



**構造化データは良いが、
非構造化データが扱えない**



**ビッグデータソリューションは多くあるが、
人がデータを解釈するのと同じレベルで扱いきれていない**



**ビジネスの現場で使いたいように
各種データを自由に使えていない**

ソーシャル メディアデータ



ソーシャルメディアにおいて参加者が書き込むもの

マルチメディア データ



WEB上の配信サイト等で提供される音声や動画

ウェブサイト データ



ECサイトやブログ等で蓄積される購入履歴やブログエントリー等

カスタマー データ



CRMシステムで管理されるDM等販促データや会員カードデータ等

ビッグデータ



センサーデータ



GPS・ICカード・RFID等において検知される、乗車履歴、温度、加速度等

オフィス データ



オフィスのPCで作成されるオフィス文書やEメール



ログデータ

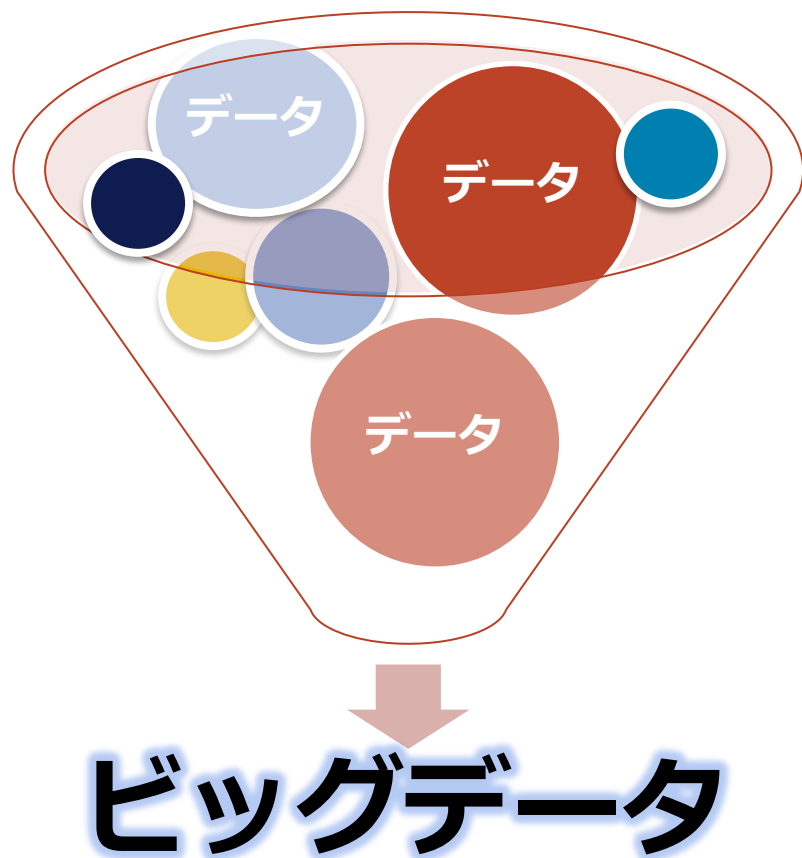
WEBサーバにおいて自動生成されるアクセスログ、エラーログ等

オペレーション データ



販売管理等業務システムにおいて生成されるPOSデータや取引データ等

出所：総務省公開資料を元に弊社作成



ビジネス面

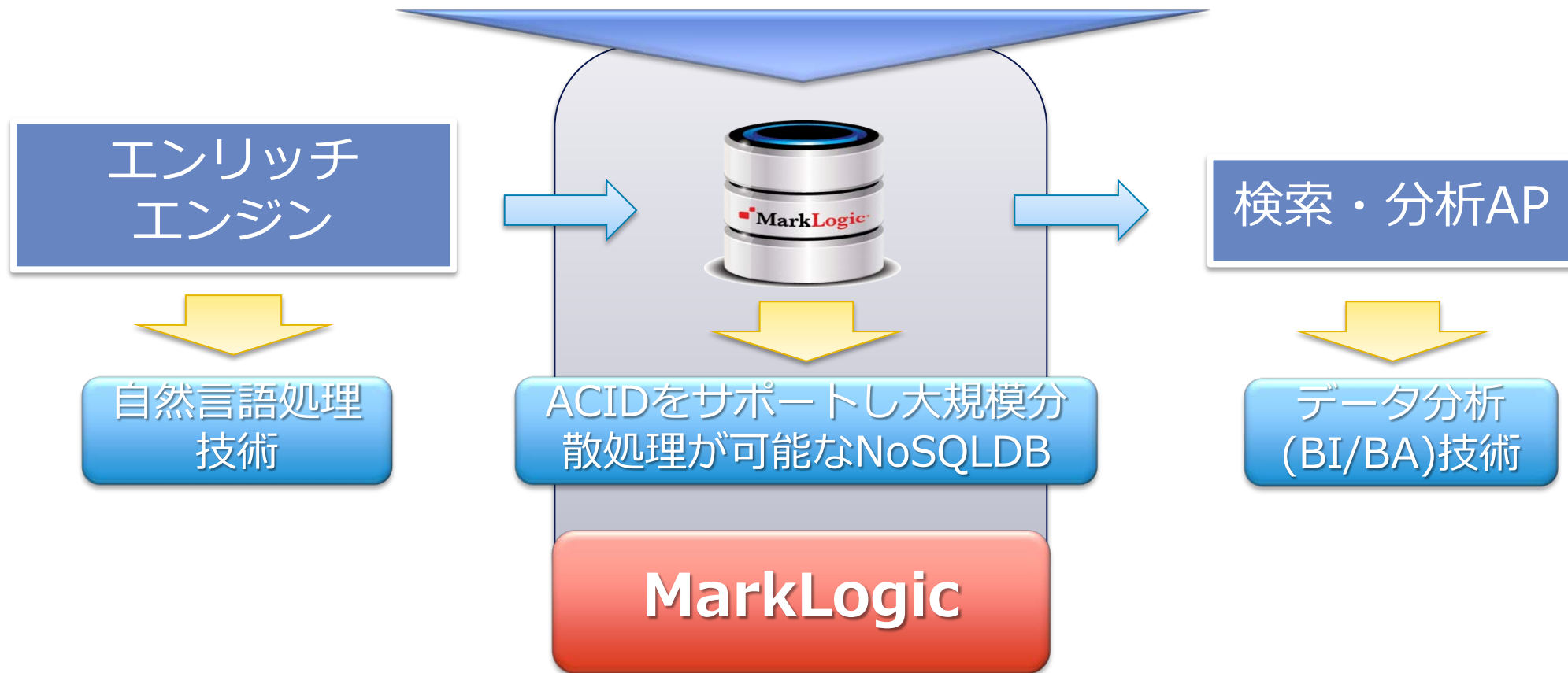
多量のみならず、むしろ多様なデータをまとめて扱うことで意味が出ることが多いのではないか

ハードルが高く、なかなか実現しなかった

技術面

多様なデータを扱うにあたり、データベースの制約や、データ自体を取扱う周辺技術が未整備

「非構造化データ活用基盤」として、従来のRDBが管理する定型化されたデータに加え、「当初には想定されていないデータ構造」や「データ構造の変更」、「テキスト文書などの非定型データ」をリアルタイムに管理し、データの意味を踏まえた検索やデータ間の複雑な関連を管理することができる、柔軟で拡張性に優れたデータ管理基盤のコア要素としてのMarkLogic Serverの活用を行います。



インターネット上のECサイトや掲示板等、審査先に関するあらゆる情報を意味解釈・タグ付けを行い、多様なデータを活用可能な状態でMarkLogic Serverに格納し、お客様審査業務の高度化を図る。

インターネット情報



加盟店



加盟店ホームページ

インターネット情報を収集
(クローリング)



ソーシャル情報他

Twitter

不特定多数



加盟店ホームページ

(加盟店が発信する情報)

- ・ 違法商材の有無
- ・ 特商法表記
- ・ URLリンク切れ 等

ソーシャル情報

(加盟店のコントロール外の情報)

- ・ 口コミ (Twitter・掲示板等)
- ・ 地図情報 (Google Map)
- ・ 不芳ニュース 等

加盟店発信情報に加えて、不特定多数によるソーシャル情報も検索

大量情報の収集手間の自動化による業務効率化



- ・ 加盟店情報の参照
- ・ 評判の参照

- 不芳加盟店の検出 (審査トリガー)
- ・ 加盟店情報の参照
- ・ 評判の参照

途上管理

