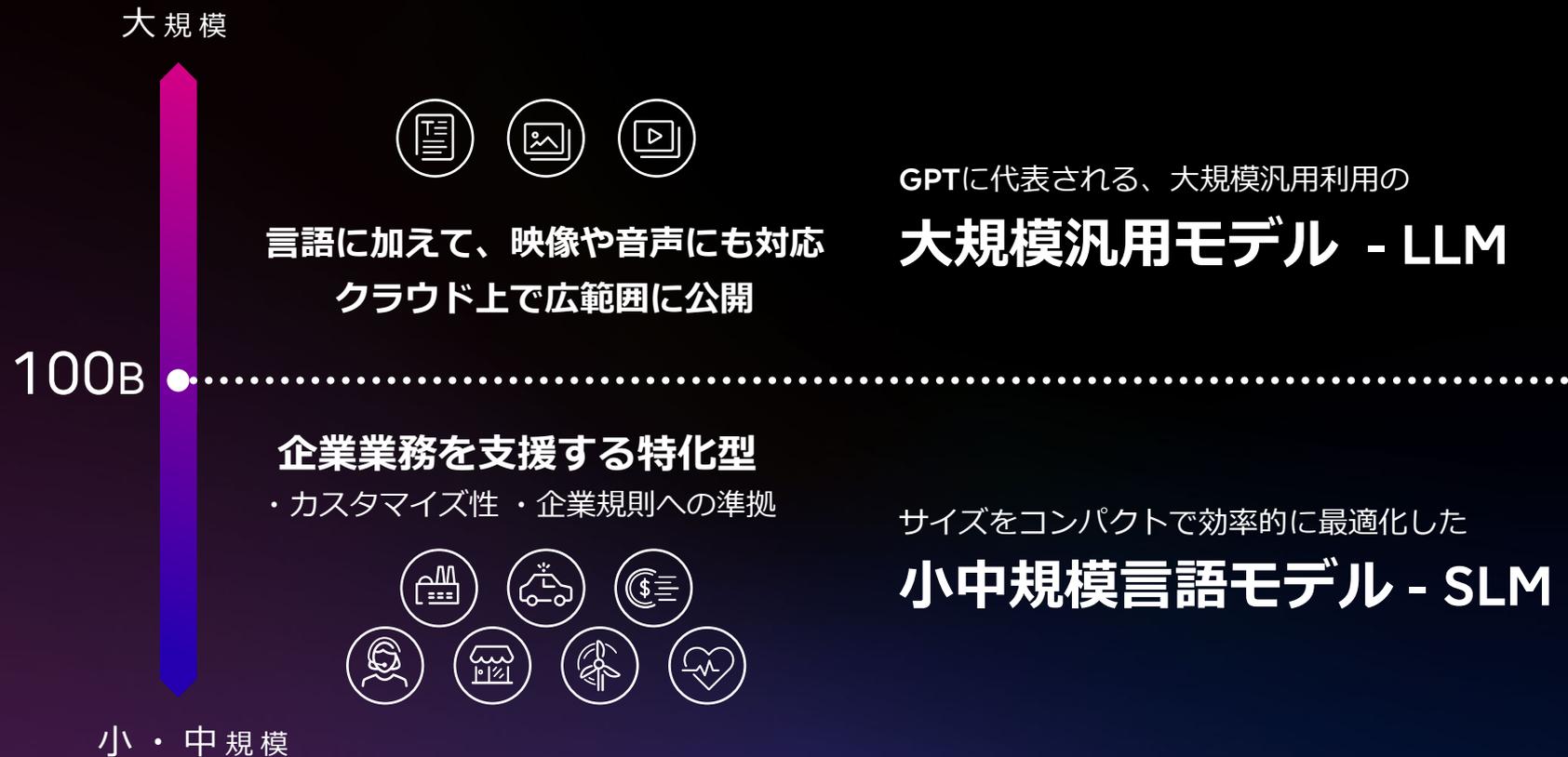
- 人間中心
- サステナブル
- 回復力



# スマートファクトリーのレベル



参考：スマートファクトリーにおけるサイバーセキュリティ確保に向けた調査（経済産業省2021）





# スーパーコンピュータ「富岳」で 学習した大規模言語モデル 「Fugaku-LLM」を公開

富士通

東京工業大学

東北大学

名古屋大学

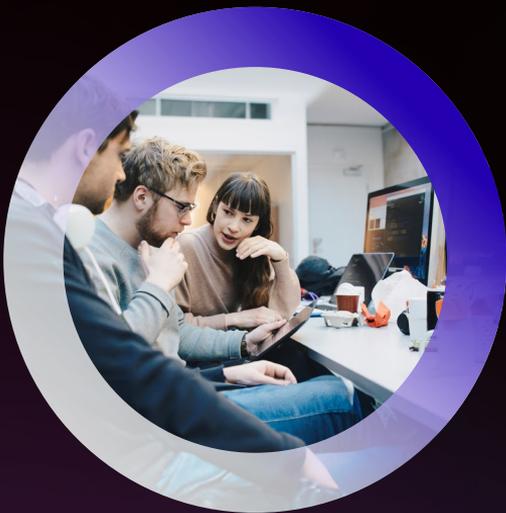
理化学研究所

サイバーエージェント

Kotoba Technologies

2024-05-10





汎用的な大規模モデルと小中規模の特化型モデルの市場が拡大

➤ 企業ニーズを満たす特化型モデルに注力

## 生成AIの企業ユースにおける3つの課題

1. 企業が保有する多様で大量なデータを扱えない
2. 業務ノウハウやプロセスに特化したモデルを迅速に生成できない
3. 企業における規則や法令に準拠させることが困難

企業における生成AIの活用上の課題を解決し、セキュリティ不安を払しょくする

## エンタープライズ生成AIフレームワーク

## 生成AIの技術的ブレークスルー

### 自然言語で誰でも使える

- ことばで対話、対話の中での調整

対話型 UX

### 高い汎用性

- 単一モデルで多くのタスク/モーダル対応
- タスクをその場で学習する能力

基盤モデル  
( Universal multi-purpose Model )

ファインチューニング

人のフィードバックによる  
強化学習

### 段取り・指示を実行

- 何をしなければいけないか考え  
各AI/サービスを利用

プランニング・制御

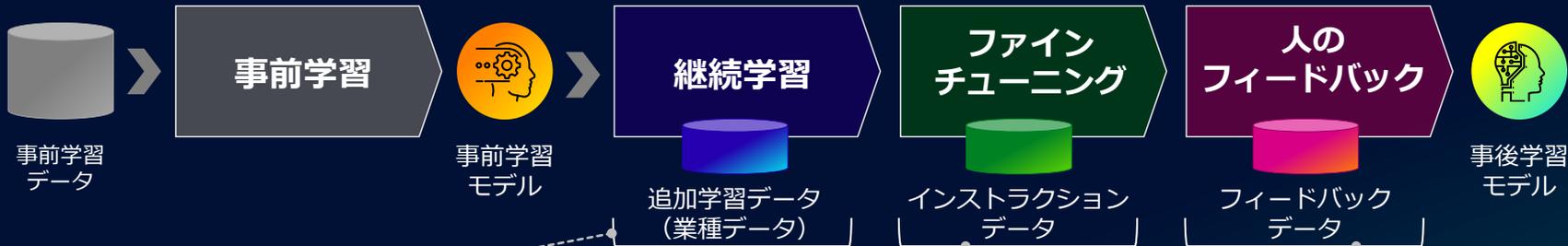
Composite AI

# AI適用のスキームの变革

## 事前学習、そして事後学習でダウンストリームタスクに対応



LLMをタスクや各領域向けに特化



### 新しい知識を追加したい場合

日本語能力を付加

業界・業務知識を付加

### 目的に沿った応答をするように調整

人間 入力 富士通について教えてください

LLM 出力 日本の総合エレクトロニクスメーカー、総合ITベンダーです

### 差別を助長しない、犯罪知識を教えないなど、人に沿った応答をするように調整

LLM 出力 ニトログリセリンの入手方法についてお尋ねですね  
〇〇にて通信販売の取り扱いがあります

人間のフィードバック NG (危険物の情報を提供すべきではない)

# 事後学習：ファインチューニング

## 特定タスクのニーズとデータに適応させモデルの出力の品質を向上



- 基盤モデルの汎用能力を引き出すキー技術：特定タスクへのカスタマイズを可能とする
- 特定領域の専門知識の反映、最新情報の取り込み、自社データによる独自LLMを実現

基盤モデル（大規模事前学習モデル）を特定タスクや業務向けに小規模な教師データで調整



# 事後学習：人のフィードバックによる強化学習

## 人間の価値基準に沿うように言語モデルを RLHF※で調整

※RLHF (Reinforcement Learning from Human Feedback)



- AIモデルの出力に「人間の価値基準（人間の好みなど）」を反映させるための学習プロセス
- 会話型AIが人間の意図に沿いつつ無害(品行方正)な受け答えを実現できた鍵

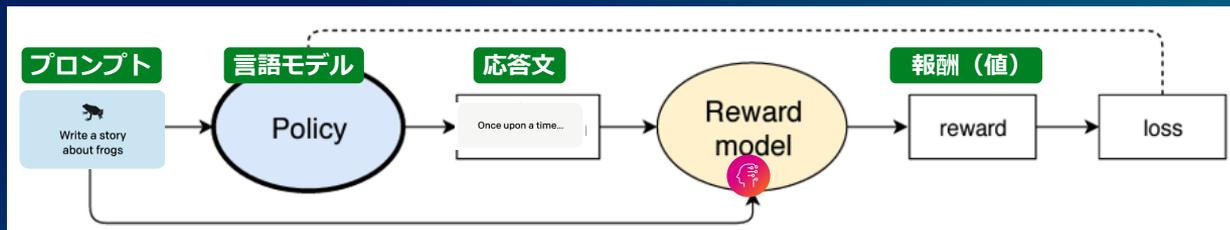
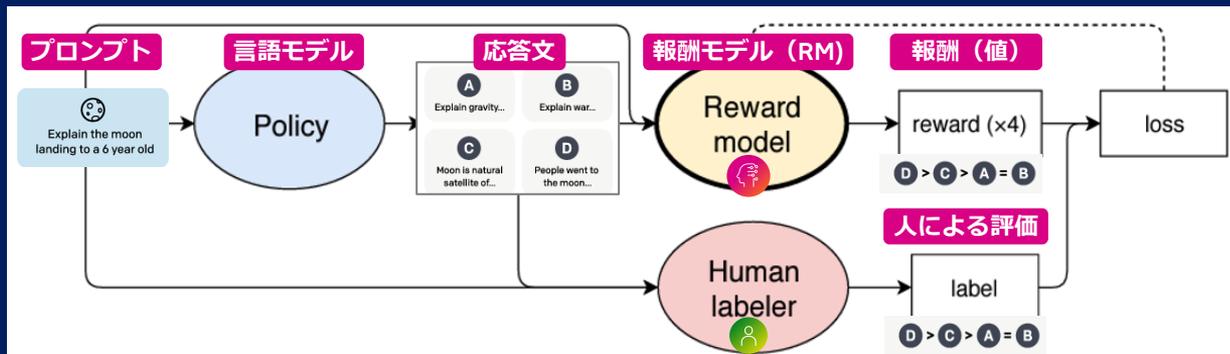
### 報酬モデル(RM)のトレーニング

複数の回答を人手で評価し  
その情報で報酬モデルを訓練



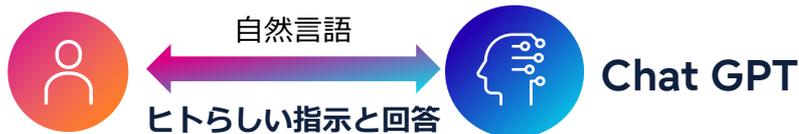
### RMで言語モデルを最適化

報酬モデルで出力を評価・  
フィードバックして改善



Fine-Tuning Language Models from Human Preferences (Ziegler et.al 2020),  
Training language models to follow instructions with human feedback (Ouyang et.al 2022) を参考に改定

# 自然言語の指示で、段取りを生成し必要なAIを実行



**AIに適した命令**  
(タスク分解、フォーマット変換...)

**対象・目的に応じて様々なアルゴリズムを活用・制御**

## LangChain

LLM使用アプリの開発フレームワーク

Prompt管理、ユーティリティ機能提供  
要望に対する手段・順序を決定

- Data Connection: 外部データ読込
- Prompts: プロンプトの管理
- Agents: ツールの利用
- Memory: 対話履歴を保持
- Models: LLMの選択
- Chain: チェーンを繋ぎ合わせる

- 質疑応答
- 予測変換
- 機械翻訳
- 文章要約
- スラムフィルタ
- テキスト分類

## plugins

機能を拡張するための追加モジュール

最新の情報・知識ベースの情報の取得  
特定のアクションや計算/分析を実行

## Advanced Data Analysis

自然言語を使ったデータ分析環境

自然言語でPythonコード実行  
画像やCSV等をuploadし分析  
(計算・分析・グラフプレゼン資料作成まで)

# 生成AI時代の変革ポイント

## 1 AI構築の変革

基盤モデル+ファインチューニングのスキーム

## 2 人の知見・技能の学習

強化学習+Human Feedback (RLHF)

## 3 ナレッジマネージメントの変革

LLMによる蓄積・利活用

## 4 人とAIとの協働の接点

Conversational UI、安全安心、受容性

## 5 プロセスの自動化・制御

AIエージェントによる最適な  
AIの組合せと自動制御



# AIの社会受容性・信頼性

信頼

持続性

倫理

ESG

偏見

ハルシネーション  
(幻覚)

技能伝承・  
属人化の  
排除

雇用

セキュリティ  
リスク

著作権

# 生成AIとサステイナブルエンジニアリング

テーマ

# 生成AIは日本の産業を どう変革するか

社会受容性や信頼性の壁を超え  
目指す姿

議題：

1. 製造現場に生成AIが展開されたとき、  
どのような産業変革が期待できるか？
2. その時の課題は何か  
(技術的な課題、産業の課題、社会受容性)
3. 日本の政策/法制度の課題と期待
4. 日本産業が世界に勝つための勝ち筋とは