SCSKのAI技術戦略
最新手法の活用とこれからの展望

SCSK株式会社
技術戦略本部
先進技術部　技術開発課

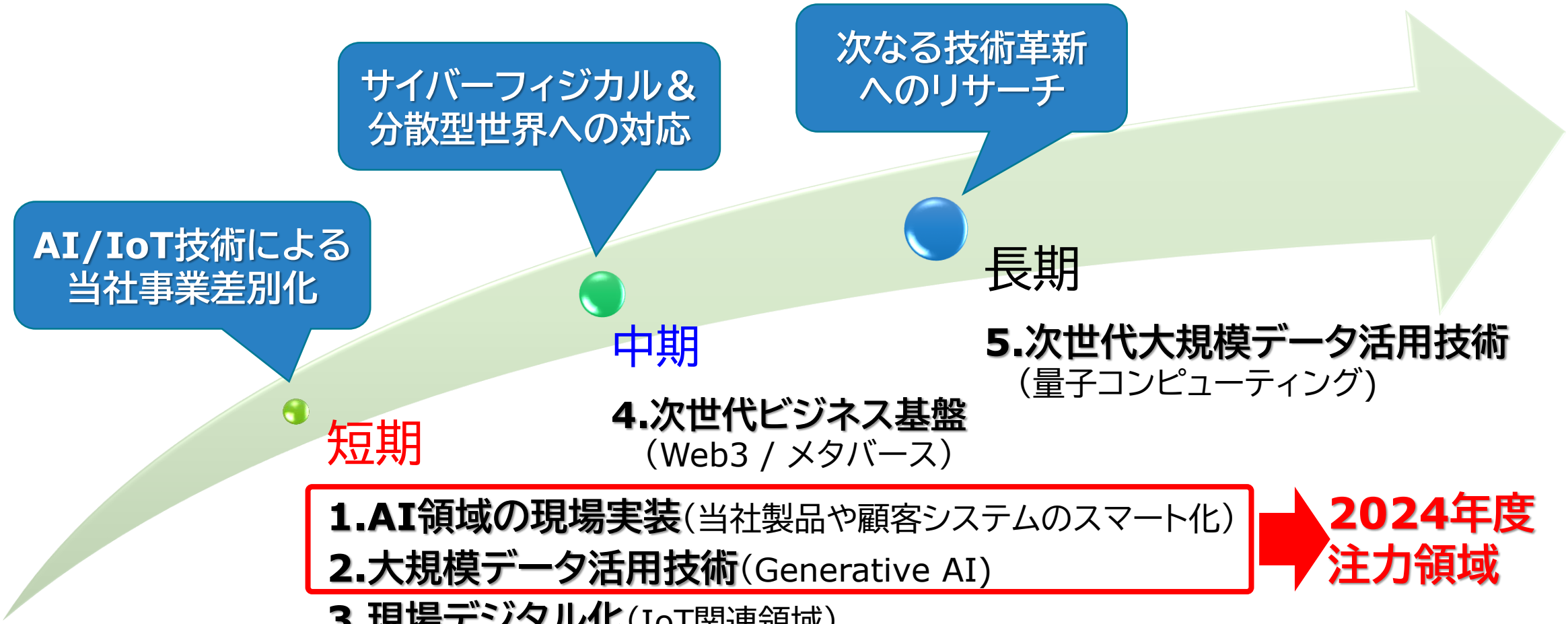
2024年6月27日

目次

2. SCSKでの先進技術の取組み方針

先進技術の取組み方針

既存技術と先進技術を融合し、お客様の価値向上に貢献する



3. 2023年度のSCSKの生成AIの取り組み

2023年度プレスリリース（SCSK-GAI）

News Release

2023年5月22日
SCSK株式会社

生成系 AI「SCSK Generative AI」を全役職員が業務利用開始
～生成系 AI の安全・安心な利活用に向けて～

SCSK株式会社（本社：東京都江東区、代表取締役　執行役員　社長：當麻　隆昭、以下　SCSK）は、自社専用環境に生成系 AI「SCSK Generative AI」（以下　SCSK-GAI）を構築し、全役職員が業務での利用を開始しました。役職員が生成系 AI を安全・安心に利活用できる環境およびガイドラインを整備し、生成系 AI を積極的に活用することで、業務の効率化・生産性向上や製品・サービスへの適用、新規事業創出を目指します。

1. SCSK Generative AI について

SCSKは「技術ドリブン推進」を中期経営計画の経営基盤強化に掲げ、推進組織として技術戦略本部を設置し、生成系 AI をはじめとする先進技術の獲得・利活用による新たな価値創出・事業開拓、社会実装に向けた高度先進技術者の拡充に取り組んでおります。SCSKは、これまでにも自然言語処理 AI に関する研究開発・特許取得を行っており、知的生産性を飛躍的に向上させることが期待される生成系 AI の全役職員の積極的な活用にむけて、セキュアなクラウド環境に SCSK-GAI を構築しました。また、生成系 AI を安全・安心に利活用していくために、「生成系 AI 利用ガイドライン」を整備するなど、全役職員が適切な管理のもとで生成系 AI を利活用するメリットを享受できる取り組み、生成系 AI を用いた新たなアイデア創出の機会を提供していきます。
SCSK-GAI は、マイクロソフトが提供する Azure OpenAI Service を活用して、チャット形式で利用します。利用者が入力した情報の二次利用や第三者提供がされない仕様にしており、かつ社内ネットワークからのみ利用可能としたことで、セキュリティを担保しています。

News Release

2023年7月6日
SCSK株式会社

生成 AI「SCSK-GAI」を活用した質疑応答支援システムの概念検証を開始

SCSK株式会社（本社：東京都江東区、代表取締役　執行役員　社長：當麻　隆昭、以下　SCSK）は、自社専用環境に生成 AI「SCSK Generative AI」（以下　SCSK-GAI）を活用した質疑応答支援システム（以下　本システム）を開発し、概念検証を開始しました。

1. 背景と概念検証の目的

生成 AI の業務への活用により、業務効率、生産性の向上が期待されていますが、公開済みの一般的な情報をもとに学習している汎用的な生成 AI では、企業が公開していない情報や各社に特化した情報を扱う事が難しく、適切な回答結果が得られないという課題があります。特に、社外からの問い合わせ対応業務において生成 AI を活用するためには、企業内に保持している多岐に渡る非公開の資料を元にした学習、さらに質問内容に応じて適切な情報を抽出したうえでの回答作成が必要となります。
今回の概念検証は、本システムの有効性の確認と実用化に向けた課題を明らかにすることを目的としています。

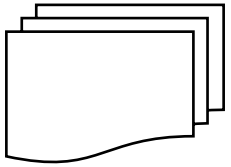
2. 本システムの概要

本システムは、公開済みの情報を学習している汎用的な生成 AI とは異なり、経営資料や社内文書などの非公開情報を事前に解析処理し、専用のデータストアの作成を行います。そのうえで、入力された質問文を解析しそのデータストアと照合し、回答作成に必要なテキスト要素を、SCSK-GAI に連携します。SCSK-GAI は、このテキスト要素から自然な回答文を生成し表示します。あわせて、回答の根拠となった参照ファイルも表示します。本システムは、SCSK専用のセキュアなクラウド環境に構築しており、経営資料や社内文書など機密情報を含む情報セキュリティを担保します。

4. 2024年度のSCSKの生成AIの取り組み

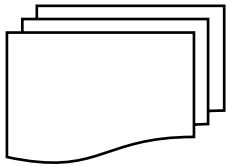
LLM学習の流れ

コーパスデータ



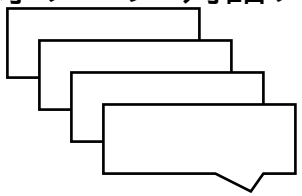
Step1
Pre-Training（事前学習）
使用データ：大規模コーパス
手法：自己教師あり学習
目的：知識や言語理解の獲得

コーパスデータ



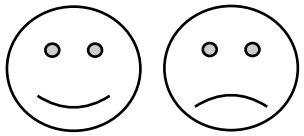
Step2
Continual Pre-Training（継続事前学習）
使用データ：学ばせたい新知識を含むコーパス
手法：事前学習済みのモデルに追加で事前学習
目的：新しい知識や言語理解の獲得

指示データ・対話データ



Step3
Supervised Fine-Tuning（教師ありファインチューニング）
使用データ：指示データや対話データのようなラベル付きデータ
手法：教師あり学習
目的：特定のタスクやドメインへの適応・言語モデルの性能の改善

フィードバックデータ



Step 4
Reinforcement Learning from Human Feedback（RLHF）
使用データ：人間のフィードバックデータ
手法：強化学習
目的：出力をより人間らしく見えるようにする

（継続）事前学習・SFTのメモリ使用量

- LLMモデルのフルパラメータチューニングに必要なメモリ量（Llama2 7B）
- 前提条件
  精度：float32
  オプティマイザー:adam
  モデルサイズ：7＊4 ＝ 28GB
  Forward passメモリ：Forward memory（バッチサイズ・シーケンス長などにより変わる）
- 各ステップの使用メモリ
  1. Load model
     - 28GB（モデルサイズ）

  2. Forward pass
     - 28GB（モデルサイズ）+Forward memory

  3. Gradient calculation
     - 28GB（モデルサイズ）+28GB（勾配）

  4. Optimizer Step
     - 56GB（モデルサイズ＋勾配）+56GB（分散＋モーメンタム）

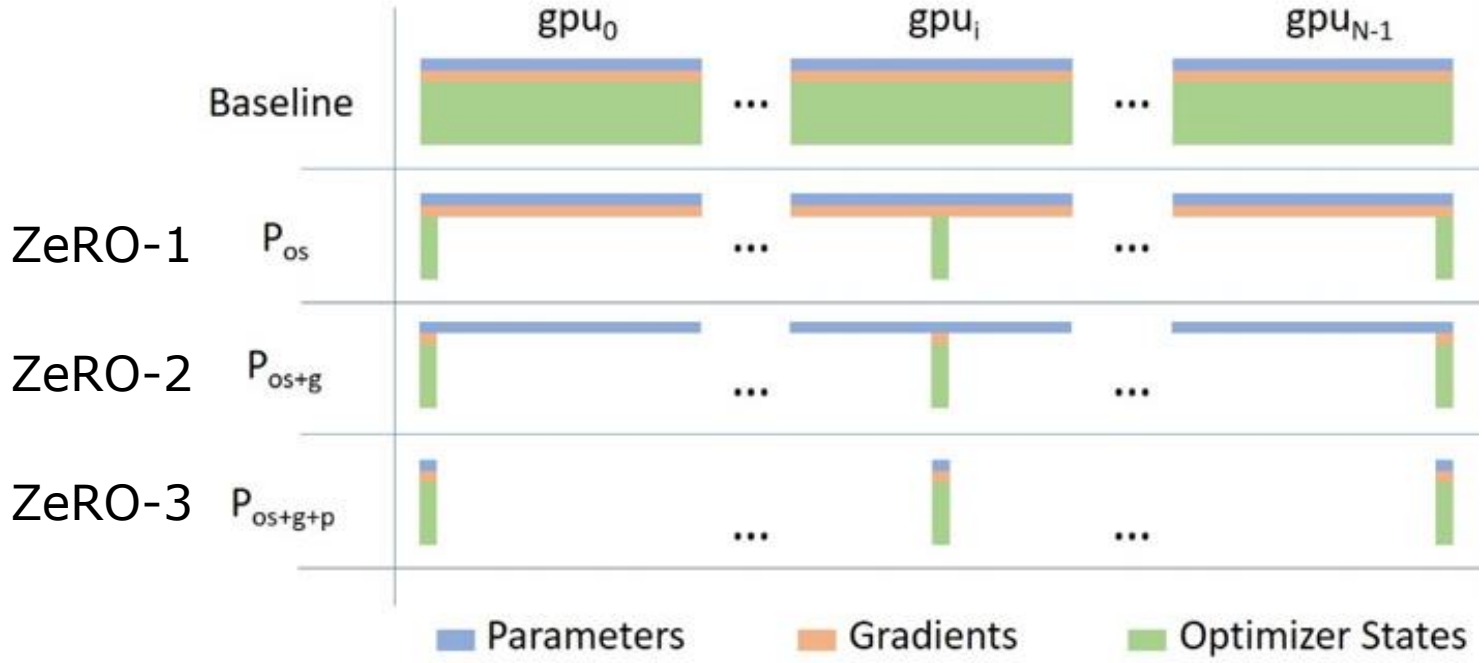
  5. 2に戻る
     - 112GB（モデルサイズ＋勾配+分散＋モーメンタム）
     +Forward memory

HPE社製GPUサーバ構成

製品名	HPE CRAY XD6500（XD670）
GPU	NVIDIA H100 GPU(80GB) x 8GPU (HGX/SXM5)
CPU	Intel Xeon Platinum 8480+ （56core/2.00GHz）x 2CPU（Total: 112core）
Memory	2048GB（64GB 2Rx4 DDR5-4800 x 32）
Storage（OS）	M.2 NVMe SSD 1.92TB x 2
Storage（DATA）	U.2 NVMe RI SSD 3.84TB x 8
Network	InfiniBand HDR/Ethernet 200Gb 2-port QSFP56 Adapter x 4 Ethernet 100Gb 1-port QSFP28 x 1
管理ネットワーク	10GbE(RJ45)×2port Management LAN(1GbE)×1port
OS	Ubuntu
Support	HPE TechCare Essential 5年（24x7）
電源	3000Wパワーサプライ x 6（IEC C20-C19 x 6）
サイズ	ラックマウント 5U

（継続）事前学習・SFTの分散学習

分散学習使用時のLLMモデルのフルパラメータチューニングに必要なメモリ量（Llama2 7B）



ZeRO と DeepSpeed:Microsoft Researchより

GPU数：8

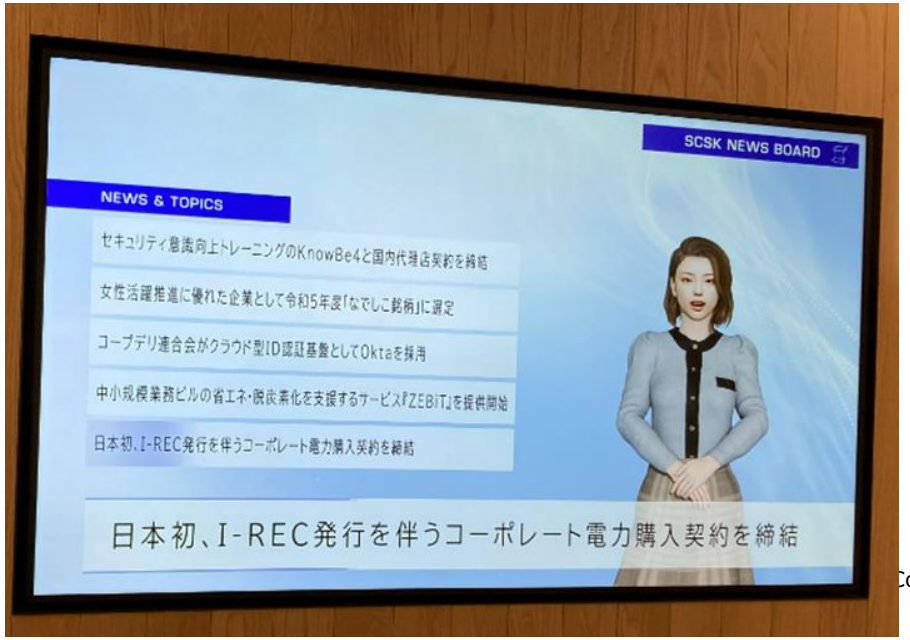
- ZeRO-1
  メモリ：28GB（メモリサイズ）+56GB（分散・モーメンタム）/8（GPU数）+28GB（勾配） +Forward_memory
  ＝63GB+Forward_memory
- ZeRO-2
  メモリ：28GB（メモリサイズ）+{56GB（分散・モーメンタム）+28GB（勾配）}/8（GPU数） +Forward_memory
  ＝38.5GB+Forward_memory

5. Next Step

GPUサーバーの活用

1. **LLMの研究開発（AI）**
   - 大規模言語モデル（LLM）のファインチューニングをはじめとしたLLMの研究開発
     - 個別LLM(Large Language Model)
     - SLM(Small Language Model)
     - AIエージェント

2. **ブランディング拠点向け展示（XR技術）**
   - ブランディング拠点を訪問されたお客様を、AIとXR技術を融合したアバターでお迎えする仕組みを開発
   - 展示コンテンツ（AIニュース、デジタルヒューマン）のモデルのリアルタイム更新



PCCCワークショップ2024　SCSK展示コーナーのご案内

ご紹介ソリューション

- ✓ SCSK　AI　ソリューション
- ✓ SCSK水冷対応データーセンター
- ✓ Insight Edge社*
  生成AI時代の新しいVoC分析ツール「Voiceek」

*住友商事株式会社100子会社のDX技術専門会社

SCSK

夢ある未来を、共に創る。