



Framework for location-aware search engine

Pasi Fränti

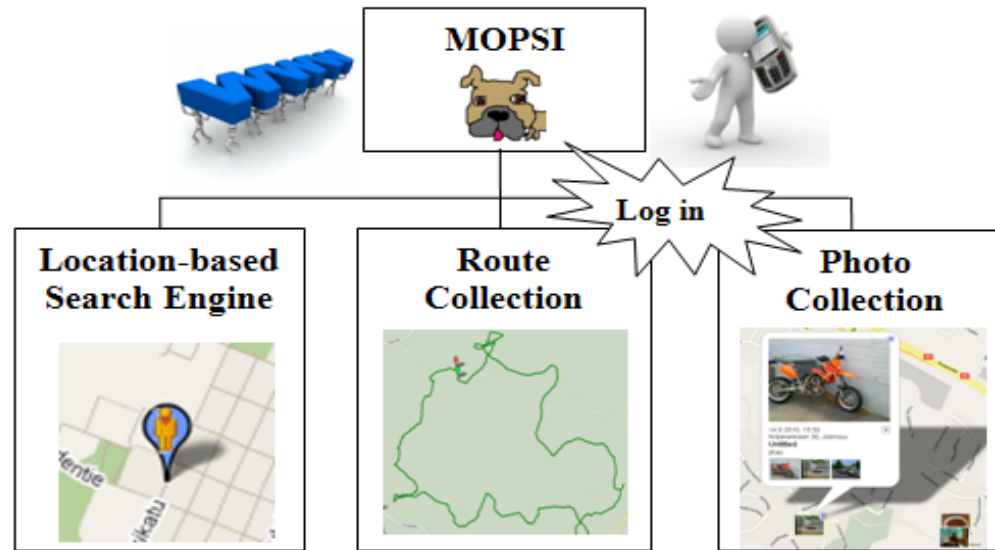
17.1.2019

A. Tabarcea, N. Gali and P. Fränti, "Framework for location-aware search engine", *Journal of Location Based Services*, 11 (1), 50-74, November 2017.

Mopsi



Mopsi overview



Recommendation System



- Service (bus, friend)
- Text (search query, photo description)

Web Content Mining



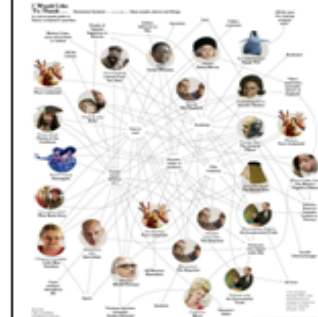
- Meta searching
- Service title detection
- Document processing

Route Pattern



- Route reduction
- Route segmentation
- Activity area

Social network



- Facebook

Location-based Game



- Orienteering
- Killer-game

Data collection in Mopsi

Other users:



MOPSI
webpage



Service
directory

Data collector:



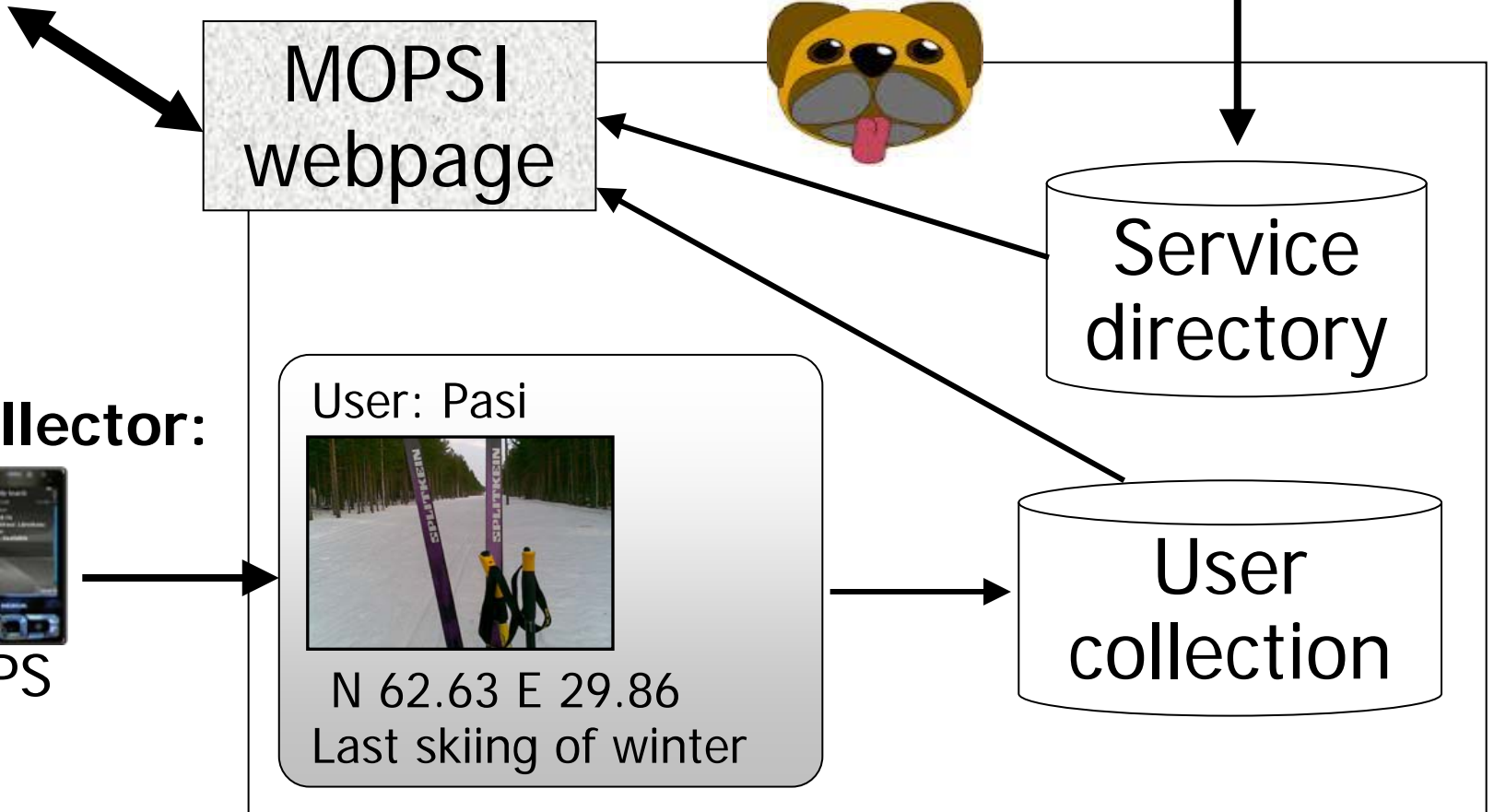
GPS

User: Pasi



N 62.63 E 29.86
Last skiing of winter

User
collection



Four aspects of relevance

P. Fränti, J. Chen, A. Tabarcea

Four aspects of relevance in location-based media: content, time, location and network"

Int. Conf. on Web Information Systems & Technologies (WEBIST), 2011



1. Content

- Text description
- Keywords (tags)



2. Time

- Recency of data
- Season (not relevant in July)



3. Location

- Distance to user

4. User and his network

- User profile
- Social network



User: Pasi



Last skiing of winter

Date: 4.4.2010

Location: N 62.63 E 29.86


Arppentie 5, Joensuu

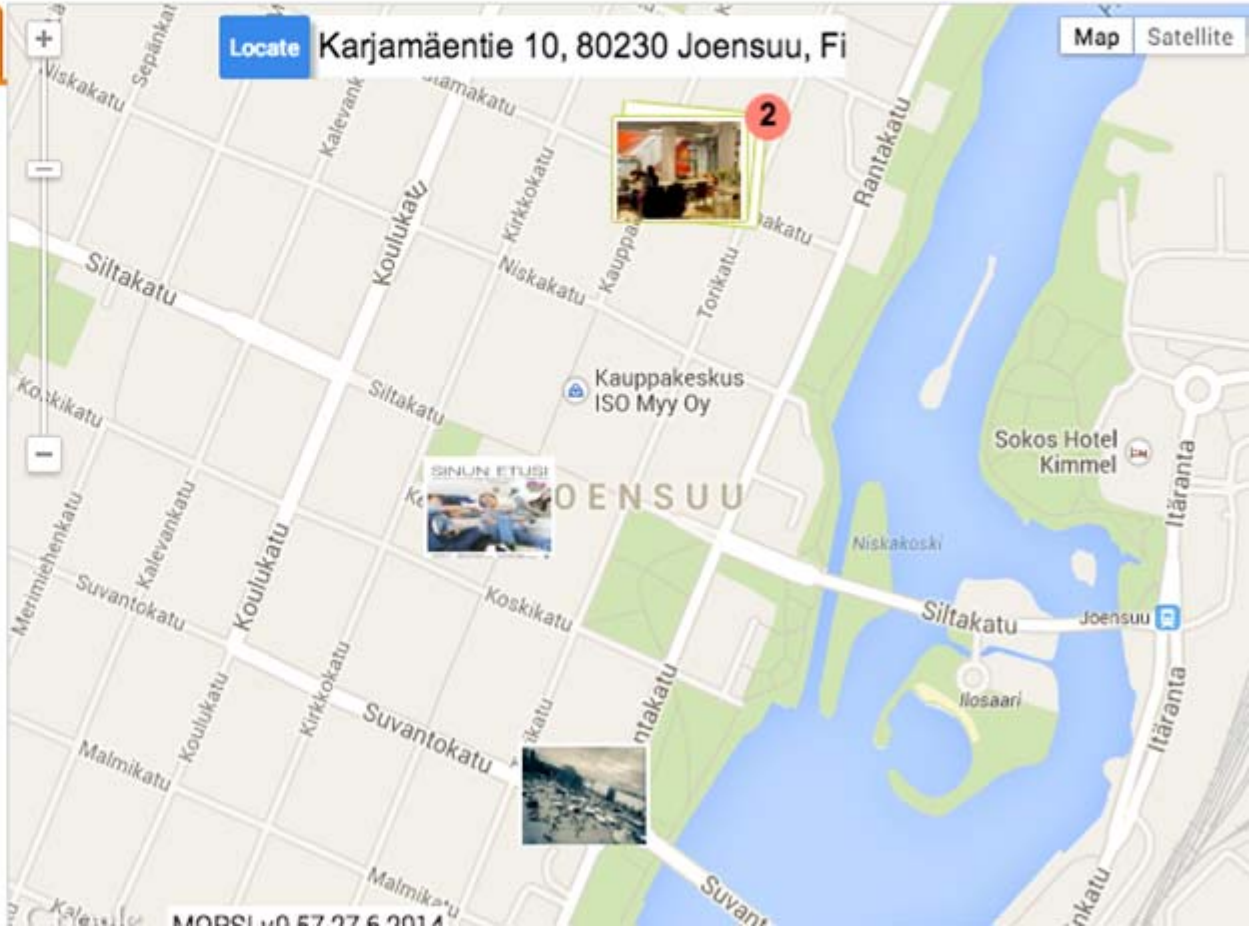
Mopsi search

Search

   Mopsi service  Photo  Web search

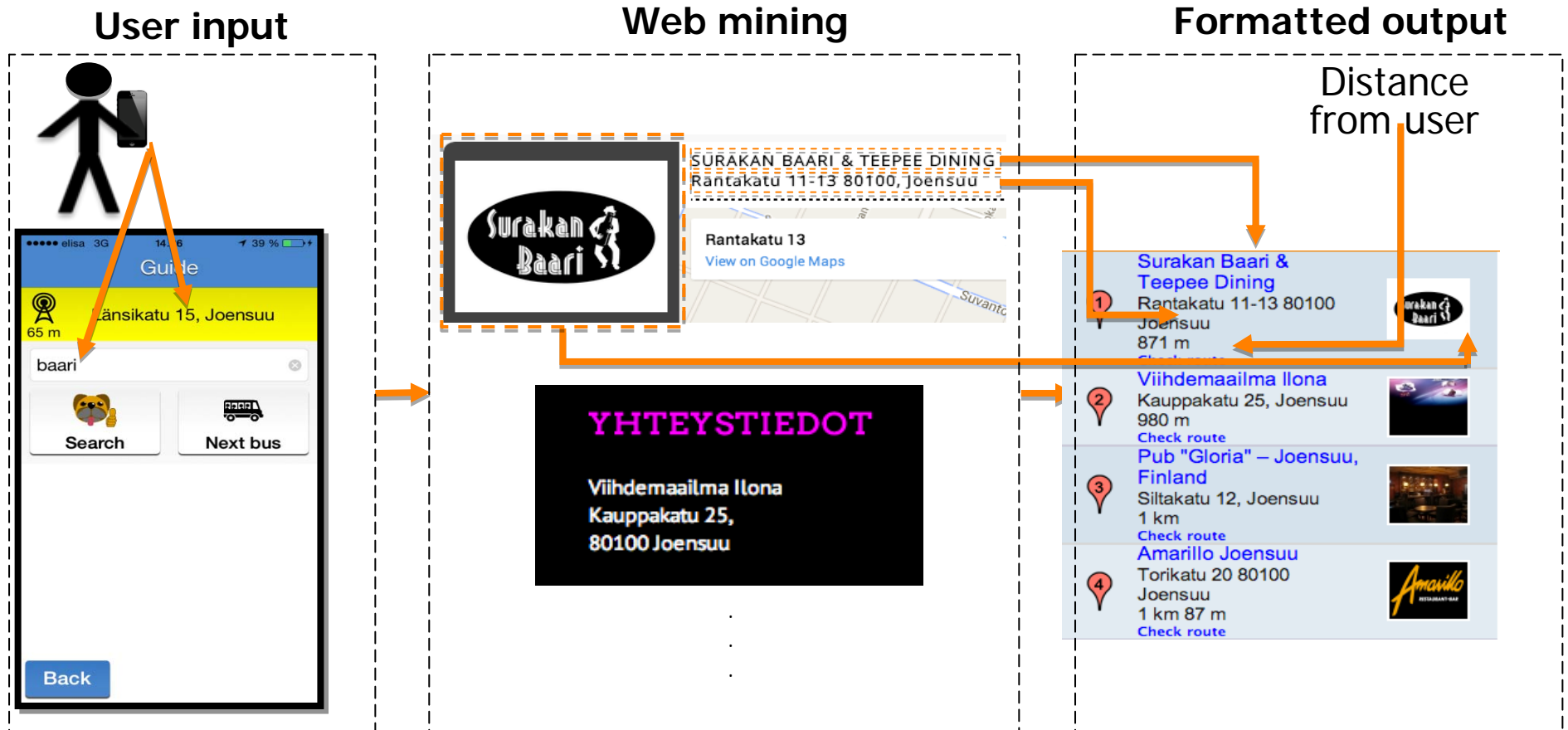


- 1** [Surakan Baari](#)
surakanbaari.fi*
Rantakatu 11-13 80100
Joensuu
3 km 213 m
[Check route](#)
[Upgrade](#) 
- 2** [Bar Play Joensuu S-kanava*](#)
Kauppakatu 23 80100
Joensuu
3 km 494 m
[Check route](#)
[Upgrade](#) 
- 3** [Super Smoothie Joensuu Cafe & Salad Bar smoothie bar*](#)
Torikatu 31 Joensuu
3 km 580 m
[Check route](#)
[Upgrade](#) 
- 4** [Jet Set Sport Bar*](#)
Kauppakatu 35 80100
Joensuu
3 km 678 m
[Check route](#)
[Upgrade](#) 



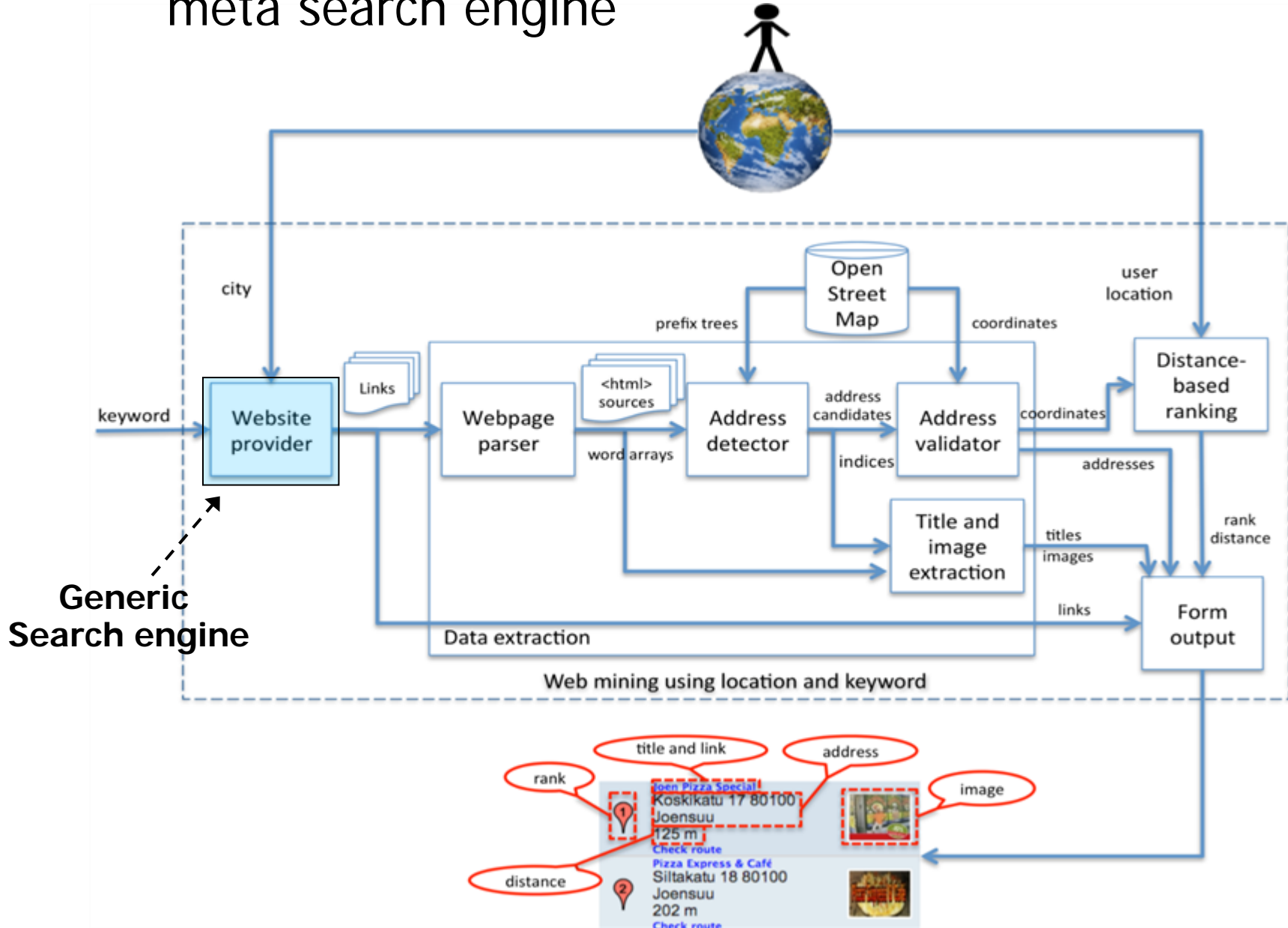
General workflow

meta search engine



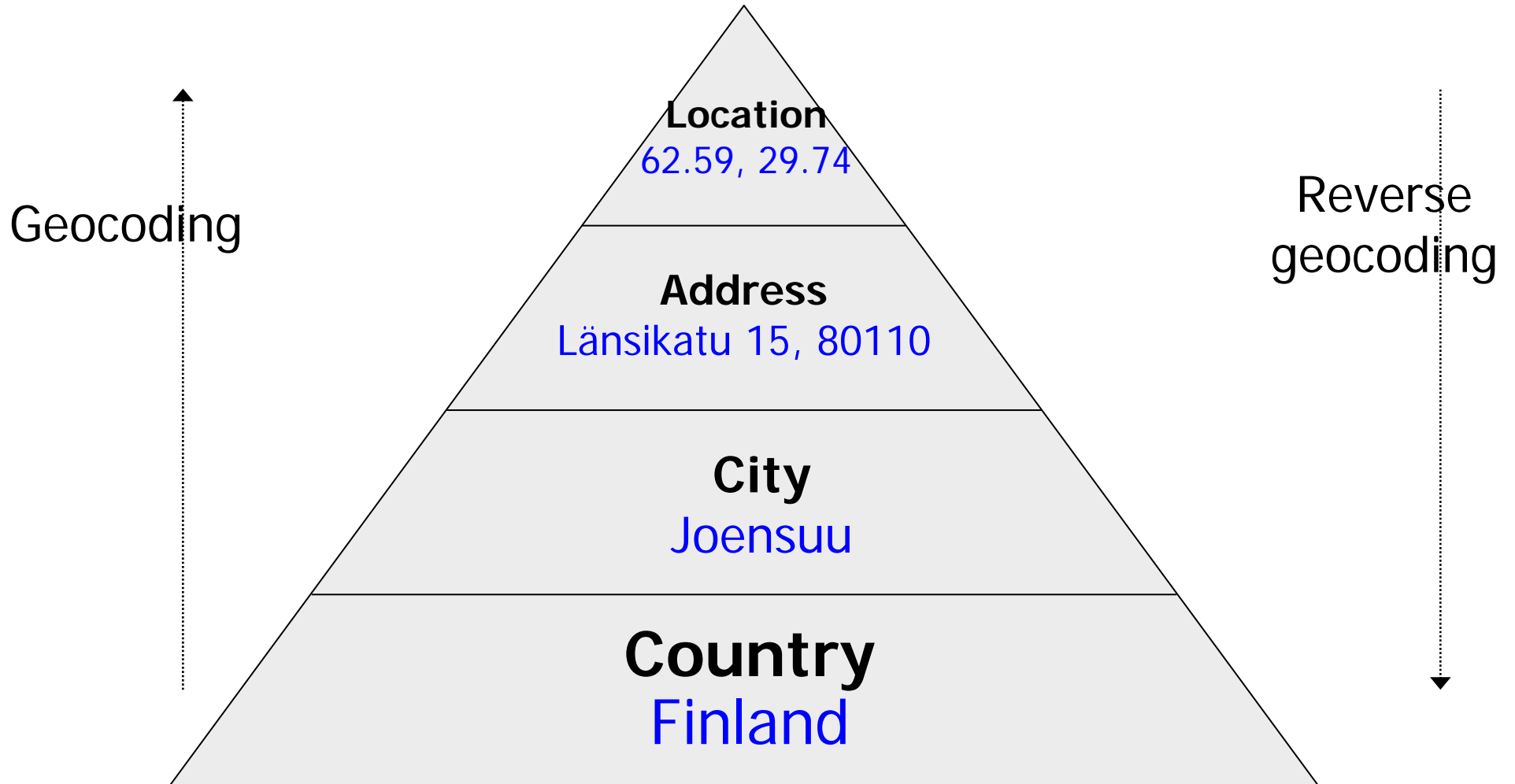
System architecture

meta search engine

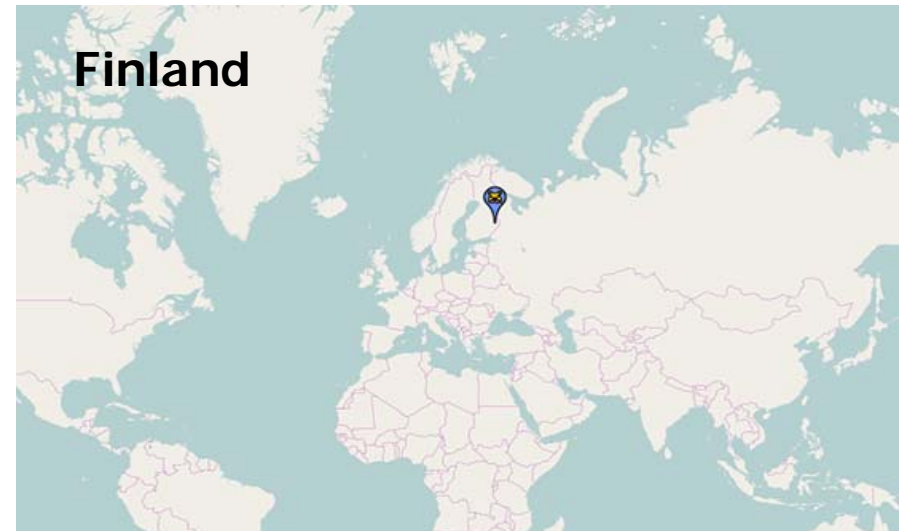
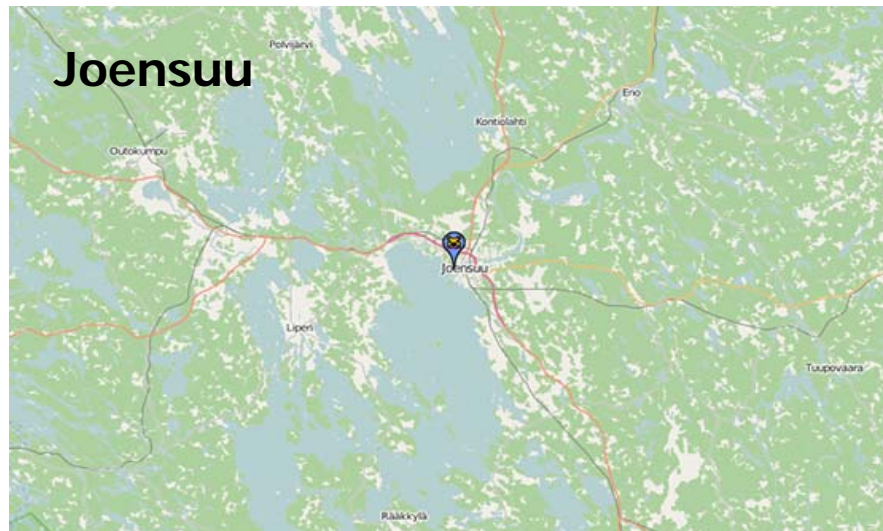
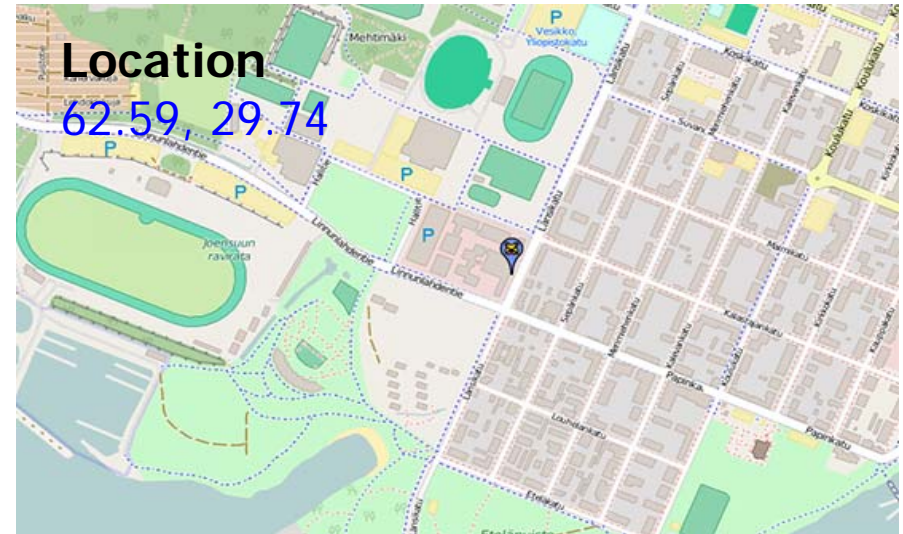
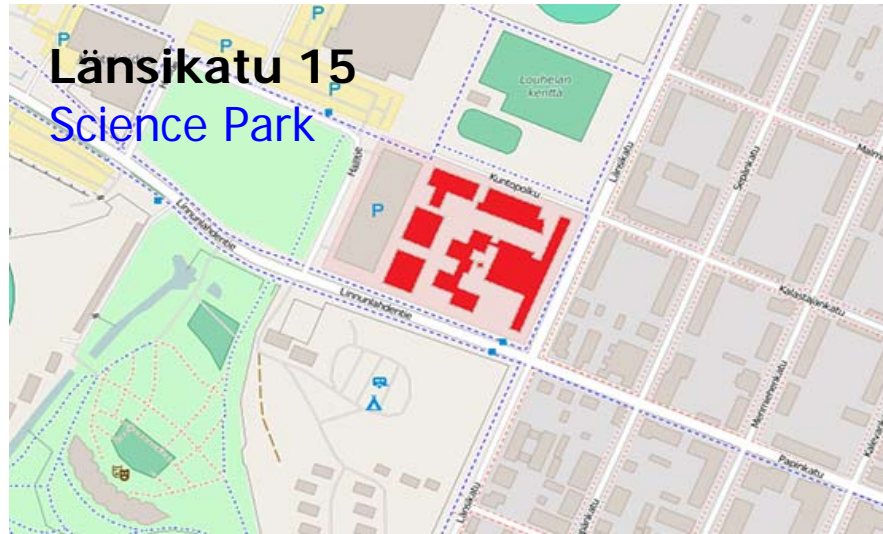


Location

Location hierarchy



Levels of location



Location in web page

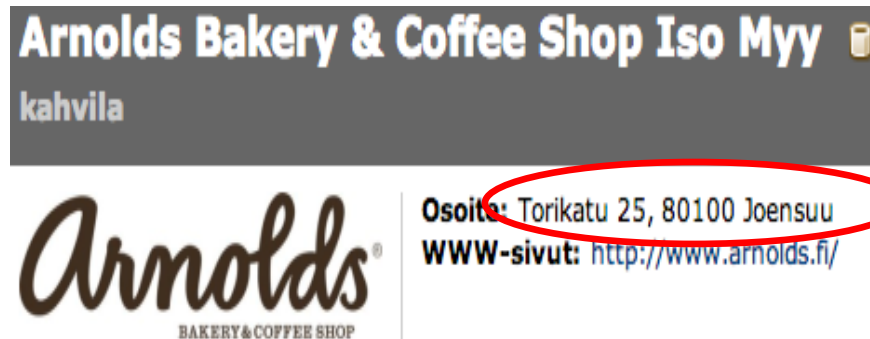
Address tag or geo-tag:

<META name="geo.position" content="62.35; 29.44">

- <0.1% of Finnish websites used geo-tags in **2004** [Vänskä 2004]
- <1% of the websites related to the Oldenburg, Germany used explicit localization in **2008** [Ahlers and Boll, 2008]
- **7%** of Mopsi service websites in May **2015**

Postal address:

- Most service websites have address

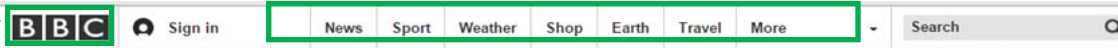


Parsing web page

Content of Web Page

Hypertext Markup Language (HTML, XHTML)

Logo image



Navigation bar

Title

Raspberry Pi 3 adds wi-fi and Bluetooth

2 hours ago | Technology



UK astronaut Tim Peake took a Raspberry Pi to the International Space Station

The Raspberry Pi has become the most popular British computer yet made.

The title was formerly held by the Amstrad PCW which is believed to have sold a total of eight million units.
Sales of the Raspberry Pi will surpass that figure this month, said the Raspberry Pi project founder Eben Upton.

Keywords

The Raspberry Pi

Text

Top Stories

Oscars 2016: DiCaprio finally wins

1 hour ago

UN to expand Syria aid as truce holds

30 minutes ago

Pakistan hangs killer of state governor

4 hours ago

Features & Analysis



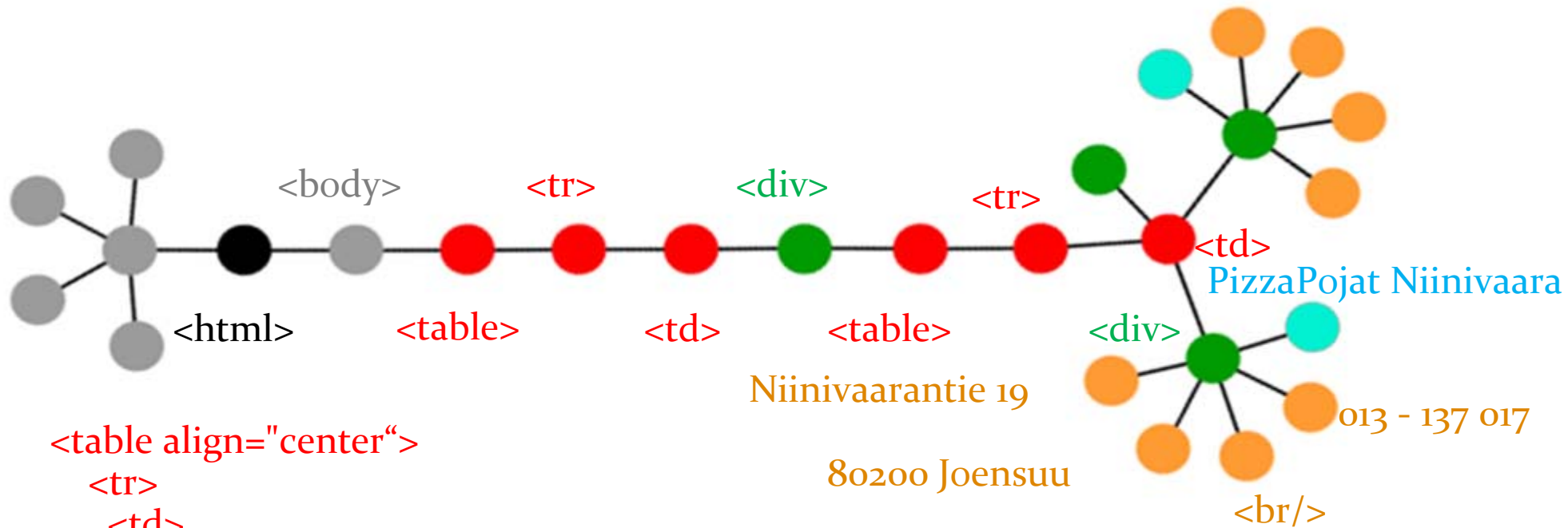
He said yes!

Eight women who proposed to their partners



Images

Another example of DOM tree



```
<table align="center">
  <tr>
    <td>
      <div id="footerleft">
        <h3>PizzaPojat Niinivaara</h3>
        <p>Niinivaarantie 19</p>
        <p>80200 Joensuu</p>
        <br />
        <p>013 - 137 017</p>
      </div>
    </td>
  </tr>
</table>
```

PizzaPojat Niinivaara
Niinivaarantie 19
80200 Joensuu
013 - 137 017

Web site functionality

Single service



[DELI CHINA](#)

[NEWS](#)

[LUNCH MENU](#)

[MENU](#)

[DRINKS LIST](#)

Welcome to our culinary journey in the restaurant Deli Chinaan Joensuu!

Restaurant Deli China is Finland's first Chinese restaurants. Restaurant was founded in 1989 and has been a very popular restaurant in Joensuu among the more than twenty years.

[DELI CHINA](#)

Opening hours

Mon 11-19, Tue-Fri 11-21

Sat 12-21, Sun 12-19

Koskikatu 5, Joensuu
Carelicum in the courtyard
of P. (013) 120 288
sales (at) deli- [china](mailto:china@deli-china.fi) . en

Deli Chinan popularity is largely based on a number of different food culture in the mix. The flavors are essentially the vietnamilaislähtöisen Thi Hanh Nguyen's handwriting. Familiarly known as "Hani" has served as the restaurant's chef de cuisine since its inception. Also, Finnish chef [Timo](#) has worked in a restaurant for nearly as long. Chinese food, the restaurant offers other Asian delicacies, mainly Thai and Vietnamese flavors.

Deli Chinassa invest freshness of the food. Fresh vegetables chopped by hand every morning before the restaurant opened. All fresh vegetables and meats purchased in Joensuu wholesalers. We do not use meat frozen foods, and prefer domestic raw materials.

The restaurant is open for lunch on weekdays from 11 to 14, with a choice of 3 different food options for the six-week round the list. [À la carte](#) dishes are served on

Service directory

kebabile.com

Kebab-haku

Hae

[Etusivu](#)

[Ravintolat](#)

[Lisää ravintola](#)

[Extrat](#)

[Info](#)

[Arvostelut](#) > [Joensuu](#)

Joensuu 7 ravintolaa

0 arvostelua vuoden aikana (yht. 509 arvostelua)

▪ [Aakkosjärjestys](#) ▪ [Paremmuusjärjestys](#)

Bosbor kebab

KKK

Kauppakatu 23
80100 Joensuu
Veronan vieressä
puh. 013 224344

Avoinna: Ma-To 10.00-24.00 Pe 10.00-05.30 La 12.00-05.30 Su 14.00-24.00

Liha	7.6	Kastike	7.5	Lisukkeet	6.8	Kotiinkuljetus	Ei	Anniskelu	Ei
Palvelu	7.8	Siisteys	6.9	Hinnoittelu	6.8	Pankkikortti	Kyllä		

67 arvostelua vuoden aikana (yht. 67 arvostelua)

▪ [Lue asiakkaiden kommentteja \(39 kpl\)](#) ▪ [Arvostele tämä ravintola](#)

Fiesta

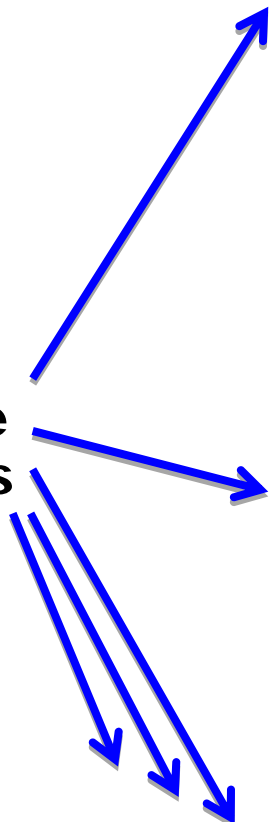
KKK

Suvantokatu 8
80100 Joensuu
Keskustan Citymarketin vastapäätä
puh. 013-120 004
<http://www.pizzeriafiesta.fi/>

Liha	7.4	Kastike	6.8	Lisukkeet	6.4	Kotiinkuljetus	Kyllä	Anniskelu	Ei
Palvelu	7.1	Siisteys	6.9	Hinnoittelu	7.6	Pankkikortti	Kyllä		

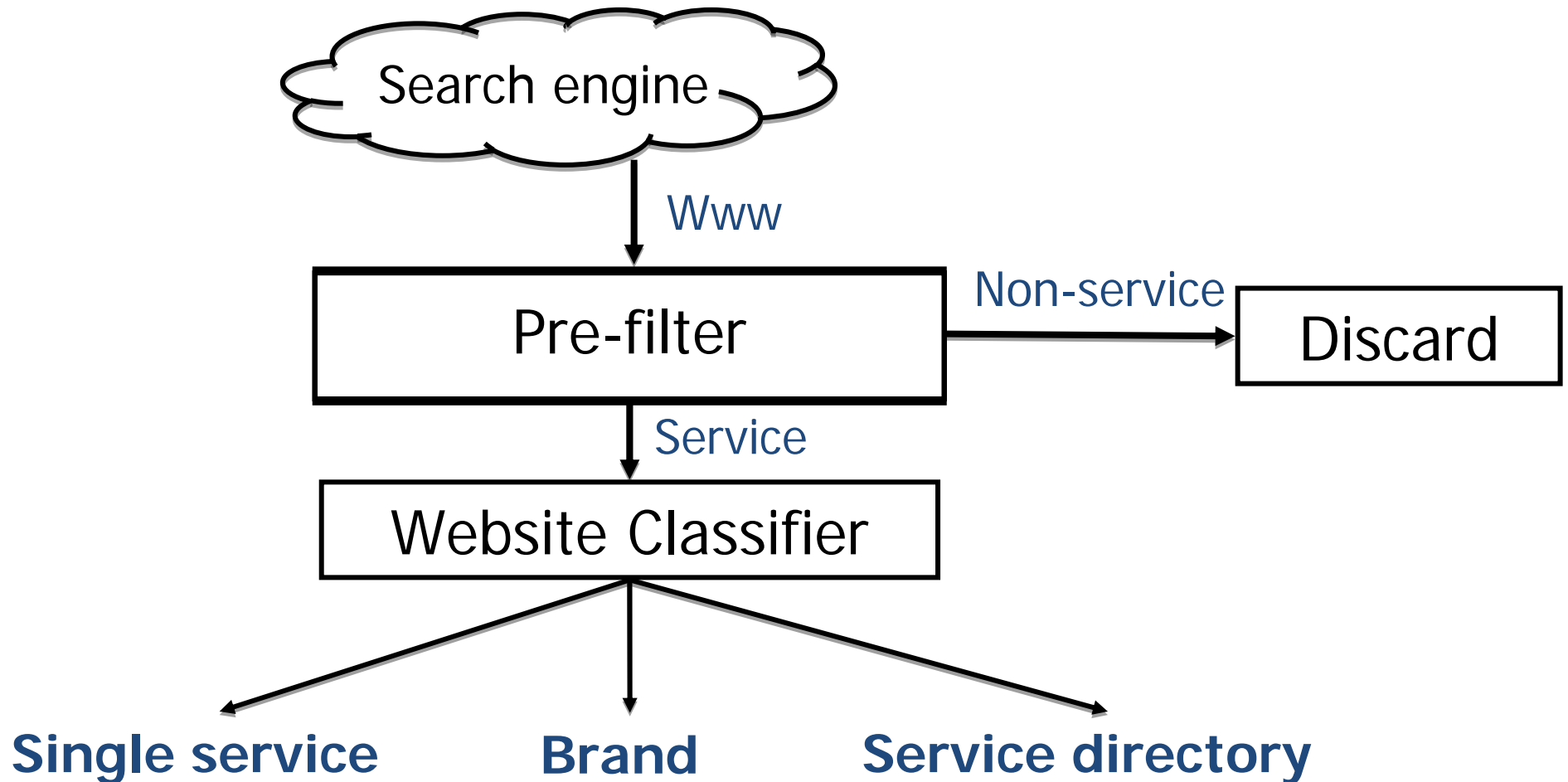
122 arvostelua vuoden aikana (yht. 146 arvostelua)

Multiple
Services



Detecting function of the web page

N. Gali, R. Mariescu-Istodor and P. Fränti, "Functional Classification of Websites"
Int. Symposium on Information and Communication Technology (SoICT),
Nha Trang, Vietnam, 34-41, December 2017



Address detection:

Address detection

Pizza MASTER

★ ETUSIVU ★ JOENSUU ★ LIEKSA ★ KONTIOLAHTI
★ OTA YHTEYTTÄ

JOENSUU

LIEKSA

KONTIOLAHTI

Yhteystiedot
Niskakatu 11
30100 Joensuu
Puh. 0400 281700

KOTIIN-KULJETUS!

ma-to 10:30-22:00
pe-la 10:30-04:30
su 12:00-22:00

PIZZA MASTER JOENSUU

Yhteystiedot
Piellisentie 10-12
81700 Lieksa
Puh. 0400 288481

MAUKAS LOUNAS!

ma-to 11:00-21:00
pe-la 10:30-05:00
su 12:00-21:00

PIZZA MASTER LIEKSA

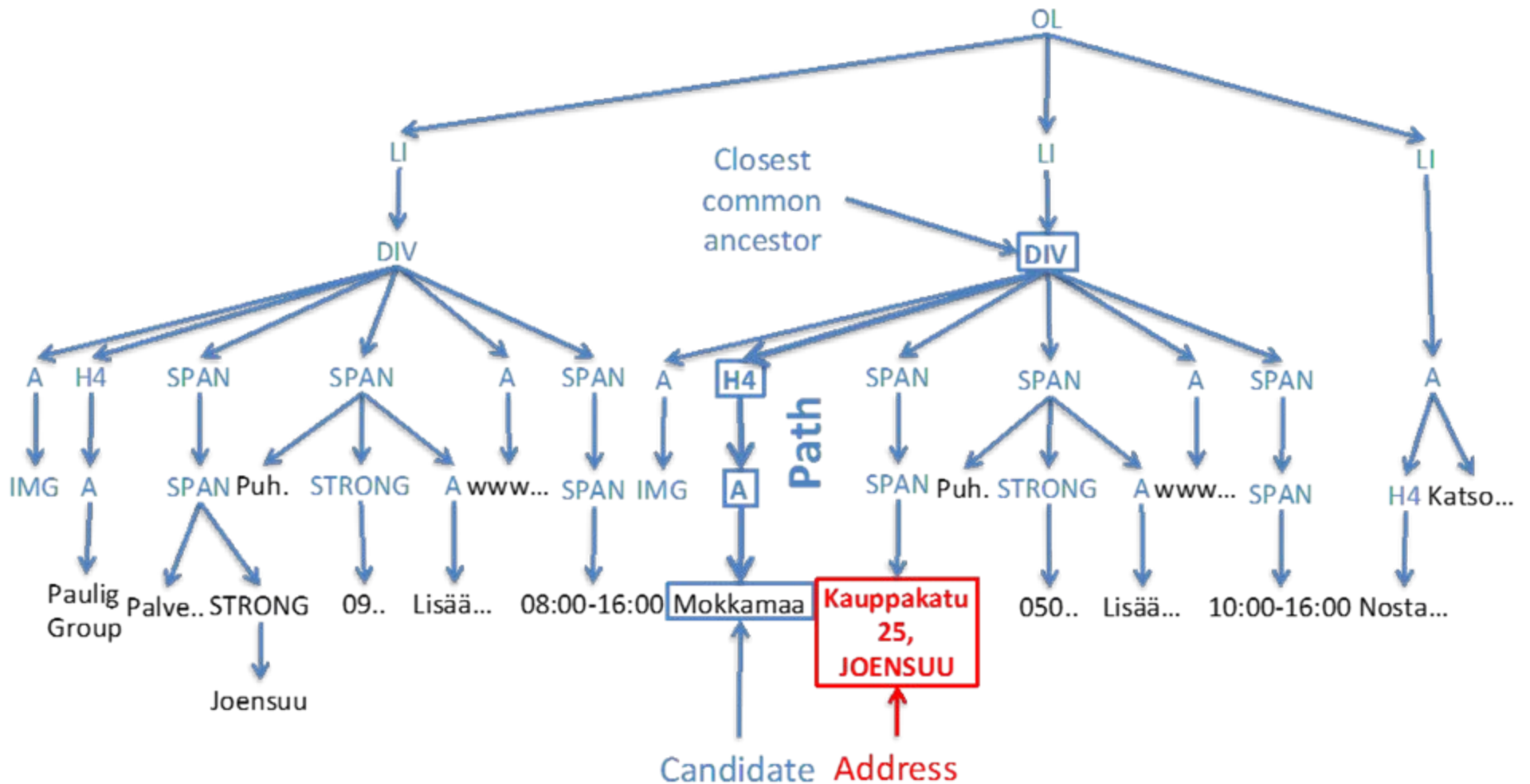
Yhteystiedot
Keskuskatu 21
81100 Kontiolahti
Puh. 0400 732 030

ma-la 11:00-21:00
su 12:00-21:00

PIZZA MASTER KONTIOLAHTI

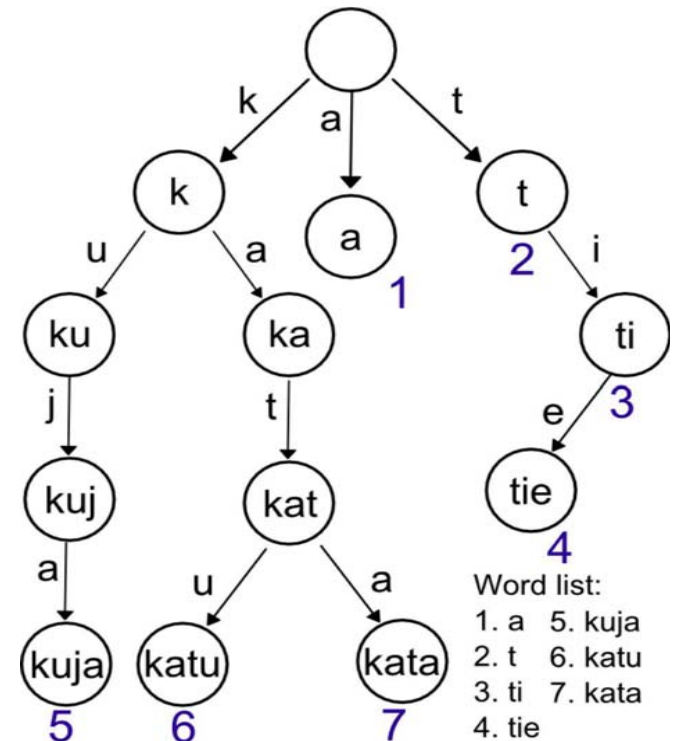
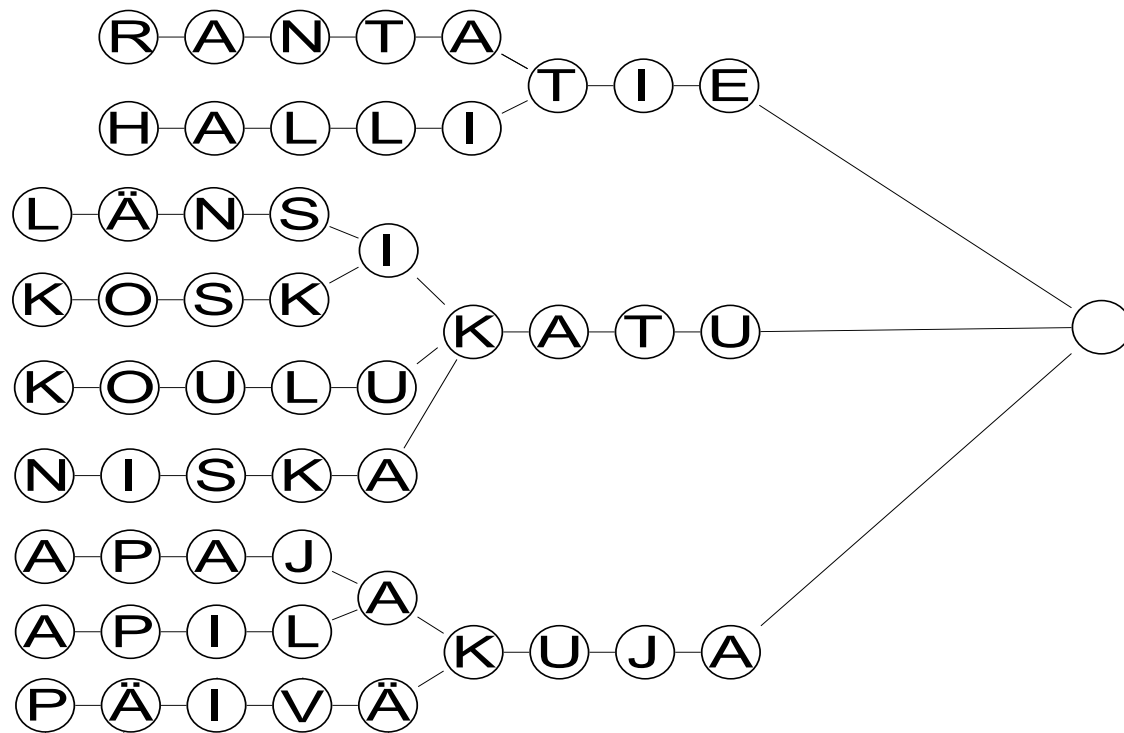
Addresses

DOM tree with address



Detecting address from web

- Analysis of text content of web page
- Matching strings with address database
- Address database stored as **prefix tree**
- Both street number and postal code required



Source of addresses in Mopsi

- Gazetteer for Finland
- OpenStreetMap address data for the rest of world

Metria

Name:

Metria (name)

Type: building:yes

Last Updated: 2015-01-28 20:00

Admin Level: 15

Rank: Other: 30

Coverage: Polygon

Centre Point: 62.60468345,29.7407824768039

OSM: way [88026221](#)

Extra Tags:

Address

Metria (Type: building:yes, 0 [GOTO](#))

Yliopistokatu (Type: highway:tertiary, way [53207737](#), 15, 0 [GOTO](#))

Otsola (Type: place:suburb, node [254754708](#), 15, 0.0117815178075672 [GOTO](#))

Käpykangas (Type: place:suburb, node [254754861](#), 15, 0.0161240029437463 [GOTO](#))

Kanervala (Type: place:suburb, node [254754827](#), 15, 0.0172025860457091 [GOTO](#))

Linnunlahti (Type: place:suburb, node [254754761](#), 15, 0.0159025646651721 [GOTO](#))

Joensuu (Type: place:city, relation [2521759](#), 8, 0.587314067837988 [GOTO](#))

Hasanniemi (Type: place:hamlet, node [631779773](#), 15, 0.0149325182959863 [GOTO](#))

Kuhasalo (Type: place:hamlet, node [631754170](#), 15, 0.0256165398391372 [GOTO](#))

Joensuu (Type: place:city, node [30969521](#), 15, 0.0277955933431901 [GOTO](#))

Mäntylä (Type: place:hamlet, node [255445403](#), 15, 0.0281146102389867 [GOTO](#))

Aittaranta (Type: place:hamlet, node [625884242](#), 15, 0.0283466788109317 [GOTO](#))

Vehkalahti (Type: place:hamlet, node [631553597](#), 15, 0.0412701170922747 [GOTO](#))

Mehtimäki (Type: place:hamlet, node [631779775](#), 15, 0.00490071602319891 [GOTO](#))

North Karelia (Type: boundary:administrative, relation [1999428](#), 6, 0.36936005058504 [GOTO](#))

Regional State Administrative Agency for Eastern Finland (Type: boundary:administrative, relation [2579841](#), 5, 1.27271150703735 [GOTO](#))

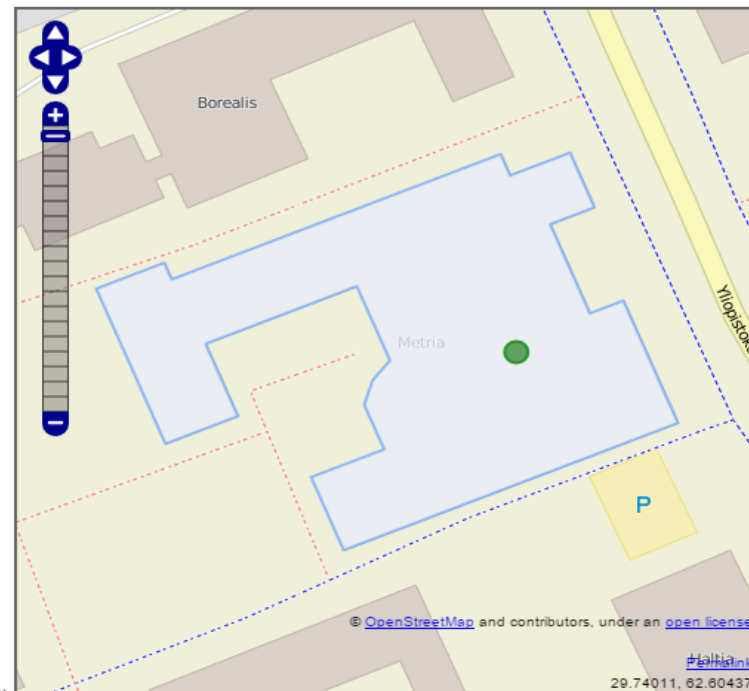
Eastern Finland (Type: boundary:administrative, relation [1702527](#), 4, 1.38580775162333 [GOTO](#))

80101 (Type: place:postcode, 0.00897074416198504 [GOTO](#))

80130 (Type: place:postcode, 0.00950706875080295 [GOTO](#))

Finland (Type: place:country, relation [54224](#), 2, 4.28800463755999 [GOTO](#))

fi (Type: place:country_code, 0 [GOTO](#))



Address matching using Gazetteer

Kaislakatu 8, 80130, Kanervalala, Joensuu, Finland

Torikatu 25, 80100 Joensuu, Finland

Parppeentie 6, 82900 Ilomantsi, Finland

Aleksanterinkatu 25, 15140 Lahti, Finland

Vene 18, 10140 Tallinn, Estonia

Carrer de la Marina, 266-270, Barcelona, Spain

2 Rue Pasteur, 06500 Menton, France

Pulchowk Rd, Lalitpur 44600, Nepal

20 Chả Cá, Hàng Đào, Hoàn Kiếm District, Hanoi, Vietnam

East Coast Park Service Road 1, Singapore

Statistics of prefix trees

	Joensuu	Paris	Singapore	Stockholm
Total number of street names	1145	6472	5316	4584
Total number of tree nodes	5430	38 323	24 062	40 189
Total number of leaf nodes	1131	4836	3274	3940
Maximum depth	20	57	63	51
Average depth of leaf node	12.5	18.5	15.5	15.1
Average width at each level	271	672	382	788
Size (MB)	0.42	3.33	2.09	3.50

Result of address detection

All websites	Location found in...		
	Geo tag	Address tag	Text content
364	10	13	145
100%	3%	4%	40%

All websites	Address is...		
	Correct	Different	None
364	119	26	219
100%	33%	7%	60%

Title extraction:

Two methods

Method A: Title Tag Analyzer (TTA)

N. Gali and P. Fränti, "Content-based title extraction from web page", *Int. Conf. on Web Information Systems & Technologies (WEBIST'16)*, Vol.2, 204-210, Rome, Italy, April 2016.

Method B: Titler

N. Gali, R. Marinescu-Istodor and P. Fränti, "Using linguistic features to automatically extract web page title", *Expert Systems with Applications*, 79, 296-312, 2017.

Web Page Title

The title can be in three different places:

<title>Wentworth House Hotel Bath Hotels - Cheap Hotels in Bath, Somerset, UK</title>

- Title Tag (91 %)
- Logo image (89 %)
- Web page body (93 %)

The screenshot shows the homepage of the Wentworth House Hotel website. The header features the hotel's name in a cursive font. A navigation menu includes links for Home, Booking, Direction, Rooms, Dining, Offers, Gallery, Bath, T&C, and Contact. A central banner displays a photograph of the hotel building, with a red circle highlighting the text 'Wentworth House Hotel' and a brief description. To the right, a 'Secure Booking' widget allows users to select arrival dates and the number of nights. Below the banner, a horizontal bar lists amenities: FREE ON SITE PARKING, FREE WIFI, and FULL ENGLISH BREAKFAST. The main content area is divided into three columns: 'WELCOME' with introductory text, 'EXPLORE BATH' with a photo of Bath Abbey, and 'GALLERY' with two smaller photos. The footer contains contact information and social media links.

Title and Meta Tags

The obvious source

But includes also additional information

```
<title> Piato Restaurant – 123 Blues Point Road, McMahons  
Point, Sydney | Visit Piato and experience the life & flavour  
of Europe. North Sydney Functions. North Sydney  
Restaurants. </title>
```

```
<title> Joensuu Keskusta | Intersport - Sport to the people  
</title>
```

Segmentation is needed!



Workflow of method A

N. Gali and P. Fränti, "Content-based title extraction from web page", *Int. Conf. on Web Information Systems & Technologies (WEBIST'16)*, Vol.2, 204-210, Rome, Italy, April 2016.



1. Placement in title & meta tags
2. Popularity in header tags
3. Position in the web link

Web page

Extract title & meta tags from the page

Segment content by delimiters

Construct candidate list

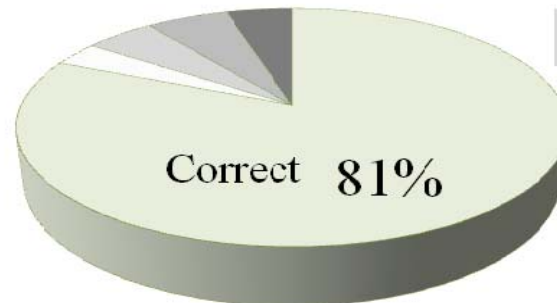
Score candidate segments

Title The coronet

Qualitative Analysis of TTA

Title	Ground truth	Content of Title tag	Selected string
Correct	3 Weeds Hotel	3 Weeds Hotel Unique Pub Bars Restaurant Party Venue Inner West Sydney	3 Weeds Hotel
Short	Irish Channel Restaurant & Pub	Irish Channel - Restaurant & Pub 500 H St NW DC (202) 216-0046	Irish Channel
Long	Secret Garden Bed & Breakfast	Secret Garden Bed & Breakfast (formerly Whitegates Guest House), near Keynsham, Bristol: Rooms, Prices and Guest Information	Secret Garden Bed & Breakfast (formerly Whitegates Guest House)
No title	Rio Pool	Hot Tubs, hot tub hire, swimming pools, Bristol, Gloucester	swimming pools
Incorrect	Slice and Dice	Home Prepared Food Swansea Slice and Dice UK	Swansea

Long (5 %)
[CATEGORY NAME] (3 %)



No title (6 %)

[CATEGORY NAME] (5 %)

Results with Mopsi Services

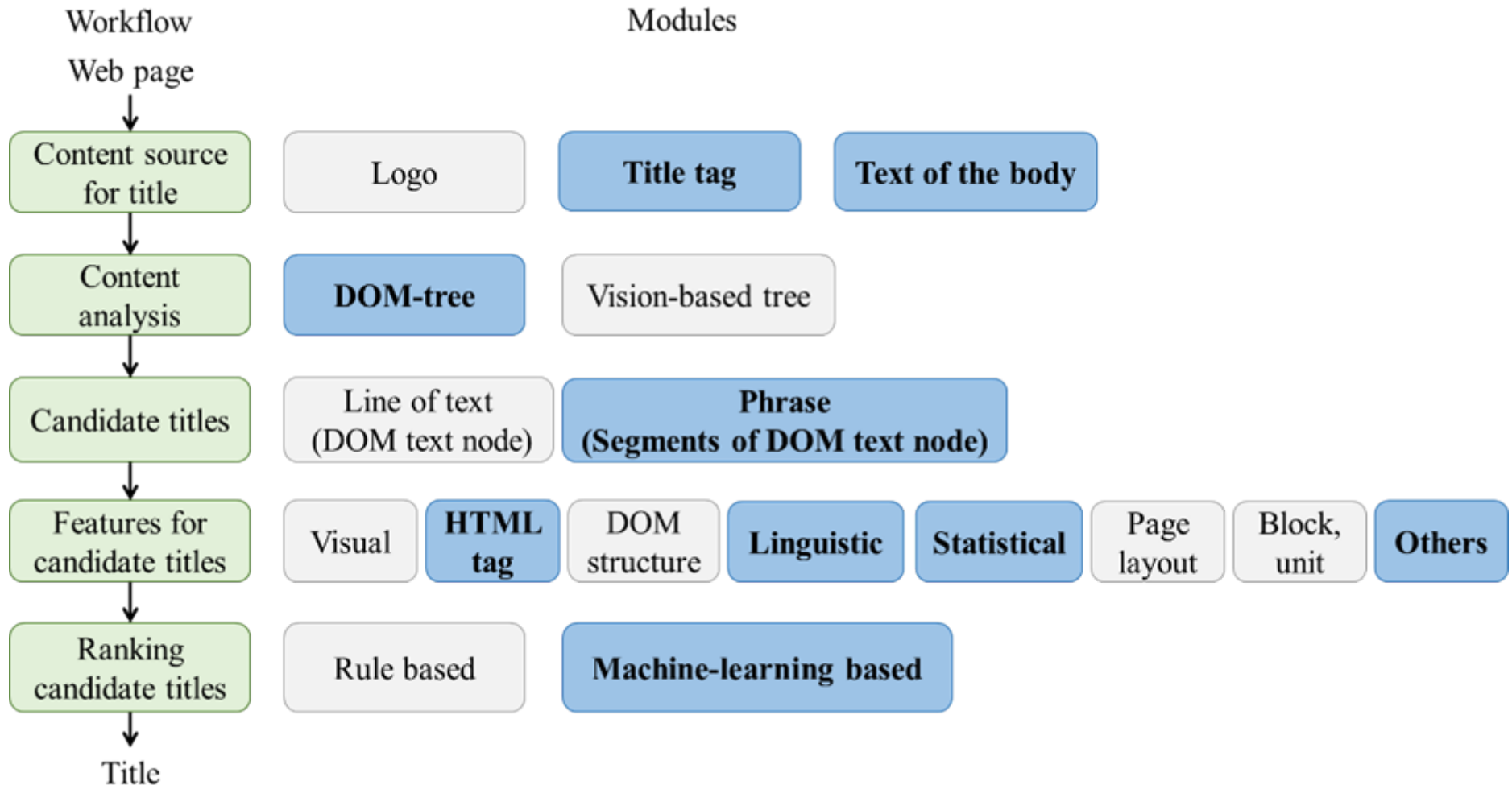


Annotated titles

Method	Rouge-1			Jaccard	Dice
	Precision	Recall	F-score		
Baseline (Title Tag)	0.71	0.33	0.41	0.44	0.54
TitleFinder (Moham.et al. 2012)	0.35	0.47	0.37	0.37	0.43
Styling (Changuel et al. 2009)	0.14	0.21	0.15	0.22	0.28
TTA (Gali and Fränti 2016)	0.52	0.59	0.52	0.54	0.62

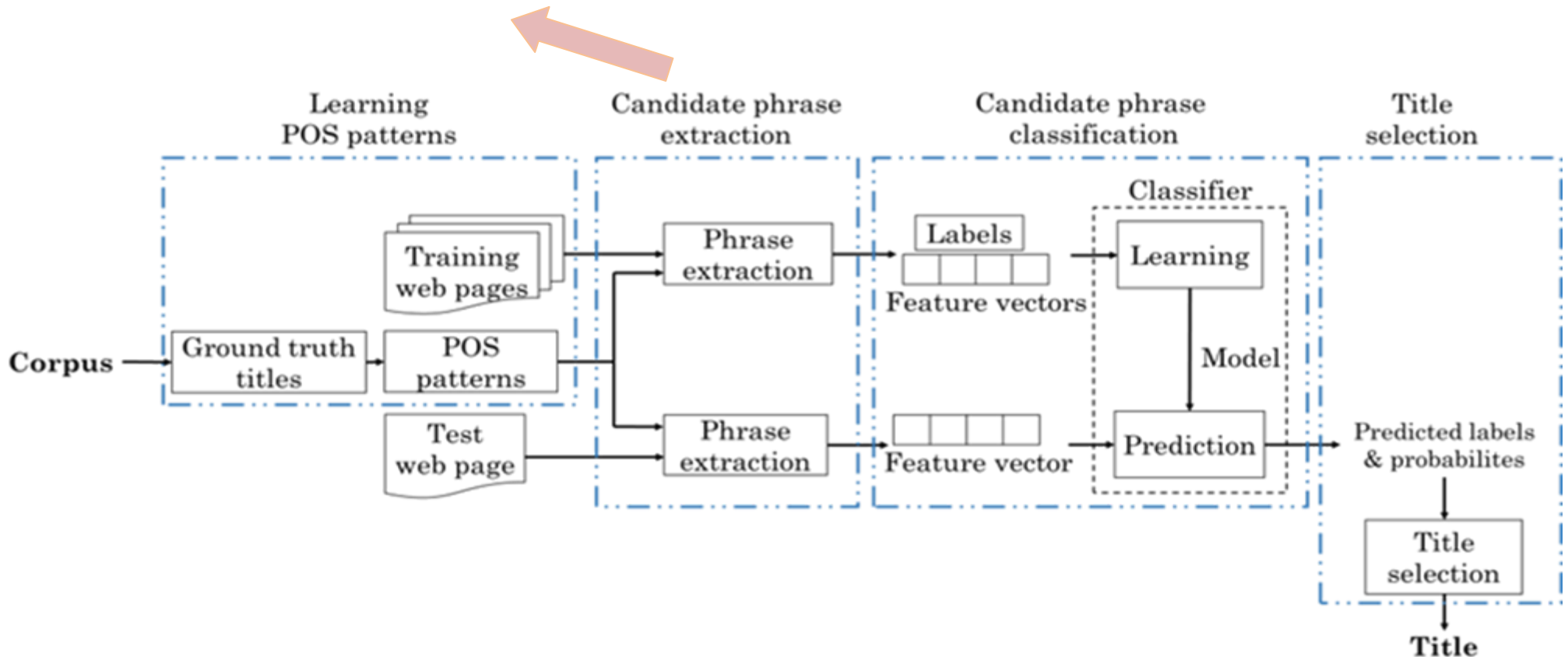
Workflow of method B

N. Gali, R. Mariescu-Istodor and P. Frănti, "Using linguistic features to automatically extract web page title", *Expert Systems with Applications*, 79, 296-312, 2017.



Representative title

Content of text nodes → N-grams (n=1...6) → Filter by part-of-speech (POS) patterns



POS tagging of phrases

NNP=Proper noun, singular
NNPS=Proper noun, plural
NN=Noun, singular or mass
VBG=Verb, gerund
VB=Verb, base form
PRP=Personal pronoun
DT=Determiner
CC=Coordinating conjunction
JJ=Adjective

NNP

Navigation

VBG

NNP

VB

PRP IN

Feeling Social? Find us on

NNP

Facebook

NNP

NNP

NNP

NNP

NNPS

NN

Sydney Waterfront Restaurant Restaurant Milsons Point

NNP

NNP

VBZ

DT

JJ

NNP

NN

NN

Aqua Dining offers a quintessential Sydney dining experience

IN

JJ

NN

NNS

WDT

NN

IN

NNP

NNP

IN

DT

with unrivalled harbour views that sweep from Luna Park to the

NN

JJ

NNP

NNP

NNP

CC

DT

NNP

NNP

world famous Sydney Harbour Bridge and the Sydney Opera

NNP

House.

Comparison

Mopsi services

Method	Rouge-1			Jaccard	Dice
	Precision	Recall	F-score		
Baseline	0.33	0.71	0.41	0.44	0.54
Google	0.34	0.74	0.43	0.46	0.56
TitleFinder (Mohammadzadeh et al., 2012)	0.35	0.47	0.37	0.37	0.43
Styling (Changuel et al., 2009)	0.14	0.21	0.15	0.22	0.28
TTA (Gali & Fränti, 2016)	0.52	0.59	0.52	0.54	0.62
Titler (KNN)	0.59	0.56	0.55	0.59	0.66

Method A

Method B

What about logo images?

~89 % of web pages have their title within a logo image

Needs to detect logo image

Apply OCR

Challenging !!!

LOCAL BISTRO

Santa's Reindeers

Savon Kinot



Representative image:

N. Gali, A. Tabarcea, and P. Fränti, "Extracting representative image from web page",
Int. Conf. on Web Information Systems & Technologies (WEBIST'15), 411-419
Lisbon, Portugal, May 2015.

Image categories

Banner



Formatting



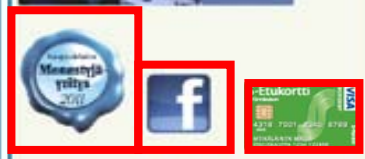
Logo



Representative



Icons



Advertisement

Tarjoimme Kreikkalaisen lounaan ma-pe kello 11-14. Tule tutustumaan lounaslistaamme.

Kaikki ruokamme saat myös kätevästi kotipakettiin. Isompiin tilauksiin tarvittaessa myös kujetus.

Mikäli sinulla on erityistoiveita ruokailun suhteen, soita Lakkselle ravintolaamme.

Tule tutustumaan Kreikkalaisen keittiön aarteisiin Ravintola Kreetaan Kuopiossa, Maaherrankatu 23, Joensuussa, Iso Myyn 2. krs Kauppakatu 28.

Overall extraction process

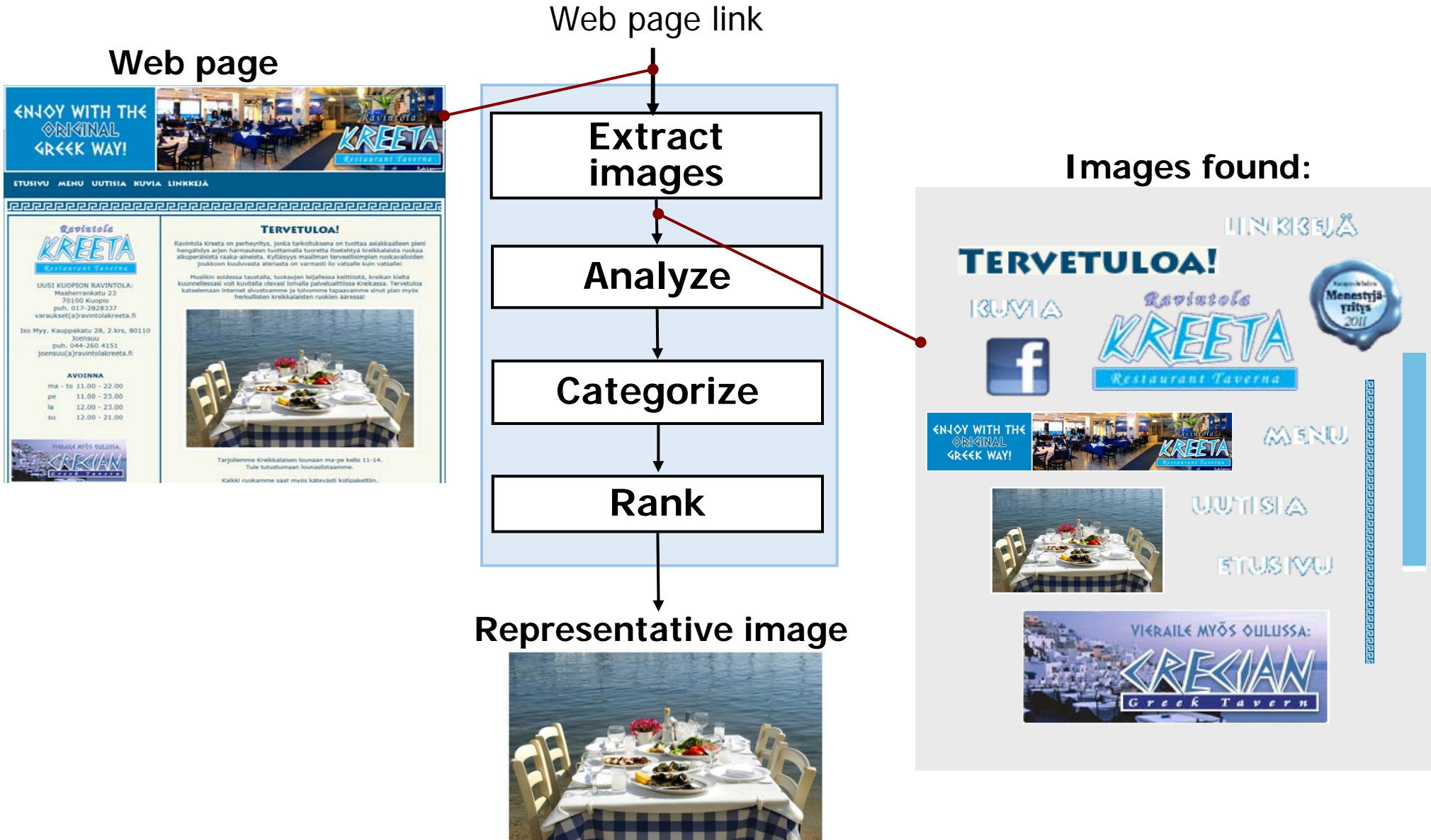


Image features used



src	http://www.ravintolakreeta.fi///images/banner.jpg
alt	--
title	--
from	css
format	jpg
width	945
height	202
size	190,890 px
aspect ratio	4.67
parent tag	<div>
class	header

Summary of the rules

Category	Features	Keywords
Representative	Not in other category	
Logo		logo
Banner	Ratio > 1.8	Banner, header, Footer, button
Advertisement		Free, adserver, now, buy, join, click, affiliate, adv, hits, counter
Formatting and Icons	Width < 100 px Height < 100 px	Background, bg, spirit, templates

Mopsi WebIma dataset



<http://cs.uef.fi/mopsi/data/>

Summary of data collected:

Websites: 1002

Images: 2363

Per page: Min=1, Average=2.36, Max=154

Collection details:

Who: 117 volunteers

When: September 2014

What: Pages of own choice or Mopsi search

How: Select 1-3 most representative images

Issues: Some level of subjectivity unavoidable

Results summary

	Accuracy	Extracted Images
WebIma	64%	99%
Google+	48%	92%
Facebook	39%	90%

- **Lightweight** method suitable for real time applications
- **Unsupervised**: No training, no user feedback needed
- In use in MOPSI: *Search* and *Service upgrade*

Recommendation system

Mopsi search

MOPSI

See what's around

☰

cafe

Search

Keyword search

MOPSI

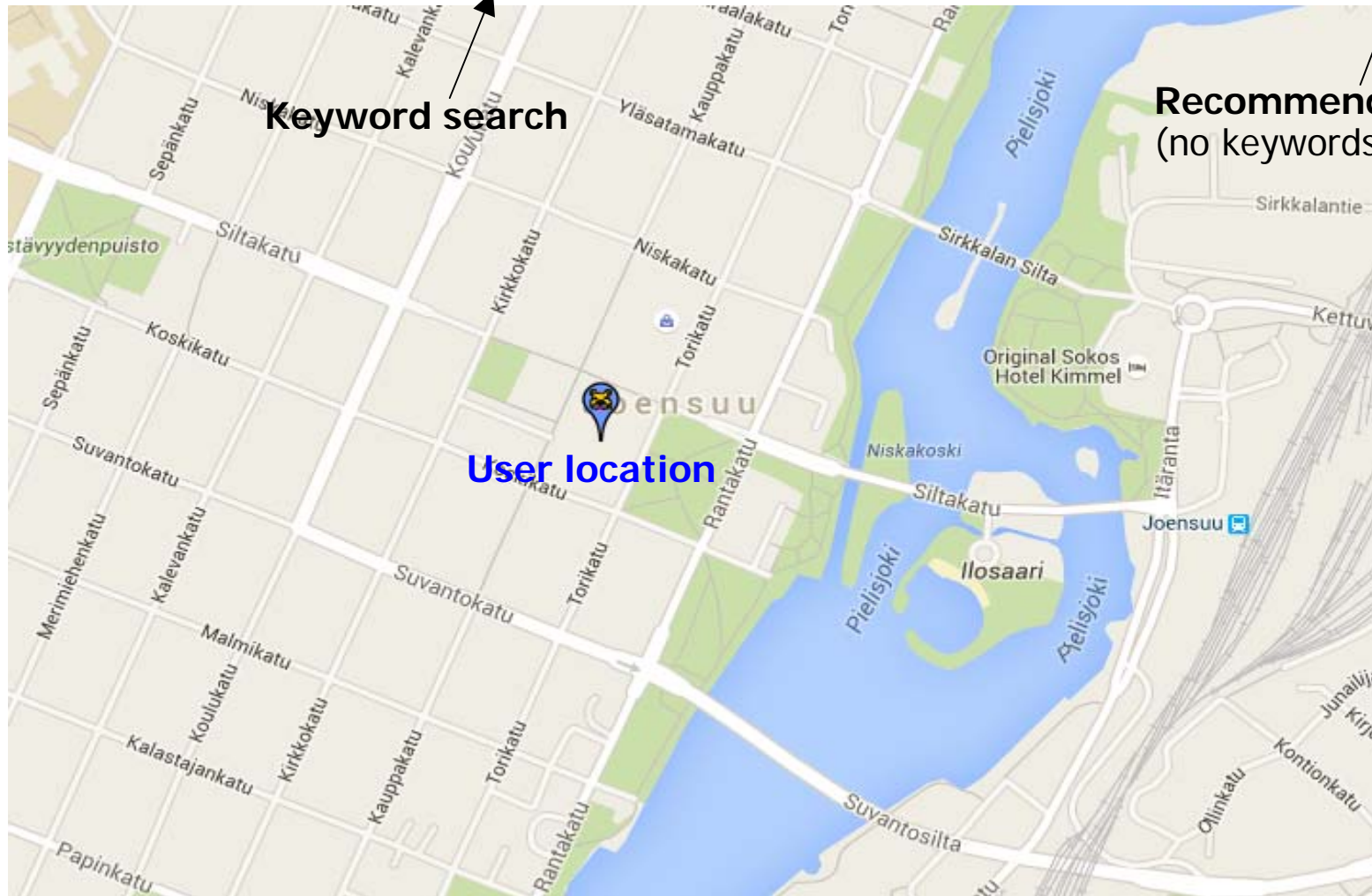
See what's around

☰

search using a keyword

Recommend

Recommendation
(no keywords)



Location-aware recommendation

Input:

- User
- Location
- Time
- Keyword (optional)

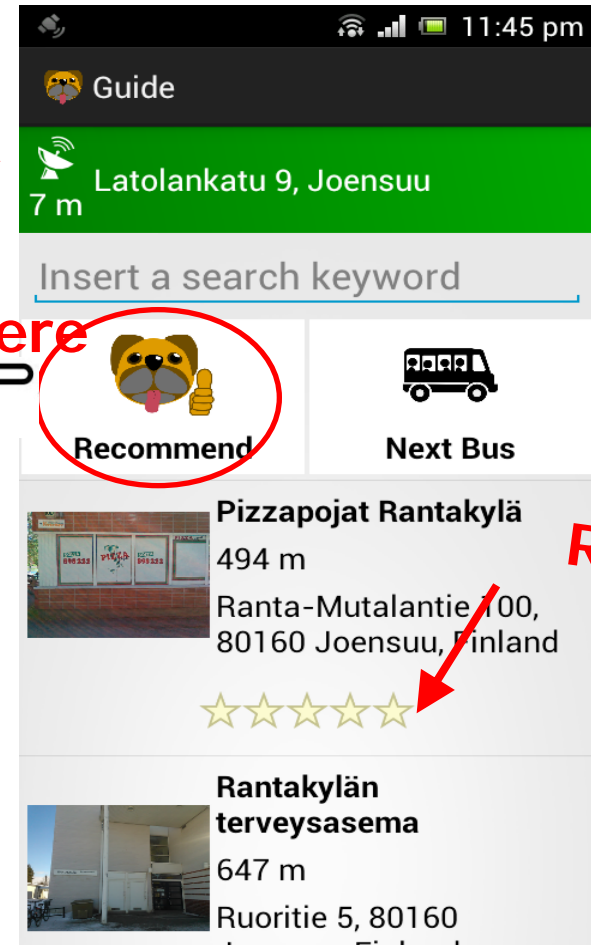
Recommendations:

- Nearby services
- Photos of other users

Location



Press here



Results

Industrial zone

Rahkeentie

Vilkku kahvio

Location 0.55
 Time 1.00
 Relevance 0.84
 User network 0.57 ★★ ★
Total: 2.95

Keywords	Nearby	Recently	Total
kahvila	0.32	0.00	0.27
Total (Max)	0.32	0.00	0.27
Total (Av)	0.32	0.00	0.27

Kuurnankulma

Location 0.60
 Time 1.00
 Relevance 0.32
 User network 1.00 ★★ ★★ ★★ ★★
Total: 2.92

Keywords	Nearby	Recently	Total
lounas	0.02	0.00	0.02
Total (Max)	0.02	0.00	0.02
Total (Av)	0.02	0.00	0.02

Heinosen leipomo

Location 0.95
 Time 1.00
 Relevance 0.84
 User network 0.00
Total: 2.79

Keywords	Nearby	Recently	Total
kahvila	0.32	0.00	0.27
Total (Max)	0.32	0.00	0.27
Total (Av)	0.32	0.00	0.27

1 ★★ ★★ ★
Vilkku Kahvio
 Pamilonkatu 33, Joensuu
 740 m
 Recom info
 Check route

2 ★★ ★★ ★★
Kuurnankulma
 Paukkajantie 2-4, Joensuu
 762 m
 Recom info
 Check route

3 ★★ ★★ ★★
Kahvila Heinosen leipomo
 Hiiskoskentie 13, Joensuu
 306 m
 Recom info
 Check route

4 ★★ ★★ ★★
La Dolce Vita
 Kuurnankatu 6, Joensuu
 625 m
 Recom info
 Check route

5 ★★ ★★ ★★
Kahvila Huili & Javerstok grilli
 Kuurnankatu 14, Joensuu
 420 m
 Recom info
 Check route

6 ★★ ★★ ★★
Skarppi - ylioppilaskunnan sauna
 Kaislaku 10, Joensuu
 758 m
 Recom info

Map showing locations of Vilkku kahvio, Kuurnankulma, and Heinosen leipomo. Distances are marked: Vilkku kahvio (740 m), Kuurnankulma (762 m), and Heinosen leipomo (306 m). Search results for 'kahvila' and 'lounas' are shown in tables.

Solutions for recommendation

Recommendation:

- User statistics
- Location
- Time

K. Waga, A. Tabarcea and P. Fränti, "Recommendation of points of interest from user generated data collection", *IEEE Int. Conf. on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom'12)*, Pittsburgh, USA, 2012.

User network:

- Similarity of users
- Local knowledge

P. Fränti, K. Waga, and C. Khurana, "Can social network be used for location-aware recommendation?", *Int. Conf. on Web Information Systems & Technologies (WEBIST'15)*, 558-565, Lisbon, Portugal, May 2015.

Conclusions

Key challenges:

- Detecting location and text summary

Is it effective?

- **40%** of websites contain useful location

When it works?

- GOOD: Service web page
- NOT SO GOOD: Blogs, news stories...

Publications

1. A. Tabarcea, N. Gali and P. Fränti, "Framework for location-aware search engine", *Journal of Location Based Services*, 11 (1), 50-74, November 2017.
2. N. Gali, R. Mariescu-Istodor and P. Fränti, "Using linguistic features to automatically extract web page title", *Expert Systems with Applications*, 79, 296-312, 2017.
3. N. Gali, R. Mariescu-Istodor and P. Fränti, "Functional Classification of Websites", *Int. Symposium on Information and Communication Technology (SoICT)*, Nha Trang, Vietnam, 34-41, December 2017
4. N. Gali, R. Mariescu-Istodor and P. Fränti, "Similarity measures for title matching", *IAPR Int. Conf. on Pattern Recognition, (ICPR'16)*, Cancun, Mexico, 1549-1554, December 2016.
5. N. Gali and P. Fränti, "Content-based title extraction from web page" , *Int. Conf. on Web Information Systems and Technologies (WEBIST 2016)*, Rome, Italy, vol. 2, 204-210, April 2016.
6. M. Rezaei, N. Gali, and P. Fränti, "CIRank:a method for keyword extraction from web pages using clustering and distribution of nouns", *IEEE/WIC/ACM Int. Joint Conf. on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 79-84, December 2015.
7. P. Fränti, K. Waga, and C. Khurana, "Can social network be used for location-aware recommendation", *Int. Conf. on Web Information Systems & Technologies (WEBIST'15)*, 558-565, 2015.
8. N. Gali, A. Tabarcea, and P. Fränti, "Extracting representative image from web page", *Int. Conf. on Web Information Systems & Technologies (WEBIST'15)*, 411-419, 2015
9. K. Waga, A. Tabarcea and P. Fränti, "Recommendation of points of interest from user generated data collection", *IEEE Int. Conf. on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom'12)*, Pittsburgh, USA, 2012.
10. P. Fränti, J. Chen, A. Tabarcea, "Four aspects of relevance in location-based media: content, time, location and network" *Int. Conf. on Web Information Systems & Technologies (WEBIST)*, 2011
11. A. Tabarcea, V. Hautamäki, P. Fränti, "Ad-hoc georeferencing of web-pages using street-name prefix trees", *Int. Conf. on Web Information Systems & Technologies (WEBIST'10)*, Valencia, Spain, vol.1, 237-244, April 2010.

PhD theses

1. Radu Mariescu-Istodor, "Efficient management and search of GPS routes", [PhD thesis](#), School of computing, Univ. Eastern Finland, August 2017.
2. Najlaa Gali, "Summarizing the content of web pages", [PhD thesis](#), School of computing, Univ. Eastern Finland, June 2017.
3. Mohammad Rezaei, "Clustering validation", [PhD thesis](#), School of computing, Univ. Eastern Finland, June 2016.
4. Karol Waga, "Processing, analysis and recommendation of location data", [PhD thesis](#), School of computing, Univ. Eastern Finland, June 2015.
5. Andrei Tabarcea, "Location-based web search and mobile applications", [PhD thesis](#), School of computing, Univ. Eastern Finland, 2014.
6. Minjie Chen, "Efficient processing and compression of map images and routes", [PhD thesis](#), School of computing, Univ. Eastern Finland, August 2012.
7. Qinpei Zhao, "Cluster validity in clustering methods", [PhD thesis](#), School of computing, Univ. Eastern Finland, June 2012.