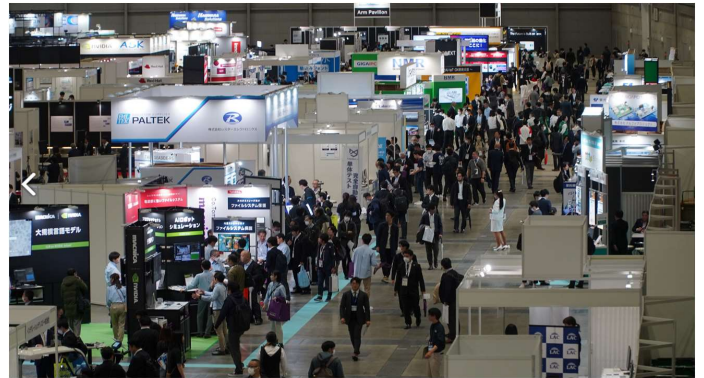


## エッジテクノロジーに新たなプラスで加速するDXと事業変革

【事業変革期を迎えた今、エッジテクノロジーに新たなプラスで顧客起点の価値創出を実現するイベントへ】と銘打ち、2023.11.15～17日までパシフィコ横浜で開催された。IoT、AI、エッジコンピューティングなどが主要なテーマで、今後はエッジの技術がキーになるという主旨のセミナーであった。3日間で33,000人超の来場者があった。



セッション1 10:00～10:45

エヌビディア

エンタープライズ事業本部 事業本部長 井崎 武士 氏

## 生成AIの開発・活用を加速するNVIDIAの取り組み

【講演概要】 スクリプトに基づく高い品質の生成画像、あたかも人と対話をしているようなChatBotなど多くの生成AIが昨年登場した。複雑なシステムもプログラムも必要とせず、話し言葉で利用できるシステムの簡便さから、従来のAIと比較し広く大衆に浸透し、いわばAIの民主化が急速に進んでいる。その一方、既存サービスのみならず、自社データや業界データを用いた独自生成AIモデル開発も活発になってきている。本講演では、生成AIの活用事例や生成AIの開発および活用を支援するNVIDIAのソリューションについて紹介があった。

## 【講演内容】

GPUで個々の処理は早くなったが、それだけ早くなってもシステム全体の効率化にはならない。DPU（データプロセッサユニット）、ネットワークの速度も上げる必要がある。さらにライブラリの開発も必要だ。NVIDIAはハードの開発会社とされているが、ソフト技術者も沢山いる。生成AI、特にChatGPTの衝撃はすごかった。2か月で一億人のユーザが出現した。これはインターフェースが良かったからだ。GPTで7兆ドルの経済押し上げ効果があったと見積もられている。

香川県三豊市ではAIを利用し、市民からの質問の回答をAIチャットでやるようにした。これは効果を上げている。

**このように、AIの利用がうまくいったか、すぐに判断できるアプリにAIを利用するのが良い。**

**ChatGPTの学習の多くは出典元にWikiをつかっているようだ。**いかにも本当らしい答えが返ってくるが気をつけないといけない。ハルシネーション（チャットAIなどで、もっともらしいウソ（=事実とは異なる内容や、文脈と無関係な内容）の出力が生成されることを指す。）の弊害が出てくる。だから、利用方法については企業内でも内部で利用についてガイドラインが必要になる。これは国でも同じ事が起きる。

LLM（大規模言語モデル）ではトランスフォーマ技術以前と以後では性能に大きな差が出ている。数百億だったパラメータが数千億になってきており、並列処理しないと動いてくれない。また、LLMの暴走が始まっており、肥大化したデータのシュリンクが必要だ。

そのため**データ発生場所でデータ処理をするエッジ化の技術が重要になってきている。**

## 【感想】

200人超の大盛況のセミナーであった。それにしても技術の進歩はすごい。プレゼンタが自らデモして見せたが、AI（Music Gen）に「侍映画に使えるような音楽を作曲してくれ。」とプロンプトを書くと、即座にそれらしい音楽を作曲して見せた。著作権云々の議論はあるだろうが、この波についていけない人と使い倒す人の違いはアツと言う間に大きくなるだろうな、と言う印象をもった。生成AIの技術は急速に進歩しているし、誰にも止められない。

### Edge活用を加速する「Dell NativeEdge」

【講演概要】 エッジ環境では、センサー、ロボット、ビデオソース、そして近年爆発的に普及したその他のデバイス、システム、「モノ」など、これまで存在しなかったデータソースによって大量のデータが生成されている。エッジは、生産性を向上させ、市場における競争上の差別化を達成するための重要な領域となってきている。しかしながら、セキュリティ管理や分散型システム運用への対応などエッジ環境の独自の課題があり、「Dell NativeEdge」がこれらをどのように解決して行くのか、その紹介があった。

#### 【講演内容】

IoTが進むにつれ、データセンターへのアクセスの集中が莫大になってきている。エッジでデータが発生し、これがネット経由でデータセンターへ集まる。これを非効率と思う人が76%。電力不足になる原因でもある。

ITの新しいオペレーションでは多拠点分散処理を検討する必要がある。その一つの解はエッジでの分散処理方式。DELLではエッジ分散処理を効率的に行うため、エッジ処理機器群を準備した。これでネットワーク途中のサプライチェーンも含めた、全体の効率化も可能となった。

・ Dell NativeEdge | Dell 日本

<https://www.dell.com/ja-ip/dt/solutions/edge-computing/edge-platform.htm#tab0=0>

#### 【感想】

もっともな主張ではある。DELLとしてはこのコンセプトで新しいユーザを取り込もうと言うのだろう。しかし、これはクラウドでのサービス構築が基本になってきている現下の流れと、どう折り合いをつけるのが課題となりそう。分散か集中か状況によって、これはいつも振れている。

### 生成系AIを『理解』したいあなたに贈る AWS の生成系AI入門

【講演概要】 皆様の中には上司から「生成系AIうちでも導入しよう！よろしく」と言われた方がいらっしゃるのではないだろうか？『でもどこから始めればわかんない、生成系AIコワイ』多くの方がそのように感じられていると思う。そんな方に向けてこのセッションではAIおよび生成系AIの基礎からお話する。またすでに使っているお客様の声も交えて活用方法を紹介する。生成系AIにチャレンジする技術部門の方にとって、先人の知恵を得られられるようなセッションとなれば幸い。Let's think step by step. という主旨で行われたセッション。

#### 【講演内容】

Amazonには各種AIサービスを取り揃えている。

- ① Amazon Kendra 機械学習を利用して、非構造化データから、より関連性の高い回答を検索する。
- ② Amazon Sagemaker Canvas コードを記述しなくても、機械学習を使用して予測を生成できる。
- ③ Amazon Quicksight さまざまなソースのデータを結合し、単一のデータダッシュボードにAWSデータ、サードパーティーデータ、スプレッドシートデータ、SaaSデータ、B2Bデータなどをまとめることができる。これらのサービスには複雑な指示は不要で、普通の文書指示で動かすことができる。東京電力では営業レポート等をこのサービスを使って作成し、効率化が図られた。

#### 【感想】

生成AIの品ぞろえはここにあるだけでも、機能別に各種揃っている。大企業でもこれらを既に業務に使っているようで、効率化のカギと考えているようだ。

典型的な業務の流れは「データ検索⇒提出文書の作成⇒ダッシュボードで業務分析」という一連の業務処理。これらの殆どがAIで処理され、作業者の独自性はいかに効率的、かつ独創的なプロンプトを書くかで決まる。

## ChatGPTはどう言葉の意味を理解しているのか？

【講演概要】 GPT3とGPT4は、トランスフォーマがアテンションメカニズムを使って急速に進化させたニューラルネットワークNLPモデルだ。ChatGPTは基礎モデルをチューニングし、人間のフィードバックと強化学習をモデル化して作られ、ChatGPTは驚異的な対話能力に到達した。単語間の距離をグラフ化するWord2Vectorや、時間の経過を理解できるようになったLSTMなど、自然言語処理には進化の変曲点がある。そうした生成AIはどのように進化してきたのか。本セッションでは、ChatGPTの使い方の概要やビジネスへの影響といった話題ではなく、自然言語処理における生成AIの基本的な動作原理を真正面から解説し、機械がどのように言葉の意味を理解し、幻覚を見るのかについて議論がなされた。

## 【講演内容】

これまでは人間が問題を出し、機械が答え、これを人間がスコアを付け、評価しシステムをチューニングしていた。今回、生成AIが飛躍的に進歩した理由は、問題のスコア付け、評価も機械が自分で行うようになり、効率が飛躍的に上がったためである。

言語処理を行列式の演算に移すことによって、理論的にも生成AIの動きを改良できる。

## 【感想】

これは本格的な生成AIの技術解説で、なかなか理解は難しかった。数式を使って詳細に説明されたが、資料もなかったため、メモも取り切れていない。下記のURLを参考に追加の理解をお勧めする。

・技術と科学の未来講座MaruLabo

<https://www.marulabo.net/>

要は生成AIの技術も数学（=行列式）で記述され、それによって機械処理も容易になり、急速に進んだと言う主旨の講演であった。

## IoT・エッジ領域で生成系AIを活用する戦略と、これからのチャレンジ

【講演概要】 生成系AIは、従来の機械学習や深層学習の分類作業と違い「AIが直接新たな価値を生む技術」として注目を集めているが、実行には高い計算能力が求められる。特にIoTで用いられる小型デバイスやエッジコンピューティングで生成型AIを使うには、どんな方法があるだろうか。このセッションでは、生成系AIの概要やプロンプトエンジニアリングといった基礎の振り返りから、マイコン等との連携手法の現状を紹介し、今後私たちが生成系AIを組み込み領域で活かす方向性はどうかを見出す主旨の講演であった。

## 【講演内容】

ChatGPTは知っている人は50%以上だが、実際に使っている人は10%以下。使わない人は流れに遅れるばかりだ。

生成AIを使いこなすカギは LLM（大規模言語モデルの中身）×PE（プロンプトエンジニアリング）である。

巨大なデータを人間がハンドリングするのは無理がある。例えば、実際の応用例だが、群衆の写真をスキャンして、その中に人が何人いるかAIは正確に勘定する。こんなことを人がやる必要はない。

生成AIの基盤モデルではすでに人間の常識レベルの学習をしている。例えばレシート画像を読み込ませて、「支払い金額はいくらか？」という問いに、レシートの合計金額を読み取って、支払金額を答えてくる。

このレベルになっている。（支払金額と合計金額が同じ意味とすることを指示しなくても理解している。）

生成AIを賢く使うには、LLMに組み込んである範囲でユーズケースを限定し、目的をはっきりさせればうまく使える。

既にLLMもハードに組み込まないと応用に限界がきている。（=Edge技術）ポケットはネット経由のためエラーが起きると再送しないとイケない。これは非効率だし、使い勝手が悪い。ハードへの組み込みが必要ということだ。

ビルゲーツの言葉らしいが：

『効率の良い仕事にAIを組み込めば更に効率は良くなる。効率の悪い仕事に組み込むと効率は更に悪くなる。』

### 【感想】

ビルゲーツの言葉には強く同意する。知らない仕事の効率化はできない。よく知っている仕事であれば、改善したのか否かの判定もできる。知らない業務にAIを使えば、AIの迷路に入り込む。よく理解している業務から効率化を図るのが正解だ。

### 【総括】

生成AIの進歩は想像もつかない速さで進んでいる。チャットのみならず、画像、音楽にまで応用は広がっている。それらの進歩についていかないとどうなるか、結果は見えているような気がする。それに合わせて、生成AIをクラウドで使う事にも問題がある。ネットワークの負荷と同時に、データセンターの巨大化に伴う電力負荷である。これを解決する技術がエッジ技術になると言う主張には納得できるが、そのうちにまた、センター処理、クラウド処理の流れに戻るだろう。特に量子コンピュータが実用化されれば、これをエッジで使うと言う訳にはいかない。流れは繰り返さざるを得ない。

以上