



XML Consortium

2009年度XMLコンソーシアムWeek

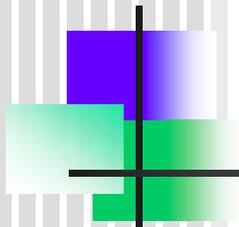
次世代Web活用部会

雨量情報のセマンティック処理

2010-03-18

XMLコンソーシアム 次世代Web活用部会

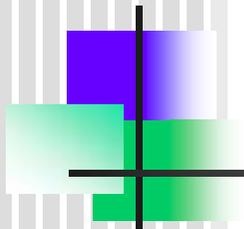
日本ユニシス株式会社	小林 茂
株式会社 日立システム	湯本 正典
東芝ソリューション株式会社	西 一嘉



メンバー一覧(五十音順、敬称略)



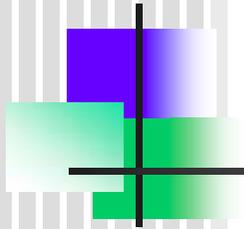
- 荒本 道隆 アドソル日進(株)
- 小林 茂 日本ユニシス(株)
- 妹尾 正身 NTTソフトウェア(株)
- 西 一嘉 東芝ソリューション(株)
- 湯本 正典 (株)日立システムアンドサービス



目次



1. はじめに
2. モデルの導出
3. 実装の紹介
4. まとめ



XML Consortium

はじめに

次世代Web活用部会の紹介

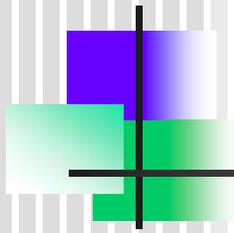


- 次世代Web活用部会での活動
 - Web2.0部会, メタデータ活用部会などを経て。
 - ソーシャルとセマンティックに絡むこと
 - 次世代のWeb記述であるHTML5
 - クラウドも
- セマンティック実証チーム
 - セマンティックWebに関心がある
 - 世の中でなかなかセマンティックWebがブレイクしない
 - セマンティックWebのメリットが容易にわかる事例があると良い
 - 事例を通してセマンティックWebの技術を具体化してわかりやすく紹介したい
 - RDF, RDFS, OWL, SPARQLなどを使う
 - 具体的なモデルを考えて, 何ができるかを示したい
 - 構築のための手順とかを提示できると良い
 - 高度な技術の前に, まずは具体的に

実証実験プロジェクトへの参加



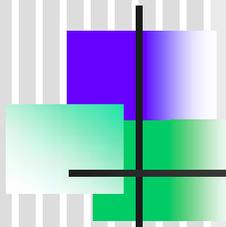
- あまり複雑にすると実装できないし、説明しにくくなる。学ぶ方にも難しくなる
- 皆がよく理解できるものが良い
- 防災情報XMLを用いた実証実験を題材に
- UIは関西部会に任せて、セマンティック技術の紹介に専念できる



雨量情報のサービス



- 部会合同でデモシナリオを検討している際に
- 雨量から、過去の似た事例を探すことに
- 直接雨量の数値を与るのではなく
- 雨に関する言い回し(言葉)から、雨量に置き換えられたら面白いとの意見がでた
- そこでセマンティック技術を使って、言い回しから雨量を返すWebサービスの提供とした



デモシステムの画面



XML Consortium

XML Consortium 気象庁防災XML実証実験Project - Microsoft Internet Explorer

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)

アドレス(A) http://localhost:12010/MainForm.aspx

XML Consortium 気象庁防災XML実証実験Project

気象警報・注意報 台風情報

【発表時刻: 2011-02-23T18:45:00+09:00】 富山県では、大雪、風雪、落雷、高波、なだれ、低温、着氷、電線等への着雪に注意してください。【東部】[発表]大雪注意報。[継続]風雪注意報。[継続]雷注意報。[継続]波浪注意報。[継続]なだれ注意報。[継続]低温注意報。[継続]着氷注意報。[継続]着雪注意報。【西部】[発表]大雪注意報。[継続]雷注意報。[継続]なだれ注意報。[継続]低温注意報。[継続]着雪注意報。【発表時刻: 2008-09-29T16:25:00+09:00】 台風第15号は台北の北にあって北に進んでいます。先島諸島では大しけとなっています。引き続き高波に警戒してください。また台風の上とともに前線の活動が活発となるため、西日本や東海地方では大雨に注意が必要です。

2009年12月						
日	月	火	水	木	金	土
29	30	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2
3	4	5	6	7	8	9

雨量の言い回しをテキストで入力

一日の雨量

雨量の言い回しのテキスト

過去情報の中から検索された写真を表示

地域情報過去検索

日付(From)

日付(To)

雨量(mm) ~

自由入力 土砂降り

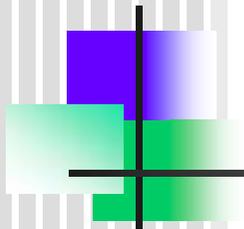
雨量(mm) 100 ~ 100

自由入力 土砂降り

該当写真件数 3

2009/12/01

8月5日の台風の時撮影した写真です



雨量情報サービスの考え方



- できるだけ単純なインターフェースとする
- 地域により同じ言い回しでも異なった値をもつことがある
- 入力 は、言い回しの文字列と地域情報
- 出力 は、雨量
 - 通常の雨量検索とあわせるため降水量ではなく、雨量とする
 - 1日の雨量

雨量情報取得Web API

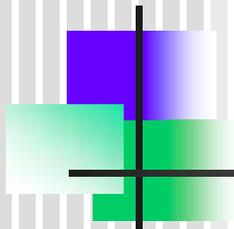


- HTTP GETでリクエスト

http://.../nextweb/raininfo?area=大阪府&q=土砂降り

- XMLで返す

```
<?xml version="1.0" encoding="UTF-8"?>
<response q="土砂降り">
  <area name="大阪府">
    <rain unit="mm">100</rain>
  </area>
</response>
```

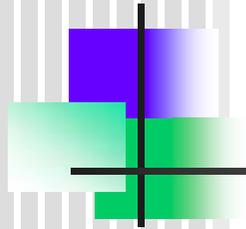


使用するセマンティック技術



規格名	規格の名称と内容
RDF	Resource Description Framework リソースを記述する(メタデータ)ためのフレームワーク
RDFS	RDF Schema 語彙を定義し, 語彙間の簡単な関係を定義するための規格
OWL	Web Ontology Language オントロジーを記述するための言語仕様
SPARQL	Query Language for RDF RDFデータをクエリーするための言語仕様

いずれもW3Cで標準化

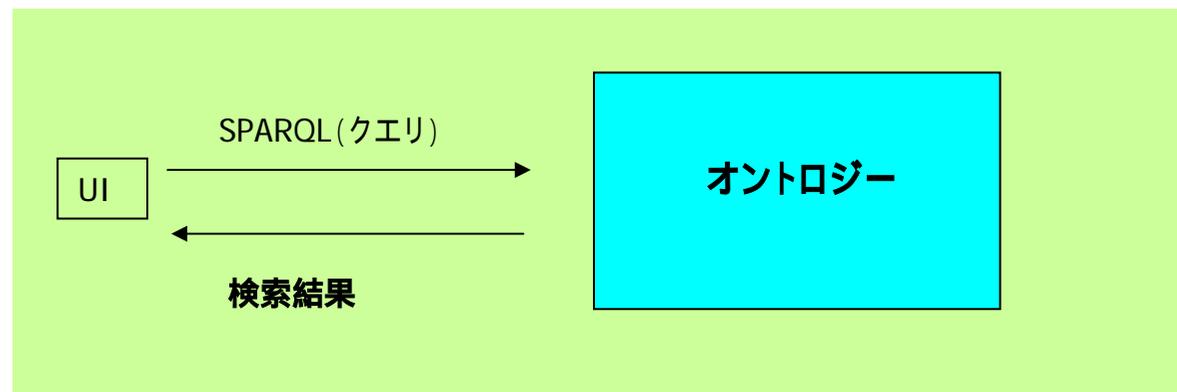


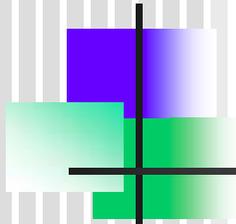
モデルの構築

- セマンティックス実現のための技術要素

	クエリ	リソースを記述する枠組み
記述言語	SPARQL	RDF、RDFS、OWL
従来の構造化	SQL	(RDB ?)

- 設計の概念



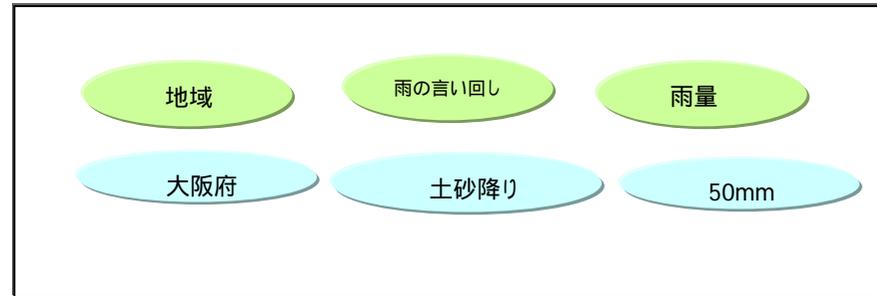
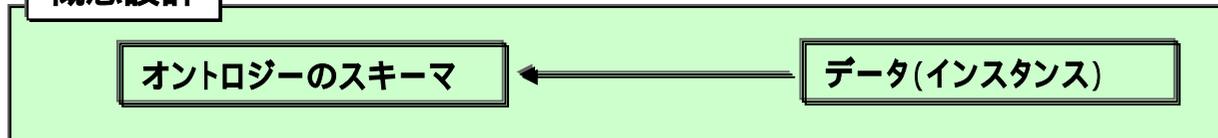


セマンティック検索機能の設計手順

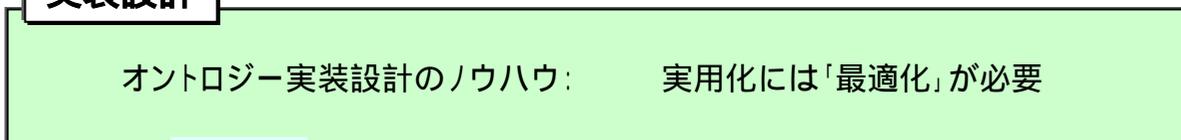


設計の手順

概念設計

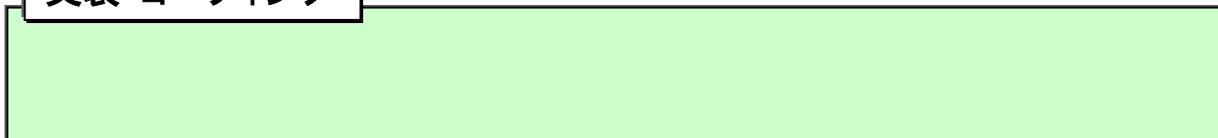


実装設計



cf.(比較)
RDBでも第3正規化してもそのまま実装するわけではない。

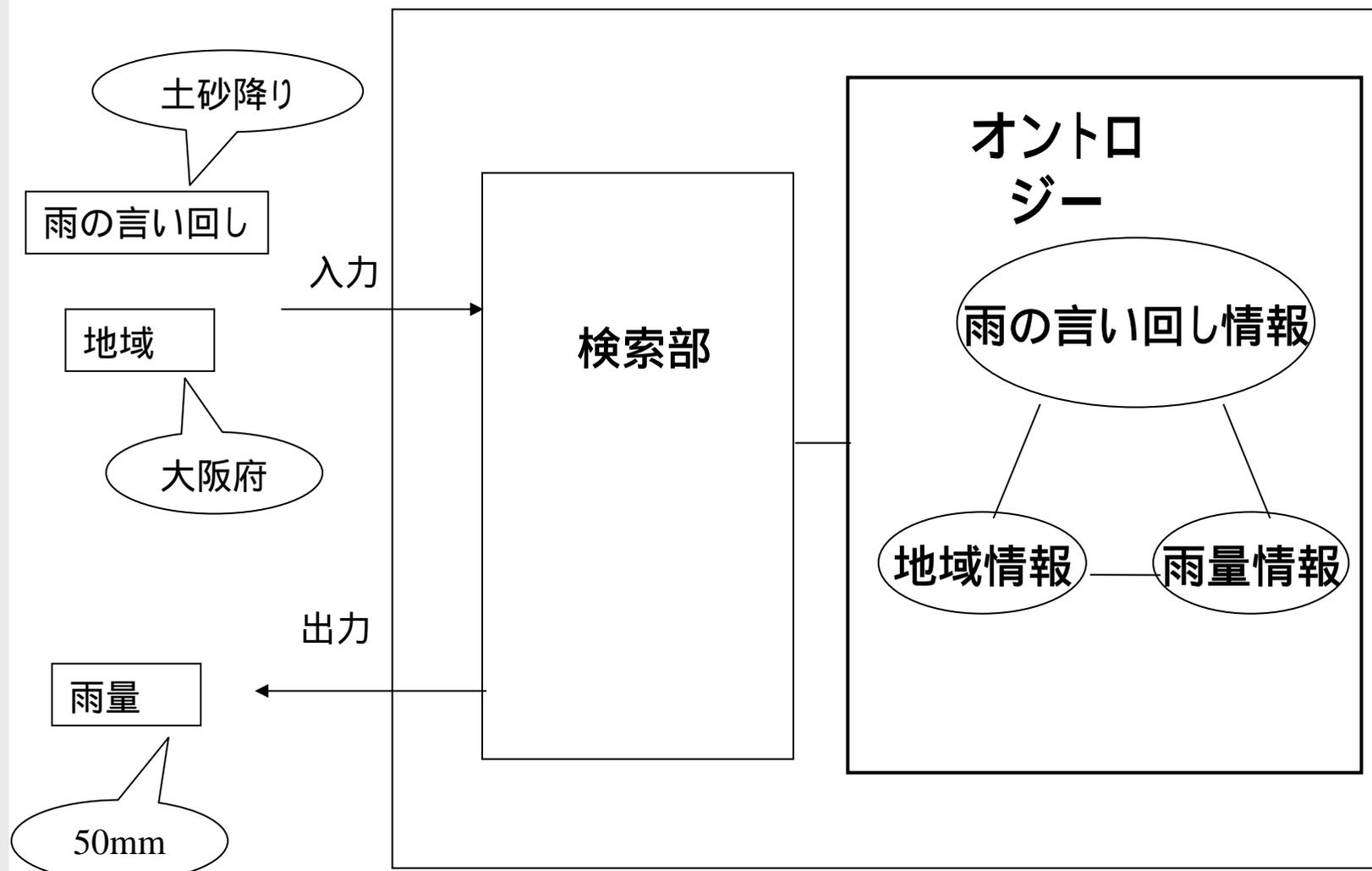
実装・コーディング

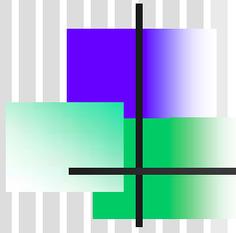


セマンティック検索実装



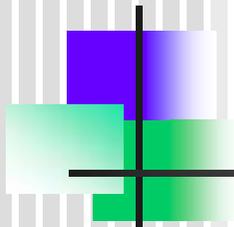
用語の「意味」(セマンティクス)を解釈して検索する。





XML Consortium

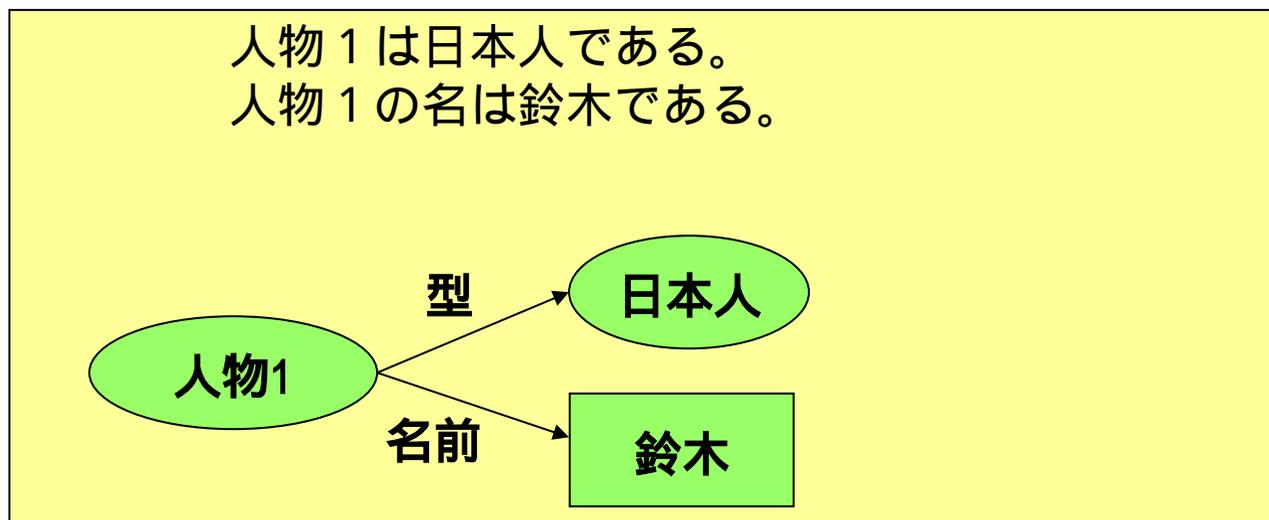
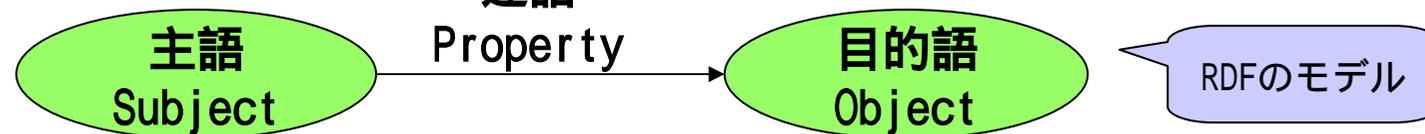
実装の紹介

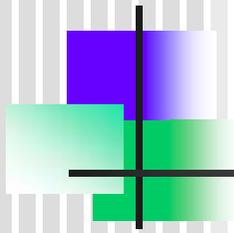


RDFとは



- リソースを記述するための枠組み
- Resource Description Framework
- メタデータを主語, 述語, 目的語の3つのリソースの組(トリプル)によって表現する





SPARQLとは



- データ記述言語
- Query Language for RDF
- トリプルがたくさんあるときに、その一部の条件を指定して、その条件を満たすトリプルを見つける。
- RDBのSQLに相当し、SELECT句、WHERE句から構成される。

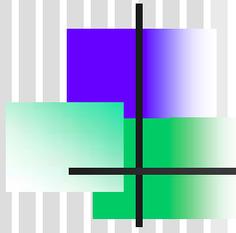
実装



- Jena (Semantic Web Framework for Java) を使用

Jena : HP Labs (<http://www.hpl.hp.com/>)

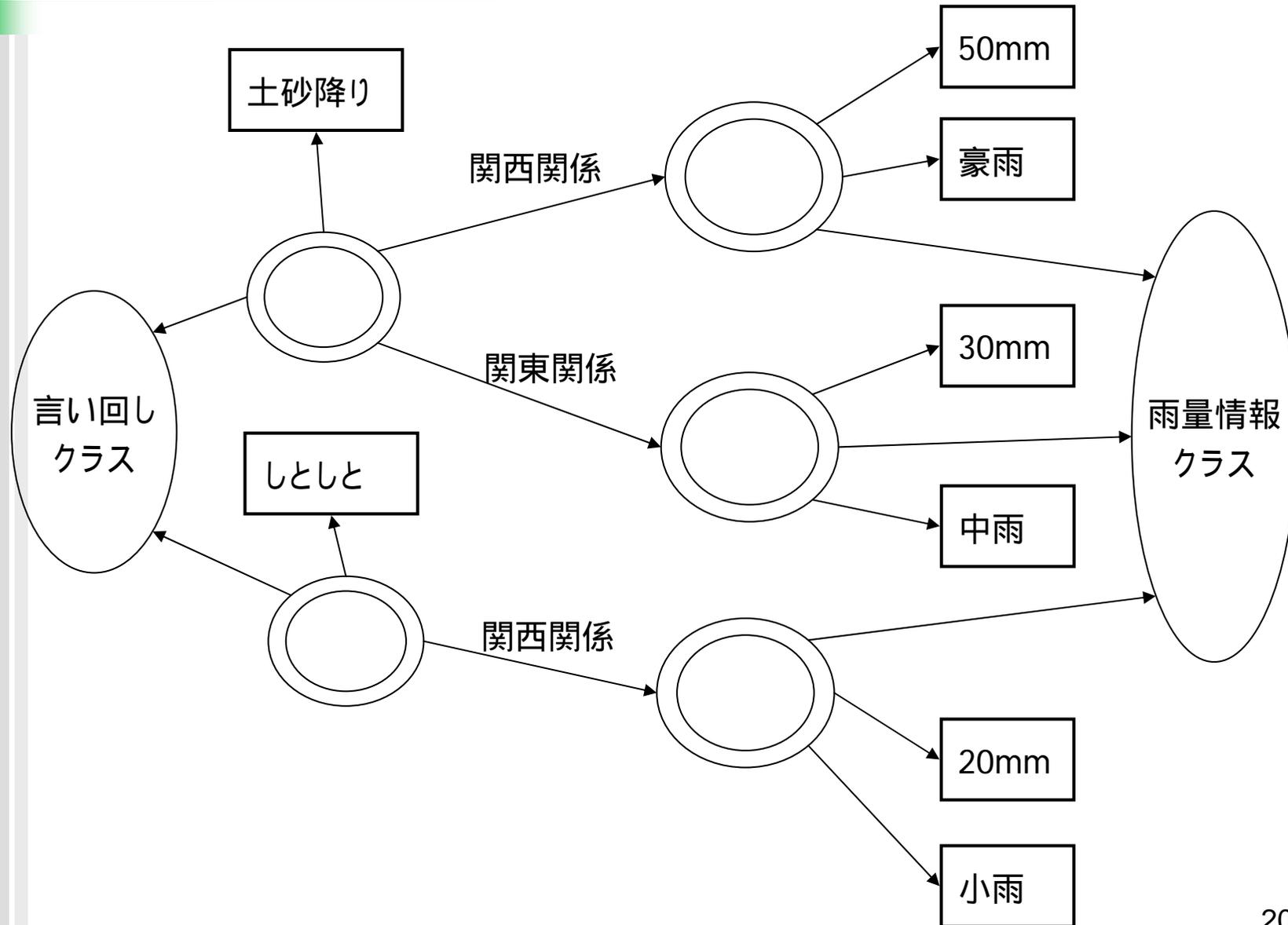
- オントロジーは人手で作成

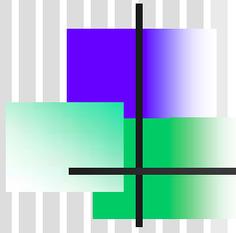


オントロジーイメージ



XML Consortium

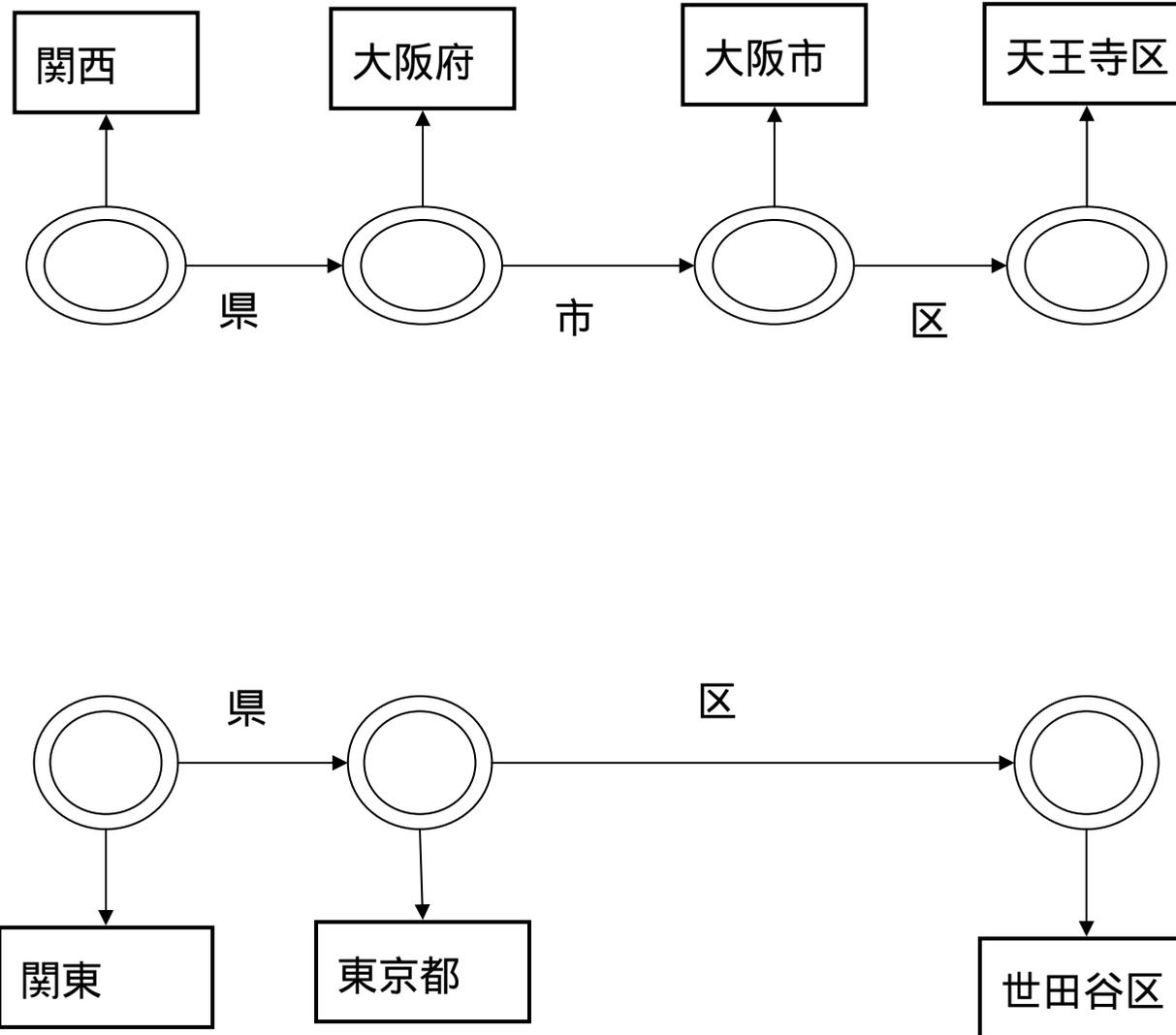


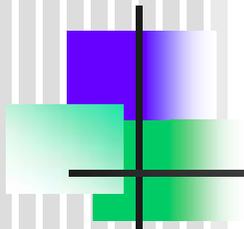


オントロジーイメージ



XML Consortium



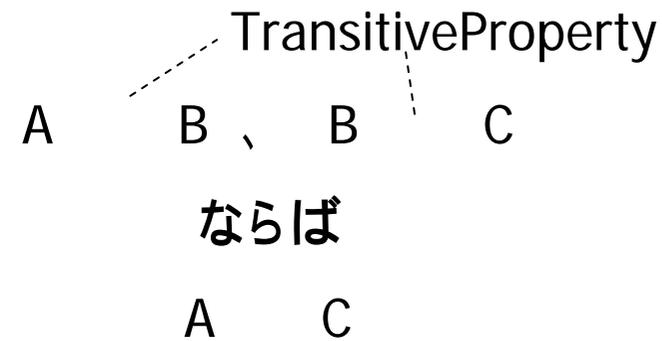


OWLを使用してみた

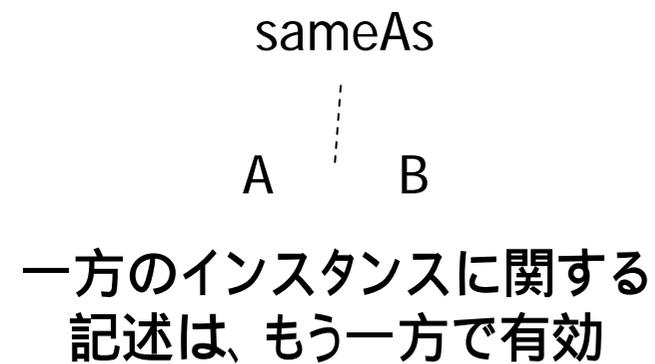


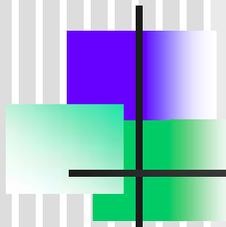
- TransitivePropertyとsameAs

owl:TransitiveProperty



owl:sameAs

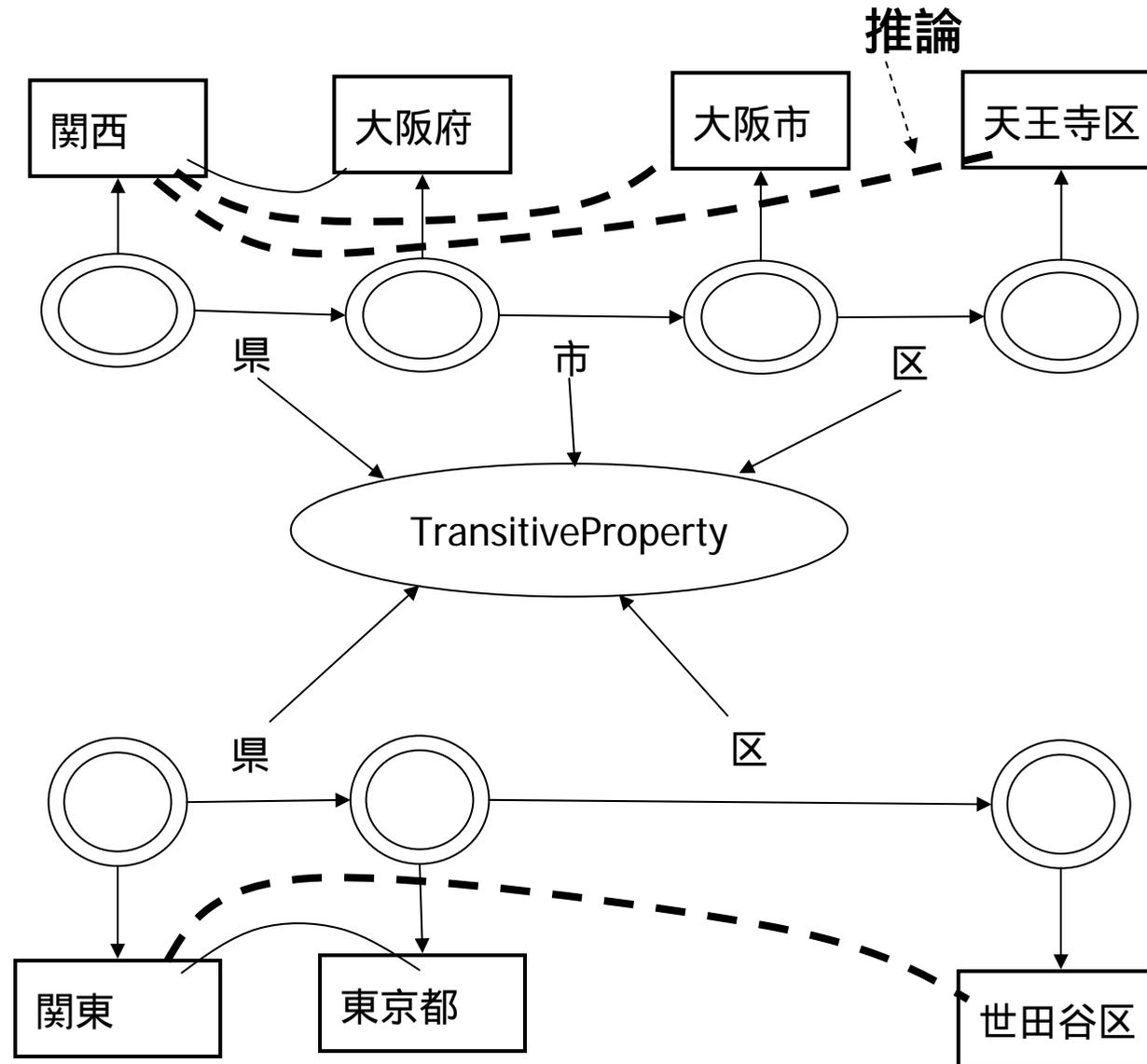


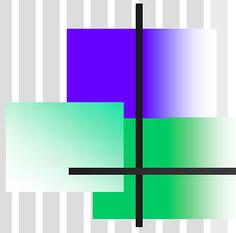


TransitivePropertyを使用してみた



XML Consortium

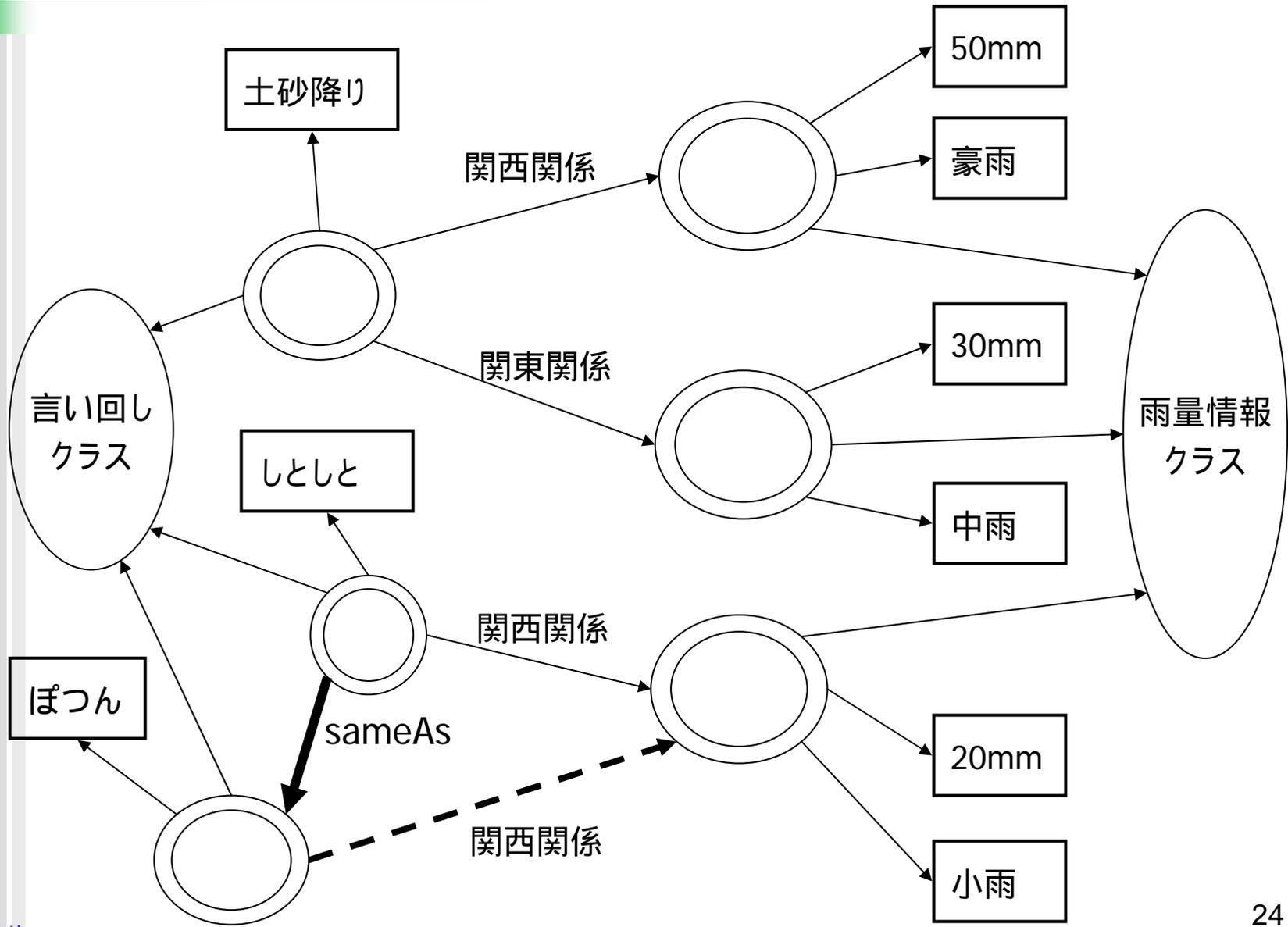


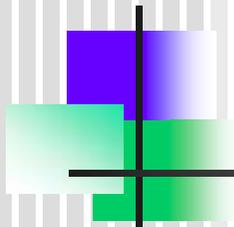


sameAsを使用してみた



XML Consortium





出力結果



```
<response q="土砂降り">
  <area name="大阪府">
    <rain unit="mm">50</rain>
    <rainFormalName>豪雨</rainFormalName>
  </area>
</response>
<response q="土砂降り">
  <area name="大阪市">
    <rain unit="mm">50</rain>
    <rainFormalName>豪雨</rainFormalName>
  </area>
</response>
<response q="土砂降り">
  <area name="天王寺区">
    <rain unit="mm">50</rain>
    <rainFormalName>豪雨</rainFormalName>
  </area>
</response>
```

出力結果

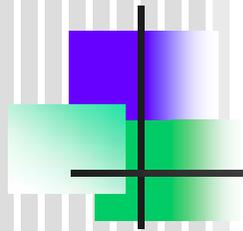
```
<response q="土砂降り">
  <area name="東京都">
    <rain unit="mm">30</rain>
    <rainFormalName>中雨</rainFormalName>
  </area>
</response>

<response q="土砂降り">
  <area name="世田谷区">
    <rain unit="mm">30</rain>
    <rainFormalName>中雨</rainFormalName>
  </area>
</response>
```

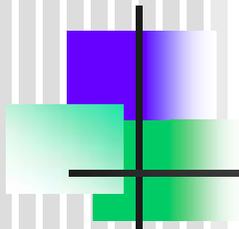
出力結果

```
<response q="しとしと">
  <area name="大阪府">
    <rain unit="mm">20</rain>
    <rainFormalName>少雨</rainFormalName>
  </area>
</response>

<response q="ぽつん">
  <area name="大阪府">
    <rain unit="mm">20</rain>
    <rainFormalName>少雨</rainFormalName>
  </area>
</response>
```



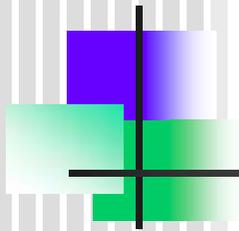
まとめ



まとめ



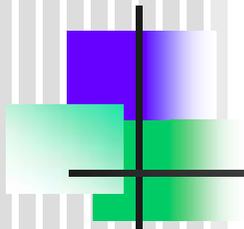
- 今回の成果
 - 次世代Web活用部会で議論してきた内容をもとに、セマンティック技術を使って、言い回しから雨量等の情報を返すWebサービスを実装した。
 - 実装まで行ったことにより、課題も得た。



まとめ



- セマンティック技術を使うと
 - 表現が違って同じ意味となる情報や、表現が同じでも条件によって違う意味となる情報を適切に検索して取得できる。
 - 仕様の追加(変更)への対応
 - プログラムロジックを変更するのではなく、オントロジー、SPARQL式を変えることで対応。



課題



- オントロジー設計の容易化
- 性能検証、要件、改善方法
- パブリックなオントロジーとの連携
 - DBpedia等
- 設計と実装のギャップの検証
- データにメタデータをつける問題