

最新 Watson Explorer 活用法 (事例紹介・デモ)

日本 アイ・ビー・エム 株式会社
アナリティクス事業部 テクニカルセールス
河村 忠明

2,500,000,000,000,000,000 (250京)

2.5 quintillion

バイトの新しいデータが
日々生成されています

今日の専門知識はこれまで以上に重要ですが、トップの専門家でさえ維持できません



データは毎年2倍のスピードで増えています



Health data
will grow
99%



Insurance data will
grow
94%



Utilities data
will grow
99%



Manufacturing data
will grow
99%

88%

84%

82%

82%

非構造化データ
の割合

80%以上が、非構造化データ（コンテンツ）で占められています

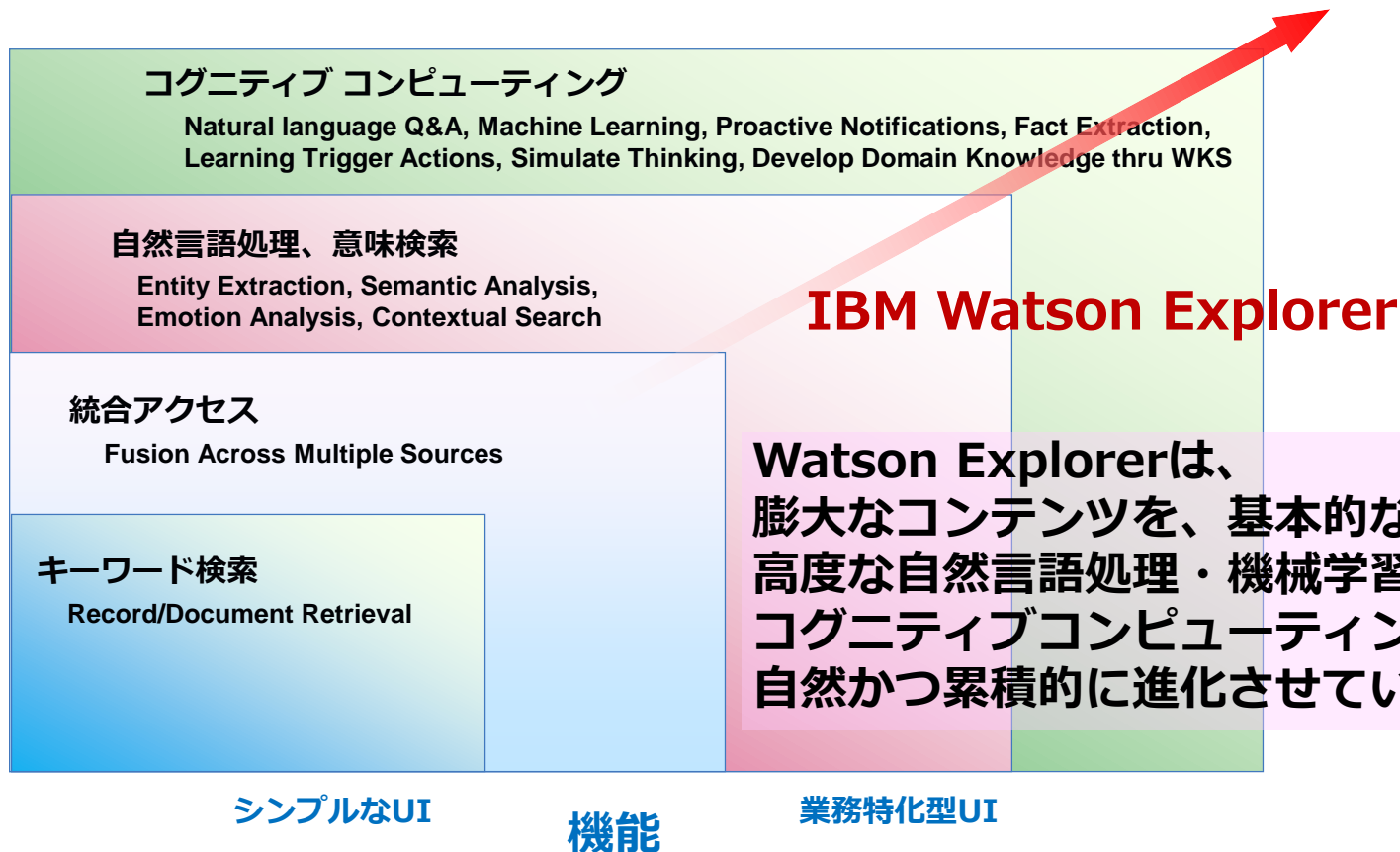
通常の組織はこのコンテンツのわずか8%しか活用できていません*

本日の内容

- コンテンツソリューションの進化
～ 検索からコグニティブまで ～
- Watson Explorer による検索、事例
- Watson Explorer による分析、事例
- IBM Watson との連携事例

コンテンツソリューションの進化 ～検索からコグニティブまで～

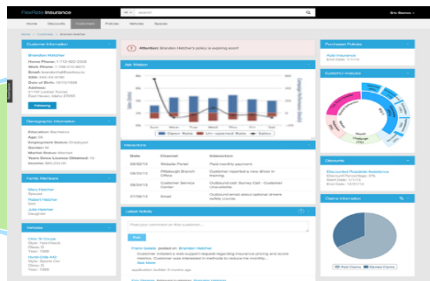
ビジネス価値



お客様の業務課題を解決する Watson Explorer

Enterprise Edition

(Foundational ComponentsのみのEdition)



Advanced Edition

(2つのComponentsを提供するEdition)



① 統合型検索エンジン

(名称 : Foundational Components)

ビジネスに関わるあらゆる情報を全方的に
統合・検索・視覚化
を行い探索します

② 高度なテキストマイニング

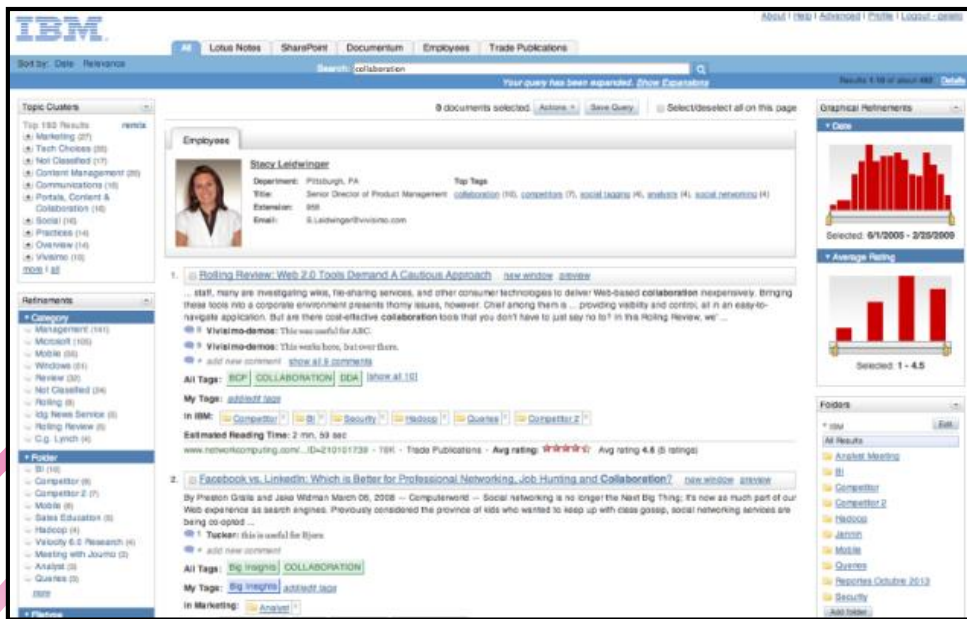
(名称 : Analytical Components)

自然言語処理で重要な情報を
分析・視覚化・暗黙知の発見
を行います

Watson Explorer による検索 (Foundational Components)

- ・ 定型・非定型情報の統合
- ・ オンプレ/クラウド、内部/外部データの統合
- ・ 分析画面、WatsonAPIの統合
- ・ 様々な社内コンテンツサーバーに対応

様々な企業内コンテンツの高速な横断検索を実現



File
Server



Web



RDB



Lotus
Notes



Share
Point



Email



CRM

Sales
Force

BOX

Office365

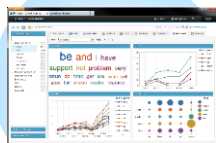
高度な統合画面

Watson Explorer: An intelligent hub for hybrid cloud environments

企業内データ



分析画面



File Server



Web



RDB



Lotus Notes



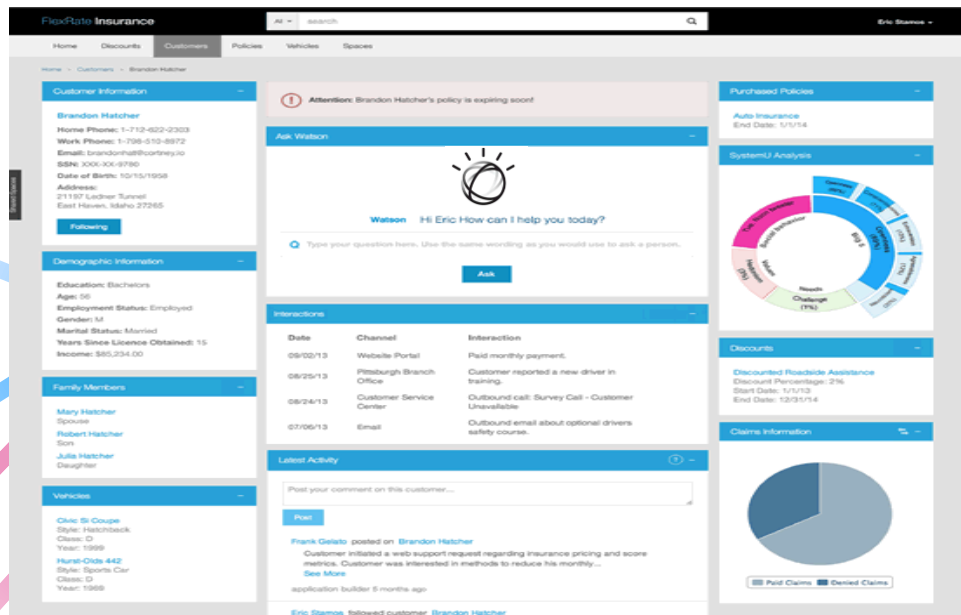
Share Point



Email



CRM



クラウド



ソーシャルデータ



Sales Force

BOX

Office365

検索 (Foundational Components) の事例

メンテナンスを改善、運用効率の向上

大手航空機プロバイダー



課題

- ・地上の技術サポート担当者は、**データがシステム全体に散在している**ため、技術的な問題の診断と修復が困難
- ・技術的解決時間が延長される事により、財政的罰金とメンテナンス費用が発生

ソリューション

- ・Watson Explorerにより、**多数のリポジトリ**（保守記録、顧客の通話レポートなど）**からのデータを組み合わせたWebベースの電子記録システム**を提供し
- ・テクニカルサポートの専門家やその他のフィールドユーザーが問題を診断し修正する事を簡素化

効果

- ・コール解決時間の短縮（**50分から15分に短縮**）
- ・サポート費用を**3600万ドル削減**
- ・人員を追加することなく、初年度に50台の追加飛行機を立ち上げ

医療機器会社における現場是正措置の迅速化

大手医療機器会社



課題

- ・FDA(食品医薬品局)の要件を満たすため、品質エンジニアによる**現場での**是正措置の**時間短縮**が必要
- ・内部システムや外部Webサイトを含む、**8つ以上のデータサイロ**に格納された**数テラバイトの製品**および顧客苦情データへのアクセスが必要

ソリューション

- ・エンジニアに**360度視点での製品品質と苦情情報を提供**するデータ探索プラットフォームを提供
- ・重要なレガシー情報ソースをクロールおよび索引付けして、**ユーザー自身が**検索およびファセットナビゲーションを使用して**情報を探索**できるように整備

効果

- ・現場の是正措置のための時間を大幅に短縮
- ・顧客の苦情や製品の問題を診断する能力の向上
- ・製品の問題や苦情に対する、エンジニアとの共同作業を実現

データ管理および検索を介して統一された製品情報を取得

大手化粧品会社



課題

- ・ 4つの部門と3つの地域にまたがる**複数のシステムからの製品情報を統合**し、情報管理とクロスファンクショナルの取り組みを改善する
- ・ 内部のレポートと計画を改善してグローバルな最適化をサポートし、コストを削減する
- ・ データの一貫性を提供しながらローカライズを容易にする

ソリューション

- ・ Watson ExplorerおよびMaster Data Managementを導入して、**複数のソースから製品情報を統合**した単一アプリケーションの設計、構築、デプロイ、およびサポートを実現

効果

- ・ **統一された情報アプリケーション**による戦略的な利点と生産性の向上
- ・ データ品質の向上による大幅な節減
- ・ 欠落/間違ったデータの**解決に要する時間の短縮**

企業間の安全な検索

消費者製品研究所



課題

- ・ 幅広いデータ・ソースと内部環境を備えており、これを異なる役割を担う同社のユーザーが迅速かつ安全に情報検索できるようにする必要がある

ソリューション

- ・ 外部向け Web サイト、ならびに全機関に対して情報を提供する社内ポータルで、研究資産の安全な意味検索と分類を実施できるようにすべく、Watson Explorerを実装

効果

- ・ キーワード検索のみならず、コンテンツ分類と文脈検索の導入も実現
- ・ 特定の照会条件の結果のみでなく、ユーザーの探しているコンセプトに基づいて、的確な検索結果が得られるようになる

ポータルサイトを通した安全な検索

グローバル食品飲料企業



課題

- ・ Connections、SharePoint など、異なるプラットフォームでの社内ポータル検索
- ・ 業務機能に基づく独自の検索スペース

ソリューション

- ・ Connections、SharePoint に対して統合的な横断検索を実現
- ・ 関連性の高い検索が可能なクロス・リポジトリ検索を提供
- ・ 新たな企業拠点や地域の追加についての企業計画を作成

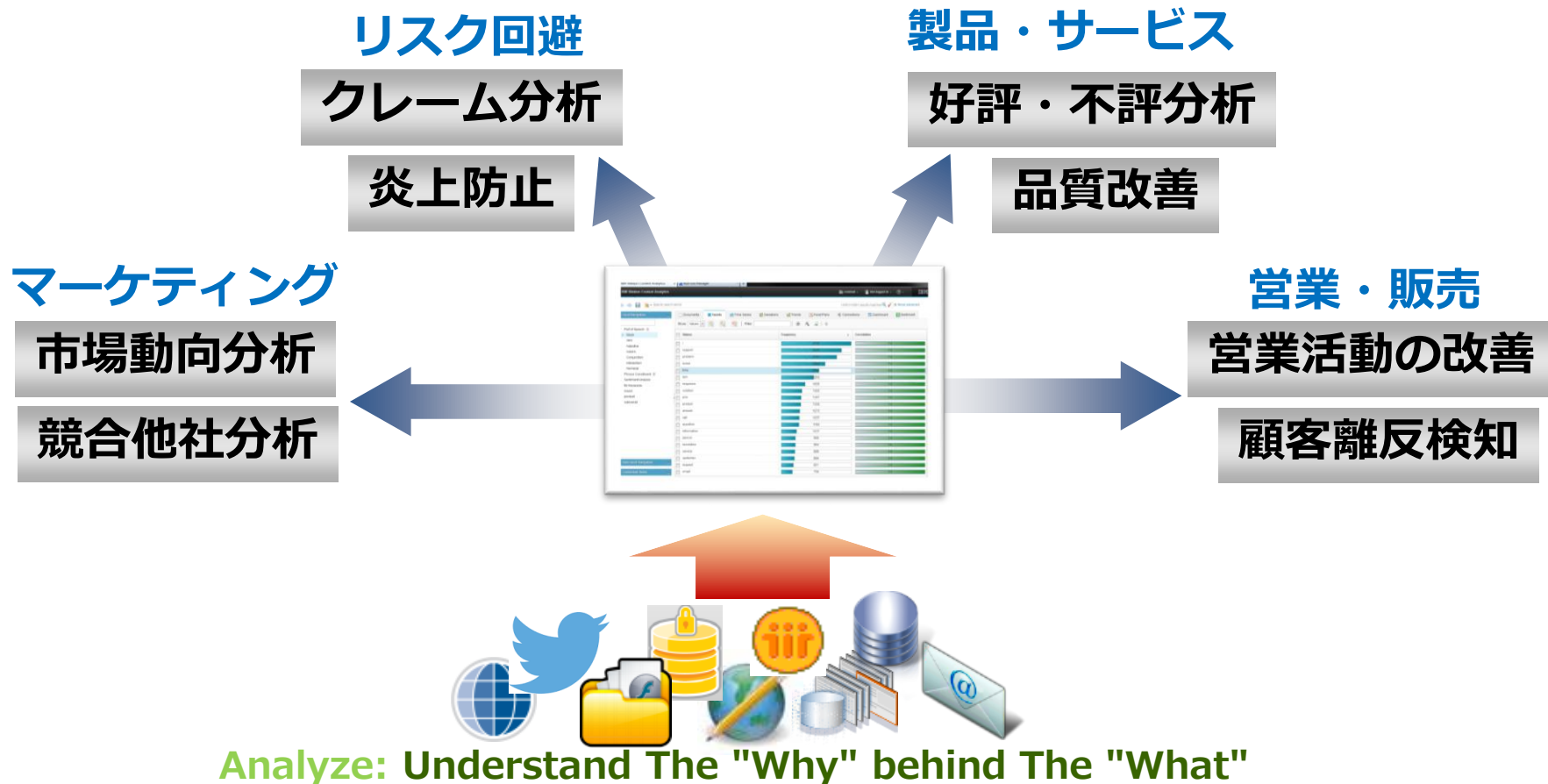
効果

- ・ 様々なリポジトリの検索にかかる時間やコストを削減
- ・ ポータル統合により、コンテンツの完全な把握が実現

Watson Explorer による分析 (Analytical Components)

- ・ 大量テキストの分析、定量化
- ・ 早期検知、知見抽出
- ・ 顧客分析

テキストマイニングはどのようなことに役立つのか



非構造化データからアクションへ

Watson Explorer Deep natural-language analysis

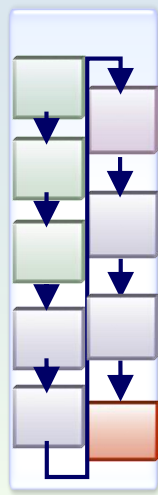
非構造化 データ

[機種名]
ThinkPad

質問：ThinkPad
T42でメモリ増設を
したいと思うのですが、
512MBのメモリでも
大丈夫でしょうか。

1GBの時には認識さ
れませんでした。

抽出・ 分類(タグ付け)



UIMA
フレームワーク

コーパス (索引)

機種	ThinkPad
モデル	T42
部品	メモリ
操作	増設
単位	512M
分類大	質問
分類中	対応可否
単位	1GB
分類大	報告
分類中	エラー

分析



分析結果

T42ではメモリ
増設の要望が多い

T42では1GBメモリ
が認識されない



対応

1GBメモリ
モデルの発売

抽出・分類(タグ付け) ～UIMA フレームワーク～



テキスト解析の要素技術

UIMAフレームワーク

言語判定
アノテーター

形態素解析
アノテーター

品詞特定
アノテーター

係り受け解析
アノテーター

辞書キーワード認識
アノテーター

固有名詞抽出
アノテーター

パターン認識
アノテーター

ルールベース
アノテーター

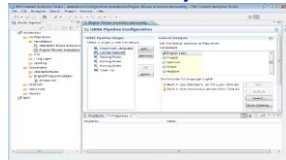
カスタム
アノテーター

パターン・エディター

Watson Knowledge Studio



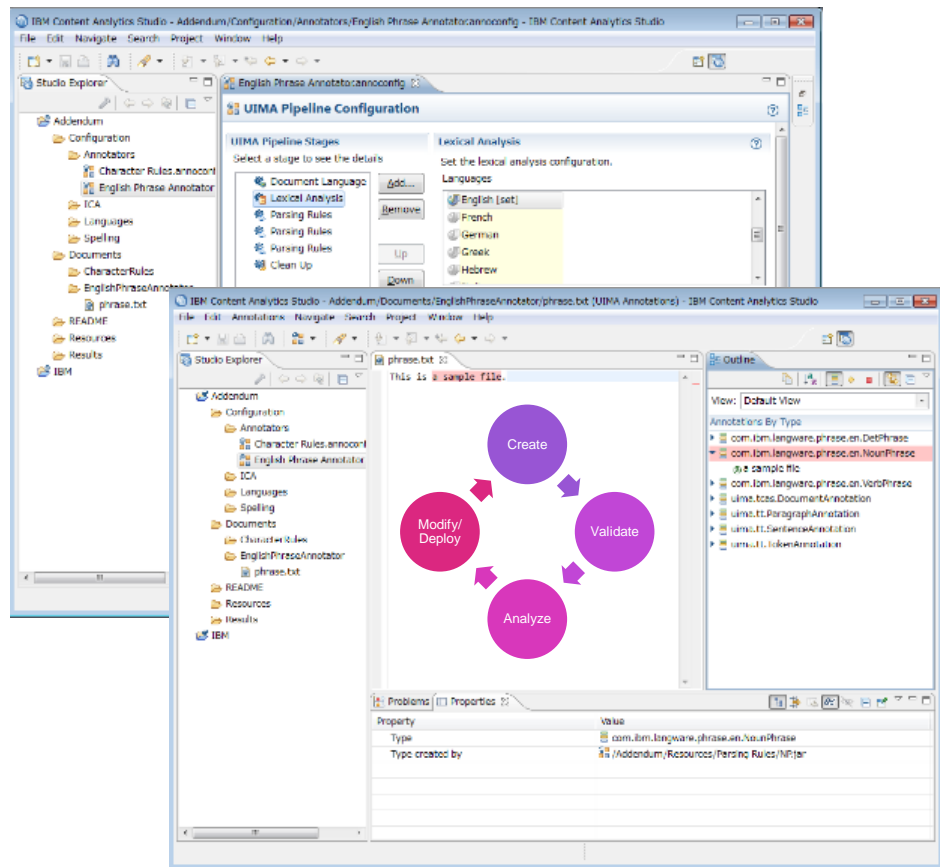
Content Analytics Studio



辞書・エディター

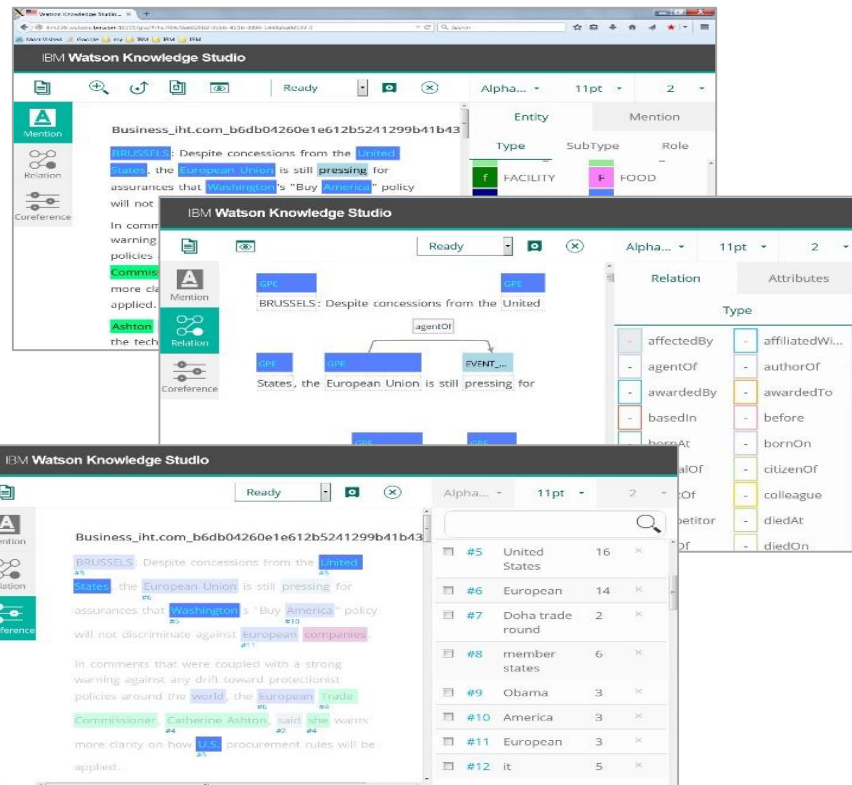
Content Analytics Studio

- 各ビジネス要件に基づき、実データをサンプルにして、高度な辞書・ルール設計が可能
- 各ビジネス要件に対応するコーパス作成の規則を1つのアノテータパッケージに集約できる
- Studioの操作は、簡単なドラッグ&ドロップのGUIインターフェースで行う
- 作成されたアノテータパッケージは、WatsonExplorerでのみ使用可能



Watson Knowledge Studio

- 基本要素を予め定義し、後は推論と統計モデルを使って機械学習により自動識別を行う
- 各ビジネス要件に対応するコーパス作成の規則を1つのアノテータパッケージに集約できる
- Studioの操作は、簡単なドラッグ&ドロップのGUIインターフェースで行う
- 作成されたアノテータパッケージは、WatsonExplorer、WatsonAPI(Discovery, NLU)で使用可能



テキストマイニングはどのようなことに役立つのか



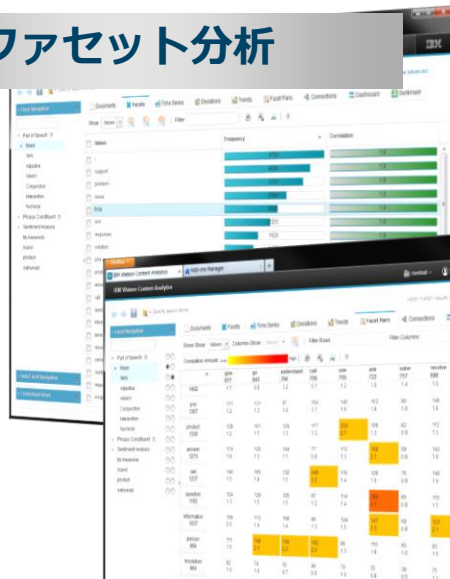
Analyze: Understand The "Why" behind The "What"

知見を抽出する分析

2007年に東京基礎研で誕生してから様々な企業や用途で使われ改善を重ねてきました。単純なレポート用途ではなく、具体的なアクションにつながる高度な分析機能を提供します。

相関計算によるBIGDATA テキスト分析

ファセット分析



偏差分析



トレンド分析



ファセットペア分析



ネットワーク分析



評判分析



グローバル対応 ～多言語サポート～

分析17ヶ国語、検索22ヶ国語に対応

基本言語処理

分析機能	アラビア語 中国語 チェコ語 デンマーク語	オランダ語 英語 フランス語 ドイツ語 ヘブライ語	イタリア語 日本語 韓国語 ポーランド語	ポルトガル語 ロシア語 スペイン語 トルコ語
検索機能	アラビア語 中国語 チェコ語 デンマーク語 オランダ語 英語	フィンランド語 フランス語 ドイツ語 ギリシャ語 ヘブライ語	イタリア語 日本語 韓国語 ノルウェー語 ポーランド語	ポルトガル語 ロシア語 スペイン語 スウェーデン語 タイ語 トルコ語

分析 (Analytical Components) の事例

欠陥を特定するための洞察を抽出

大手自動車メーカー



課題

- ・ 欠陥を事前に特定し、人命・法的支出・ブランドイメージを守るための**分析ソリューション**が必要

ソリューション

- ・ National Highway Safety Administration (NHTSA) データに基づく Watson Explorerソリューションを提供
- ・ 11の情報源からの**740万レコード**の情報と、10年間をカバーする**100万のNHTSA情報**でソリューション構築を開始
- ・ データ・サイロと苦情の分析に基づくトレンドを統合し、包括的な分析により詳細を掘り下げる事を可能にした

効果

- ・ 欠陥を特定して未然にリコールを行う事により、人命を救い、**何百万ドルもの罰金とブランドイメージの損失を防ぐ**システムを実現した

お客様の声を聞き、他社への乗り換えを削減

大手通信企業



課題

- ・顧客満足度を向上による、市場シェアの確保
- ・「お客様の声」を理解し、契約解消を未然に防ぐ機会を特定する事が必要
- ・新しいサービスを速やかに立ち上げるために、迅速な問題対応が必要

ソリューション

- ・他社へ乗り換える可能性の高い顧客を速やかに発見
- ・顧客の問題を特定し、セルフサービス Web サイトへ掲載することが望まれる FAQ 候補を提案
- ・顧客体験に関連する傾向やパターン、特徴的な製品やサービスを抽出

効果

- ・乗り換えそうな顧客発見の精度を 50% アップ
- ・固定客への製品モデルやサービスのアップグレード率を改善
- ・セルフサービスの FAQ システムを改善
- ・新しい提案やサービスに向け、お客様の声をモニタリング

顧客応対を分析しエージェントのトレーニングの質を向上

ビジネス・サービス・プロバイダー



課題

・エージェントのトレーニングを改善し、より良いカスタマーサポートを提供するため、顧客サービス上のやりとりで得られる大量のデータを分析する必要があり

ソリューション

・ Watson Explorer により、電話や E メール、Web 上の対話で得られた統合ログに基づいて顧客サービスでの難しいやりとりを分析
・ 問い合わせの中で使用されるキーワードを特定し、エージェントの応対を評価

効果

・ 顧客サービス上のやりとりを分析し、難しいトピックやエージェントのスキル不足を特定して、トレーニングを改善
・ 主要なエージェントが電話を転送することなく取り扱う問い合わせ件数が 20% アップ
・ 製品の問題や非効率性に関する新しい洞察を統合し、管理者がサービスやトレーニング、製品開発の質を向上できるように改善

世界3万人の従業員に企業情報への透過的アクセスを提供

大手製薬会社



課題

・買収した企業のシステムを含め世界中の異なる情報システムに保存されている膨大な量の顧客データ、患者データ、およびリサーチデータへのアクセスを提供することにより、情報およびデータ共有機能の変革が必要

効果

- ・研究開発の効率を90%向上させ検索時間を50%短縮し、初年度だけで数百万ドルを節約
- ・新規人員の要件を1.2%削減し、年間1340万ドルの削減
- ・売上高を1.4%向上させるための販売生産性の向上

特許情報・技術文書の検索・分析

知財部門・製造部門



課題

- ・ 競合他社や業界全体の技術動向を素早く捉える為、**BigDataに潜んでいる**経済的効果を生む**先進技術の予兆**を捉えたい
- ・ 様々な技術分野で、競合他社の**特徴的な特許技術**を検出し、自社との差を明確化する一方で類似特許も見つけたい
- ・ 読解が難しい特許の**コアアイディア**を短い時間で把握したい

ソリューション

- ・ 出願傾向の分析・戦略的な出願のされ方をモデル化した分析を実施
- ・ 多様な切り口で、**相関の高い言葉**を検出し、分析請求項の書かれ方をモデル化し、コアアイディアを抽出
- ・ コネクションビューによる**発明の構造化表示**機能

効果

- ・ 経済的効果を生み出す技術への**予兆を検出**
- ・ 電気自動車等カテゴリー別の**新語を検出**
- ・ **分野別特徴や類似特許を検出**

保険業界における事例

支払い査定支援、お客様の声分析



富国生命

- ・診断書査定自動コード化システム

明治安田生命

- ・支払査定支援システム

三井住友海上

- ・お客様の声分析システム

かんぽ生命

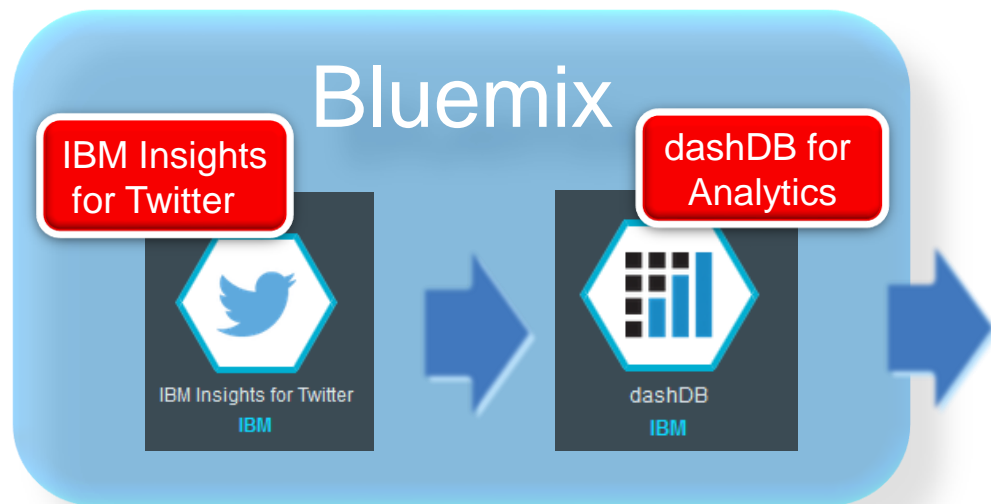
- ・査定担当者への判断支援

損保ジャパン日本興亜ひまわり生命

- ・支払い業務支援システム

Bluemixを利用したSNS分析

- Bluemix上のTwitter全量取得サービスと、分析用SQLDBサービス
を利用し、直近又は過去のTwitterデータの分析が可能



Watson Explorer

ワークショップ、セッション、および資料は、IBMまたはセッション発表者によって準備され、それぞれ独自の見解を反映したものです。それらは情報提供の目的のみで提供されており、いかなる参加者に対しても法律的またはその他の指導や助言を意図したものではなく、またそのような結果を生むものでもありません。本講演資料に含まれている情報については、完全性と正確性を期するよう努力しましたが、「現状のまま」提供され、明示または暗示にかかわらずいかなる保証も伴わないものとします。本講演資料またはその他の資料の使用によって、あるいはその他の関連によって、いかなる損害が生じた場合も、IBMは責任を負わないものとします。本講演資料に含まれている内容は、IBMまたはそのサプライヤーやライセンス交付者からいかなる保証または表明を引きだすことを意図したものでも、IBMソフトウェアの使用を規定する適用ライセンス契約の条項を変更することを意図したものでもなく、またそのような結果を生むものでもありません。

本講演資料でIBM製品、プログラム、またはサービスに言及していても、IBMが営業活動を行っているすべての国でそれらが使用可能であることを暗示するものではありません。本講演資料で言及している製品リリース日付や製品機能は、市場機会またはその他の要因に基づいてIBM独自の決定権をもっているいつでも変更できるものとし、いかなる方法においても将来の製品または機能が使用可能になると確約することを意図したものではありません。本講演資料に含まれている内容は、参加者が開始する活動によって特定の販売、売上高の向上、またはその他の結果が生じると述べる、または暗示することを意図したものでも、またそのような結果を生むものでもありません。パフォーマンスは、管理された環境において標準的なIBMベンチマークを使用した測定と予測に基づいています。ユーザーが経験する実際のスループットやパフォーマンスは、ユーザーのジョブ・ストリームにおけるマルチプログラミングの量、入出力構成、ストレージ構成、および処理されるワークロードなどの考慮事項を含む、数多くの要因に応じて変化します。したがって、個々のユーザーがここで述べられているものと同様の結果を得られると確約するものではありません。

記述されているすべてのお客事例は、それらのお客様がどのようにIBM製品を使用したか、またそれらのお客様が達成した結果の実例として示されたものです。実際の環境コストおよびパフォーマンス特性は、お客様ごとに異なる場合があります。

IBM、IBM ロゴ、ibm.com、IBM Watsonは、世界の多くの国で登録されたInternational Business Machines Corporationの商標です。他の製品名およびサービス名等は、それぞれIBMまたは各社の商標である場合があります。現時点での IBM の商標リストについては、www.ibm.com/legal/copytrade.shtmlをご覧ください。