

「セマンティックWeb技術に触れてみよう！  
RDF/SPARQLハンズオン勉強会」  
～ オープンデータからLinkedDataまでを総ざらい～

# LODについて

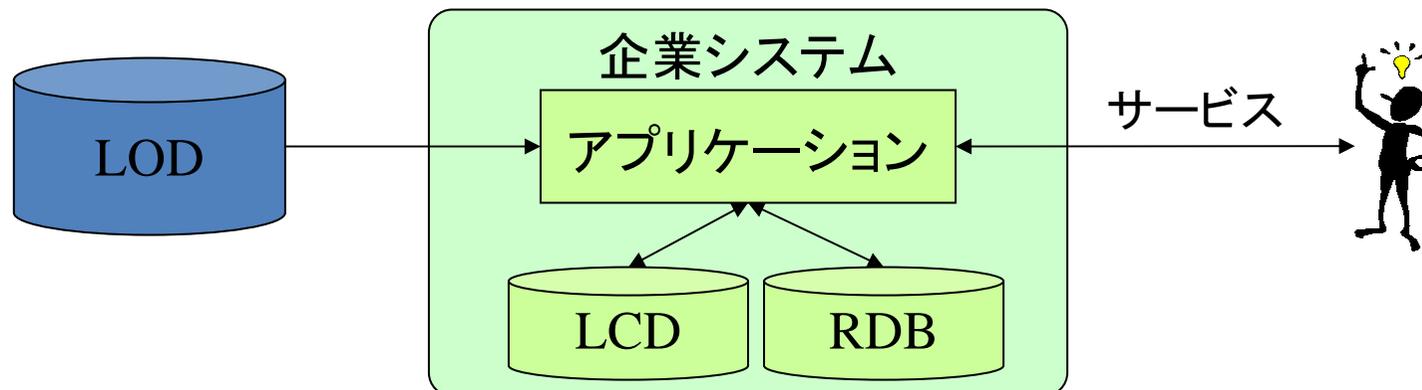
2013/12/21  
コンテキスト・コンピューティング研究部会  
サブリーダー 小林 茂

- オープンデータについて
  - オープンガバメント
- セマンティックWeb技術(RDF, SPARQL)
  - RDFとは
  - RDFの表現形式: タートル, RDFa, マイクロデータ
  - RDFグラフへの問い合わせ: SPARQL
  - 利用環境(SPARQL Timeliner, SparqlEPCU, LinkData)
- LODについて
  - Linked Data, LOD, LCD, LODクラウド図, ...
  - 事例

- RDFグラフで繋ったデータをLinked Dataと呼ぶ
  - RDFで公開されるデータ
  - 外部から参照可能
- Linked Data であるための4条件  
Tim Berners-Leeが発信(2006-07-27)
  - 事柄の名前にURIを使うこと
    - 全てのモノ, コトにURIを付与する
  - 名前の参照が HTTP URI でできること
    - DOI(Digital Object Identifier)やURNではなくHTTPで
  - URIを参照したときに関連情報が得られること
    - そのURIの内容が理解できるように
  - 外部へのリンクを含めること
    - 他にリンクすることで, より多くの発見が期待できる

# LODとLCD

- LODとは、オープンなLinked Dataのこと
  - LODは、本来Linked Open Dataであるが、Linking Open Data と言うことがある
    - 公開されたLinked Dataを集め、広げる動きを指すときに使う
    - 繋がりが増大する動きを強調するためか(?)
- LCDとは、クローズドなLinked Dataのこと
- 企業内の非公開のデータをLinked Data (LCD)として持ち、外部で公開されたLODを活用して、企業システムの強化、新規ビジネスの創出に繋げる



- オープンデータは、生活や社会の基盤となりうる生のデータを二次利用可能な形で公開するという考え方
- その実現にLODを用いるという進め方
- オープンガバメント(政府主導)の枠を超えて進める方向
  - 政府も民間も個人も同じ立場で
- Web上に公開されたデータを皆で繋げて、新しい価値を生み出す動き

# LODの到達度

- ★ 1. 形式に関係なくオープンライセンスで公開する
  - 表データを画像で公開
- ★★ 2. コンピュータで処理可能な構造化データで公開する
  - Excelで公開(ある1社で利用する形式)
- ★★★ 3. オープンなデータ形式で公開する
  - CSVデータで公開
- ★★★★ 4. URIを利用してモノに名前を付ける
  - RDFデータで公開する
- ★★★★★ 5. 他のデータへリンクする

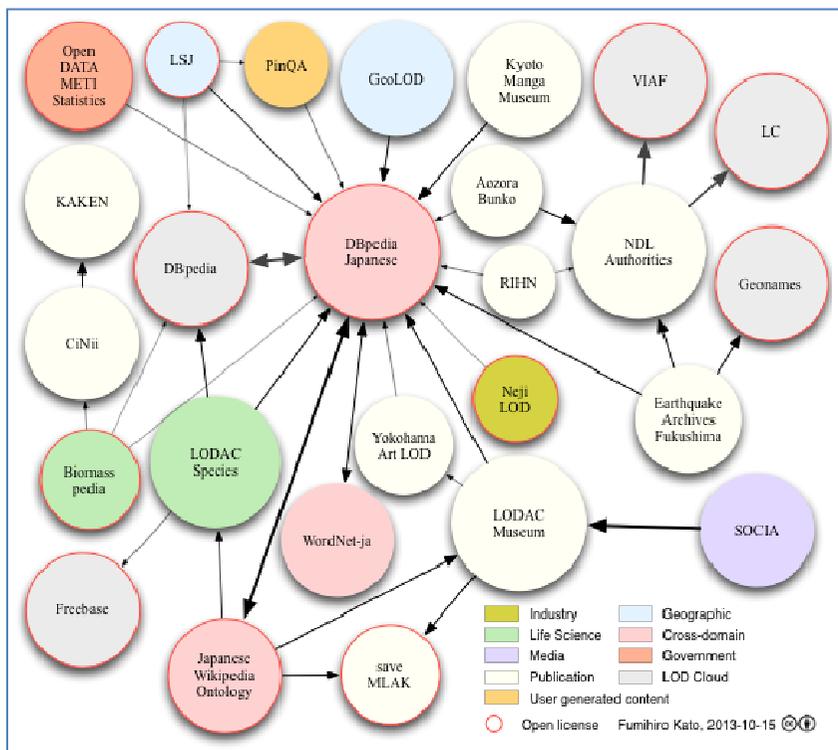


出典: <http://www.w3.org/DesignIssues/LinkedData.html>

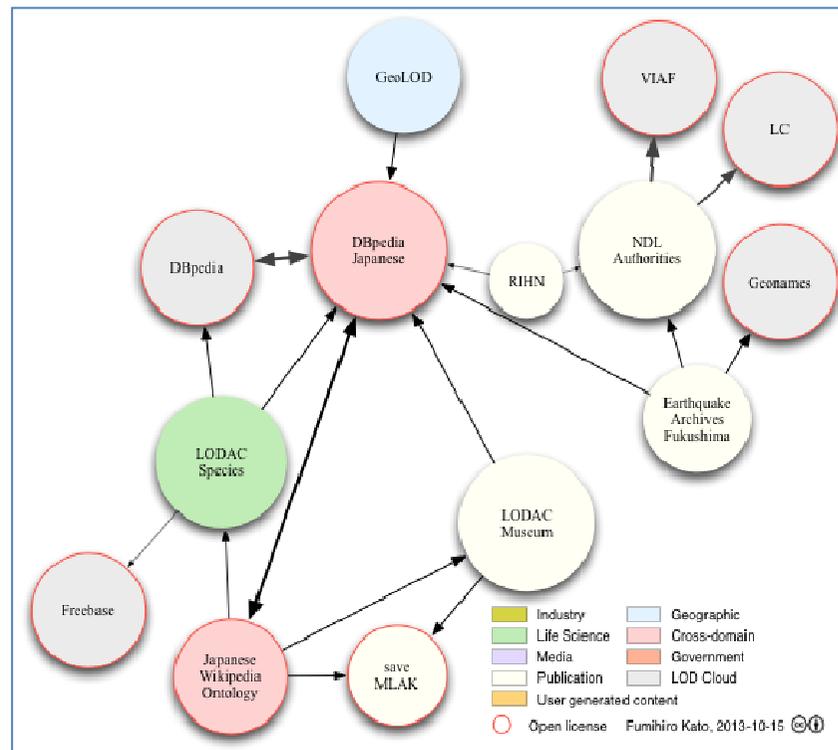


# 日本版LODクラウド図

2013-10-15



本家より緩い条件で



本家の図に入るための条件では

出典: <http://linkedopendata.jp/?p=411>

# DBpedia

- WikipediaにあるInfoboxのデータをLOD化
- 英語版が中心だが、日本語版Wikipediaも参照

The image shows two overlapping browser windows. The top window is a DBpedia page for Mount Fuji, displaying a table with properties and values. The bottom window is the English Wikipedia page for Mount Fuji, with a red box highlighting the Infobox. Callout boxes with arrows point to the DBpedia page, the Wikipedia page, and the Infobox.

出典: [http://en.wikipedia.org/wiki/Mount\\_Fuji](http://en.wikipedia.org/wiki/Mount_Fuji)

DBpedia

[http://dbpedia.org/resource/Mount\\_Fuji](http://dbpedia.org/resource/Mount_Fuji)

Wikipedia

Infobox

# 日本語DBpedia

## 日本語DBpedia

About: 夏目漱石

An Entity of Type : [Artist](#), from Named Graph : <http://ja.dbpedia.org>, within Data Space : [ja.dbpedia.org](http://ja.dbpedia.org)

夏目 漱石(なつめ そうせき、1867年2月9日(慶応3年1月5日) - 1916年(大正5年)12月9日)は、日本の小説家、評論家、英文学者。本名、金之助(きんのすけ)。江戸の牛込馬場下横町(現在の東京都新宿区喜久井町)出身。俳号は愚陀仏。大学時代に正岡子規と出会い、俳句を学ぶ。帝国大学(後の東京帝国大学、現在の東京大学)英文科卒業後、松山で愛媛県尋常中学教師、熊本で五高教授などを務めた後、イギリスへ留学。帰国後、東京帝大講師として英文学を講じながら、「吾輩は猫である」を雑誌『ホトギス』に発表。これが評判になり「坊っちゃん」「倫敦塔」などを書く。その後朝日新聞社に入社し、「虞美人草」「三四郎」などを掲載。当初は余裕派と呼ばれた。「修善寺の大患」後は、『行人』『こゝろ』『硝子戸の中』などを執筆。「則天去私(そくてんきよし)」の境地に達したといわれる。晩年は胃潰瘍に悩まされ、「明暗」が絶筆となった。

Property	Value
dbpedia-owl:abstract	夏目 漱石(なつめ そうせき、1867年2月9日(慶応3年1月5日) - 1916年(大正5年)12月9日)は、日本の小説家、評論家、英文学者。本名、金之助(きんのすけ)。江戸の牛込馬場下横町(現在の東京都新宿区喜久井町)出身。俳号は愚陀仏。大学時代に正岡子規と出会い、俳句を学ぶ。帝国大学(後の東京帝国大学、現在の東京大学)英文科卒業後、松山で愛媛県尋常中学教師、熊本で五高教授などを務めた後、イギリスへ留学。帰国後、東京帝大講師として英文学を講じながら、「吾輩は猫である」を雑誌『ホトギス』に発表。これが評判になり「坊っちゃん」「倫敦塔」などを書く。その後朝日新聞社に入社し、「虞美人草」「三四郎」などを掲載。当初は余裕派と呼ばれた。「修善寺の大患」後は、『行人』『こゝろ』『硝子戸の中』などを執筆。「則天去私(そくてんきよし)」の境地に達したといわれる。晩年は胃潰瘍に悩まされ、「明暗」が絶筆となった。
dbpedia-owl:activeYearsEndYear	1916-01-01 00:00:00 (xsd:date)
dbpedia-owl:activeYearsStartYear	1905-01-01 00:00:00 (xsd:date)
dbpedia-owl:almaMater	dbpedia-ja:帝国大学
dbpedia-owl:birthName	夏目金之助
dbpedia-owl:birthPlace	dbpedia-ja:武蔵国
dbpedia-owl:deathPlace	dbpedia-ja:東京
dbpedia-owl:education	dbpedia-ja:学士
dbpedia-owl:genre	dbpedia-ja:小説

<http://ja.dbpedia.org/resource/夏目漱石>

[http://dbpedia.org/resource/Natsume\\_Soseki](http://dbpedia.org/resource/Natsume_Soseki)

## 日本語Wikipedia

年2月9日(慶応3年1月5日)は、日本の小説家、評論家のすけ)。江戸の牛込馬場下横町(現在の東京都新宿区喜久井町)出身。俳号は愚陀仏。大学時代に正岡子規と出会い、俳句を学ぶ。帝国大学(後の東京帝国大学、現在の東京大学)英文科卒業後、松山で愛媛県尋常中学教師、熊本で五高教授などを務めた後、イギリスへ留学。帰国後、東京帝大講師として英文学を講じながら、「吾輩は猫である」を雑誌『ホトギス』に発表。これが評判になり「坊っちゃん」「倫敦塔」などを書く。その後朝日新聞社に入社し、「虞美人草」「三四郎」などを掲載。当初は余裕派と呼ばれた。「修善寺の大患」後は、『行人』『こゝろ』『硝子戸の中』などを執筆。「則天去私(そくてんきよし)」の境地に達したといわれる。晩年は胃潰瘍に悩まされ、「明暗」が絶筆となった。

1912年9月13日(明治天皇の大喪の礼の日)<sup>[1]</sup>

誕生 夏目金之助  
1867年2月9日  
● 日本・武蔵国江戸牛込馬場下横町

死没 1916年12月9日(満49歳没)  
● 日本・東京

<http://ja.wikipedia.org/wiki/夏目漱石>

[http://en.wikipedia.org/wiki/Natsume\\_Soseki](http://en.wikipedia.org/wiki/Natsume_Soseki)

# SPARQL: クエリーと結果

## DBpediaのSPARQLサイト

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dbpprop: <http://dbpedia.org/property/>
SELECT ?musician ?abstract
FROM <http://dbpedia.org/>
WHERE {
  ?musician <http://dbpedia.org/resource/Date_of_birth>
  dbpprop:abstract ?abstract
  FILTER ( lang(?abstract) = "ja" ) .
}
LIMIT 5
OFFSET 100
```

<http://dbpedia.org/sparql/>

## SPARQL問合せ結果

musician	abstract
<a href="http://dbpedia.org/resource/Sachi_Tainaka">http://dbpedia.org/resource/Sachi_Tainaka</a>	“ <a href="#">タイナカ 彩智</a> ( <a href="#">タイナカ さち</a> , 1986年4月30日 - )は、日本の女性シンガーソングライター。兵庫県加古川市生まれの兵庫県神戸市育ち。旧芸名はタイナカ サチ、本名は田井中 彩智 (読み同じ)。血液型A型。身長は156.5cm (本人のブログにて)。神戸学院大学附属高等学校卒業。2010年12月、自身のレーベルtutu records (チュチュレコース) を設立するとともに、名前の表記を現在のタイナカ彩智に変更した。”@ja
<a href="http://dbpedia.org/resource/Yamantaka_Eye">http://dbpedia.org/resource/Yamantaka_Eye</a>	“ <a href="#">山塚 アイ</a> ( <a href="#">やまつか アイ</a> , YAMATSUKA EYE, 1964年 (昭和39年) 2月13日 - )は日本の音楽家。兵庫県神戸市三宮出身。”@ja
<a href="http://dbpedia.org/resource/Isao_Tomita">http://dbpedia.org/resource/Isao_Tomita</a>	“ <a href="#">富田 勲</a> ( <a href="#">とみた いさお</a> , 1932年4月22日 - )は、日本の作曲家、編曲家、シンセサイザー奏者である。”@ja
<a href="http://dbpedia.org/resource/Akira_Sakata">http://dbpedia.org/resource/Akira_Sakata</a>	“ <a href="#">坂田 明</a> ( <a href="#">さかた あきら</a> , 1945年2月21日 - )は、広島県呉市広長浜出身のジャズサクソフ奏者、タレント、俳優。ミジコ研究家としても知られている。息子もミュージシャンで、元・バンド「Polaris」のドラム担当の坂田学。”@ja
<a href="http://dbpedia.org/resource/Kyoko_Fukada">http://dbpedia.org/resource/Kyoko_Fukada</a>	“ <a href="#">深田 恭子</a> ( <a href="#">ふかた きょうこ</a> , 1982年11月2日 - )は、日本の女優、歌手、タレント。本名同じ。愛称は、深キョン、きょーこりん。東京都北区出身。ホリプロ所属。堀越高等学校卒業。高校の同級生には加藤あい・椎名法子・仲根かすみ・Sowelu・小高早紀・上良早紀・藤原竜也らがいる。第21回ホリプロタレントスカウトキャラバングランプリ受賞をきっかけに芸能界入り。ドラマ『神様、もう少しだけ』でブレイクし、その後、数々のドラマや映画、CMに出演。歌手としてCDもリリースしている。代表作はドラマ『富豪刑事』、映画『下妻物語』など。”@ja

# SPARQL式

```
PREFIX dterms: <http://purl.org/dc/terms/>
PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>
SELECT ?musician ?abstract
FROM <http://dbpedia.org/>
WHERE {
  ?musician dterms:subject <http://dbpedia.org/resource/Category:Japanese_musicians> ;
  dbpedia-owl:abstract ?abstract
  FILTER ( lang(?abstract) = "ja" ) .
}
LIMIT 5
OFFSET 100
```

<http://dbpedia.org/sparql/>

## WHERE句での記法

- 1つの文では、3つ組を指定する
  - 文はピリオドで終わる
    - URIは、<>で囲む
  - リテラルは、引用符""で囲む
    - 変数は、?で始まる

### •主語が同じ場合の記法

A B1 C1 .	↔	A B1 C1 ;
A B2 C2 .		B2 C2 .

### •フィルタリングできる

# SPARQL問合せ結果

The screenshot shows two browser windows. The left window displays the SPARQL query results in HTML format, with a yellow box labeled "HTMLで表示" (Display in HTML) overlaid on it. The right window displays the same results in XML format, with a yellow box labeled "XML" overlaid on it. Below the XML window, a cyan box contains the text "SPARQL Query Results XML Format : W3C 勧告 2008-01-15".

musician	abstract
<a href="http://dbpedia.org/resource/Sachi_Tainaka">http://dbpedia.org/resource/Sachi_Tainaka</a>	“タイナカ 彩智(タイナカさち、1986年4月30日 - )は、日本の女優。兵庫県神戸市育ち。旧芸名A型。身長は156.5cm(本人のブログにてのレーベルtutu records(チュチュレコード)に変更した。”@ja
<a href="http://dbpedia.org/resource/Yamantaka_Eye">http://dbpedia.org/resource/Yamantaka_Eye</a>	“山塚 アイ(やまつか アイ、YAMATSUKA EYE、1964年(昭和39年)2月21日 - )は、日本の女優。東京都北区出身。加藤あい・椎名法子・仲根かすみ・Soweluの4人組のユニット「4人娘」のメンバー。数々のドラマや映画、映画『下妻物語』など。”@ja
<a href="http://dbpedia.org/resource/Isao_Tomita">http://dbpedia.org/resource/Isao_Tomita</a>	“富田 勲(とみた いさお、1932年4月22日 - )は、日本の作曲家、編曲家、指揮者。東京都出身。数々のドラマや映画、映画『下妻物語』など。”@ja
<a href="http://dbpedia.org/resource/Akira_Sakata">http://dbpedia.org/resource/Akira_Sakata</a>	“坂田 明(さかた あきら、1945年2月21日 - )は、日本の俳優、脚本家、演出家。広島県呉市出身。数々のドラマや映画、映画『下妻物語』など。”@ja
<a href="http://dbpedia.org/resource/Kyoko_Fukada">http://dbpedia.org/resource/Kyoko_Fukada</a>	“深田 恭子(ふかだ きょうこ、1982年11月2日 - )は、日本の女優。東京都北区出身。加藤あい・椎名法子・仲根かすみ・Soweluの4人組のユニット「4人娘」のメンバー。数々のドラマや映画、映画『下妻物語』など。”@ja

```
<sparql xmlns="http://www.w3.org/2005/sparql-results#" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.w3.org/2005/sparql-results# http://www.w3.org/2005/sparql-results.xsd">
  <head>
    <variable name="musician"/>
    <variable name="abstract"/>
  </head>
  <results distinct="false" ordered="true">
    <result>
      <binding name="musician"><uri>http://dbpedia.org/resource/Sachi_Tainaka</uri></binding>
      <binding name="abstract"><literal xml:lang="ja">タイナカ 彩智(タイナカ さち、1986年4月30日 - )は、日本の女優。兵庫県神戸市育ち。旧芸名A型。身長は156.5cm(本人のブログにてのレーベルtutu records(チュチュレコード)に変更した。”@ja</literal></binding>
    </result>
    <result>
      <binding name="musician"><uri>http://dbpedia.org/resource/Yamantaka_Eye</uri></binding>
      <binding name="abstract"><literal xml:lang="ja">山塚 アイ(やまつか アイ、YAMATSUKA EYE、1964年(昭和39年)2月21日 - )は、日本の女優。東京都北区出身。加藤あい・椎名法子・仲根かすみ・Soweluの4人組のユニット「4人娘」のメンバー。数々のドラマや映画、映画『下妻物語』など。”@ja</literal></binding>
    </result>
    <result>
      <binding name="musician"><uri>http://dbpedia.org/resource/Isao_Tomita</uri></binding>
      <binding name="abstract"><literal xml:lang="ja">富田 勲(とみた いさお、1932年4月22日 - )は、日本の作曲家、編曲家、指揮者。東京都出身。数々のドラマや映画、映画『下妻物語』など。”@ja</literal></binding>
    </result>
    <result>
      <binding name="musician"><uri>http://dbpedia.org/resource/Akira_Sakata</uri></binding>
      <binding name="abstract"><literal xml:lang="ja">坂田 明(さかた あきら、1945年2月21日 - )は、広島県呉市出身。数々のドラマや映画、映画『下妻物語』など。”@ja</literal></binding>
    </result>
    <result>
      <binding name="musician"><uri>http://dbpedia.org/resource/Kyoko_Fukada</uri></binding>
      <binding name="abstract"><literal xml:lang="ja">深田 恭子(ふかだ きょうこ、1982年11月2日 - )は、日本の女優。東京都北区出身。加藤あい・椎名法子・仲根かすみ・Soweluの4人組のユニット「4人娘」のメンバー。数々のドラマや映画、映画『下妻物語』など。”@ja</literal></binding>
    </result>
  </results>
</sparql>
```

# SparqlEPCU

LOD公開の  
ためのツール



- **LOD作成及び活用の支援サイト**
  - プログラミング無しで、LODデータを格納・検索できる
  - LODアプリサイトをJavaScriptで簡単に作成できる
- **提供機能**
  - RDF格納サイト
  - Sparql検索Webページ
  - RDF登録・検索のためのJavaScriptライブラリ
  - LODアプリサイトのサンプル



<http://lodcu.cs.chubu.ac.jp/SparqlEPCU/>

# SPARQL Timeliner

LOD公開の  
ためのツール



- SPARQLエンドポイントとクエリー式を指定し、コンテンツを表示させるWebアプリの基盤
- 関連する場所を地図上で示し、関連する期日をタイムライン上で示す
- クエリー式では定められた変数に対して応える式を書く
- コンテンツを左右に移動させると、対応する地図、タイムラインが同期する

SPARQL エンドポイント

SPARQL クエリー式

LOADボタンで実行指示

コンテンツ表示

タイムライン

地図

<http://uedayou.net/SPARQLTimeliner/>

15

# Link Data

LOD公開の  
ためのツール



- 理研(理化学研究所)が提供するサービス
- 表形式データをRDFに変換するツールを提供
- RDF公開の場を提供

	A	B
1		
2		
3		
4		



<http://linkdata.org/>

16

- オープンデータ概論 Ver.3 (国際大学GLOCOM 林雅之)  
<http://ssslide.com/www.slideshare.net/mhayasi/201311-28083066>
- オープンデータの技術よりなお話 (LODI 加藤文彦)  
<http://ssslide.com/www.slideshare.net/fumihiro/20130620-23239372>
- SparqlEPCUでLODをマッシュアップ (中部大学 年岡晃一)  
<http://www.slideshare.net/KohichiToshioka/sparql-epculod-28065847>
- お手軽LOD可視化ツール SPARQL Timeliner (ATR Creative 上田洋)  
<http://www.slideshare.net/uedayou/linked-open-datasparql-timeliner>
- Linkdata.org (理化学研究所 下山紗代子, 豊田哲郎)  
<http://www.slideshare.net/sayokoshimoyama/lodlinkdataorg>