

Web Of Linked Data (WoLE 2012)
International Semantic Web Conference (ISWC 2012)
Boston, USA, November 11th, 2012

Facetted Browsing of Extracted Fusion Tables Data for Digital Cities

Gianluca Quercini¹, Jochen Setz², Daniel Sonntag², Chantal Reynaud¹

¹ **Université Paris-Sud XI, Laboratoire de Recherche en Informatique (LRI)**

² **DFKI - German Research Centre for Artificial Intelligence**

Overview

- ◆ **Facetted browser** over a repository of RDF data on points of interest (POIs) of digital cities.
 - ◆ Accessing information organized into *facets*.
- ◆ Creation of the repository in a **semi-automatic** way.
- ◆ Data are obtained from **Google Fusion Tables** (GFT)
 - ◆ Annotation of tables.
 - ◆ Data extraction.
 - ◆ RDF conversion.
- ◆ Experimental results.
- ◆ Concluding remarks.

Facetted Browser: Interface

POI types

DIGITAL CITIES

Resource Types

Museum Restaurant

Facets

City

2Adelaide
23Melbourne
1Sydney

Country

346(missing this field)
26Australia
3Canada
1Egypt

1

hasCategory

23(missing this field)
3Hotel-restaurant

AGGREGATED • MAP • COMPACT • FULL

26 DigitalCitiesResource filtered from 2346 originally ([Reset All Filters](#))

sorted by: [hasTopic](#); then by... • grouped as sorted

Data presentation

Art Gallery of New South Wales



Address:
1 Nicholson St Sydney, New South Wales
2011 Australia
DBpedia (City):
<http://dbpedia.org/resource/Sydney>
Categories:

The Wheatsheaf Hotel



Address:
33 George St Adelaide, South Australia 5031
Australia
DBpedia (City):
<http://dbpedia.org/resource/Adelaide>
Categories:
Hotel-restaurant

Earl of Leicester Hotel



Address:
89 Leicester St Adelaide, South Australia
5063 Australia
DBpedia (City):
<http://dbpedia.org/resource/Adelaide>
Categories:
Hotel-restaurant

Biero (The Precinct)



Address:
Wicklow Ln Melbourne, Victoria 3000
Australia

Panama Dining Room & Bar



Address:
237 Smith St Melbourne, Victoria 3066
Australia

Oscar's Alehouse



Address:
4 Bayview Rd Melbourne, Victoria 3160
Australia

Facetted Browser: Map View

POI types →

DIGITAL CITIES

Resource Types

Museum Restaurant

Facets →

City

2 Adelaide
23 Melbourne
1 Sydney

Country

346 (missing this field)
26 Australia
3 Canada
1 Egypt

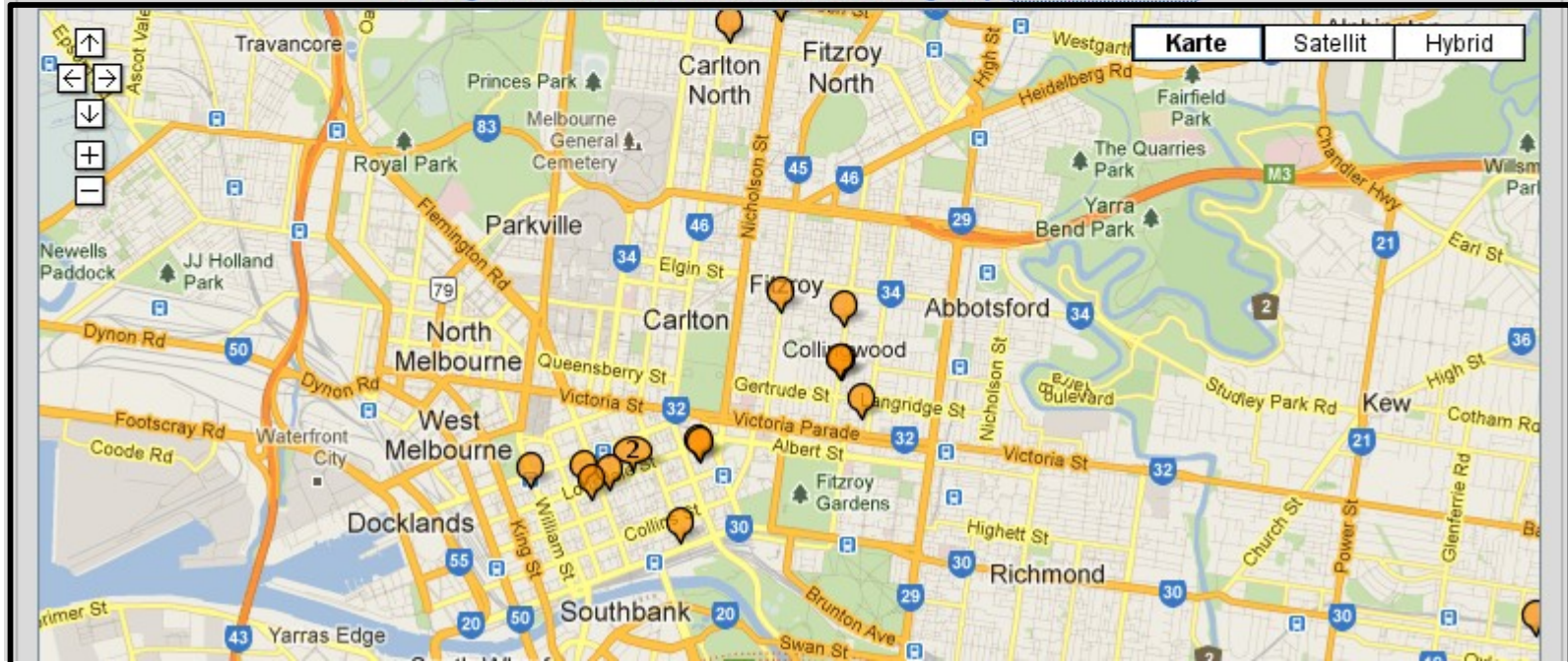
1 hasCategory

23 (missing this field)
3 Hotel-restaurant

AGGREGATED • MAP • COMPACT • FULL

26 DigitalCitiesResource filtered from 2346 originally (Reset All Filters)

Data presentation →



Facetted Browser: Link to the LOD cloud

Art Gallery of New South Wales




Address:
1 Nicholson St Sydney, New South Wales
2011 Australia

DBpedia (City):
<http://dbpedia.org/resource/Sydney>

Categories:

DBpedia (City): <http://dbpedia.org/resource/Adelaide> | <http://dbpedia.org/resource/Adelaide>


Sydney



Data: <http://dbpedia.org/resource/Sydney>
Flickr Links: null (from DBPedia)
Photos provided by Flickr

Categories:
restaurant

Alehouse



Address:
Melbourne, Victoria 3160
Australia


DBpedia (City):
[resource/Melbourne](http://dbpedia.org/resource/Melbourne)

Categories:

Tiger

Sydney

Comment



Sydney is the most populous city in Australia and the state capital of New South Wales. Sydney is classified as an Alpha+ World City indicating its influence in the region and throughout the world. Sydney is located on Australia's south-east coast of the Tasman Sea. As of June 2010, the greater metropolitan area had an approximate population of 4.6 million people. Inhabitants of Sydney are called Sydneysiders, comprising a cosmopolitan and international population.

[First](#) [Comments](#) [External Wikis](#)

Address:
18 Crossley St Melbourne, Victoria 3053
Australia

Address:
14 Cambridge St Melbourne, Victoria 3000
Australia

Repository Creation

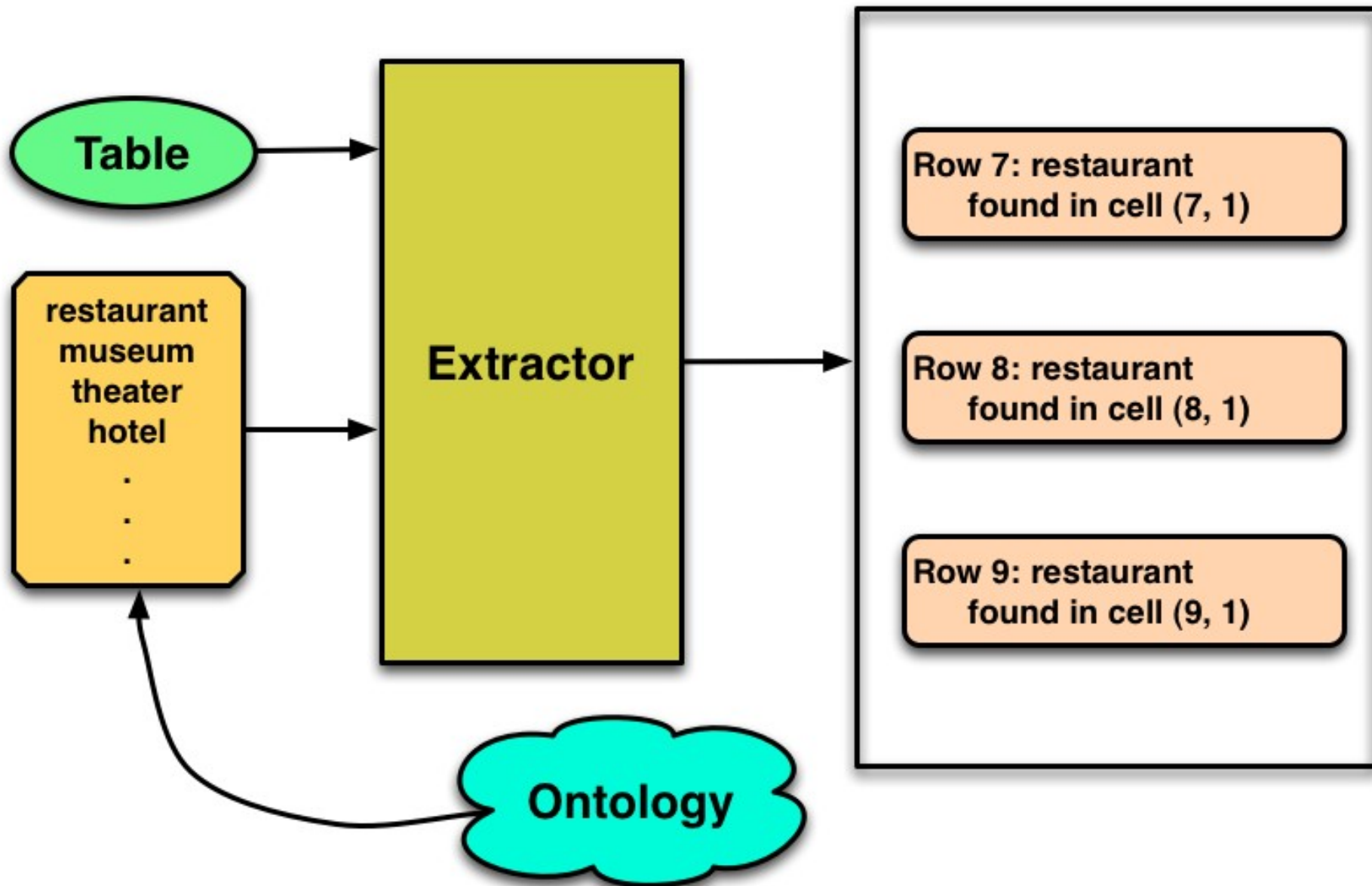
- ◆ **Goal:** obtain data on POIs of a given set of **types** (e.g. "restaurant", "museum", "theatre" ...).
- ◆ Data source: **Google Fusion Tables (GFT)**.
 - ◆ Web service provided by Google for managing and integrating tables.
 - ◆ Contains **thousands** of tables (the exact number is unknown).
 - ◆ Advantages of GFT tables over other tabular data (e.g. HTML tables, spreadsheets...)
 - ◆ **Simple and neat structure.** No column branches into sub-columns.
 - ◆ **Typed columns** (*TEXT, NUMBER, DATE, LOCATION*). Annotating is easier.
 - ◆ **API.** To query tables by using SQL.
- ◆ **Three steps:**
 - ◆ (Manual) creation of an **ontology**, describing major city POIs.
 - ◆ Selection of GFT tables with data on major POIs.
 - ◆ Annotation and extraction.

Repository Creation: Selection of Tables (1)

- ◆ Google indexes the tables in GFT.
- ◆ A search for "restaurant" retrieves all tables in GFT that contain the keyword "restaurant".
- ◆ But tables can be heterogeneous...

	TITLE	DESCRIPTION	ADRESS
Temples	Kankouji Temple	The temple has approximately 600 years history, an...	Uwano 267, Minamiuonuma-shi...
	Untoan Temple	Untoan is a temple of the Soto school of Zen Buddh...	660 Unto, Minamiuonuma, Niigata...
	Bishamon Temple	Fukoji is a "designated cultural asset" by the cit...	Bishamon-do, Urasa Fukoiji Temple grounds ...
Hotels	GOUZOKU PALACE RYUGON	RYUGON IS A TRADITIONAL JAPANESE STYLE HOTEL...	79, SAKADO, MINAMIUONUMA-SHI...
	HOTEL KOJYOKAN	Welcome to the Kojyokan. Our hotel is located in t...	1873, ISHIUCHI, MINAMIUONUMA-SI...
	LODGE MASHU	GET TO THE ISHIUCHI MARUYAMA SKI TRAIL JUST...	Go to Koide using route 17. Pass the bowlings...
Restaurants	Azumaken Restaurant	Hmmm... 	Get on route 291 towards Koide. Turn left at the S...
	Café West Restaurant	Once a year, it can't hurt <img src="images\...	Drive on route 17 towards Koide. The place is on t...
	Early's Restaurant	Oh gee, these burgers! <img src="images\pic5....	2035-2, ISHIUCHI, MINAMIUONUMA-SHI, NIIGATA

Repository Creation: Selection of Tables (2)



Repository Creation: Selection of Tables (3)

TITLE	DESCRIPTION	ADDRESS
Kankouji Temple	The temple has approximately 600 years history, an...	U w a n o 2 6 7 , Minamiuonuma-shi...
Untoan Temple	Untoan is a temple of the Soto school of Zen Buddh...	6 6 0 U n t o , Minamiuonuma, Niigata...
Bishamon Temple	Fukoji is a "designated cultural asset" by the cit...	Bishamon-do, Urasa Fukoji Temple grounds ...
GOUZOKU PALACE RYUGON	RYUGON IS A TRADITIONAL JAPANESE STYLE HOTEL...	7 9 , S A K A D O , MINAMIUONUMA-SHI...
HOTEL KOJYOKAN	Welcome to the Kojyokan. Our hotel is located in t...	1 8 7 3 , I S H I U C H I , MINAMIUONUMA-SL...
LODGE MASHU	GET TO THE ISHIUCHI MARUYAMA SKI TRAIL JUST...	2035-2, ISHIUCHI, MINAMIUONUMA-SHI, NIIGATA
Azumaken Restaurant	H m m m ... < b r / > < i m g src="images\pic1.jpg"/>	Get on route 291 towards Koide. Turn left at the S...
Café West Restaurant	Once a year, it can't hurt 17. Pass the bowlings...	



Get snippets

[Hotel Kojyokan, Ishiuchi, Shiozawa, Niigata Prefecture](#)

Welcome to the **Kojyokan** website. Our **hotel** is located in the heart of the beautiful Snow Country area of Shiozawa town in ...

www.mars.jstar.ne.jp/~kojyokan/e

[Minamiuonuma Accommodation - Hotel Kojyokan, Minamiuonuma, Niigata ...](#)

Hotel Kojyokan - ski resort accommodation in Minamiuonuma, Niigata, Japan

www.snowjapan.com/e/hotel/viewhotel.php?hotid=50

[HOTEL KOJYOKAN - niigata](#)

HOTEL KOJYOKAN Getting to **Hotel Kojyokan** is easy. Located near the Ishiuchi Maruyama resort and close by to the famous resorts in Yuzawa Town, we can be reached in under 2 ...

web.travel.rakuten.co.jp/portal/my/info_page_e.Eng?f_no=16820

[HOTEL KOJYOKAN - Rakuten Travel](#)

Japan Niigata Uonuma, Tokamachi, Tsunan, Muikamachi ,**HOTEL KOJYOKAN**,

travel.rakuten.com/hotelinfo/20/16820

[Hotel Kojyokan - ACCESS: How to Find Us](#)

Getting to **Hotel Kojyokan** is easy. Located near the Ishiuchi Maruyama resort and close by to the famous resorts in Yuzawa Town, we can ...

www.mars.jstar.ne.jp/~kojyokan/e/access.html

[SnowJapan Route Finder - Find directions to Hotel Kojyokan](#)

Snow Japan - the ultimate guide to winter sports in Japan

www.snowjapan.com/e/route-finder/accommodation-route-finder.php?hotid=50

[Uonuma, Tokamachi, Tsunan, Muikamachi hotel - Rakuten Travel](#)

Getting to **Hotel Kojyokan** is easy. Located near the Ishiuchi Maruyama resort and close by to the famous resorts in Yuzawa Town, we can be reached in under 2 hours from central ...

travel.rakuten.co.jp/group/tiku_en/03japanniigataminami.html

Text classifier
(SVM)

Hotel

Repository Creation : Annotation and Extraction

- ◆ Identification of the columns that correspond to properties of a POI.
 - ◆ Set of heuristics, based on column **headers** and **types**.
- ◆ Extraction of the data in the selected rows and columns.
- ◆ Insertion of the extracted data in the ontology.

Experimental Results: Settings

- ◆ Evaluation on the selection of tables.
- ◆ Extracted data from **40** GFT tables on:
 - ◆ 1500 restaurants
 - ◆ 500 museums
 - ◆ 160 theatres
 - ◆ 67 hotels
 - ◆ 109 schools
- ◆ Used two text classifiers: **SVM** and **Bayes**
- ◆ Evaluation metrics:
 - ◆ **Precision** : number of cells correctly identified over the number of cells identified.
 - ◆ **Recall** : number of cells correctly identified over the total number of cells.
 - ◆ **F-measure** : harmonic mean of precision and recall

Experimental Results

Restaurants

Method	P	R	F
SVM	0.89	0.69	0.78
Bayes	0.59	0.80	0.68
TIN	0.64	0.1	0.18
TIS	1.0	0.19	0.32

Museums

Method	P	R	F
SVM	0.83	0.82	0.83
Bayes	0.45	0.93	0.61
TIN	1.0	0.61	0.76
TIS	1.0	0.59	0.75

Theatres

Method	P	R	F
SVM	0.83	0.76	0.80
Bayes	0.34	0.89	0.5
TIN	1.0	0.18	0.31
TIS	1.0	0.29	0.45

Hotels

Method	P	R	F
SVM	0.74	0.89	0.81
Bayes	0.23	0.92	0.37
TIN	0.15	0.07	0.10
TIS	0.96	0.60	0.74

Schools

Method	P	R	F
SVM	0.96	0.91	0.94
Bayes	0.75	0.96	0.85
TIN	0.78	0.56	0.66
TIS	0.98	0.88	0.93

Concluding Remarks

- ◆ Facetted Browser
 - ◆ Tighter integration with the LOD.
 - ◆ Better user interface
 - ◆ More focus on the map interface.
- ◆ Data extraction
 - ◆ Focus on scalability
 - ◆ Extraction from other tabular data (spreadsheets, HTML tables/lists)
 - ◆ Extraction from unstructured text.