



81kHz



56kHz



29kHz



82kHz



55kHz



93kHz



101kHz

  
informatica

# インテリジェント データウェアハウス

マスターデータ管理を活用して  
インテリジェントな  
ビジネスインテリジェンスを実現

# インテリジェンスが もたらす価値

企業で生み出されるデータ量が増えるほど、ビジネスインテリジェンスシステムが企業の意思決定に及ぼす影響は大きくなります。また、競争力のある資産としてデータの有用性が高まるにつれて、ビジネスインテリジェンスの重要性はかつないほど高まっています。

しかし、BIソフトウェア、データウェアハウス、データ統合プロジェクトに対して相応の投資をしているにも関わらず、アナリスト部門はいまだに高い精度のレポートをタイムリーに提供できていないのが現状です。それは、なぜなのでしょう。扱うデータが整理されていなければスピードは失われ、提供するレポートに対する信頼性も低下してしまいます。

これが、ビジネスインテリジェンスを謳いながら、その実態は全くインテリジェントではないという結果を招いています。



多くの企業が業務システムで利用しているデータは精度が低く、一貫性がなく、また多くが重複しているため、自社の顧客、サプライヤー、商品、従業員の正確な情報をデータから読み取ることができなくなります。

結果として、レポートとそこから導かれる意思決定もまた、不正確で信頼性に欠けるものになります。レポートとデータをよりどころにしているユーザーが聞いたら、大きなショックを受けることでしょう。

# 低品質のデータの影響からは誰も逃れられない

**経営者**は、レポートの精度が低い  
ため、時間の制約がある中で、正しい  
戦略的な意思決定を下すことができ  
なくなります。

**ビジネスアナリスト**は、ほとんどの  
時間を手作業でのデータ再同期に  
費やしているため、販売実績、収益  
配分、リスク緩和を適切に分析する  
ための時間を確保することができま  
せん。

**IT技術者とデータアーキテクト**は、  
ミスの発生しやすい手作業での数多  
くのプロジェクトによって高品質で  
正確な標準化されたデータを作成し  
なければならず、その作業は困難を  
極めます。また、合併や買収といっ  
た大きな変化が発生すれば、時間の  
かかる悪夢のような日々が延々と続  
くこととなります。

**コンプライアンス担当者**は、規制  
当局に対して正確さを保証できる  
レポートを提出することができま  
せん。これは本質的に管理上の問題  
であり、十分に回避可能であるにも  
関わらず、会社に巨額の制裁金が科  
せられる危険性があることを意味し  
ています。

本電子ブックは、データをどのように準備すれば、ビジネスインテリジェンスイニシアチブから最大限の価値を引き出すことができるのかについて説明します。このままの状態が続けば、低品質なデータによる低品質なインテリジェンスのために、正しい意思決定ができないという悪循環に陥ることでしょう。

データウェアハウスが  
これほど乱雑になって  
しまった理由は？

```
CS211 Beginning PASCAL MWF Allow from .hp.com <STRONG>all</STRONG> document requests.
CS314 Compiler Design TR Require group Wise According to <CITE>Corporate Policy 42-A-1</CITE>:
CS451 Advanced UNIX MWF <BLOCKQUOTE>Documents can only be released from the
</PRE> <H3>Description List</H3> </Limit> accompanied by appropriate payment in the form of
order.</BLOCKQUOTE><P>
Fill out your request as in this example:<P>
ns. <SAMP>Introduction to Standards</SAMP><BR>
<SAMP>Document #15226</SAMP><BR>
<SAMP>$425.00</SAMP><P>
t. Be sure to include the appropriate <VAR>document n
<HR> 29kHz
p -i $query $DATABASE ; >>>
{ print "Search was unsuccessful\n" ; } 29kHz
answer\n" ; }
failed
default /top/homepage.html
ect /box/upper_left.html 0,0 199,99
ect /box/upper_right.html 200,0 399,99
ect /box/lower_left.html 0,100 199,199
rect /box/lower_right.html 200,100 399,
29kHz
M1G)I9&%Y+"!397!T96UB97(@,C,L(#$Y.30@,3<Z-3
M;VX@,"XY,BXV#0H-"E1H:7,@:7,@86X@3DY44"!N97
M:6-R;W-09G0@5VEN9&]W<R`S+C$@;W(@5VEN9&]W<F
M('5S92!I="!T;R!R96%D(&%N9"!P;W-T(%5S96YE='
M9"!E;6%I;"`H=FEA(%—5%'@;W(@34%022D-"@T*4F
M#0I&;W(@5TE.4T]#2R`H=C$N,"!0<B!H:6=H97(I(&
<Limit GET PUT POST>
order deny,allow
deny from all
allow from 167.142
</Limit>
if (length($query)>0) {
print <<EOM;
Search for <B>$query</B> yields these entr
```

ビジネスをとりまくデータの様相は、かつてBIシステムを構築した当時のものとは大きく様変わりしています。しかし、データウェアハウスはカオス化した現在のデータを扱うための仕組みを持っていません。そのため、次の2つの課題に直面しています。



## 1

### 複数のソースから流れ込むデータ

企業データは、数十、場合によっては数百ものシステム上に存在していますが、様々な部署の様々なユーザーが、それぞれの顧客やアカウント、製品に関する情報をこれらのシステムに入力するため、データの形式も精度のレベルも全く統一性のないものになります。データは重複し、一貫性がなく、正確さにも欠けますが、データウェアハウスにはこうした状態を解消するための仕組みがありません。

これが国内の組織だけの問題であっても解決は容易なことではありませんが、実際には世界中の組織が関わる非常に複雑で大規模な問題です。また、一般的にBIツールはこのような複雑な事態を想定して構築されたものではないので、不完全な部分はいつもビジネスアナリストが引き受けることになります。

さらに悪いことに、ビジネスアナリストに課せられた膨大な作業は、問題を解決するものではありません。このため、どれほど努力をしても、また次回レポートを作成するときには同じようなデータの混乱に悩まされるのです。

## 2

### データの非一貫性と重複

一般的にBIシステムはGEとGeneral Electric社の相違を認識するようには構築されていないため、ソースシステムにおけるデータの非一貫性はBIシステムにそのまま伝わります。したがってデータに間違いがあれば、間違ったレポートが生成されてしまいます。

財務諸表の精度は、ある顧客との取引が別々のシステムで処理された場合でも、正しい照合が行われていることを前提としています。データを正確にまとめることができなければ、分析（及び会計）業務を的確に行うことができません。

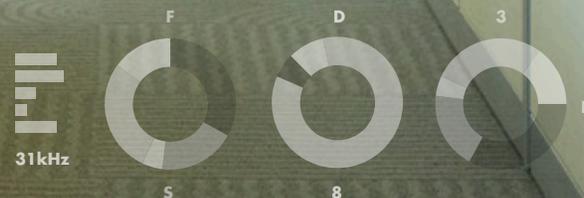
顧客についての一貫したデータがなければ、営業面でもマーケティング面でもマイナスです。サプライヤーについての一貫したデータがなければ、サプライチェーンは非効率的になります。商品についての一貫したデータがなければ、製品展開のタイミングが遅れてしまいます。

ビジネスインテリジェンスは、ビジネスに対する確固たる評価が信頼できるものであることが前提となっています。しかし、データを信頼することができなければ、インテリジェンスを信頼することができません。



 **General Electric**

# この混乱が もたらす障害



信頼できる一元化された情報源がなければ、インテリジェンスを必要とするすべての部署にとって障害となります。

## 営業部門

競争の激しい市場において、営業部門は、顧客獲得に役立つあらゆる情報を知りたいと考えています。必然的に、情報不足は競争力の著しい低下を招きます。

例えば、ブラジルに本社のある化粧品製造・小売企業であるNatura社は、自社の営業部門が、顧客との関係ややり取りについて、正確な情報を持っていないことに気が付きました。

## マーケティング部門

マーケティング部門がリソースを戦略的に割り当てるには、顧客ベースや見込客を様々な属性によって分類する必要があります。さらに重要な点として、各マーケティングキャンペーンの成功度合いを追跡管理する必要があります。

ハイテク企業のCitrix社はチャネルパートナー、顧客、見込客についての質の高いデータがなく、salesforce.com、Marketo、MyCitrix.com上にこれらのディメンションの情報が集約されていなかったため、どれだけのリードが実際の商談に結び付いたのかを把握できていませんでした。

## 人事部門

HP社では、同社のシステムに保存された顧客、パートナー、営業担当者に関するデータが断片化、重複していたため、人事部門から営業担当者に誤って報酬を支払うという事態が発生しました。

従業員データが一元化されていないことによる代償は高くつきます。特に、従業員への報酬がビジネスに直接結びついている場合は顕著です。

## 財務部門

ある大手バイオテクノロジー企業は、同社の官公庁向け料金を設定する際に、業務内容や要件についての過去の情報が信頼できないものであることに気が付きました。顧客との関係についての正確な情報を提供するレポートを利用できないことは、料金の再交渉で直接的な障害になりました。

大局的に見れば、重複や矛盾を抱えたままでは、財務諸表の完全性が損なわれることとなります。信頼性の欠如は、やがて経営者の意思決定の質を低下させることとなります。この意思決定は、企業の財政状態の健全性に対する視点と常に結びついています。

## コンプライアンス

米国連邦政府と州の条例に準拠するため、企業は事業を追跡管理し、報告しなければなりません。結果として、罰則や刑罰を回避するためには、かつてないレベルのデータの正確さが必要になります。

例えば、製薬会社の依頼に応じて臨床試験を実施する医薬品の研究業務受託機関であるQuintiles社は、米食品医薬品局（FDA）による審査に常に適合しなければなりません。適合しない場合、FDAによる医薬品の承認は大幅に遅れ、多額の損失につながります。

規制の厳しさは業界によって異なりますが、コンプライアンス報告の義務が免除される業界はありません。また確かに、これらの義務を当然のこととして受け入れるだけの余裕がある企業もありません。

## 研究開発

ダブリンに拠点を置く研究業務受託機関であるICON社は、臨床試験の研究結果を随時、顧客である製薬会社に報告する必要があります。この報告は、試験の効果と、有害な事象があった場合にはその警告を顧客に伝えるものです。

治験責任医師、サイト、ロケーションなどの重要な要素についての情報が不正確・不完全であれば、顧客企業は治験のリスクを監視し、治験責任医師の業務を追跡することも、正確なレポートを作成することもできません。

## その他

**BIシステムが正しい方法で正しい情報を扱っていないければ、正確なデータに依存するリスク、収益、販売、供給に対する分析、報告、意思決定にとっての大きな障害となります。**

# DIとDQでは なぜ不十分なのか

089FG

12AFG

12AFG

0A33



データウェアハウスの構築では、当然ながらデータ統合およびデータ品質という**2つの重要なツール**に依存することになります。これらのツールは大きなメリットをもたらしますが、残念ながらビジネスインテリジェンスをよりスマートにするものではありません。その理由は、以下の通りです。

## 1

### データ統合

データ統合は様々なソースのデータを単一のシステムにまとめるために重要ですが、これはデータからデータウェアハウスへの一方向の接続であり、BIシステムによって制御されるものです。BIシステムはデータの調整を目的として構築されたものではありません。

さらに重要なことは、統合はハードコーディングされているということです。このため、スキーマの変更や次元の更新に際して、ETL開発者は膨大な量の手作業に追われることになります。膨大な量のスクリプトを調べて変更を行い、これらの次元のリネージと履歴を追跡しなければなりません。

## 2

### データ品質

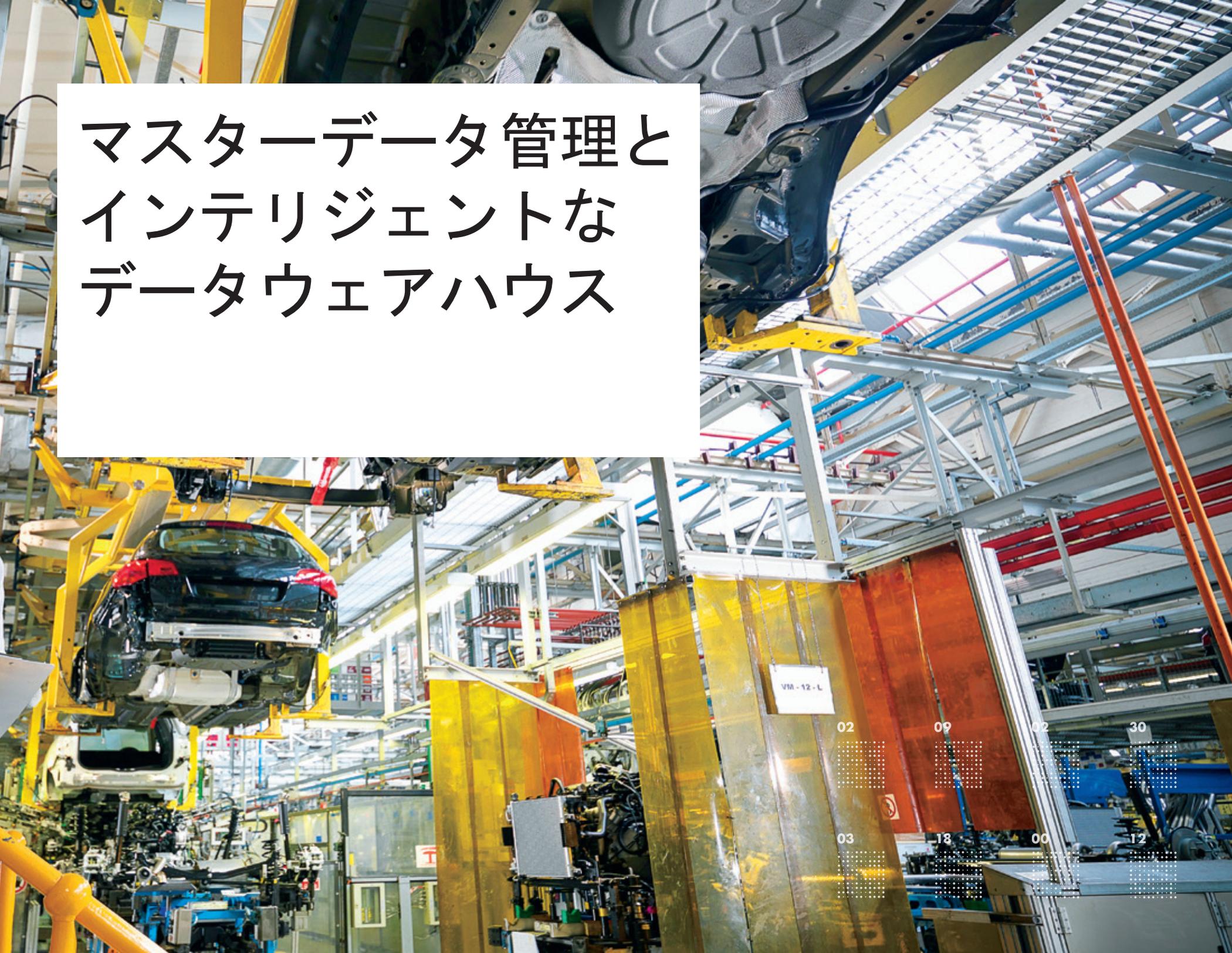
データ資産の精度と価値を高めるため、データ品質ツールは、データをデータウェアハウスに格納する前のクレンジング作業で使われます。

データウェアハウスに格納する前にデータクレンジングを行うということは、ソースシステムは不正確で不完全なデータであふれているということです。データ品質ソフトウェアは、全ての高品質で正しいデータのレポジトリを維持するものではないため、これらのソースシステムを使用する場合は、依然としてクレンジングされていないデータに依存することになります。

例えば、2人の人物が同じデータ（月間の営業利益データ）を見ている場合、それぞれが異なるソースシステムを見ているとすれば、正しいバージョンが2つ存在することになります。また、これらの2つのバージョンは、実際の売上の件数ではなく、システムのアップデートの頻度によって変化します。

つまり、これらのツールは一貫性のあるデータウェアハウスにとって不可欠なものですが、BIの課題というパズルを解消するためには、まだ足りないピースがあります。

# マスターデータ管理と インテリジェントな データウェアハウス



データウェアハウスの障害となっている制限事項を下記にまとめます。

データソースは統合されているが、データ自体はいまだに重複や非一貫性という欠陥がある。

データはクレンジングされているが、高品質なデータは実際には全ての場所に保存されているわけではない。

データは変更できるが、これらの変更を実装し記録するには膨大な量の、ミスが発生しやすい手作業が必要である。

インテリジェンスのパズルは、これら3つの制限事項を解決できたときに完成します。

## データプロフェショナルのための万能薬

マスターデータ管理（MDM）は、管理されたプロセスにより、コードフリーで保守負荷の少ないソリューションを提供します。これはBIシステムやDIツール、DQプロセスでは実現できないものです。

1

データを統合し、ディメンションに関する単一の信頼できる視点を提供。これによってデータは安心して活用できるようになります。これは、従来のBIシステムでは対応できない非一貫性や重複の問題（GEとGeneral Electricが同じ会社であることを認識できない等）を全て解決します。

2

データを定期的・自動的にクレンジング。このようにして健全で高品質になった全てのデータを一元化されたレポジトリに保存し、接続された全てのソースシステムを更新します。端末システムである既存のBIとは違い、MDMは双方向のデータフローを可能にします。

3

データが変更されると異常を示す警告が直ちに生成されるので、変更を検証、保存、追跡できます。これは、データの履歴とリネージに対する信頼性が確保されるということです。またETL開発者は、全ての変更を反映してデータを準備するために交替制で働く必要もありません。

重複がなく、統合され、規制に準拠し、クレンジングされた、唯一の信頼できるバージョンが、いつでも最新の状態で利用できるようになります。

# MDMは、信頼できるデータをどのように提供するのか

## MDMがないDW

米国の会計システム

General Electric Companyのトランザクション  
**\$400<sub>MM</sub>**



EUの会計システム

GEのトランザクション  
**\$300<sub>MM</sub>**



General Electric Companyのトランザクション  
**\$400<sub>MM</sub>**



GEのトランザクション  
**\$300<sub>MM</sub>**

2つのデータウェアハウジングシナリオ：左の図は、MDMシステムを使用せずに信頼できるマスターデータを作成し、データウェアハウスに保存しようとした場合に生じる不正確なレポートのリスクを示しています。

## MDMを導入したDW

米国の会計システム

General Electric Companyのトランザクション  
**\$400<sub>MM</sub>**



EUの会計システム

GEのトランザクション  
**\$300<sub>MM</sub>**



完全なレコード：  
General Electric Company  
社のトランザクション  
**\$700<sub>MM</sub>**

右の図は、マスターデータをデータウェアハウスに保存する前にデータの問題を解決することで、MDMシステムが正確なレポートをどのようにサポートするのを示しています。

「当社の第一の目標は、販売パイプラインを強化することです。  
**MDM**を導入したことで、リードコンバージョン率を**20%**改善  
することができました。データへの全体的な視点と予測分析の  
力を手にしたことは、未来を予知する能力にも匹敵します。」

マーケティング部シニアマネージャー、Dagmar Garcia氏



## MDMの技術的なメリット

BIの観点から見たMDMによる最大のメリットは、3つの技術的項目に対して極めて効率的に対応できることです。

1

### 共通ディメンション

MDMシステムでは、共通データ構造にデータを挿入する際に通常発生するデータ形式の変換作業がありません。従って、社内のデータウェアハウスリソースは、データの統合ではなくビジネスに関連する質問に答えることに集中できます。

2

### 徐々に変化するディメンション

MDMは、データウェアハウス開発者が徐々に変化するディメンションのアプローチを採用した場合に生じるパフォーマンスの問題を解決します。履歴追跡オプションによってリファレンスデータエンティティに対する全ての変更が記録され、データウェアハウスは、徐々に変化するディメンションとして、データセグメンテーションの変化を追跡することに集中できるようになります。

あるレコードについて、ユーザーが徐々に変化するのではないディメンションの履歴を照会したい場合は、MDMシステムの履歴テーブルまでドリルダウンして確認できます。

3

### データリネージ

全てのレコードの全てのフィールドについての詳細なデータリネージ（セルレベルまで）が記録されるので、MDMでは時間のかかる手作業でのデータ追跡が不要になります。

例えば、「レコードの顧客名の値がなぜJohn Smithなのか」を知るために全てのステージングテーブルを照会する必要はなく、MDMシステムにアクセスしてその値が格納されているソースシステムを特定すれば良いということです。また、レコード結合の履歴も追跡します。

# MDMを原動力 とするデータ ウェアハウスの力

(なぜその力が大きいのか)



02

10

Y

BB

8

S

R

33

F

D

BB

8

S

データウェアハウスのソースがクリーンで、そのデータが常に更新されているということは、分析およびインテリジェンスシステムが消費すべき最良のデータに基づいて運用されているということです。そのデータとは、信頼でき、監査性があり、正確で、単一の統合された真のデータであり、これ以外に真のデータと言えるものはありません。

これは、膨大な時間と多額の費用を投資して構築されたBIシステムから作成されたレポートやデータを利用するユーザーにとっては、すばらしいニュースです。

**経営者**は、レポートに含まれる情報が明確な全体像を示していることを知っているので、安心して質の高い意思決定を行うことができます。

ビジネスアナリストは間違いや矛盾を見つけることに全ての時間を費やすことがなくなり、本来の仕事（データの分析）に専念できるようになります。また、直観的に得たものをより簡単に検証し、より徹底してその仮説をテストすることができます。

**IT技術者とデータアーキテクト**は、データを提供する仕組みではなくデータ自体の品質改善に集中できるようになります。また、改善と統合を継続的に実現するシステムを使用しているため、スキーマの変更を毎回ハードコーディングすることはありません。

**コンプライアンス担当者**は、規制当局が求めるレベルの精度でレポートを作成することができます。より重要な点、それはより高品質のレポートは、担当者が目標とするコンプライアンス基準達成のための大きな力になるということです。

本当に信頼できる正確な情報に基づいて運用されるビジネスインテリジェンスシステムのメリットは計り知れません。また**MDM**は、真のインテリジェンスがもたらす理想の世界へと**BI**システムを導く重要な役割を担っています。よりインテリジェントな意思決定のため、ビジネスインテリジェンスには質の高いデータが不可欠です。

# マスターデータ管理について

InformaticaのMDM製品およびマルチドメイン機能は、規模を問わずあらゆる業界の企業において大きな注目を集めています。弊社はまた、ForresterのMDM Waveレポートでリーダーに位置づけられています。



詳しくは「The Forrester Wave: マスターデータ管理ソリューション2014年Q1」（英語）をご覧ください。

# Informatica について

Informaticaは、ビジネスにとって重要な製品、サプライヤー、顧客に関する完全で正確なデータを提供することで、データを第一に考えるお客様を支援しています。

[お問い合わせはこちら](#)

