

# 技術革新 IT



1 ラボでのディスカッション風景(左), 社外プレゼンス事例(右)

## 1 トップデータサイエンティストとノウハウを集約し事業に貢献する Lumada Data Science Lab.

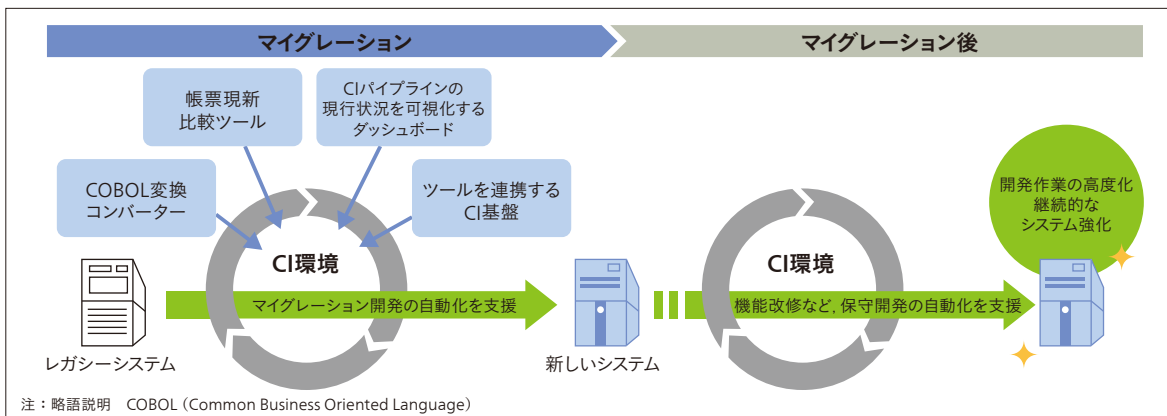
社会価値・環境価値・経済価値の向上をめざす日立の研究開発では、Lumadaを基盤としたグローバルなイノベーション創出や、ステークホルダーとともに成長するエコシステム構築を進めている。2019年4月に中央研究所に開設したオープンな研究開発拠点「協創の森」に、2020年4月には、データサイエンティストのトップ人材として、高度なデータサイエンスと技術の業務適用に不可欠なOT (Operational Technology) の深い知見を有する研究者および事業部のエンジニアなど約100名を結集し、個々のスキルと知見を生かしてコラボレーションする新組織「Lumada Data Science Lab.」を設立した。

多様なデータサイエンティストの連携強化とオープンイノベーションにより、複雑で高度な顧客課題への対応を通じて価値創生を図るとともに、個々の案件によって得られたナレッジを共有する仕組みを立ち上げることで、データサイエンスビジネスの拡大を推進している。

また、AI (Artificial Intelligence) ・アナリティクス分野における人材育成もミッションに掲げており、社内外メンバーによるハッカソン開催、トップカンファレンスでの発表、国際コンペ上位入賞、エンジニアリング記事投稿などにより、社内外における求心力を発揮し、日立のデータサイエンスを主導していく。

## 2 DX加速に向けたマイグレーション自動化技術

新型コロナウイルス感染拡大の影響もあり、企業・



2 マイグレーション自動化技術

産業のデジタルトランスフォーメーション (DX: Digital Transformation) が急務である。データ活用と継続的な業務革新に向けては、レガシーIT資産の刷新が必要である。刷新の方法の一つに言語をモダン化するマイグレーションがあるが、DX加速に向けてはマイグレーションの迅速化が必要である。

そこで、継続的インテグレーション (CI:Continuous Integration) 環境をベースとするマイグレーション開発に特化した自動化技術を開発した。特に、テストの3~5割を占めるとも言われる目視作業が必要なシステム帳票の現新一致確認作業に対して、画像処理技術を応用して自動化を図り、パイロット事例においてテスト工数を3分の1まで削減できることを確認した。また、CI環境をベースとしているため、マイグレーション後の継続的なシステム強化の環境を実現できる。

日立が強みとしている既存システム仕様再生技術と組み合わせ、既存資産を有効活用したレガシーシステムのDX-ready化に貢献していく。

### 3 業務効率化を支える非定型な文書からの自動情報抽出技術

非定型な文書からの情報抽出業務では、文書中のテキストに対して前処理を行い、情報検索基盤での管理や検索なども用いて文書の内容を把握し、文書の中から抽出した記述をデータベースへ登録するという作業が行われている。しかし、抽出結果を手手で1件ずつ確認する必要があり、これまではその作業に多大な工数を要していた。

今回開発した情報抽出技術は、大量のテキストデー

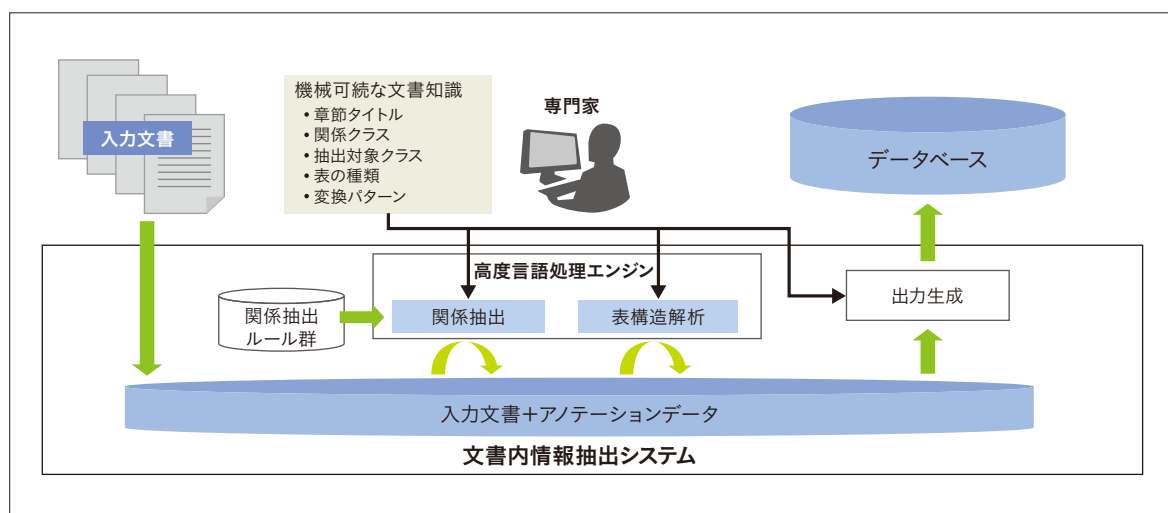
タを日本語の文法構造に基づいて解析し、単語間の関係性を抽出する「関係抽出技術」と、表構造の自動認識を可能にする「表構造解析技術」で構成されている。これらの技術により、文章の構造パターンや表の構造を基に、非定型で複雑な文書から必要な情報を取り出し、データベースに登録するまでの一連の作業を自動化する。

現在は、日本語文書を中心にさまざまな分野で本格適用を開始しており、今後、英語文書へと適用範囲の拡大をめざす。

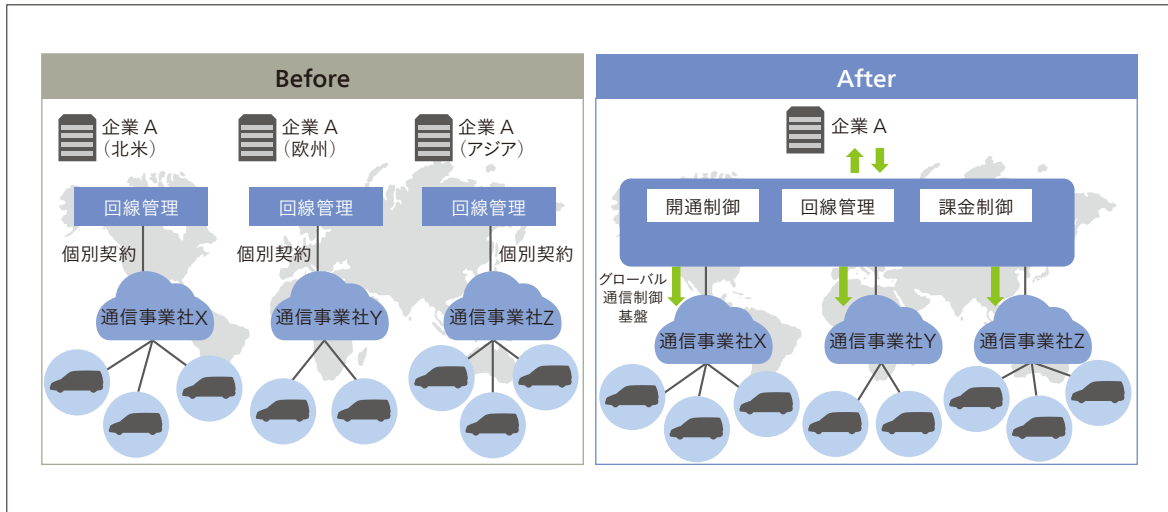
### 4 Hitachi Global Data Integration ソリューション

グローバル企業が世界各地の機器からIoT (Internet of Things) データを収集・活用するためには、各地でのIoT回線の契約・運用と、増加する機器のデータを確実に収集・蓄積するシステムが必要である。そこで日立は、グローバル通信制御基盤技術と、スケーラブルなデータ収集・蓄積基盤技術の特徴とする、Hitachi Global Data Integrationソリューションを開発した。

グローバル通信制御基盤技術では、各地の通信事業者のIoT基盤と接続インタフェース (API: Application Programming Interface) を介して接続し、一元的な回線の制御・管理を行う。これにより、各地の機器におけるモバイル回線の開通・停止や通信状態、課金の管理が可能となる。また、スケーラブルなデータ収集・蓄積基盤技術では、データを収集するデータ受付処理と顧客がデータにアクセスするAPI受付処理を、水平



3 文書内情報抽出システム



4 一元的な回線の制御・管理を実現するグローバル通信制御基盤

スケリング可能な複数マシンで受理し、低負荷のマシンに振り分けて実行する。これにより、増加するデータにあわせた高スループットを実現する。

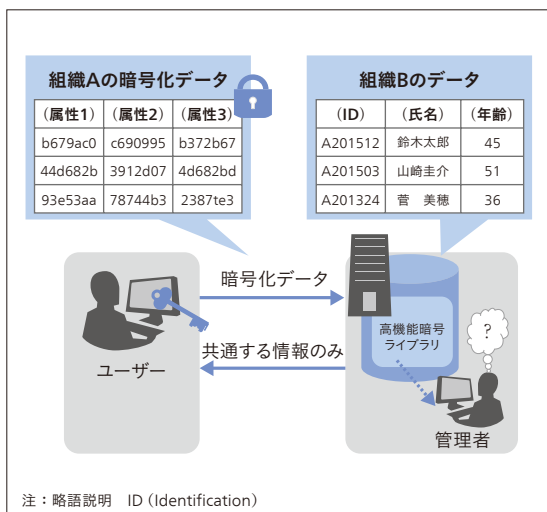
今後、Hitachi Global Data Integrationで顧客のグローバルIoTビジネスの立ち上げ、推進を加速し、世界各地の持続的な産業発展に貢献していく。

各組織が互いにデータを秘密にしなが、目的の達成に必要なデータのみを共有するのは困難である。この問題を解決するため、高機能暗号と暗号実装技術を活用し、複数組織が保有する情報から、暗号化したまま共通する情報のみを抽出するPSI (Private Set Intersection) 技術を開発した。

今後は、共通する情報だけでなく類似する情報の抽出にも取り組むなど、機能拡張に取り組む。また、顧客協創を通じて本技術の社会実装を推進していく。

## 5 データを暗号化したまま 共通集合を抽出するPSI技術

IoTやAIなどの情報処理技術が著しく進展する中で、組織の競争力の源泉がデータの利活用に関する事業に移行し、複数組織でデータを通じたつながりが新たな付加価値をもたらす社会が到来しつつある。一方、複数組織間で共有するデータに機微情報が含まれる場合、開示するデータは目的の達成に必要なものに限定しなければならない。



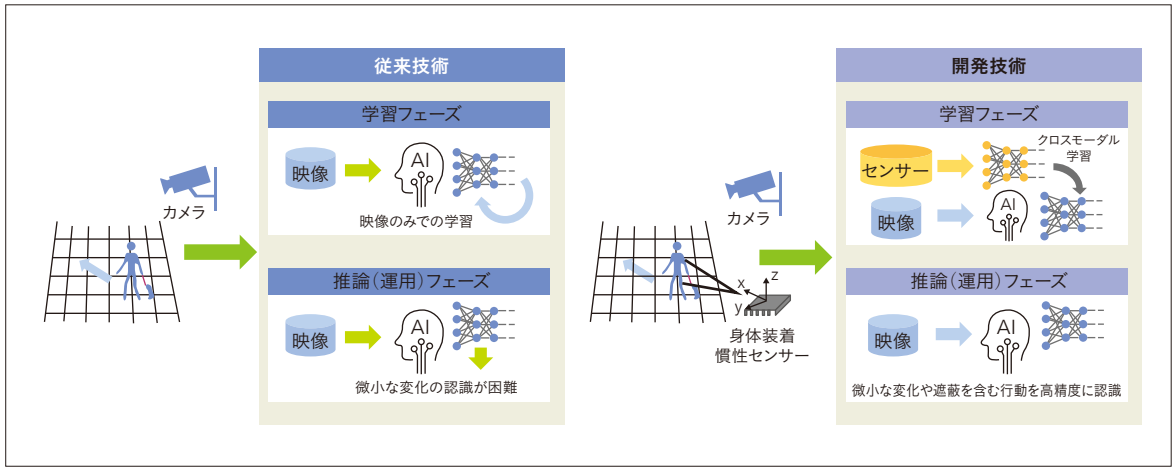
5 PSI技術の活用イメージ

## 6 人の小さな動きを認識できるAI技術

防犯カメラ映像からAIを用いて人物を認識することで、安全・安心を支援する取り組みの普及が進んでいる。これらの技術は、工場作業者の動作解析による安全支援業務や、公共エリアでの不審行動の発見などに用いられている一方で、動きが少ない行動や人混みの中での認識には課題があった。

従来は、映像だけでAIが学習するため微小な動作を教えることが困難であった。そこで、カメラの映像情報と複数種類の身体装着センサーから得られる信号情報を組み合わせてAIに学習させるクロスモーダル学習技術を開発した。センサーと映像を教師として学習したAIは、カメラ映像のみからでも微小な行動の変化を捉え、体の一部が遮蔽されていても高精度に行動認識が可能となった。映像のみで学習した行動認識と比較して、金庫の開閉やスマートフォン操作といった行動の認識精度を最大53%向上した。

今後、本技術を映像監視システムに適用し、不審行



6 センサー情報で映像認識を強化するクロスモーダル学習技術

動の検知や急病人発見のソリューションとして製品展開を図る。

発した。本方式では、従来方式では全く検出できなかった重複区間を80%以上の精度で検出できる。現在、日立の音声書き起こし支援サービスに本方式を適用し、コールセンターなどでの対話状況を分析する実証実験を行っている。

今後、本技術を軸に音声書き起こし支援サービス事業を拡大し、顧客企業のサービス品質向上に貢献する。

## 7 会議の状況を分析する 発話区間検出技術

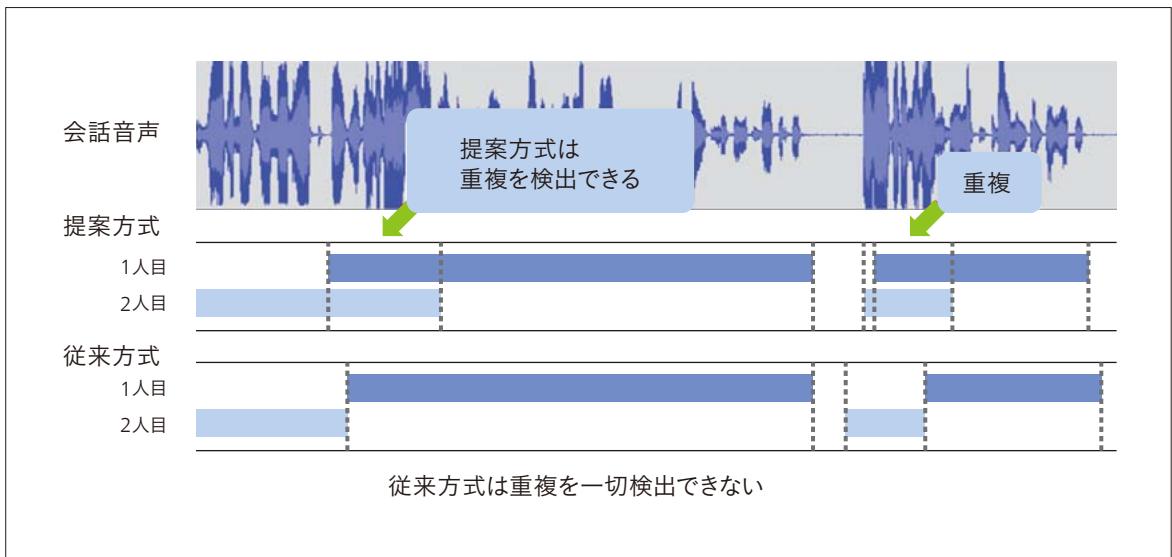
新型コロナウイルス感染症の影響で、企業はテレワークへの移行を加速させ、オンライン会議の導入が進んだ。こうしたオンライン会議において、議事録を作成したり会話を分析したりするために、音声認識の導入が始まっている。しかし、多くの音声認識は一人の発言のみを対象とするため、会議などの分析には話し手の交替を識別する技術が別途必要となる。従来技術では、複数の話し手が重複して同時に話す区間を検出できないため、会議への適用は困難だった。

そこで重複を含んだ会話音声から学習したニューラルネットワークによって発話区間を検出する方式を開

## 8 Selective Detail Decodingによる 学習型画像圧縮技術

IoTによる生成データ量は、2025年には約79.4ゼットバイト<sup>※1)</sup>に達すると予想されている<sup>※2)</sup>。これに対して日立は、増え続けるデータを活用できるよう画像データに着目し、深層学習ベースの画像圧縮技術を研究している。

画像には、例えば花や草木のように一枚一枚の葉の



7 従来方式と提案方式の発話区間検出結果の違い





圧縮前

圧縮後 (JPEG)

圧縮後 (本技術)

8 同程度のデータ量に圧縮した場合の画像比較

詳細な構造は重要でないが、テクスチャ(パターンや質感など)が重要な画像内の領域のほか、一部、小さな文字など詳細な構造が重要な画像内の領域が存在すると考えられる。そこで、視覚的画質(人が視覚的に評価する画質)への特化を狙い、画像内の領域に応じて圧縮・伸長処理を最適化するSelective Detail Decoding技術を開発した。これにより、極めて少ないデータ量の条件下において、視覚的画質を向上させる。

今後、日立は研究を加速し、大量データを活用するIoTソリューションへの応用を検討していく。

なお、本成果は東京大学大学院情報理工学系研究科電子情報学専攻相澤研究室との共同研究の結果得られたものである。また、国立研究開発法人産業技術総合研究所のAI橋渡しクラウド(ABCI: AI Bridging Cloud Infrastructure)を利用し得られたものである。

※1) ゼットは10の21乗

※2) 出典:

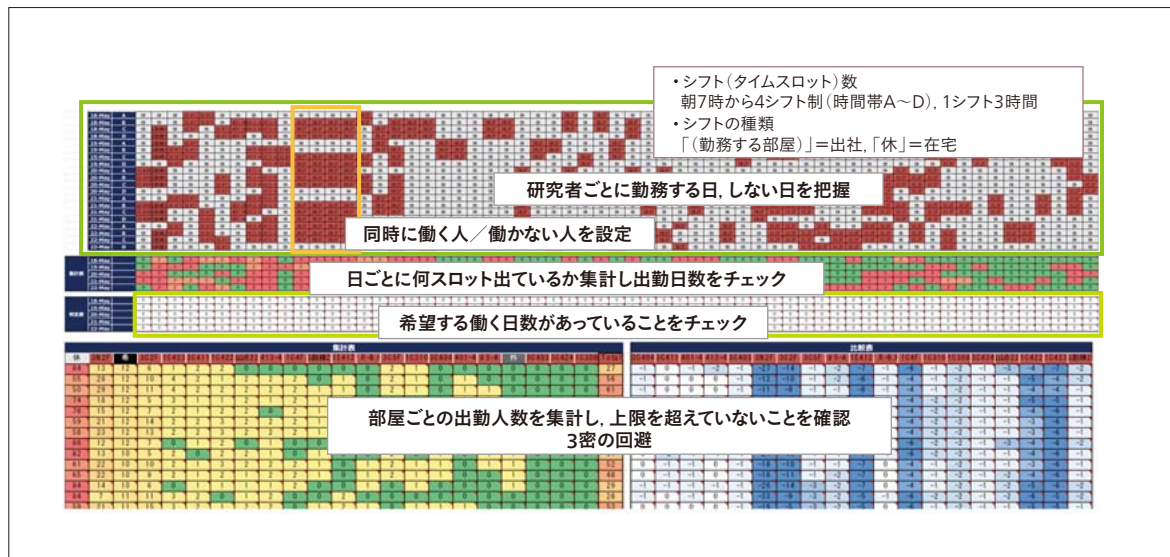
“The Growth in Connected IoT Devices Is Expected to Generate 79.4ZB of Data in 2025, According to a New IDC Forecast”, IDC,

<https://www.idc.com/getdoc.jsp?containerId=prUS45213219>

## 9 新型コロナウイルス時代に対応したCMOSアニーリングの事業化推進

大規模なスケジューリング最適化やポートフォリオ最適化などを実現するために、組合せ最適化問題の実用解を高速に探索するCMOS (Complementary Metal Oxide Semiconductor) アニーリング技術の開発を行っている。

日立の研究開発グループでは、新型コロナウイルスによる緊急事態宣言下の2020年5月から、中央研究所に所属する約360名の研究者を対象にシフト出社制を導入し、このシフト作成にCMOSアニーリングに



9 CMOSアニーリングを用いて作成したシフト表

よる高速シフト作成ソリューションを活用した。研究活動は特別な設備を使用する必要があるため、出勤時間を4シフト制とし、所属チームや研究内容、実験の進捗、実験設備の使用状況、希望出勤頻度、通勤時間など条件を細かく設定することで、数百人規模の3密回避のシフトを効率的に作成し、新型コロナウイルスへの感染リスクを最小限に抑えながら、円滑に研究活動を継続することができた。

このCMOSアニーリングを用いた高速シフト作成ソリューションは2020年下期よりサービスを開始した。

## 10 IT運用最適化サービスにおけるAI活用IT運用技術

ITシステムの管理コスト削減を目的として、AIや機械学習をシステム運用管理へ適用し、人の作業をAIで支援するAIOpsへの投資が増加している。特に工数が多いイベント運用とレポート運用はニーズが高い。日立は、IT運用最適化サービスを立ち上げ、それぞれの運用向けに対処引当自動化技術と特異点自動検出技術を開発した。

対処引当自動化技術は、対処ごとに引き当てるメッセージ文が異なる課題へ対応するため、機械学習を用いてITシステムが通知するイベントに対して、単語一致率を類似度として対処ごとに閾値を学習し、新着イベントと対処の類似度を計算して閾値を超えた対処を引き当てる。検証の結果、本開発技術により引当精度95%以上を達成した。

特異点自動検出技術は、従来IT熟練者が手作業で行っていたITシステムの性能異常を抽出する作業を

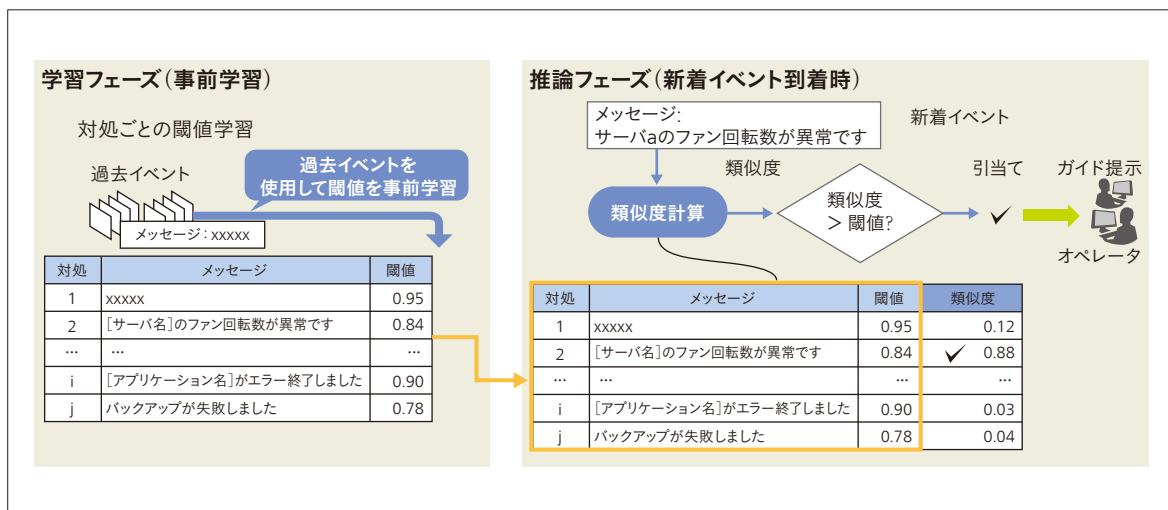
自動化する。機械学習により性能値の時間的パターンを学習し、通常と異なる特異点を自動検出する。ITシステムで頻出する時間的パターンを組み込むことで計算時間を短縮し、従来1週間を要したサーバ数百台規模のITシステムの特異点抽出作業を、数時間程度へ短縮できる。

今後は、対処引当を契機とした対処自動実行やレポート生成の短時間化への取り組みを図る。

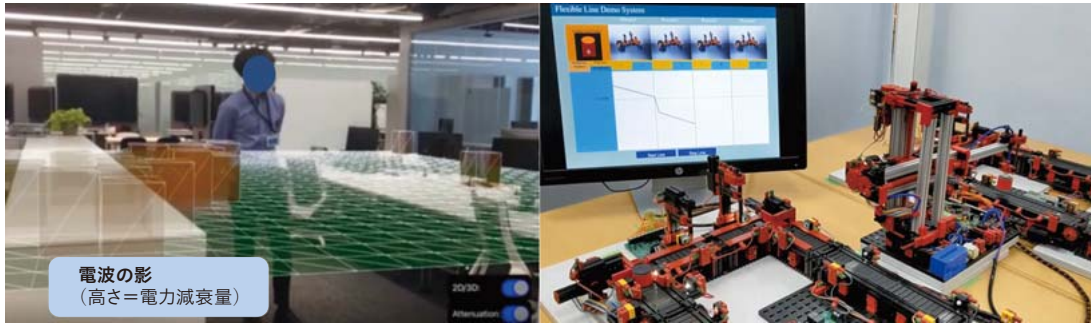
## 11 5Gプラットフォームに向けた5Gエンジニアリング技術

労働者不足をはじめとする社会課題の解決をめざし、5G (5th Generation) を活用したデジタルソリューションの導入と運用を容易にするエッジコンピューティング運用管理技術を開発している。今回、製造現場などの適用に耐える高品質な無線通信環境の迅速な提供と継続的な運用を可能にする5Gエンジニアリング技術を開発した。

現場機器の制御通信や映像通信は、顧客により要求品質が異なり、現場無線環境も絶えず変化する。そこで、GPU (Graphical Processing Unit) 活用と回折解析の散乱波近似計算による高速電波伝搬シミュレータによって、秒オーダーの現場無線環境リアルタイム把握を実現した。また、顧客要件と現場無線環境に合わせた適切なQoS (Quality of Service) 設定による仮想ネットワークスライシングと経路冗長化を特徴とするコネクティビティ技術により、信頼性の高い通信 (パケットロス率 $10^{-6}$ 、遅延50 ms) を迅速に提供可能にした。



10 対処引当自動化技術



高速シミュレータを活用した現場無線環境リアルタイム把握 (AR可視化)      コネクティビティ技術を活用した高信頼無線通信技術による製造ラインの無線化デモ

注：略語説明 AR (Augmented Reality)

## 11 5Gエンジニアリング技術の活用事例

今後、協創の森に構築した5G実証環境を用いてさまざまなアプリ検証を進め、5Gソリューション事業の創生をめざす。

の制約のため、高負荷なAI処理が困難であるのに対し、ディープニューラルネットワーク (DNN: Deep Neural Network) の推論モデルを自動的に軽量・縮約して配備することにより、性能制約のある機器でのリアルタイムAI処理を可能とした。

これらの技術で、従来1時間以上を要していた処理機能の配備を、専門知識なしに1分以内で実施可能とした。今後、さまざまなアプリケーションへの適用を進めていく。

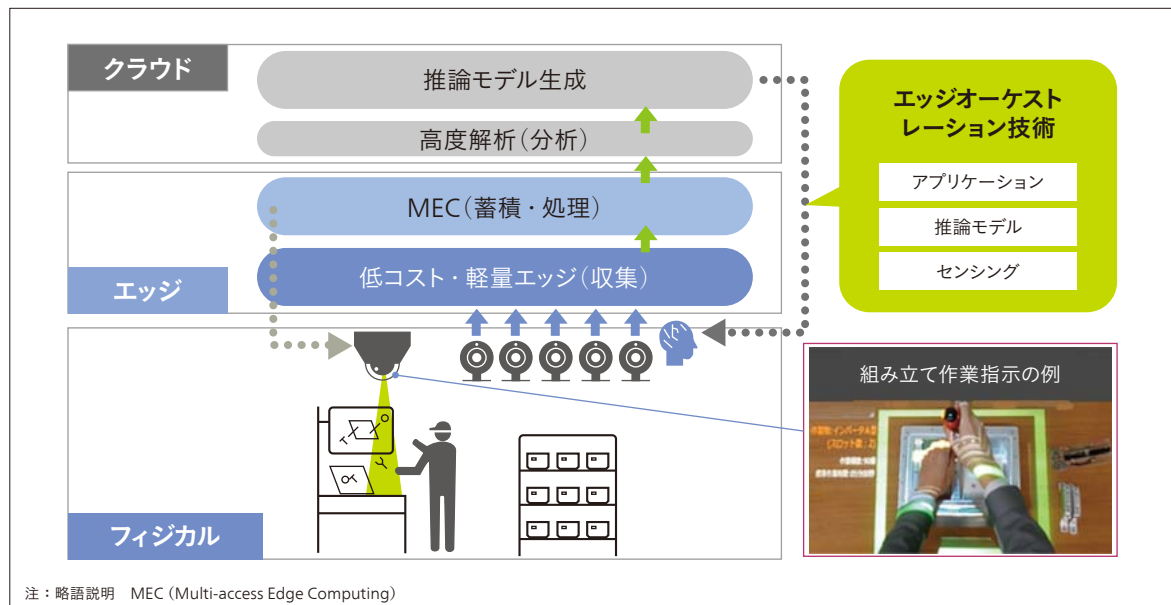
## 12 5G向け エッジコンピューティング技術

労働者不足をはじめとする社会課題の解決をめざし、5Gを活用したデジタルソリューションの導入と運用を容易にする5G向けエッジコンピューティング技術を開発した。

現場環境に設置する機器やネットワークには多くの制約があるため、システム構築に多大な設計工数が必要となる。今回、顧客要件や現場OT環境の制約を把握したうえで、映像による作業状況分析などの機能を遠隔で配備するエッジオーケストレーション技術を開発した。また、現場OT環境では機器の計算性能など

## 13 保全文章の機械的な解析と可視化を実現するナレッジ活用ソリューション

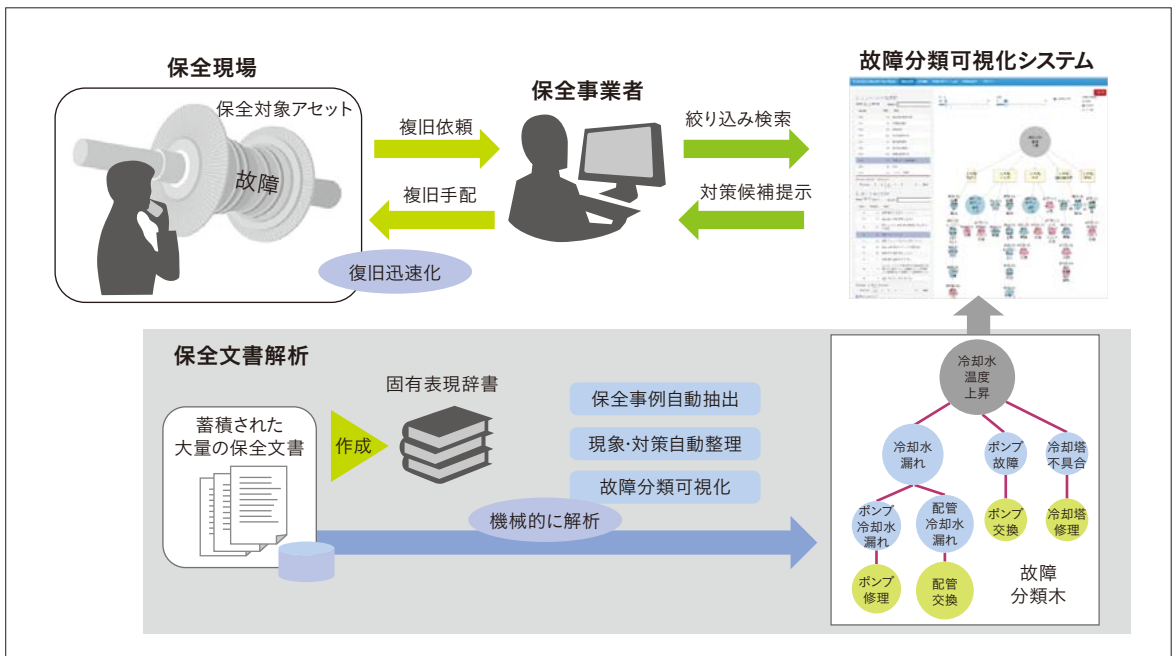
電力アセットなどの保全の現場では、作業内容を記録した保全文書を作成・蓄積し、ナレッジとして故障発生時における故障対策の立案業務などに活用している。しかし、これらの保全文書はいわゆるダークデー



注：略語説明 MEC (Multi-access Edge Computing)

## 12 5G向けエッジコンピューティング技術の概要





13 故障分類可視化システムを用いたアセット保守のイメージ

タと呼ばれる解析が困難な非構造データであり、人手で一件ずつ精読する必要があり、活用の手間やコストが問題となっていた。

そこでナレッジ活用ソリューションとして、保全文書を機械的に解析して故障の現象と対策を抽出し、ツリー状に整理した故障分類木を利用者に提示することで、故障復旧の迅速化を支援する故障分類可視化システムを開発した。

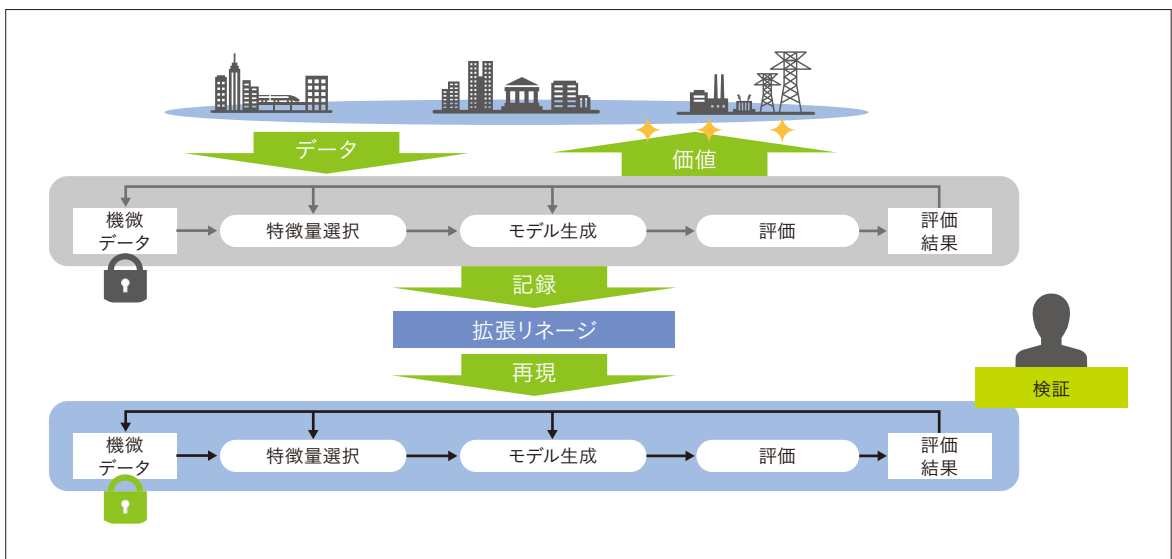
本システムでは、保全文書に固有の表現を日立のスマートディクショナリ基盤<sup>※</sup>で抽出し、保全事例の抽出、現象と対策の関係性分析、故障分類木による可視化を行う。故障分類木は、頂上にある検索クエリの故障現象（灰色）に対して、過去の事例で同時に生じた故

障現象（青色）とその対策（黄緑色）を整理し、その分類を分岐で表現して可視化したものであり、大量の保全文書の内容を横断的かつ迅速に把握可能とする。

※) スマートディクショナリ基盤は「15.ダークデータから意味のある情報を抽出するスマートディクショナリ基盤」(129ページ)を参照

## 14 機微データの安全な分析を実現する拡張リネージによるE2Eデータ管理技術

近年、多様なデータからAIなどの分析技術を用いて新たな価値を創出し、その価値を提供するデータ駆動型社会に期待が集まっている。例えば、公共分野では



14 E2Eデータ管理技術の概要



よりよい健康と福祉の実現に向けて、患者の個人情報などを分析することで、医療介護などの社会保障サービスを拡充する取り組みが進んでいる。

しかし、個人情報などの機微データはどのように加工してデータを利用したかという分析プロセスの透明性が必要であり、これまで人が膨大な時間をかけて試行錯誤を含む分析プロセスを調査して検証する必要があった。

これに対し、データの関係性に加えて特徴量選択やモデル生成、評価にわたる一連の分析プロセスを拡張リネージとして記録して、分析プロセスを再現可能なE2E(End to End)データ管理技術を開発した。今後、社会保障や交通などの分析プロセスに本技術を適用し、公共サービスの質向上による安全・安心な社会の実現に貢献していく。

るニーズが高まっており、マテリアルズインフォマティクス(MI: Materials Informatics)におけるニーズはその典型例である。

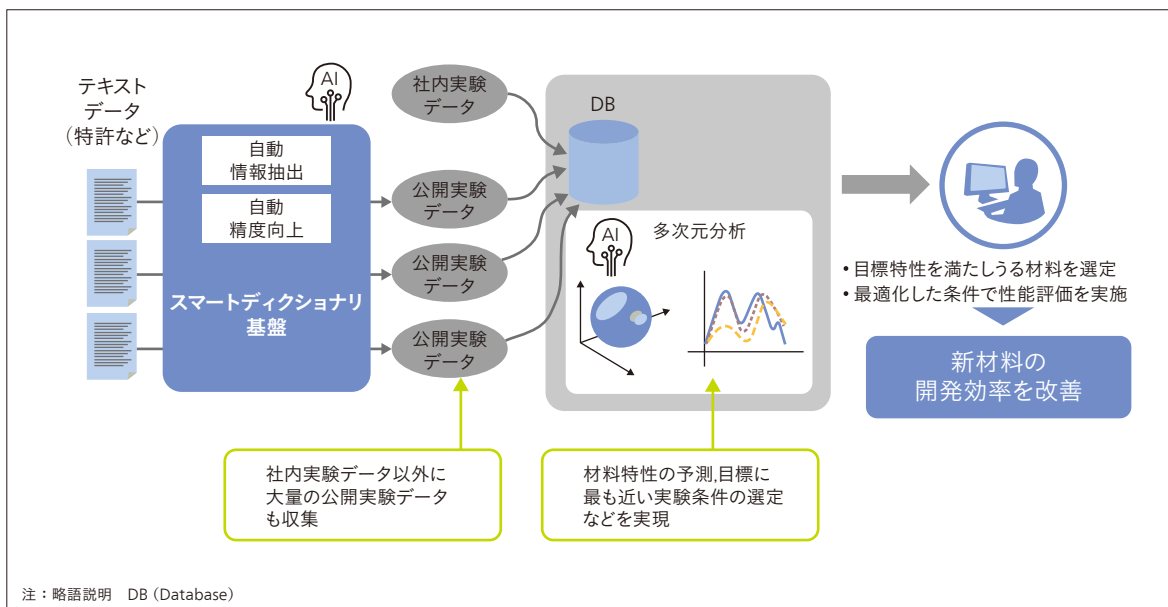
MIでは、実験データを大量に蓄積・分析することで、目標特性を満たしうる材料の選定や最適化した条件での性能評価を可能にし、新材料の開発効率を改善する。実験データ量が多いほど改善効果が高いため、特許などのテキストデータからも実験データを収集する要望が高まっており、高精度な情報抽出手法の確立が急務になっている。

機械学習を用いたテキストデータからの自動情報抽出技術は学術レベルで提案されてきているが、精度向上に要する工数が大きいことが課題であった。スマートディクショナリ基盤技術では、自動情報抽出機能だけでなく、精度低下要因の特定と補正をAIにより自動実行する自動精度向上機能を備えることで、実ビジネスへの適用を可能にした。

今後はさまざまなパートナーとの実証実験を通じて本基盤の有効性を証明するとともに、MIだけでなく保守支援サービスなどの幅広い分野における新事業開拓に貢献していく。

## 15 ダークデータから 意味のある情報を抽出する スマートディクショナリ基盤技術

テキストデータなどの従来活用されていなかったデータ(ダークデータ)から業務に役立つ知識を抽出す



15 MIへのスマートディクショナリ基盤の適用イメージ