

企業内ソーシャル・データの分析・活用

「IBM Watson Explorer Content Analytics」を用いた「IBM Connections」分析の勧め

TwitterなどのSNSデータを分析してマーケティングに活用する事例が増加しています。近年広がりつつある企業内SNSのデータも、企業・業界・顧客の現状を把握する上で分析価値の高いデータと考えられます。

本稿では、IBMの企業内SNSソフトウェア「IBM Connections」のデータをテキスト・マイニング・ソフトウェア「IBM Watson Explorer Content Analytics」で分析する手法とその価値を解説します。

▶▶ 1. 企業内ソーシャル・データの価値

TwitterやFacebookに代表されるSNSが広く一般に浸透し、SNSの人をつなげる力・高い情報伝搬能力を企業内コラボレーションでも生かそうという企業内SNSが徐々に広がりつつあります。企業内SNS導入の効果としてまずイメージされるのは、ファイル共有やつぶやき、ディスカッション機能による情報共有の効率化や、プロフィール機能による専門家発見のスピードアップです。しかし、それだけではありません。

マネジメント層が投稿したブログは今の会社の方向性、営業員がアップしたファイルは今お客様に求められている情報、新入社員が掲示板に投稿した質問は今若手が注

目している事柄の表出です。そういったものの集積である企業内ソーシャル・データには、いわば会社の“今”が凝縮されています。そのようなデータを入手できることも、企業内SNS導入の大きな効果の一つです。

近年、TwitterなどのSNSデータを分析して消費者の“今”を知りマーケティングに生かそうという企業が増加しており、IBMでも多くのお客様のお手伝いをしています。そういったプロジェクトに携わる中で、アンケートなどで意図的に抽出された回答とは異なる、“消費者が自発的に発した声から得られる洞察”の価値を強く感じます。

同様のことが、企業内ソーシャル・データでも可能であると考えています。次章以降では、企業内SNSソフトウェア「IBM Connections」(以下、Connections)と、

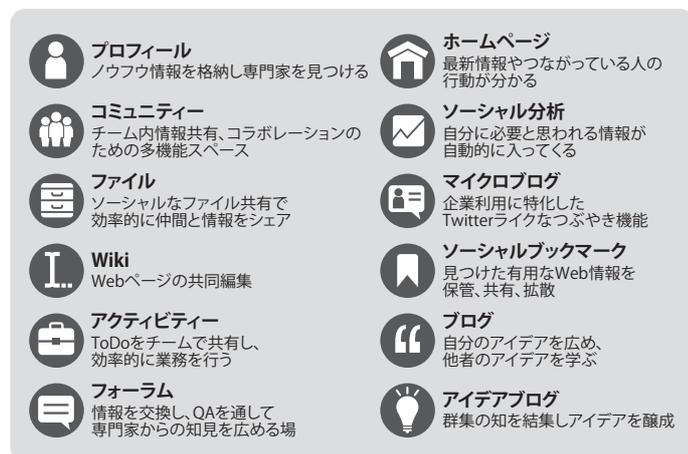


図1. Connections機能概要

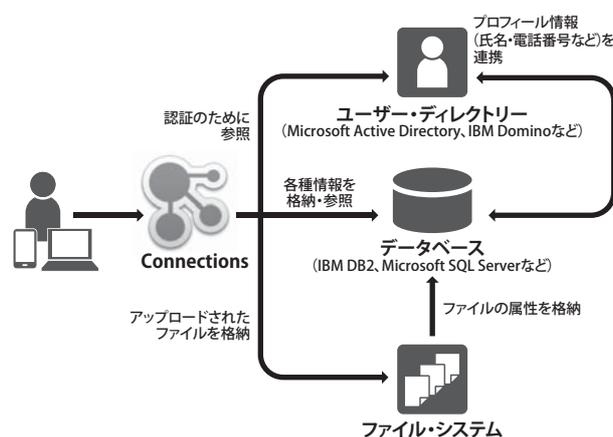


図2. Connectionsのアーキテクチャ

テキスト・マイニング・ソフトウェア「IBM Watson Explorer Content Analytics」(以下、WCA)を連携させて、社員・顧客・業界の“今”を把握し次の一手につなげる仕組みとその価値を解説します。

2. ConnectionsとWCAの概要

企業内ソーシャル・データ分析のベースとなるConnectionsとWCAの概要を紹介します。

●Connections

Connectionsは、IBM WebSphere Application Server (以下、WAS)上で稼働するWebアプリケーション群です。社員の所属・連絡先・専門性を示すプロフィール機能、ディスカッションやQAの場として使えるフォーラム機能、単なるファイル共有だけでなくバージョン管理も行えるファイル機能、特定のユーザーだけの共同作業を行えるコミュニティー機能など、さまざまなニーズに応える機能を備えています(図1)。

ユーザーがアップロードしたファイルはサーバーのファイル・システム上に、ファイルのメタデータ(アップロードしたユーザー名、「いいね」の数など)やブログ・Wiki

などのテキスト・データはデータベースに格納されます。また、ユーザー認証やプロフィール情報を取得するために、ユーザー・ディレクトリーを参照します(図2)。

●WCA

WCAは、Connectionsと同様Webアプリケーション・サーバーを基盤として動き、テキスト・データを形態素解析し、単語間の相関関係や出現頻度、時系列などを分析します。文書ファイル・データベース・Webサイトなどさまざまな形式のデータを扱うことができます[1]。テキスト・マイニング結果はWCA画面上でグラフィカルに表示したり(図3)、他の分析ツール(「IBM Cognos BI」など)に取り込んでさらに深い分析を行うことが可能です。組み込みのWebアプリケーション・サーバーを使用してクイックに構築することも、WAS上にインストールしてWASの高度な負荷分散/冗長化機能を活用することもできます。WCAは、対象データを取り込むコンポーネント、取り込んだデータを解析し索引付けするコンポーネント、解析結果をユーザーに表示するコンポーネントといった複数のコンポーネントの集合体です。



図3. WCAの分析画面例

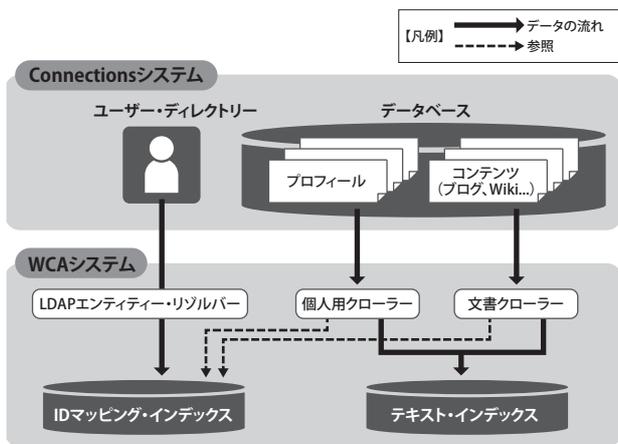


図4. ソーシャル分析用データ生成の仕組み

3. WCAを用いたConnectionsの分析

Connections上に投稿されたテキスト・データは、先述のとおりすべてデータベースに格納されるため、それだけを分析することは難しいことではありません。しかし、ConnectionsのようなSNSの価値は、すべてのデータが「人」に紐付いていることです。誰がこの知識を持っているのか、どういう人がこの情報に興味があるのかといったことまで把握できてこそ、ソーシャル・データを分析する価値があります。

WCAはV3.5から、ソーシャル分析機能が実装されました[2]。これは、単なるテキスト・マイニングだけではなく、情報を投稿したりコメントを付けたりしたユーザーが誰なのかといった情報を含めて解析し、そのデー

タを基に特定分野の専門家を抽出したり、ユーザー間のつながりを可視化したり、ユーザーに関心のありそうな情報をレコメンドしたりします。

ソーシャル分析用のデータは、一般的なテキスト・データを収集する「文書クローラー」に加えて、Connections上のユーザー情報(プロフィール)を収集する「個人用クローラー」と、ユーザー・ディレクトリー上の情報を収集してIDマッピング・インデックスを作成する「LDAPエンティティー・リゾルバー」が動作して生成されます(図4)。

個人用クローラーが収集した情報はテキスト・データと同様に解析および索引付けされ、誰がどんな情報を投稿したか、誰と誰がつながっているのかといったユーザー・ベースの分析に使用されます。また、Connections上のデータに内部的に紐付けられるユーザーIDは複数の形式があるため、「abcde」というIDと紐付いている文書も“uid=tyamada,o=ibm,o=com”というIDと紐付いている文書も『山田太郎』が作成したものといった名前解決をするために、LDAPエンティティー・リゾルバーが作成したIDマッピング・インデックスが使用されます。

こうしてデータを生成した上でソーシャル・データ分析を行います。WCA上で分析を行う際には、まず「コレクション」と呼ばれる“箱”を作成します。コレクションにはクローラーやLDAPエンティティー・リゾルバー、



図5. エンタープライズ検索コレクション画面例

クロールして取得したデータ、索引付け用の辞書、生成されたインデックスなどすべてが含まれ、一つのWCAサーバー上に複数作成することができます。コレクションごとに個別のURLが用意され、ユーザーはWebブラウザからWCAの画面にアクセスして分析を行います。一つのコレクションで複数のデータを対象とした分析(例: Connectionsと社内ポータルデータを同じ画面で分析する)を行うことも、同じデータを対象として全く違う目的の分析を行う場合に別のコレクションに分ける(例: 経営企画が興味のある切り口の辞書でConnectionsを分析できるコレクションと、製品開発部が興味のある切り口の辞書でConnectionsを分析できるコレクションを分ける)という使い方もできます。

コレクションには、「エンタープライズ検索コレクション」と「コンテンツ分析コレクション」の大きく分けて2種類があります。それぞれのコレクションで使用できる機能が異なりますが、ソーシャル分析機能はどちらのコレクションでも使用でき、ソーシャル分析結果の使用目的に応じて使い分けられます。使い分けの判断基準となる、それぞれのコレクションの特徴を以下で説明します。

●エンタープライズ検索コレクション

大量のデータを効率良く検索するために最適化されており、探している情報を素早く見つけることができます。また、複数の検索コレクションで横断検索が可能です。

エンタープライズ検索コレクションでConnections

を検索したときのイメージが図5です。検索を行う前は、枠線で囲っている「ファセット・ツリー」には何も表示されません。何らかの検索を行うとその結果に紐付いているタグやユーザー名を動的に表示し、欲しい情報にたどり着くための絞り込みを助けてくれます。

●コンテンツ分析コレクション

テキスト・マイニングのための機能が多く利用可能で、分析によって新たな知見を得ることを目的とします。図3のような単語の出現頻度や相関関係の可視化などはコンテンツ分析コレクションでしかできません。

コンテンツ分析コレクションには2種類の画面があります。一つはソーシャル分析要素が多い画面(図6)、もう一つはテキスト・マイニング要素が多い画面(図7)です。前者には図5と同様、右端に「関連人物」「タグ・クラウド」などが表示され、「おすすめ」「ソーシャル・ネットワーク」「エキスパート」というタブもあります。これらはソーシャル分析機能をONにしたときに使える要素です。後者ではそれらが表示されない代わりに、図3で紹介したような深い分析用のタブがあります。ただし後者でも、データとユーザーの紐付き解析は可能です。同じコレクションに対してアクセスするURLを変えるだけで両方の画面が参照できます。

図5、図6に共通するソーシャル分析要素をピックアップしたのが図8です。「ソーシャル・ネットワーク」タブには、入力したキーワードに関連するユーザーの相関図



図6. ソーシャル分析要素が多いコンテンツ分析コレクション画面例

が表示されます。これによって、キーワードに関する専門家集団とその中心人物を把握することができます。「エキスパート」タブには専門家がランキング形式で表示されます。「情報を発信した>他のユーザーの投稿にコメントした>上位の専門家と関係が深い」というように、ユーザーとキーワードとの関連度を重み付けした上でランキングが作成されます。「おすすめ」タブには、データと人の紐付きを見た上で、キーワードと関連が深いと予測されるコンテンツが表示されます。画面右端の「関連人物」「知り合い」「タグ・クラウド」では、データと紐付いている人やタグ情報が一目で把握できるようになっています。

ここまで見てきたような分析画面を表示させるためのコレクションの作成やConnectionsとの連携は、WCAの画面から簡単に行えます。初回のデータ処理にかかる時間は、Connectionsに蓄積されたデータ量によりますが、設定作業自体には1時間もかからないでしょう。企業内SNSソフトウェアや分析ソフトウェアは世の中に数多くありますが、これほど簡単にかつリッチに企業内ソーシャル・データの分析を行える組み合わせはないでしょう。

▶▶ 4. 企業内ソーシャル・データ分析のメリット

こうして分析の基盤が整ったら、活用の可能性は無限

大です。一般ユーザーにとってエンタープライズ検索コレクションは、より効率的に質の高い情報にアクセスするための大きな助けとなり、Connectionsの活用を促進すると考えられます。しかし、冒頭で述べたように、社員・顧客・業界の“今”をつかむためには、ぜひコンテンツ分析コレクションも活用してください。

例えば、商品企画部が新しい商品を考えるとき、消費者のニーズを探るべくアンケートやインタビューなどの市場調査が行われますが、日々顧客と接している営業や販売担当員の持っている情報にも有益な示唆があるはず。そういった社員がもらしたつぶやきやフォーラムに投稿した意見、顧客のために作成した資料などを拾い上げることで、「商品名とどんな単語の相関が強いかなを見て現在の評判・課題の洗い出しをする」「その商品の注目度がいつ上がっているかを時系列で見て季節やキャンペーンの影響力を評価する」「その商品の専門家を見つけ出してさらに詳しい調査や意見を求める」といったことが可能になります。第3章で述べた通り、複数のデータを同時に分析することができるため、Connections以外のデータ（営業日報、コールセンターへの問い合わせ記録、顧客アンケート結果、Twitterのような外部のSNSデータなど）を組み合わせることでより幅広い分析を行うこともできます。



図7. テキスト・マイニング要素が多いコンテンツ分析コレクション画面例

もちろんこういった活用をするためには、Connections上に多くの情報がアップされていなければなりません。企業内SNSが活発に使われるようになるためには、新しいツールに慣れるだけでなく、企業文化・社員の意識の変化も必要になるケースが多く、「うちの社員はなかなか情報をアップしないから分析までたどり着くのは無理だ」と思われる企業もあるかもしれません。しかし、それは逆です。企業内SNSが活用されるようになってから分析を始めるのではなく、「ただ情報共有を便利にするだけではなく、社員からの情報や声を分析して社員・顧客・業界の“今”をつかむことは、ひいては業績向上につながる。だからみんな情報をアップしてほしい」というモードで最初から始めるのです。ユーザーがなかなか情報をアップしない理由として、「あえてSNSで共有する必要性が分からないから、わざわざ手間をかけたくない」という声を数多く耳にしますが、明確な目的を提示することはその障壁を取り除くのに有効な手段となります。「データ分析を目的に提示して企業内SNSの活用を促す」→「多くのデータが集まり有益な分析が行える」→「分析結果にユーザーが価値を感じてさらにSNSを活用する」→「より多くのデータが集まって分析の精度が上がる」というのが理想的な循環です。

本稿で紹介したように、ConnectionsとWCAとを連

携させることで大きなメリットを得ることができ、しかも非常に容易に開始できます。企業内SNSをすでに活用されている企業はもちろん、これから始めようとしている企業、定着に苦勞している企業も、企業内SNSの活用と合わせてソーシャル・データ分析に着手し、ソーシャル活用で得られる効果を最大限に享受してください。

[参考文献]

- [1] Software Product Compatibility Reports:Watson Explorer 11.0 – Related software for a specific product, <http://www.ibm.com/software/reports/compatibility/clarity-reports/report/html/prereqsForProduct?deliverableId=D1909460F42811E49F743BB1368AEB6A&osPlatforms=AIX|Linux|Windows&duComponentIds=SO01&optionalCapIds=30|22|221|132|20|26#sw-1>
- [2] Release Notes:problems and solutions in Watson Content Analytics Version 3.5, <http://www.ibm.com/support/docview.wss?uid=swg27037841#ga>



日本アイ・ピー・エム システムズ・エンジニアリング株式会社
オートメーション・ソリューション
アドバイザー・ITスペシャリスト

三輪 麻衣子
Maiko Miwa

2008年入社。入社以来、ICS(旧Lotus)製品を中心にIBMグループ内外への技術支援を担当。近年はソーシャル・モバイル・アナリティクス活用の研究に注力している。主な担当製品はIBM Connections、IBM Notes Traveler、IBM MaaS360、IBM Watson Explorerなど。

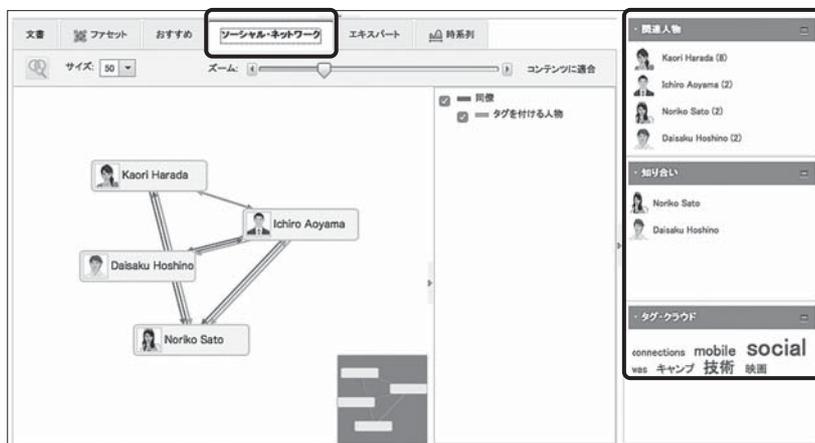


図8. ソーシャル分析要素の画面例