

# 技術文書から知見を引き出し 企業価値に変えるために

IBM Watson Explorerが実現するテクニカルな知見の共有基盤



## はじめに

### 従来にない幅広い技術への対応力が求められる時代

製造業をはじめとする、企業における専門的な技術を扱う部門は今、共通して技術者の世代交代を迎える中で、退職していくベテラン技術者のテクニックや知見を若い世代に継承していくことが急務となっています。ベテランがいた時は経験の浅い技術者は支援を受けることができましたが、いなくなってしまうと頼れるものがない一方で、成果は高い水準を求められます。これはフィールド・サポート系の部門などに当てはまる話ですが、そのほかにも技術を扱う部門の例として、イノベーションを求められるR&D部門があります。R&Dにおいては、既存の技術に甘んじることなく新しい技術を取り入れていく必要があります。そのため今の時代のR&Dのエンジニアは、元来の自分の専門領域を超えて、幅広くかつ最新の専門知識を企業内部だけでなく外部から取得しなければいけないという多大な負荷を課せられています。そうした広範な技術情報の習得を、各自の自己啓発や企業内の教育体制のみに頼っていたのでは限界があり、ここでは、エンジニアが真に必要な情報を効率的に収集、探索、分析して活用できる仕組みを整えていく必要があります。

## いかにして技術文書から重要な知識や知見、 ノウハウを抽出するか

ベテランの知識を伝えていかなければいけない課題に対する解決方法の一つに、ベテランの頭の中にある知見をヒアリングしそれを記録するようなことが考えられます。とはいえ、広く伝承・共有化すべき知識や知見、ノウハウは、必ずしも個々の人材の頭の中にしかないわけではありません。まずは、過去にベテランによって現場で作られ蓄積されてきたさまざまな報告書や手順書、ドキュメント、論文といった技術文書から重要な知識や知見、ノウハウを抽出することが、効率的な伝承・共有化を実現する最短の道と言えます。そして、そのための有用な手法を備えているのが、構造化及び非構造化データ(=コンテンツ)を迅速に可視化、マイニングし、業務ナレッジを引き出して企業価値に転換する次世代のコンテンツ探索・分析プラットフォーム、IBM Watson Explorer(以下、Watson Explorer)なのです。Watson Explorerは次の2つのステップで技術文書から知見を引き出します。

### ステップ1 企業内外に散在する文書を横串でキーワード全文検索

活用できていない大量の技術文書が企業内にあるとき、また活用したい公的文書がインターネット上にあるとき、Watson Explorerは非常に簡単な設定でそれらを横串検索できるシステムを迅速に構築します。例えばエアバス社は約10年前からWatson Explorerにより企業内のファイルサーバーやERPシステム、文書管理システム等多種多様な情報源を横串検索して、顧客企業へのクレーム対応を迅速に行い、ビジネス損失を削減しています。Watson Explorerのキーワード全文検索は、高度な自然言語処理技術を用いて検索結果の文書をすべて自動タグ付けします。生成されたタグは次の絞り込みに役立つだけでなく、気づきを与えてくれる効果もあります。

キーワードによる全文検索テクノロジーは、こうして企業内の大量の情報を可視化できるメリットがありますが、業務の専門知識に乏しい新人ユーザーにとっては以下の課題が生じることになります。

●キーワードが一致さえすれば関係のない文書まで大量にヒットしてしまい、その後の絞り込みが容易でなく、結局必要な情報にたどり着けない。

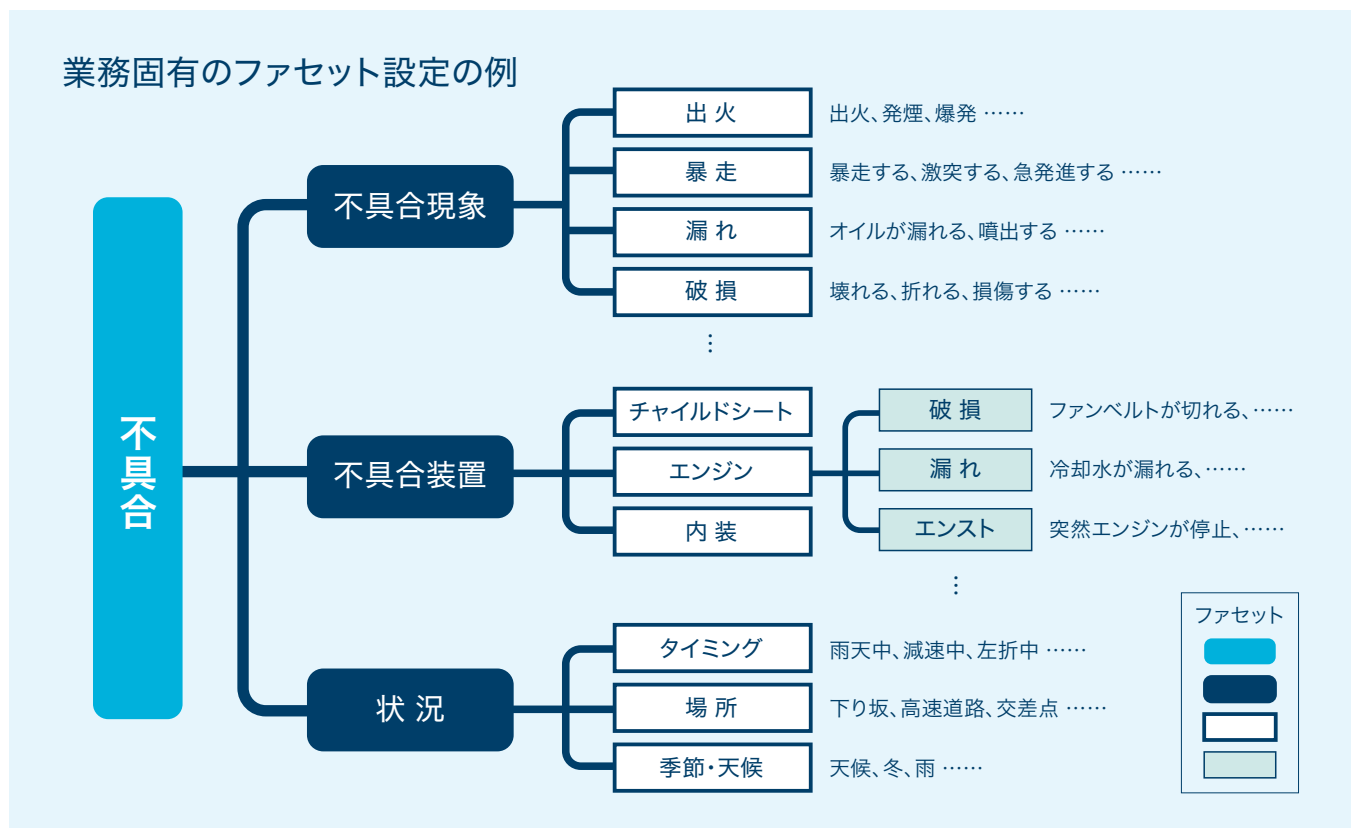
●専門知識に乏しいユーザーはそもそも適切な検索キーワードを思いつかないことも多い。

そこでこの課題に対し、情報探索の高度化を図れるのがWatson Explorerの分類・分析のテクノロジーです。

## ステップ2 知見を発見しネクストアクションにつなげるテキスト分析

### 高度な自然言語処理、ビジネス要件に対応できる文書の自動分類技術

Watson Explorerのテキスト分析機能は、情報探索するユーザーの業務要件に対応できるコーパス(索引)により大量の文書の分類(タグ付け)がなされます。このタグをファセットと呼んでいますが、こうした付随情報があるおかげで、テクニカルタームを思いつかない新人でも知りたい情報を含む文書だけに絞り込まれた検索結果を得ることができるのです。

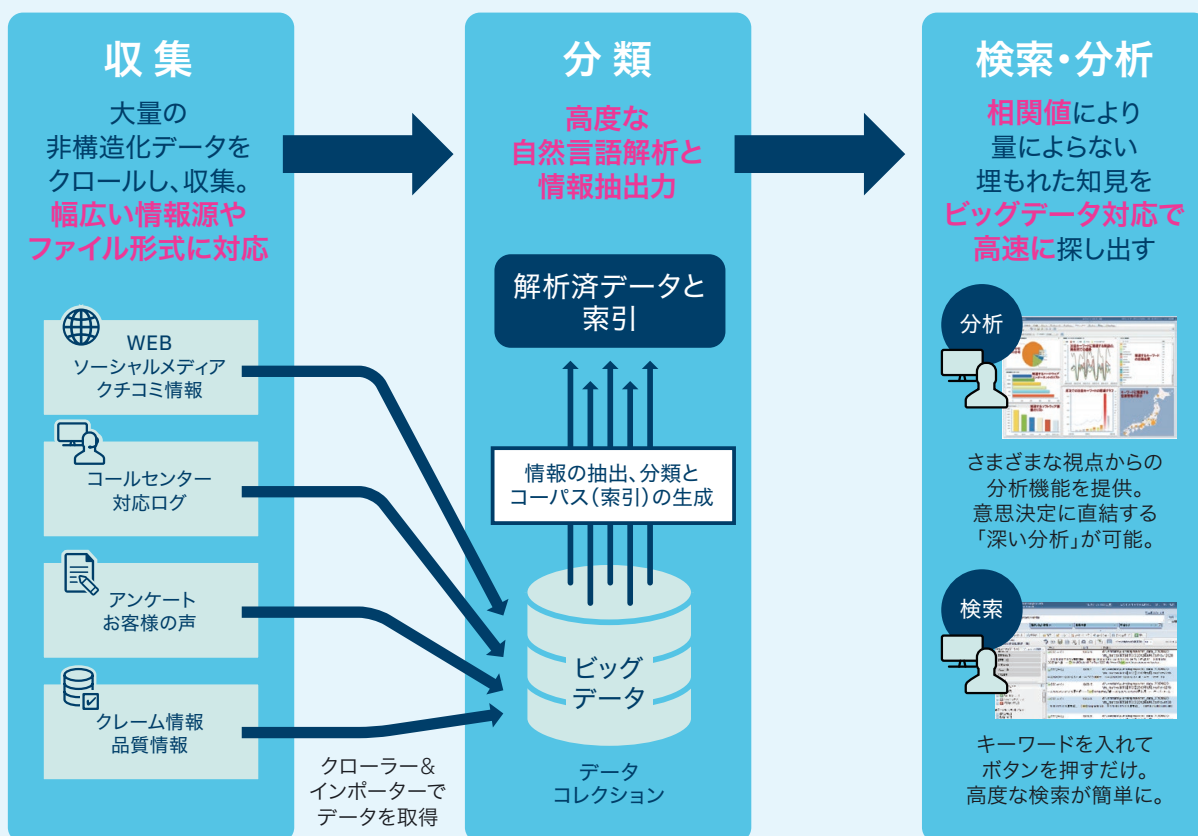




## 「相関値」を用いた 発見型分析アプローチへ

IBMでは、高度な自然言語処理能力を備えたテキスト分析のソリューションとしてWatson Explorerを提供しています。日本IBMの東京基礎研究所で開発され、10年の歳月を経て製品化されたコンテンツ(文書)探索・分析プラットフォームです。Watson Explorerの最大の特長は、「相関値」により特定の条件下に出現が偏っている特徴的なワードを自動で検出することで、それらを含む文書から知見を獲得する、発見型分析手法を提供している点です。それに対して出現頻度による一般的な統計的アプローチでは、こうした気づきにつながる特徴を捉えることは難しく、常にありきたりの事象だけを確認するにとどまります。また、Watson Explorerは、社内外の多様なデータソースから大量の非構造化データをクロール(巡回)して取得する「収集」から、高度な自然言語解析と情報抽出による「分類」、埋もれた知見の高速な「検索・分析」に至る一連のプロセスをトータルでサポート。伝承・共有化すべき知識や知見、ノウハウの獲得を支えます。

### Watson Explorer: 知見の獲得を支えるしくみとテクノロジー





## 機械学習、AI技術の活用

### 文書の自動分類/類似文書検索/AI辞書登録支援

2018年2月に発表されたIBM Watson Explorer V12(正式製品名称:IBM Watson Explorer Deep Analytics Edition)では、「oneWEX」と呼ばれる新たなコンポーネントが加わり、11月のバージョン12.0.2でさらに大きく進化しました。以下の3つの機械学習およびAIテクノロジーベースの新機能を提供した点が大きな特長です。

#### 1. 教師あり学習に基づく文書の自動分類

これまでルールベースや事前記述で行われてきた文書の分類を自動化します。手本(教師データ)を投入した学習作業についてもユーザー自身が容易に行える管理ユーザー・インターフェースを提供します。

#### 2. 類似文書検索

ユーザーが選択した文書と類似した文書を類似度の高い順にリスト表示することで、情報を効果的に取得することを可能とします。

#### 3. AI辞書登録支援ツール

手間のかかる辞書登録作業をAIが支援。ワードを登録するとそれと類似した文脈で使われているワード(=辞書登録候補)をAIが類推して提示します。

その他にも、oneWEXでは分析の初心者でも洞察の獲得にたどりつけるよう、ガイド型のユーザー・インターフェースを新たに搭載しました。組織における分析ユーザーの層を広げることができるでしょう。



## ユースケース

### 自動車の故障対応にみる技術文書探索の流れ

実際にどのような手順を経て、Watson Explorerは大量の技術文書データの中から重要な知識や知見、ノウハウにたどりつくことができるのでしょうか。ユースケースをもとにその流れを追っていきます。

#### 課題の確認

ある自動車メーカーで故障対応担当として配属された新人技術者 新山氏は、「車種A」が走行不能になるという故障を受け、原因の調査を開始しました。そして、過去にベテランが作成した故障報告書や修理履歴書、またマニュアルなどの数十万件もの技術文書を学習させたWatson Explorerでファセット・ペア分析をしてみると、「車両A」は「ロアアーム」と非常に相関が高いという結果が出ました。

この事象は新山氏にも理解できる知識の範囲内で、十分に納得できる答えです。しかし「車種A」と「ロアアーム」の組み合わせにヒットする技術文書は1万件にも上り、まだ具体的な故障原因や対処法を得ることはできません。

#### 未知だった事実の発見

そこで新山氏は、Watson Explorerのタブをクリックして画面をコネクション分析に切り替え、「車両A」と「ロアアーム」の組み合わせに対して高い相関を示す別の要素を探索しました。すると瞬時に洗い出されたのが「ボールジョイント」です。これは新山氏の知識の範囲外で、まさに新たな“発見”です。

## 事象間の因果関係を抽出

さらに新山氏はファセット分析の画面に切り替え、「車両A」「ロアアーム」「ボールジョイント」に相関の高い故障を検索してみました。そこから見えてきたのが、「ボールジョイント(揺れる)」→「ロアアーム(はずれる)」→「車種A(走行不能)」という因果関係です。そして、この分析結果を条件に加えてふたたび検索を行ったところ、該当する技術文書を100件程度にまで絞り込むことができました。これなら内容を直接読んで確認できる件数であり、新山氏はベテランが書き残した記録から「ロアアームのボールジョイント部に問題がある」という原因に確信を持つと共に、その対処法を学ぶことができました。ここに至るまでの時間はわずか20分でした。

## 機械学習による自動分類

さらに、先述したoneWEXとして実装された最新の機械学習テクノロジーにより、新たな不具合報告書を学習済みの分類器に通すだけで、原因を予測して関連するタグが自動で発行されます。それにより新山氏は手作業による文書分類を行うことなく、大量の技術文書から探している情報を素早く見つけ出すことができます。

## 自動車メーカーの故障原因分析の例

### ファセット・ペア分析

「車種」と「名詞」で見る。  
「車種A」と「ロアアーム」の相関が特に高いことに気づく

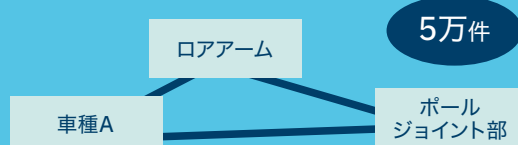
	車種A	車種B	車種C	車種D
ロアアーム	■			
エアサス				
サスペンション			■	
制動装置				

100万件

絞り込み

### コネクション分析

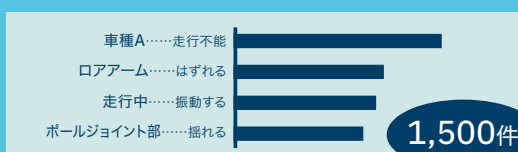
「車種A」「ロアアーム」「ボールジョイント」の  
組み合わせの相関が高い



絞り込み

### ファセット分析

どんなクレームが多いのか。  
絞り込むと、なんとなく問題が見え始めた



絞り込み

目で追える件数に絞り込んでから  
文書の内容を確認。  
問題が明確になった

100件

ロアアームがはずれ、  
かじ取り不能

ボールジョイント部に  
ガタがあった

前輪のロアアームが  
はずれた

### 知識発見& アクション

ロアアームの  
ボールジョイント部に  
問題がある。  
至急改善が必要だ



## まとめ

### テクノロジーで知の継承が可能に

紹介したユースケースから言えるように、絞り込み条件と言葉との関連の強さを表す「相関値」を用いて、目的とする技術文書の探索(絞り込み)を行うのが Watson Explorer の特長です。従来の単純なキーワードを用いた検索技術でありがちな「ありきたりのものしか見つからない」という結果に陥ることなく、若手技術者は探索の途中プロセスで必ず新たな「気づき」や「発見」を得ることができます。実はこれこそが熟練技術者が持っていた“暗黙知”そのものなのです。

このように若手技術者は、技術文書を探索する途中プロセスおよび結果として絞り込まれた技術文書から多くの知識や知見、ノウハウを修得することができます。そして、それに基づいて彼らが実施したアクションの結果は、また新たな知識(=教師データ)としてシステムに蓄積され、次の世代へと伝承されていきます。

なお、Watson Explorer は API 経由で Web アプリケーションやモバイル・アプリケーションから呼び出すことが可能であり、さまざまな担当業務に最適化されたユーザー・エクスペリエンスを設計することで、より簡便かつ効率的に知識や知見、ノウハウの伝承を図ることが可能となります。