

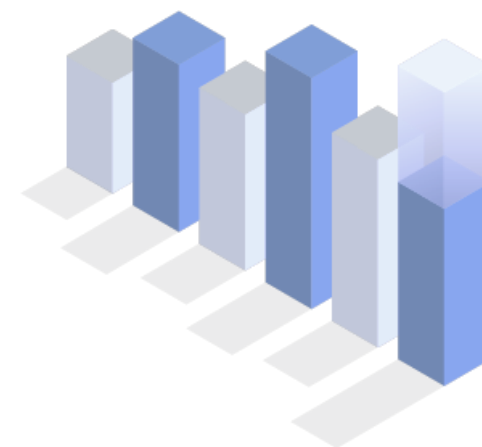
AI OCRとデータプラットフォーム連携で加速する 現場データの活用アプローチ

株式会社DATAFLUCT 東芝デジタルソリューションズ株式会社



Index

- 01 | 本資料の趣旨
- 02 | 日本企業におけるこれまでのAI OCR 活用実態
- 03 | 各業種におけるAI OCRの導入検討状況と課題
- 04 | 業種別の課題とニーズ
- 05 | 東芝デジタルソリューションズ×DATAFLUCTのアプローチ
- 06 | AI OCR×AirLakeによる課題解決
- 07 | ソリューションの全体像
- 08 | AI OCR概要 / AirLake概要
- 09 | ソリューション利用イメージ
- 10 | ユースケース（製造・金融・医療・社会インフラ・保険・介護）
- 11 | 最後に / お問い合わせ
- 12 | 会社概要



本資料の趣旨

Introduction

OCR技術は、AIの進化とともに読み取り精度が向上し「データ化」の課題を克服する一方で、さらなる進化が求められています。

本資料では、AI OCRによるデータ化に留まらない、現場データの有効活用を加速させる方法として、データプラットフォーム連携という新しいソリューションをご提案します。

AI OCRによる読み取りや活用が難しかったフォーマットのデータにも焦点を当て、本ソリューションを用いた新たな活用アプローチを解説します。

最後に多業界の活用事例を通じ、企業がどのようにして効果的にデータを活用し、業務の最適化とビジネス成長を実現できるかをご紹介します。

日本企業におけるこれまでのAI OCR 活用実態



STEP 1

OCRによる文字認識

OCR（光学文字認識）技術の進歩により、**印刷文字・手書き文字を高精度で認識可能に**

RPA（ロボティック・プロセス・オートメーション）を組み合わせることで、**データ処理の自動化と業務プロセスの効率化**を大幅に向上



STEP 2

加工・分析

読み取ったテキストに対して、**特徴抽出や意味付け**（カテゴリ分類）を行うことで、集計やレポーティングが可能に



STEP 3

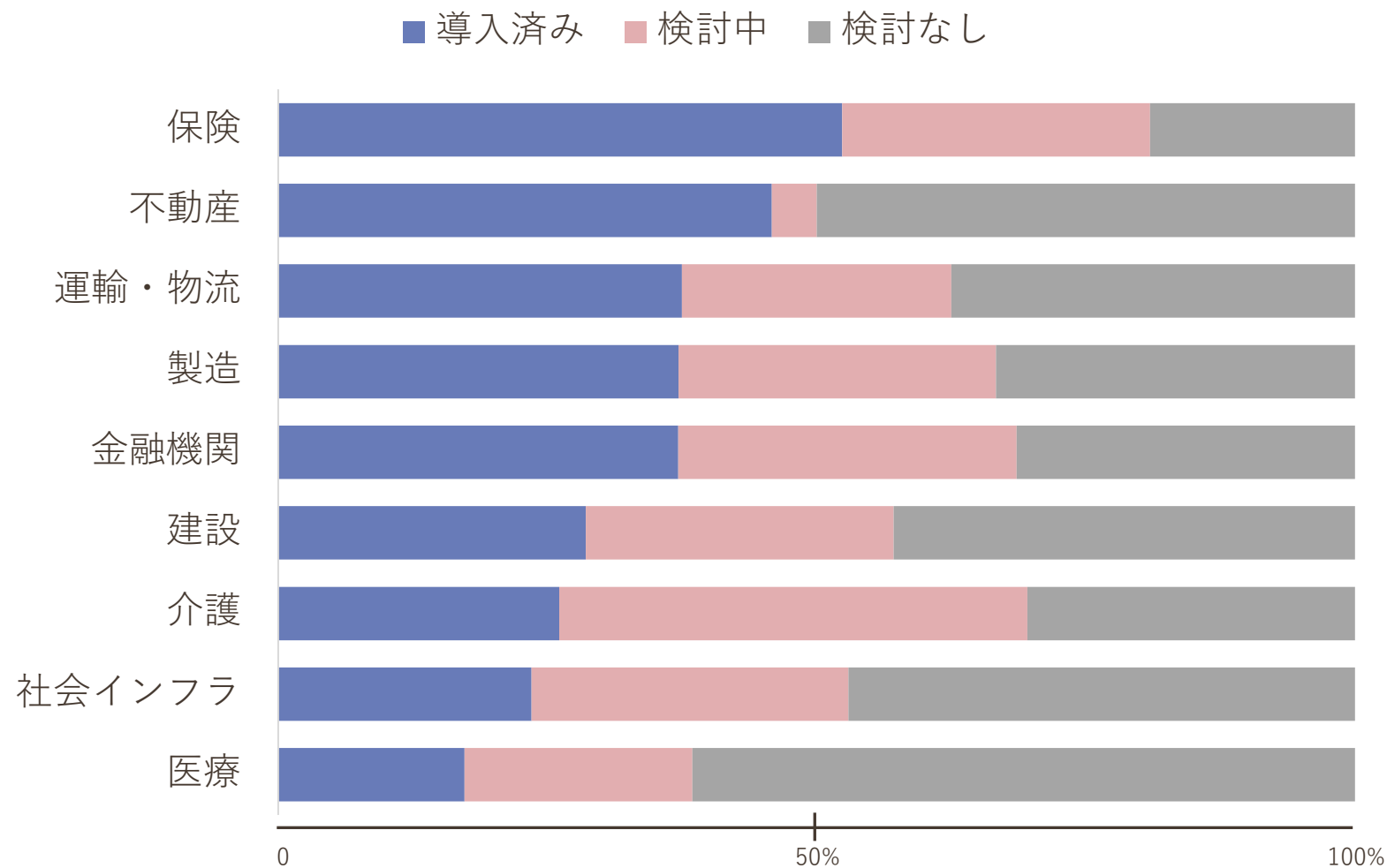
検索・提案

データが構造化されることで、**社内ナレッジを検索可能に**

分析結果を元にした**業務改善の提案**や自動通知が可能に

近年ではDXの実現が求められる中で、文字認識による書類のデータ化だけでなく、**新たなデータ活用やプロセスの抜本的な再設計**が求められる

各業種におけるAI OCRの導入検討状況と課題



これまで紙を中心とした保険業、不動産業を筆頭に約半数の企業がAI OCRを導入済み。

一方、介護・社会インフラ・医療では、導入が進んでいるのは30%にも満たず、まだまだOCRを含むデータ活用に取り組めていない実態がうかがえます。

※TDSLによるエキスパートインタビュー調査結果。

OCRを含むデータ活用ニーズのある企業がヒアリング対象。(ただし、単純な帳票の読み取りのみを希望する企業は除外。)

業種別の課題とニーズ

OCRのニーズが大きい業種は多いが、業種特有のフォーマットの読み取り及び、その先のデータ活用にはまだ至っていない企業が多いことが明らかになっています。

業種	OCRニーズ	業種全体のOCR導入概況	課題とニーズ
製造	大	×	紙の製造記録や設計資料の電子化のニーズが高い。
社会インフラ	中	×	紙の製造記録や設計資料の電子化のニーズが高い。インフラの施設利用は長期的になる場合が多いが、ナレッジが蓄積されておらず、効率的な施設運用ができていない。
金融	大	△	財務諸表の読み取りなど審査業務支援にニーズが高い。 海外では財務文書の読み取り事例あり。
保険	大	○	定型読み取りではすでにOCRが浸透。 健康診断書など非定型読み取りに課題あり。自社システム連携のニーズが高い。
医療	大	×	紙で流通する紹介状、検査結果、お薬手帳の読み取りにニーズあり。ICTに不慣れな従事者が多く、OCRシステム導入自体に課題あり。
介護	大	×	人手不足が深刻で、介護記録や診療情報の読み取りにニーズあり。 高齢のスタッフが多く、使いやすさが重要。
不動産・建設	中	△	図面やCADデータなどの不動産・建設書類の読み取り、データ活用ニーズが高い。
運輸・物流	中	△	カメラやハンディスキャナでの伝票読み取りでOCRが浸透。しかし、全てのフォーマットに対応できておらず、ドライバーに負荷が掛かっているのが課題。

※TDSL調べ。

OCRを含むデータ活用ニーズのある企業がヒアリング対象。(ただし、単純な帳票の読み取りのみを希望する企業は除外。)

東芝デジタルソリューションズ×DATAFLUCTのアプローチ

TOSHIBA

- 製造業として長年にわたり培ってきた幅広い事業領域の知見や実績
- ものづくりで磨いた高い技術力
(文字認識・画像解析・音声解析)



DATAFLUCT

- データに関するフルスタックのテクノロジー活用能力
- 新規事業開発ノウハウ及び、
様々な会社のDX支援実績

東芝デジタルソリューションズ株式会社(以降、TDSL)のこれまで培った知見と最先端のAI OCR技術、株式会社DATAFLUCTの様々な業種企業へのDX支援実績及び、データ基盤技術を融合させ、データ化の先の現場業務DXを実現するソリューションを提供します。

AI OCR × AirLakeによる課題解決：連携時の価値

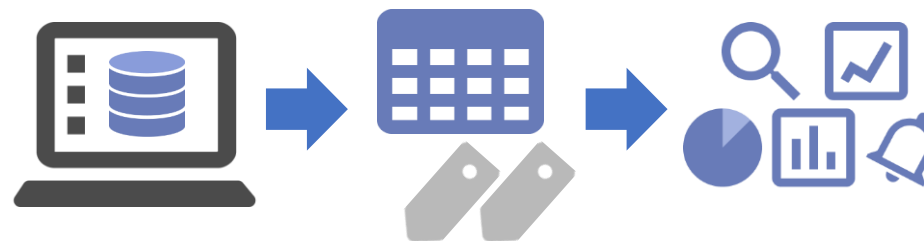
TOSHIBA

高度なAI OCR技術で、定型読み取りだけでなく非定形の読み取りも可能に



DATAFLUCT

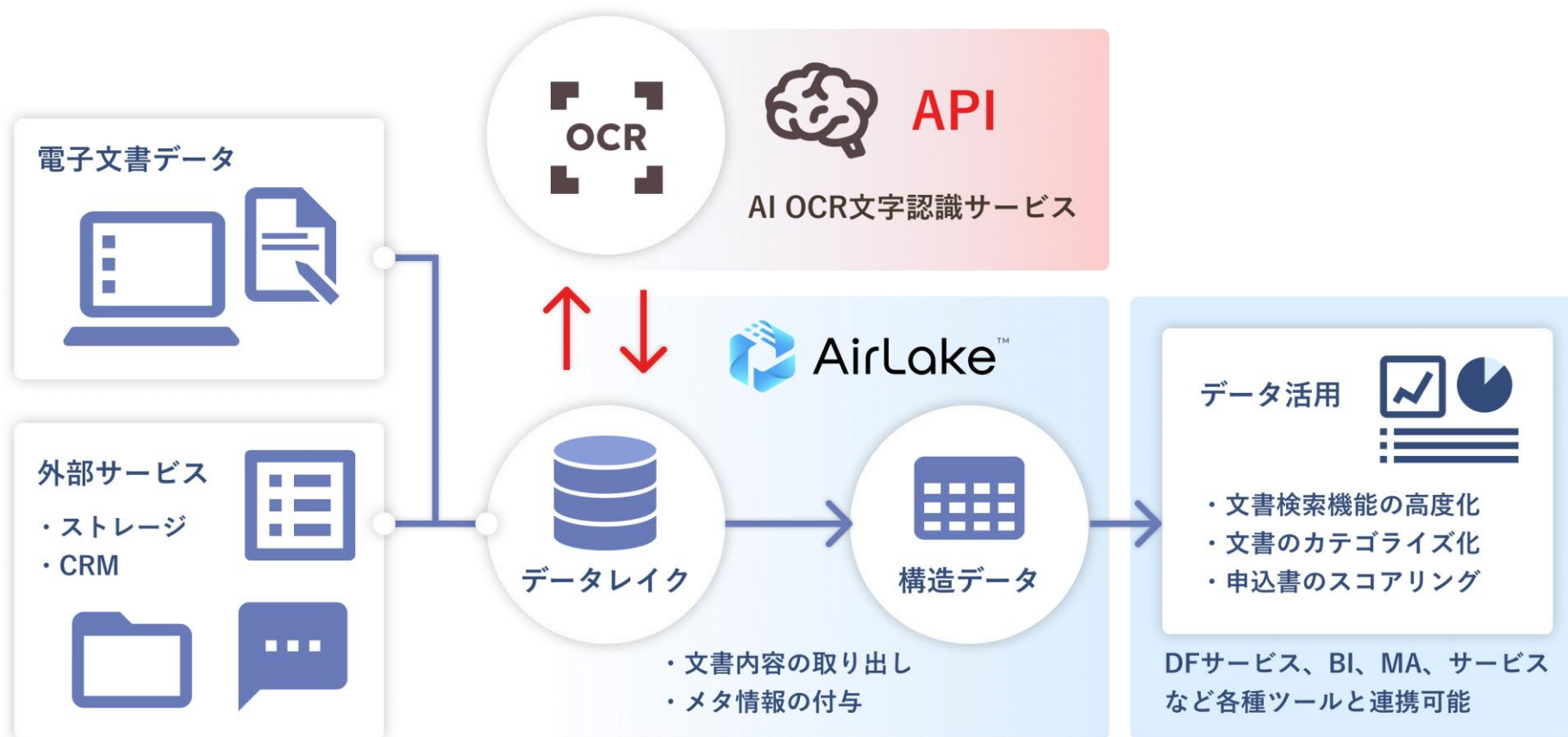
高精度なデータの意味付けや
既存システム連携、BI連携の実現



TDSLの「AI OCR文字認識サービス」（以降、AI OCR）とDATAFLUCTのAirLakeを連携することで、非定形の読み取りから、読み取り後のデータ活用を見据えたタグ付与処理や検索、可視化、AI活用までワンストップで提供が可能になります。

ソリューションの全体像

AI OCRとDATAFLUCTのノーコードで非構造化データを変換するデータ基盤「AirLake」を連携させることで、デジタル化に留まらず、データ活用までシームレスにリーチ可能になります。



AI OCR 概要

業界最高水準の精度の文字認識サービス

～最新のAI技術を活用した認識エンジンで、さまざまな書式の帳票に対して高い認識精度を実現～

Point 1

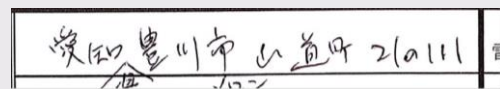
業界最高水準の高い認識精度 ※1

以下の実際の運用で目にする文書に対しても高精度に認識可能

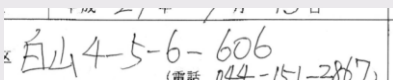
- 「接触文字」「枠線にかかった文字」「つづけ字」など、手書きの文書で見られる様々な文字
- 取り消し線や訂正印の読み飛ばし
- イメージスキャンの際の画像劣化



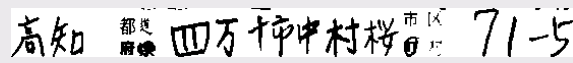
→神奈川県川崎市幸区堀川町72-34



→愛知県豊川市山道町21の111



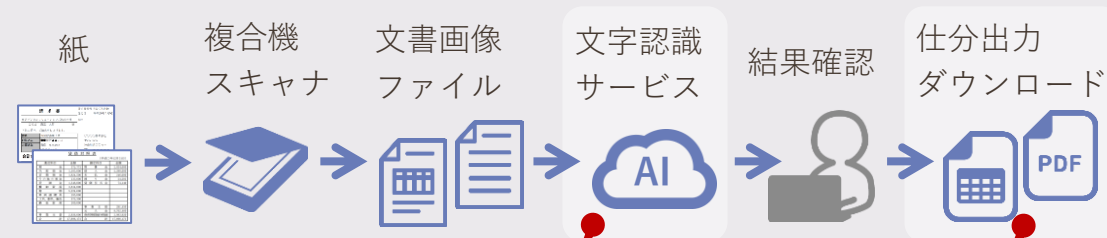
→白山4-5-6-606



→高知県四万十市中村桜町71-5

Point 2

業務用途に応じた4つの読取モードと帳票識別



業務用途に応じた4つのモードを利用可能

1. 読取範囲を指定した「定型読取」
2. キーワード指定による「非定型読取」
3. 請求書専用の「プリセット読取」
4. 文書全体を読み取る「全テキスト読取」

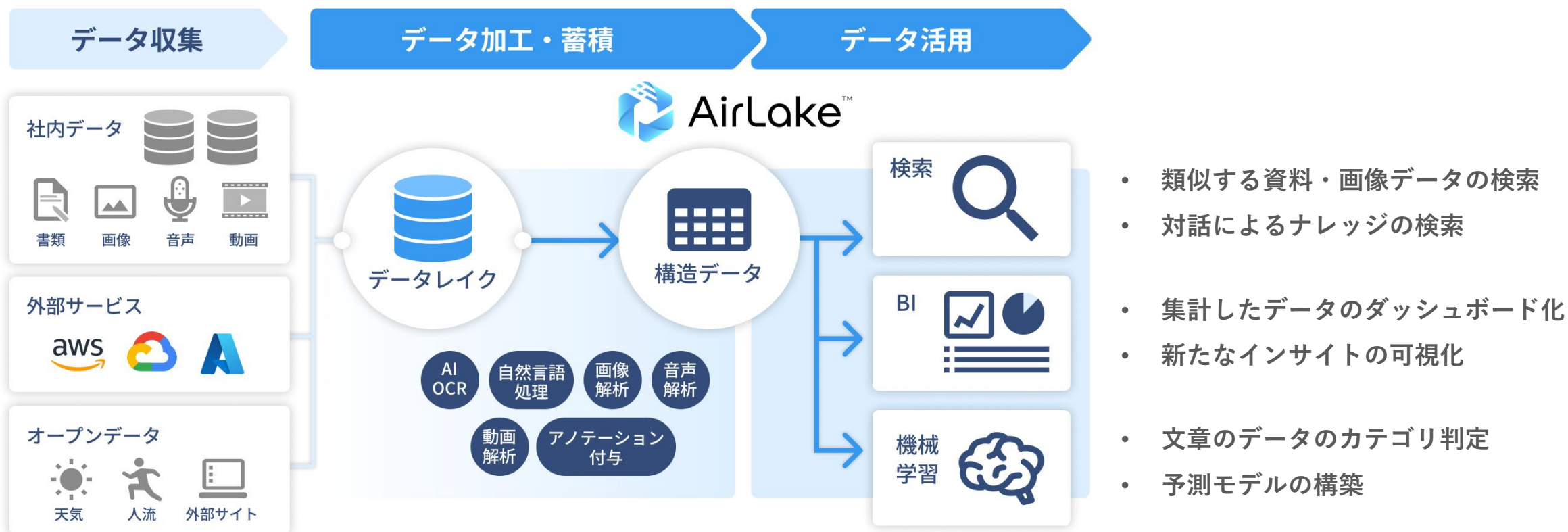
事前仕分が不要となる「帳票識別機能」を、高精度文字認識と組み合わせることで、更なる業務効率化を実現

※1：P6813704（座標推定）、P7043670（近接文字抽出）、P2022-121020（重複読み、読み飛ばしの抑制）など多数の特許、複数の国際学会発表実績といった先進的な技術を保有

AirLake 概要

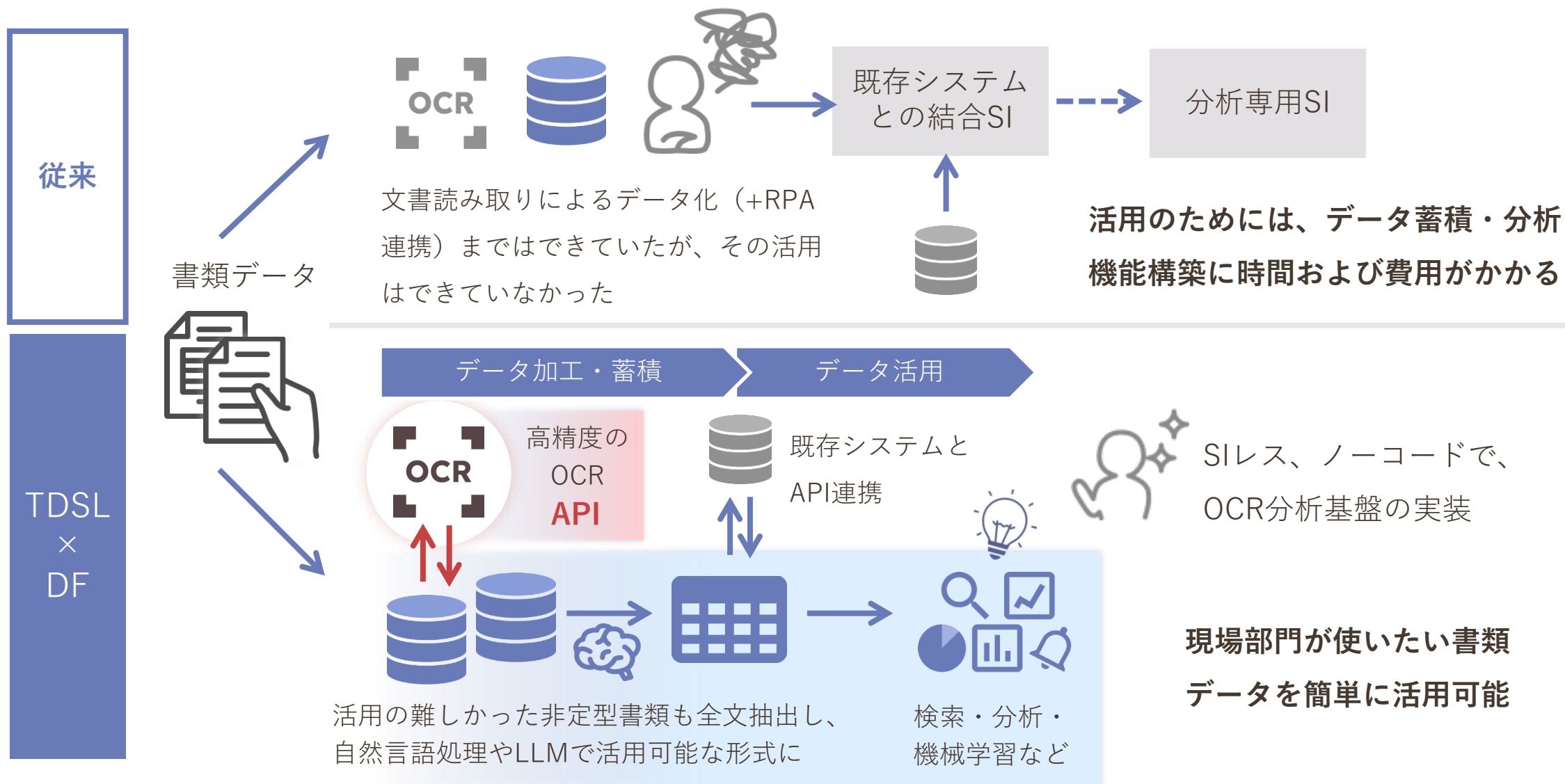
マルチモーダルなデータの取扱いを容易に。

データ活用に必要なプロセスをプラットフォームサービスとしてトータル提供



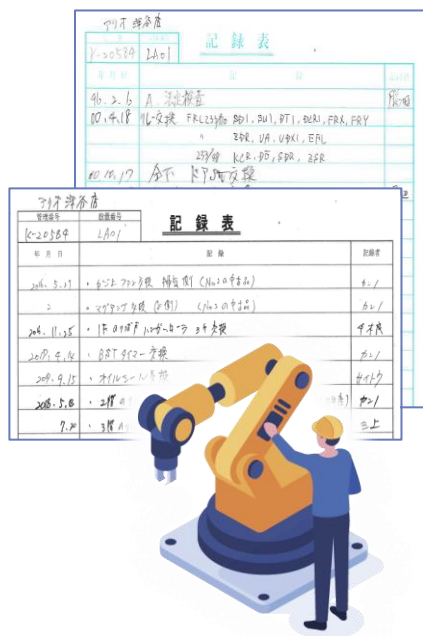
収集したデータを加工・蓄積し、活用するまでの業務を効率化

ソリューション利用イメージ



ユースケース①：製造

紙でしか存在しない手書きの点検表・記録表をデータ化し分析



Point

- 人によって表現が異なる手書きへの柔軟な対応
- 取り消し線や訂正印、複数行記載などの様々な記載方法への対応力

ソリューション

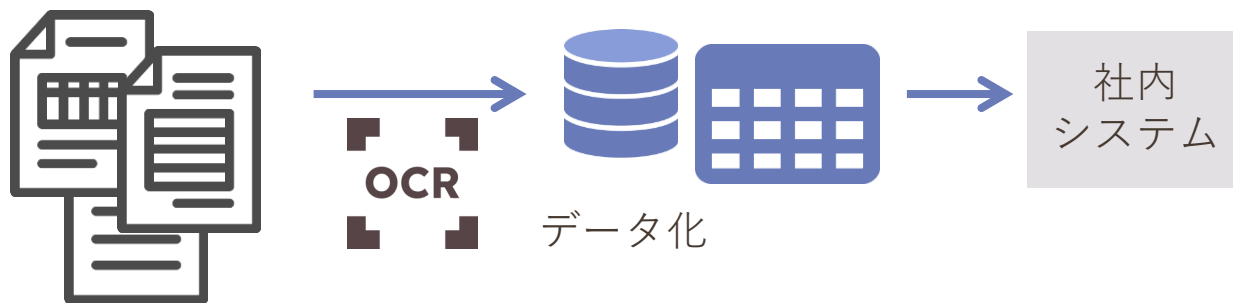
傾向分析を行い、よりプロアクティブな対応施策を実現

- 製品毎の保守対応回数や地域、利用場所特性の把握
- 保守部品の確保
- 保守人員の最適配置
- 品質向上、顧客満足度向上の施策を実現



ユースケース②：金融

複数社の取引で扱う非定型帳票(申込書、与信審査書など)を
データ化し社内システムへ連携



フォーマットの
異なる帳票ファイル



Point

- 項目定義の容易さ
- 非定型帳票にありがちな項目名の表記揺れの吸収性の高さ
- 社内システムへの自動反映等、後続業務への連携性の高さ

ソリューション

データ化にかかる時間の削減

各社で異なる帳票フォーマットの表記揺れは人力で吸収しており、数分～10分程度の工数がかかっていた



読み取る項目を指定するだけで、
数秒でシステム登録し、集計作業を実施可能

顧客満足度向上、データドリブンな提案



- 与信審査期間の短縮
- 投資提案のための分析

ユースケース③：医療

手書きの問診表・紹介状・診察情報を電子カルテに自動
で読み込むことで、ミスや待ち時間を低減



問診表・紹介状・
診察情報など



Point

- ・ 接触文字や読みづらい文字であっても、内容を正確に読み取ることのできる精度の高さ
- ・ 複数行記入や取り消し線など、手書きならではの記載方法への柔軟な対応

ソリューション

文字の読み取り精度向上

書き写しミスが許されないため、データ入力に時間がかかっており、OCR商品では精度上導入に踏み切れなかった

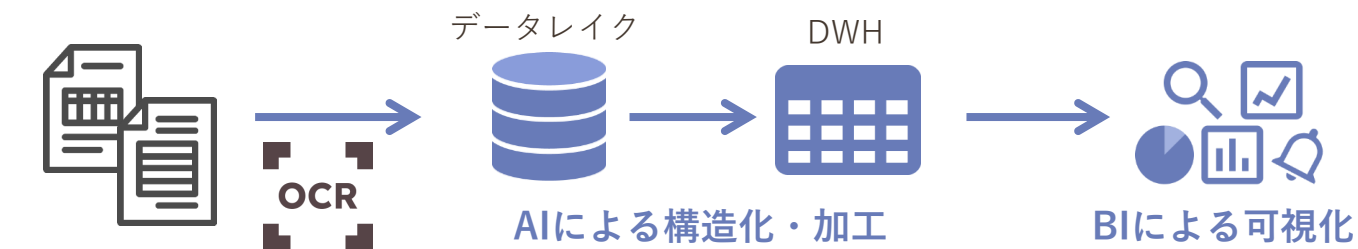
- ➡ データ入力時間が減り、患者さんの待ち時間を削減
- ➡ 読み取ったデータを連携するだけで電子カルテ化が可能のため、ICTに不慣れな医療現場においても簡単にデータ入力が可能に

過去症例や服薬データの分析

- ➡ 病気の原因予測や服薬指導のサポートなど、新人とベテランの経験値を補完

ユースケース④：社会インフラ

施設情報・工事記録・点検情報などのインフラや大規模施設に関する情報を一元管理し、データドリブンなインフラ運用を実現



設計資料
災害関連資料



Point

- 手書き書類やExcelなど、様々な形式のデータを構造化
- 全文抽出の中から、顧客企業にとってノウハウとなる内容を抽出するAI(自然言語処理モデル)を構築して運用

ソリューション

ノウハウを抽出し属人化解消へ

多くの書類が手書きやExcelによる管理で、データアクセスや分析が煩雑かつ属人化され、経験則にのみ基づいた運用



構造化しデータを一元管理することで、ユーザーが直接見たい情報を見たい時に閲覧することが可能に



ノウハウを抽出して検索でき、ナレッジの共有が可能に

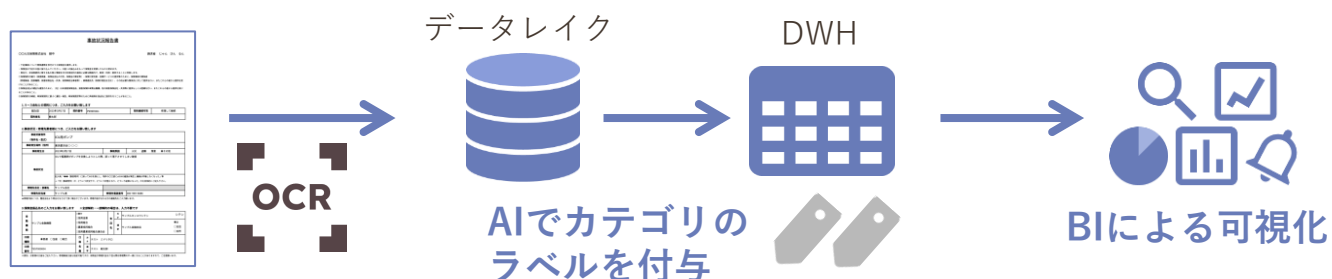
例)

- 安全基盤情報の推計
- 設計積算時の最適案の算出



ユースケース⑤：保険

BtoBの保険契約において、事故発生時に作成される事故報告書から、事故の原因把握や予防施策の立案を支援



事故報告書
(手書きスキャンPDF)



Point

「事故状況欄」に記載された自由記述テキストに対して自然言語処理モデルを活用し、原因・故障箇所・事故状況などのカテゴリを付与し、より詳細な解析を可能に

カテゴリごとの集計

単語同士の繋がりを可視化

ソリューション

データドリブンの事故対策立案

手書きの事故報告書からは、保険金支払いのために必要な情報以外のデータは蓄積されておらず、事故原因の傾向の把握や予防施策が立案できない状況



AI OCRを利用し、事故報告書の項目をデータ化し可視化することで、事故の原因や傾向を詳細に把握し、データに基づいた施策を立案することが可能に



ユースケース⑥：介護

介護記録や診療情報の紙データの分析を行うことで、食事や運動、薬の投与などのケアプランの作成をAIがサポート



介護記録や診療情報



Point

- 現場で記載を行う読みづらい手書き文字でも、正確に読み取ることのできる精度の高さ
- 複数行記入や取り消し線など、手書きならではの記載方法への柔軟な対応

ソリューション

データ化によるケアプラン作成へのAI活用

- 紙データでは、夜勤日勤の引継ぎや過去の状態確認が大変だった
- ICTに不慣れな従業員が多く、OCR導入検討は難しかった



手書きの特性含めて読み取りが可能のため、修正作業も少なく簡単に利用することが可能



過去のデータを分析して、個人の症状に合わせたケアプランなどをAIが作成支援できるようになり、引継ぎや新人教育がスムーズに

最後に

多様化し続ける現場環境や社会情勢を捉え、現場でのデータ活用を促進するには、データを集約・構造化・分析し、データドリブンで意思決定ができる環境・文化をつくる必要があります。

東芝デジタルソリューションズが培ってきた読取精度の高いAI OCRとDATAFLUCTが提供するマルチモーダルデータ基盤の組み合わせは、データ基盤作りから高度なデータ活用までのライフサイクル全体をサポートできる強力なソリューションです。データ活用による様々な課題解決に挑戦する皆さまを、全力で支援します。

お問い合わせ

株式会社DATAFLUCT

Mail : info@datafluct.com

<https://datafluct.com/contact/>



東芝デジタルソリューションズ株式会社

Mail : tdsl-OCR-cloud@ml.toshiba.co.jp

<https://www.global.toshiba/jp/products-solutions/ai-iot/mojigazou.html>

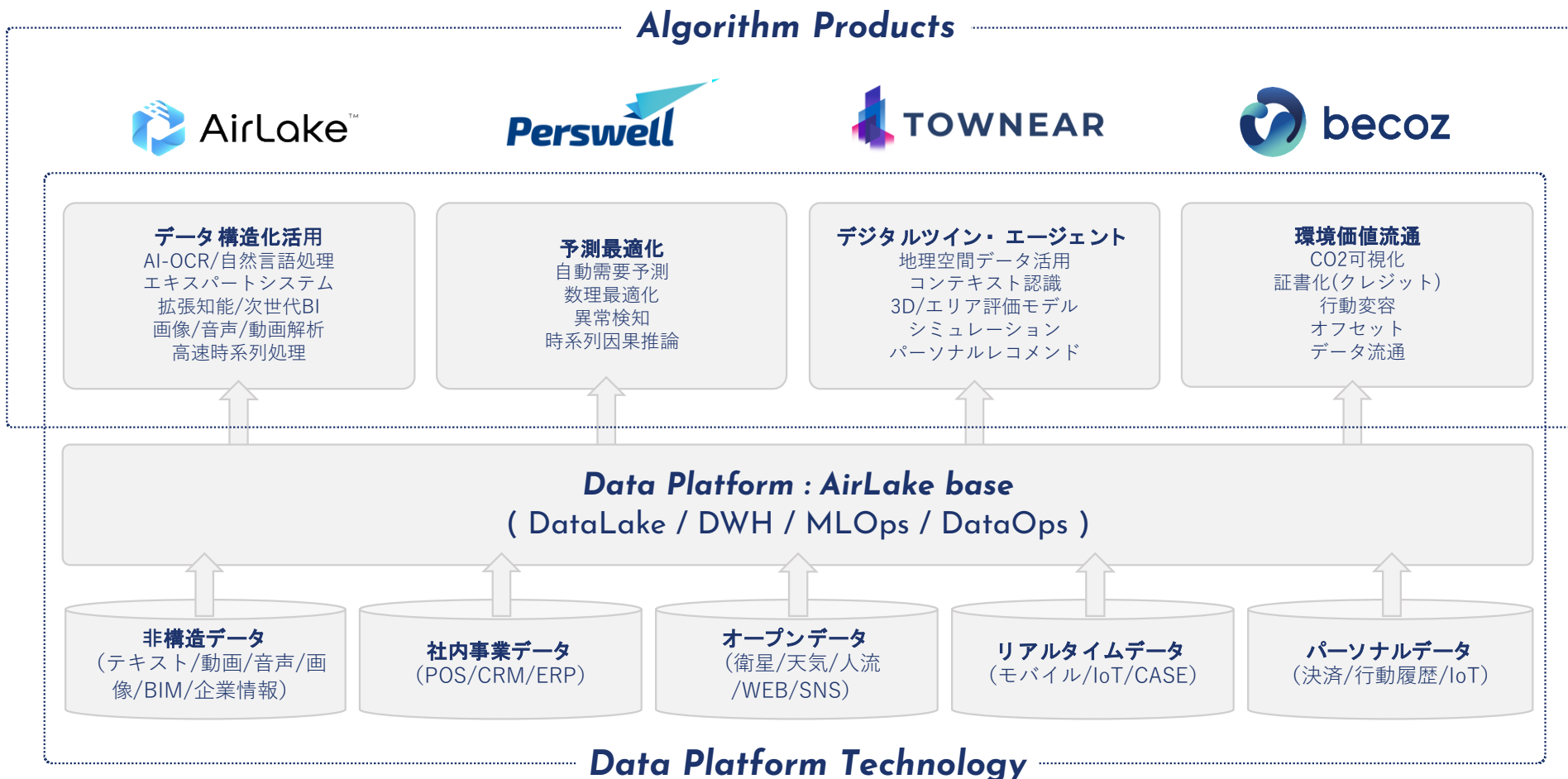


会社概要



DATAFLUCTは、「データを商いに」というビジョンのもと、データから新たなビジネスや価値を生み出すビジネスパートナーとして、さまざまな産業にデータ活用を届けています。幅広い分析実績をもち、技術やデータに拘らず業界をこえて

持続可能なアルゴリズムの社会実装を経営理念としています。多彩な自社サービス開発の知見を生かしたDX支援のほか、データ活用でSDGsに貢献することを目指し、ビジネスと社会貢献を両立させる新規事業開発にも取り組んでいます。



会社概要

TOSHIBA 東芝デジタルソリューションズ株式会社

東芝グループは、製造業として長年にわたり培ってきた幅広い事業領域の知見や実績と情報処理やデジタル・AI技術の強みを融合し、サイバーフィジカルシステムテクノロジー企業として成長することを目指します。
東芝デジタルソリューションズ株式会社は、



デジタルソリューション事業領域の中核企業としてデジタルトランスフォーメーションを推進し、各業種向けシステムインテグレーション、共通業務ソリューションおよびAI・IoTを活用したサービスソリューションを提供します。

