



XML Consortium

XMLDBの特性とそのユーザ像

XMLDB勉強会

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

XML DB の特性

XMLは、構造と値が一体化されたテキストデータ

階層化による複雑なデータ表現が可能

いつでもデータの改造が容易

データの原本はDBでなく
このテキストファイル



対象をよりリアルに表現できる

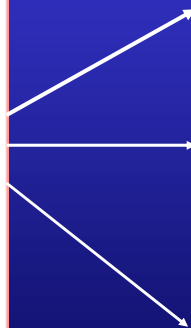
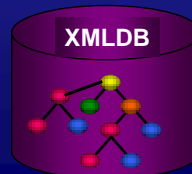
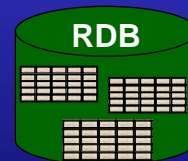
まずはDB設計を意識しなくてもよい
(システムのプロでなくてもよい)

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<健康診断データ>
  <健康者>
    <個人データ>
      <名前>加藤大介</名前>
      <性別>男</性別>
      <年齢>50</年齢>
      <身長>174</身長>
    </個人データ>
    <体重>63</体重>
  </健康者>
  <上>110</上>
  <下>80</下>
  </血圧>
  <腎>正常</腎>
  <心臓>正常</心臓>
  <腎臓>50歳以上に1回高血圧検査必要</腎臓>
  <遠視近視等>高血圧検査<
    <遠視>正常</遠視>
    <近視近視等>正常</近視近視等>
    <心電図>軽い不整脈</心電図>
  </遠視近視等>
  </健康者>
</健康診断データ>
```

XMLのデータ処理

ファイル処理

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<健康診断データ>
  <健康者>
    <個人データ>
      <名前>加藤大介</名前>
      <性別>男</性別>
      <年齢>50</年齢>
      <身長>174</身長>
    </個人データ>
    <体重>63</体重>
  </健康者>
  <上>110</上>
  <下>80</下>
  </血圧>
  <腎>正常</腎>
  <心臓>正常</心臓>
  <腎臓>50歳以上に1回高血圧検査必要</腎臓>
  <遠視近視等>高血圧検査<
    <遠視>正常</遠視>
    <近視近視等>正常</近視近視等>
    <心電図>軽い不整脈</心電図>
  </遠視近視等>
  </健康者>
</健康診断データ>
```

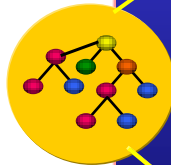


意味の構造からシステムの構造への変換

XML=意味の構造

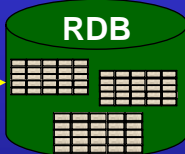
DB=システムの構造

```
<?xml version="1.0" encoding="Shift_JIS" ?>  
-<健康診断データ>  
-<健康者>  
-<個人データ>  
-<名前>加藤大介</名前>  
-<性別>男</性別>  
-<年齢>50</年齢>  
-<身長>174</身長>  
-</個人データ>  
-<検査>  
-<上>110</上>  
-<下>80</下>  
-</血圧>  
-<骨>正常</骨>  
-<心臓>正常</心臓>  
-<腎臓>正常</腎臓>  
-<糖尿病検査>正常</糖尿病検査>  
-<心電図>軽い不整脈</心電図>  
-</検査結果>  
-</健康診断データ>
```



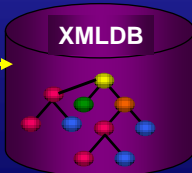
人為的変換

テーブル設計とマッピング



機械的変換

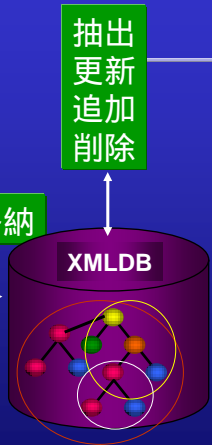
|| スキーマレス



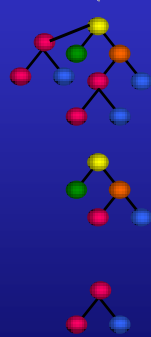
XMLDBの特性

```
<?xml version="1.0" encoding="Shift_JIS" ?>  
-<健康診断データ>  
-<健康者>  
-<個人データ>  
-<名前>加藤大介</名前>  
-<性別>男</性別>  
-<年齢>50</年齢>  
-<身長>174</身長>  
-</個人データ>  
-<検査>  
-<上>110</上>  
-<下>80</下>  
-</血圧>  
-<骨>正常</骨>  
-<心臓>正常</心臓>  
-<腎臓>正常</腎臓>  
-<糖尿病検査>正常</糖尿病検査>  
-<心電図>軽い不整脈</心電図>  
-</検査結果>  
-</健康診断データ>
```

スキーマレスに格納



抽出
更新
追加
削除

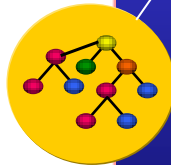


ファイル
ノード
エレメント

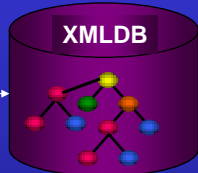
トランザクション管理

機械的変換はユーザは楽な分、製品は苦勞する

```
<?xml version="1.0" encoding="Shift_JIS" ?>  
-<健康診断データ>  
-<健康者>  
-<病入者>  
-<名前>加藤大介</名前>  
-<性別>男</性別>  
-<年齢>50</年齢>  
-<身長>174</身長>  
-<体重>63</体重>  
-<血圧>  
-<上>110</上>  
-<下>80</下>  
-<血糖>  
-<糖>正常</糖>  
-<脂質>正常</脂質>  
-<腎臓>50歳以上かつ併高血圧検査必要</腎臓>  
-<遠視近視>遠視</遠視近視>  
-<近視>正常</近視>  
-<聴覚検査済>正常</聴覚検査済>  
-<心電図>軽い不整脈</心電図>  
-</健康者>  
-</健康診断データ>
```



機械的変換

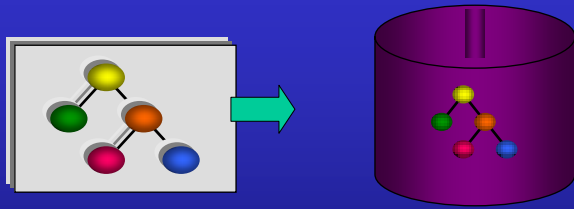


XMLDB各社は、この方法にしのぎを削る。

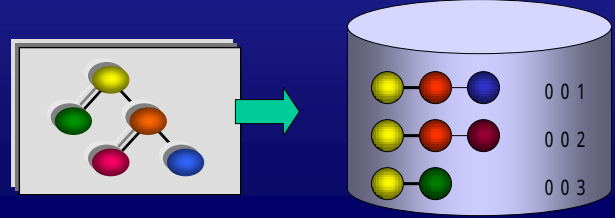


どうやって機械的変換を実現するか

XMLを構造的なDBにそのまま容れる(データプロセス指向)。



すべてのノードにインデックスを行う(全文検索指向)。



XML DB ユーザ像

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。

XML普及の圧力(1999年当時)

海外規格、業界標準、官公庁規格



B2B
B2G

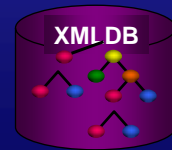
ポリティカルなXML利用

企業内システム

草の根XML利用

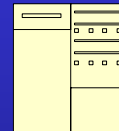


企業内先覚ユーザ



XMLDBを必要としないシステム : B2B、B2G

1. すぐデータ変換して他システムへ

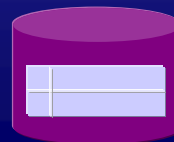


2. 処理の粒度がファイルで十分

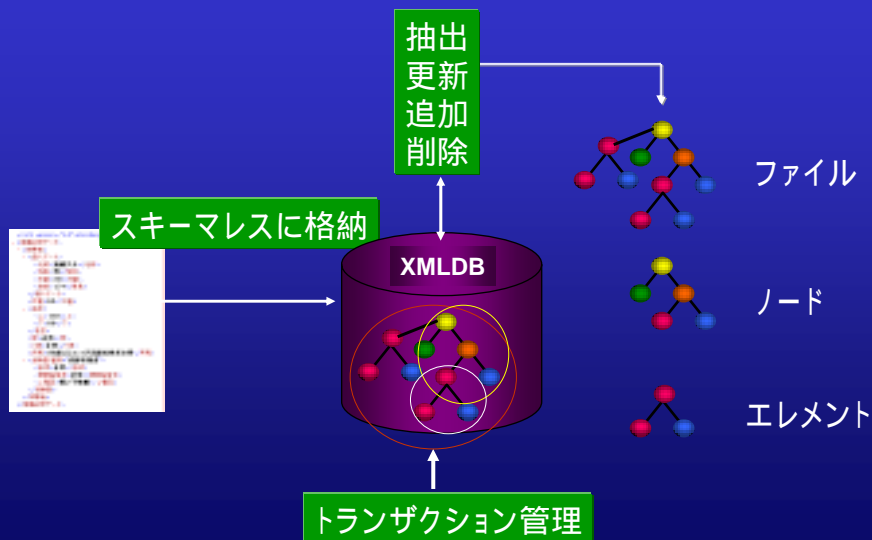
ファイル管理



RDBのCLOGカラムへ



XMLDBを必要とするシステム：任意の粒度で処理したい



XMLDBが登場した1999年から2000年のXMLDBの利用

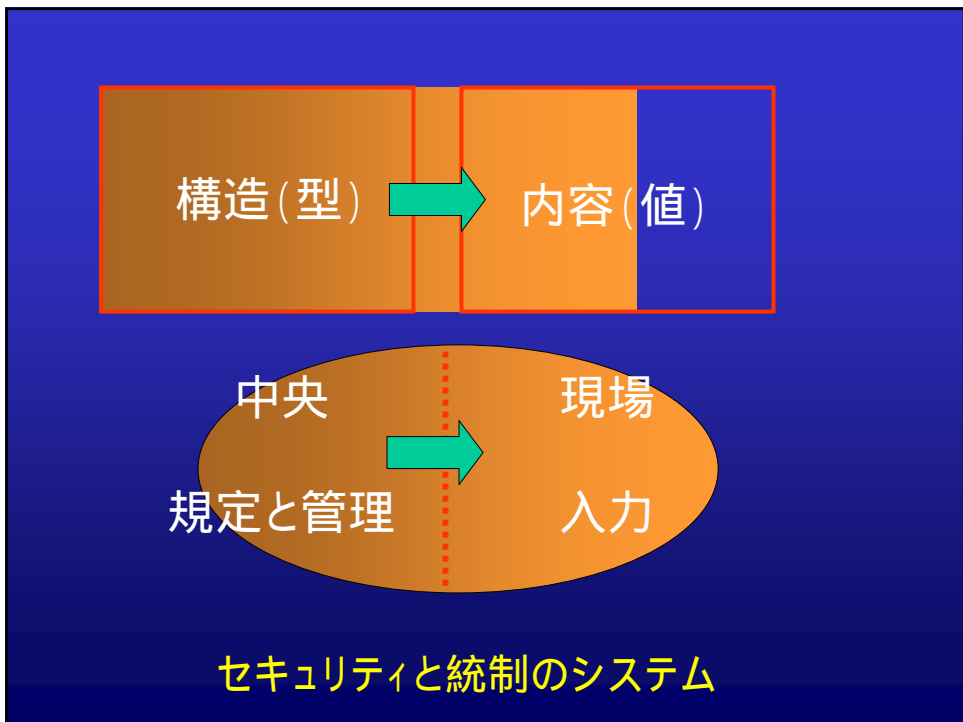
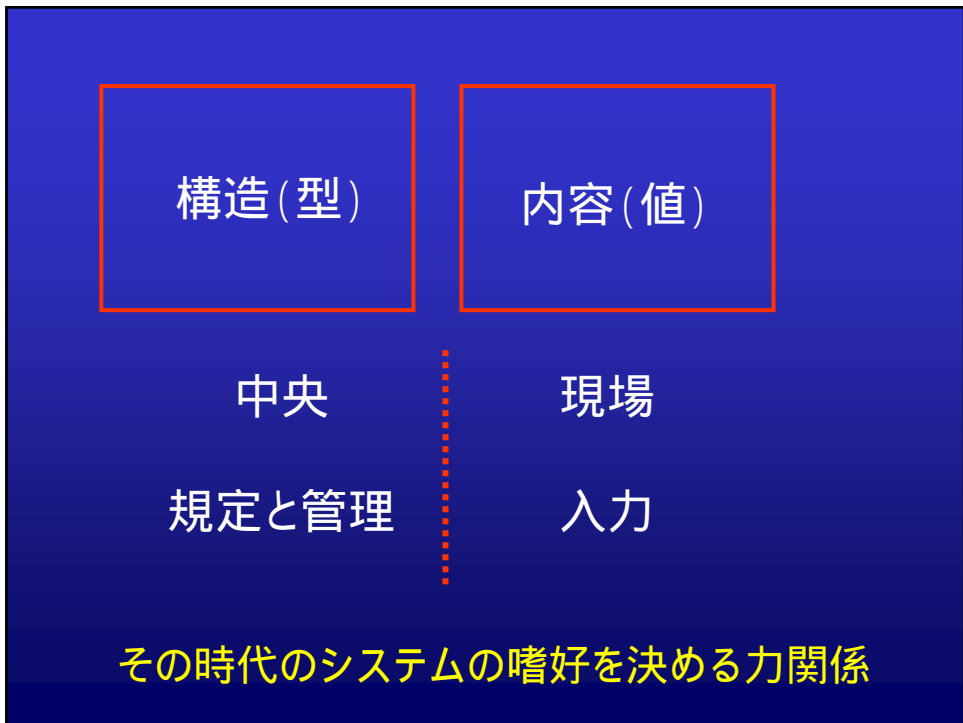
- ・ 製造業のエンジニアが
- ・ 自己裁量できる製品仕様などをXMLで表現し
- ・ XMLDBに容れてシステム化をはかった。

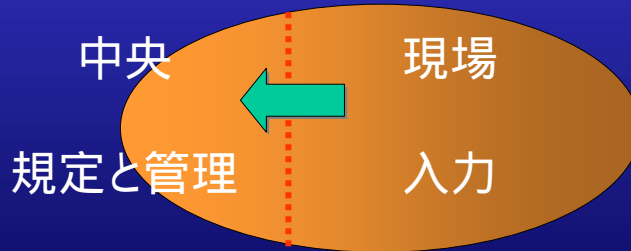
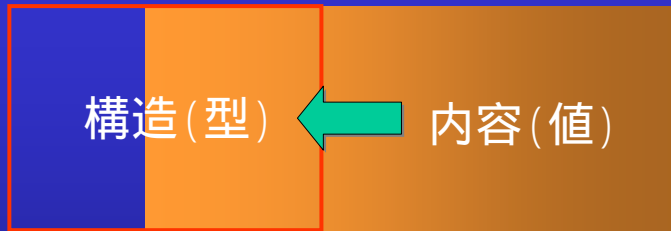
B2Bでは無く、社内ナレッジ管理。
欧米より、日本、韓国で普及した。
現場指向が強い？
文化の違い？

XML利用形態の性格

	B2B, B2G	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

XMLDBはシステムアイディアの起爆剤





情報共有、CRM、ナレッジ
構造と内容の決定権が高いほど創造的である

Excel 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

新しい発見やニーズがでたら



	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

自由に列と行を挿入し、項目と値を追加して対処。

Excelは企業内ナレッジの宝庫

ExcelのデータはDB化されないの
でせっき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" >5,000</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>
```

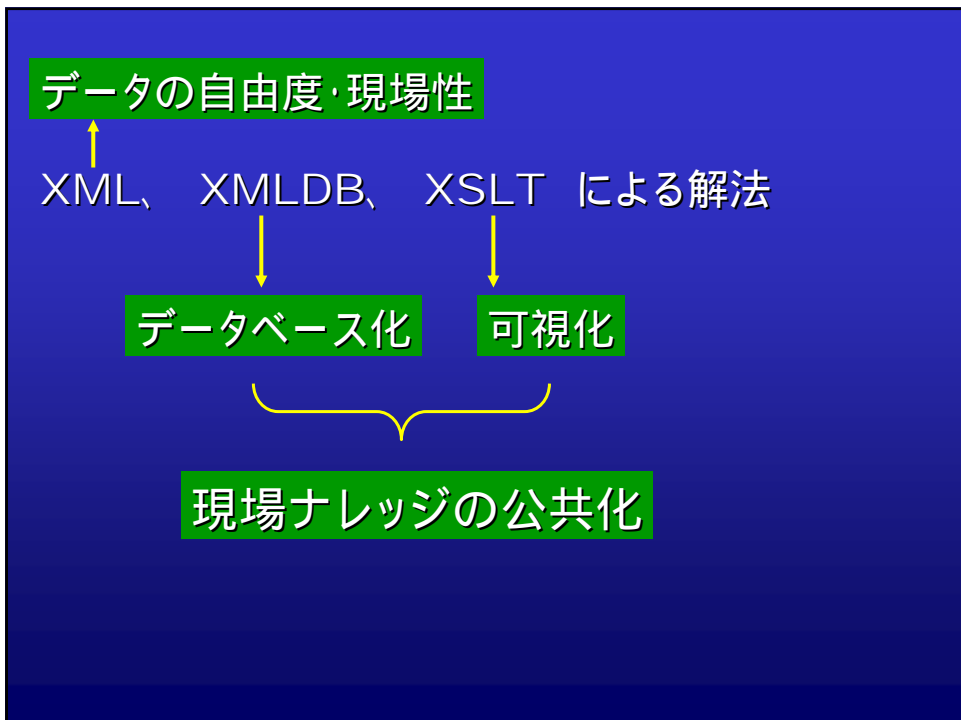
	単位	仕様	ZSX200	ZSX300
車体重量	KG	標準	2,000	3,200
		寒冷地	2,300	3,500
エンジン出力	CC		5,000	6,500

**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>

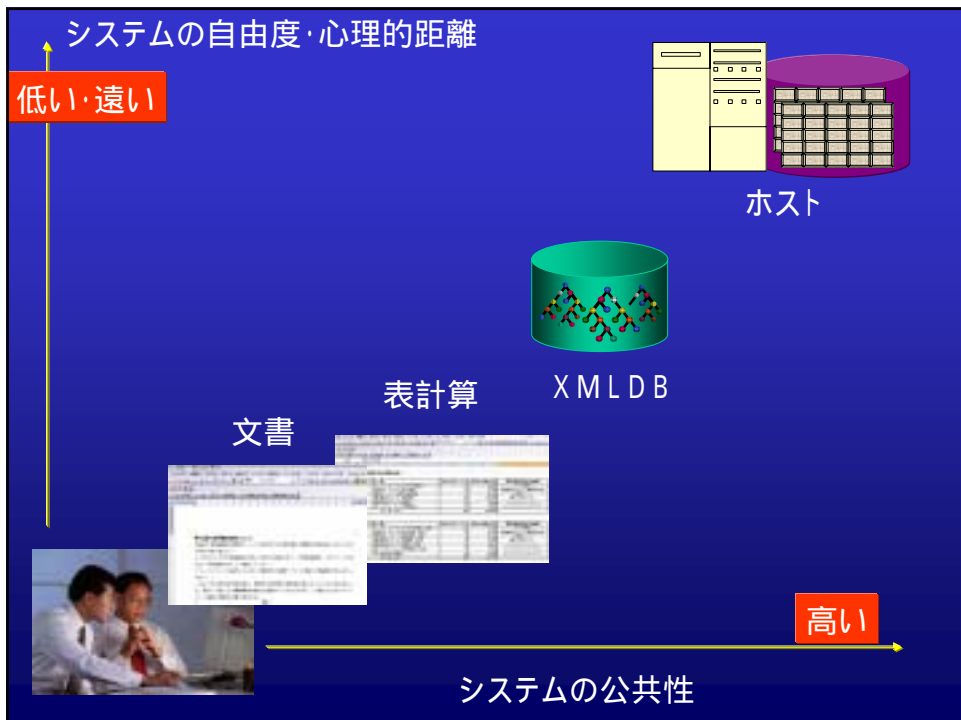
```



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

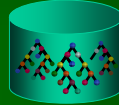


非定型データ

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齢の不健康な食生活からもたらされる体力不足が目立つ。
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、個性的な倦怠感と精神的な焦りをうったえる方が多かった。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

Cyber Luxeon 事例紹介

エンジン整備作業記録の電子化

石川島播磨重工業殿

宇宙航空事業本部 生産管理部
宮下 幹夫様による「XMLDB勉強会
(2007年1月10日)講演より抜粋

システム化の背景

整備作業、指示記録の電子化ニーズ

- ・顧客である航空会社及びその所轄機関へ管理精度向上に向けた取り組みが必要。
- ・管理制度向上のため文章をデータとして扱えるシステムが必要とされた。
- ・従来からEXCEL(VBマクロ)で管理していたが、データ量の増加に対応出来なくなった。

システム化の要件

作業指示記録の電子化ニーズ

階層型DB

整備(指示と記録)情報はエンジンを構成する各階層レベルに対応することが必要であり、必然的に『DBの階層構造化が可能なこと』が条件。

関連システムとのリンク

重複入力を防止するために、現在稼働中システムとのリンクがリアルタイムに可能であること。



XMLDB導入成果

1. 整備情報管理
エンジン構成を表す階層構造は最小で3、多い場合は数十階層に細分化
XMLデータベースは理論的には無制限であり、且つ容易に構造変更が可能
当初は40階層まで検討したが、管理コストを考慮して現在は5階層程度
2. スキーマレスによる項目変更へのフレキシビリティの高さ
短期間での開発においては、確定済み仕様と検討中の仕様を混在させながら進めざるを得ないケースが多い。
今回も同様の状況下での開発を進めたが、項目増減(DB変更)が日程に影響を与えることなく開発を可能にした。
3. リソースの有効活用
既に既存システムの端末が数多く設置されていることから、現有機器を使用することを前提にWEB系システムとし、これと親和性の高いXMLデータベースを採用することでリソースの有効活用を実現。
4. 電子帳票化対応
XML+FO&PDFを使っでの帳票設計・出力・保管が容易に、且つ多種多様な整備記録や作業記録の電子帳票化要求への対応が可能。

今後の計画

1. レスpons向上への取り組み
データの増加と共に、レスponsが悪くなるのはどんなデータベースでも同じ。
構造変更が簡単に出来る強みを生かし、レスpons向上に向けた構造化を検討。
2. 手順書 / 記録の電子化推進
作業現場には多くの機種と、それらに対応した数多くの手順書が存在している。
手順書自体は構造化された文書であり、その内容は大きく『指示』と『記録』。
ドキュメント(指示)とデータ(記録)の統合はXMLの主機能であり、今後発展性のある課題。
3. 既存システムとの再編
稼働中システムは汎用系～サーバー系(RDB/XML)と多岐に亘っており、それぞれの強みを生かした再編(棲み分け)とリンクの推進。
(統合は理想的な姿ではあるが、データの特徴により必ずしもベストではない)

社内業務規定集のWEB化

福岡銀行殿

福岡銀行 業務規定集、マニュアルのWEB化

銀行内の規則や規定、銀行業務の預金・為替・融資・ローン・外国為替などに対する事務手続文書、リファレンスマニュアルのWEB化

従来は紙ベースで60冊(1万数千ページ)

- ・管理コスト(印刷コスト、配布コスト、保管スペース)
- ・改訂と差替のタイムラグ、部分差し替えの手間
- ・参照性、検索性が弱く使い勝手が悪い

基幹システム刷新のタイミングで規定文書の電子化を検討

文書の単なるPDF化ではなく、
規程類の各コンテンツ(文章や図等)を再利用が可能な部品
単位のXMLデータ化し、Cyber Luxeonに格納してDB管理。
全行員はWEBから閲覧、参照する。



「福岡銀行 規程集電子化システム」

現在閲覧している事務取扱要領に関連する項目はリンクで繋がれており、1回のクリックで容易に確かめることができる

ワード、エクセル、PDFなどで作成されたコンテンツはXMLオブジェクトとしてCyber Luxeonに格納。それぞれ関連付けが行われる。

このデータはネットワークを介して各店舗のPC上でブラウズできる。

XMLデータなので表形式をフロー形式に切り替えるなど多様な表示方法が可能。

DBサーバーはPCサーバ1台で運用、開発はJava,全開発工期1年半。

規程類は複雑な構造をしているが、XMLで構造を定義できればどんな複雑な構造データでもCyber LuxeonはDB化できる。

成果

- ・印刷・配布・差し替えにかかるコストが年間で5000万円程度削減。
- ・サーバでデータが更新されると同時に各店舗で最新が閲覧可能。
- ・データをリンクさせることにより、リファレンス性も大幅に改善。
- ・閲覧数は伸び続け、1日6,000件以上。
- ・ログから閲覧箇所と頻度とがわかるため、現在の業務で重点的に必要とされている規程の種類が、本部側から一目で把握できる。



今、必要とされている銀行のサービスの形が見えてくる。
これをCRMシステムと連携させ、顧客満足度向上のための、より戦略的なシステムに発展させていく予定。

その他の事例

プラントドキュメントのXML管理

大手プラントメーカー

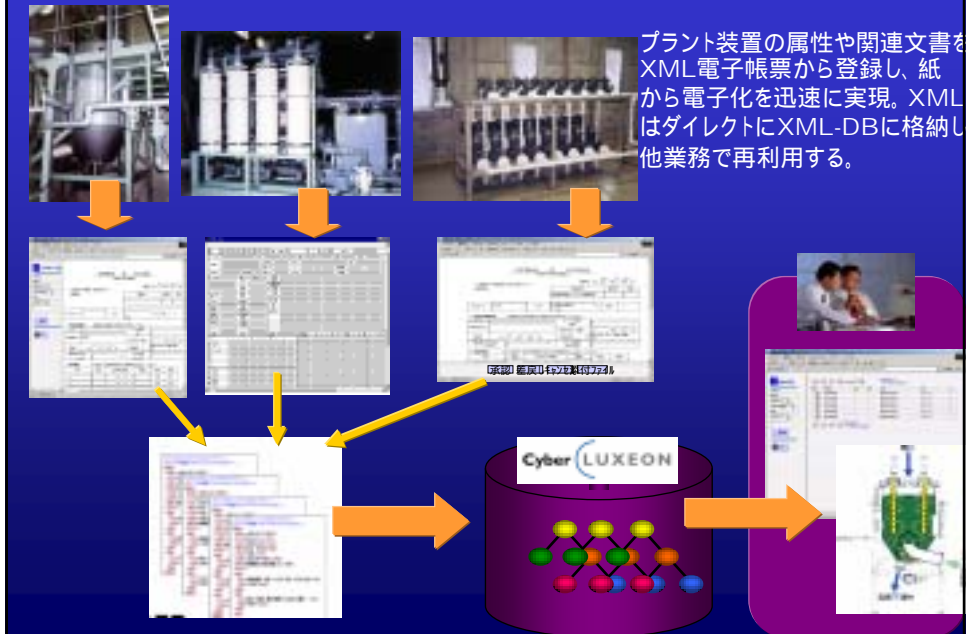
用途 : プラントデータ管理システム

特徴 : プラント工事に関わるドキュメントや部品素材をXMLで表現しXIS (6サーバー) に格納して管理。

意義 : プラントにかかわる図面や各種資料などの複雑な構造をXMLで表現。データ入力にはXML電子帳票を千シート以上作成しワークフロー化。入力値はXML化され、そのままCyber Luxeonに格納管理。

実行環境 : Cyber Luxeon on Windowsサーバ (2CPU x 2)

プラント装置の関連データ・文書を即・電子化

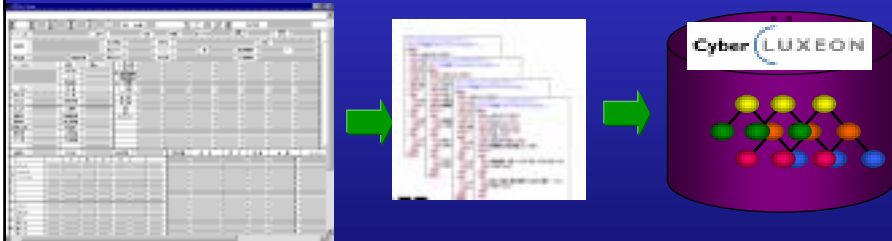


複雑なデータもXML-DBがDB化

XML対応電子帳票との連携

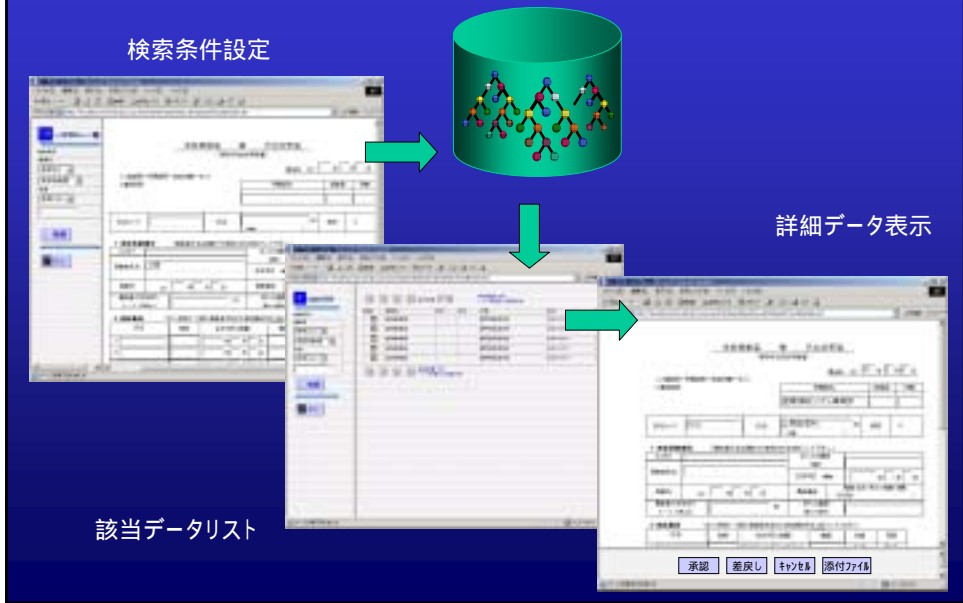
電子帳票への入力値が自動的にXML化

ダイレクトにDB化

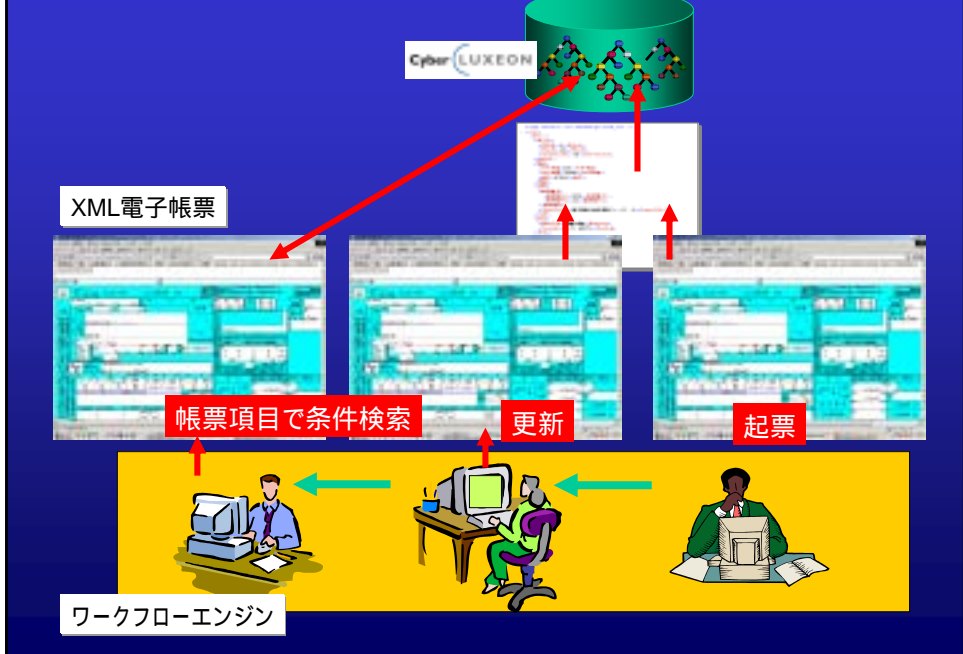


電子申請、電子政府、業務帳票の電子化

電子フォームがそのまま検索画面、表示画面に。



電子帳票とワークフローによる企業データ統合



自社製品仕様のXML統合と見積支援

大手住宅設備メーカー

用途 : 自社住宅設備製品の仕様管理と見積支援システム

内容 : 多品種複雑な製品仕様をXMLで柔軟に定義し、そのままCyber Luxeonに格納して管理。

意義 : 設備製品の組み合わせのノウハウまでXML DB化して営業見積もりの支援機能を実現した。
1日10,000人程が利用する大規模運用。

実行環境 : Cyber Luxeon on SUNサーバー (2CPU x 2)

プログラム仕様のXML管理

大手精密機器メーカー、大手機械メーカー

用途 : 機器組み込みプログラム仕様管理システム

内容 : 文書で記述されたプログラム仕様をXMLに変換しそのままCyber Luxeonに格納管理。
内部で利用されるコンポーネントやデータの関係性をWEBで可視化する。

意義 : 仕様書の部品化、再利用化によるプログラム開発の生産性向上

実行環境 : Cyber Luxeon on Windowsサーバ (1CPU x 1)

自社製品へのクレーム情報管理

大手精密機器メーカー

用途 : 自社製品クレーム管理システム

特徴 : 自社製品に関わるクレームデータをXMLで表現し管理。

意義 : クレーム情報というあらかじめ項目が見極めがたいデータのDB化のためにXMLを利用し、イレギュラーな項目を後づけで継ぎ足していくためにXML-DBに格納管理する。将来はCRMシステムに発展していく。

実行環境 : Cyber Luxen on Windowsサーバ (1CPU x 4)

システムのログデータ管理

大手通信業社

用途 : システムのログデータ管理

特徴 : システムのログデータをXMLで記述してCyber Luxeonに格納管理。

意義 : 項目が多様であるあるシステムのログをXML化して、高速検索する。市販されているほとんどのXML-DBを実際にベンチマークした上で、検索、格納、開発環境のバランスに富んだCyberLuxeonを採用した。

実行環境 : Cyber Luxen on Linuxサーバ (2CPU x 1)

プラント設備保全システムのXML化

大手石油メーカー

用途 : 自社石油プラント設備保全システムでの最適XML DB化

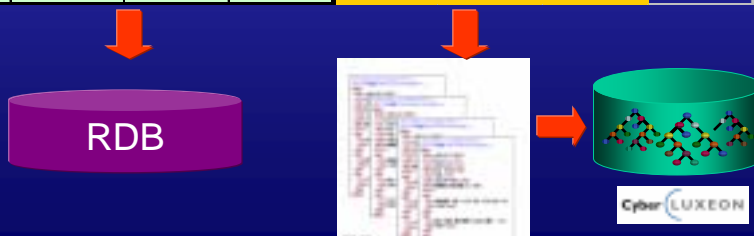
特徴 : 設備の基本属性はRDBで管理するがテーブルからはみ出る設備ごとの個別属性はXMLで定義し、Cyber Luxeonに格納して管理。またXMLにて外部ERPシステムとのデータ連携を行う。

意義 : RDBを補完する柔軟なXML DBの使用例。

実行環境 : Cyber Luxeon on Windowsサーバ (1CPU x 1)

RDBとXML DBの併用

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



自社製品仕様のXML統合とWEBカタログ化

大手金属メーカー

用途 : 自社製品の仕様管理とWEBカタログ化

内容 : 多品種複雑な製品仕様をXMLで柔軟に定義し、そのままCyber Luxeonに格納して管理。

意義 : 製品属性をXMLで、図面をSVGで統合してWEBカタログ化。その際製品選のナレッジまでシステム化に成功。WEBカタログをCD化して販促にも利用。

実行環境 : Cyber Luxeon on Windowsサーバー (2CPU x 1)

製品情報の統合DBに。

CyberTech



WEBカタログサイト



まとめ

XMLDBのユーザ像

XMLDBユーザの特質

- ・業務への真摯な情熱(負けず嫌い)
- ・現状への問題意識(常に現状に不満足)
- ・チャレンジブル(新しもの好き)
- ・リーダーシップ(妥協できない)
- ・ビジョンを追う(思い入れの強さ)
- ・構造的思考(性格は割りとアバウト)
- ・現場志向(在野志向)

熱いエンジニア

団塊の世代



株式会社 サイバーテック
営業統括部 部長

取締役 加藤 哲義

Tel : 03-5577-8000

e-Mail : katou@cybertech.co.jp

<http://www.cybertech.co.jp>

本書に記載されている商品名、会社名などの固有名詞は、各社の商標及び登録商標です。
本書の内容の一部またはすべてを、当社の事前の許可なく複製や再利用することを禁じます。

