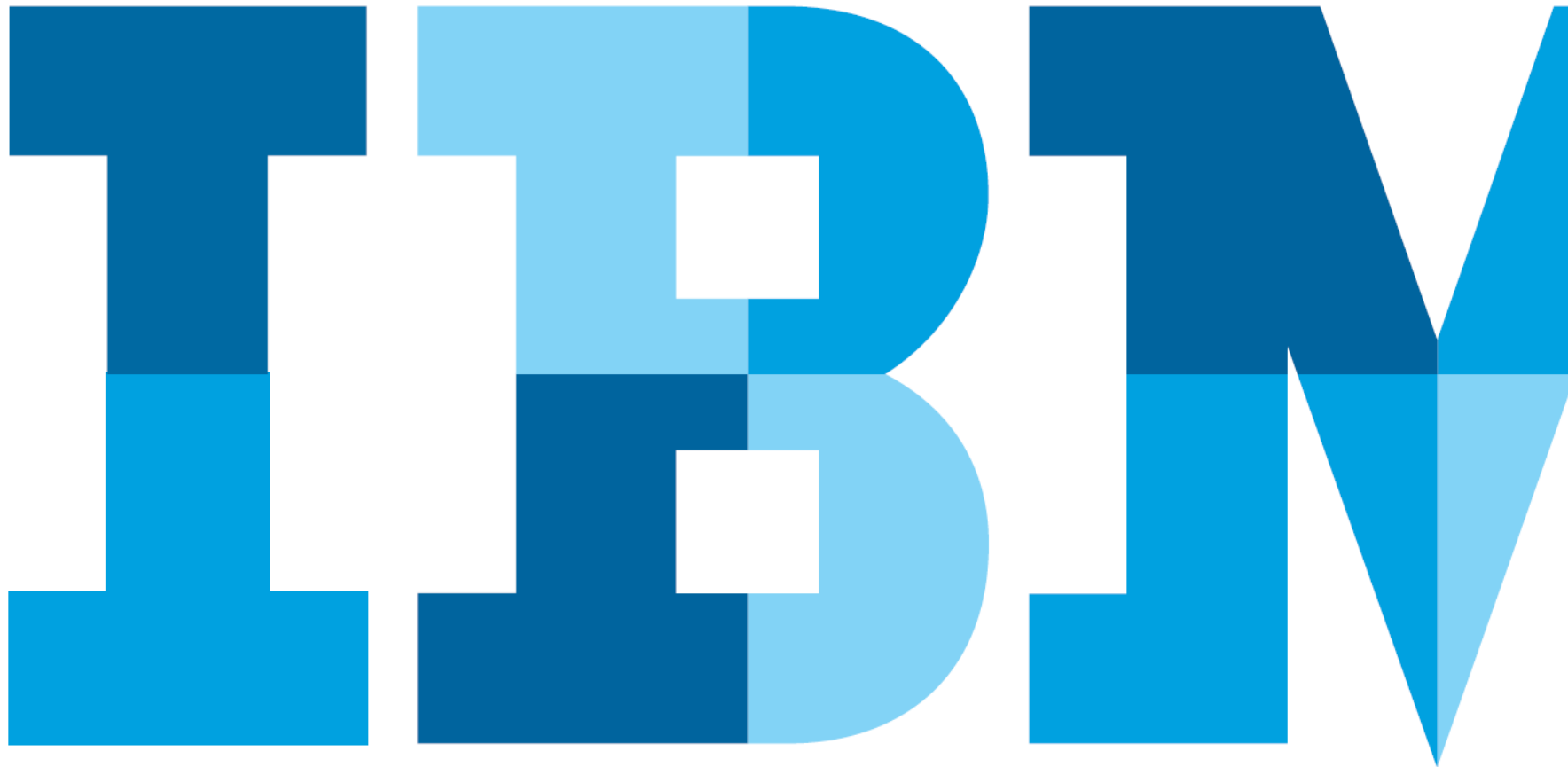


# 非構造化データの管理

非構造化データの課題を解決するためのクイック・ヒント



## はじめに

「非構造化データの問題 (The Unstructured Data Problem)」というタイトルの eBook で、企業が非構造化データを無視できない理由について説明されています。インターネットに接続されたデバイスで生成される非構造化データの量は増加の一途をたどり、その速度は増すばかりです。現在、企業は、ビッグデータの生成に起因する課題を解決するソリューションを探し始めなければならない状況にあります。

### 課題

非構造化データから洞察を得るには多くの障害があります。多くの企業では、その答えは1つのソリューションの中にあるのではなく、社内ツールと外部サービスの組み合わせにあります。最高レベルのデータ管理プログラムを利用することで、俊敏性や ROI (投資収益率) が高まります。最大の効果を得るためには、リーダーがすべての選択肢を検討する必要があります。データのモニタリング、収集、抽出、重複排除、正規化、そしてストレージについて検討することにより、戦略を作成します。これから、データに関心を持つリーダーが計画を立てる上で重要となる、いくつかの側面を見ていきましょう。

### ソリューション

現在、かなりの数のデータ・モニタリングおよび分析プラットフォームが提供されていますが、その大半はソーシャル・メディア・ネットワークに重点を置いています。しかし、その他のソース (ブログ、レビュー・サイト、Eメール、ビデオ・サイト、フォーラムなど) からデータを抽出して処理する機能も備えるプラットフォームも数多くあります。これらのプラットフォームは大企業向けに設計される傾向が強く、ほぼリアルタイムのソーシャル・メディア・モニタリング、検索とディスカバリー、トレンド検出、感情分析、データ可視化、データ・アナリティクスなど、データ・ソースを分析するためのさまざまな機能が含まれています。主なデータ・モニタリングおよび分析プラットフォームの例としては、Bottlenose、Brandwatch、DataSift、Engagor、Gnip、Nexalogy、Viralheat があります。

- 最も重要で、最初に解決しなければならない非構造化データの問題は何でしょうか。
- 最も価値が高く、重要なビジネス洞察を提供する非構造化データのタイプはどのようなものでしょうか。
- 企業の成功にとって重要なのはどのような情報でしょうか。

## データ管理の計画

前述のプラットフォームの大半は、企業が非構造化データから洞察を得るためのさまざまな機能を提供する一方で、主に外部のデータ・ソースに重点を置き、比較的高額な場合があります。大企業や中堅企業で生成されるデータの多くは、Eメール、テキスト文書、デジタル画像、スプレッドシート、その他の内部ファイルのデータです。ソーシャル・メディア、非構造化データ、その他の外部データ・ソース (Twitter など) から洞察を得ることは重要ですが、企業は内部の非構造化データのためのデータ管理計画を実施する必要もあります。また、データの保管場所、保管期間、分析方法などを決める必要もあります。

予算、スケジュール、ユース・ケースは企業ごとに異なります。非構造化データの質を高め、そのデータから洞察を得るために API を使用することで、多数の機能を備えたデータ・モニタリングおよび分析プラットフォームを利用するよりもはるかにコストを低くすることができます。また API により、「多目的用途に対応する」ソリューションとは対照的な、カスタマイズされスリム化された手法で非構造化データを処理できます。API を利用すると、データ・モニタリングおよび分析機能を既存のアプリケーションに統合できるだけでなく、企業の特定のニーズに低コストで対応するカスタムの非構造化データ・ソリューションを構築することもできます。

## ストレージ

非構造化データと構造化データに関して、ストレージは重要な考慮事項です。それほど費用のかからない、API を提供するクラウド・ストレージ・プロバイダーは、ここで挙げきれないほど多数存在します。クラウド・ストレージ・サービスは、Amazon、Bitcasa、Box、Dropbox、EMC、Google、Microsoft、Rackspace、Red Hat、Smartfile などによって提供されています。

## データのモニタリングおよび分析

非構造化データのモニタリングおよび分析は、考慮すべき重要なソリューションです。IBM Watson は、ほぼリアルタイムのソーシャル・メディア・モニタリング、感情分析、競合インテリジェンス、抽出機能 (テキスト、キーワード、エンティティ、関係など) を実現するために利用できる API を提供しています。GoodData、Keen Io、Tableau など、API を提供するデータ可視化およびアナリティクス・プラットフォームがあります。

ダッシュボード・プラットフォームは、オンデマンドのデータ・モニタリングおよび分析でよく利用される別の方法です。Cyfe、New Relic、Geckoboard、Leftronic、Klipfolio などの多数のダッシュボード・プラットフォームが API を提供しています。Bottlenose、Brandwatch、DataSift、Engagor、Gnip、Nexalogy、Viralheat といったデータ・モニタリングおよび分析プラットフォームのすべてが同様に API を提供している点に注目すべきです。

## まとめ

上記のソリューションを検討する前に、企業は非構造化データの優先順位や利用目的を把握する必要があります。最も重要で、最初に解決しなければならない非構造化データの問題は何でしょうか。最も価値が高く、重要なビジネス洞察を提供する非構造化データのタイプはどのようなものでしょうか。企業の成功にとって重要なのはどのような情報でしょうか。

多くの企業がこれらの質問への答えを既に手にし、非構造化データを利用して市場での競争優位を得ようとしています。もし貴社がまだ、非構造化データの問題を解決し重要なビジネス洞察を得るためのソリューションを見つけていないのならば、今すぐ探し始めるべきです。競合企業は既に非構造化データを有利に活用しているかもしれません。対応は待たないでください。

## Looking for a partner in data analysis?

Ask the right questions

- ✓ Technology
- ✓ Vendor positioning
- ✓ Customer sentiment

Get the Buyer's Checklist

### Janet Wagner について

Janet Wagner 氏はオハイオ州トレドを拠点とするデータ・ジャーナリストであり、フル・スタック開発者です。彼女の専門は API、データ可視化、機械学習、およびデータ主導型ジャーナリズムです。

Twitter @webcodepro や Google+ でフォローするか、E メールをお送りください。

### IBM Watson について

2014 年 1 月、IBM は、クラウドで提供するコグニティブ・コンピューティング・テクノロジーの開発と商用化に特化したビジネス、IBM Watson ユニットを立ち上げました。この動きは、学習によって向上し、膨大なビッグデータから洞察を引き出す新しいクラスのソフトウェア、サービス、アプリケーションの提供に向けた IBM の戦略的変更を示すものです。IBM Watson の詳細については、[ibm.co/watsonecosystem](http://ibm.co/watsonecosystem) をご覧ください。

© Copyright IBM Corporation 2015

IBM Corporation  
Software Group  
Route 100  
Somers, NY 10589

Produced in the United States of America  
October 2015

IBM、IBM ロゴ、ibm.com および 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、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

本書の情報は最初の発行日の時点で得られるものであり、予告なしに変更される場合があります。本書の情報は最初の発行日の時点で得られるものであり、予告なしに変更される場合があります。すべての製品が、IBM が営業を行っているすべての国において利用可能なものではありません。

本書に掲載されている情報は特定物として現存するままの状態を提供され、第三者の権利の侵害の保証、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任なしで提供されています。

IBM 製品は、IBM 所定の契約書の条項に基づき保証されます。

適切なセキュリティ実施について: IT システム・セキュリティには、企業内外からの不正アクセスの防止、検出、および対応によって、システムや情報を保護することが求められます。不正アクセスにより、情報の改ざん、破壊もしくは悪用を招くおそれがあり、またはシステムの損傷や、他のシステムへの攻撃を含む悪用につながるおそれがあります。完全に安全と見なすことができる IT システムまたは IT 製品は存在せず、また単一の製品またはセキュリティ対策が、不正アクセスを防止する上で、完全に有効となることもありません。IBM のシステムおよび製品は、包括的なセキュリティの取り組みの一部となるように設計されており、これらには必ず追加の運用手順が伴います。また、最高の効果を得るために、他のシステム、製品、またはサービスを必要とする場合があります。IBM は、何者かの悪意のある行為または違法行為によって、システム、製品、またはサービスのいずれも影響を受けないことを保証していません。



Please Recycle