

日本IBM

「IA」なくして「AI」なし AI実用化のための実践的アプローチ

AI活用に挑む多くの企業が“PoC疲れ”に陥っている。サイロ化したデータとその分析環境に課題があるからだ。データの準備に多大な工数を取られ、AIモデルの品質が上がらない。分析には高度なスキルも求められる。この壁を乗り越えるためには、統合的なIA(情報アーキテクチャ)基盤が必要である。その実践的アプローチを日本IBMの田中孝氏が解説する。

日本アイ・ビー・エム株式会社 クラウド&コグニティブ・ソフトウェア事業本部 Data and AIテクニカルセールス部長 田中 孝氏



AI実用化を阻む2つのハードル 統合的なIA基盤の整備が重要に

多くの企業がAIを戦略的に活用し、データに隠された洞察をビジネス価値につなげようとしている。国内AIシステム市場も右肩上がりでの拡大を続けている。しかし、すべての企業が期待通りの成果を上げられるわけではない。AI実用化の前には、高いハードルが聳えている。1つは、本番環境へのデプロイへと進む際のハードルだ。AIの技術検証に取り組み手応えを感じることもあるが、本番環境の構築へと進むGOサインが出ない。実際、AI活用に取り組む企業の半数以上がこのハードルを越えられず、“PoC疲れ”に陥っているという。もう1つ

は、全社展開のハードルだ。1つめのハードルをクリアし、部門レベルでは一定の成果を出すものの全社にAI活用を浸透させることができない。ビジネス成果も限定的になり、AI活用の機運がしばんでいく。

「2つのハードルを乗り越えるためには『データ』『人材』『信頼』という3つの要件に目を向ける必要がある」。こう話すのは日本IBMの田中孝氏だ。

データはAIの生命線である。より高品質な学習データをスピーディに調達することが重要である。そのデータをAIで分析・利活用するためには、専門スキルを持つAI人材が不可欠だ。しかし、AI人材は慢性的に不足している。誰でもAIを使える環境を整えることで、AI人材は

より高度な分析業務にシフトさせる必要がある。加えて「AIシステムが本当に正しい結果を導き出してくれるのか」という利用側の懐疑を払拭しなければ、AIに対する信頼も確立できない。この3つの要件を整えることで、AI実用化に向けた取り組みが大きく前進する。「まず分散しているデータのサイロ化を解消し、データとAIのライフサイクル管理を高度化する統合的なIA(情報アーキテクチャ)基盤を構築することが重要です。IAなくしてAIの実用化は達成できない」と田中氏は主張する。

“使うAI”と“作るAI”で AI適用範囲が大きく広がる

このIAを実現する多様なソリューションを体系化したものが「IBM Data and AI ポートフォリオ」である。AIシステム「IBM Watson」の価値を最大限に高める開発・分析環境を実現する。「様々な場所で発生するあらゆるタイプのデータにシンプルにアクセスできるようになり、そのデータのクレンジングやカタログ化が容易になります。さらにスピーディなAIモデルの構築と洞察の発見を支援し、分析結果に対する説明性も担保しやすくなります」と田中氏は説明する。

より実践的なAI活用を加速する仕組

“使うAI”と“作るAI”の機能概要

汎用的なユースケースには“使うAI”が、専門性の高い分析要件には“作るAI”が有効

アプローチ	活用方法	必要作業		
		モデル構築	学習データ	アプリ組み込み
“使うAI”	典型的なユースケースについて定義済みAIモデルを活用する	不要	不要	要
		不要	一部要 (固有部分のみ学習)	要
		不要	要	要
“作るAI”	実現したいユースケースに応じたAIモデルを独自に開発する	一部要 (ライブラリ等を活用)	要	要
		要	要	要

みもIBM Watsonとして提供している。それが“使うAI”と“作るAI”だ。「“使うAI”は、AIモデルの構築が不要。典型的なユースケースをもとに定義済みAIモデルを活用して、業務にあったデータを学習させることで、自然言語処理や音声認識など用途に応じたAIモデルをスピーディに利用することができます」(田中氏)。

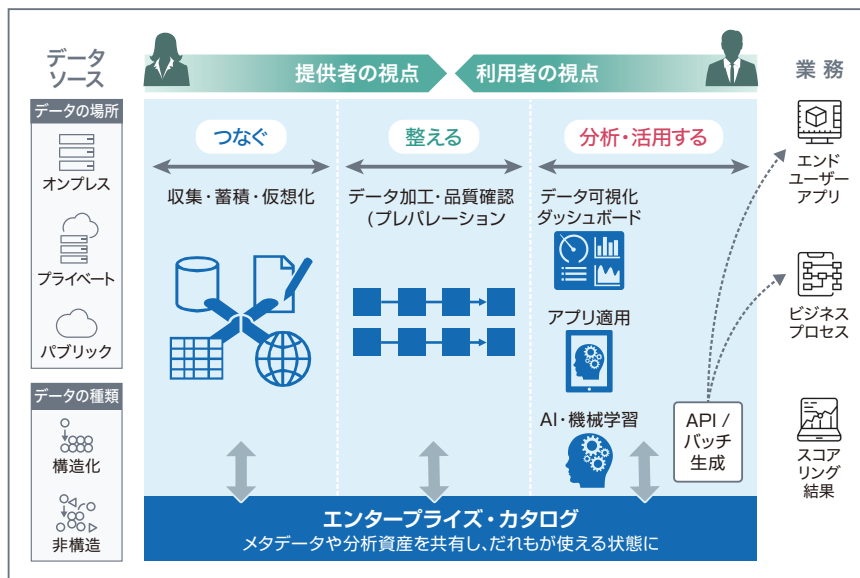
専門性の高い個別の分析要件には“作るAI”が適している。実現したいユースケースに応じて、カスタムAIを独自に開発できる。AIモデルの構築と学習データの整備が必要になるが、オープンソース技術を活用した統合開発環境で多様なライブラリ群を利用できるため、構築・稼働・評価の作業を効率的に行える。「2つを用途に応じて使い分け、組み合わせることで、AI適用範囲が拡大し成果も上げやすくなるでしょう」と田中氏は語る。2019年6月にはカスタムAIの開発・管理をより効率化する「AutoAI」をリリースした。データ前処理、モデル選定、パラメータチューニングなど含めたモデル最適化を自動的に行い、AIモデルの開発を高速化するものだ。AIモデルの構築と管理がさらに容易になり、分析結果に対する説明性・信頼性も向上するという。

データの加工・品質向上を支える 次世代データプラットフォーム

ビジネス成果を高めるためには、データの品質と鮮度が重要な要素になる。これがAIモデルの精度を左右するからだ。「常に高品質な新しいデータを適用し、構築したAIモデルを高度化させることが欠かせません」と田中氏は訴える。実際、データの収集、学習データとして

次世代データプラットフォームの必要性

多様な分析ニーズに対応し、データをつなぐ・整える・分析・活用する基盤が欠かせない



利用するための加工・準備作業は、AI活用プロセス全体の8割を占めるといふ。継続的学習サイクルを迅速かつ効率的に回す仕組みが必要である。

これを支援するため、IBMは次世代データプラットフォーム「IBM Cloud Pak for Data」を提供している。データを収集・加工・分析するためのソフトウェアを100以上コンテナ化して実装し、Red Hat OpenShiftが稼働するすべての場所でIBMソフトウェアを利用できる。幅広いデータとAIのサービスをIBM Cloudをはじめパブリック・クラウドやプライベート・クラウド(オンプレミス)にデプロイでき、オンプレミスからクラウドまで一貫性のある管理が可能になる。

次世代データプラットフォームはAIシステムに変革をもたらす。「データ分析に関わるすべてのユーザーが統合された1つのユーザーインターフェースでデータを活用できるのです。データの来歴や構造だけでなく、分析プロジェクトで作成されたAIモデルまでカタログ化して管

理することも可能です。データ仮想化により、必要な時に必要なデータにアクセスできるため、データ活用の試行サイクルも大幅にスピードアップします」と田中氏はメリットを述べる。

ニーズに合わせて、必要なツールや環境も柔軟に追加可能だ。DockerやKubernetesを活用したコンテナ化により、マルチクラウドに対応したデータ資産の活用が進み、リソースも最適化できる。この次世代データプラットフォームはIBM Data and AIポートフォリオの基盤としての役割も担っている。“使うAI”と“作るAI”、そしてAutoAIとも連携し、より高度なAIプラットフォームへ進化する。AIのビジネス実装と全社展開を成功させるためには、データ活用とAIモデルの構築・評価のPDCAサイクルを高速に回していく仕組みが欠かせない。IBMはデータ収集から価値の創出・実行までデータ活用のフルサイクルを支援する統合プラットフォームの提供を通じ、多くの企業のAI実用化に貢献していく。