FUJITSU

shaping tomorrow with you

Knowledge Graph in Japan: Open data beyond

Sep. 22, 2016@Beijing

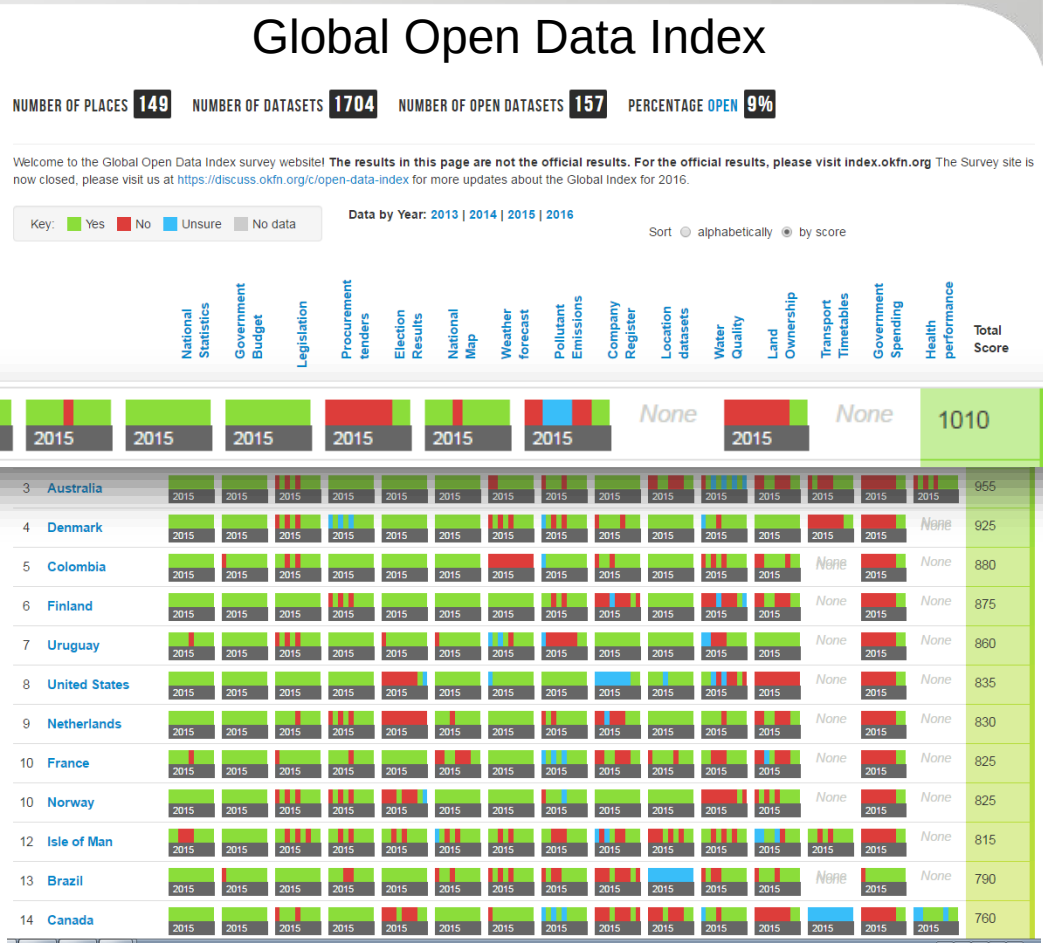
Fujitsu Laboratories Ltd.
AI Research Center
Nobuyuki Igata

株式会社富士通研究所
人工知能研究センター
井形　伸之

Copyright 2016 FUJITSU LABORATORIES LIMITED

Background: Open Data

- G8 (Fr, US, UK, DE, JP, IT, CA, RU) Open Data Charter released in June 2013
  - Identified 14 high-value categories
  - Dec.2015
  - Release data free of charge, under an open and unrestrictive license
- Over 40 countries in the world have websites for data disclosure
- data strategy



from http://global.census.okfn.org/ by Open Knowledge Foundation

Copyright 2016 FUJITSU LABORATORIES LIMITED

Activities of Open Data in Japan

FUJITSU

- Central Governments (中央政府)

from http://data.go.jp/ by

- Local Governments (地方自治体)
  - Over 200 local city (total 1,750) started open data activities
    - Catalog site, Hackathon using their

- Sciences (科学系)
  - JST, JAXA, National Diet Library, Academic Society published their own data with open license
  - Open Science
    - Cabinet Office, Gov. of Japan (内閣府)organized an expert committee to discuss how to achieve Open Science movement

- Public Services (公共サービス)
  - Tokyo Metro (東京メトロ) has opened their subway service situation and big contest (2014)
  - NHK has opened their TV show time table with LOD format (2015)
  - Metropolitan Police Department (警視庁) has opened a crime information that occurred in Tokyo city (2015)

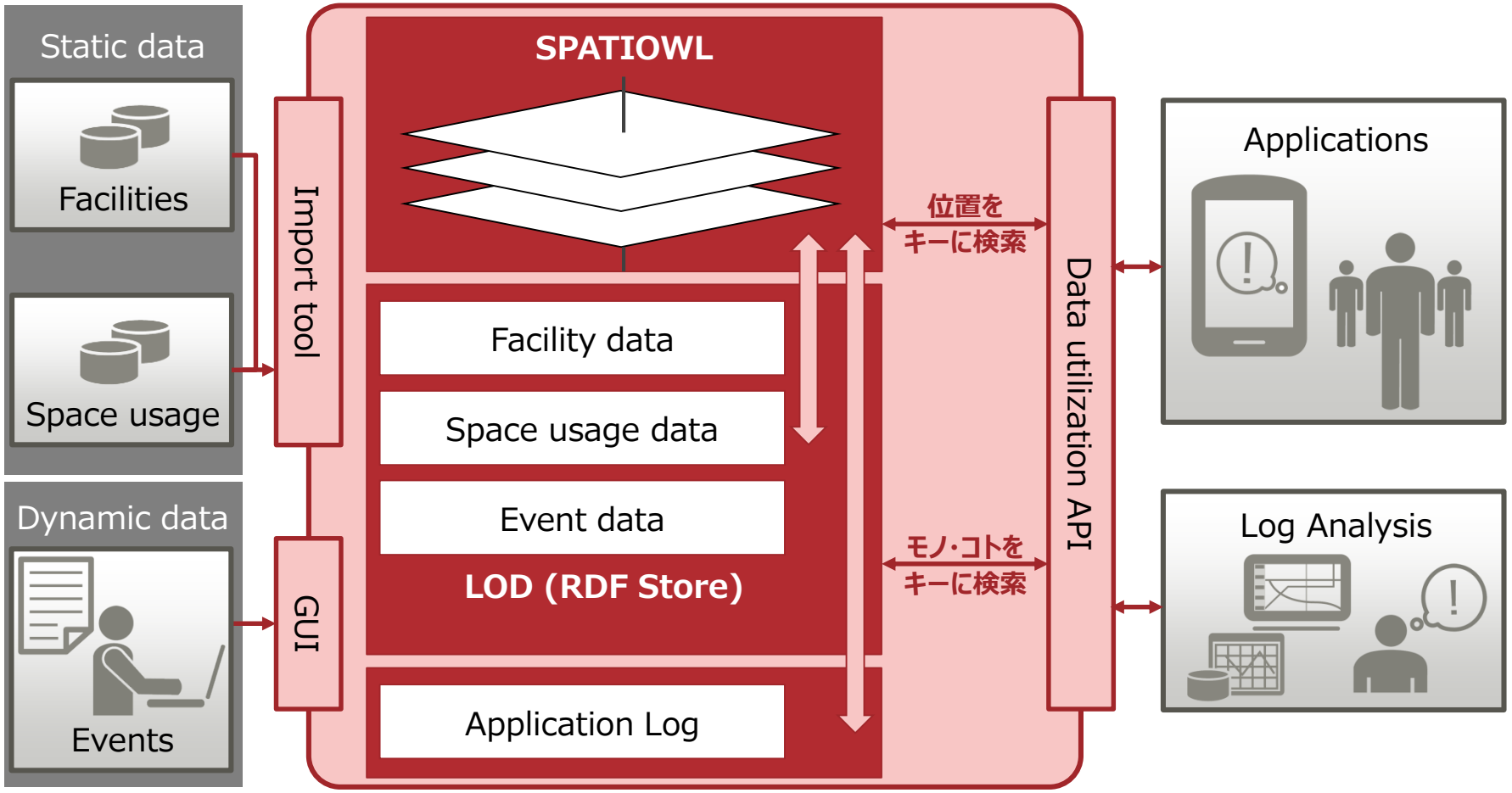
Copyright 2016 FUJITSU LABORATORIES LIMITED

Case of Local Governments: Kawasaki City

Copyright 2016 FUJITSU LABORATORIES LIMITED

Case of Local Governments: Kawasaki City

- Kawasaki City has published various information to support families with small children such as facilities, events, so on.
- Fujitsu Ltd. developed Data utilization Platform and an application for parents.



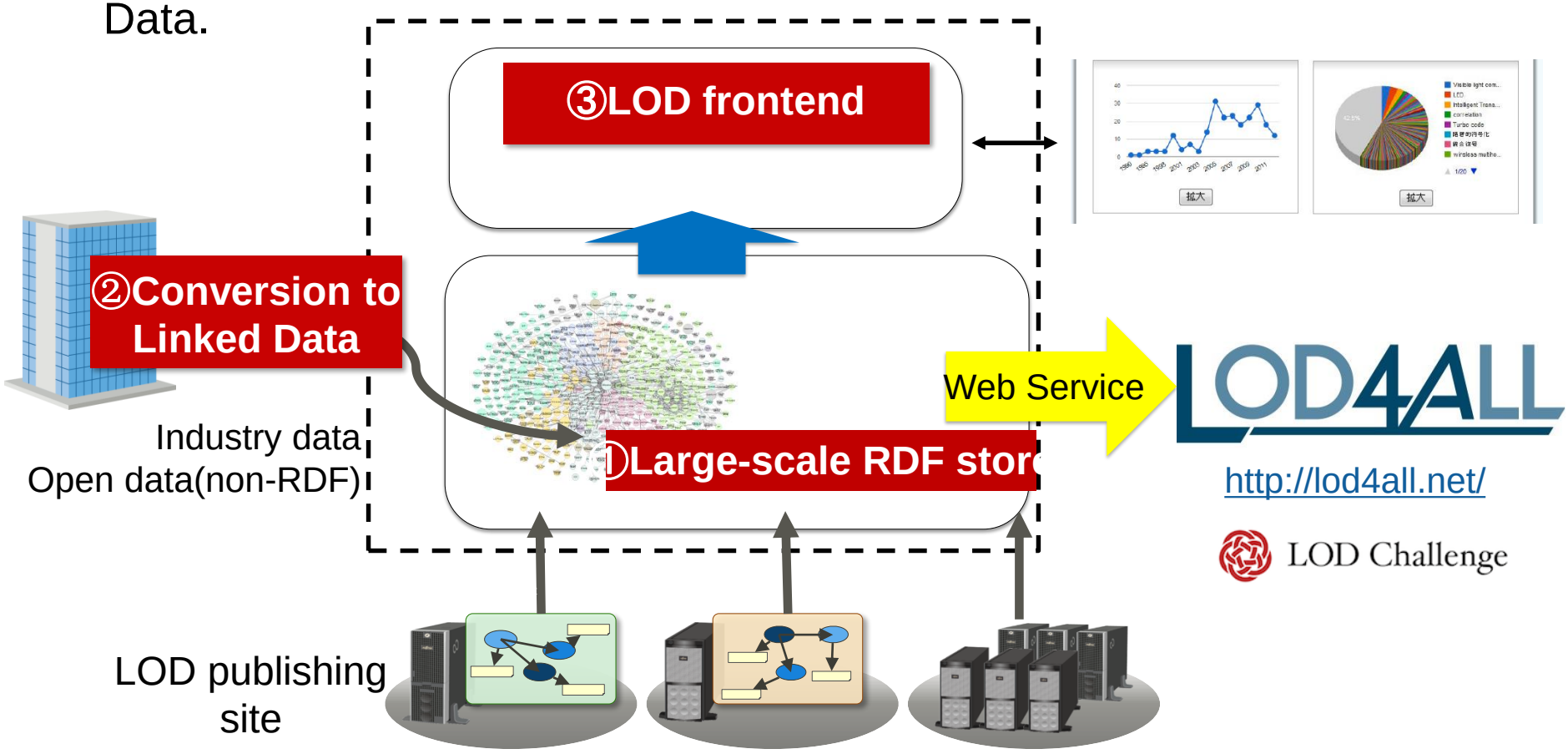
Copyright 2016 FUJITSU LABORATORIES LIMITED

Technologies related LOD by Fujitsu Labs.

FUJITSU

R&D for Data utilization platform with LOD technologies

- Cross over data utilization among LOD, Open Data (non-RDF), Industry Data.

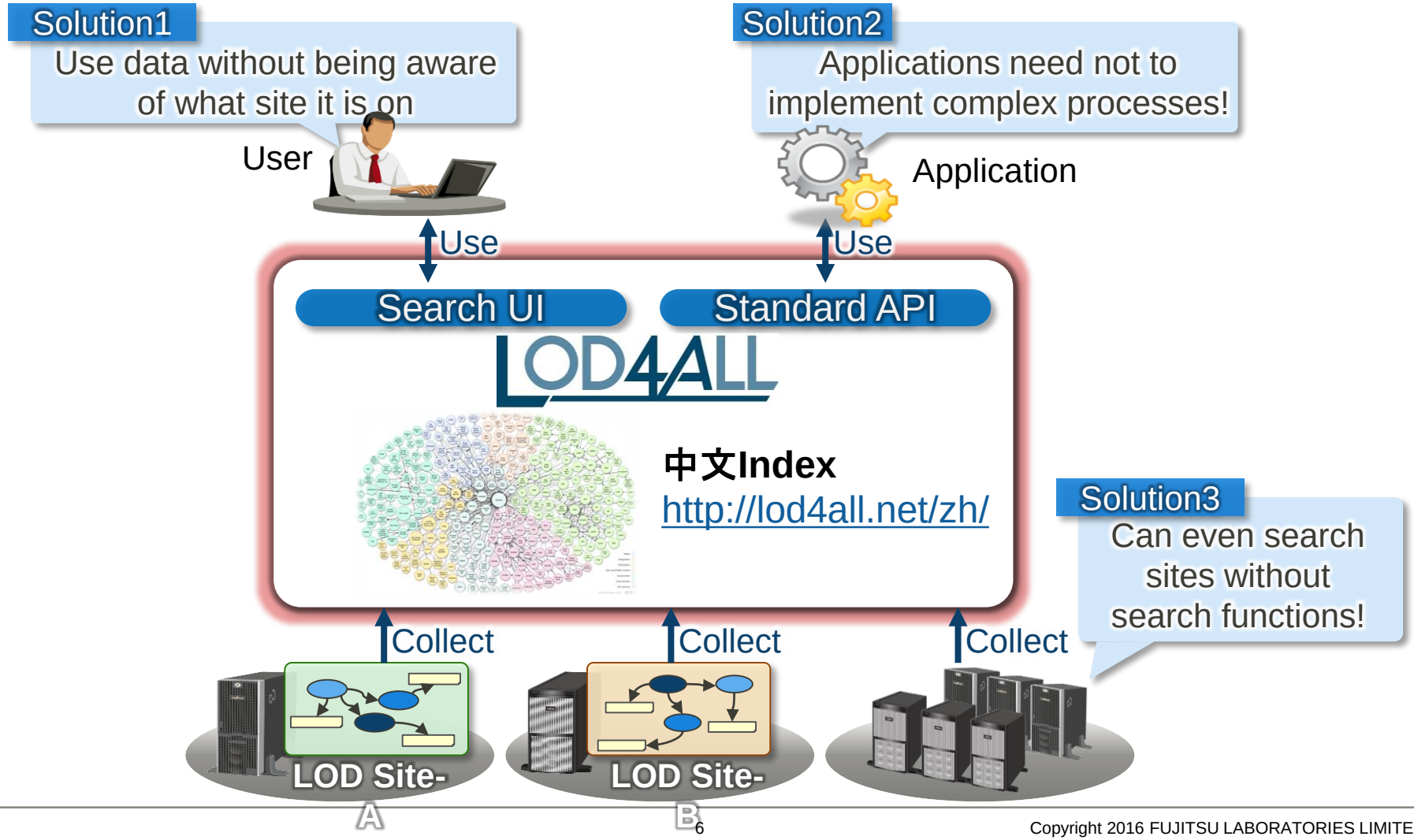


Copyright 2016 FUJITSU LABORATORIES LIMITED

LOD Search Engin LOD4ALL

FUJITSU

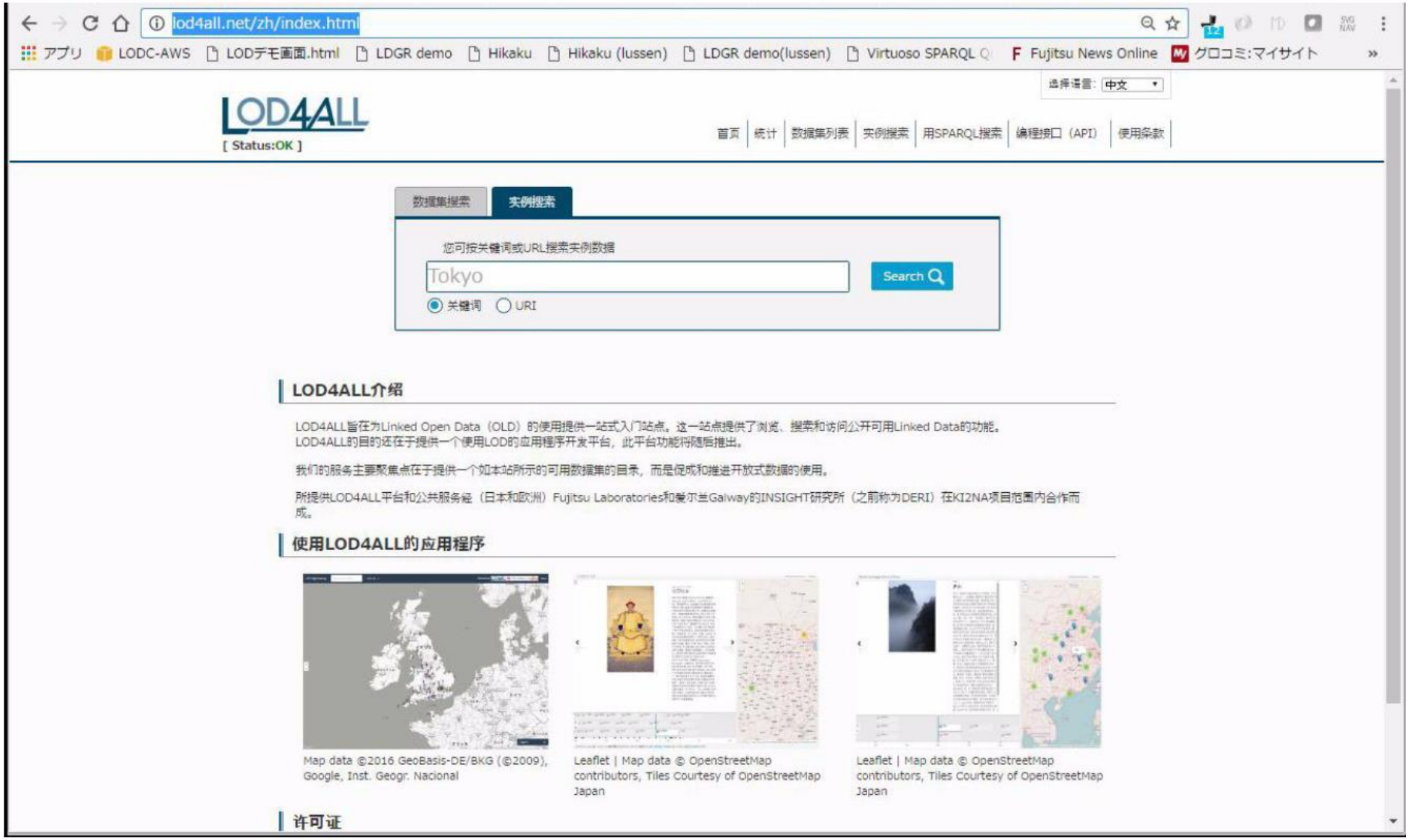
- Crawl LOD from around the world (several tens of billion data items), and provide a high-speed search service



Copyright 2016 FUJITSU LABORATORIES LIMITED

LOD Search Engin LOD4ALL

FUJITSU

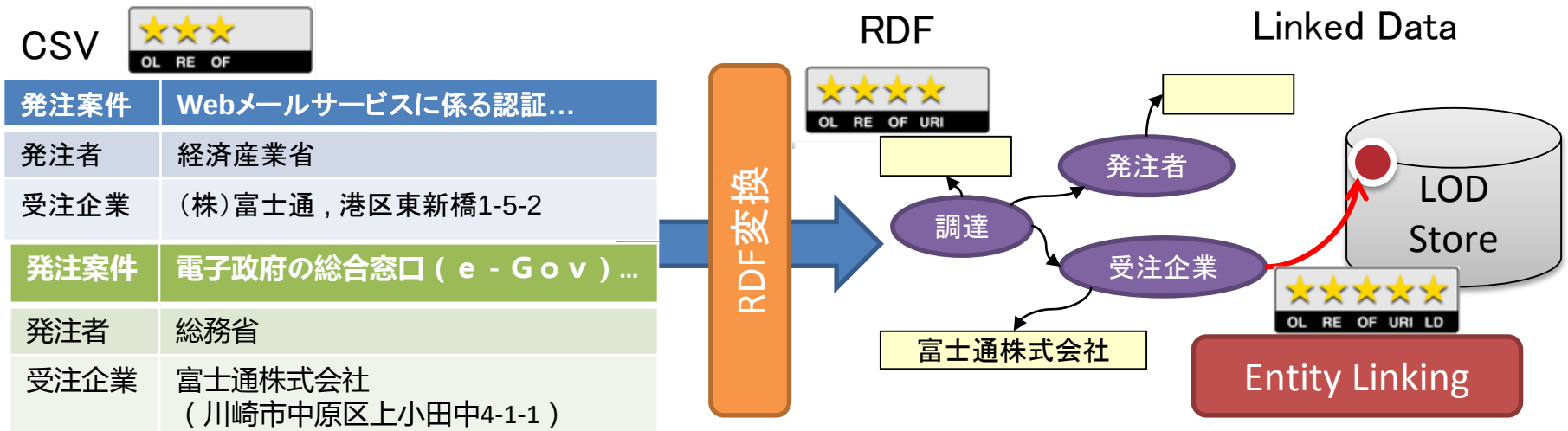


Copyright 2016 FUJITSU LABORATORIES LIMITED

Conversion to Linked Data

- Convert from CSV/XML(3★) to Linked Data(5★)
  - Step1: CSV/XML to RDF
  - Step2: Linking to other datasets
- Use cases：
  - Open/Close data (non-RDF) Link to LOD Could

Process of Linked Data Conversion



Copyright 2016 FUJITSU LABORATORIES LIMITED

Case of LOD Conversion: Corporation Number

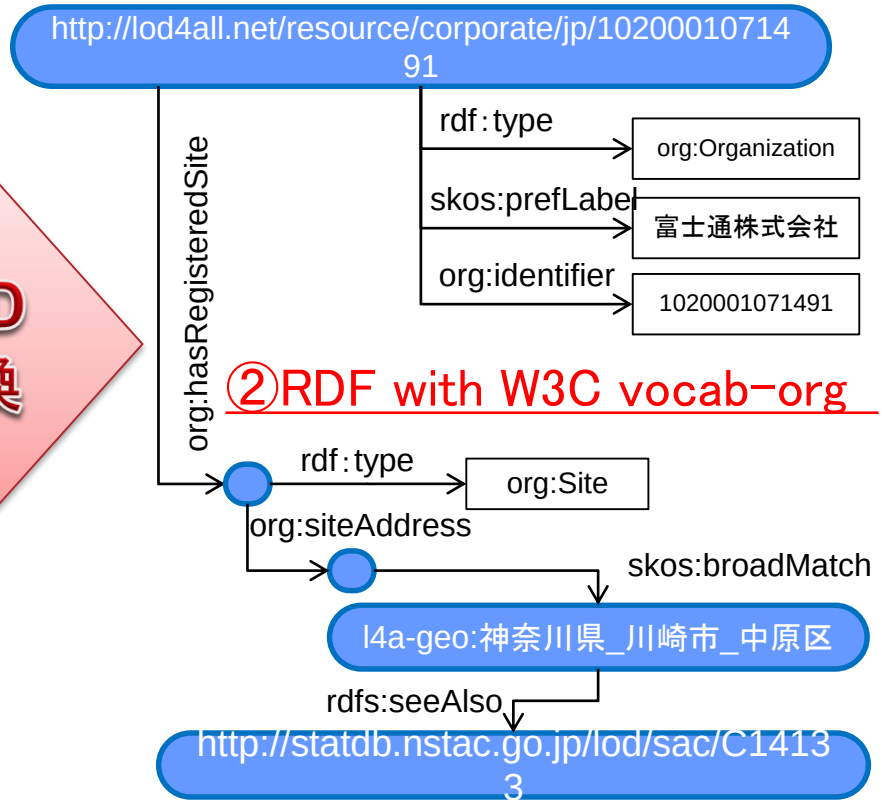
FUJITSU

- National Tax Agency, Japan has published a dataset of corporation numbers covering 4.4 million companies in Japan.
- Fujitsu Laboratories converted it to LOD and re-published in

```
<corporation>
    <sequenceNumber>5203</sequenceNumber>
    <corporateNumber>1020001071491</corporateNumber>
    <process>71</process>
    <correct>0</correct>
    <updateDate>2015-11-06</updateDate>
    <changeDate>2015-10-15</changeDate>
    <name>富士通株式会社</name>
    <nameImageId/>
    <kind>301</kind>
    <prefectureName>神奈川県</prefectureName>
    <cityName>川崎市中原区</cityName>
    <streetNumber>上小田中４丁目１番１号</streetNumber>
    <addressImageId/>
    <prefectureCode>14</prefectureCode>
    <cityCode>133</cityCode>
    <postCode>2110053</postCode>
    <addressOutside/>
    <addressOutsideImageId/>
    <closeDate/>
    <closeCause/>
    <successorCorporateNumber/>
    <changeCause>平成２７年１０月１日埼玉県熊谷市中曽根１３７６富士通
ワイヤレスシステムズ株式会社（9030001085368）を合併</changeCause>
    <assignmentDate>2015-10-05</assignmentDate>
  </corporation>
```

LOD 変換

①URL付与(http IRI参照可)



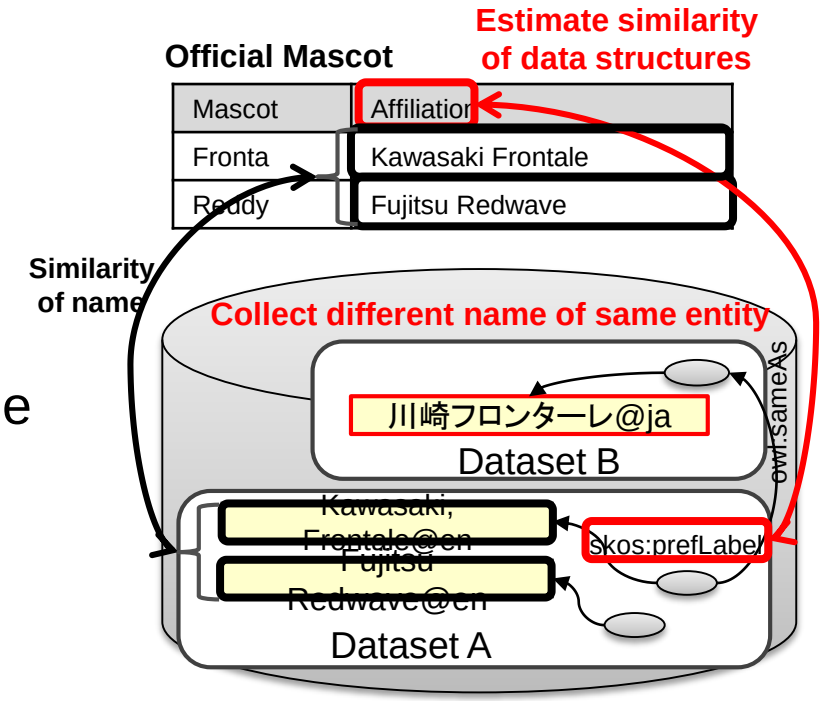
③Linking to other LOD(5★)

Original: XML format(3★)
http://www.houjin-bangou.nta.go.jp/download/

Copyright 2016 FUJITSU LABORATORIES LIMITED

Entity Linking

- Aim
  - Find the same entity (person, organization, place, etc.) from other dataset
  - Linking as “owl:sameAs”
- Issues
  - Differences in data structures and name for same entities
  - Same name but Different entities
- Technology
  - Use similarity in name to measure the similarity of data structures
  - Use the data structures in LOD to collect different names about the same entity
  - Leverages machine learning to judge the entity identity



Chinese-language microblog entity linking evaluation at the NLP&CC2013

Participants	Accuracy
Fujitsu R&D Center	**90.9%**
CAS:中国科学院	88.5%
Tsinghua University:清華大学	87.9%
Harbin Institute of Technology	84.5%

Copyright 2016 FUJITSU LABORATORIES LIMITED

Integrated Corporation Number & LOD datasets

FUJITSU

LOD4ALL

富士通株式会社

URI: h[redacted]urce/corporate/jp/102[redacted]

DBpedia

Corporation Number

DBpedia

DBpedia:代表画像

FUJITSU

拡大

DBpedia:所在地情報(住所)

abstract	中原区(なかはらく)は、川崎 を構成する7区のうちのひと である。人口・世帯数は、川 市で最大となっている。また ッカーJリーグの川崎フロン レの本拠地である等々力陸 競技場や川崎市市民ミュー ムなどのある等々力緑地が

拡大

国土地理院:地図(住所)

法人番号:基本３情報

number	1020001071491
name	富士通株式会社
pref	神奈川県
city	川崎市中原区
street	上小田中4丁目1番1号

拡大

法人番号:履歴情報

件数:2件

date	event	o
2015-10-15	吸収合併	平成27年10月1日埼玉県熊谷市中曽根1376富士通ワイヤレスシステムズ株式会社(9030001085368)を合併
2015-10-15	吸収合併	平成27年10月1日埼玉県熊谷市中曽根1376富士通ワイヤレスシステムズ株

拡大

法人番号:同一名称の法人リスト

件数:0件

id	ad

DBpedia:基本情報

富士通株式会社(ふじつう、英: Fujitsu Limited)は、日本の[redacted]ロニクスメーカーであり、総合ITベンダーでもある。ITサービス[redacted]世界上位。大手電機8社の一角(日立製作所、東芝、三菱電[redacted]ック、ソニー、シャープ、富士通、日本電気)。通信システム、[redacted]テムおよび電子デバイスの製造・販売ならびにそれらに関す[redacted]提供を行っている。

富士通株式会社

富士通株式会社

0020

拡大

DBpedia:グループ会社

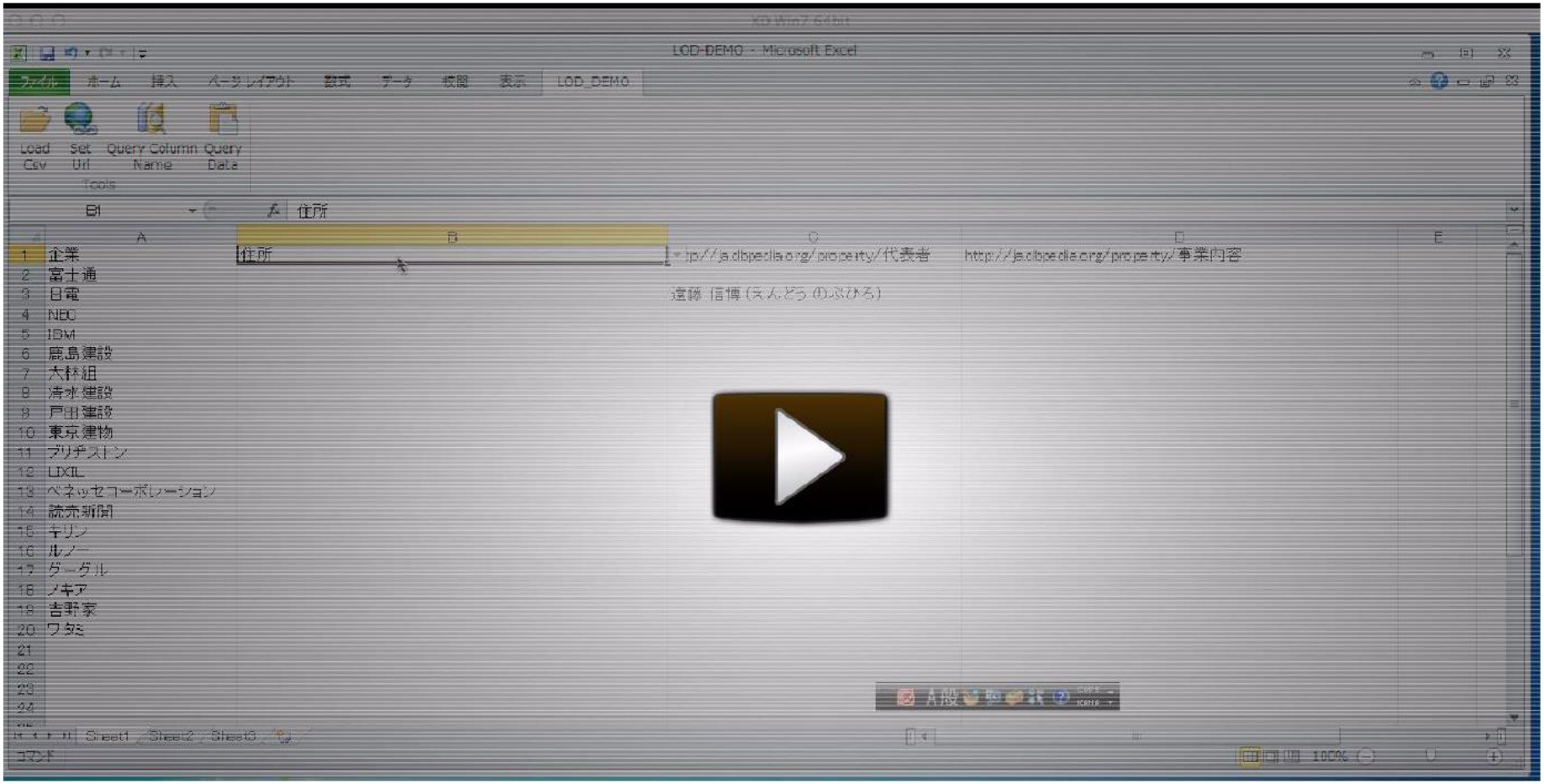
label	g
富士通エフ・オー・エム	http://
富士通アドバンストエンジニアリング	http://
富士通フロンテック	http://
富士通研究所	http://
大興電子通信	http://

拡大

Copyright 2016 FUJITSU LABORATORIES LIMITED

Use cases of Corporation LOD from Excel

- User input company name (富士通, IBM) and attribute name (住所, 社長)
- System automatically fill attribute values.

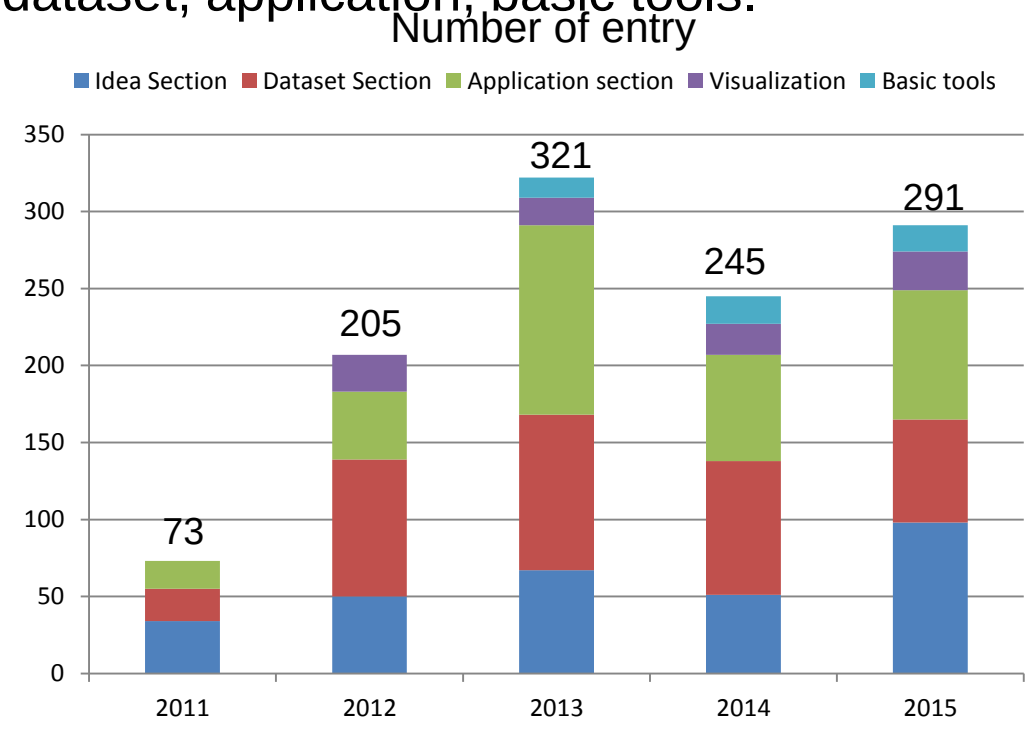


Copyright 2016 FUJITSU LABORATORIES LIMITED

Social Activities: LOD Challenge Japan

- The biggest contest for Open data in Japan
  - hosted by Open Community: Academia, Industry, Schooler, Programmer , etc.
  - Annual Event from Year 2011
    - Entry period from October to January and Awards Ceremony in March
    - Many related events (like Hackathon) during the entry period.
  - Everyone can entry their idea, dataset, application, basic tools.

http://lodc.jp/



Copyright 2016 FUJITSU LABORATORIES LIMITED

Social Activities: LOD Challenge Japan

FUJITSU

Company

Open data (non-RDF)

Motivation

Data Curator

Input

Dev elop

Programmer

LinkData

http://linkdata.org/

Data distribution and Conversion Service

Register

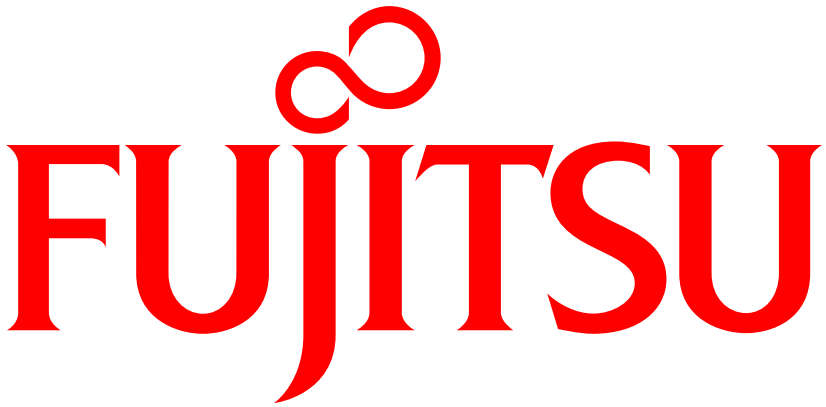
LOD4ALL

http://lod4all.net/

LOD Search & API Service

Field of Open Innovation

Copyright 2016 FUJITSU LABORATORIES LIMITED



shaping tomorrow with you