

# PostgreSQL の XML 機能解説 と将来拡張への提言

2009年11月21日 土曜日  
11:00 ~ 11:50  
トラック B

吉田 晃伸、安部 みのぶ



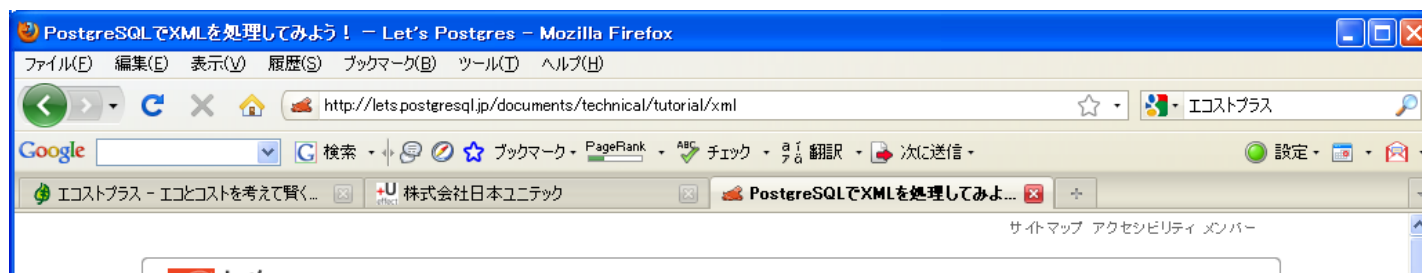
# PostgreSQL のポータルサイト "Let's Postgres"





# PostgreSQL でXMLを処理してみよう！

URL  <http://lets.postgresql.jp/documents/technical/tutorial/xml>



| マップ | 読み物 | ニュース | イベント

## ■ PostgreSQLでXMLを処理してみよう！(最終回)(1)

第6回 (最終回) PostgreSQLによるXML処理落ち穂拾い

-XML Schema検証アプリケーションとの連携、PostgreSQL のXML名前空間対応-

響 楽 人

たXMLデータを取り出す時に、様々な条件を付けて、XMLデータの特定の一部分にアクセスすることのできるXPathの記述法について説明します。PostgreSQLでは8.3から標準でXPath式を評価する機能が追加されています。

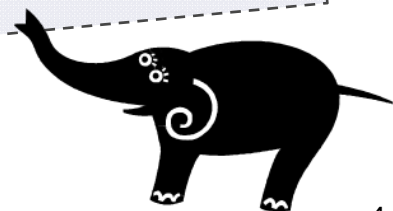
■ PostgreSQLでXMLを処理してみよう！(第3回) 作成者 Yuki Kamado ー 最終変更日時 2009年07月30日 10時03分

SQLでは、SELECT文によって、条件での絞り込みや集計を行ってデータを取り出すことができますが、PostgreSQL8.3では、この機能をXMLデータに対しても適用することができます。この記事では、RDBとしてのPostgreSQLの特徴を活かし、XMLデータとRDBデータをうまく組み合わせたデータベース利用について考えます。さらに、PostgreSQLのインデックス機能利用し、XMLデータの検索を高速化する

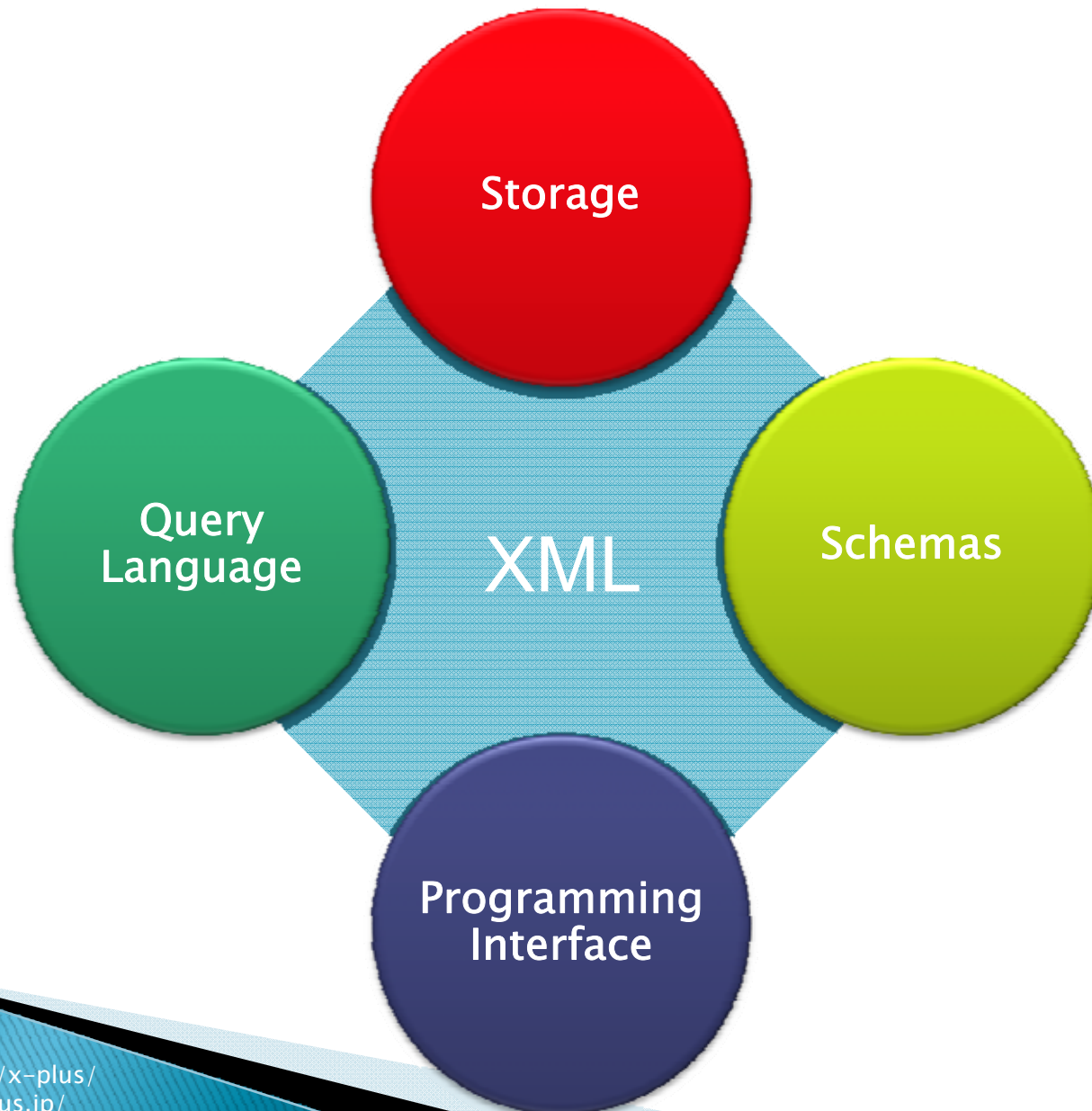
完了

# Table of Contents

- a. XMLとデータベースの関係概論
  - XMLを扱う際に必要なこと
  - XMLとデータベース
- b. PostgreSQLとXMLの関係
  - PostgreSQLのXMLサポートの歴史
  - PostgreSQL8.3 - XML対応の詳細
- c. 具体的な機能紹介とデモンストレーション
  - 1. XpathによるXMLデータの検索機能
  - 2. GINインデックスによる高速検索機能
  - 3. xml2追加モジュールのXSLTプロセッサ機能
  - 4. リレーショナルデータからXMLを生成するSQL/XML機能
- d. 機能拡張提言
  - PostgreSQLでサポートしている機能の範囲
  - スキーマ検証の重要性
  - 将来拡張への提言



# XMLを扱う際に必要な要素



# XMLとデータベース

## ▶ XMLをデータベースで扱う主な方法

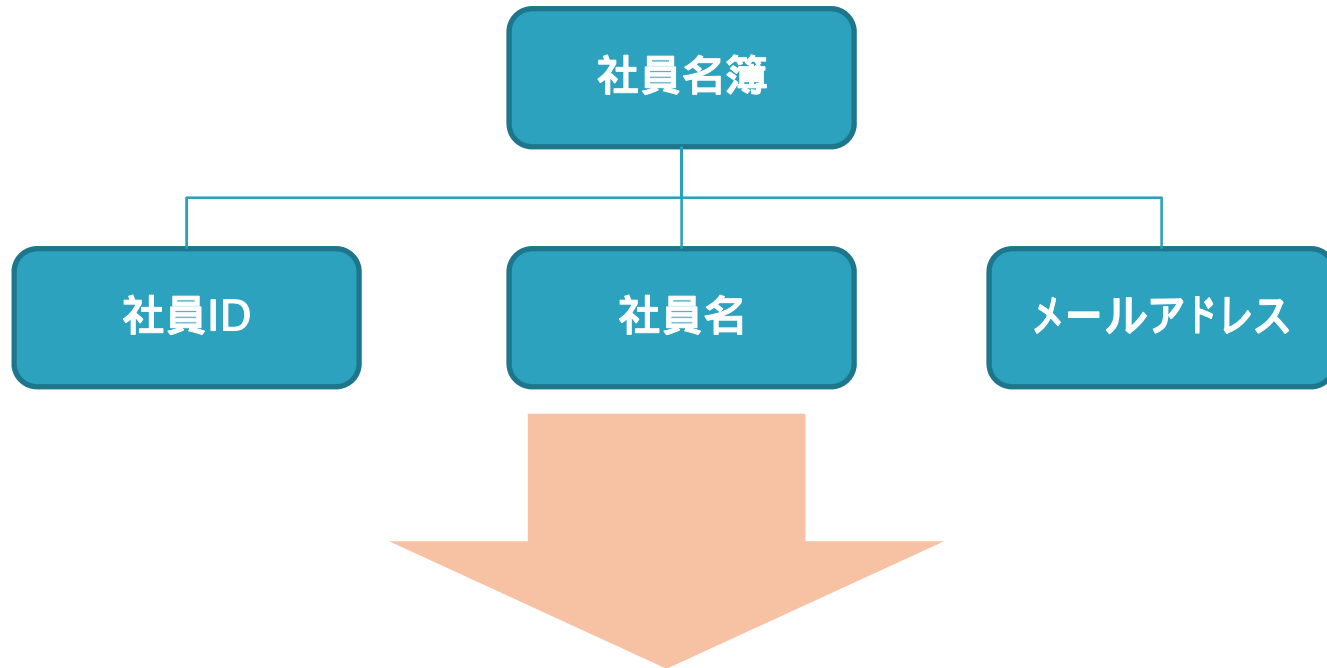
### A) RDB

1. テーブルマッピング
2. データ型としてのサポート

### B) ネイティブXMLデータベース

# XMLとデータベース

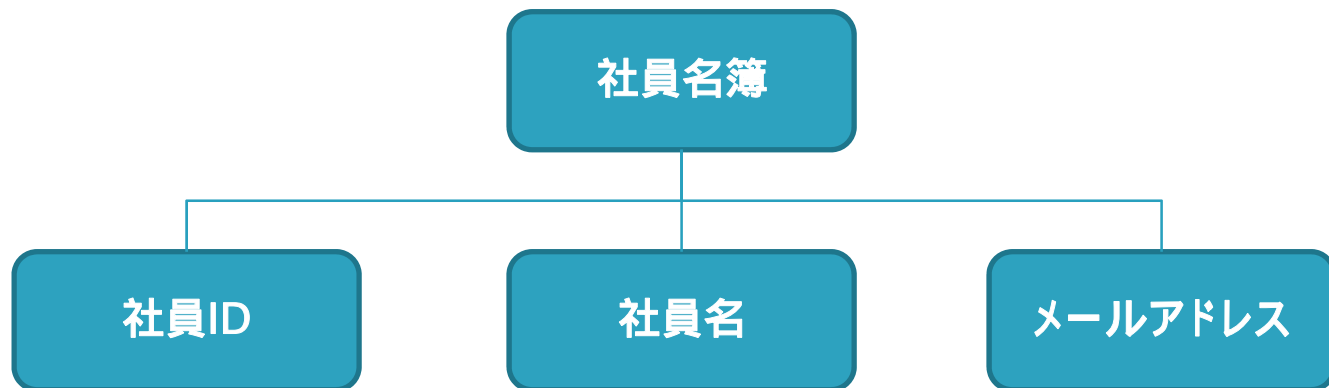
## RDB - テーブルマッピング






社員ID	社員名	メールアドレス
12345678	吉田 晃伸	ayoshida@XXX.co.jp
98765432	安部 みのぶ	mabe@XXX.co.jp
8383838383	響 楽人	ghibiki@XXX.co.jp

# XMLとデータベース

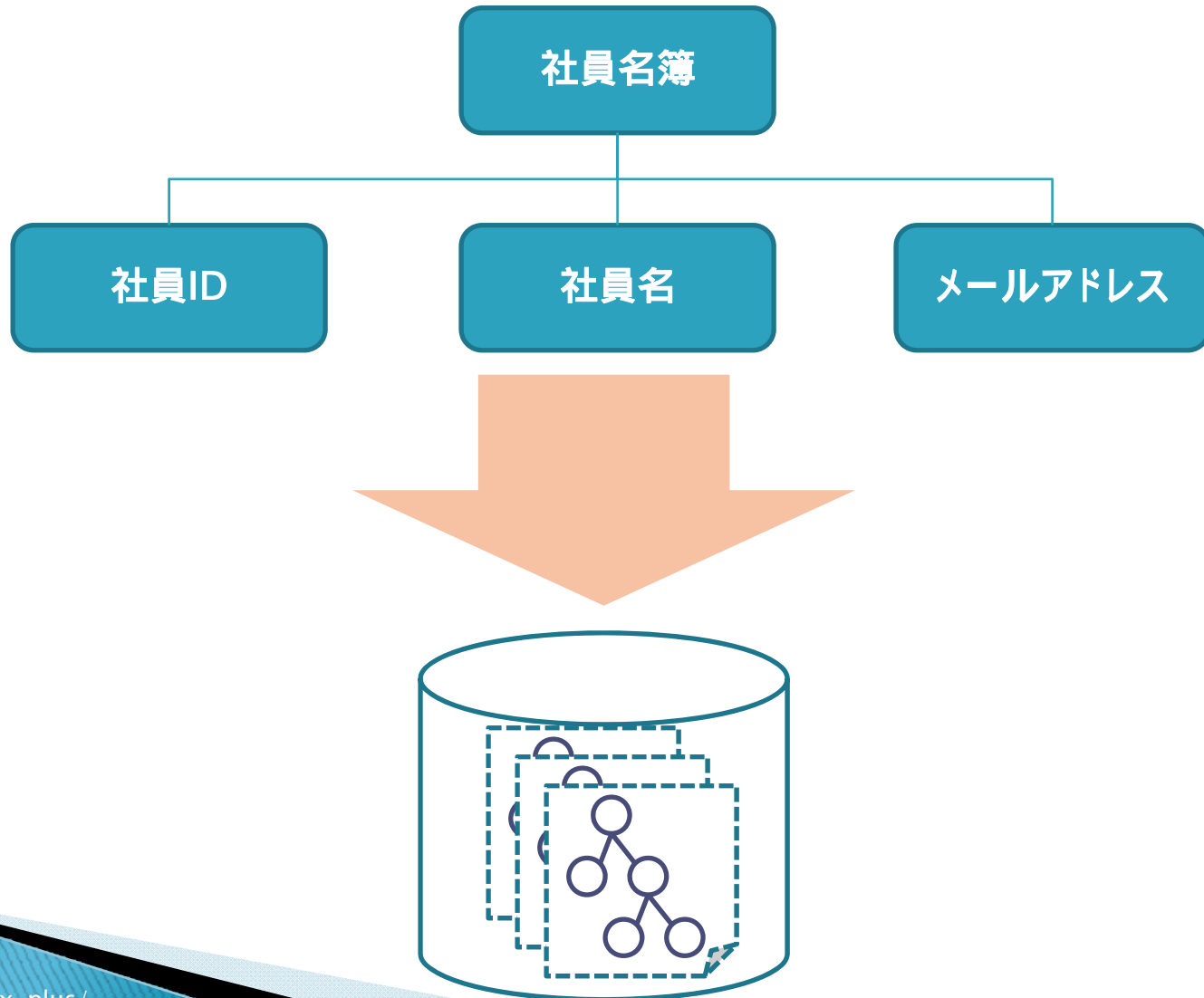
## RDB - データ型としてのサポート



社員ID	XMLデータ
12345678	
98765432	
8383838383	

# XMLとデータベース

## ネイティブXMLデータベース



# XMLとデータベース

## ▶ XMLをデータベースで扱う主な方法

### A) RDB

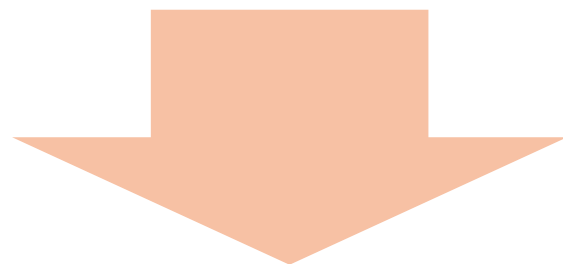
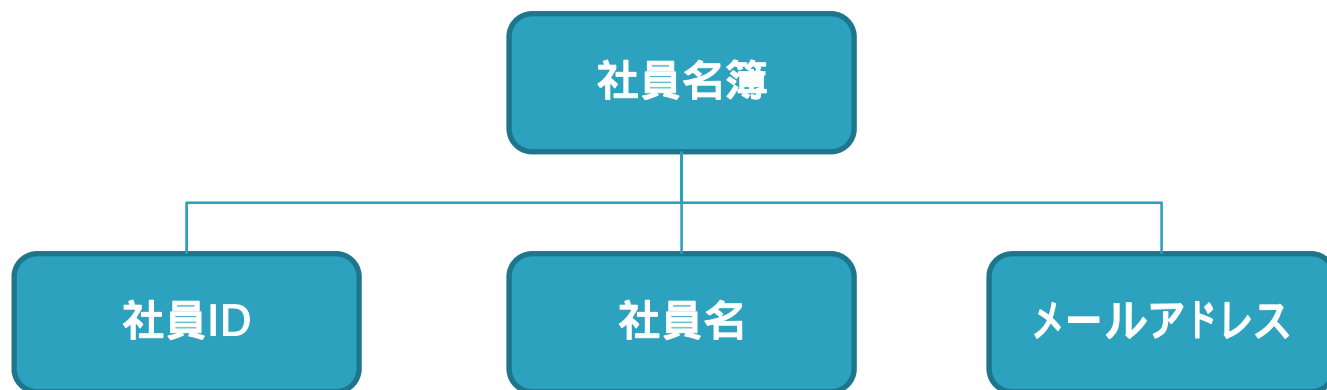
1. テーブルマッピング
2. データ型としてのサポート




### B) ネイティブXMLデータベース



# PostgreSQLでXMLを扱う方法

## RDB - データ型としてのサポート



社員ID	XMLデータ
12345678	
98765432	
8383838383	

# PostgreSQLのXMLサポートの歴史

標準機能でのサポート

2008年2月

8.3から標準機能としてXML型や検索機能をサポート

8.4でさらなる機能拡張？！

拡張機能でのサポート

2002年2月

7.2で初めてXML拡張モジュール導入

2005年1月

8.0にXML2モジュール追加(※)

※XpathとXSLT機能

ミドルウェアでのサポート

2001年 XMLPGSQL開発

その他、多数

# PostgreSQL 8.4 — XML対応の詳細

- ▶ データ型「XML型」のサポート
- ▶ XMLデータ検索「XPath」のサポート
- ▶ RDB→XML変換「SQL/XML」のサポート

# PostgreSQL 8.4 — XML対応の詳細

## ▶ データ型「XML型」のサポート

```
CREATE TABLE meibo(id integer, detail xml );
```

id (社員ID) ( integer )	detail【詳細】 ( xml )
12345678	<name>吉田 晃伸</name> <email>ayoshida@---</email>
98765432	<name>安部みのぶ</name> <email>mabe@---</email>
8383838383	<name>響 楽人</name> <email>ghibiki@---</email>

# PostgreSQL 8.4 — XML対応の詳細

## ▶ XMLデータ検索「Xpath」のサポート

```
SELECT
xpath('/employee[name/text()='吉田 晃伸'], detail)
FROM meibo;
```

id (社員ID) ( integer )	detail【詳細】 ( xml )
12345678	<code>&lt;name&gt;吉田 晃伸&lt;/name&gt;</code> <code>&lt;email&gt;ayoshida@---&lt;/email&gt;</code>
98765432	<code>&lt;name&gt;安部みのぶ&lt;/name&gt;</code> <code>&lt;email&gt;mabe@---&lt;/email&gt;</code>
8383838383	<code>&lt;name&gt;響 楽人&lt;/name&gt;</code> <code>&lt;email&gt;ghibiki@---&lt;/email&gt;</code>

# PostgreSQL 8.4 — XML対応の詳細

## ▶ RDB→XML変換「SQL/XML」のサポート

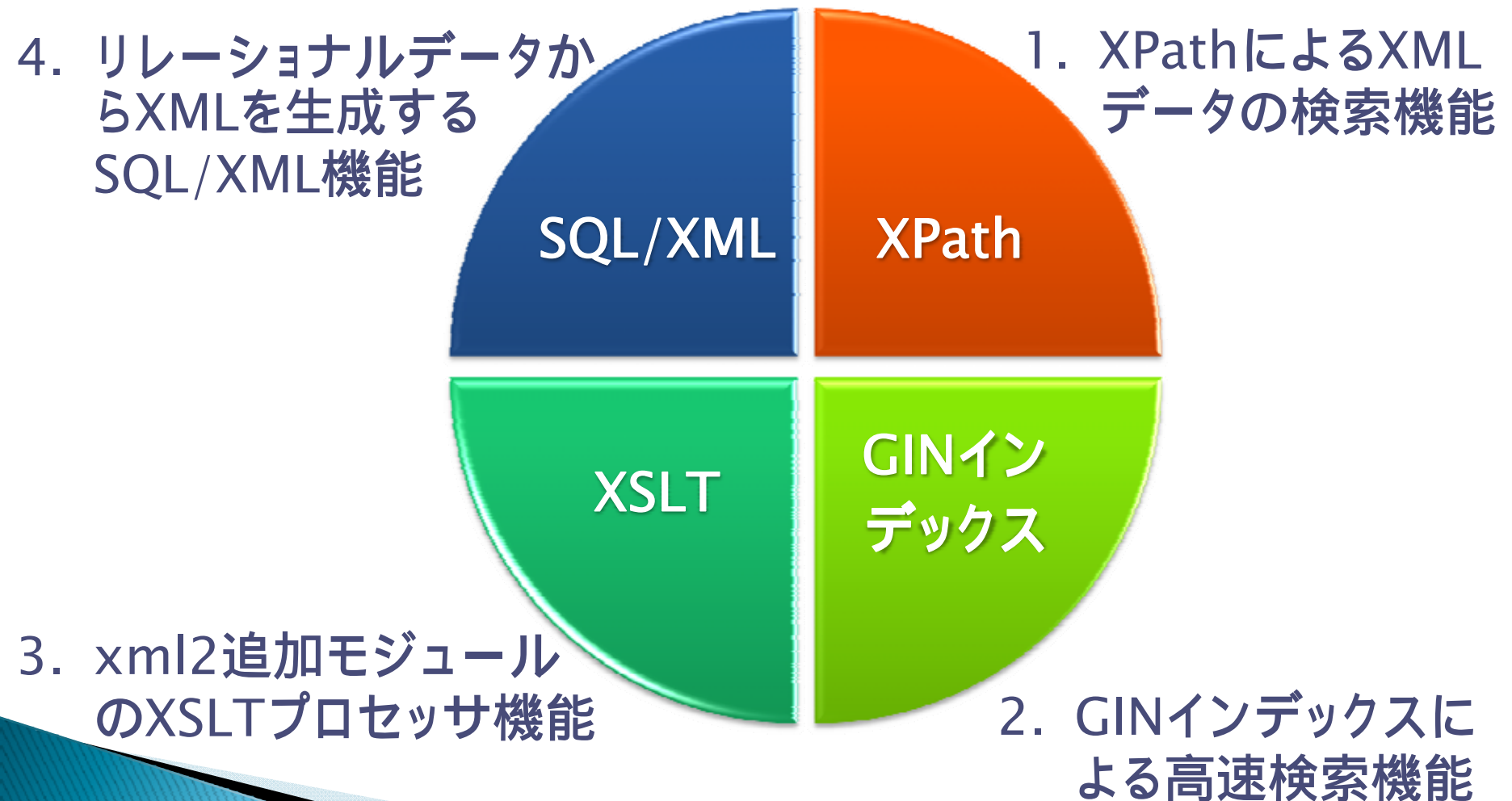
Id (社員ID)	Employee (社員名)	email (メールアドレス)
12345678	吉田 晃伸	ayoshida@XXX.co.jp
98765432	安部 みのぶ	mabe@XXX.co.jp
8383838383	響 楽人	ghibiki@XXX.co.jp

```
select xmlelement (name email,email) from meibo;
```



```
<email>ayoshida@XXX.co.jp</email>  
<email>mabe@XXX.co.jp</email>  
<email>ghibiki@XXX.co.jp</email>
```

# PostgreSQLの具体的なXML機能紹介、 およびデモンストレーション



# デモで使用するデータについて

## 社員情報テーブルを想定

基本情報      リレーショナル      職務経歴      XML

id	name	department	age	email	work_history
1031	生田 靖	IT事業統括部	45	ikuta@abcd.com	<work-record>
1088	前田 孝	ネットワーク 管理課	41	maeda@abcd.com	
1101	鈴木 花子	開発一課	33	suzuki@abcd.com	

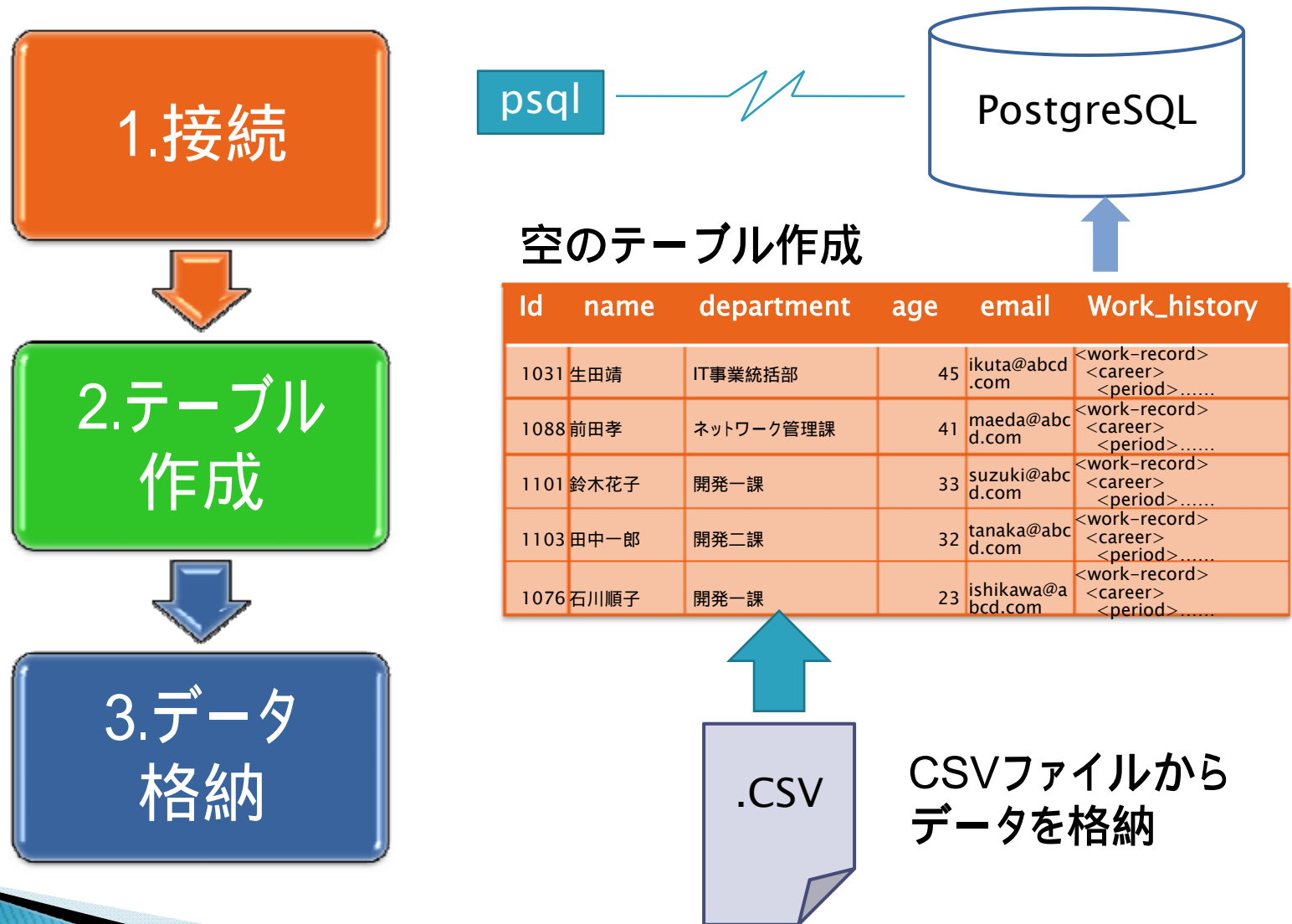
詳細

```
<work-record>
  <career>
    <period>
      <start>1970-04-01</start>
      <end>1980-03-31</end>
    </period>
    <content>製造業向け購買管理システムの開発、大手流通業向け販
    売管理システム開発</content>
    <role>プログラマー</role>
  </career>
  <career>
    <period>
      <start>1980-04-01</start>
      <end>1985-03-31</end>
    </period>
    <content>金融機関向け社内Webシステム開発</content>
    <role>プロジェクトリーダー</role>
  </career>
  <career>
    <period>
      <start>1985-04-01</start>
      <end/>
    </period>
    <content>IT戦略部にて企画提案業務</content>
    <role>マネージャー</role>
  </career>
</work-record>
```

### (使用するシステム)

- PostgreSQL 8.4.1
- OS: Windows XP

# デモを始める前に...



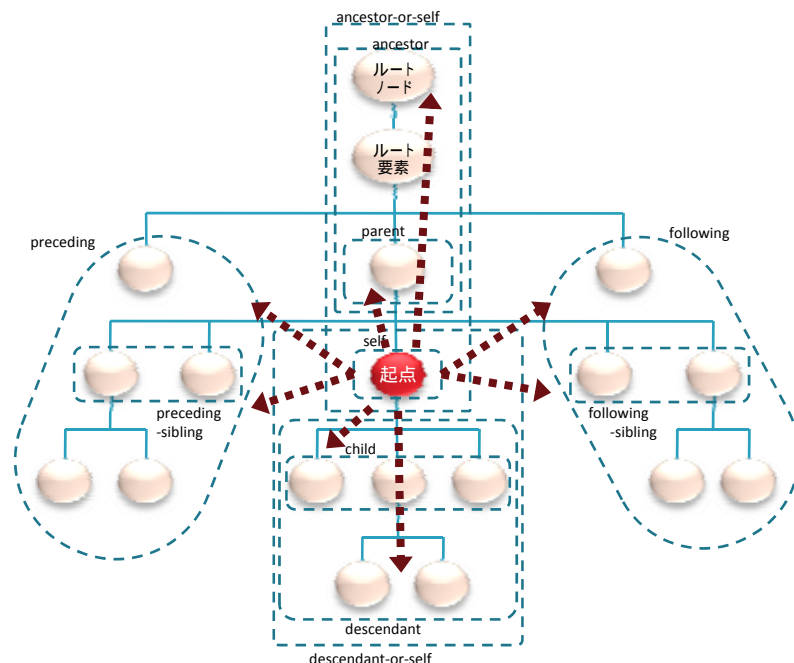
# 1.XPathによるXMLデータの検索機能

## 構造情報

```
<work-record>
  <career>
    <period>
      <start>1970-04-01</start>
      <end>1980-03-31</end>
    </period>
    <content>製造業向け購買管理システムの開発、
    大手流通業向け販売管理システム開発</content>
    <role>プログラマー</role>
  </career>
  <career>
    <period>
      <start>1980-04-01</start>
      <end>1985-03-31</end>
    </period>
    <content>金融機関向け社内Webシステム開発
  </content>
    <role>プロジェクトリーダー</role>
  </career>
  <career>
    <period>
      <start>1985-04-01</start>
      <end/>
    </period>
    <content>IT戦略部にて企画提案業務
  </content>
    <role>マネージャー</role>
  </career>
</work-record>
```

値

## XMLのツリー構造モデル



## ロケーションパス

ロケーションステップ

ロケーションステップ

/work-record/career[role/text()="マネージャー"]

ノードテスト

述語

式

# XPath:デモ

- ▶ 「マネージャー」を経験したことがある人のid,name,<role>を抽出

Id	name	department	age	email	Work_history
1031	生田靖	IT事業統括部	45	ikuta@abcd.com	<pre>&lt;work-record&gt; &lt;career&gt;   &lt;period&gt;     &lt;start&gt;1970-04-01&lt;/start&gt;     &lt;end&gt;1980-03-31&lt;/end&gt;   &lt;/period&gt;   &lt;content&gt;製造業向け購買管理システムの開発、大手流通業向け販売管理システム開発&lt;/content&gt;   &lt;role&gt;プログラマー&lt;/role&gt; &lt;/career&gt; &lt;career&gt;   &lt;period&gt;     &lt;start&gt;1980-04-01&lt;/start&gt;     &lt;end&gt;1985-03-31&lt;/end&gt;   &lt;/period&gt;   &lt;content&gt;金融機関向け社内Webシステム開発&lt;/content&gt;   &lt;role&gt;プロジェクトリーダー&lt;/role&gt; &lt;/career&gt; &lt;career&gt;   &lt;period&gt;     &lt;start&gt;1985-04-01&lt;/start&gt;     &lt;end&gt;   &lt;/period&gt;   &lt;content&gt;IT戦略部にて企画提案業務&lt;/content&gt;   &lt;role&gt;マネージャー&lt;/role&gt; &lt;/career&gt; &lt;/work-record&gt;</pre>

- ▶ XML型のままでは比較演算子「=」で検索値と比較することができない

**XPath式** → **変換** → **テキスト配列型**

# 2.GINインデックスによる高速検索機能

↓  
検索処理の高速化

一般的には、

B-Tree

XML(配列)型には、、

GIN

なぜ、GINインデックス？

XPath式「xpath('/work-record/career/role/text()',work\_history)」

結果

xpath

-----  
{プログラマー,プロジェクトリーダー,マネージャー}  
{サブリーダー  
{プロジェクト  
{運用メンバー  
{運用メンバー}

ノード集合≡配列型

# GIN: デモ

Id	name	department	age	email	Work_history
1	...	...	...	...	...
2	...	...	...	...	...
3	...	...	...	...	...
...	...	...	...	...	...
299	...	...	...	...	...
300	...	...	...	...	...

XPath式をINDEX登録

```
<work-record>
<career>
  <period>
    <start>1970-04-01</start>
    <end>1980-03-31</end>
  </period>
  <content>製造業向け購買管理システムの開発、大手流通業向け販売管理システム開発</content>
  <role>プログラマー</role>
</career>
<career>
  <period>
    <start>1980-04-01</start>
    <end>1985-03-31</end>
  </period>
  <content>金融機関向け社内Webシステム開発</content>
  <role>プロジェクトリーダー</role>
</career>
<career>
  <period>
    <start>1985-04-01</start>
    <end></end>
  </period>
  <content>IT戦略部にて企画提案業務</content>
  <role>マネージャー</role>
</career>
</work-record>
```

INDEXの値

INDEXの値

INDEXの値

1. Meiboテーブルのデータを300件にする

2. インデックスをつける

3. 検索する

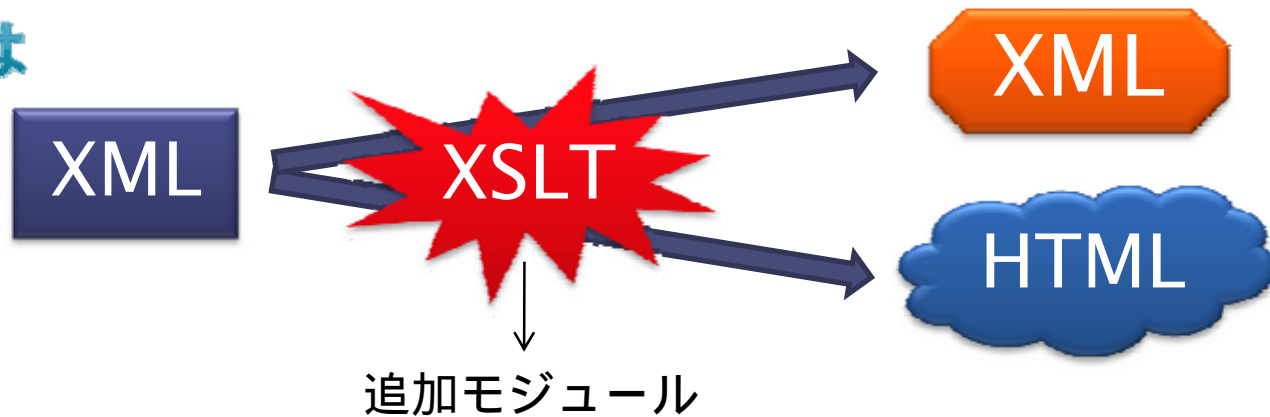
4. インデックスを削除する

5. もう一度検索する

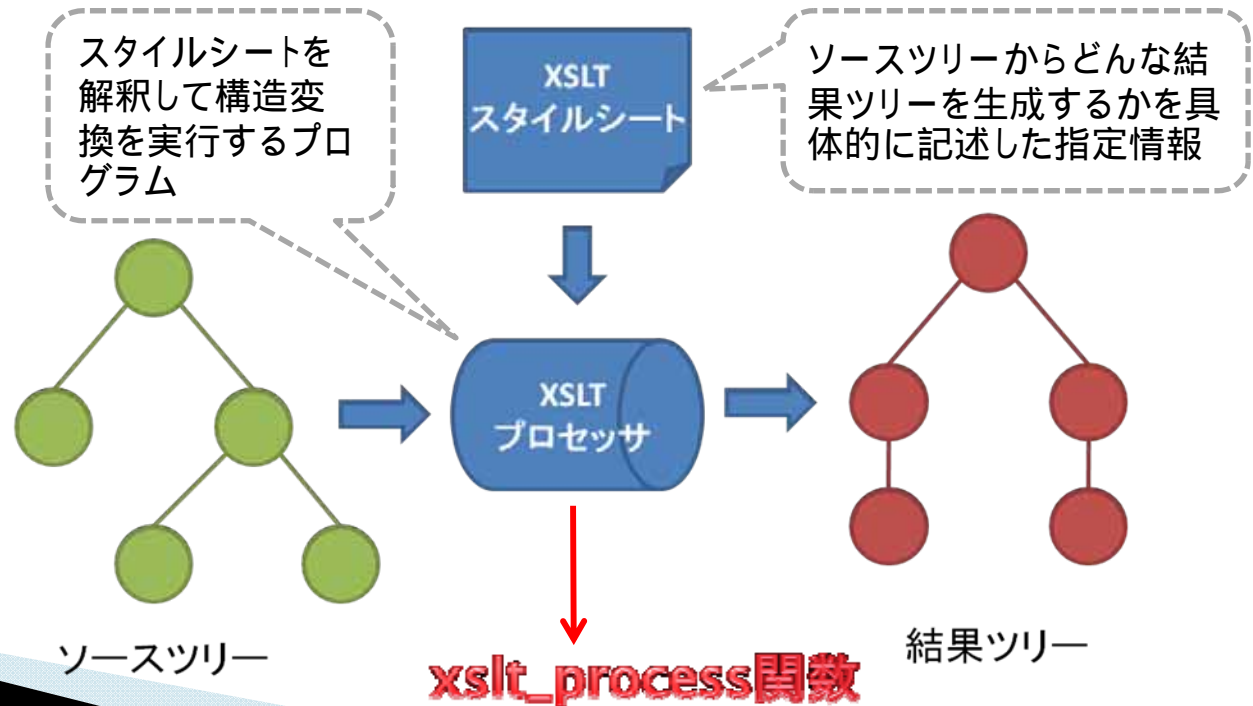
比較演算子「@>」を使って

# 3.xml2追加モジュールのXSLTプロセッサ機能

XSLTとは



変換の仕組み



# XSLT: デモ

## シナリオ

Id	name	departm ent	age	email	Work_history
1031	生田靖	IT事業統括部	45	ikuta@abcd. com	<div>&lt;work-record&gt;<div>&lt;career&gt;<div>&lt;period&gt;<div>&lt;start&gt;1970-04-01&lt;/start&gt;<div>&lt;end&gt;1980-03-31&lt;/end&gt;<div>&lt;/period&gt;<div>&lt;content&gt;製造業向け購買管理システムの開発、大手流通業向け販売管理システム開発&lt;/content&gt;<div>&lt;role&gt;プログラマー&lt;/role&gt;<div>&lt;/career&gt;<div>経歴(carrier) 1</div></div></div></div></div></div></div></div></div><div>出来上がりイメージ</div><div><div>2&lt;start&gt;1980-04-01&lt;/start&gt;<div>&lt;end&gt;1985-03-31&lt;/end&gt;<div>&gt;金融機関向け社内Webシステム開発&lt;/content&gt;<div>プロジェクトリーダー&lt;/role&gt;<div>経歴(carrier) 2</div></div></div></div></div></div><div>経歴(carrier) 3</div><div>&lt;/work-record&gt;</div></div>

Windows Internet Explorer

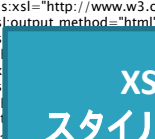
C:\Wpssqlmi\result01.html

ファイル(F) 編集(E) 表示(V) お気に入り(I) ツール(T) ヘルプ(H)

### 職務経歴一覧表

始	至	内容	役割
1970-04-01	1980-03-31	製造業向け購買管理システムの開発、大手流通業向け販売管理システム開発	プログラマー
1980-04-01	1985-03-31	金融機関向け社内Webシステム開発	プロジェクトリーダー
1985-04-01		IT戦略部にて企画提案業務	マネージャー

# XSLT:デモ



XSLT  
スタイルシート

## xslt\_process()

WHERE name = '生田靖'

## 結果ツリー

## 1. スタイルシートを作る

## 2. xslt\_processを実行

### 3. ファイルへ出力する

## 4. ブラウザで開く

# 4.リレーショナルデータからXMLを生成するSQL/XML機能

ISOのSQL:2003標準(XML関連機能)

PostgreSQLで対応済みのもの

生成	関数	機能
	xmlelement()	与えられた名前、属性、および内容を持つXML要素を生成する
	xmlforest()	与えられた名前と内容を使用し、要素のXMLフォレスト(シーケンス)を生成する
	xmlattributes()	XML属性を生成する
	xmlcomment()	特定のテキストを内容としてXMLコメントを含んだXML値を作成する
	xmlconcat()	XMLの内容断片を含む単一の値を作成するための個別のXML値のリストを結合する
集約	xmlpi()	XML処理命令を作成する
	xmlroot()	XML値のルートノードの属性を変更する
	xmlagg()	xmlconcatが行うように、入力値を集約関数呼び出しに連結する
変換	xmlparse()	文字列型の値からXML型の値への変換を行う
	xmlserialize()	XML型の値から文字列型の値への変換を行う

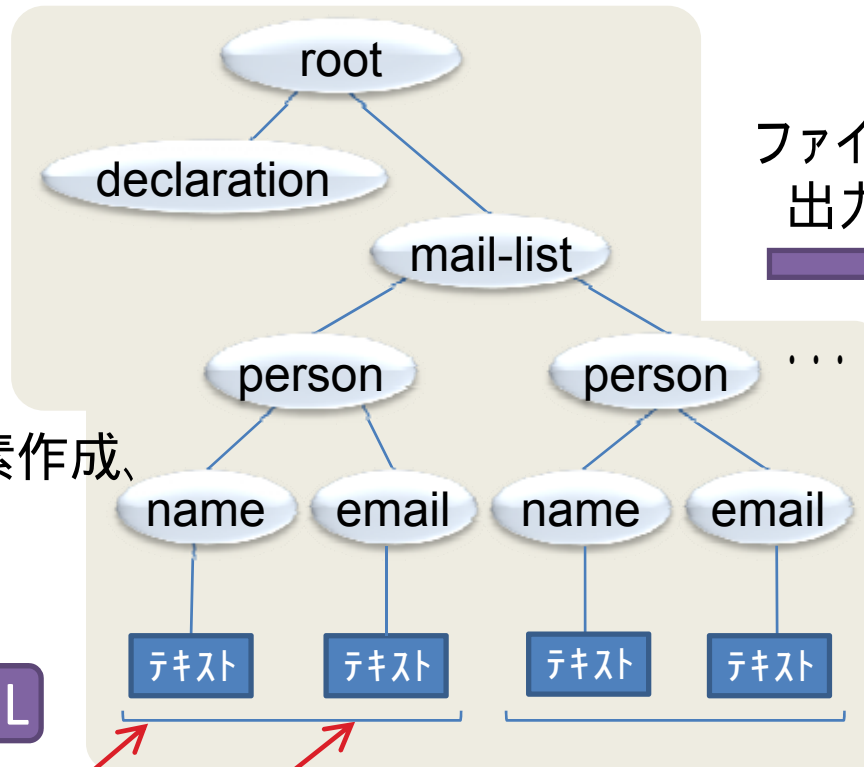
(PostgreSQL 8.4.1文書 8.13.1、9.14.1より)

# SQL/XML: デモ

## シナリオ

データ抽出・要素作成、  
ツリー構築

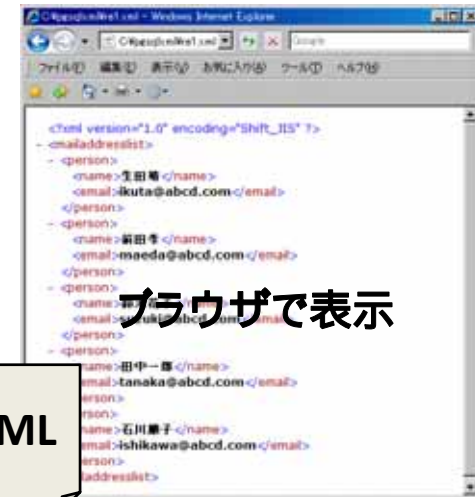
SQL/XML



ファイル  
出力

XML

ブラウザで表示



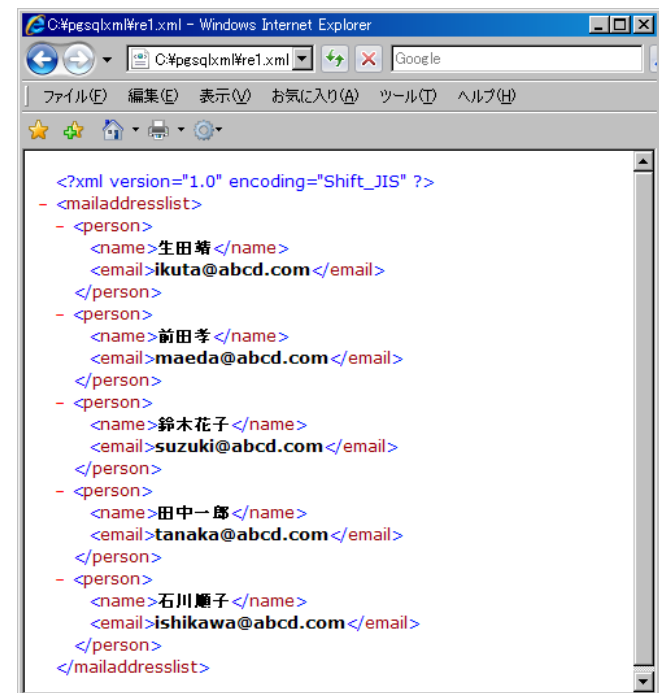
id	name	email	work_history
1031	生田靖	ikuta@abcd.com	<work-record> <career>...

# SQL/XML: デモ

## ▶ リレーショナルデータからXMLを作るSELECT文

◦ select xmlroot(xmlelement(name mail\_list,  
(select xmlagg(  
    xmlelement(name person,  
        xmlelement(name name, name),  
        xmlelement(name email, email)  
    )) from meibo)),version '1.0',standalone yes);

サブクエリー



```
<?xml version="1.0" encoding="Shift_JIS" ?>
- <mailaddresslist>
-   <person>
      <name>生田 靖</name>
      <email>ikuta@abcd.com</email>
    </person>
-   <person>
      <name>前田 孝</name>
      <email>maeda@abcd.com</email>
    </person>
-   <person>
      <name>鈴木 花子</name>
      <email>suzuki@abcd.com</email>
    </person>
-   <person>
      <name>田中 一郎</name>
      <email>tanaka@abcd.com</email>
    </person>
-   <person>
      <name>石川 順子</name>
      <email>ishikawa@abcd.com</email>
    </person>
  </mailaddresslist>
```

# 補足情報

- ✓ XPath、XSLTの詳細
- ✓ XQL/XMLサポートの詳細
- ✓ XML名前空間



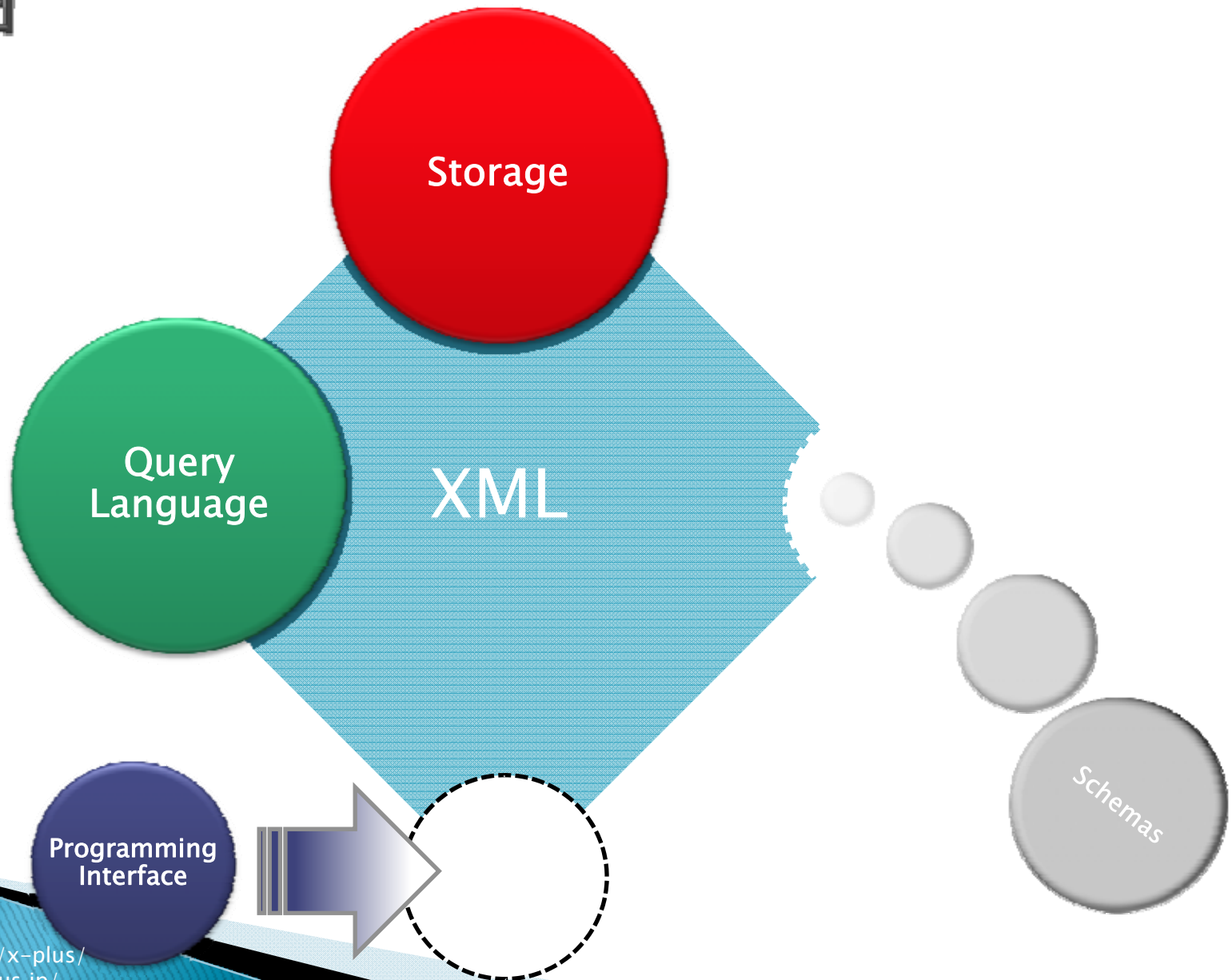
**PostgreSQLでXMLを試してみよう！  
連載(全6回)**

Postgresql XML

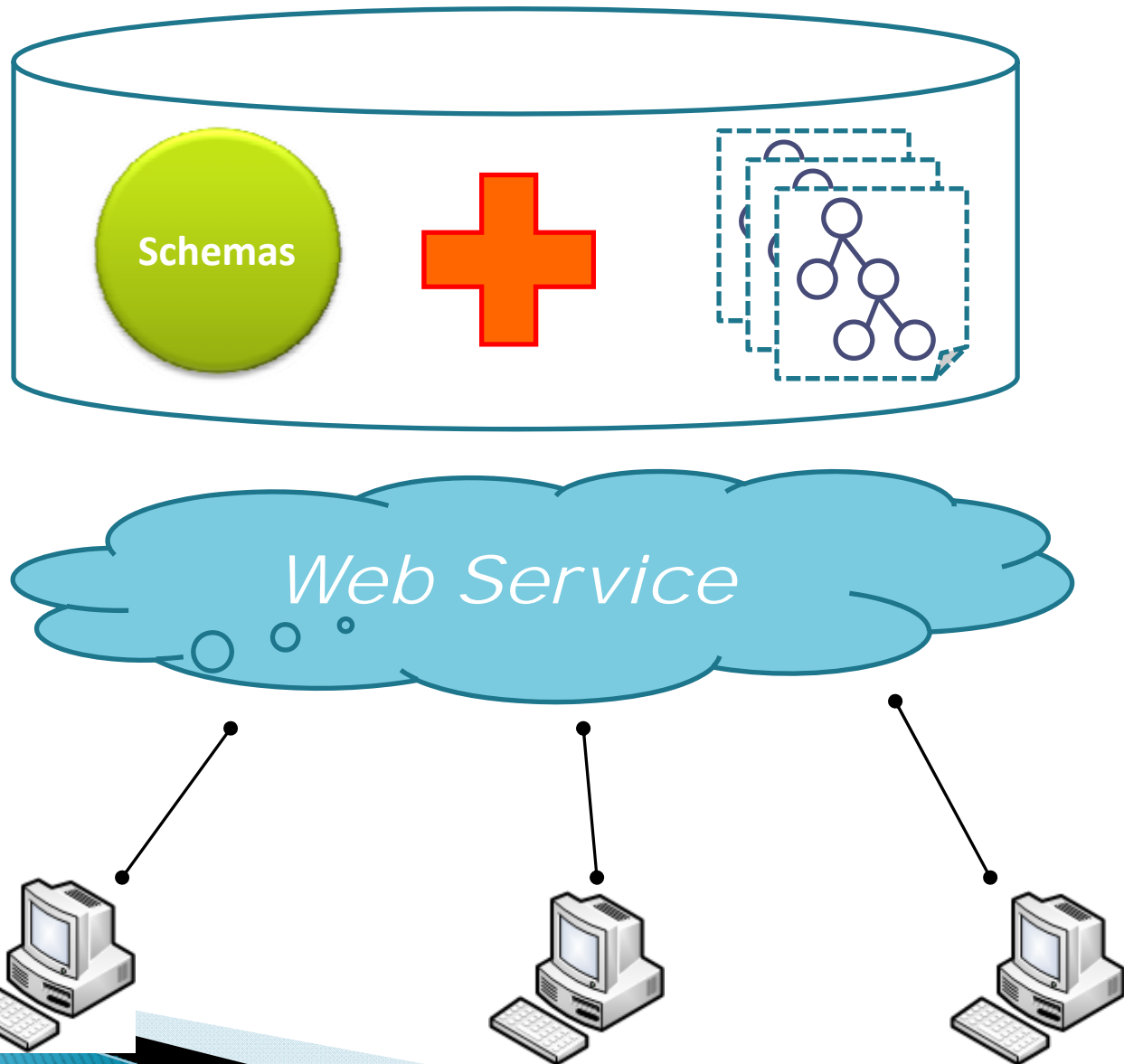
検索



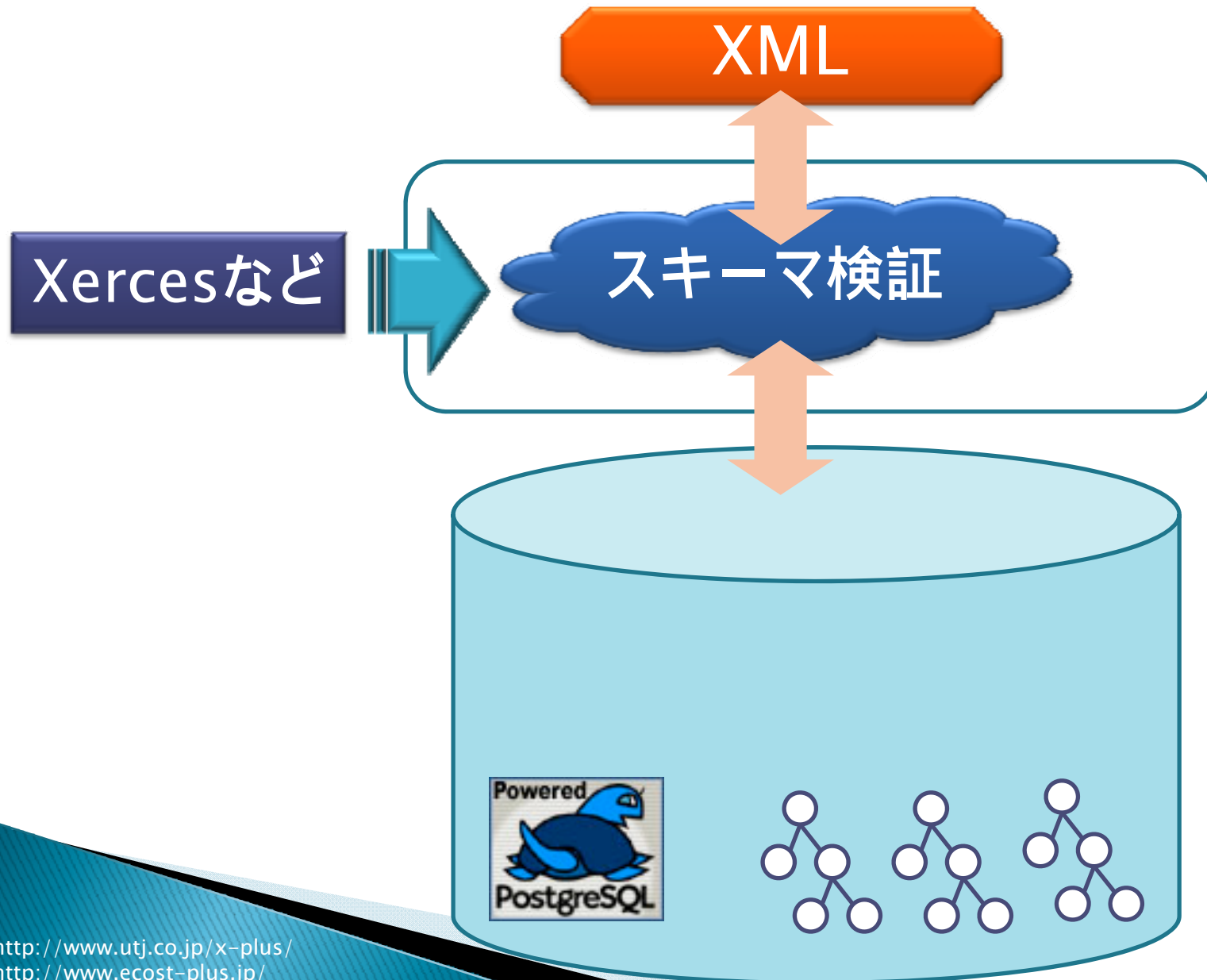
# PostgreSQLでサポートしているXML機能 の範囲



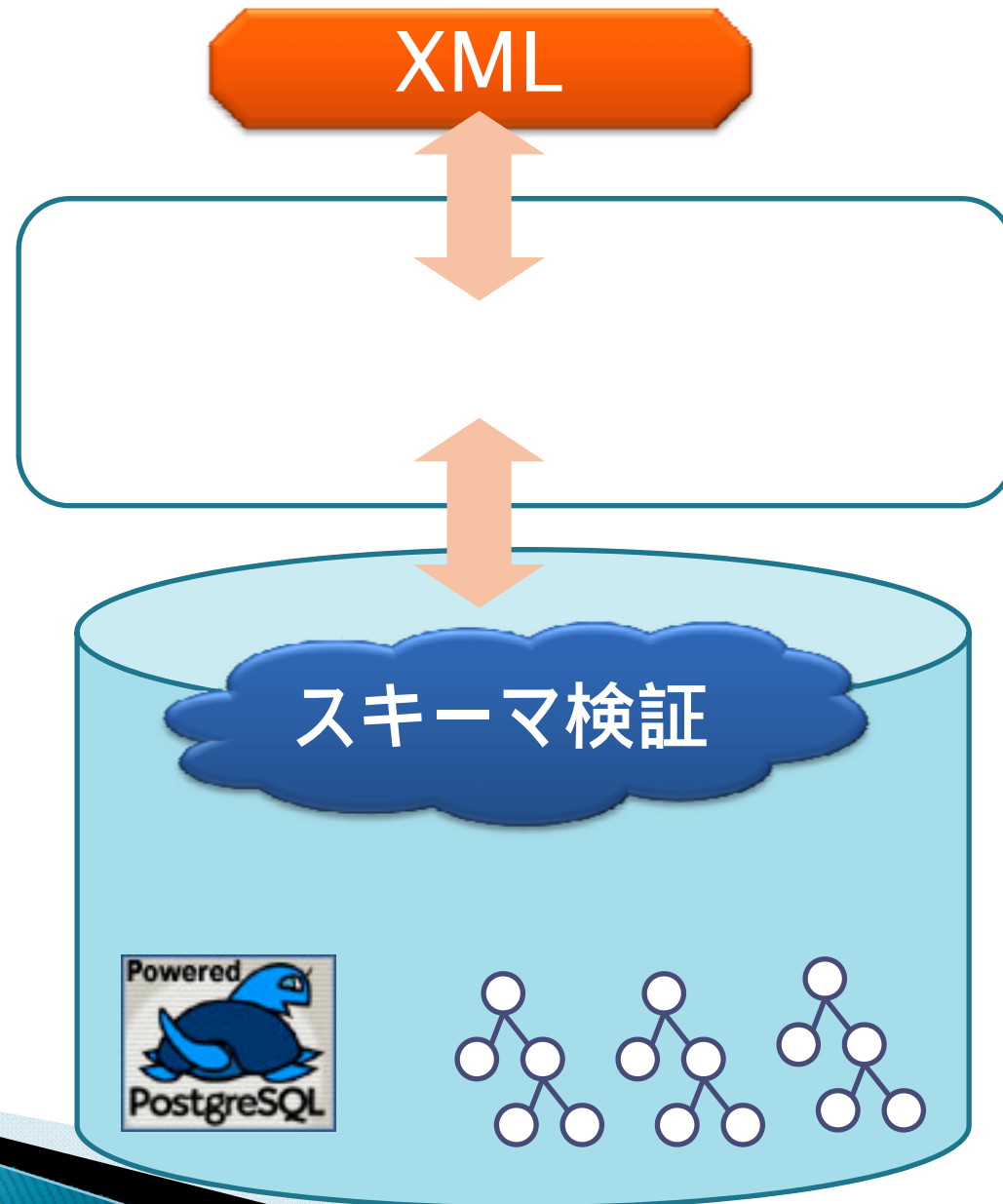
# スキーマ検証の重要性



# スキーマ検証をアプリケーションで行う



# スキーマ検証をデータベースで行う



# 講演のまとめ

## PostgreSQLのサポートするXML機能

- ▶ データ型「XML型」のサポート
- ▶ XMLデータ検索「XPath」のサポート
- ▶ RDB→XML変換「SQL/XML」のサポート

将来拡張への提言

# XMLスキーマ検証 機能の導入を！！

