

# RDF構築 SPARQLハンズオン

2013/12/21  
クラウド・テクノロジー活用部会  
リーダー 荒本道隆

## ・Java JDK 6以降

- <http://www.oracle.com/technetwork/java/javase/downloads/index.html>
- SDKがオススメ

## ・Eclipse

- <http://www.eclipse.org/downloads/>
- Standardがオススメ
- Windowsは、こんなbatファイルを作っておくと、便利

```
set PATH=C:¥Java¥jdk1.7.0_45¥bin  
start C:¥eclipse¥eclipse¥eclipse.exe -data C:¥eclipse¥workspace
```

- Workspace を切り替えることで、他プロジェクトと完全に切り離せる
- JDKの切り替えも簡単 & 確実
- ショートカットでもOK

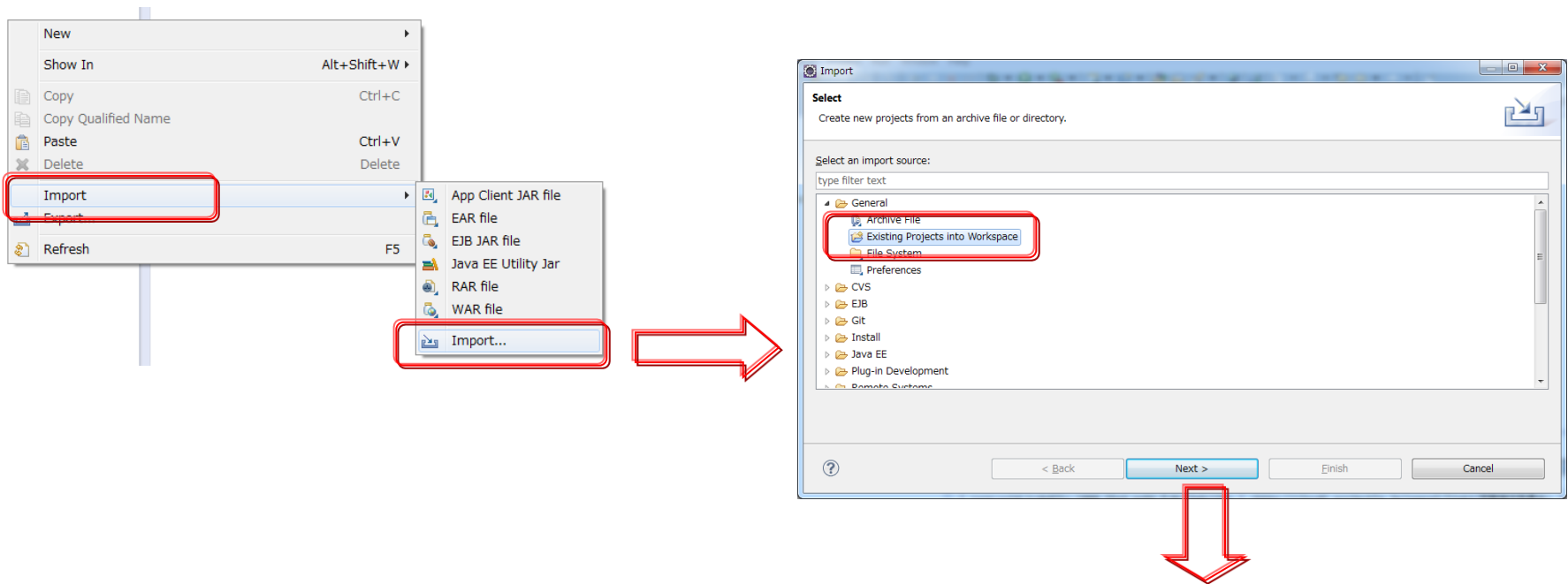
- ・RDFの環境を各自のPC上に構築
  - ・好きな件数で色々使ってみよう
    - ・まずは1件だけで挙動を把握
    - ・データ件数を増やして集計処理を実行
  - ・高負荷のSPARQLを思う存分に実行
- ・PC上でうまく動かない場合は
  - ・回避策を準備済み
- ・今回は検索(SELECT)のみですが
  - ・SPARQLにはINSERT, DELETE もあります
    - ・ <http://www.asahi-net.or.jp/~ax2s-kmtm/internet/rdf/REC-sparql11-update-20130321.html>
- ・SPARQLに慣れたら
  - ・ Dbpediaなどにアクセスしてみる
    - ・ <http://ja.dbpedia.org/>



# 使用するファイルのダウンロード

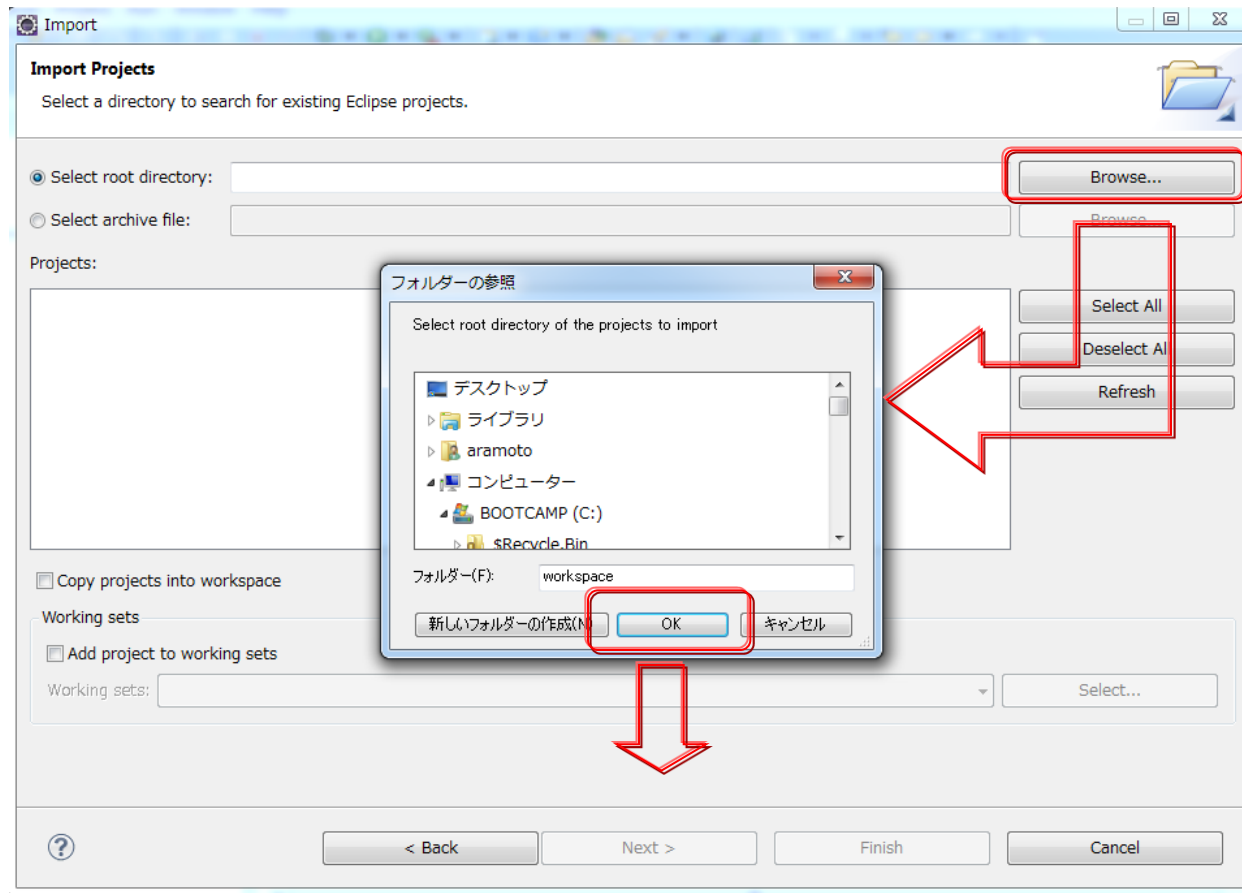
## ・sparql 実行環境

- ・ <http://cloud.aitc.jp/20131221/>
- ・ ダウンロードが終わったら、workspace内に解凍
  - ・ 解凍後は755MByte
- ・ Eclipse でプロジェクトをインポート



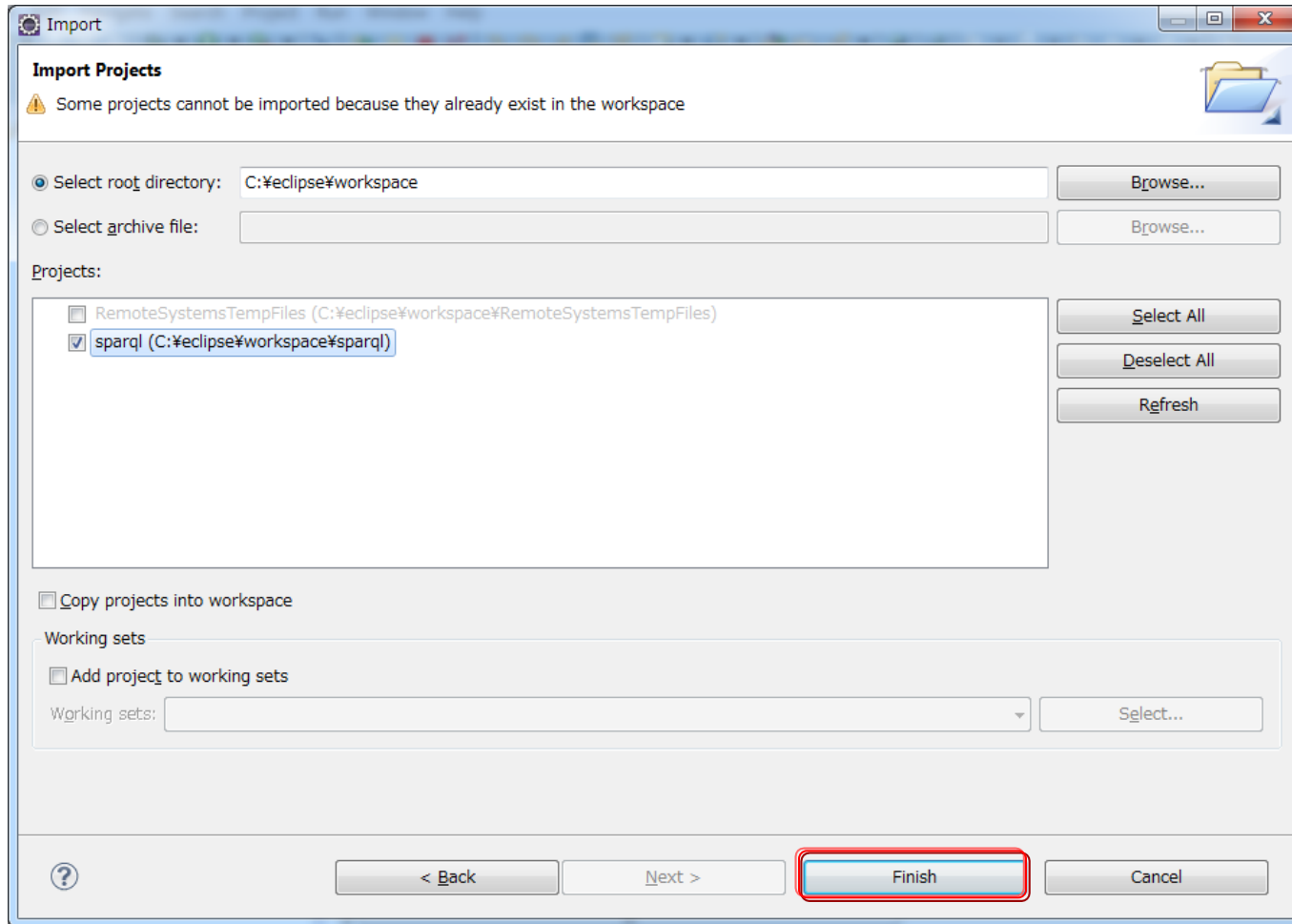
# 使用するファイルのダウンロード

## ・Eclipse でプロジェクトをインポートの続き



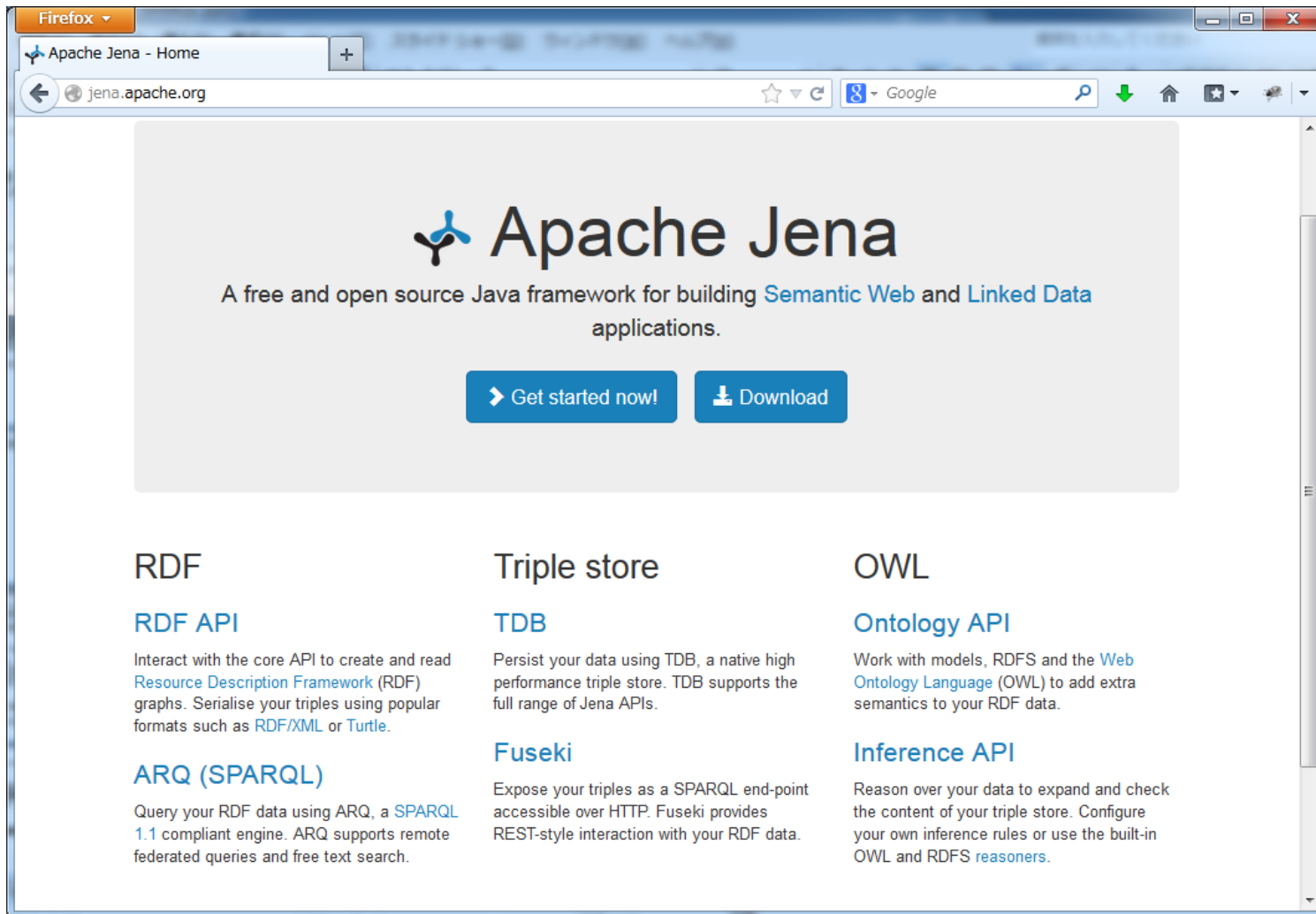
# 使用するファイルのダウンロード

## ・Eclipse でプロジェクトをインポートの続き



## ・Apache Jena

- <http://jena.apache.org/>



## ・SPARQL 1.1に対応

- ・ 2008年1月15日勧告版

- ・ <http://www.asahi-net.or.jp/~ax2s-kmttn/internet/rdf/rdf-sparql-query.html>

## ・ファイルシステム上に、RDFを構築できる

- ・ RDBが不要なので、非常に簡単

- ・ メモリキャッシュが効く

- ・ 2回目以降の問い合わせが高速

- ・ 別JavaVMから更新しないように

- ・ 注意！

- ・ エラーハンドリングが不十分でcloseを忘れると

- ・ **不整合が起こって2度とアクセスできなくなってしまう**

## ・JavaでRDFを登録できる

- ・ RDFの書き方が分からなくても、何とかなる



## ・気象庁XMLを配信するためのfeedメッセージ

- <http://xml.kishou.go.jp/>
- ・ 2013年1月1日～12月2日の途中まで、131,617件(391件/日)
- ・ Atom形式
  - ・ RSS(RDF site summary)から派生したもの
  - ・ IETFにおいてRFC4287として仕様が公開されている
- ・ feedの中に、気象庁XML本文のURLが書いてある
  - ・ feed内の要素「title, id, updated, author, link, content」
- ・ RDFに合わない部分を強引に変換

## ・気象台情報

- <http://www.jma.go.jp/jma/kishou/link/link2.html>
- ・ 気象官署名, 郵便番号, 所在地等, 電話番号
- ・ 所在地から、Geocoding APIで緯度・経度を取得
  - <http://www.geocoding.jp/api>
  - ・ 皆が一斉に実行すると負荷がかかるので、事前に取得済み
- ・ JenaのAPIを使って登録

# 気象庁XMLで配信されているデータ

## XML化対象電文

気象庁情報カタログ(<http://www.jma.go.jp/jma/kishou/177/jmh/catalogue.html>)における資料(情報)名による一覧

表 1.1  
(平成25年5月28日現在)

選番	対象分野	情報分類	資料(情報)名	現在のページ名 (オンライン配信用)	現在のデータ形式	新たに提供されるXML形式		備考		
						提供開始	サイズ(Byte) 平均 最大			
1	気象	観測	特殊気象観測	特殊	半角カナ		3	5		
2	気象	観測	特殊気象観測	特殊, プライベイト	半角カナ		3	5		
3	気象	観測	特殊気象観測	特殊	半角カナ		3	5		
4	気象	観測	特殊気象観測	特殊	半角カナ		3	5		
5	気象	観測	海外観測データ	海外(ファイルで提供中)	現XML		10	14		
6	気象	解析・予報	全国気象予報(総合情報, 上送特報)	全国気象予報11	かな漢字		5	6		
7	気象	解析・予報	全国気象予報(総合情報, 気象情報, 気象予報)と気象予報(総合情報)	全国気象予報12	かな漢字		5	6		
8	気象	解析・予報	全国気象予報(位置詳細)	全国気象予報21	かな漢字		5	6		
9	気象	解析・予報	全国気象予報(電文新形式)	WFXML (9=00-05)	コード		20	100		
10	気象	解析・予報	全国気象予報(電文新形式)	WFXML (9=00-05)	コード		25	120		
11	気象	解析・予報	全国海上警報(電文)	海上警報1	半角カナ		25	110		
12	気象	解析・予報	全国海上警報(電文)	海上警報2	半角カナ		25	110		
13	気象	解析・予報	地方海上警報	地方海上警報1	かな漢字		15	100		
14	気象	予報	地方海上予報	地方海上予報1	かな漢字		50	200		
15	気象	予報	気象特別警報・警報・注意報	特別警報1, 特別警報2, 特別警報3, 特別警報4, 特別警報5, 特別警報6, 特別警報7, 特別警報8, 特別警報9, 特別警報10, 特別警報11, 特別警報12, 特別警報13, 特別警報14, 特別警報15, 特別警報16, 特別警報17, 特別警報18, 特別警報19, 特別警報20, 特別警報21, 特別警報22, 特別警報23, 特別警報24, 特別警報25, 特別警報26, 特別警報27, 特別警報28, 特別警報29, 特別警報30, 特別警報31, 特別警報32, 特別警報33, 特別警報34, 特別警報35, 特別警報36, 特別警報37, 特別警報38, 特別警報39, 特別警報40, 特別警報41, 特別警報42, 特別警報43, 特別警報44, 特別警報45, 特別警報46, 特別警報47, 特別警報48, 特別警報49, 特別警報50, 特別警報51, 特別警報52, 特別警報53, 特別警報54, 特別警報55, 特別警報56, 特別警報57, 特別警報58, 特別警報59, 特別警報60, 特別警報61, 特別警報62, 特別警報63, 特別警報64, 特別警報65, 特別警報66, 特別警報67, 特別警報68, 特別警報69, 特別警報70, 特別警報71, 特別警報72, 特別警報73, 特別警報74, 特別警報75, 特別警報76, 特別警報77, 特別警報78, 特別警報79, 特別警報80, 特別警報81, 特別警報82, 特別警報83, 特別警報84, 特別警報85, 特別警報86, 特別警報87, 特別警報88, 特別警報89, 特別警報90, 特別警報91, 特別警報92, 特別警報93, 特別警報94, 特別警報95, 特別警報96, 特別警報97, 特別警報98, 特別警報99, 特別警報100	かな漢字(コード), 現XML		平成25年5月27日 提供開始 ※ただし、特別警報13は 平成25年7月18日提供 開始	200	1,300	
16	気象	予報	指定河川洪水予報	指定河川洪水予報 (9=00-00), 指定河川洪水予報 (9=01-00)	かな漢字		15	300		
17	気象	予報	土砂災害警報情報	土砂災害警報情報	かな漢字		15	30		
18	気象	予報	記録的短時間大雨情報	記録的短時間大雨情報	かな漢字		5	15		
19	気象	予報	電磁注意情報	電磁注意情報	かな漢字		5	10		
20	気象	予報	全国気象情報	全国気象情報	かな漢字		3	6		
21	気象	予報	地方気象情報	地方気象情報	かな漢字		3	6		
22	気象	予報	前線気象情報	前線気象情報	かな漢字		3	6		
23	気象	予報	天気解説	天気解説	かな漢字		3	6		
24	気象	予報	前線天気予報	前線天気予報	かな漢字, 現XML		40	100		
25	気象	予報	地域別天気予報	地域別天気予報	かな漢字		3	6		
26	気象	予報	全国週間天気予報	全国週間天気予報	かな漢字		3	6		
27	気象	予報	地方週間天気予報	地方週間天気予報	かな漢字		20	60		
28	気象	予報	前線週間天気予報	前線週間天気予報	かな漢字, かな漢字, 現XML		3	6		
29	気象	予報	スモッグ気象情報	スモッグ気象情報	かな漢字		3	6		
30	気象	気候	全国気象情報(気候的)に関する情報	全国気象情報(気候的)に関する情報	かな漢字		15	20		
31	気象	気候	地方気象情報(気候的)に関する情報	地方気象情報(気候的)に関する情報	かな漢字		15	20		
32	気象	気候	前線気象情報(気候的)に関する情報	前線気象情報(気候的)に関する情報	かな漢字		15	40		
33	地震・津波	観測	震度速報	震度速報	コード		4	50		
34	地震・津波	観測	震源に関する情報	震源に関する情報	かな漢字(コード), かな漢字(コード)		3	4		
35	地震・津波	観測	震源・震度に関する情報	震源・震度に関する情報	かな漢字(コード), かな漢字(コード), かな漢字(コード), かな漢字(コード), かな漢字(コード), かな漢字(コード)		10	1,000	XMLでは1つの電文に統合して配信する。	
36	地震・津波	観測	各地の震度に関する情報	各地の震度に関する情報	コード, コード					
37	地震・津波	予報	緊急地震速報(警報・予報)	緊急地震速報(警報・予報)	コード, かな漢字		15	100		
38	地震・津波	予報	緊急地震速報(警報)	緊急地震速報(警報)	コード, かな漢字		15	100		
39	地震・津波	観測	津波情報	津波情報	かな漢字(コード), コード, かな漢字(コード), コード		25	200		
40	地震・津波	予報	津波警報・注意報	津波警報・注意報	かな漢字(コード), コード, かな漢字(コード), コード		15	50		
41	地震・津波	予報	東海地震に関する情報(東海地震予知)/注意情報	東海地震に関する情報(東海地震予知)/注意情報	かな漢字(コード)		3	10		
42	地震・津波	観測	東海地震に関する情報(東海地震予知)/注意情報	東海地震に関する情報(東海地震予知)/注意情報	かな漢字(コード)		3	10		
43	地震・津波	その他	地震・津波・火山関連のお知らせ	地震・津波・火山関連のお知らせ	かな漢字		4	15		
44	火山	観測	火山観測	火山観測	かな漢字		4	10		
45	火山	観測	火山の状況に関する観測情報	火山の状況に関する観測情報	かな漢字(コード)		20	80		
46	火山	予報	噴火警報・予報	噴火警報・予報	かな漢字(コード)		8	300		
47	火山	予報	火山地方海上警報	火山地方海上警報	かな漢字		5	10		
48	海洋	解析・予報	前線海水予報	前線海水予報	かな漢字		2	3		
49	海洋	観測	全国潮位情報	全国潮位情報	かな漢字		4	6		
50	海洋	観測	地方潮位情報	地方潮位情報	かな漢字		4	6		
51	海洋	観測	前線潮位情報	前線潮位情報	かな漢字		4	6		
52	気候	予報	全国空質予報	全国空質予報	かな漢字		20	50		
53	気候	予報	地方空質予報	地方空質予報	かな漢字		15	25		
54	気候	予報	緊急大気汚染警報	緊急大気汚染警報	かな漢字		10	20		
55	地震・津波	その他	緊急地震速報の配信システム電文	緊急地震速報の配信システム電文	コード, かな漢字		1	3		
56	気象	予報	全国スモッグ気象情報	全国スモッグ気象情報	かな漢字		3	6		
57	気象	予報	地方スモッグ気象情報	地方スモッグ気象情報	かな漢字		2	5		
58	気象	予報	前線スモッグ気象情報	前線スモッグ気象情報	かな漢字		2	5		
59	気象	予報	地方スモッグ気象情報	地方スモッグ気象情報	かな漢字		2	5		
60	地震・津波	観測	緊急地震速報に関する情報	緊急地震速報に関する情報	かな漢字(コード), コード		25	200		
61	地震・津波	予報	緊急地震速報に関する情報	緊急地震速報に関する情報	かな漢字(コード), コード		25	200		
62	地震・津波	予報	緊急地震速報に関する情報	緊急地震速報に関する情報	かな漢字		20	100		

注1)「現在のデータ形式」について、それぞれの形式の説明は用語集を参照のこと。ほか、「現XML」は現フォーマットによるXML形式の電文を示す。「(コード)」はコード形式の情報も含まれていることを示す。  
 注2)「現在のデータ形式」は、基本的に全て「新たに提供されるXML形式」に移行する。提供開始後、一定期間の並行運用期間を経て、「現在のデータ形式」による配信は停止する計画である。

[http://xml.kishou.go.jp/jmaxml\\_20130628\\_format\\_v1\\_1\\_hyo1\\_1.pdf](http://xml.kishou.go.jp/jmaxml_20130628_format_v1_1_hyo1_1.pdf)

# 気象庁XMLで配信されているデータ

16	気象	予報	指定河川洪水予報	カセウクス'イii (ii=80-99), フケンクス'イii (ii=01-20)	かな漢字
17	気象	予報	土砂災害警戒情報	ドシヤケイカイ VXWWW40	かな漢字 現XML
18	気象	予報	記録的短時間大雨情報	キロウアメ1	かな漢字
19	気象	予報	竜巻注意情報	タツマキシヨウホウ1	かな漢字
20	気象	予報	全般気象情報	ゼンコクシヨウホウ1	かな漢字
21	気象	予報	地方気象情報	チホウシヨウホウ1	かな漢字
22	気象	予報	府県気象情報	フケンシヨウホウ1	かな漢字
23	気象	予報	天気概況	フケンガイキヨウ1	かな漢字
24	気象	予報	府県天気予報, 地域時系列予報	ヨホウ1, VPF40, KYYY50	かな漢字, 現XML, バイナリー
25	気象	予報	全般週間天気予報	ゼンバンシユウカン1	かな漢字
26	気象	予報	地方週間天気予報	チホウシユウカン1	かな漢字
27	気象	予報	府県週間天気予報	フケンシユウカン1, シュウカンシライト'1 VPEW40, VPEW41	かな漢字, かな漢字, 現XML, 現XML
28	気象	予報	スモッグ気象情報	オセンジヨウホウ1	かな漢字
29	気象・気候	予報	全般気象情報(社会的に影響の大きい天候に関する情報)	ゼンバンテンコウ1	かな漢字
30	気象・気候	予報	地方気象情報(社会的に影響の大きい天候に関する情報)	チホウテンコウ1	かな漢字
31	気象・気候	予報	府県気象情報(社会的に影響の大きい天候に関する情報)	フケンテンコウ1	かな漢字
32	地震・津波	観測	震度速報	シント'クホウ3	コード
33	地震・津波	観測・解析	震源に関する情報	シンケン'ソクホウ1, シンケン'ソクホウ11	かな漢字(コード), かな漢字(コード)
34	地震・津波	観測・解析	震源・震度に関する情報	ゼン'コクサイスE1, ゼン'コクサイスE3, ゼン'コクサイスE11, ゼン'コクサイスE13, ゼン'コクサイスE15, ゼン'コクサイスE3	かな漢字(コード), コード, かな漢字(コード), コード, かな漢字(コード), コード
35	地震・津波	観測・解析	各地の震度に関する情報	チキチテンシント'3, チキチテンシント'13	コード, コード
36	地震・津波	予報	緊急地震速報(警報・予報)	ナウキヤスト3, ナウキヤスト4	コード, かな漢字
37	地震・津波	予報	緊急地震速報(警報)	キンキユウジシン3, キンキユウジシン4	コード, かな漢字
38	地震・津波	観測	津波情報	ツナミシヨウホウ1, ツナミシヨウホウ3, ツナミシヨウホウ11, ツナミシヨウホウ13	かな漢字(コード), コード, かな漢字(コード), コード
39	地震・津波	予報	津波警報・注意報	ツナミヨホウ1, ツナミヨホウ3, ツナミヨホウ6, ツナミヨホウ11, ツナミヨホウ13, ツナミヨホウ16	かな漢字(コード), コード, かな漢字(コード), かな漢字(コード), コード, かな漢字(コード)
40	地震・津波	観測・解析・予報・情報	東海地震に関連する情報(東海地震予知/注意情報)	トウカイヨチ1	かな漢字(コード)
41	地震・津波	観測・解説	東海地震に関連する情報(東海地震に関連する調査情報)	トウカイカンソク1	かな漢字(コード)
42	地震・津波・火山	その他	地震・津波・火山関連のお知らせ	サイスマオシラセ1	かな漢字
43	火山	観測	火山観測報	カザンケンシヨウ1	かな漢字
44	火山	観測・解析・解説	火山の状況に関する解説情報	カザンカイセツ1	かな漢字(コード)
45	火山	予報	噴火警報・予報	カザンシヨウホウ1	かな漢字(コード)
46	火山	予報	火山地方海上警報	チホウウミカザン'ii (ii=01-11)	かな漢字
47	海洋	解析・予報	府県海水予報	カイヒヨウ1	かな漢字
48	海洋	観測・解析	全般潮位情報	ゼンバンチヨウイ1	かな漢字
49	海洋	観測・解析	地方潮位情報	チホウチヨウイ1	かな漢字
50	海洋	観測・解析	府県潮位情報	フケンチヨウイ1	かな漢字
51	気候	予報	全般季節予報	キセツヨホウ1	かな漢字
52	気候	予報	地方季節予報	チホウキセツヨホウ1	かな漢字
53	気候	予報	異常天候早期警戒情報	チホウソウケイ1	かな漢字

[http://xml.kishou.go.jp/jmaxml\\_20130628\\_format\\_v1\\_1\\_hyo1\\_1.pdf](http://xml.kishou.go.jp/jmaxml_20130628_format_v1_1_hyo1_1.pdf)

# 元feedメッセージ

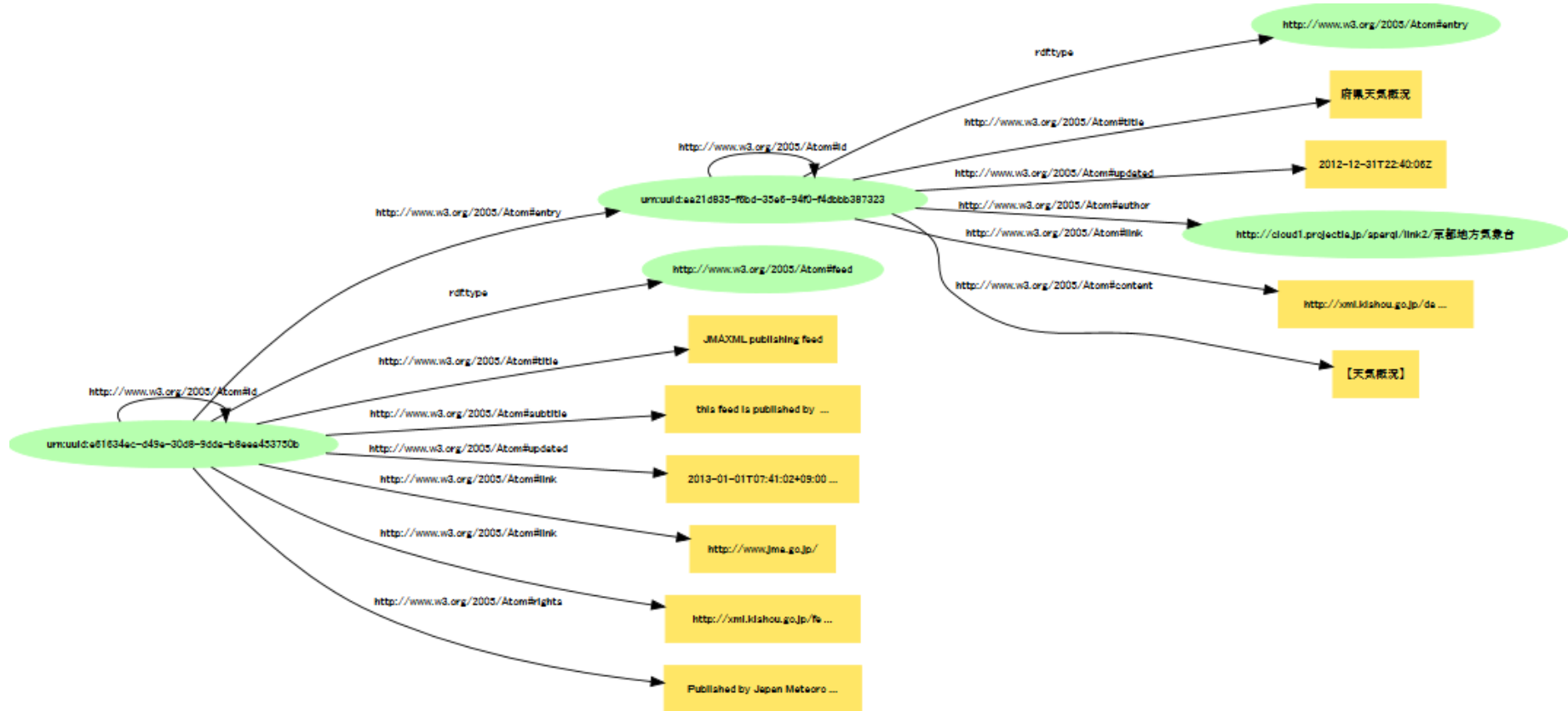
```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom" xml:lang="ja">
  <title>JMAXML publishing feed</title>
  <subtitle>this feed is published by JMA</subtitle>
  <updated>2013-01-01T07:41:02+09:00</updated>
  <id>urn:uuid:e61634ec-d49e-30d8-9dda-b8eea453750b</id>
  <link href="http://www.jma.go.jp/" rel="related" />
  <link href="http://xml.kishou.go.jp/feed/regular.xml" rel="self" />
  <rights>Published by Japan Meteorological Agency</rights>

  <entry>
    <title>府県天気概況</title>
    <id>urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323</id>
    <updated>2012-12-31T22:40:06Z</updated>
    <author>
      <name>京都地方気象台</name>
    </author>
    <link
      href="http://xml.kishou.go.jp/data/aa21d835-f6bd-35e6-94f0-f4dbbb387323.xml"
      type="application/xml" />
    <content type="text">【天気概況】</content>
  </entry>
</feed>
```

# 変換後feedメッセージ

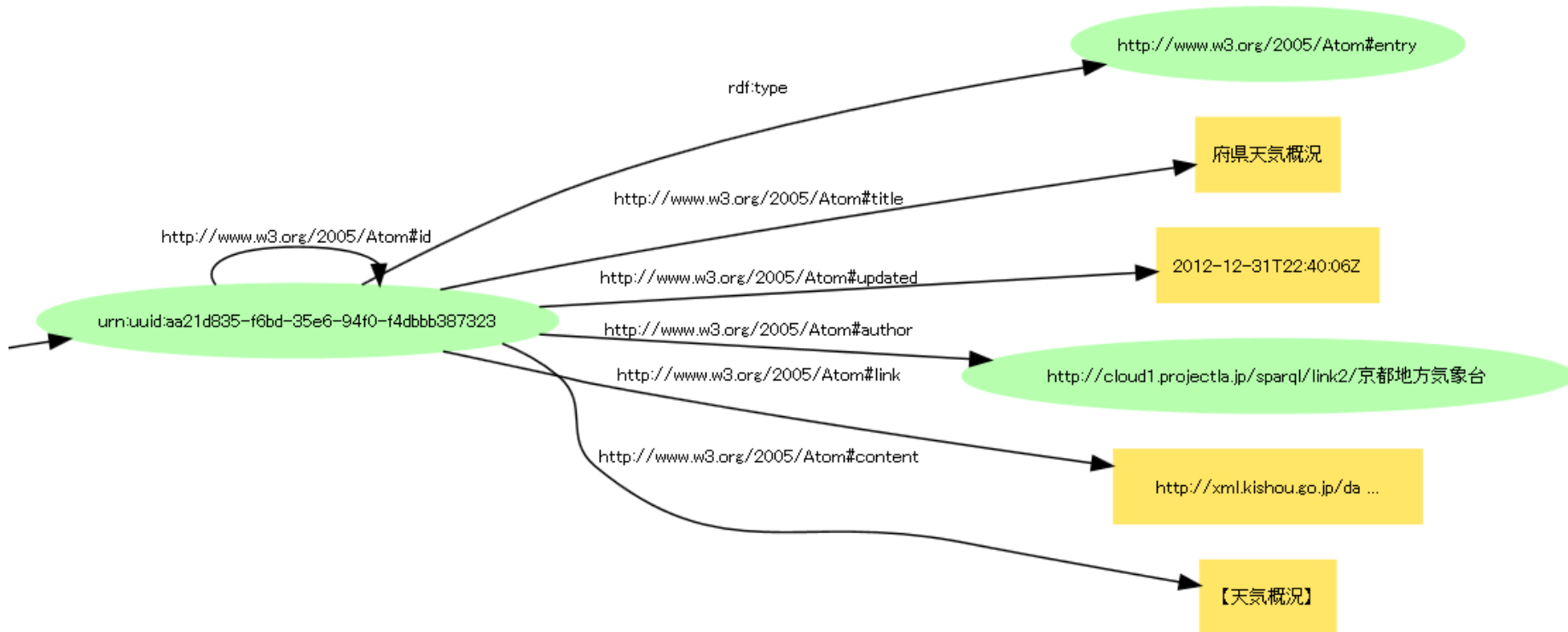
```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <entry rdf:about="urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323"
    xmlns="http://www.w3.org/2005/Atom#">
    <title>府県天気概況</title>
    <id>urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323</id>
    <updated>2012-12-31T22:40:06Z</updated>
    <author rdf:resource="http://cloud1.projectla.jp/sparql/link2/京都地方気象台"></author>
    <link>http://xml.kishou.go.jp/data/aa21d835-f6bd-35e6-94f0-f4dbbb387323.xml
    </link>
    <content>【天気概況】</content>
  </entry>
  <feed xmlns="http://www.w3.org/2005/Atom#" rdf:about="urn:uuid:e61634ec-d49e-30d8-9dda-
b8eea453750b"
    xml:lang="ja">
    <title>JMAXML publishing feed</title>
    <subtitle>this feed is published by JMA</subtitle>
    <updated>2013-01-01T07:41:02+09:00</updated>
    <id>urn:uuid:e61634ec-d49e-30d8-9dda-b8eea453750b</id>
    <link>http://www.jma.go.jp/</link>
    <link>http://xml.kishou.go.jp/feed/regular.xml</link>
    <rights>Published by Japan Meteorological Agency</rights>
    <entry rdf:resource="urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323" />
  </feed>
</rdf:RDF>
```

# このRDFを可視化すると



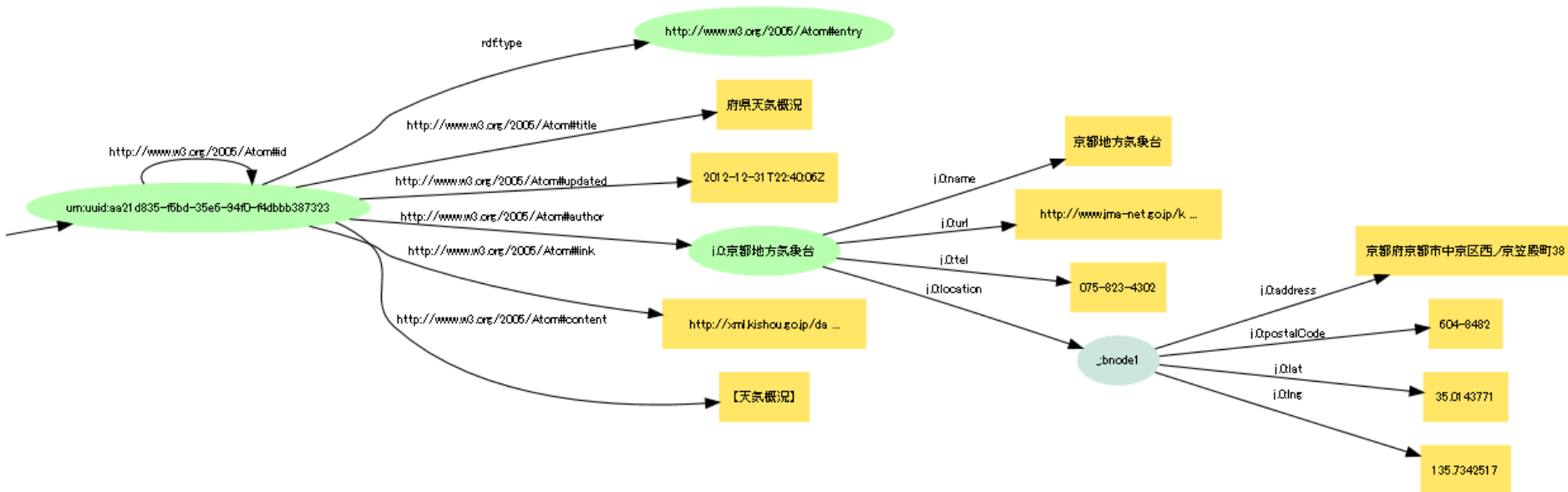
<http://www.kanzaki.com/works/2005/rap/graph-check>

# 使いそうな部分のみを拡大表示



<http://www.kanzaki.com/works/2005/rap/graph-check>

# 京都地方気象台の情報も付加



<http://www.kanzaki.com/works/2005/rap/graph-check>

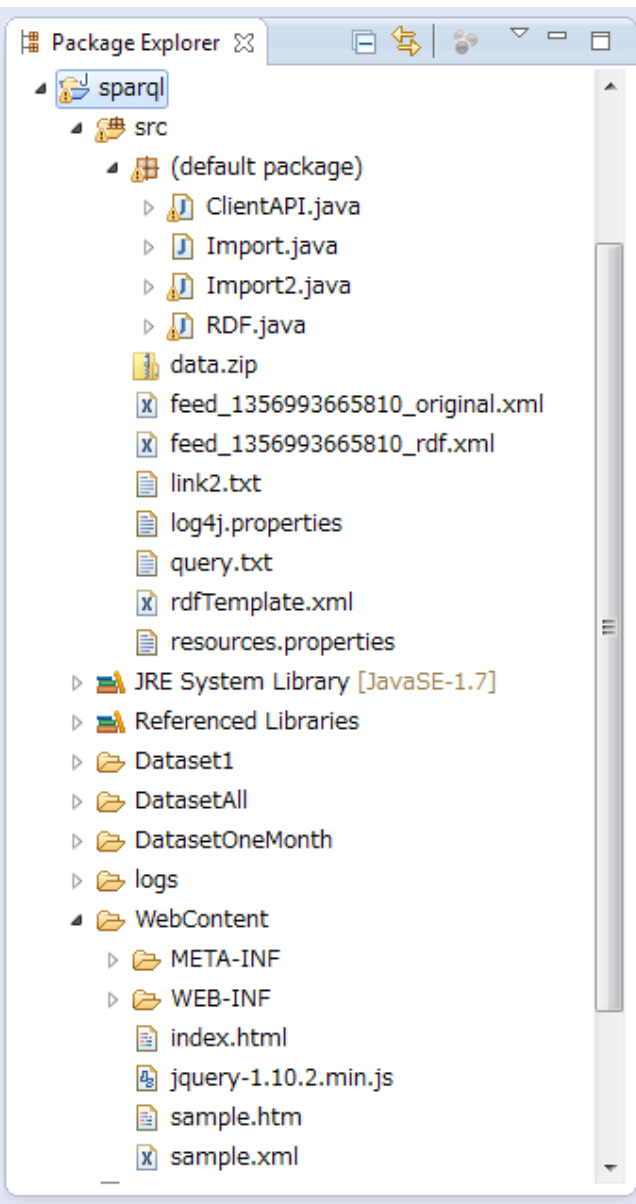


# この時のRDF

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://www.w3.org/2005/Atom#"
  xmlns:j.0="http://cloud1.projectla.jp/sparql/link2/">
<feed rdf:about="urn:uuid:e61634ec-d49e-30d8-9dda-b8eea453750b">
  <entry>
    <entry rdf:about="urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323">
      <title>府県天気概況</title>
      <id>urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323</id>
      <updated>2012-12-31T22:40:06Z</updated>
      <author>
        <rdf:Description rdf:about="http://cloud1.projectla.jp/sparql/link2/京都地方気象台">
          <j.0:name>京都地方気象台</j.0:name>
          <j.0:url>http://www.jma-net.go.jp/kyoto/index.html</j.0:url>
          <j.0:tel>075-823-4302</j.0:tel>
          <j.0:location rdf:parseType="Resource">
            <j.0:address>京都府京都市中京区西ノ京笠殿町38</j.0:address>
            <j.0:postalCode>604-8482</j.0:postalCode>
            <j.0:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#double">35.0143771</j.0:lat>
            <j.0:lng rdf:datatype="http://www.w3.org/2001/XMLSchema#double">135.7342517</j.0:lng>
          </j.0:location>
        </rdf:Description>
      </author>
      <link>http://xml.kishou.go.jp/data/aa21d835-f6bd-35e6-94f0-f4dbbb387323.xml</link>
      <content>【天気概況】</content>
    </entry>
  </entry>
  <title xml:lang="ja">JMAXML publishing feed</title>
  <id xml:lang="ja">urn:uuid:e61634ec-d49e-30d8-9dda-b8eea453750b</id>
  <updated xml:lang="ja">2013-01-01T07:41:02+09:00</updated>
  <link xml:lang="ja">http://www.jma.go.jp/</link>
  <link xml:lang="ja">http://xml.kishou.go.jp/feed/regular.xml</link>
  <subtitle xml:lang="ja">this feed is published by JMA</subtitle>
  <rights xml:lang="ja">Published by Japan Meteorological Agency</rights>
</feed>
</rdf:RDF>
```

- ・AtomはそのままではRDFには入りません
  - ・今回は、強引に変換して登録しています
  - ・他のfeedで同じ変換をしてもダメの可能性大
- ・SPARQLにコメントアウト機能はありません
  - ・今回は、お手製で機能追加しました
  - ・他の環境で実行する時は注意してください
- ・10万件をインポートするにはとても時間がかかります
  - ・検索も、書き方次第でとても時間がかかります
    - ・Webでやると、キャッシュが効くので分かり難い

# ファイルの説明



ClientAPI.java	HTTPでSPARQLリクエストを受け付けるServlet
Import.java	気象庁XMLのfeedをRDFに変換してJenaに登録
Import2.java	気象台情報をJenaに登録
<b>RDF.java</b>	<b>SPARQLを実行するコマンド</b>
data.zip	2013/01/01～2013/12/02までのfeedを圧縮したもの
feed_135699366581_0_original.xml	feedデータのオリジナル
feed_135699366581_0_rdf.xml	feed_135699366581_0_original.xml をRDF形式に変換したもの
link2.txt	気象台情報
<b>query.txt</b>	<b>RDF.javaで実行するSPARQL</b>
resource.properties	1件だけ、1ヶ月分、1年分を切り替える
Dataset1	1件だけデータを登録したもの
DatasetAll	2013/01/01～2013/12/02までのデータを登録したもの
DatasetOneMonth	2013/01/01～2013/01/31までのデータを登録したもの
logs	ログの出力先
<b>Sample.html</b>	<b>1件分のRDFを図化したもの</b>
Sample.xml	1件分のRDF

- ・query.txt に実行したいSPARQLを書く
- ・RDF.java をアクティブにして、実行
  - ・ [Ctrl]+[F-11]もしくは、メニューの「Run」→「Run」
- ・2回目以降は、ショートカットが早い
  - ・ query.txt を編集して保存 ([Ctrl]+[S])
  - ・ そのまま実行 ([Ctrl]+[F-11])
- ・query.txt の書き方
  - ・ 試行錯誤し易いように、コメントアウト機能を独自実装
    - ・ 1文字目が # の行はコメントとして扱う
    - ・ 行の途中にある # は、そのまま実行される

# そろそろSPARQLの話が始まります

- ・SPARQLを思いつくままに書いて実行
  - ・自分のPC上で動くので、誰にも迷惑はかかりません
- ・環境がうまく動かない場合は
  - ・そんな事もあるかと、テスト用サイトを準備済み
  - ・ <http://cloud1.projectla.jp/handson/sparql1/>



ハルミン

## ・まずは最初の一歩

```
SELECT * WHERE {?s ?p ?o} LIMIT 100
```

三つ組みをそれぞれs,p,oと  
いう変数に割り当てる  
変数名は何でもいい

とんでもない件数の  
場合もあるので、100  
件だけ取得

## ・これと同じ意味

```
SELECT ?s ?p ?o WHERE {?s ?p ?o} LIMIT 100
```

表示する順

取得する順

## ・特定のものだけを取得

```
SELECT * WHERE {<urn:uuid:aa21d835-f6bd-35e6-94f0-f4dbbb387323> ?p ?o}  
LIMIT 100
```

```
SELECT * WHERE {?s <http://www.w3.org/2005/Atom#title> ?o} LIMIT 100
```

```
SELECT * WHERE {?s ?p "府県天気概況"} LIMIT 100
```

## ・PREFIXを使って省略

```
PREFIX atom: <http://www.w3.org/2005/Atom#>  
SELECT * WHERE {?s atom:title ?o} LIMIT 100
```

## ・レスポンスの形式を変更

```
SELECT ?s ?p ("府県天気概況" AS ?o) WHERE {?s ?p "府県天気概況"}  
LIMIT 100
```

固定のレスポンスを変数oに割り当てる  
カッコが無いとエラー

## ・フィルタによる絞り込み

```
SELECT * WHERE {?s ?p ?o . FILTER(?p = <http://www.w3.org/2005/Atom#title>)} LIMIT 100
```

ピリオドで区切る

フィルタ条件を指定

```
SELECT * WHERE {?s ?p ?o .  
  FILTER(?p = <http://www.w3.org/2005/Atom#title> && ?o = "府県天気概況")}  
  LIMIT 100
```

```
SELECT * WHERE {?s ?p ?o . FILTER(?o < 100)} LIMIT 100
```

数値型であれば、大小比較もOK

```
SELECT * WHERE {?s ?p ?o . FILTER(REGEX(?o ,"天気概況"))} LIMIT 100
```

```
SELECT * WHERE {?s ?p ?o . FILTER(REGEX(?o ,"^[0-9]"))} LIMIT 100
```

正規表現も使用可能



## ・件数を知りたい

```
SELECT (COUNT(*) AS ?count) WHERE {?s ?p ?o}
```

countという変数に割り当てる  
カッコが無いとエラー

## ・条件に合うものの件数

```
SELECT (COUNT(*) AS ?count)  
WHERE {?s <http://www.w3.org/2005/Atom#title> ?o}
```

## ・要素で集計をする

```
SELECT ?p (COUNT(?o) AS ?count) WHERE {?s ?p ?o} GROUP BY ?p
```

## ・RDBでおなじみの機能を使ってみる

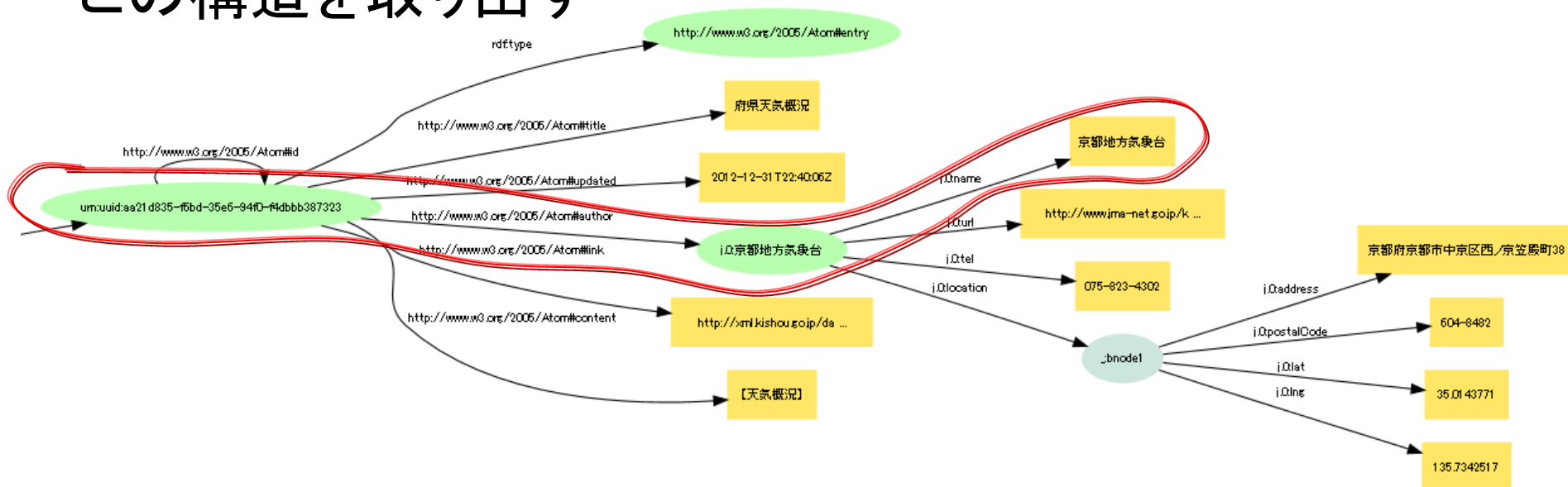
```
SELECT DISTINCT ?p WHERE {?s ?p ?o} LIMIT 100
```

```
SELECT ?p (COUNT(?o) AS ?count) WHERE {?s ?p ?o}  
GROUP BY ?p ORDER BY ?count
```

```
SELECT ?p (COUNT(?o) AS ?count) WHERE {?s ?p ?o}  
GROUP BY ?p ORDER BY DESC(?count)
```

# RDFの構造を活用 - 1

## この構造を取り出す



PREFIX atom: <http://www.w3.org/2005/Atom#>

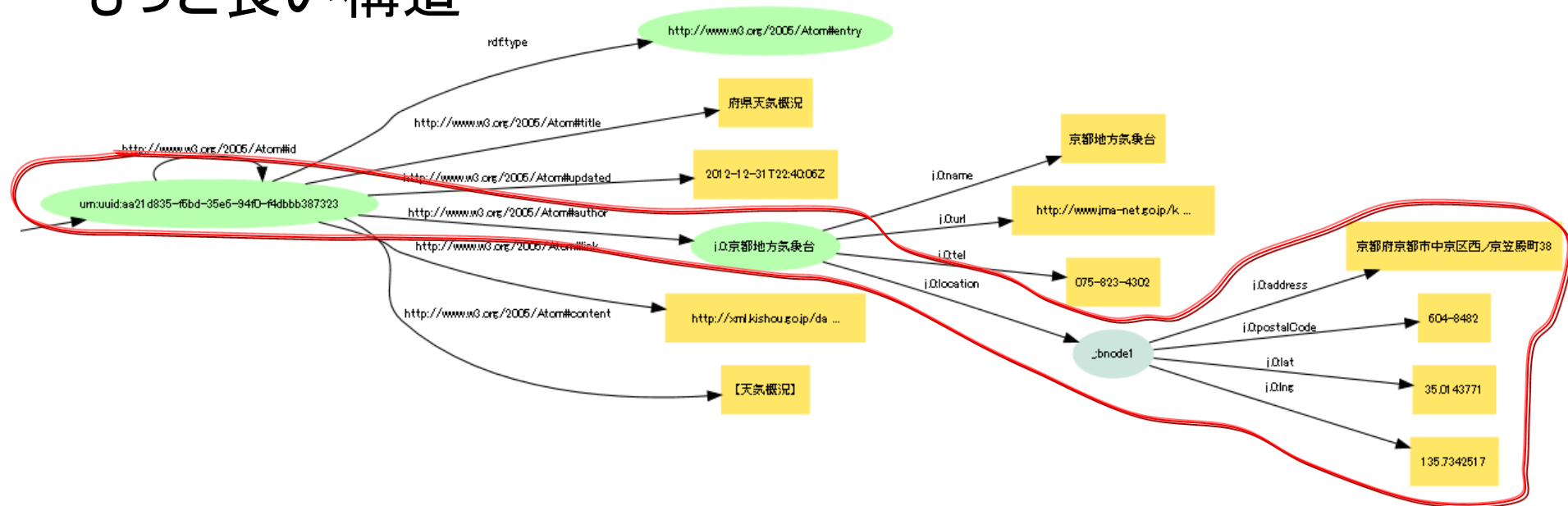
PREFIX j:0: <http://cloud1.projectla.jp/sparql/link2/>

```
SELECT * WHERE {?s atom:author ?o . ?o j:0:name ?o2 } LIMIT 100
```

<http://www.kanzaki.com/works/2005/rap/graph-check>

# RDFの構造を活用 - 2

## もっと長い構造



PREFIX atom: <http://www.w3.org/2005/Atom#>

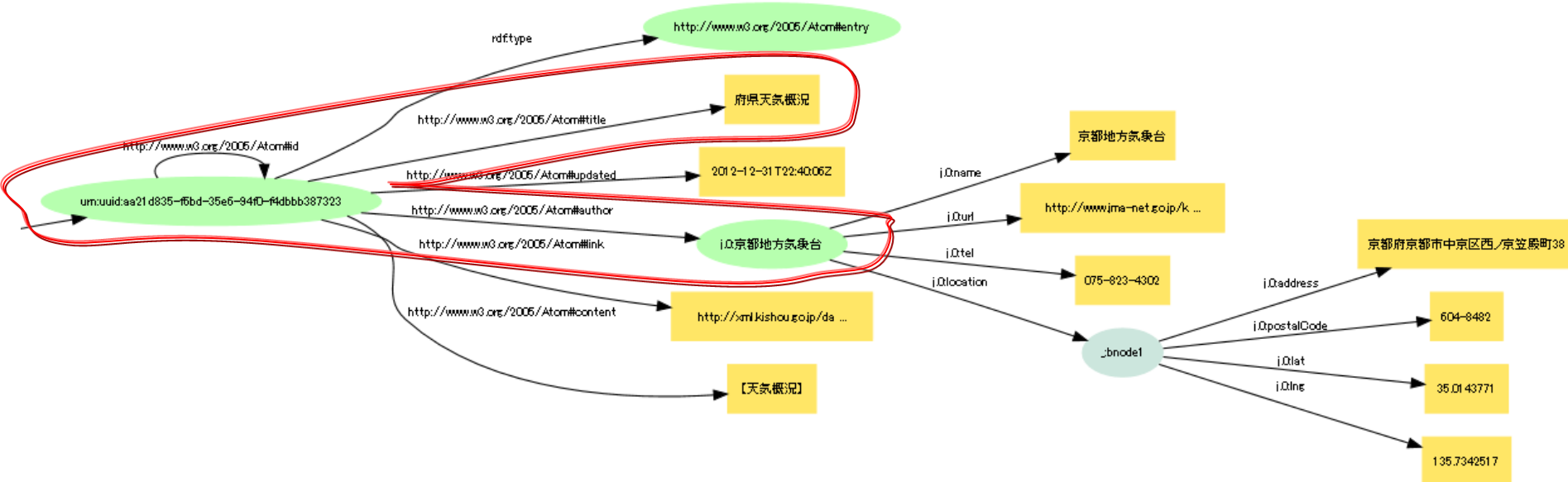
PREFIX j.0: <http://cloud1.projectla.jp/sparql/link2/>

```
SELECT * WHERE {?s atom:author ?o . ?o j.0:location ?o2 . ?o2 ?p3 ?o3}
LIMIT 100
```

<http://www.kanzaki.com/works/2005/rap/graph-check>

# RDFの構造を活用 - 3

## 「府県天気概況」を出している気象台



```
PREFIX atom: <http://www.w3.org/2005/Atom#>  
PREFIX j.0: <http://cloud1.projectla.jp/sparql/link2/>
```

```
SELECT * WHERE { ?s atom:title "府県天気概況" . ?s atom:author ?o2 }  
LIMIT 100
```

<http://www.kanzaki.com/works/2005/rap/graph-check>

# 1ヶ月分のデータを使用

## •resources.properties を修正

```
# data.path=./Dataset1  
data.path=./DatasetOneMonth  
# data.path=./DatasetAll
```

## •ブラウザでやっている人

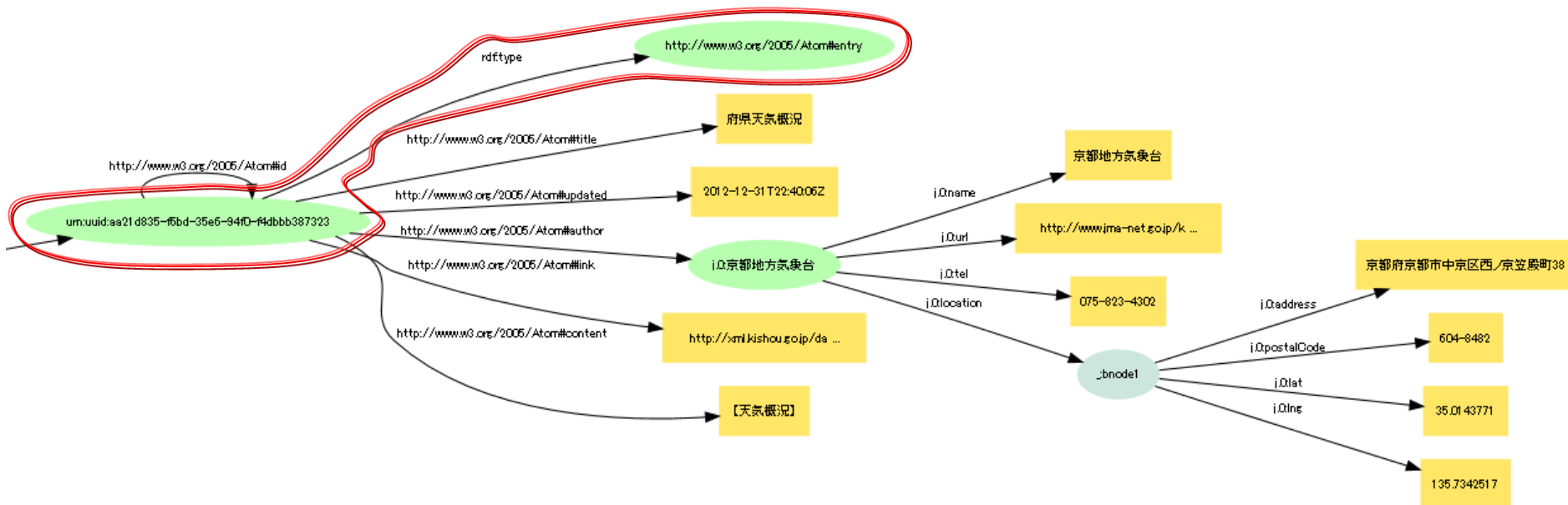
- ・ 別ブラウザで1ヶ月版を開く
  - <http://cloud1.projectla.jp/handson/sparqlOneMonth/>
- 全件見たい場合は「LIMIT 100」を外す
  - サーバとネットワークとクライアントに負荷がかかります



ハルミン

# 1ヶ月分のデーター1

・entryデータが何件あるか？



PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

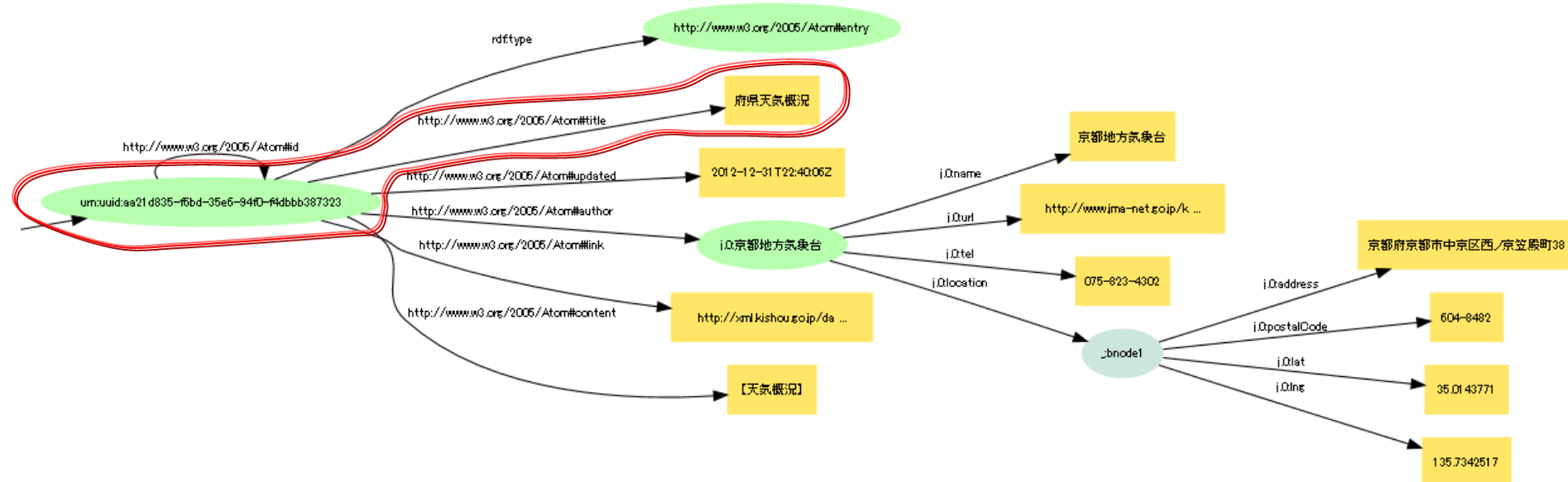
PREFIX atom: <http://www.w3.org/2005/Atom#>

```
SELECT (COUNT(*) AS ?count) WHERE {?s rdf:type atom:entry}
```

<http://www.kanzaki.com/works/2005/rap/graph-check>

# 1ヶ月分のデーター2

・どんなデータがあるか？



PREFIX atom: <http://www.w3.org/2005/Atom#>

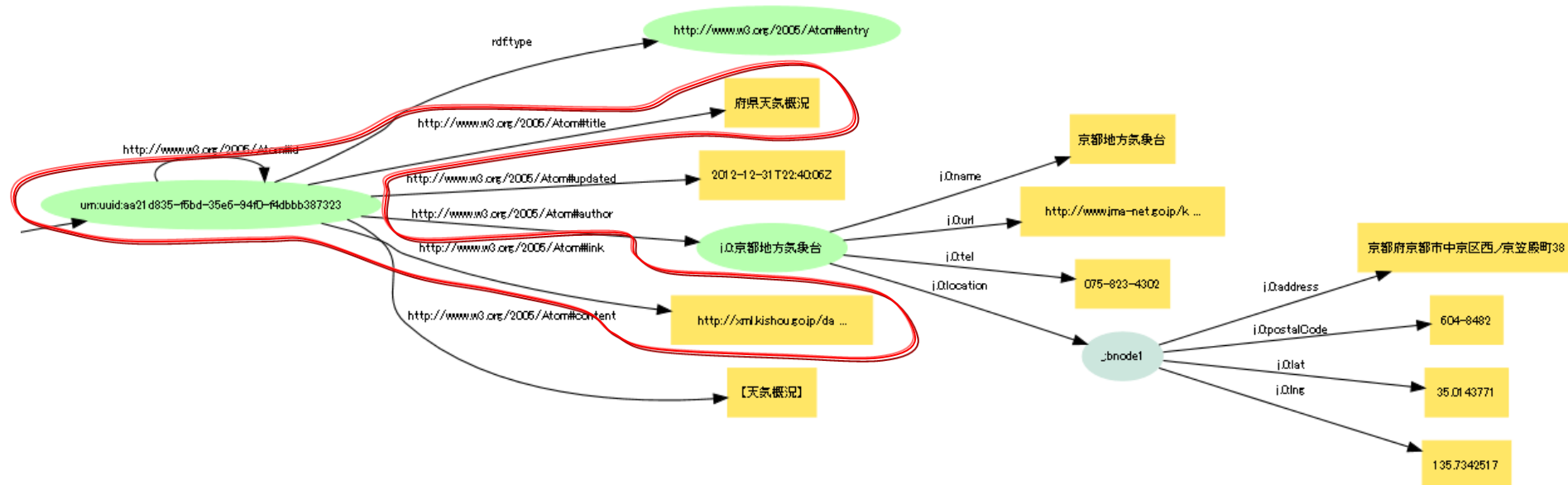
SELECT DISTINCT ?o WHERE {?s atom:title ?o}

<http://www.kanzaki.com/works/2005/rap/graph-check>



# 1ヶ月分のデーター3

## ・気になるtitleのcontent



PREFIX atom: <http://www.w3.org/2005/Atom#>

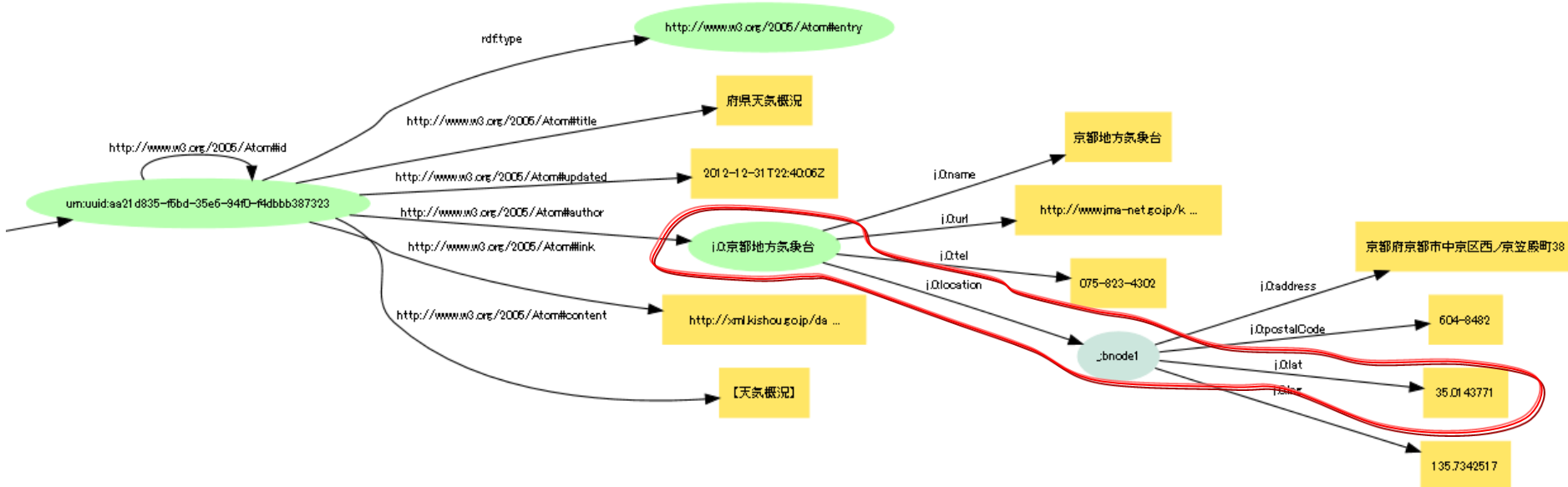
SELECT \* WHERE

{?s atom:title "異常天候早期警戒情報" . ?s atom:content ?o2} LIMIT 100

<http://www.kanzaki.com/works/2005/rap/graph-check>

# 1ヶ月分のデーター4

## ・全気象台を北から順にリスト



PREFIX atom: <<http://www.w3.org/2005/Atom#>>

PREFIX j.0: <<http://cloud1.projectla.jp/sparql/link2/>>

```
SELECT ?s WHERE {?s j.0:location ?o . ?o j.0:lat ?o2}
ORDER BY DESC(?o2)
```

<http://www.kanzaki.com/works/2005/rap/graph-check>

## • 様々なクエリを書いてみる

- 各気象台の出したentry数を多い順にリスト
- 各気象台の出したentryのcontentに「雪」が入っている数
- 以下のエリア内の気象台から出された「気象警報・注意報」数
  - 45.41024, 138.28024 ~ 41.27589, 146.5495
- その他、面白いクエリを考えてみてください
  - 参考資料: SPARQL1.1
  - <http://www.asahi-net.or.jp/~ax2s-kmtm/internet/rdf/rdf-sparql-query.html>
  - [http://docs.oracle.com/cd/E49329\\_01/appdev.121/b72469/sdo\\_rdf\\_concepts.htm#CIHEAEC](http://docs.oracle.com/cd/E49329_01/appdev.121/b72469/sdo_rdf_concepts.htm#CIHEAEC)

## • 約1年分のデータに対しても、同じ事をやってみる

- resources.properties を修正

```
# data.path=./Dataset1  
# data.path=./DatasetOneMonth  
data.path=./DatasetAll
```

## • 後で、各自、今日イチのクエリを発表

- Webサイト上で実行してログで確認
- <http://cloud1.projectla.jp/handson/sparqlAll/>



ハルミン