

## 「富士通 研究戦略説明会」質疑応答議事録

---

1. 日 時：2024年6月4日（火曜日）10時～12時30分（説明会およびデモ見学会）
  2. 場 所：ステーションコンファレンス川崎
  3. 登壇者：富士通研究所長 所長 岡本 EVP  
人工知能研究所 園田所長  
先端技術開発本部 新庄本部長
- 

### 【質疑応答】

#### ■質問者 A

Q1. 生成 AI は NTT 「tsuzumi」、NEC 「cotomi」 など展開している中、富士通の AI の独自性、競争力優位性があれば教えてください。

A1. （園田）今日ご紹介したフレームワークにあります RAG の部分は、大規模なデータを扱えるという点では世界 No.1 の精度を保っています。企業で生成 AI を使うときの課題の一つだと思っており、ここが一つ差異化になると思っています。

生成 AI 混合技術も、企業に特化するところの情報を使ってベストな生成 AI を選びルーティングするということで、企業の様々なユースケースに対応できるような生成 AI 混合技術になっていますので、企業向けにきめ細やかに必要な生成 AI をカスタマイズなしで提供できるところに特長があると考えています。

Q2. 特化型 AI を事業展開するにあたり、クライアントとして想定している業種業態などユースケースがあれば教えてください。また目標の事業規模があれば教えてください。

A2. （園田）具体的なユースケースで考えているのは、ソフトウェアエンジニアリング領域、ドキュメント関係の分析です。今回展示しているデモは契約書のチェックをしています。企業で使うドキュメント全般の特化型。またネットワークのログ解析などもやりたいと考えています。もう一つ特徴的なのは映像系。当社は長いことビジョン系の AI をやってきていますので、ビジョン系のデータも扱えるようにしていこうとしています。例えば監視系の映像データを長期間監視して、その中で生成 AI がいろいろな解釈をするようなソリューションを提供していこうとしています。それも今回展示しています。今回展示しているデモは、力を入れているものということになります。

（岡本）生成 AI に関しては、商談の数で言うと 300 から 400 の間くらいなので、AI の中でも事業規模を大きくしていきたいと思っています。

#### ■質問者 B

Q1. 世界初の真偽判定統合分析システムとありますが、これは完成しているのでしょうか。まだ完成していない場合、いつ頃までに完成する予定でしょうか。

A1. （岡本）これは現在研究開発中です。このシステムは当社単独でやるというよりも、かなり多くのステークホルダーを巻き込みながら構築していきたいと考えています。本日は発表のタイミングもあり間に合わなかったのですが、後ほど詳しく発表するタイミングが来ると思います。

## ■質問者 C

Q1. AI Computing Brokerにて消費電力を劇的に削減する仕組みとして、GPUの稼働率を上げることは理解しましたが、GPUが遊んでいるときに消費している電力を、効率的にジョブを振り分けることで遊んでいる時間を削減し、その分の電気が浮くという理解でよろしいでしょうか。

A1. (岡本) はい、その理解で合っています。ジョブに関して空いているところにいかに割り込ませるか、というかなり難しい技術なのですが、AI Computing (Broker) の技術を使って振り分けを行っています。これに関してデモ展示も行っているのです、是非ご覧いただければと思います。

Q2. AI Computing Brokerの年間10TWh以上の削減の過程はどうなるのでしょうか？例えば、日本国内の現時点のGPUに対して技術を適用した場合など、計算の過程を教えてください。

A2. (岡本) 情報ソースが複数から来ているところがあり、計算の過程に関しては別途ご提供させていただければと思います。

※広報 IR 室より以下を回答済

2023年度におけるNvidia製GPUのうち、生成AI向けの出荷台数は55万台。最新のGPU1器あたりの消費電力が約1000Wとすると、これは約5TWhの消費電力になる。

出荷台数について、CAGR(年平均成長率)を年間50%(TrendForce)と仮定すると、4年後には200万台を超えることから、この技術が普及すれば、その消費電力約20TWhの半分を削減可能と見積もっている。

## ■質問者 D

Q1. 生成AIに関して、汎用の対話型LLMより特化型生成AIモデルの方がマーケット的には大きいのでしょうか。質問の背景としては、生成AIのマーケット構造が、規模の大きさも含めて今後どのように広がっていき、その中で富士通はなぜ特化型の方に注力していくのかも含めてお聞きしたい。

A1. (岡本) 汎用型の基盤モデルという一般のLLMに対しては、やはりインターネットユースからユーザーを広げていく戦略だと考えており、本日、園田が説明しました通り、当社はエンタープライズに向けたLLMに特化して開発していくということです。私達の見立てとしては、LLMの企業ユースはかなりあると考えていますので、企業向け、エンタープライズ向けのLLMに関しても相当な市場規模があるだろうとみています。

## ■質問者 E

Q1. 本日の説明会のテーマとして「技術の融合」を掲げていましたが、そのために研究所ではどのような新しい仕掛けをしているのでしょうか。また、グローバルでの拠点間の連携について教えてください。

A1. (岡本) 現在、Global OneTeamで、それぞれの研究テーマに関して拠点を跨って進めているのが非常に大きな特長です。その際、各拠点の特長を活かしていくために、例えば2年前に設立したインドの拠点では、AIや量子の分野の理数系に強い人材や、FUJITSU-MONAKAのソフトウェア部隊の研究者などもインドに集めて取り組んでいます。技術融合について

は、マトリックスで如何に効率を上げて研究を進めていくかが非常に重要で、AI の研究者をコンピューティングに移したり、あるいはバーチャルに組織を作ってプロジェクトを推進したり、様々な形でプロジェクトを推進しています。

Q2. インドの拠点は現在何名で、今後人数を増やす計画はありますか。

A2. (岡本) 研究所に関するところは 90 名程ですが、ネットワークの部隊を合わせると 400 名弱です。今後インドの研究所については、拡大していく方向で検討しています。

#### ■質問者 F

Q1. エンタープライズ向け生成 AI サービスの差別化要素としてはナレッジグラフ拡張 RAG と生成 AI 混合技術とのことでしたが、同様の技術を国内外でやっているライバルは他にどういったところがあって、そういう競合の技術に対して、どのあたりについてキャッチアップが難しいのか教えてください。

A1. (園田) ナレッジグラフ拡張 RAG については、例えば Microsoft とネオスが組んで、おそらくナレッジグラフのようなものを使って RAG を作っていくだろうと感じています。ナレッジグラフと LLM の相性は非常に良いと思っていて、当社はずっとナレッジグラフの研究をしていますし、グラフに関する技術開発もインドを中心にやってきており、大規模なグラフを分析する技術を持っています。これはナレッジグラフが大規模になってきたときの分析、スキーマーに対してグラフを検索することが必要になりますが、こういう技術開発を先行していますので、キャッチアップというよりもリードしていきたいと考えています。

混合技術のルーティングのようなベストなモデルを選ぶところは、中身はわかりませんが SambaNova や Sakana AI など、同じような技術は出てきています。MOE (Mixture of experts) など、いくつかのモデルを組み合わせることでよい精度が出てくることはわかっていて、当社は特に企業に特化するということで、うまくルーティングする工夫をしています。企業に特化するところでリードしていきたいと思っています。いろいろなユースケースがありますので、特に企業向けの特化型を作るときに何を選ぶべきかをうまく抽出するところで、混合の技術を尖らせていきたいと思っています。

#### ■質問者 G

Q1. エンタープライズ生成 AI フレームワークについて、一からモデルを作るわけではなく、既存のモデルを混合して特化していくものと理解したのですが、他社でも例えば「tsuzumi」や「cotomi」など、日本企業向け特化モデルを作る流れがあると思います。軸となる一つのモデルを一から作るのではなく、このような戦略を取られた背景を教えてください。

A1. (園田) 先ほども申し上げましたが、いくつかのモデルを組み合わせることで精度が上がる事が確認できています。また、特化型のコンセプトは大規模で汎用的なモデルではなく、一部に特化させてコストやレスポンスを改善できることにあると考えており、この辺りが特化型を選択した理由です。その中で様々なものをマージしたり、パイプラインを同時に動かしたりして様々な機能を実現し、お客様のニーズやユースケースに対応できると考えています。さらにもう一つの特長として、生成 AI だけではなく既存の AI、例えば生

成 AI が苦手と言われる最適化や予測についても混合する取り組みを行っています。今回デモでご紹介するインシデントの例では、最適化と予測を合成していて、様々な機能に対応するために混合という選択をとっています。これにより様々なサービスや価値を創出できると考えています。

Q2. 特化型生成 AI の自動生成において、最適な既存モデルが無い場合は必要なモデルを自動生成するというお話がありましたが、どのようにして実現しているのでしょうか。

A2. (園田) いくつかのパターンがあります。生成 AI に特化させるために、ファインチューニングやプロンプトなどを自動化させるものや、生成 AI ではないですがデータから予測モデルを自動生成する AutoML と呼ばれる技術も保有しています。また、昨年秋に発表した要件学習技術については、最適化の要件を対話型でインプットすることで最適化のモデルを自動で生成する技術です。これらの技術によって、データを使って予測したり、要件をインプットして最適化のモデルを作ったり、特化したい情報をインプットして特化型モデルを作ったりするなどして、自動で生成する仕組みになっています。

#### ■質問者 H

Q1. 生成 AI は特化型モデルを進めるということですが、特化型モデルを利用していく環境はメガクラウドのようなインフラで使うのか、企業特化で使うということからプライベートな環境に導入して利用するイメージなのか、教えてください。

A1. (園田) 当然オンプレも考えています。お客様からオンプレの要望もありますし、VPC (Virtual Private Cloud) 環境で使うことも考えています。お客様がクラウドでよいということであればそこに配置するし、基本的にはどこでも使えるようにしていきたいと考えています。

Q2. FUJITSU-MONAKA について、ポスト富岳の話もありましたが、研究開発に莫大な投資が必要になると思います。回収という観点では海外市場も視野に入れられていると思いますが、海外を含めた HPC、AI の、FUJITSU-MONAKA の研究開発で国内とは違う点を意識していることがあれば教えてください。

A2. (新庄) 当然国内だけではなく海外も含めて市場を考えています。一つはオープンソースコミュニティ、これは当然海外が中心と言ったほうが良いと思いますが、そういうところに入っていったら、どういうところに需要があり、ユースケースが求められているかを探っていく、そういうところで性能を発揮できるようにしていきたいと考えています。

#### ■質問者 I

Q1. ルールメイキングの話がありましたが、欧州で AI 規制法が成立してアメリカや日本でも法制化の動きが進んでいると思います。国によって関心や重点を置いているポイントが異なると思いますが、そういうところへの対応の難しさについて教えてください。

A1. (岡本) 各国の規制に対する違いをどのように吸収していくかについては、今はそれに対してどう備えるかの準備期間だと考えています。当社はグローバルにビジネスを行っていますし、研究開発の部隊もいます。研究は Global One Team でやっていますが、それぞれ

のリージョンの規制への対応は、各国でそれぞれに対してどういうことが行われているのかを理解していく必要があります。「Fujitsu Kozuchi」の上でもどのようなルールがどのような影響を及ぼしているのかを判定する AI 倫理影響評価ツールキットを公開していますが、そういうツールも活用していきたいと考えています。

Q2. 偽情報の真偽判定統合分析システムの説明がありましたが、生成 AI は技術の進展が早く次々にリスクが出てくることへの対応の難しさについても教えてください。

A2. (岡本) 真偽判定統合分析システムはいろいろなステークホルダーが関係しています。私たちが技術だけで対応できる場所とそうでないところがあると思いますので、人や認定機関などいろいろな方々に加わって頂き、エコシステムを作っていくことが非常に重要になると考えています。詳細に関しましては後日発表する機会が出てくると思います。

#### ■質問者 J

Q1. ナレッジグラフ技術や混合技術などにおいて世界でリードしたいという話でしたが、これらの技術は日本語に特化しているのでしょうか？どんな言語でも通じるのでしょうか？今後の展開先が日本国内かグローバルなのか教えて下さい。

A1. (岡本) グローバルに戦っていくつもりです。グローバルに研究チームがいるので、日本語でもやりますが、英語などその他必要な言語に対応していく予定です。

以 上