



# XMLDB勉強会

## 活動の意義と目標

リーダー 加藤哲義  
(株式会社サイバーテック)

katou@cybertech.co.jp

2006.12.11

### 1. 経緯

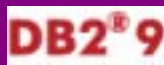
- 2006. 8.10. 第1回設立準備委員会
- 9. 5. 第2回設立準備委員会 設立趣意書提示
- 10.4. 第1回XMLDB勉強会 (XMLDBセミナー併催)
- 11. 2. 第2回XMLDB勉強会
- 12. 6. 第3回XMLDB勉強会・マーケティンググループ
- 12. 7. 第3回XMLDB勉強会・技術グループ

⋮

2007年2月まで

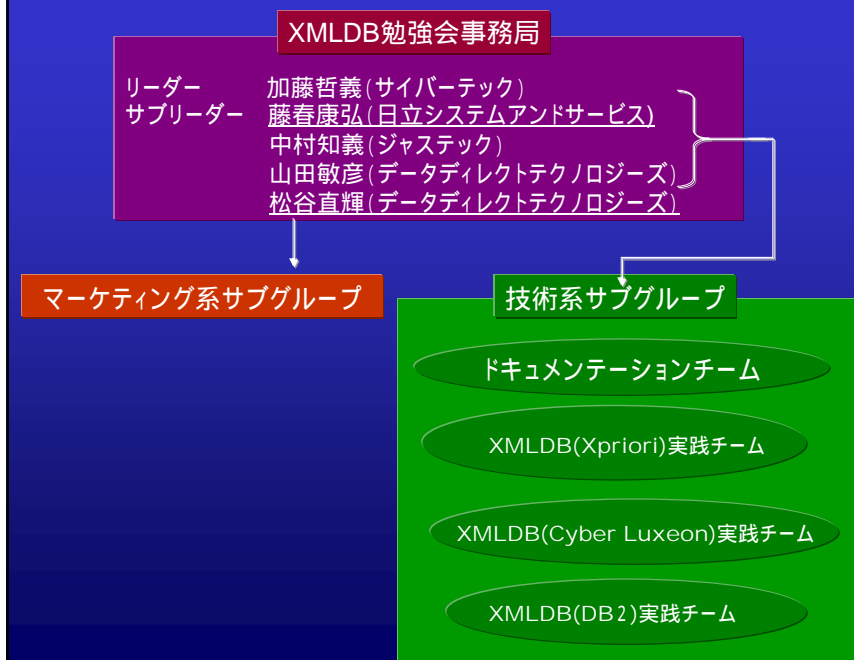
## 準備委員会からXMLDB勉強会へ

準備委員会 : サイバーテック、データディレクトテクノロジーズ、  
東芝ソリューション、日本アイ・ビー・エム、  
日本オラクル、日立製作所、三井物産、  
メディアフュージョン  
XML技術者育成推進委員会、  
XMLコンソーシアム運営委員会代表



ユーザの理想(思い込み) VS ベンダーの現実感(製品の都合)

## 2. 体制



### 3. 活動概要

#### マーケティング系サブグループ

ベンダー訪問を行い、

その製品の特長とユーザ事例や、実ユーザの生の声を聞く。  
それをもとにディスカッション。  
訪問記を残す。

#### 訪問先

- ・日本IBM (12/6)
- ・サイバーテック (1/10)
- ・日本オラクル (2/7)



日本IBM訪問風景 2006.12.6

### 3. 活動概要

#### 技術系サブグループ

ドキュメンテーションチーム

XMLDB(Xpriori)実践チーム

XMLDB(Cyber Luxeon)実践チーム

XMLDB(DB2)実践チーム

W3C「XQuery Use Cases」翻訳  
「W3Cの歩き方」みたいなものもつくれたら

実際にXMLDBを使って勉強しながらXMLDB導入ステップの考察、  
利用方法を考えていく。



#### 4 . 技術系サブグループの雑談の中にヒントあり！

RDBエンジニアからみたXMLDBへのとまどい、つまずきがあるよね。

別のものとして考えるべきだろう。

XMLDBアレルギーとその処方箋をつくっては、

直感的にXMLDBで解決できそう、だが、不安（性能、安定感・・・）

Xpath, Xquery, XSLT・・・勉強することが多すぎる、敷居が高い

性能を引き出すためのXQueryのお作法は？

XMLDBのよさはアバウトさだ

⋮

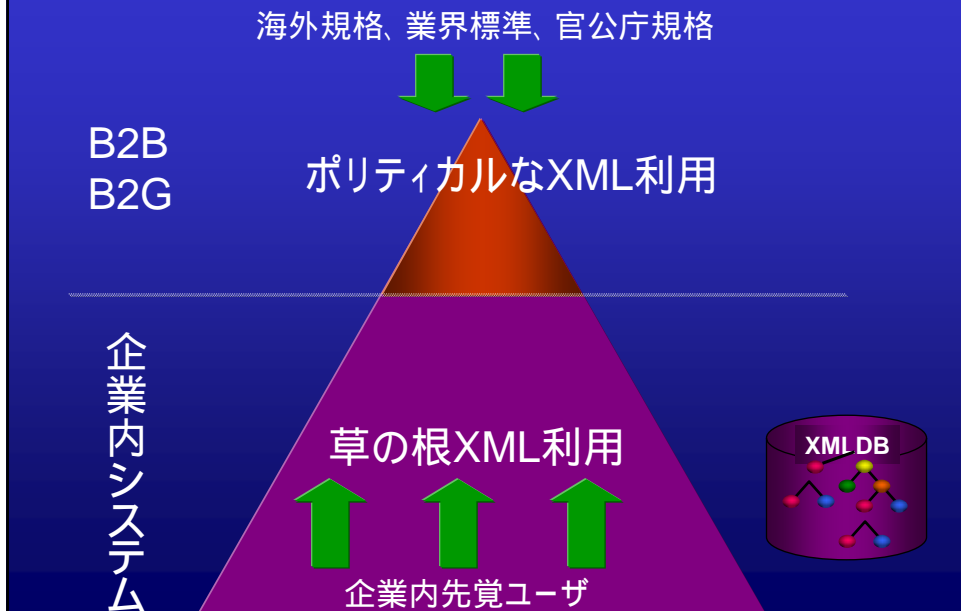
XMLDBってそもそもなんなの？

XMLDB、ってそもそもなんなの？

「XMLDB徒然草」

7年の振り返りとヒント

## XML普及の圧力(1999年当時)



## XMLDBが登場した1999年から2000年のXMLDBの利用

- ・ 製造業のエンジニアが
- ・ 自己裁量できる製品仕様などをXMLで表現し
- ・ XMLDBに容れてシステム化をはかった。

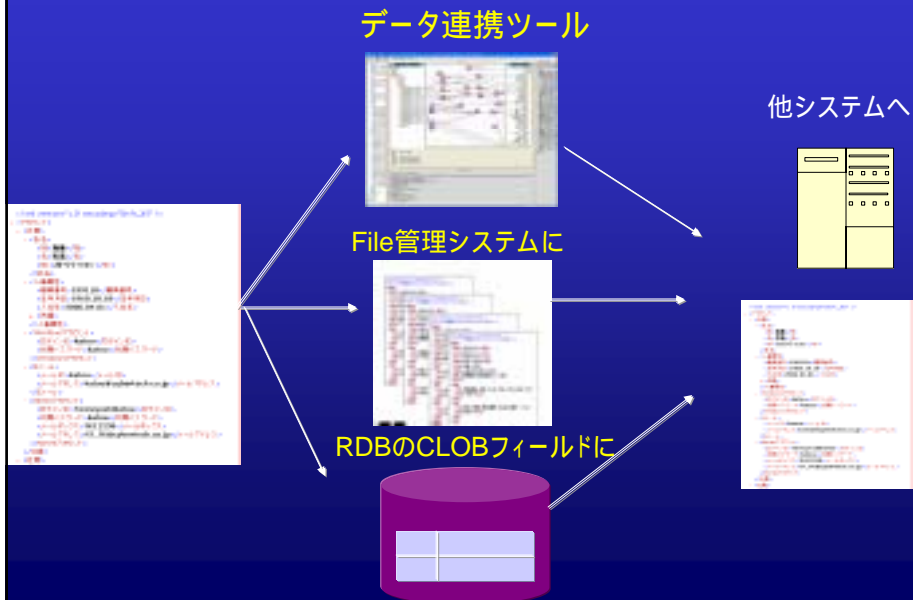
B2Bでは無く、社内ナレッジ管理。  
欧米より、日本、韓国で普及した。  
現場指向が強い？  
文化の違い？

## 利用形態の差異

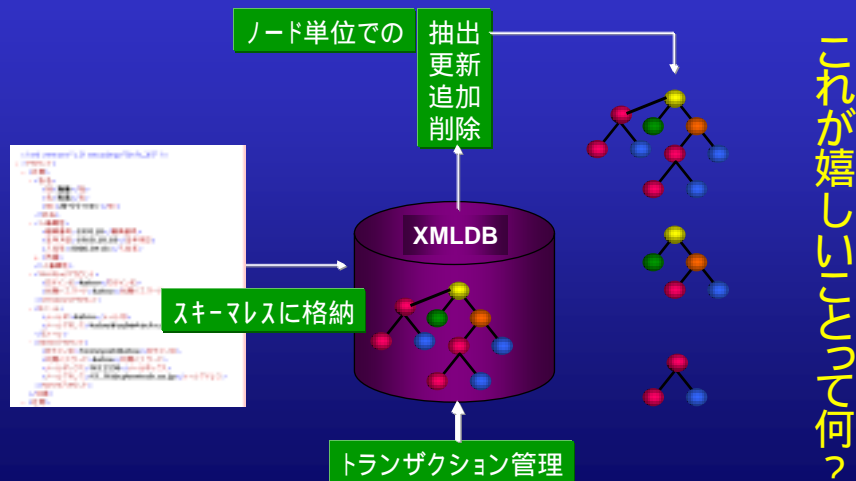
	企業間利用	企業内利用
影響範囲	業界、グループ傘下	自社、自部門
決裁者	海外、トップ判断	担当判断、部門長決裁
データスキーマ	既存、標準仕様	自社仕様
推進者	外圧、政策的	社内技術的先覚者
スキーマ変更	一旦決まれば固定的	変更しがち
XML-DB	あえて必要でない	必要

XMLを利用するシステムは必ずしも  
XMLDBを利用するわけではない。

非XMLDB ファイルでためてファイルで出す。



XMLDBを使うと処理の粒度がXMLファイル~どのノードでも自由



## XMLDBはシステムアイデアの起爆剤

Excel 項目(構造)と値が一体で織り込まれている。

	単位	A型機種	B型機種
車体重量	KG	2,000	3,200
エンジン出力	CC	5,000	6,500



	単位	仕様	A型機種	B型機種
車体重量	KG	標準	2,000	3,200
		寒冷地	2,300	3,500
エンジン出力	CC		5,000	6,500

新しい発見やニーズがでたら  
自由に列と行を挿入し、項目と値を追加して対処。



ExcelのデータはDB化されないので  
せっかくのナレッジが公共化されない。

自由度と公共性の対立  
構造を制限しないと共有化できない。

	単位	A型機種	B型機種
車体重量	KG	2,000	3,200
エンジン出力	CC	5,000	6,500

項目と値が一体で織り込まれている。

```
<データ>
<機種>
  <名称>A型機種</名称>
  <車体重量 単位="KG" >2,000</車体重量>
  <エンジン出力 単位="CC" >5,000</エンジン出力>
<機種>
<機種>
  <名称>B型機種</名称>
  <車体重量 単位="KG" >3,200</車体重量>
  <エンジン出力 単位="CC" >6,500</エンジン出力>
<機種>
</データ>
```

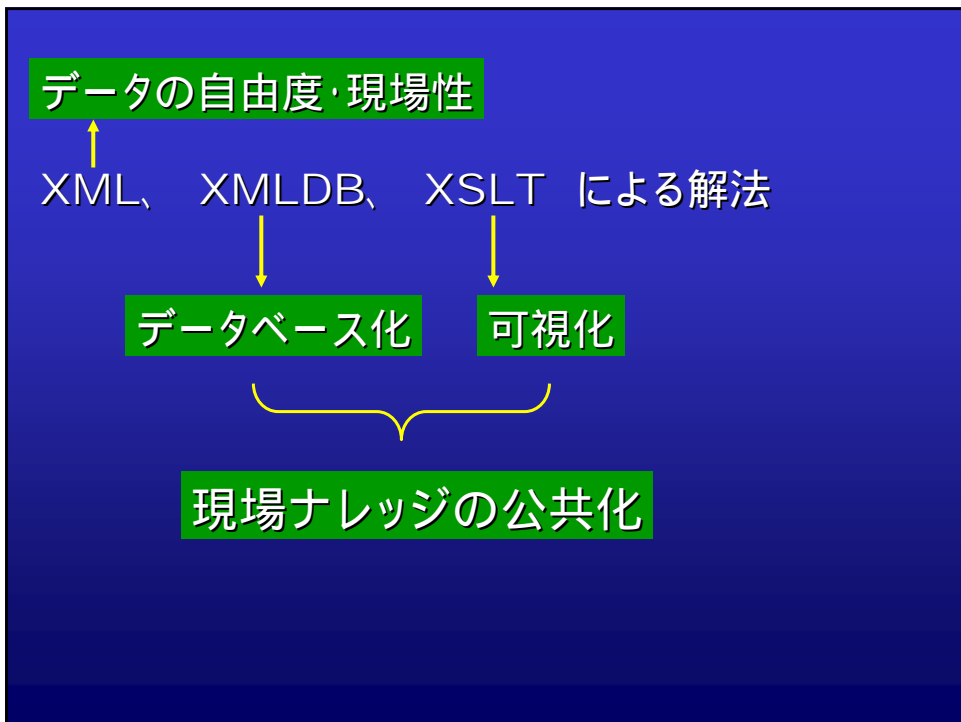
	単位	仕様	ZSX200	ZSX300
車体重量	KG	標準	2,000	3,200
		寒冷地	2,300	3,500
エンジン出力	CC		5,000	6,500

新しい発見やニーズがでたら  
 自由に行を挿入し、タグと値を追加して対応。

```

<データ>
<機種>
<名称>A型機種</名称>
<車体重量 単位="KG" >
  <標準>2,000</標準>
  <寒冷地>2,300</寒冷地>
</車体重量>
<エンジン出力 単位="CC" >5,000</エンジン出力>
<機種>
<機種>
<名称>B型機種</名称>
<車体重量 単位="KG" >
  <標準>2,300</標準>
  <寒冷地>3,500</寒冷地>
</車体重量>
<エンジン出力 単位="CC" >6,500</エンジン出力>
<機種>
</データ>
  
```



構造と内容の決定権が高いほど創造的である

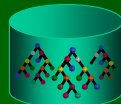


製造業の設計エンジニア

非定型データ

**社員の身体測定結果について**  
全般的に高年齢層の精神的ストレスの過大および若年層の不健康な食生活からもたらされる体力不足が自立。  
20代から30代の肥満度は昨年より減少の傾向にあり、平均肥満度は - 3ポイントであるが、体内脂肪はかえって増加している。  
ファーストフードの普及にともない未婚男女の栄養バランスの偏りに関連性があるかもしれない。  
一方40代の男性は内科所見上、慢性的な倦怠感と精神的な焦りをうったえる方が多かつた。長引く不況による職場環境の悪化が原因の一つかもしれない。今後ならなかケアしていく施設の用意が必要と思われる。

半定型データ



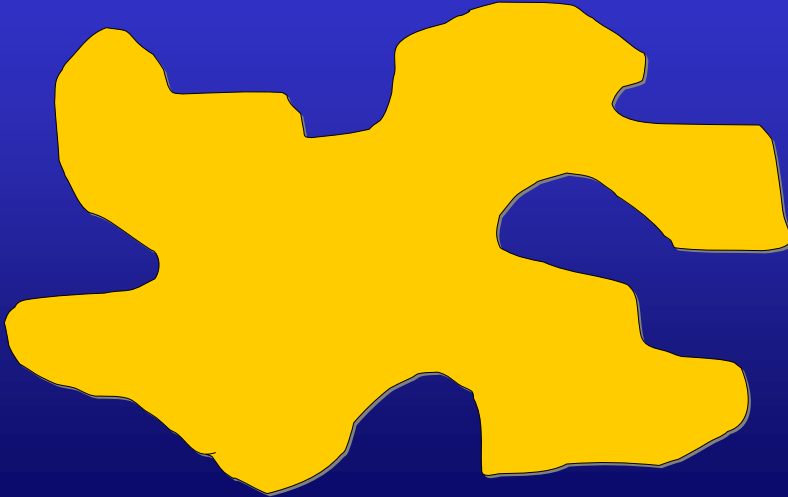
定型データ



一見ドキュメントだが構造的があり、その構造的検索、更新で再利用価値のあるもの

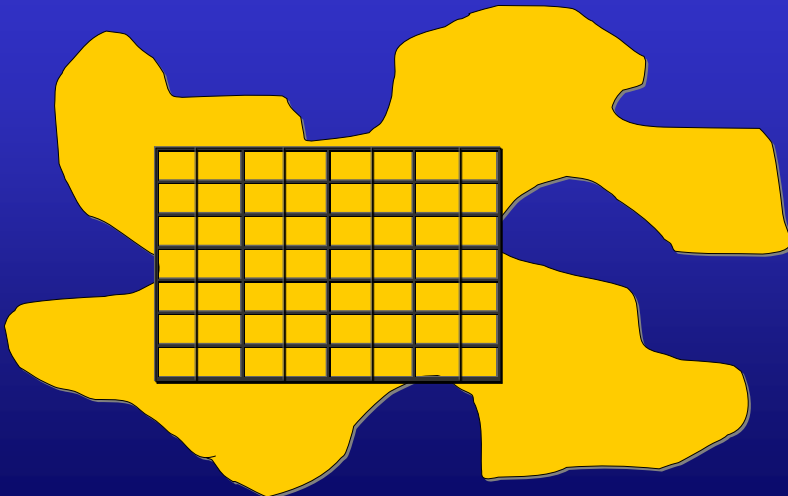
XMLは複雑なデータ構造が表現でき、その可視性が高い。

現実のビジネスやナレッジは不定形



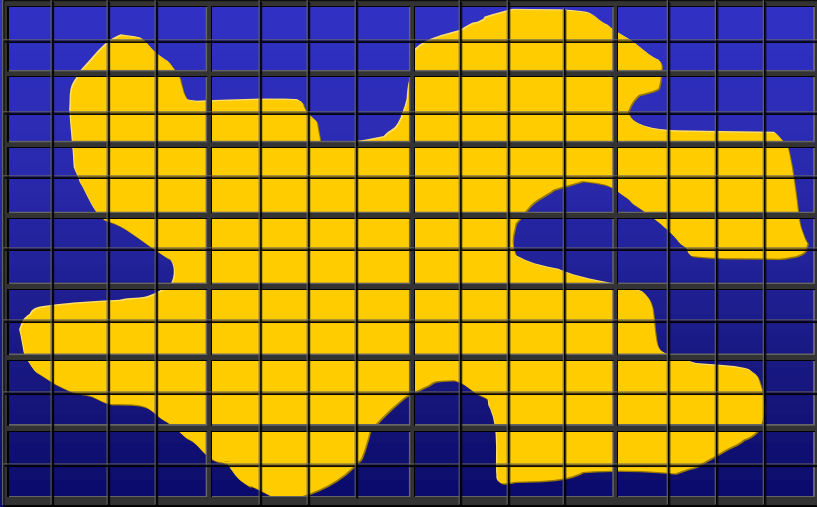
リレーショナルデータベースで定義

(1) 部分(個性、レアケース、不確定部分)を無視



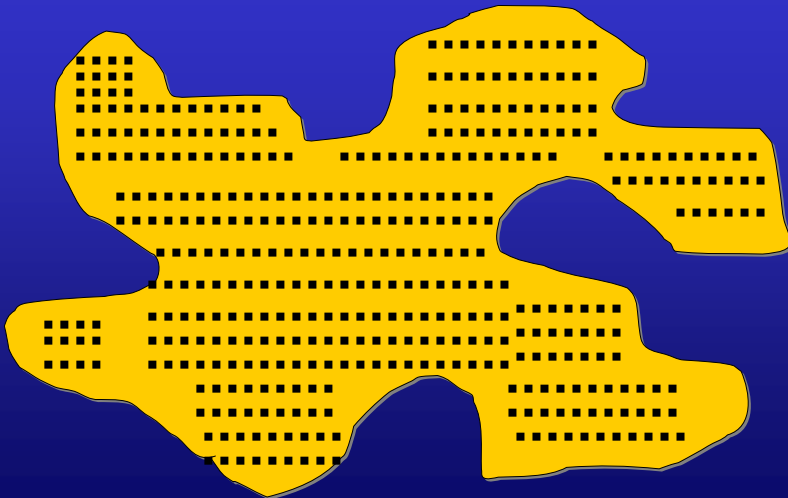
## リレーショナルデータベースで定義

(2) スカスカのテーブル 非現実的



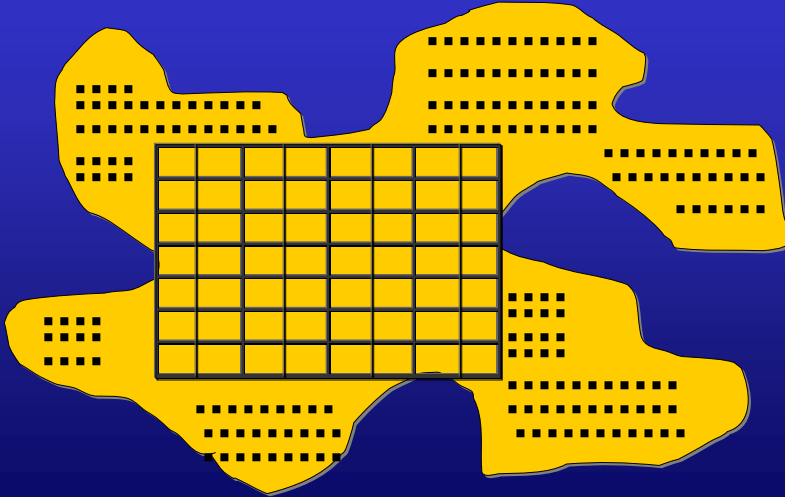
## XMLで定義

現実を忠実にデータ化



# XMLとリレーショナルデータベースの混成

現実を効率的にデータ化



XMLは複雑なデータ構造が表現でき、その可視性が高い

・複雑な階層を持つモデルが定義可能

仕様書、分子構造、図面管理、配線管理、部品構成  
会社組織、アカウント、マニュアル、.....

・レコード単位に項目がまちまちなモデルも最適な  
データ設計ができる

設備製品	共通属性			個別属性			
	属性1	属性2	属性3	?	?	?	?
あ				■	■		
い				■	■		
う				■	■	■	
え				■	■	■	
お				■	■	■	
か				■	■		
き				■	■		
く				■	■		
け				■	■	■	

## XMLはデータ構造の見直し、拡張が容易

・あらかじめテーブル設計やフィールド定義がしがたいモデルのDB化

- ・例外ケース
- ・想定外ケース
- ・ユーザーから仕様が段階的出てくる

・つくりながら分かる、運用をはじめてから分かるモデル

- ・クレームデータ
- ・マーケティングデータ
- ・実験データ
- ・評価基準



## XMLDBを選択するポイント

その製品の**逆鱗**を知ること

XMLDBは諸刃の刀

その製品は何がトレードオフされているか？

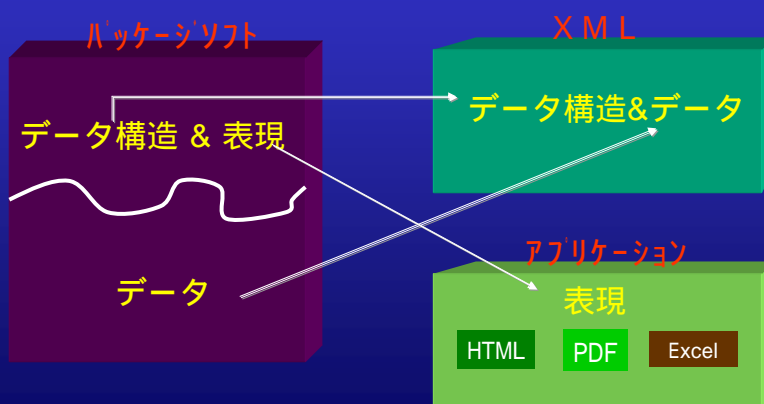
## XMLDBシステムの開発スタイル

データと、その表示は疎結合  
分けて開発する。

XML

XSL

XMLはデータがアプリケーションソフトから独立的





## XMLDBシステムの開発スタイル

### 段階的、スパイラル、プロトタイプから

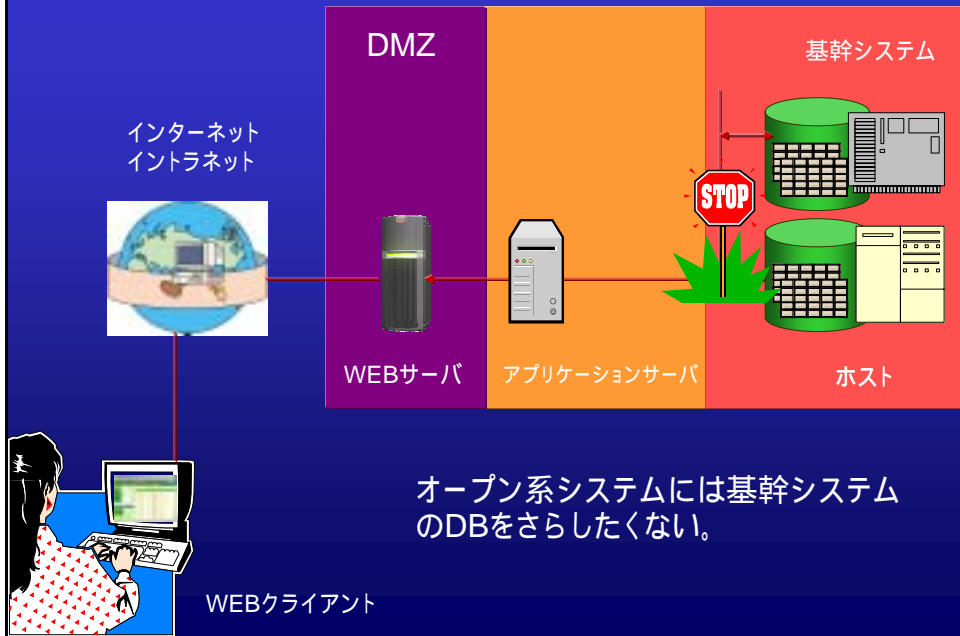
仕様をまとめる段階で明らかになっている範囲でデータ構造を設計し、それをもとにサンプルデータとシステムを作成し、テストを通して仕様を再検討して最適なシステムを作り上げていく。

プロトタイプを進化させながら段階的にデータ構造、仕様、サイジングを見極め、実用システムへ高めて行く。

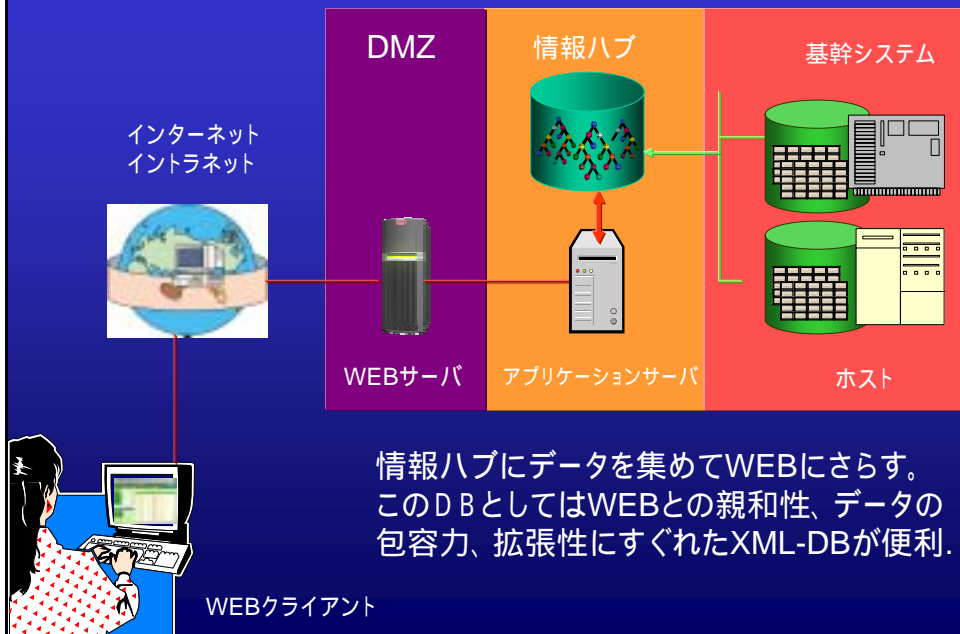
## XMLDBシステムの成功判断

当初の目的を超える効果を得たかどうか？

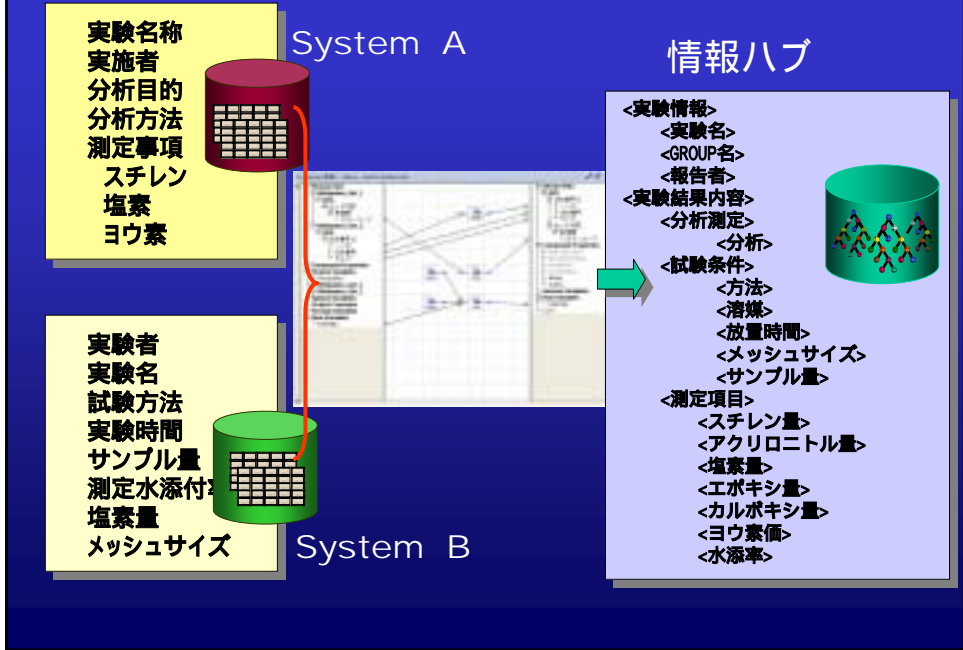
## WEBシステムが参照するDBは



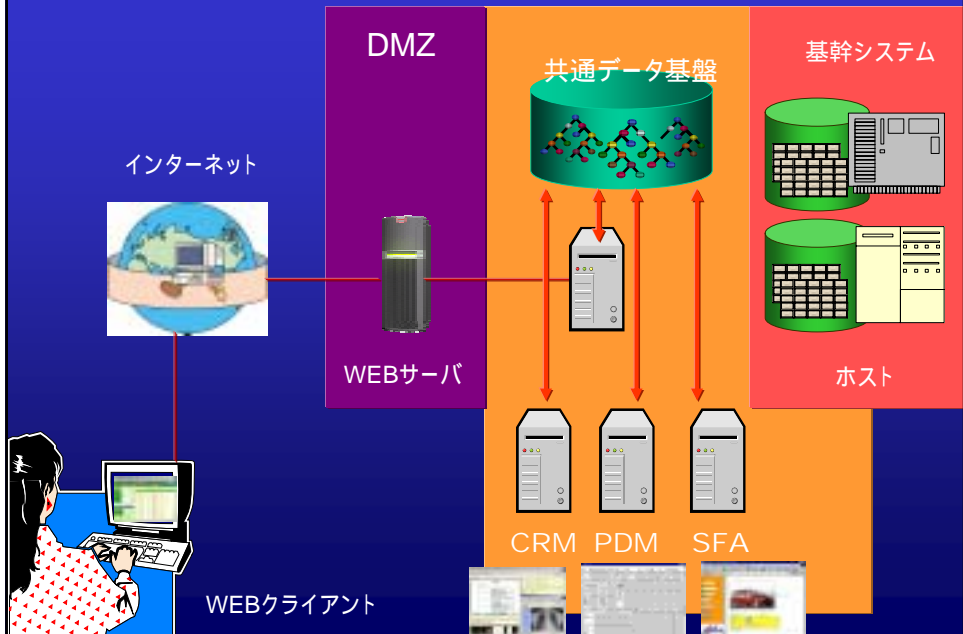
## WEBシステムが参照するDBは



# 各データスキーマを束ね易いのがXML-DB



# 情報ハブからオープン系システムのデータ基盤に成長



## RDB

### 利用目的 = データ構造

利用目的通りのデータ構造。無駄が無い。

## XMLDB

### 利用目的 > データ構造

データを再編成して当初の用途以上の用途を生み出す。

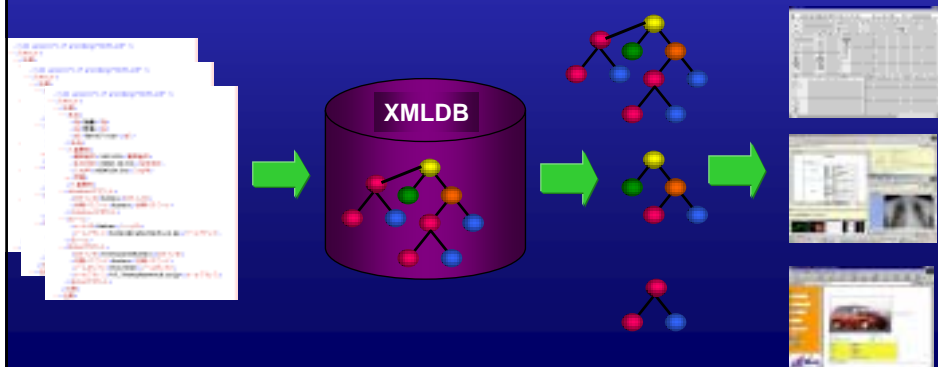
### 利用目的 < データ構造

使わないタグは無視、無いタグは階層的に作る。

XMLDBは、包容力ある(アバウトな)データの容器。

使う時に用途に合わせてデータを最適化して使えばいい。

そのためにデータの粒度が選択自由な器である。



## XMLDBを利用する目的

情報システムに**現場性**をどこまで**リアル**に取り込むか

### まとめ

XMLDBの意義を考えることは  
XMLそのものの真価を見直すことになる。