

DBpedia Overview

Jens Lehmann

AKSW Research Group
Institute of Computer Science

UNIVERSITÄT LEIPZIG

Transforming Wikipedia into a Knowledge base

- ☺ Wikipedia is the **7th most popular website** (according to Alexa.com)
- ☺ Maybe the finest example of truly **collaboratively created content** (>8M articles, >200 languages, >300.000 authors)
- ☺ Covers many possible topics and domains, articles are a result of a “**community consensus**”
- ⊖ Many **inconsistencies** can be found on different pages/language versions
- ⊖ **Not very well integrated** with other data sources
- ⊖ **Lacks structured representations** of content which facilitate querying and search



Transforming Wikipedia into a Knowledge base




- Simple Questions – hard to answer:
 - What have Innsbruck and Leipzig in common?
 - Who are mayors of central European towns elevated more than 1000m?
 - Which films are longer than 4 hours and had a budget of less than \$1 Million?
- The information required to answer these is contained in Wikipedia!
- How can we reveal structure and semantics of Wikipedia content?

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
 - other language versions
 - other Wikipedia pages
 - To the Web
 - Redirects
 - Disambiguations

Busan Metropolitan City



A view of the **Geumjeong** district in Busan

Korean name	
Hangul	부산 광역시
Hanja	釜山廣域市
Revised Romanization	Busan
McCune-Reischauer	Pusan
Kwangyŏkshi	
Short name	
Hangul	부산
Hanja	釜山
Revised Romanization	Busan
McCune-Reischauer	Pusan
Statistics	
Area	763.46 km ² (295 sq mi)
Population (2006)	3,635,389 ^[1]
Population density	4,762/km ² (12,334/sq mi)
Government	
Mayor	Hur Nam-sik
Administrative divisions	
	15 wards (Gu), 1 county (Gun)
Region	
	Yeongnam
Dialect	
	Gyeongsang

14,544 people have donated (show more)

Busan

From Wikipedia, the free encyclopedia

"Busan" redirects here. For the Vedic Hindu god, see **Pusan**.

Busan Metropolitan City, also known as **Pusan**^[a] is the largest port city in the Republic of Korea. With a population of about 3.6 million, Busan is also South Korea's second largest metropolis, after Seoul. The metropolis built-up areas of the city are situated in a number of narrow valleys between the Nakdong River and Suyeong River, with mountains separating some of the various districts.

Busan was the host of the 2002 Asian Games and APEC 2005. It was also one of the host cities of the 2002 FIFA World Cup, and is the center for international competition in Korea. Administratively, it is designated as a Metropolitan City. On October 14, 2005, the city officially announced its bid to host the 2020 Summer Olympics Games.

History

Gaehŏnŏng-gak existed in the second and third centuries as a subdivision of Jinhan. It was absorbed by Silla and renamed Gaecheonŏng. The word Gaecheon means rugged mountain, probably referring to **Haengyeongsan**, located at the center of the city. The grave goods excavated from mounded burials at **Bolcheon-dong** indicate that a complex chieftom ruled by powerful individuals was present in the Busan area just as the Three Kingdoms of Korea were forming, c. A.D. 300-400. The mounded burials of **Bolcheon-dong** were built along the top of a ridge that overlooks a wide area that make up parts of modern-day Dongnae-gu and Yeongju-gu. Archaeologists excavated more than 250 iron weapons and ingots from Burial No. 39, a wooden chamber tomb at **Bolcheon-dong**.

In 757, Gaecheonŏng was again renamed Dongnae, which it is still called. From the beginning of the fifteenth century, the Korean government designated Busan a trading port with the Japanese and allowed their settlement. Other Japanese settlements in Ulsan and Jinhae diminished later, but the Busan settlement, called **Waegwan** at the time, continued until Japan invaded Korea in 1592. After the war, diplomatic relations with the new shogunate in Japan were established in 1607, and Busan Waegwan was permitted to be reconstructed. The Japanese settlement, though relocated into **Choryang** later, continued to exist until Korea was exposed to modern diplomacy in 1876. In 1876, Busan became the first international port in Korea.

During the Japanese occupation, Busan developed into a hub trading port with Japan. Busan was the only city in Korea to adopt the steam tramway before electrification was introduced in 1924. And Busan is the only city in Korea, with the exemption of Daejeu, that was never taken by the North Korean Communists during the Korean War. As a result the city was a refuge camp site for Koreans during the war.

Busan was one of the few areas in Korea that remained under the control of South Korean troops established a defensive perimeter around the city known as the **Pusan Perimeter** in the summer and autumn of 1950. Since then, like Seoul, the city has been a self-governing metropolis and has built a strong urban character.

Administrative divisions

In 1957 Busan adopted a division system with the creation of 6 gu: Busanjin District, Dong-gu, Dongnae-gu, Jung-gu, Seogu, and Yeongju-gu. Today, Busan is divided into 15 gu (districts) and 1 gun (county).

- Buk-gu (북구, 北區)
- Busanjin-gu (부산진구, 釜山鎮區)
- Dono-gu (동구, 東區)

Economy

Transportation and shipping are among the most high profile aspects of the local economy. Since 1978, Busan has opened three container ports including Jangsaengpo, Shinusanda and Gammam. Busan is renowned as one of the world's largest ports and can handle up to 6.4 million TEU shipping containers per year.

The **Busan-Jinhae Free Economic Zone Authority**, one of two such administrations (the other in the harbor of Incheon), was created to continue the tradition of Busan's status as an international trading centre. The port attracts ships from all over the globe and aspires to become a regional financial centre. Currently, Busan is ranked internationally as the third largest seaport in terms of cargo volume and efficiency by the **ALFA**.

Shopping and Commerce

Commercial areas are dispersed throughout the city in major intersections and adjacent to university campuses, but the two largest central business districts in Busan are Seomyeon and Gwangbok-dong/Nampo-dong. There are also four substantial shopping areas of note: Seomyeon, Gwangbok-dong, Busan Dae Haeap in Jangjeon-dong and Haeundae.

Seomyeon is the crossroads of Busan. The local subway station serves two lines and is one of the busiest in the city. The local head offices of Korean and international banks are located in Seomyeon. It is recognized as the ascendant shopping and entertainment districts with many cafes, bars, restaurants, department stores, shops, and offices. Directly adjacent to Seomyeon is **Bujeon Market**, the largest traditional market in the city.

The **Gwangbok-dong**, **Nampo-dong**, and **Jung-gu** -dong areas form the old central business district with many offices, cafes, bars, shops and restaurants. Some of the restaurants in this district are locally famous with family recipes passed on through the generations. **Jajalchi Market** (near part of the very active port) is an area of narrow streets, stalls and is well known for its fish market. The **Gulje Market** is also located nearby. **Jungang-dong** is the home of many international law offices, the old Immigration Office, and the international ferry terminal serving Japanese routes. **Lotte World II** is currently under construction along the road between Jangjeong-dong 7-Che and 8-Che. When completed, **Lotte World II** will house a hotel, department store, entertainment complex, and a skyscraper that will be one of the tallest buildings in the world.

Sister cities

Busan shares the title of sister city with several coastal cities or provinces across the world^[11]

- Kaohsiung, Taiwan (1978)
- Los Angeles, USA (1978)
- Shimonoseki, Japan (1978)
- Barcelona, Spain (1983)
- Rio de Janeiro, Brazil (1985)

See also

- List of Korea-related topics
- List of cities in South Korea
- Centum City
- Pusan International Film Festival
- Gwangsan District

External links

- Busan Metropolitan Government
- Port of Busan
- Tourist information
- GaBjip (Busan page on Busan)
- Hotels in Busan
- Major Busan
- Los Angeles baseball team
- Pusan International Film Festival
- Pusan National University

Regions and administrative divisions of South Korea

Regions and administrative divisions of South Korea

Categories: All articles with unsourced statements | Articles with unsourced statements since August 2007 | Busan | Cities in South Korea | Coastal cities

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations

2011/05/04

Semantic Wikis and the Web of Data - Jens Lehmann

in other languages

- Български
- Dansk
- Deutsch
- Eesti
- Español
- Esperanto
- Français
- Galego
- 古文 / 文言文
- 한국어
- Հայերեն
- Ido
- Bahasa Indonesia
- Italiano
- Latviešu
- Lietuvių
- मराठी
- Nederlands
- 日本語
- Norsk (bokmål)
- Norsk (nynorsk)
- Polski
- Português
- Română
- Русский
- Simple English
- Suomi
- Svenska
- ไทย
- Tiếng Việt
- Türkçe
- 中文

Administrative divisions

- Busan-gu (부산진구, 釜山鎭區)
- Dong-gu (동구, 東區)
- Jung-gu (중구, 中區)
- Sogu (서구, 西區)
- Yeongdo-gu (영도구, 嶺島區)

Sister cities

- Kaohsiung, Taiwan (1966)
- Los Angeles, USA (1967)
- Shimizu, Japan (1976)
- Barcelona, Spain (1983)
- Rio de Janeiro, Brazil (1985)

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- **Categories**
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations

Categories: [Busan](#) | [Cities in South Korea](#) | [Coastal cities](#)

The screenshot shows the Wikipedia article for Busan. At the top, the title "Busan" is displayed. Below it is the abstract: "Busan Metropolitan City, also known as Pusan is the largest port city in the Republic of Korea. With a population of 3.66 million, Busan is also South Korea's second largest metropolis, after Seoul. The most densely built-up areas of the city are situated in a number of narrow valleys between the Nakdong River and Suzyong River, with mountains separating some of the various districts." The infobox on the right contains the Korean name "Busan Metropolitan City" and a map. The "Categories" section is highlighted with a red box and contains the text: "Categories: [Busan](#) | [Cities in South Korea](#) | [Coastal cities](#)". Other sections visible include "History", "Administrative divisions", "Economy", "Sister cities", and "See also".

Infobox Templates

```

{{Infobox Korean settlement
| title           = Busan Metropolitan City
| img             = Busan.jpg
| imgcaption     = A view of the [[Geumjeong]] district in Busan
| hangul         = 부산 광역시
| . . .
| area_km2       = 763.46
| pop            = 3635389
| popyear        = 2006
| mayor          = Hur Nam-sik
| divs           = 15 wards (Gu), 1 county (Gun)
| region         = [[Yeongnam]]
| dialect        = [[Gyeongsang]]
}}

```

Wikitext-Syntax



RDF representation

```

dbp:Busan      dbp:title      "Busan Metropolitan City"
dbp:Busan      dbp:hangul    "부산 광역시" @Hang
dbp:Busan      dbp:area_km2 "763.46"^^xsd:float
dbp:Busan      dbp:pop      "3635389"^^xsd:int
dbp:Busan      dbp:region   dbp:Yeongnam
dbp:Busan      dbp:dialect  dbp:Gyeongsang
. . .

```

Busan Metropolitan City



A view of the [Geumjeong](#) district in Busan

Korean name

Hangul	부산 광역시
Hanja	釜山廣域市
Revised Romanization	Busan Gwangyeoksi
McCune-Reischauer	Pusan Kwangyŏkshi

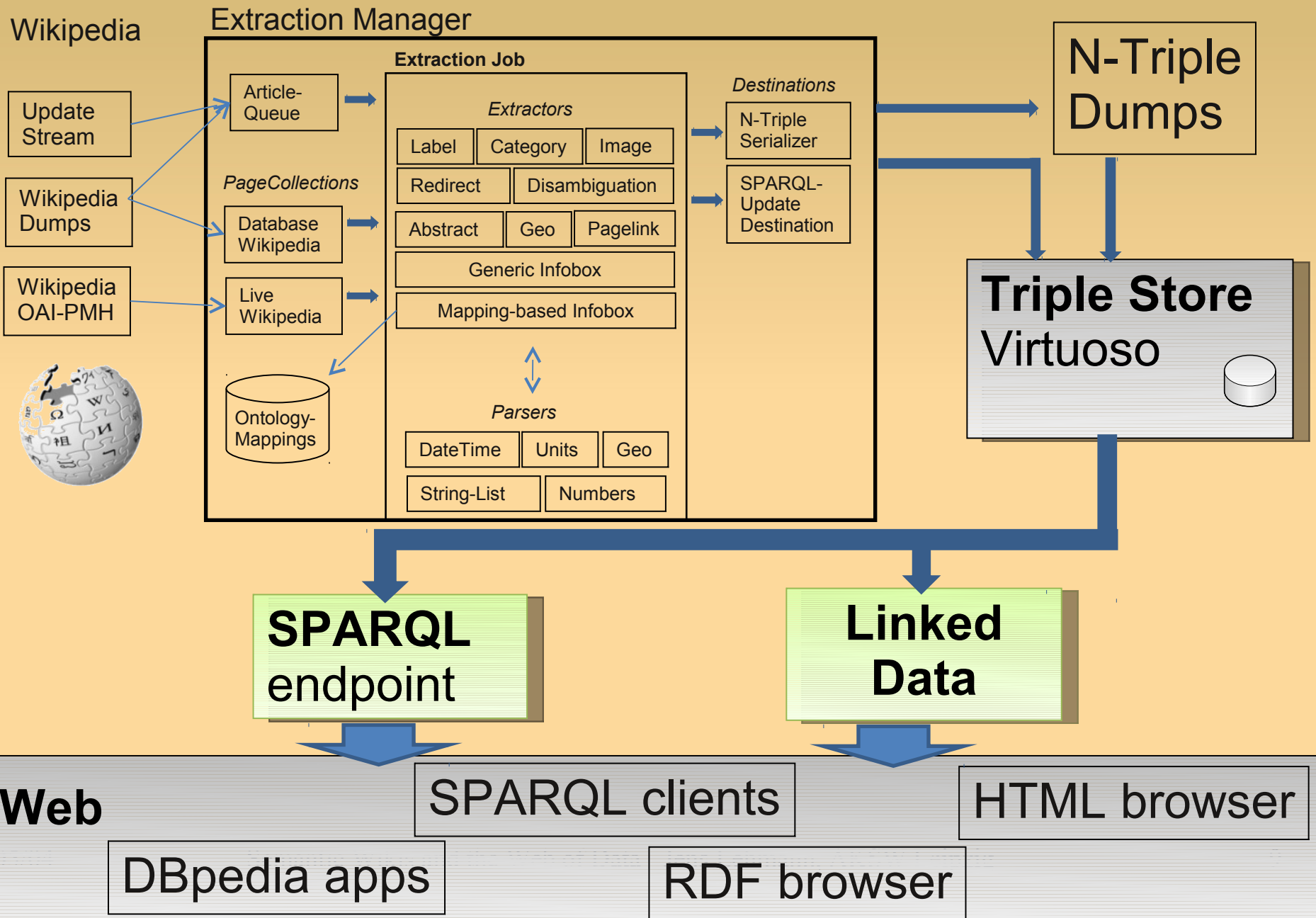
Short name

Hangul	부산
Hanja	釜山
Revised Romanization	Busan
McCune-Reischauer	Pusan

Statistics

Area	763.46 km ² (295 sq mi)
Population (2006)	3,635,389 ^[1]
Population density	4,762/km ² (12,334/sq mi)
Government	Metropolitan City
Mayor	Hur Nam-sik
Administrative divisions	15 wards (Gu), 1 county (Gun)
Region	Yeongnam
Dialect	Gyeongsang

DBpedia Architecture



Results

- Extraction in **97 languages**
- DBpedia describes **3.5 million things**, including at least 364,000 persons, 462,000 places, 99,000 music albums, 54,000 films, 148,000 organisations
- **672 million RDF triples**
- Manual tests reveal: **~87% accuracy**, (9% redundant information, 2% irrelevant, 1% errors)
- Very large **multi-domain** ontology

Dataset (En only)	Triples
Titles	2.7M
Abstracts	7.6M
External Links	3.2M
Categories	7.3M
Infoboxes (generic)	26.0M
Infoboxes (mapped)	7.0M
Yago Classes	2M
Geo-coordinates	450k
Properties	66k
Mapping to Flickr, DBLP, Eurostat, CIA-Factbook, Musicbrainz, Project Gutenberg, US Census, ...	2.5M
Mapping to OpenCyc	45k

Hierarchies

- **DBpedia Ontology Schema:**
 - manually created for DBpedia (infoboxes)
 - 275 classes + 1335 properties; 20mio triples
- **YAGO:**
 - large hierarchy linking Wikipedia leaf categories to WordNet
 - 250,000 classes
- **UMBEL** (Upper Mapping and Binding Exchange Layer):
 - 20000 classes derived from OpenCyc
- **Wikipedia Categories:**
 - Not a class hierarchy (e.g. cycles), represented using SKOS
 - 415,000+ categories

DBpedia SPARQL Endpoint

- <http://dbpedia.org/sparql>
- hosted on a **OpenLink Virtuoso** server
- can answer SPARQL queries like
 - Give me all Sitcoms that are set in NYC?
 - All tennis players from Moscow?
 - All films by Quentin Tarentino?
 - All German musicians that were born in Berlin in the 19th century?
 - All soccer players with tricot number 11, playing for a club having a stadium with over 40,000 seats and is born in a country with over 10 million inhabitants?

DBpedia SPARQL Endpoint

```
SELECT ?name ?birth ?description ?person WHERE {  
  ?person dbp:birthPlace dbp:Berlin .  
  ?person skos:subject dbp:Cat:German_musicians .  
  ?person dbp:birth ?birth .  
  ?person foaf:name ?name .  
  ?person rdfs:comment ?description .  
  FILTER (LANG(?description) = 'en') .  
} ORDER BY ?name
```



name	birth	description	person
"Moser, Edda"@de	"1938-10-27"^^xsd:date	"The German soprano Edda Moser was born on October 27, 1938 in Berlin, Germany. She is the daughter of the musicologist Hans Joachim Moser."@en	:Edda_Moser ↗
"Möbius, Ralph Christian"@de	"1950-01-09"^^xsd:date	"Rio Reiser (January 9, 1950 - August 20, 1996), was a German rock musician and singer of the famous rock group Ton Steine Scherben. He was born Ralph Christian Möbius in Berlin and died at the age of 46 in the little German town of Fresenhagen. Rio Reiser was politically active during his whole life. In the early 70ies he participated in the squatter scene, for which he wrote the famous Rauchhaussong."@en	:Rio_Reiser ↗
"Straube, Karl"@de	"1873-01-06"^^xsd:date	"Montgomery Rufus Karl/Carl Siegfried Straube (January 6, 1873, Berlin - April 27, 1950, Leipzig) was a German church musician , organist, and choral conductor, famous above all for championing the abundant organ music of Max Reger."@en	:Karl_Straube ↗
"Tricht, Käte van"@de	"1909-10-22"^^xsd:date	"Käte van Tricht (October 22, 1909–July 13, 1996), was a German organist, pianist, harpsichordist, and pedagogue."@en	:K%C3%A4te_van_Tricht ↗
"Urlaub, Farin"@de	"1963-10-27"^^xsd:date	"Jan Ulrich Max Vetter, better known as Farin Urlaub (like German "Fahr in Urlaub!" ("Go on holiday!"), after his love of travelling) was born on October 27, 1963 in what was then West Berlin, Germany. He is best known as the guitarist/vocalist for the German punk rock band Die Ärzte."@en	:Farin_Urlaub ↗
"Voormann, Klaus"@de	"1938-04-29"^^xsd:date	"Klaus Voormann (born 29 April 1938) is a German artist, musician, and record producer who was associated with the early days of The Beatles in Hamburg and later designed the cover of their album Revolver."@en	:Klaus_Voormann ↗

DBpedia Applications

- **DBpedia Mobile**: location aware mobile client for DBpedia
 - Uses current location and DBpedia to display map
 - Can navigate into other knowledge bases
- **DBpedia Query Builder**: user front end for building queries
- **DBpedia Relationship Finder** finds relation between two objects in DBpedia



DBpedia Applications (3rd party)

- **Muddy Boots (BBC)**: Annotate actors in **BBC News** with DBpedia identifiers
- **Open Calais (Reuters)**: **named entity recognition**; entities are connected via owl:sameAs to DBpedia, Freebase, Geonames
- **Faviki**: Social Bookmarking Tool uses DBpedia in backend to **group tags** etc. and **multi-language support**
- **Topbraid Composer**: ontology editor, which **links entities** to DBpedia based on their labels

The End

Thanks for your attention!