

データドリブン業務実践フレームワーク

A Framework for Implementing Data-driven Operations

森 純 一*
Junichi MORI

石谷 紗知
Sachi ISHITANI

山田 邦晴
Kuniharu YAMADA

岩田 泰士
Yasushi IWATA

東 征治
Seiji HIGASHI

西 陽平
Yohei NISHI

磯崎 郷平
Kyohei ISOZAKI

里見 佑太
Yuta SATOMI

在原 大樹
Hiroki ARIHARA

枚田 優人
Yuto HIRATA

山本 政
Masashi YAMAMOTO

抄 録

鉄鋼業を取り巻く事業環境の変化に迅速に対応するためには、従来のような一元的なアプローチではなく、データを起点とした客観的な分析や意思決定を実践することが重要である。このデータドリブン業務を効率的に実践するためには、データを活用するためのスキル、手順を整える必要がある。そこで、日本製鉄(株)におけるデータドリブン業務を実践するためのフレームワークについて述べた。

Abstract

To deal with the business environment changes surrounding the steel industry, it is important to make objective decisions based on data, rather than take unitary approaches. To effectively implement data-driven operations, skills and procedures for utilizing data need to be established. In this paper, we describe the framework for implementing the data-driven operations.

1. 緒 言

鉄鋼業では、コスト・品質面の国際競争力強化に加えて、世代交代対応・労働生産性向上を狙いとして、大量に蓄積されたデータを活用した業務・生産プロセスの改善・改革への期待が高まっている。データドリブン業務とは、これらの改善・改革を経験や勘といった主観的な判断だけに頼らず、データを起点とした客観的な分析・意思決定を、迅速かつ持続的に実施することである。

- 客観性: データに基づき客観的な判断を行っている状態にある
- 迅速性: データが素早く業務判断に活かされる状態にある
- 持続性: データ品質、データ活用モデル、業務が持続的に改善される状態にある

データに基づく人の知的作業の流れは、データを“観察”し、現状を“認識”し、状態を“予測”をすることで、適切な“制御”(意思決定)を行うプロセスを経る。本稿では、これらの知的作業を代替・支援するデータ駆動型の仕組みをデータ活用モデルと称し、その活用範囲ごとに以下の4

つに分類する(図1)。

- ①収集・蓄積したデータを可視化する“可視化モデル”
- ②現在のKPI(Key Performance Indicator)の異常・正常の判断やレベル感を認識する“認識モデル”
- ③将来のKPIの予測を行う“予測モデル”
- ④将来のKPIが改善する最適なアクションを決定する“最適化モデル”

データドリブン業務では、これらのデータ活用モデルをIT技術によりプロダクト化し、業務の中に組み込んでいく必要がある。そのためには、データ活用モデルとその入力となるデータのライフサイクルとともに、プロダクトの本番システム化と保守運用のサイクルを確立しておくことが必要である。

データライフサイクルにおいては、データドリブン業務を成立させるための、品質の高いデータを用意する。“Garbage in, Garbage out(ゴミを入力するとゴミが出力される)”という格言があるように、適切な意思決定には、品質の高いデータが必要である。そのためにも、データマネジメントを確立させ、適切なガバナンスに基づき、データを

* プロセス研究所 インテリジェントアルゴリズム研究センター 生産マネジメント研究室 研究第一課 上席主幹研究員 Ph.D. 千葉県富津市新富20-1 〒293-8511

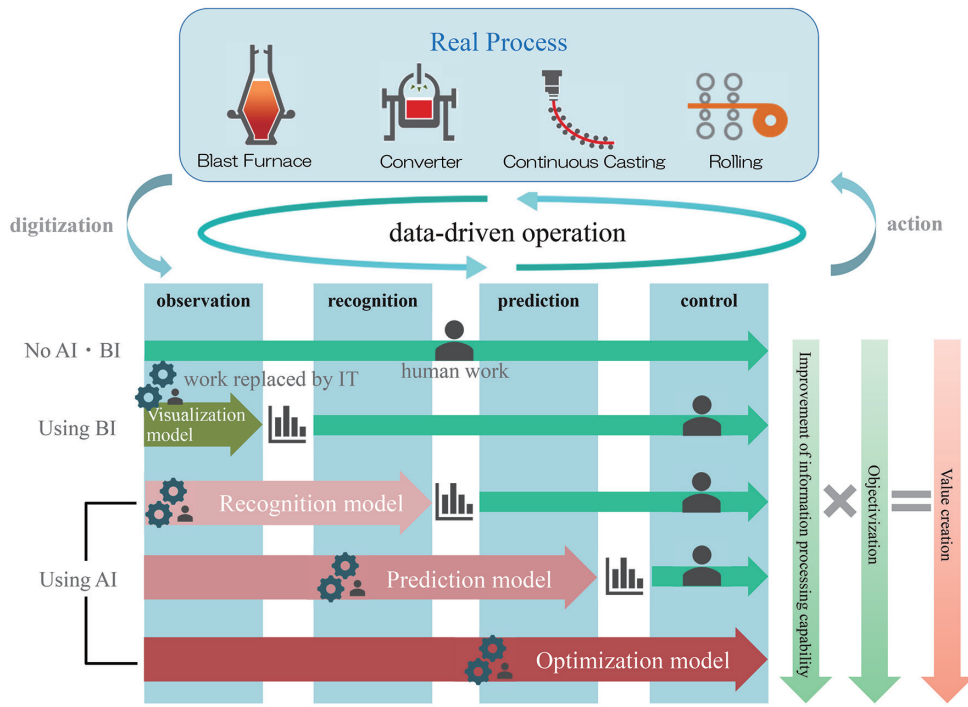


図1 データ活用モデルと活用範囲
Data utilization model and the scope of application

蓄積・提供していくことになる。しかしながら、ガバナンスに気を配りデータ蓄積したとしても、様々な用途を予め想定することは困難であり、利用する段階になってはじめて問題が発覚することも多い。そのため、問題が発覚した際には、それをフィードバックし、品質をすばやく改善する手段を提供することが必要である。モデルライフサイクルにおいては、企画段階におけるデータ準備・理解に始まり、モデリングとモデル評価を行った後、開発したモデルを業務適用していくことになる。モデルに問題が発生した場合は、それをフィードバックし、モデルの再学習やモデル構造の再検討などを通じて、モデル品質を維持・改善することが必要となる。

本番システム化においては、通常システム開発におけるタスクに加えて、モデルを組み込んだ業務・運用設計やモデルの機能・非機能性の担保、さらにはモデル品質を持続的に改善するためにリリースプロセスの自動化を図ることが重要となる。また、保守運用においても、業務上のKPIやモデル・データの品質をモニタリングし、問題がある場合は原因分析したうえでモデル・データライフサイクルへフィードバックすることが、データドリブン業務の持続性を高めるためには必要である(図2)。

日本製鉄(株)では、これまで数多くのデータ活用モデルの開発と業務適用を通じたデータドリブン業務の推進を行ってきたが、下記のとおり、各フェーズにおける様々な要因によって、データドリブン業務が健全に機能しているとは言い難い状況であった。

フェーズ 課題と要因

- 企画
 - 適切な問題設定が行われずに KPI の改善に繋がらないモデルを開発する
 - 類似の技術(社内・社外)を応用すれば解決できることに気が付かない
- 開発
 - 実現性や問題設定が見直されないまま開発にコスト(人・時間)を費やす
 - 検証項目や目標値が明確ではなく、PoC(Proof of Concept)が長期化する
- 本番システム化
 - 組織が分断され、研究成果を素早く実機化できない
 - 運用テストフェーズでモデル精度が不十分であることに気が付く
- 保守運用
 - EOSL(End of Service Life)まで使い続ける前提、変化への追従が難しい
 - モデル変化(性能劣化・データ変化)への対応
 - 事業環境変化(ハイグレード化・統廃合)や業務変化への対応
 - IT技術の変化への対応
 - モニタリングの仕組みがなく変化に気づかぬまま運用される
 - 業務ユーザーに使用を促す仕組みがなく運用が定着化しない

そこで日本製鉄では、各種研究機関・団体などでまとめられた指針を参考にしながら¹⁾、冒頭で述べたデータドリブン視点での業務・生産プロセス改善の目指すべき姿を実現するためのフレームワークを構築してきており、本稿で

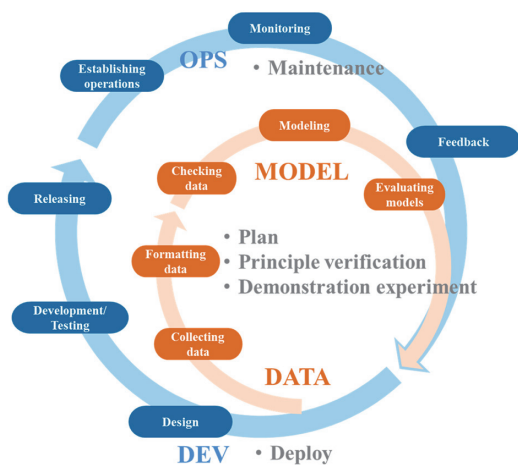


図2 データドリブン業務を持続的に改善するためのプロセス
Process for sustainable improvement of the data-driven operations

はその内容について述べる。以降、2章では、データドリブン業務を推進するために必要なスキルとチーム体制について述べる。3章では、データドリブン業務を支えるシステムを整備する段階をフェーズ分割し、各フェーズの実行内容、次のフェーズに進むための条件を明らかにする。

2. データドリブン業務を推進するチーム体制

データドリブン業務を確立・推進するうえで必要なスキルは、業務を理解したうえで課題を抽出し整理する“ビジネス力”，統計・機械学習・最適化などの“データサイエンス力”，データ処理や実装・運用するための“IT力”である（図3）。

データドリブン業務確立に向けた各フェーズで、必要なスキルセットのバランスは変化する⁴⁾。体制を組む際にはこのスキルバランスを考慮する必要がある（図4）。

データ活用人材は、これらのスキルをすべて保持していることが望ましいとされている。ただし、全スキルが高いレベルにある人材は稀であるため、個別のスキルに秀でた人材を組み合わせ、体制を整えることが一般的である。また、フェーズごとに体制が徐々に入れ替わるとしても、各スキルの専門家が、コミュニケーションを円滑に行い、スピード感を阻害しないためにも、互いのスキル領域に対して理解を深めていくことが重要である。

従来型の開発プロセスでは、企画→原理検証→実地検証→本番システム化→保守運用に組織が分断されており、企画から実用化までに多大な工期を要するだけでなく、モデル開発プロセスのサイクルを回し難い構造となっていた。組織間のロスなく開発プロセスを高速に回すためには、各プロセスを組織で分断しないよう各ロール（データサイエンスユーザー、データサイエンティスト（シチズン／エキスパート）、ITエンジニア、保守運用エンジニア）を担うメンバーがお互いのタスクを認識して協調作業する体制を構築することが必要である。なお、データサイエンスの知識・

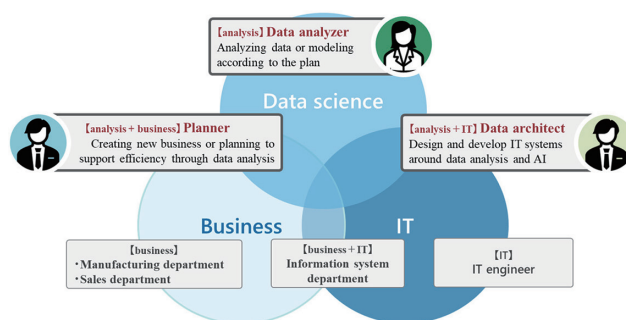


図3 データドリブン業務に必要なスキル
Skills needed for data-driven operations

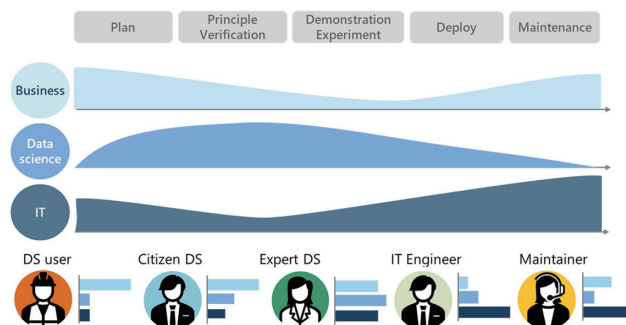


図4 フェーズごとのスキルバランス
Skill balance at each phase

スキルに応じた階層別人材（データサイエンティストユーザー／シチズンデータサイエンティスト／エキスパートデータサイエンティスト）の位置付け及び教育体制については本誌別稿 No.21 にて述べる。

3. プロダクト構築ステップ

プロダクト構築プロセスは、下記のフェーズで表現される。このフェーズは“試行錯誤的”なアプローチになっており、場合によっては後戻りしてやり直すことを許容する^{5,6)}（図5）。

このフェーズはあくまで目安であり、不確定要素が少なければ、企画と原理検証が混在し、フェーズを省略するなど、様々なパターンがあり得る。重要なのは、各ステップをなるべく高速に回しながら、あらかじめ設定したフェーズゲートにおいて“先に進む”か“やり直す”かを判断し、不確定要素を着実に減らしていくことにある。

プロダクトは、保守運用に入った後も継続的に改善や機能発展を行うことを見越しておく必要がある。そのためにも、フェーズが進むにつれ、前述のモデルとデータのライフサイクル、さらにはシステムの継続的改善（DevOps）の確立と効率化を行っていくことが重要である。以降では各ステップにおける実行内容、次のフェーズに進むための条件を明らかにする。

3.1 企画

企画フェーズでは、データに基づく業務改善の目的とア

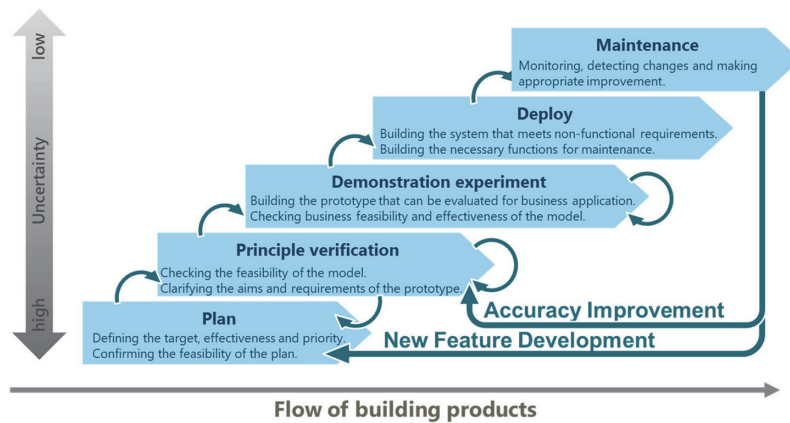


図5 プロダクト構築ステップ
Step for building products

アプローチを明確にし、関係者間で共有するための情報を整理する。ここでは、現状把握するためにデータフロー図や業務フロー図といったツールを使い、業務分析、業務ヒアリングを最初に実施する。そして、理想状態とのギャップ分析を通じて、データ活用の目的を明確にしていく。その際、目的と対応する形で、達成水準を測る KPI を設定することが重要である。KPI は、目的を達成することで得られる効果を定量的に測る指標となる。KPI は経済効果で表現できると、関係者間で合意形成しやすく評価も容易になる。しかしながら、企画初期から KPI を確定することは難しい。これはデータに基づく意思決定、さらにはアクションまでを具体的にイメージできていないことが要因である。そのため、業務分析や関係者と議論する中で具体的な取りうるアクションを想像しつつ徐々に KPI を定めていく必要がある。目的達成に対して、AI 技術が最適であるかを判断することも重要である。ヒアリングを通じて、解決すべき技術課題、検証すべき技術的要素が明確化されるなかで、当該業務への AI 利用あるいは、AI で代替すること自体のリスクも評価する。例えば、AI は精度 100% を保証できないが、AI を適用して職場を完全自動化することにどのようなリスクがあるかを評価する。

また、企画時のデータ調査は、業務理解にも役立ち、データ品質を明らかにすることで実行可否の判断基準にもなる。そのため、数値データであれば、業務データを蓄積する共通基盤とセルフ BI (Business Intelligence) ツールを組み合わせ、業務ユーザー自らがデータ調査できる環境を用意しておく。とくに機械学習モデルを構築することが目的であれば、目的変数や説明変数候補となるデータについて、欠損値や異常値の影響、データを取得できるタイミングなどをよく精査しておくといよい。

画像データの場合は、そもそも画像データの蓄積から取りかかるケースもある。そのため、共通基盤として画像とデータを時系列保存できる解析ナビシステムや、映像を全社から収集保管配信するロケーションフリー ITV システム

が整備されている。一部の画像処理システムでは、処理速度と効果／費用に応じたメニュー化もされており、このような既存の仕組みを活用することを企画段階で決定し後続の検討を省力化することも考えられる。詳細については、本誌別稿 No.8 にて説明する。

企画では、これら整理した情報をもとに関係者間でレビューを行い、得られる効果とコストを比較して、可否を判断する。ただし、この時点での効果やコスト見積りはあくまでも概算値であるため、値の大小に捉われ過ぎず、多少のリスクは許容してもよい。

3.2 原理検証

原理検証フェーズでは、プロダクトの有効性・実現可能性を検証する。なかでも不確実性が高い要素であるモデルの有効性を、時間と労力を抑え素早く検証することが重要である。そのため、検証環境準備に時間をかけず、データ取得や加工などの繰り返し処理はなるべく自動化できるように、データ処理や分析を行う共通基盤を予め整えておく。検証の対象範囲を限定するなど検証プロセスを工夫することも有効である。また、試行錯誤的に検証実験を行うため、実験管理も重要となる。いつどのような条件で検証したか、どのような結果だったかを管理し、成果の証跡として残す。それだけでなく、検証の再現性を担保するため検証環境を仮想化し、保管する方法がとられることもある。

さらには、この時点で実地検証にむけた要件整理も考慮する。原理検証は素早さを優先するため、オフライン検証とすることが多く、実業務に照らし合わせる検証が不完全になる場合がある。そのため、不完全な部分を補う追加検証を実地検証の要件として抽出しておく必要がある。

原理検証で用いるデータは、なるべく本番と同じデータとする。とくに画像データは、対象物・環境・条件が本番と一致していないと有意な検証を行うことはできない。本番データを使うことで“実環境で想定よりも精度が低い”という後続フェーズのトラブルを回避できる。例えば、製

品に記載された ID を読み取る画像認識モデルを検討した場合、実際利用する場所は屋外にもかかわらず、屋内で撮影した画像で構築したモデルを実用化すると、逆光時等の考慮が不足しており、誤判定の可能性がある。そのため、できる限り実環境で撮影した画像を用意する。一方で、この時点で同じ条件でデータを用意できないケースもある。その際は本番データとの差異から想定される技術的な課題を想定・抽出することが重要である。また、データ量や利用するアルゴリズムによっては処理時間がかかるものもあり、業務要件にあわせて適切な手法を検討することになる。とくに深層学習のようにハードウェア性能の寄与が高いものは、モデルを実行する共通基盤の仕様を鑑みて、処理時間を見積もらなければならない。

モデル評価は、誤差や正解率など一般的な精度指標だけでなく、業務影響がわかりやすい指標を併せ業務ユーザーに報告し評価する。事前に決めた KPI はここで指標として使われる。指標が達成水準を満たしていること、業務ユーザーの納得感がある結果が出せていることを確認し有効性の評価とする。また、データ蓄積／取得から結果出力までのデータ処理に無理がなく、自動化の見込みがあることや処理時間の目安も含め総合的に評価を行う。

3.3 実地検証

ユーザーが実業務で使える仕組みを整え、ユーザーによる施策の有効性の評価や業務成立性を確認する。実地検証ではデータ活用モデルの業務的価値の“検証”、業務ユーザーのフィードバックを受けた後のデータ活用モデルの“改善”や検証範囲の“拡張”を繰り返し実行することを目的としている。このため、なるべくコスト(期間×労力)をかけずに業務ユーザーにプロトタイプを提供し、早期にフィードバックを得ることが重要であり、以下の視点でプロトタイプのコア機能と評価方法を定めたいうで実地検証を行う。

- ・外部データ連携機能／アプリケーションの造り込みをしない
- ・価値検証に直接影響を与えないユーザビリティ向上のための機能開発に時間をかけない
- ・施策の有効性評価に直接関連しない非機能要件は考慮外とし、機能要件の精度向上に注力する

本ステップの初期段階では、機能面を評価する検証範囲(品種、部門など)、対象データを限定することでデータ活用価値を早期に検証し、業務ユーザーからのフィードバック獲得とモデル改善に注力する。部分最適から始まり徐々に全体最適を図っていくアプローチであることを関係者に周知しておくことも重要である。また、既存システムのデータ更新に同期したプロトタイプのリアルタイム実行を追求せず、タスクスケジューラや RPA (Robotic Process Automation) ツールによる定期実行や人介入を許容することで、

業務価値検証を早期に開始する。業務価値検証にリアルタイムなデータ連携が必要な場合に限り、ミラー環境／データベースのレプリケーションにより並行本番環境を準備する。

ユーザーへの可視化は BI ツールの標準機能を用いた実装方法を第一に考える。標準機能での実現が難しい場合は、手動での実行を許容できないかを検討する。手動検証を十分に実施したうえで、業務的価値検証には GUI (Graphical User Interface) の造り込みが不可欠と判断される場合は内製化を行う。GUI の内製化にはコスト(期間×労力)がかかるだけでなく、プロトタイプの開発サイクルを長期化してしまうリスクがある。

そこで、日本製鉄では、セルフ AI プロトタイプの雛形とガイドラインを整備し、シチズンデータサイエンティスト自身でセルフ AI モデル開発から実地検証を進められるようにすることで、素早く業務適用判断まで進められるようにしている。シチズンデータサイエンティストの中には、Python でセルフ AI モデルを開発するスキルはあるが、プロトタイプアプリ開発や実地検証の経験が少ない人もいる。このため、現状のスキルをベースにガイドライン、開発支援ツールでスキルを補強できるように、ガイドラインではシチズンデータサイエンティスト自身による実地検証とセルフ AI プロトタイプ開発プロセスを定義し、プロセスに準拠したセルフ AI プロトタイプ開発支援ツールとして、プログラム雛型(テンプレートコード)と開発手順書を提供している。

この際、提供するガイドライン、セルフ AI プロトタイプ開発支援ツール、テンプレートコードは、多くの社内事例をカバーすることが求められる。社内セルフ AI プロトタイプ開発事例を収集し、実地検証による頻出事例、セルフ AI プロトタイプの機能要件、業務要件、開発する場合の課題を調査することで、社内のセルフ AI プロトタイプ開発事例で頻出の機能、類型をカバーして、社内でも広く適用可能なガイドラインとしている。セルフ AI プロトタイプ開発支援ツール、テンプレートコードは、シチズンデータサイエンティストが理解しやすい Python の Web アプリ開発向けのフレームワークを採用し、コードのベースにしている。しかしながら、サンプルコードを元にしてもプログラミングを専門としていないシチズンデータサイエンティストにはプログラム作成量が多く、プログラムの難易度が高いと抵抗感を持たれ使われなくなってしまう。そこで、開発を容易にする工夫として、社内セルフ AI プロトタイプ事例で頻出となる機能・画面パーツを多く用意し、極力穴埋めだけでセルフ AI プロトタイプを作成できるようにすることで、シチズンデータサイエンティストのプログラム作成量が少なく抵抗感が生まれないようにしている。

テンプレートコードをベースにすることで、シチズンデータサイエンティスト単独でセルフ AI プロトタイプを開発

し、そのまま実地検証を進めることで、これまで6か月以上要していた実地検証の工期を半分以下の2〜3か月で実施できるようにしている。

製鉄所の現場作業員による業務適用評価を実施するためには、製造現場においてリアルタイムに適用テストできることが望ましい。画像処理は検出対象や制御対象によって必要なユーザーインターフェースが異なるため一件ごとに内製することになるが、画像処理基盤内に音声アラート機能、パトライト制御、グラフ表示機能を備えていることで、低コストかつ短期間で表示ガイダンス機能を構築できるようにしている。画像を含む処理判定結果を前述の解析ナビシステムに登録し解析できる機構をもっており、性能評価を容易にしている。

以上の取り組みを通して、想定した検証範囲においてKPI改善に繋がる意思決定・アクションをガイダンスできしており、検証範囲外においてもKPI改善効果の目途が立つ。あるいは検証範囲内のみでも十分な改善効果が得られる場合は、先のフェーズに進む。検証不足により判断できない場合は、検証ケースを増やす、検証範囲を拡大するなどして不確実性をできるだけ小さくする。また、本番システム化、運用のコストと、得られるKPI改善効果のバランスから進退を決定することも重要である。

3.4 本番システム化

データ活用モデルを含むプロダクトは、利用者及び他システムとの入出力により機能を発揮するため、業務システム及び業務プロセスの一部として組み込むことが必須である。ここでは、実用化にあたり、(1)業務設計、(2)システム設計、(3)データ収集、(4)運用設計の順に要点を述べる。

業務設計では、データ活用モデルの出力である可視化結果、画像認識結果や数値予測結果を業務で活用するシナリオを明確化し、そのために必要となる機能要件、目標となる速度や精度等の非機能要件を再度確認する。とくにAIを導入する場合には、AIの仕組みや制約などにに基づき、AIを利用すること自体のリスクを評価する。AIは完全な精度保証ができないため、精度のバラつきが業務に与えるリスクを理解し適切に評価する必要がある。また、実際に精度不良が発覚した際の業務フローも事前に検討すべきである。例えば、製品に記載されたIDを読み取る画像認識モデルで認識不良が発生した際には、一時的に利用者が目視で訂正を行う業務も併せて考えておく。このようなトラブルの発生頻度や作業負荷等の面から現実的な業務フローを検討するよう留意する。また、利用者とのインターフェースとして効果的な結果の見せ方を考える必要もあり、原理検証、実施検証から得られる利用者のフィードバックに基づき画面構成を検討する。

システム設計では、業務設計の内容に基づきシステム構成を検討する。AIが組み込まれたシステムの特徴は、シス

テム全体とAIモデルのライフサイクルが異なることにある。AIモデルはデータの特徴変化によって精度が劣化し、適宜再学習やアルゴリズムの変更の対応が必要になりえる。そのため、システムの保守性・再利用性の観点から、AIモデルを動作させる部分は疎結合になるよう設計する。システム処理中のエラー通知ルールも、このシステム境界に合わせて検討することで、業務システムとAIモデルの問題切り分けを明確化する。システム実行環境については、AIRON-EDGE®(本誌別稿 No.18に記載)のガイドラインに準じた実行環境を候補に、ネットワーク疎通性や速度制約、AIモデルの実行条件を踏まえ選択する。データ利活用を主とするシステムは、通常の業務システムに比べ、データ流量が膨大になるケースも多く、通信料及びストレージについて想定した設計も必要となる。モバイル端末などのクライアントデバイスを活用する場合には、機器性能や利用実績を鑑み業務要件(利用条件)や非機能要件(速度や頻度)を十分検討し選択する。

データ収集の目的は、実地検証時には存在しなかった未知データに対して本番環境でも期待通りの精度を発揮するか検証することである。検証は繰り返し行い、必要に応じてAIモデルを再学習し改善することになる。この時点では、本番で利用するデータ取得機器の設置条件や導入タイミングが決定していることから、それらに応じたデータ収集や蓄積の仕組みを検討し構築する。

運用設計では、AIモデルに付随する運用フローを具体化する。AIは環境変化により精度の低下が懸念される。例えば、工場内に設置したカメラは、利用を続ける中で、粉塵等によりレンズが汚れる可能性があり、汚れによって誤認識となる場合がある。このようなケースが想定できる場合は、推論精度が低下したとみなす条件や確認頻度をこれまでの検証結果から決める。また、精度是正を実行する際のトリガーや実行方法を検討する。

3.5 保守運用

運用保守フェーズでは、プロダクトの活用を定着化するとともに、意図した効果を発揮し続けられるように、モニタリングと改善を行う。プロダクト構築後は、業務ユーザー教育を実施するだけでなく、利用促進に繋がる仕組みを同時に整えることで業務定着を促す。また、システム全体の保守、運用と並行して、稼動中のモデルのパフォーマンスと、モデルパフォーマンスに影響するデータ特性を定常監視する。とくに、AIモデルは、過去データに基づき推論する関係上、環境変化や操業変動もしくはデータ取得機器(センサやカメラ等)の経年変化により、データ傾向や分布が変動することで精度が大幅に劣化する可能性もある。そのため、この傾向変化を検知した場合は、決められた手順で要因分析・対策を検討したうえで、モデルのパラメータ調整などの再学習操作を行い、業務ユーザーが主体となって持

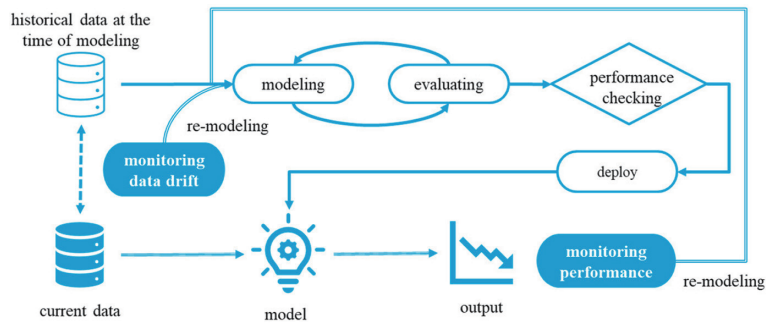


図6 データ活用モデルのライフサイクル
Life cycle of the data utilization model

持続的な効果発揮に努める必要がある（図6）。モデル精度が改善しない場合は、データサイエンティストと共に、再度、分析フェーズを実施する。効果測定は、システムが収集する情報をもとに算出できるようにするが、システム外でのアクションにより効果が発揮される場合は、業務ユーザーへのヒアリングも有効である。

AIモデルのパフォーマンスの尺度としては、連続値を予測する回帰モデルの場合は予測誤差に基づく指標（RMSE, MAE等）、カテゴリ値を予測する分類モデルの場合は正解率に基づく指標（正解率、再現率、適合率、F値、AUC, LogLoss等）が存在する。KPIやモデル活用方法に応じて適切な評価指標を予め設定することが重要である。例えば、後続の工程で顕在化する欠陥を確実に検出したい場合は、欠陥を正しく欠陥と予測することが重要であるため再現率を評価指標の一つに選ぶべきである。

モデルパフォーマンスに影響する入力データの特性変化の検知手法としては、過去データと振る舞いが異なる入出力データを検出する異常検知が有効である。例えば温度計不良のように測定値の信憑性が低い場合はそのデータを除去して再学習を行う必要がある。一方、設備トラブルのようなプロセス要因で異常データを検知した場合は、プロセスそのものの改善に努める必要がある。また、個々のデータは異常値ではなくても、データの統計的分布がモデル訓練時のデータの統計的分布とずれることによって、モデルの予測精度が劣化することがある。予測精度の劣化が顕在化する前に、この統計的分布のずれを検知して、最新データを用いた再学習といった対策を取ることが好ましい。検知手法としては、特徴量の次元削減や統計的仮説検定に基づく手法が提案されている。

また、画像系AIモデルを搭載したシステムは撮影機材の更新や周辺環境変化により調整が必要になる場合がある。このような場合、画像処理基盤が全録画している画像データを開発環境に複製して、再現テストが行えるようになっている。再調整をオフラインで行えるため、現地に画像処理要員を配置しなくても対応できる。一方、画像処理の展開速度アップのためには、現地に画像処理要員を手配する。画像処理技術者を育成しており各製鉄所への配置も

積極的に行っている。

ITシステムである以上、ハードウェア・OS・ミドルウェア・ライブラリなどのバージョンアップ対応も必要である。これらは組み合わせに制約があり、複数の構成要素を同時に更新するケースも多く、影響範囲に考慮する。バージョンアップ前後でAIモデルの挙動はライブラリ等の内部構造変更・数値計算の誤差・乱数の影響などにより、完全に一致しないことから、バージョンアップ前後での出力の比較結果を利用者に提示し、業務上の問題がないことを確認したうえで、更新することが重要である。

4. 結 言

鉄鋼需給環境の変化、気候変動、労働人口の減少を始めとした、鉄鋼業を取り巻く事業環境の変化に対応するためには、データドリブンに客観的な分析や意思決定を迅速かつ持続的に実施していくことが重要である。本稿では、このデータドリブン業務を実践するためのフレームワークとして、データドリブン業務を推進するために必要なスキルとチーム体制、支援するプロダクトの構築ステップについて説明した。スキルにはビジネス力とデータサイエンス力、IT力があり、各スキルのバランスを考慮してチーム体制を組むことが重要であることを説明した。また、プロダクト構築ステップでは、企画、原理検証、実地検証、本番システム化、保守運用の各フェーズにおける目的、タスク、手法及び次フェーズへの進退を決めるフェーズゲートについて述べた。本フレームワークは、日本製鉄システム開発部門における実業務での活用から始まり、現在は本フレームワークをデータドリブン業務実践ガイドラインとしてまとめ、年間約200人の社員が参加する全社研修にも活用することでDX人材の育成にも貢献している。

参考文献

- 1) 国立研究開発法人産業技術総合研究所：機械学習品質マネジメントガイドライン第3版。2022年8月
<https://www.cpsec.aist.go.jp/achievements/aigm/>
- 2) DAMA International：DAMA-DMBOK2：データマネジメント知識体系ガイド第2版。2018年12月

- <https://www.dama.org/cpages/body-of-knowledge>
- 3) 人工知能学会：AI マップ β2.0. 2020 年 6 月
<https://www.ai-gakkai.or.jp/resource/aimap/>
- 4) 情報処理推進機構：データサイエンティストのためのスキル
 チェックリスト／タスクリスト概説. 2020 年 7 月
<https://www.ipa.go.jp/files/000083733.pdf>
- 5) 日鉄ソリューションズ(株)：＜技術動向＞ AI 時代のデータ活

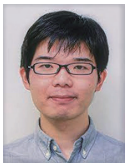
- 用は「DataOps」で. 2019 年
<https://www.nssol.nipponsteel.com/future/stories/assets/nssol-ss-prir-cas-174-01.pdf>
- 6) 日鉄ソリューションズ(株)：データ活用力を DataOps で高める. 2021 年 2 月
<https://dcross.impress.co.jp/industry/column/column20201007/index.html>



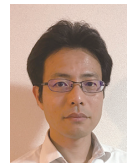
森 純一 Junichi MORI
 プロセス研究所
 インテリジェントアルゴリズム研究センター
 生産マネジメント研究室 研究第一課
 上席主幹研究員 Ph.D.
 千葉県富津市新富20-1 〒293-8511



里見佑太 Yuta SATOMI
 プロセス研究所
 インテリジェントアルゴリズム研究センター
 生産マネジメント研究室 研究第一課
 主任研究員



岩田泰士 Yasushi IWATA
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 インテリジェンス研究部 統括研究員



山本 政 Masashi YAMAMOTO
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 研究戦略・ソリューション化推進部
 専門部長 上席研究員



磯崎郷平 Kyohei ISOZAKI
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 インテリジェンス研究部 主務研究員



山田邦晴 Kuniharu YAMADA
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 インテリジェンス研究部 統括研究員



枚田優人 Yuto HIRATA
 情報システム部 情報システム企画第二室
 AIソリューション課長



西 陽平 Yohei NISHI
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 インテリジェンス研究部 主務研究員



石谷紗知 Sachi ISHITANI
 情報システム部 情報システム企画第二室
 AIソリューション課 主査



在原大樹 Hiroki ARIHARA
 日鉄ソリューションズ(株)
 技術本部 システム研究開発センター
 インテリジェンス研究部 研究員



東 征治 Seiji HIGASHI
 設備・保全技術センター
 システム制御技術部 システム制御技術室
 基盤開発課 主査