

AITC Week 2013

クラウドテクノロジー研究部会

気象庁XMLをデータベースに溜めてみた

2013/08/30

クラウド・テクノロジー研究部会

上村準也

(キヤノンソフトウェア株式会社)

1. データベースに溜める目的
2. RDB に溜めてみた
3. XMLDB に溜めてみた
4. RDF リポジトリに溜めてみた

- 実際の配信データを眺めてみたい
 - どんなことが分かるか感触をつかみたい
- まとまったデータが欲しい
 - データがテストやデバッグで必要となりそう
- XMLで可能なことを試したい
 - プロトタイプ作成のための環境

<http://cloud.projectla.jp/jmaxldb/>

- PostgreSQL 9.2 を利用
 - XML型はSQL 2003標準、JDBC 4.0で利用可
- 実際に配信される情報を眺めてみてください
 - 本番システムには組み込まないで！
 - 突然、サービス停止したりします
 - データに「重複」や「抜け」があります

- 気象庁防災情報XMLをまるごと1つのカラムに

```
CREATE TABLE reports (  
  id serial,  
  entry uuid,  
  updated timestamp,  
  report xml  
);
```

- BaseX - <http://docs.basex.org>
 - ネイティブな XMLDB
 - クライアントやAPIも付属
- [XPath/XQuery 検索ページ](#)を作りました
 - XQuery はXMLデータ問い合わせの言語
 - プログラミングレスで複雑なことができます

- Apache Jena - <http://jena.apache.org/>
 - セマンティックWebのためのフレームワーク
 - GUIやAPIの周辺プロジェクトも
- SPARQL 検索ページを作りました
 - SPARQLはRDFグラフ問い合わせの言語
 - パターンで検索することができます

- ニュースや天気予報の元ネタを知るのには新鮮な気分でした
- 情報の種類がとても多い！
- 実際は、ここからがスタート

ご清聴ありがとうございました