

ビッグデータの解析処理構築セミナー  
～データ収集蓄積から解析レポートまで  
低コスト短時間で実現するには～

**(セッション1)**  
**ビッグデータを活かすETLおよびBIツールの実力**

2015年7月2日  
株式会社エアー  
プロダクト・カンパニー  
第二営業部 兼本侑始

# 株式会社エアー：会社概要

●所在地 【本社】 〒565-0851 大阪府吹田市千里山西5-31-20  
【東京】 〒106-0032 東京都港区六本木2-2-8 KDXビル

●設立 1983年8月8日 ●代表者 代表取締役：北山洋一

●資本金 80,705,310円 ●売上高 9億6千万円（31期）

●従業員数 64名（2014年4月現在）

●事業内容 【IT関連製品の開発・販売、サービスの提供】

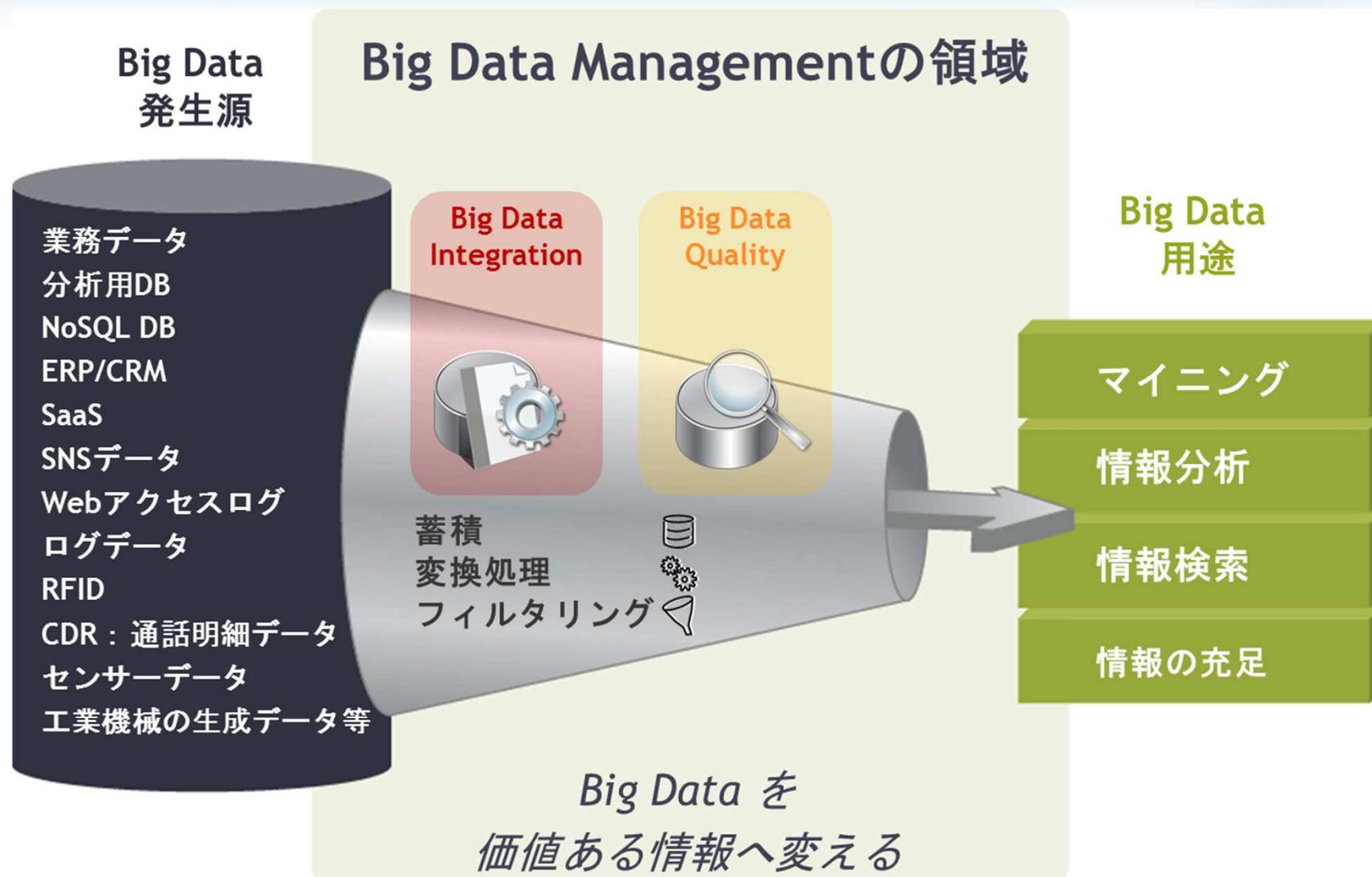
- メッセージングソリューション（WISE Walker / WISE Sync / CommuniGate Pro）
- メールセキュリティソリューション（WISE Audit / WISE Attach / WISE Alert）
- 仮想化ソリューション（Propalms TSE / Oracle SGD）
- 仮想化環境における印刷問題解決ツール（ScrewDrivers / Simplify Printing）
- 印刷インク/トナーコスト削減ツール（WISE Print）
- 印刷セキュリティソリューション（WISE Print 印刷ログ監査）
- **ビックデータ活用（Talend / Actuate）**
- 検索可能な暗号化ライブラリ（WISE Encrypt）

【ソフトウェア受託開発、システム構築サービス】

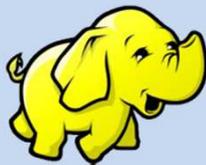
- IT関連受託開発
- コンサルティング業務
- システムインテグレーション



# ビッグデータを活用するために



# エアーのビックデータソリューション



Hadoop



Webサービス



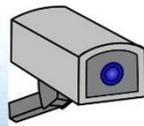
RDBMS



ログファイル



CSV, Excel



センサー (IoT)



データの整備・統合



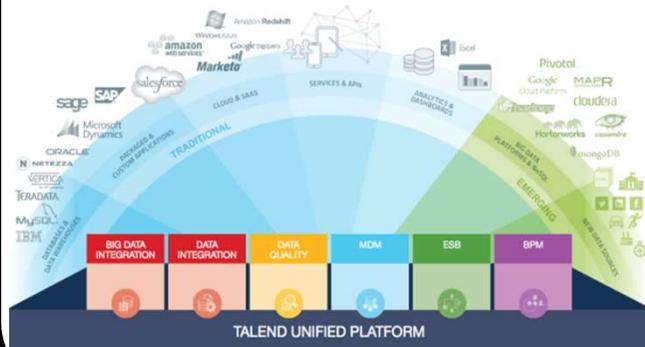
ビックデータ対応



アプリケーション連携



マスタデータ管理



## OPENTEXT<sup>™</sup> ANALYTICS



レポート



モバイル対応



Excel出力



ダッシュボード



分析



スケーラビリティ



セキュリティ



マルチソース



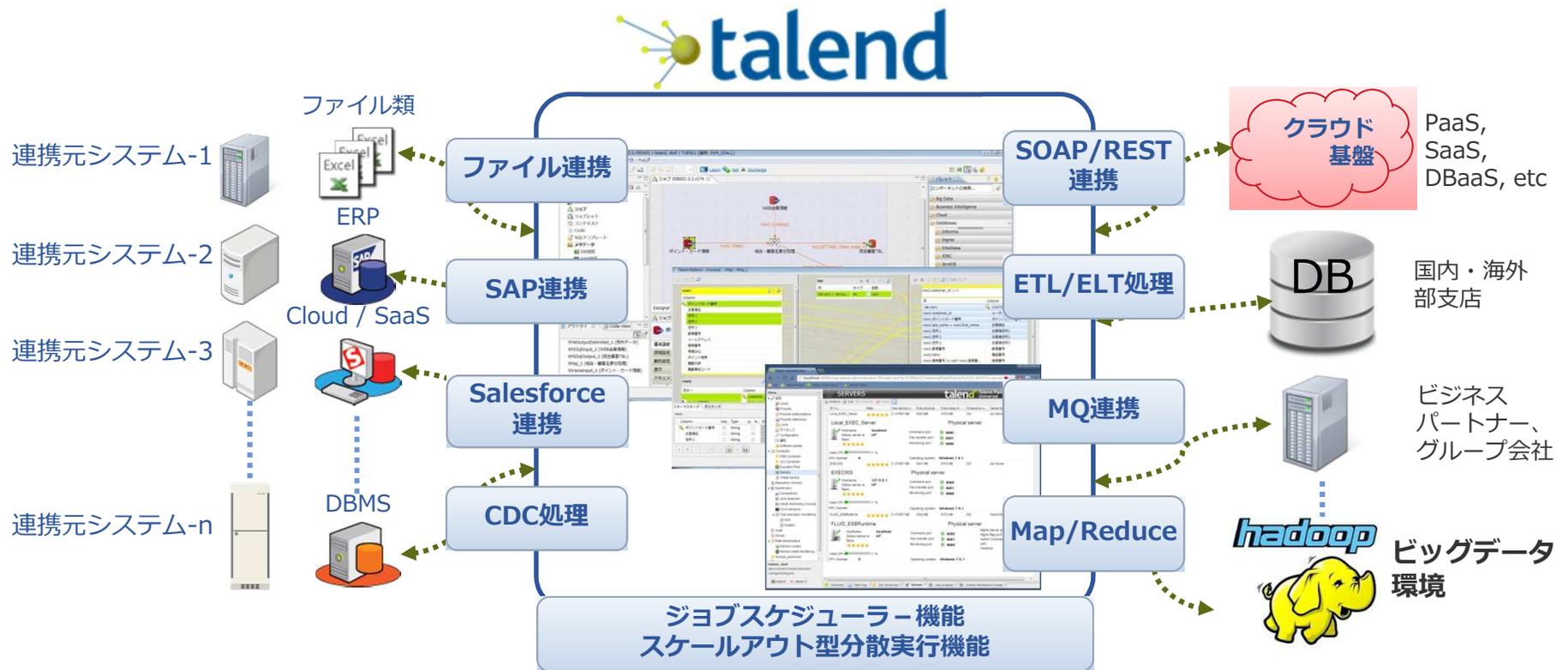
# データ連携ソリューション



AIR  
COMPANY LIMITED

# Talend : ビッグデータ統合・連携プラットフォーム

開発・運用・維持管理までEnd-to-Endでサポートする統合連携プラットフォーム



# 他社が真似できないTalendの優位性

## オープン性

- オープンソースであるが故に情報ソースへのコネクトが豊富(800以上)
- ソースコード生成型モデルであるためコードに透明性があり信頼性がある
- 多くのデータ活用ベンダー(技術)でも接続可能
- オープンソースモデルであるため、先進技術への対応が早い(先進性)

## 先進性

- YARNなどの先進技術への対応
- 先進的な情報ソース(Hadoop、No SQL、DWH、インメモリーコンピューティングなど)へのアクセス
- オープンソースをベースとした開発者コミュニティに支えられているため、最新技術への対応が早い
- コミュニティの支持によりコードの信頼性(製品の信頼性)が高い

## ビジネスモデル

- 開発者単位でのライセンス
  - 顧客は最小限のライセンスを支払えば良く、投資対効果が高い
- サブスクリプションライセンス
  - 顧客は年単位でプロジェクトを見直してライセンスコストの最適化を図れる
- ソースコード生成型で分散処理環境に迅速に対応

# Talend Studio - ジョブ開発環境

The screenshot displays the Talend Studio interface for a job named 'ジョブ JOB001 0.1.r174'. The main workspace shows a data flow diagram with the following components and connections:

- Input:** 'ポイント・カード情報' (Point Card Information) feeds into 'row1 (Main)'.
- Join:** 'row1 (Main)' and 'row2 (Lookup)' (from 'WEB会員情報') connect to a join component '結合・顧客名寄せ処理'.
- Output:** The join component outputs to 'toCUSTTABL (Main order:1)' and '例外データ' (Exception Data) via 'reject (Main order:2)'.
- Target:** 'toCUSTTABL (Main order:1)' connects to the '統合顧客TBL' (Integrated Customer Table).

The configuration window for 'ポイント・カード情報(tOracleInput\_1)' is open, showing the following details:

- Host:** context.FLUIDORA\_Server
- Database:** context.FLUIDORA\_ServiceName
- User:** context.FLUIDORA\_Login
- Schema:** 組み込み (Built-in)
- Table:** "ポイントカード情報"

The 'スキーマエディタ' (Schema Editor) is also open, comparing the source schema 'row1' with the target schema 'toCUSTTABL':

Column	Key	Type	N.	日付パターン...	Length	Precis...	Def...	Com...
ポイントカード番号	<input checked="" type="checkbox"/>	String			11	0		
お客様名	<input type="checkbox"/>	String			45	0		
住所1	<input type="checkbox"/>	String			45	0		

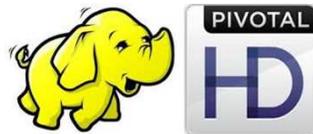
Column	Key	Type	N.	日付パターン...	Length	Precis...	Def...	Com...
CUSTID	<input checked="" type="checkbox"/>	Int					0	
ユーザーID	<input type="checkbox"/>	Integer					0	
ポイントカード番号	<input type="checkbox"/>	String			11	0		

# サポートするビッグデータ関連技術

## No SQL DB



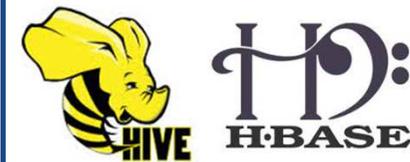
## Hadoop



## インメモリーDB、DWH



## Technology



# リアルタイムビッグデータへの対応



ストリーミングデータ用新規コンポーネント

Spark



**100x**

性能向上

**< 1 sec**

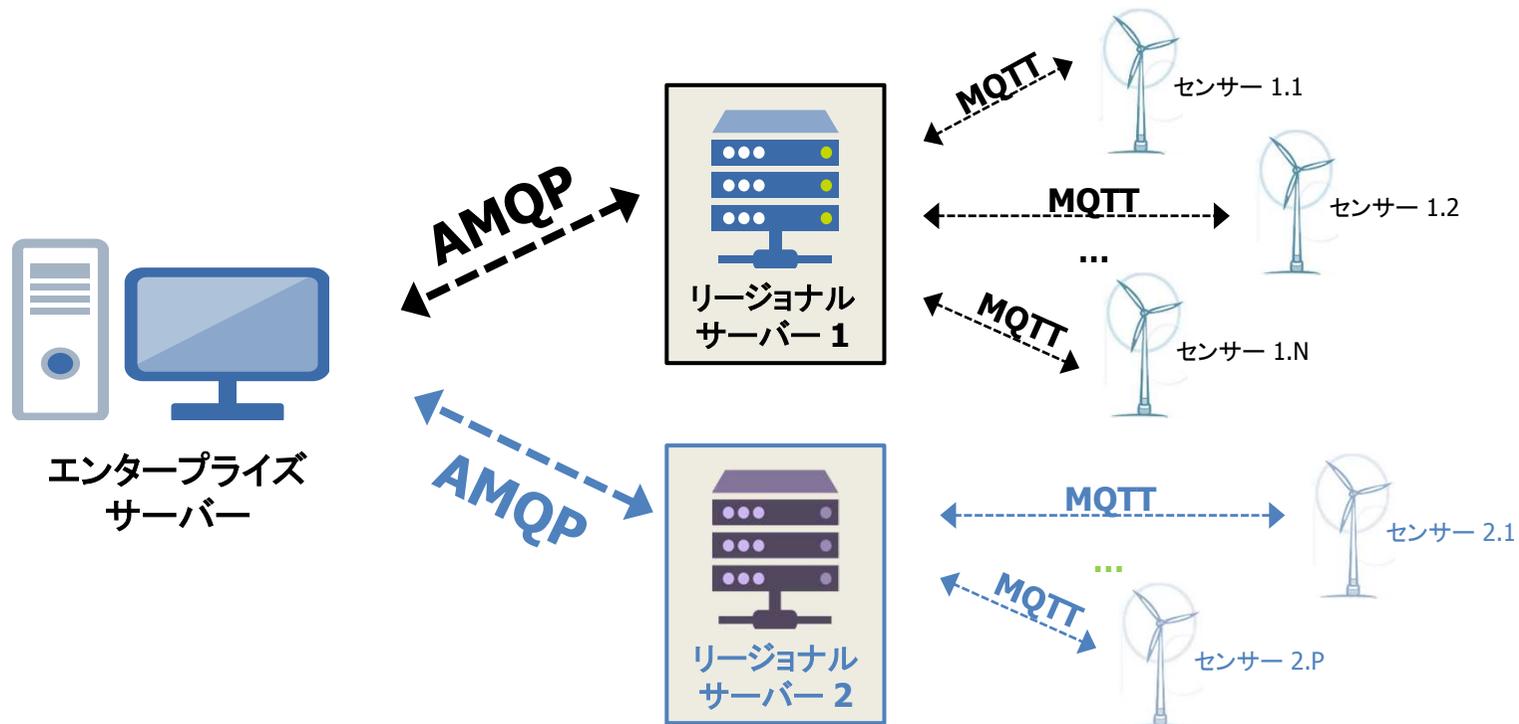
応答時間

テクニカルプレビュー版  
No SLAサポート

↓  
新ユースケースの実現

(顧客離反の防御、ダイナミックプライシング、  
リアルタイム違法取引検知、CEP,等)

# IoTプロトコルサポート



IoT用メッセージングプロトコル

- MQTT(デバイス&サーバ間)
- AMQP(サーバ&サーバ間)



**故障検知**

**在庫削減**

**顧客満足の上**

# Talendのビックデータ顧客事例

## ■ GE社(アメリカ) 大手エネルギー関連機器メーカー

「インダストリアルインターネット」: ネットワーク接続されたセンサーとウェアの統合により、数十億円のエネルギー増産を達成

### 課題:

センサーデータの収集、分析に30日かかり、タイムリーなダウンタイムの予測が困難。またエネルギーの増産が容易ではない。

### 目標:

障害が発生する前に、タイムリーなダウンタイム予測をし、タービンのピッチと方向を調整してエネルギー増産につなげたい。

### ソリューション:

全世界の22,000台以上の風力タービンをスマートセンサーとクラウドを使ってリアルタイム接続。Talend Big Dataを使ったデータ収集を行い分析へ繋げる。

### 結果:

30日かかっていたセンサーデータの収集、分析がTalendを利用することにより1日で完了し、タイムリーなダウンタイム予測、タービンのピッチと方向を調整することが可能になった。この結果、数十億円規模のエネルギー増産を達成。



# Talend: ライセンスおよび技術サービス

- 商用ライセンス
  - ◆ 年間サブスクリプション
    - Talend Enterprise Data Integration 3ユーザ ... ¥4,900,000/年間
    - Talend Enterprise Data Integration 3ユーザ ... ¥5,600,000/年間
  
- 有償トレーニング
  - ◆ Talend Open Studioトレーニング (¥160,000/2日間/1名)
  - ◆ Talend 運用管理者トレーニング (¥80,000/1日間/1名)
  
- オープンソース版Talend向け有償技術支援サービス
  - ◆ オープンソースライセンスの範囲で、システム構築および運用されるお客様を対象に弊社が技術支援を実施
  - ◆ チケット制(¥300,000/5チケット/年間～)
  
- 導入支援、ジョブ開発支援、受託開発など



# エンタープライズ レポートニングソリューション

## OpenText Analytics

# パーソナライズされた分析

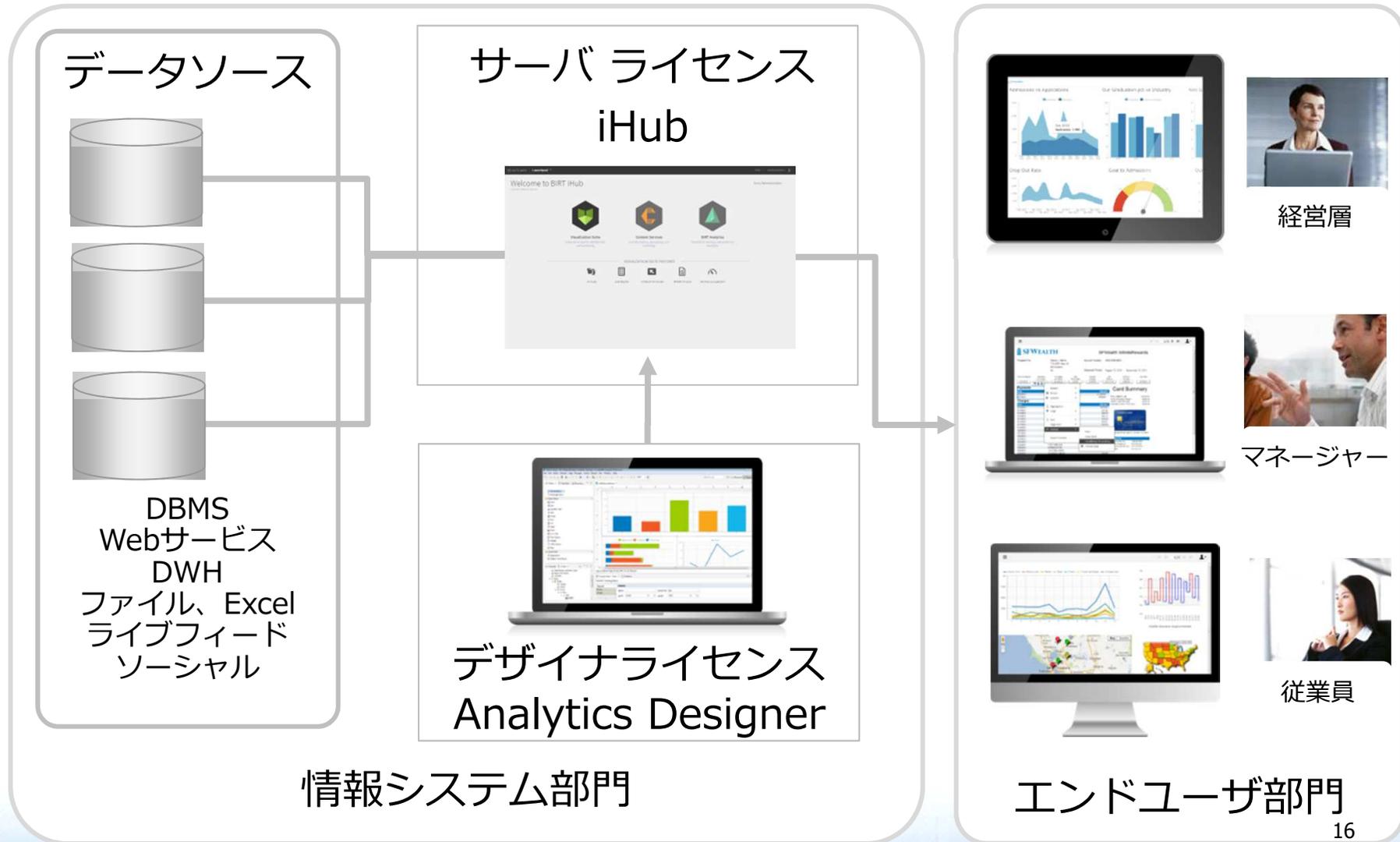
データソース



## Information Hub



# シンプルなアーキテクチャ



# Analytics Designer: 開発環境

## BIRTデザイン開発環境

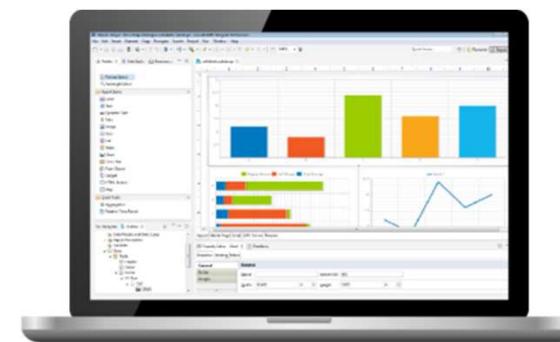
- オープンソースEclipse 標準プラグインがベース
- Eclipse BIRT : オープンソース版
- Analytics Designer : 商用版
- 直感的なGUI操作
- ライブラリによる部品共通化
- HTML 5、JavaScript、CSS

## 開発者コミュニティ

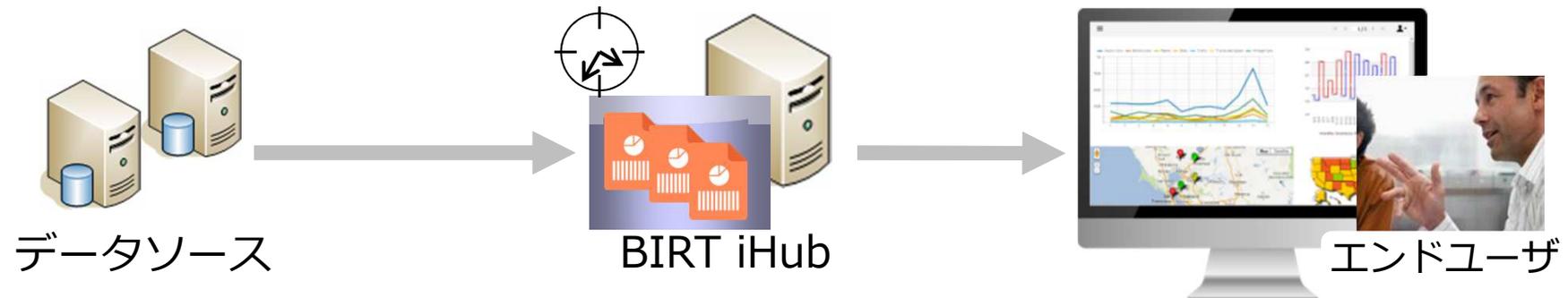
- ワールドワイドで350万以上の開発者
- コミュニティ : BIRT Developer Center

## Forrester Waveによる高い評価

- オープンソースBI分野のLeader



# Information Hub (iHub): ビジネス分析プラットフォーム



- 単一インスタンス (シングルインストール)
- コンテンツ (レポート、ダッシュボード) の生成と配布
- コンテンツ生成のスケジューリング、バージョン管理、アーカイブ
- ユーザ利用状況のロギング
- セキュリティ設定 (コンテンツ、ページ、データ単位でのアクセス制御)
- クラスタリング構成 (ロードバランシング、フェイルオーバー)
- プロセス分散による高パフォーマンスと高可用性
- オープンアーキテクチャ (SOAP、JAVA)
- APIによる外部システムとの連携

# Information Hub 機能紹介 (標準機能)

## 定型レポート、定型ダッシュボード

- プラグインやアドオンを必要としないピュア・ブラウザでの操作
- ユーザ認証、アクセス権限、アクセスログ
- 定型レポートや定型ダッシュボードの参照、アドホックレポートの実行
- レポートのエクスポート (PDF、Excel、PowerPoint 等)
- 高い表現力、容易な操作性

## マッシュアップによる他システムとの連携

- iHubのコンテンツ要素を、容易に既存システムに組み込み
- YouTube、Google Map、Flashなどの外部アプリをレポートに組み込み
- HTML、JavaScript、Flashといった標準インターフェース

## モバイル対応

- ネイティブブラウザ対応
- REST API + JavaScriptによるカスタム MobileAPP開発環境



# Interactive Viewer(オプション)

## レポートカスタマイズ、分析

- 参照しているレポートに対して、直接カスタマイズや分析操作が可能
- ソート、フィルタ、グルーピング、集計、グラフの操作
- サーバにアップロードすることで、カスタマイズしたレポートを共有
- アップロードしたレポートを定期的なバッチ実行として運用



# Big Data Analytics

OpenText Actuate Big Data Analyticsは、  
ビジネスユーザーを予測分析の達人に変身させる他に類を見ない  
ビッグデータ分析ソリューションです。

このソリューションでは、すべてのデータを統合し、  
データサイエンティストでなくてもすぐにデータ分析を実施するこ  
とができるセルフサービスモデルを用いています。

## 大規模データ統合

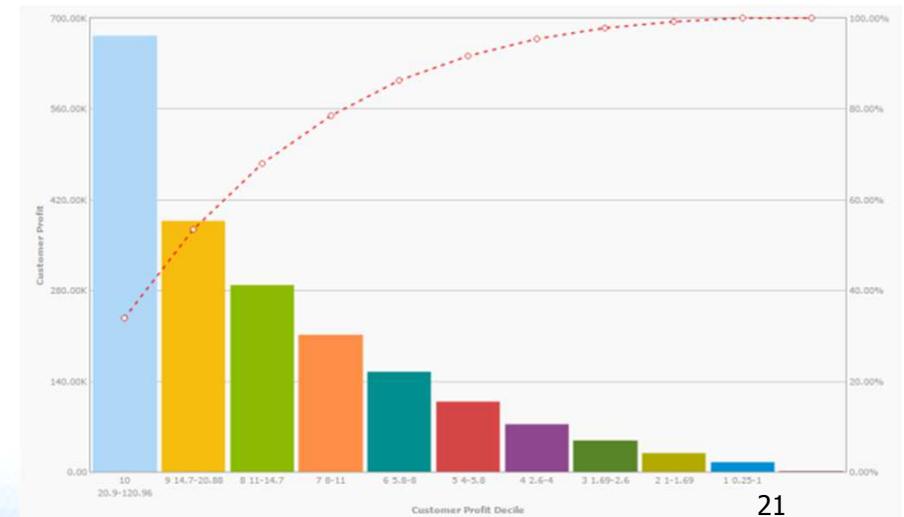
- 複数のソーシャル、インターネットソース
- 膨大なレコード数

## ビジネスユーザーのために

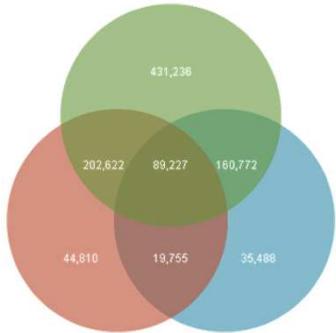
- ITスタッフの関与は、ほとんど不要
- 分析に精通しているユーザーのための高速ターンアラウンド

## ビジュアルデータディスカバリーとマイニング

- ベン図、パレート図、分布図
- プロファイル、相関分析、クラスター分析
- 時系列予測
- その他、主要なアルゴリズムを網羅



# 新しい分析ビジュアライゼーション



**埋もれた関連性の発見**  
クロス-バジェットिंग  
営業機会

優先度	項目	グループ別	売り	インデックス	コスト	パフォーマンス
3000未満	金融サービス	20,024	83,898	117.79	20.77	
E-Mail	金融サービス	47,969	231,546	102.24	5.02	
Bungalow	金融サービス	21,075	102,889	101.09	1.51	
既婚	金融サービス	60,129	295,081	100.57	1.50	
専業主婦	金融サービス	15,405	75,498	101.09	1.28	
350万～500万	金融サービス	41,488	203,647	100.54	1.13	
Terraced	金融サービス	46,403	227,949	100.47	1.04	
	金融サービス	1,499	7,212	102.58	0.89	
上級管理職	金融サービス	7,407	36,149	101.13	0.89	
販売員	金融サービス	11,899	58,099	100.82	0.84	
役員	金融サービス	40,917	201,188	100.37	0.77	
死別	金融サービス	14,858	72,915	100.57	0.65	
既婚	金融サービス	29,943	147,284	100.34	0.57	
自営業	金融サービス	15,279	75,117	100.39	0.45	
W	金融サービス	136,345	682,493	100.04	0.25	
500万～750万	金融サービス	20,721	102,394	100.06	0.20	
事務職	金融サービス	20,113	100,199	100.05	0.07	
75歳以上	金融サービス	171	847	99.64	-0.04	
100万～200万	金融サービス	50,373	248,805	99.92	-0.19	
既婚	金融サービス	83,064	314,456	99.92	-0.23	
100万未満	金融サービス	24,143	119,409	99.79	-0.32	
作家	金融サービス	140,730	724,509	99.95	-0.32	

## 要因検出

なぜそれは起こったのか?  
最も優れたプロジェクトはどのようなものか?  
最も識別できるものは何か?

方向	名前	リンク	合計	%リンク	%リンク
←	[金融サービス][口座]		1,000,000	1,000,000	100.00
←	[金融サービス][コンタクト履歴]		189,344	1,000,000	18.93
←	[金融サービス][レスポンス]		3,620	1,000,000	0.36
←	[金融サービス][キャンペーン]		90,555	1,000,000	9.06

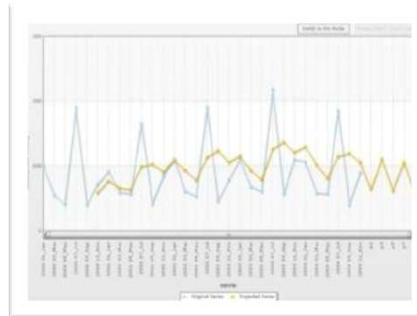
## 相関分析 - 原因と結果

関連性や類似性パターンの検出  
どのプロジェクトが成功し、  
予算内に収まったか?



## 行動の特異性

どのような振る舞いか?  
雨の日には人々はどのような  
行動をするか?



## 将来予測

時系列分析  
消費、販売数、  
新しいプロジェクトの予測



## クラスタリング

より良い分類  
利益とポテンシャルリスクに  
よってグループ化された  
トランザクションは  
どのようなものか?

# OpenText: ライセンス

## ■ サーバ環境

製品名	説明	価格
Online	レポートのオンデマンド生成と表示	¥600,000/プロセス/年間 ※最小2プロセス～
Factory	レポートのバッチ生成	¥2,400,000/プロセス/年間 ※最小1プロセス～
Interactive Viewer	レポートカスタマイズ	¥600,000/プロセス/年間 ※最小2プロセス～
Dashboard	ダッシュボードの動的作成と表示	¥600,000/プロセス/年間 ※最小4プロセス～

## ■ 開発環境

製品名	説明	価格
Analytics Designer	レポート開発環境	¥30,000/開発者/年間

## ■ ビックデータ分析

製品名	説明	価格
Big Data Analytics	オンメモリの高速データ分析	¥1,800,000/年間～ ※100,000,000レコードまで分析可能

ご清聴ありがとうございました。

株式会社エアー  
プロダクト・カンパニー  
第二営業部  
兼本侑始  
03-3587-9221(代表)  
kanemoto@air.co.jp

AIR  
COMPANY LIMITED