

S-Sense: **An Opinion Mining Tool for Market Intelligence**

Choochart Haruechaiyasak and Alisa Kongthon

Speech and Audio Technology Laboratory

National Electronics and Computer Technology Center (NECTEC)

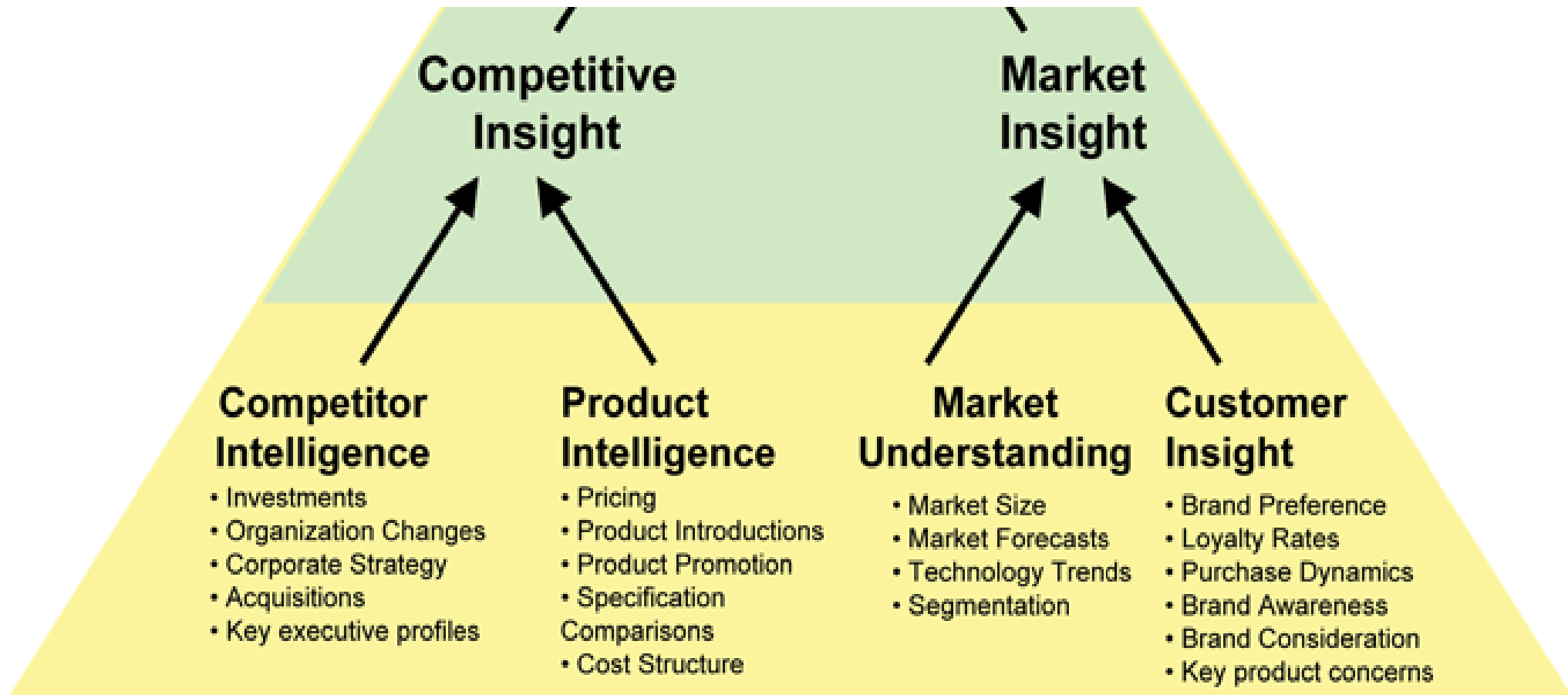
What is market intelligence ?

Marketing Intelligence is the information relevant to a company's markets, gathered and analyzed specifically for the purpose of accurate and confident **decision-making** in determining **market opportunity, market penetration strategy, and market development metrics.**

What is market intelligence ?



What is market intelligence ?



Customer insight

Customer insight involves fully understanding the customer preferences, drivers of customer behavior, brand loyalty, satisfaction rates and any other customer views that impact their behavior in relation to your firm's products or services.

Competitor Intelligence

- Investments
- Organization Changes
- Corporate Strategy
- Acquisitions
- Key executive profiles

Product Intelligence

- Pricing
- Product Introductions
- Product Promotion
- Specification Comparisons
- Cost Structure

Market Understanding

- Market Size
- Market Forecasts
- Technology Trends
- Segmentation

Customer Insight

- Brand Preference
- Loyalty Rates
- Purchase Dynamics
- Brand Awareness
- Brand Consideration
- Key product concerns

Background and motivation

- Traditional market survey is often carried out through **voluntary questionnaires or focused group interviews.**
 - limited number of participated customers
 - controlled environments
- With the **Web 2.0 or social networking websites,**
 - the amount of user-generated contents (UGC) has increased exponentially.
 - UGC often contain **opinions and/or sentiments.**
- An in-depth analysis of these opinionated texts could reveal preferences of customers towards **products and services.**

Example of online review



FlyingGar...
Walpole,
Massachusetts
2 reviews



“Great Location, Beautiful Lobby, Tiny Rooms”

●●●○○○ Reviewed July 28, 2011 **NEW**

I've stayed at this hotel many times. It's a great property with lots to love and a few drawbacks. The staff is helpful and mostly friendly. The location is primo and the lobby is beautiful. There's a good gym and the beds have great linens. And you have to love the Lavazza coffee makers and large windows...which are two walls of windows in corner rooms.

But...the rooms and the bathrooms are TINY (not so much in the Executive Towers). Most bathrooms are nicely done with marble and granite, but there's not place to put a shaving kit. The restaurants and bar are just average...so go out and enjoy the city!

I would have given the Hilton Portland 4 stars, but their expensive internet policy keeps it to three.

Room Tip: ask for a corner room or stay at the
[See more room tips](#)

Stayed April 2011, traveled on business

●●●○○○ Value

●●●○○○ Sleep Quality

●●●○○○ Cleanliness

●●●○○○ Service

Background and motivation (cont'd)

- Opinion mining and sentiment analysis is a task for **analyzing** and **summarizing** what people think about a certain topic.
- Three granularities of opinion mining:
 - **Document level** [Turney, 2002; Pang et al., 2002; Dave et al., 2003; Beineke et al., 2004]
 - **Sentence level** [Kim and Hovy, 2004; Wiebe and Riloff, 2005; Wilson et al., 2009; Yu and Hatzivassiloglou, 2003]
 - **Feature level** [Hu and Liu, 2004; Popescu and Etzioni, 2005]

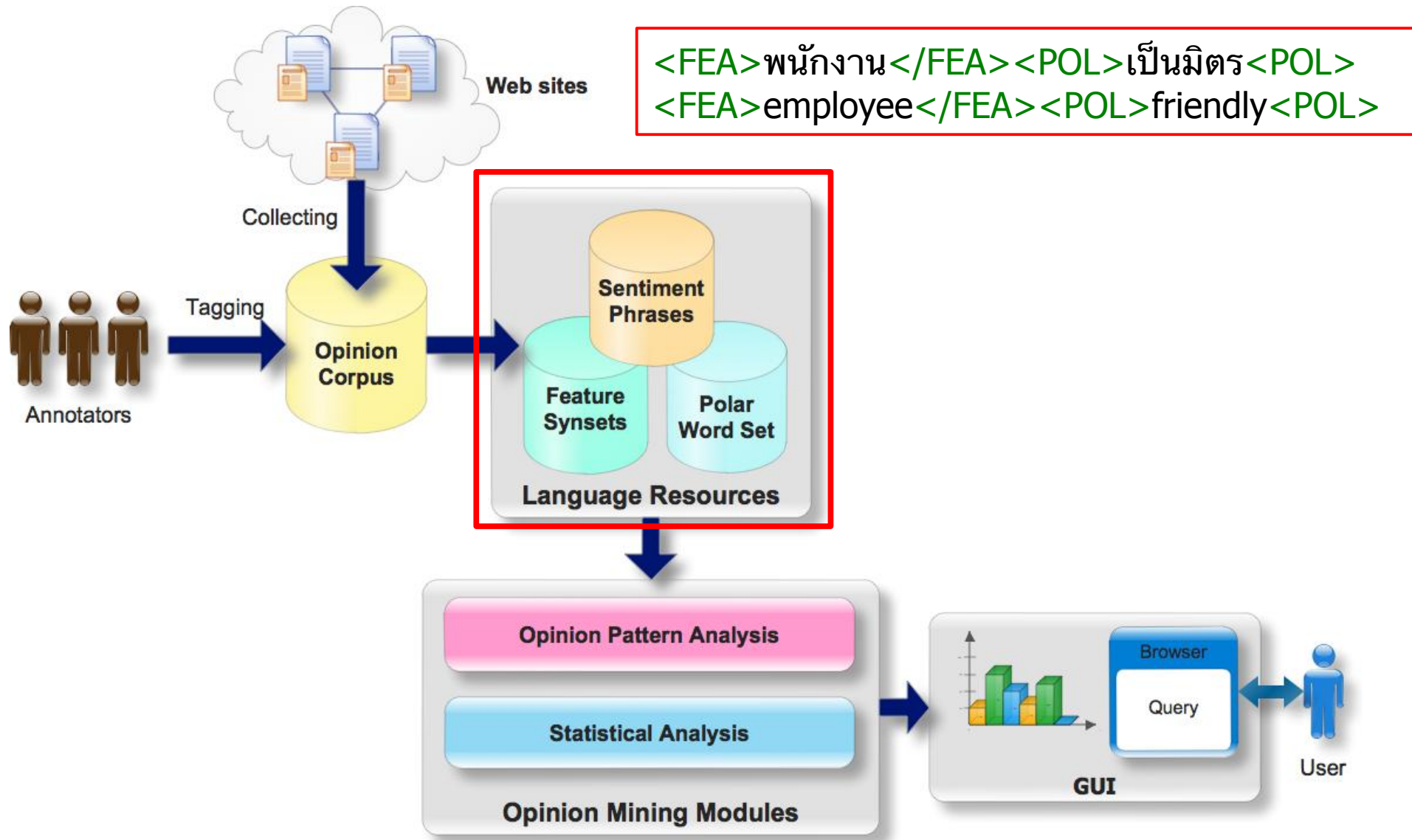
Feature Based Opinion Mining

- We focus on the **feature level** or **feature-based** opinion mining. This approach typically consists of two following steps.
 - (1) Identifying and extracting **features** of an object, topic or event from each sentence.
 - (2) Determining whether the opinions regarding the features are **positive** or **negative**.

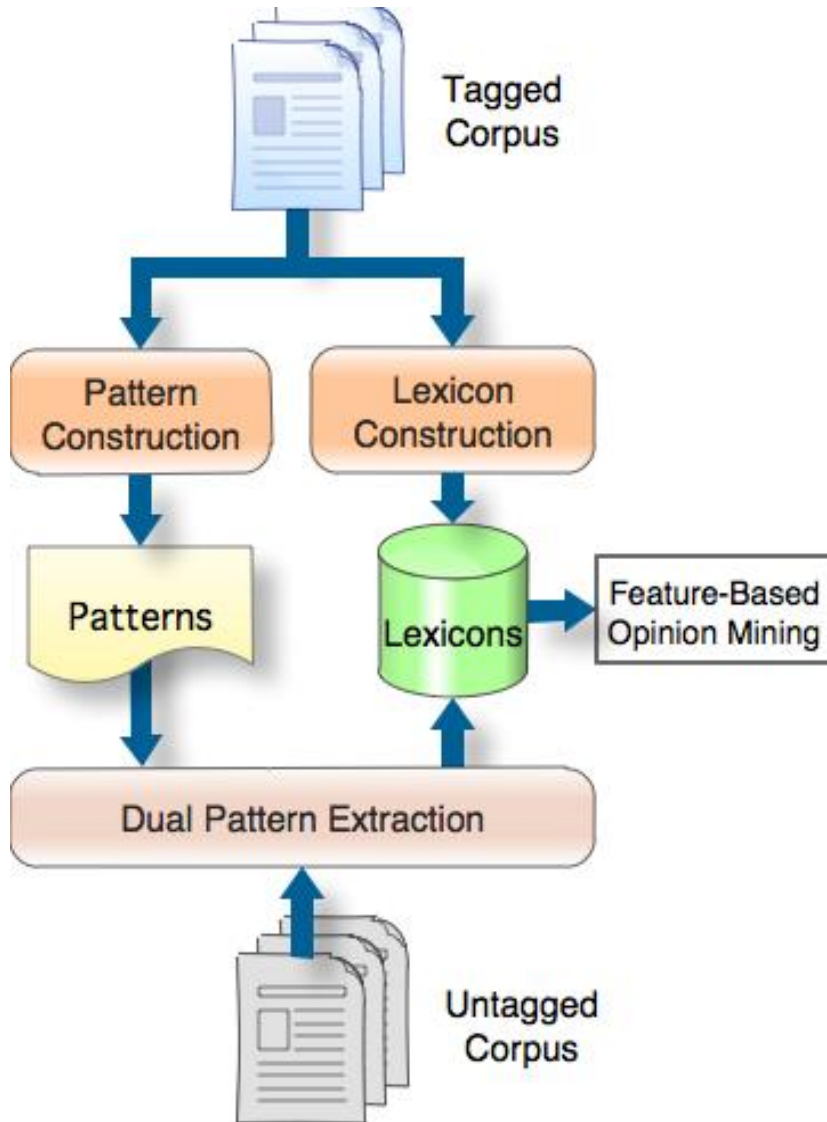
Feature Based Opinion Mining

- The **feature-based** opinion mining could provide users with some **insightful information** related to opinions on a particular topic.
- For example, on **hotel reviews**, users can view positive or negative opinions on hotel-related features such as *price, service, breakfast, room, facilities* and *activities*.
- Breaking down opinions into feature level is very essential for **decision making**.
 - Different customers could have different preferences when selecting hotels to stay for vacation.

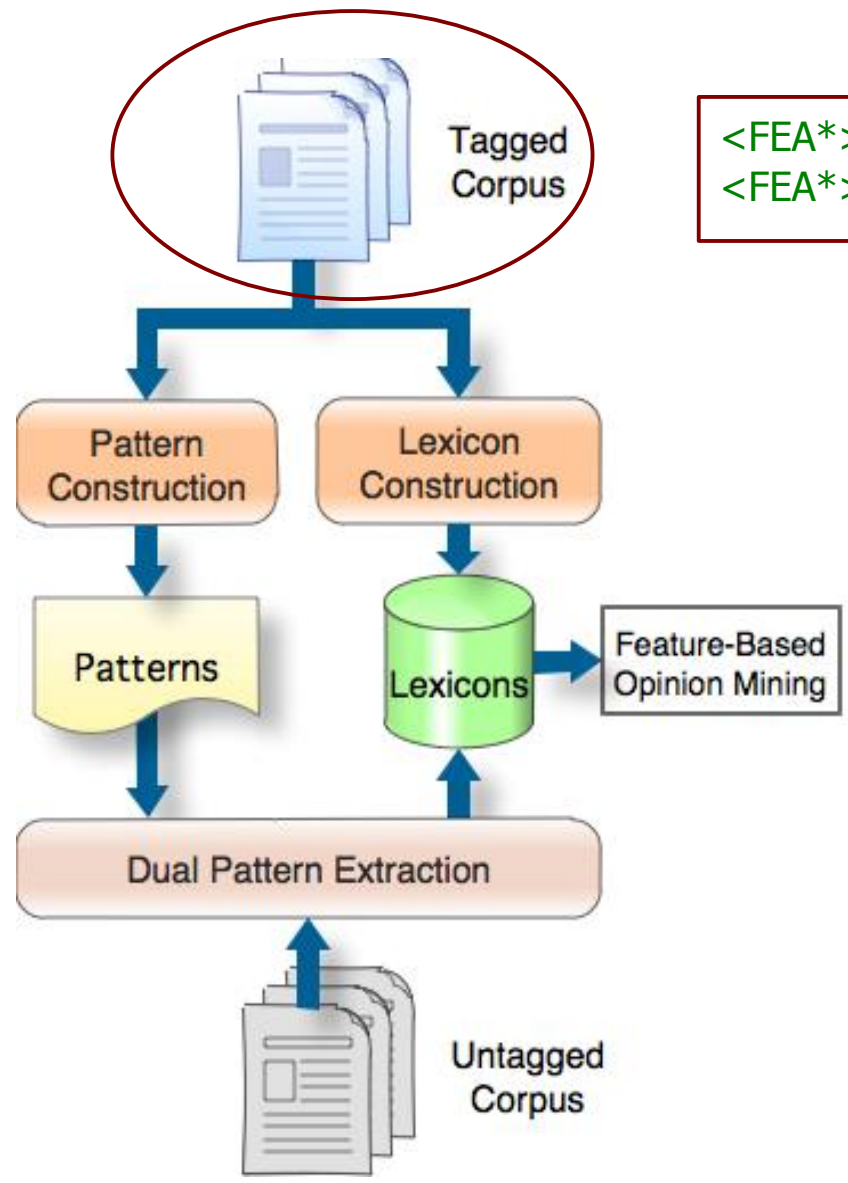
S-Sense: social sensing tool



Dual pattern extraction

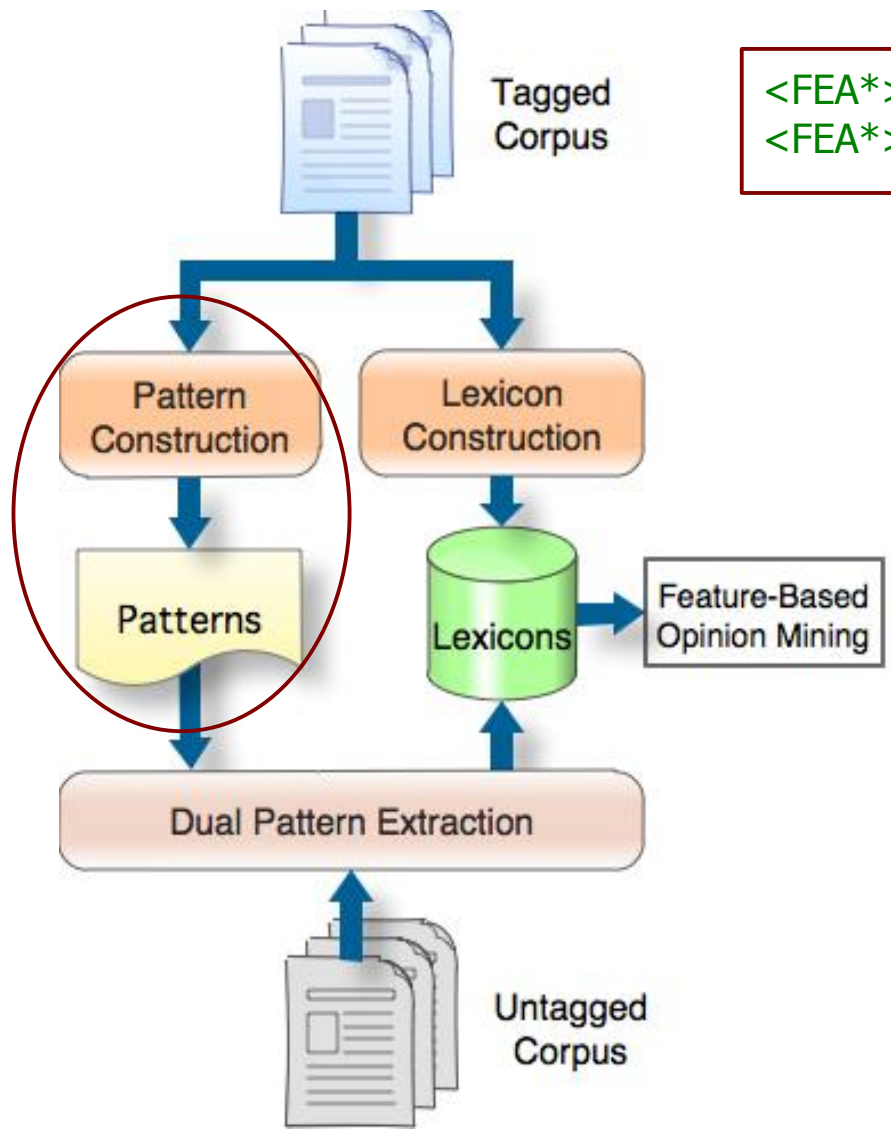


Dual pattern extraction



<FEA*>พนักงาน</FEA*><POL>เป็นมิตร<POL><DEG>มาก</DEG>
 <FEA*>employee</FEA*><POL>friendly<POL><DEG>very</DEG>

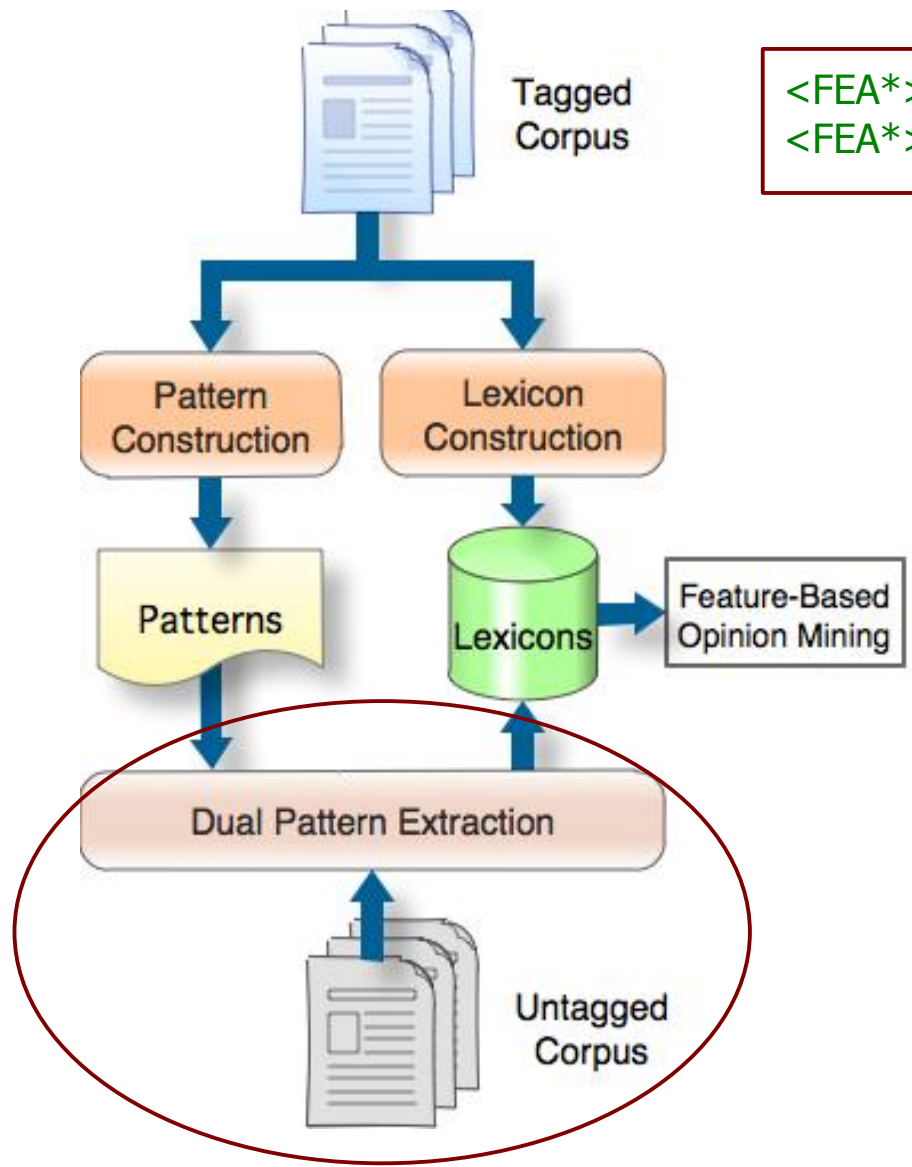
Dual pattern extraction



<FEA*>พนักงาน</FEA*><POL>เป็นมิตร<POL><DEG>มาก</DEG>
 <FEA*>employee</FEA*><POL>friendly<POL><DEG>very</DEG>

<FEA*><POL><DEG>

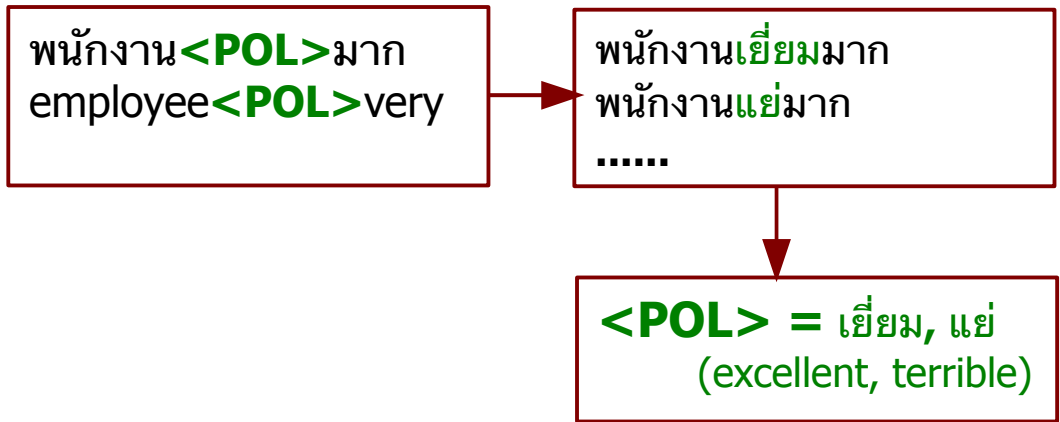
Dual pattern extraction



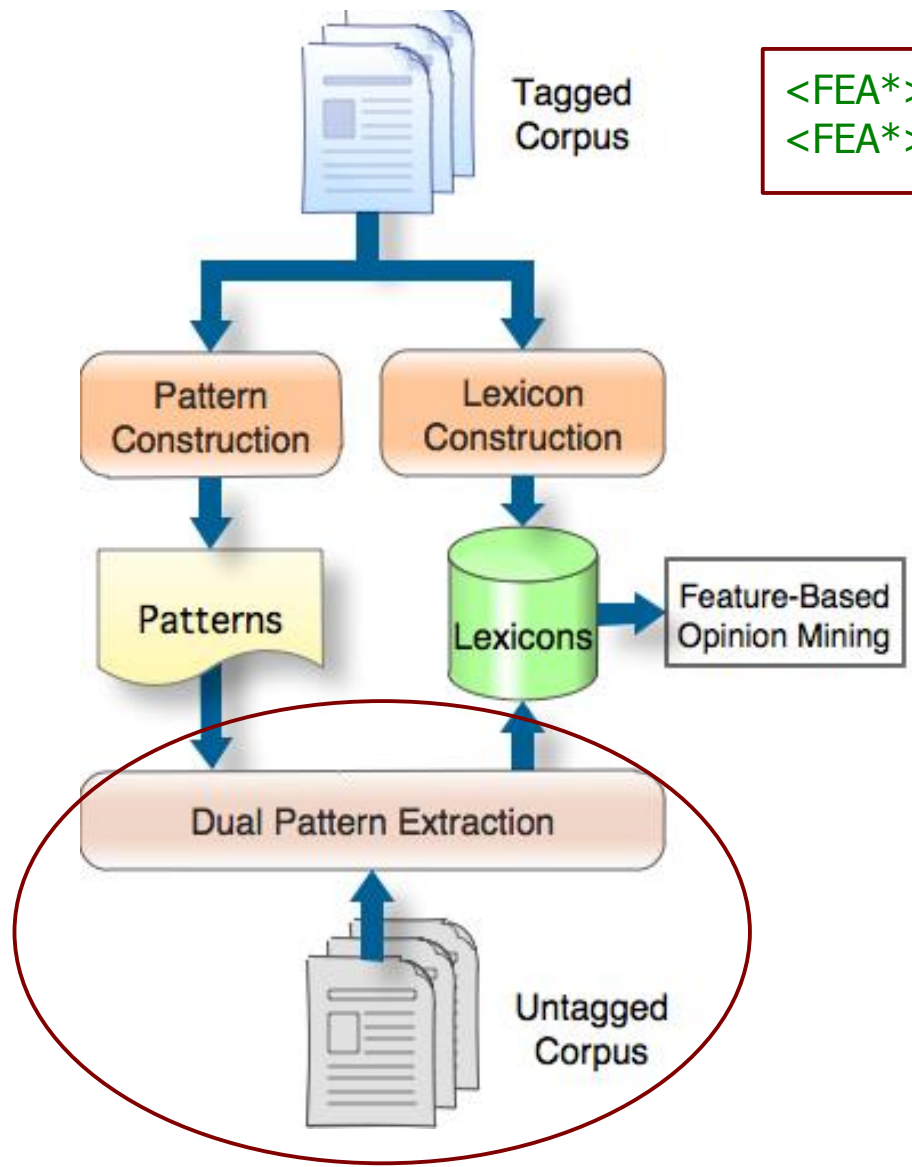
<FEA*>พนักงาน</FEA*><POL>เป็นมิตร<POL><DEG>มาก</DEG>
 <FEA*>employee</FEA*><POL>friendly<POL><DEG>very</DEG>

<FEA*><POL><DEG>

(1) Extract new polar words



Dual pattern extraction



<FEA*>พนักงาน</FEA*><POL>เป็นมิตร<POL><DEG>มาก</DEG>
 <FEA*>employee</FEA*><POL>friendly<POL><DEG>very</DEG>

<FEA*><POL><DEG>

(2) Extract new sub-features

<FEA*>เป็นมิตรมาก
 <FEA*>ดูแลดีมาก
 <FEA*>แย่มาก

พนักงานต้อนรับดูแลดีมาก
 พนักงานขับรถแย่มาก

<FEA*> = พนักงานต้อนรับ, พนักงานขับรถ
 (receptionist, driver)

A Case Study on Hotel Reviews

- Agoda website: www.agoda.com

The screenshot displays the Agoda website interface. At the top, the Agoda logo is on the left, and navigation links like 'Join', 'My Account', and 'Learn More' are on the right. A search bar with 'Agoda Search by Google™' and a 'search' button is present. The language is set to 'English'. Below the search bar, there are navigation tabs for 'Hotels', 'About Agoda', 'Rewards', 'Special Offers', and '24/7 Help'. A breadcrumb trail shows the current location: 'Home → World → Asia → Thailand → Koh Chang / Trad → Koh Chang Resortel'. The main content area features a 'Search All Hotels' sidebar on the left with filters for Country (Thailand), City (Koh Chang / Trad), and arrival/departure dates. The main search results show 'Koh Chang Resortel, Koh Chang / Trad' with a 'book now!' button. Below the hotel name are tabs for 'Hotel Info', 'Photos', and 'Customer Reviews [11]'. A photo of the resort is shown, along with a 3-star rating, location 'White Sand Beach', and address '13/9 M.4 T. Koh Chang, Koh Chang / Trad, Thailand'. At the bottom, there is a 'Check rates & availability at this hotel' section with a 'best price guarantee' icon and a 'check rates' button. The arrival and departure dates are set to June 18 and June 21, 2010, respectively, with 1 room, 2 adults, and 0 children.

A Case Study on Hotel Reviews

ผัดหวังกับ Koh Change Resortel

Guest Rating: **3.7**

Rachit B.

 Thailand

Other

Date of Stay: May 03, 2010

Value for Money:



Location:



Staff Performance:



Hotel Condition/Cleanliness:



Room Comfort/Standard:



Food/Dining:



+ ทำเลที่ตั้งอยู่ไม่ไกลจนเกินไปจากท่าเรือเฟอร์รี่ รวมถึงมีสระว่ายน้ำขนาดใหญ่

- โรงแรมมีข้อด้อยในหลายๆ จุด ยกตัวอย่างเช่น หาดเป็นหาดหิน ไม่สามารถลงไปเล่นน้ำได้, สภาพแวดล้อมภายในโรงแรม ซาดการดูแลที่ดี ทำให้ดูทรุดโทรม, สภาพห้องพักไม่ดี มีกลิ่นอับ ไม่สะอาด (เจอแมลงสาบในห้องพักด้วย), อาหารเข้าธรรมดาตามาก

โรงแรมมีทำเลที่ตั้งที่ใช้ได้พอสมควร อยู่ไม่ไกลจากท่าเรือเฟอร์รี่ แม้ว่าหาดจะเป็นหาดหิน ไม่สามารถเล่นน้ำได้ แต่ก็ถือว่ามีวิวที่ตีพอสมควร เพียงแต่ว่าตัวสภาพแวดล้อมต่างๆ ภายในโรงแรม ดูเหมือนว่าไม่ได้รับการดูแลและบูรณะให้อยู่ในสภาพดี ทำให้สิ่งแวดลอมต่างๆ ดูทรุดโทรม ไม่ว่าจะเป็นบริเวณสวนต่างๆ ที่ตกแต่งอยู่รอบๆ โรงแรม สภาพตัวอาคาร ไปจนถึงภายในห้องพัก และที่สำคัญห้องพักยังสภาพแยกว่าที่คาดคิดไว้มาก มีกลิ่นอับและไม่สะอาด รวมถึงซาดการดูแลสิ่งอำนวยความสะดวกต่างๆ ภายในห้องพัก

Experiment

- Using the tagged corpus and the extracted lexicons, we construct the most frequently occurred patterns.
 - For two main features, **breakfast** and **service**, the numbers of tagged reviews for each feature are 301 and 831, respectively.
 - We randomly split the corpus into 80% as training set and 20% as test set. We only consider the patterns which contain both features and polar words.

Example: Extracted patterns for "service"

No.	Top-ranked "service" patterns
1	<p><FEA*><POL> <พนักงาน><เป็นมิตร> <employee><friendly></p>
2	<p><FEA><POL> <บริการ><ประทับใจ> <service><impressive></p>
3	<p><FEA*><FEA><POL> <พนักงานขับรถ><บริการ><ดี> <driver><service><good></p>
4	<p><FEA*><FEA><POL><DEG> <พนักงาน><บริการ><สุภาพ><มาก> <employee><service><polite><very></p>
5	<p><FEA*><OTH><POL> <พนักงาน><ทุกคน><ยิ้มแย้มแจ่มใส> <employee><everyone><smiling></p>

Extracted lexicons: **service**

Lexicons	Examples
Features (FEA)	บริการ (service), service, เซอร์วิส (service)
Sub-features (FEA*)	bell boy, reception, receptionist, คนขับรถ (driver), พนักงานขับรถ (driver), เจ้าหน้าที่รักษาความปลอดภัย (security guard), รปภ. (security guard), บริกร (waiter), แผนกต้อนรับ (reception), แม่บ้าน (maid) พนักงานต้อนรับ (receptionist), พนักงานขนกระเป๋า (concierge)
Polar words (POL)	มีน้ำใจ (considerate)[+], สะอาด (clean)[+], ใจดี (kind)[+], ตูแฉดตลอด (courteous)[+], กระตือรือร้น (eagerly)[+], อบอุ่น (warm)[+], ช่วยเหลือ (helpful)[+], จริงใจ (sincere)[+], เอาใจใส่ (courteous)[+], น่ารัก (lovely)[+], เป็นกันเอง (friendly)[+], นิสัยดี (nice)[+], ช้า (slow)[-], คุยโม้โอ้อวด (arrogant)[-] ขาดคุณธรรม (deceitful)[-], จู้จี้จ้าน (nosy)[-] เฉยเมย (inattentive)[-], หงุดหงิด (grumpy)[-] มือไม้แข็ง (impolite)[-]

Example: Extracted patterns for "breakfast"

No.	Top-ranked "breakfast" patterns
1	<p><FEA><POL> <อาหารเช้า><อร่อย> <breakfast><delicious></p>
2	<p><FEA><AUX><POL> <อาหารเช้า><ก็><ดีเยี่ยม> <breakfast><"kor"><excellent></p>
3	<p><FEA><POL><DEG> <อาหารเช้า><แย่มาก> <breakfast><terrible><very></p>
4	<p><FEA*><POL> <ถ้วยกาแฟ><สกปรก> <coffee cup><dirty></p>
5	<p><FEA><POL><POL> <อาหารเช้า><สะอาด><มีคุณภาพ> <breakfast><clean><with quality></p>

Extracted lexicons: Breakfast

Lexicons	Examples
Features (FEA)	ABF, Breakfast, เบรคฟาสต์ (Breakfast), อาหารเช้า (Breakfast)
Sub-features (FEA*)	เมนู (menu), รสชาติ (taste), ห้องอาหาร (restaurant), คุณภาพ (quality), บริการ (service), ปริมาณ (quantity), ขนมปัง (bread), กาแฟ (coffee), ที่นั่ง (seat), พนักงาน (waiter)
Polar words (POL)	ดี (good)[+], หลากหลาย (various)[+], สด (fresh)[+], มีคุณภาพ (with quality)[+], สะอาด (clean)[+], ประทับใจ (impressive)[+], แย่ (terrible)[-], ไม่ได้เรื่อง (awful)[-], จำเจ (repeating)[-], ต้องรอคิว (long line)[-], น้อย (little)[-], น่าเบื่อ (boring)[-], คับแคบ (confined)[-]

Experimental results

Feature	Accuracy (%)	
	FEA*	POL
Breakfast	80.00	95.74
Service	82.56	89.29

- Extracting **polar words** is relatively easier since the patterns are simpler and more obvious, especially for **breakfast** feature.

– e.g., <FEA*><POL>

กาแฟอร่อย

coffee delicious

Opinion Mining On Hotel Reviews

คุณลักษณะของโรงแรม

เชียงใหม่

- เรียงตามจำนวนความคิดเห็น
 เรียงตามชื่อ

รายชื่อโรงแรมและจำนวนความคิดเห็น

โรงแรมเขื่อนทارا ดวงตะวัน(131)

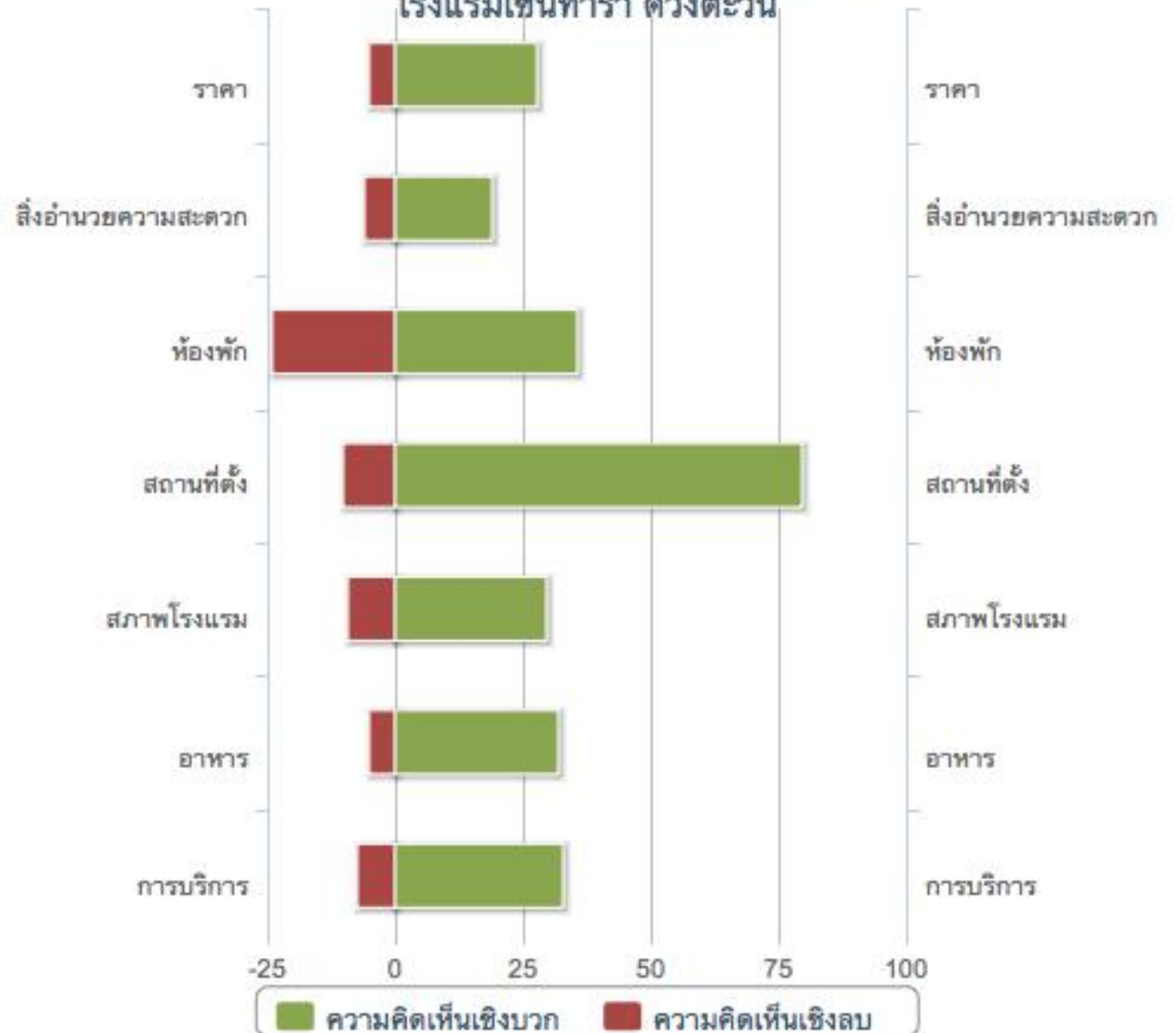
แสดงผลแบบกราฟิก

แสดงผลแบบข้อความ

- แสดงคำบอกคุณลักษณะ
 แสดงคำระบุข้อความความคิดเห็น

เปรียบเทียบคุณลักษณะโรงแรม

กราฟแสดงภาพรวมในแต่ละคุณลักษณะของ
โรงแรมเขื่อนทารา ดวงตะวัน



Opinion Mining On Hotel Reviews

คุณลักษณะของโรงแรม

เชียงใหม่

เรียงตามจำนวนความคิดเห็น

เรียงตามชื่อ

รายชื่อโรงแรมและจำนวนความคิดเห็น

โรงแรมเขื่อนทارا ดวงตะวัน(131)

แสดงผลแบบกราฟิก

แสดงผลแบบข้อความ

แสดงคำบอกคุณลักษณะ

แสดงคำระบุข้อความความคิดเห็น

เปรียบเทียบคุณลักษณะโรงแรม

พนักงานบริการดี ดูความคิดเห็นเดิม

+

ทัวร์ที่ถูกและบริการดี ดูความคิดเห็นเดิม

+

โรงแรมไม่มีใครมาสนใจ ดูความคิดเห็นเดิม

+

การบริการที่แตกต่างพอสมควร ดูความคิดเห็นเดิม

+/-

บริการไม่ทั่วถึง ดูความคิดเห็นเดิม

-

พนักงานบางคนยังไม่ค่อยสนใจ ดูความคิดเห็นเดิม

-

ไม่ได้เห็นพนักงานยิ้มแย้มแจ่มใสบริการดี ดูความคิดเห็นเดิม

-

พนักงานต้อนรับไม่ค่อยยิ้มแย้มแจ่มใส ดูความคิดเห็นเดิม

-

นาฬิกาและไม่มีมาให้คำแนะนำ ดูความคิดเห็นเดิม

-

Opinion Mining On Hotel Reviews

เชียงใหม่

- เรียงตามจำนวนความคิดเห็น
- เรียงตามชื่อ

โรงแรมที่ต้องการเปรียบเทียบและจำนวนความคิดเห็น

โรงแรมเซ็นทารา ดวงตะวัน(131)

แสดงผลแบบข้อความ

โรงแรมโลตัส ปางสวนแก้ว(74)

แสดงผลแบบข้อความ

โรงแรมอิมพีเรียล แม่ปิ้ง(90)

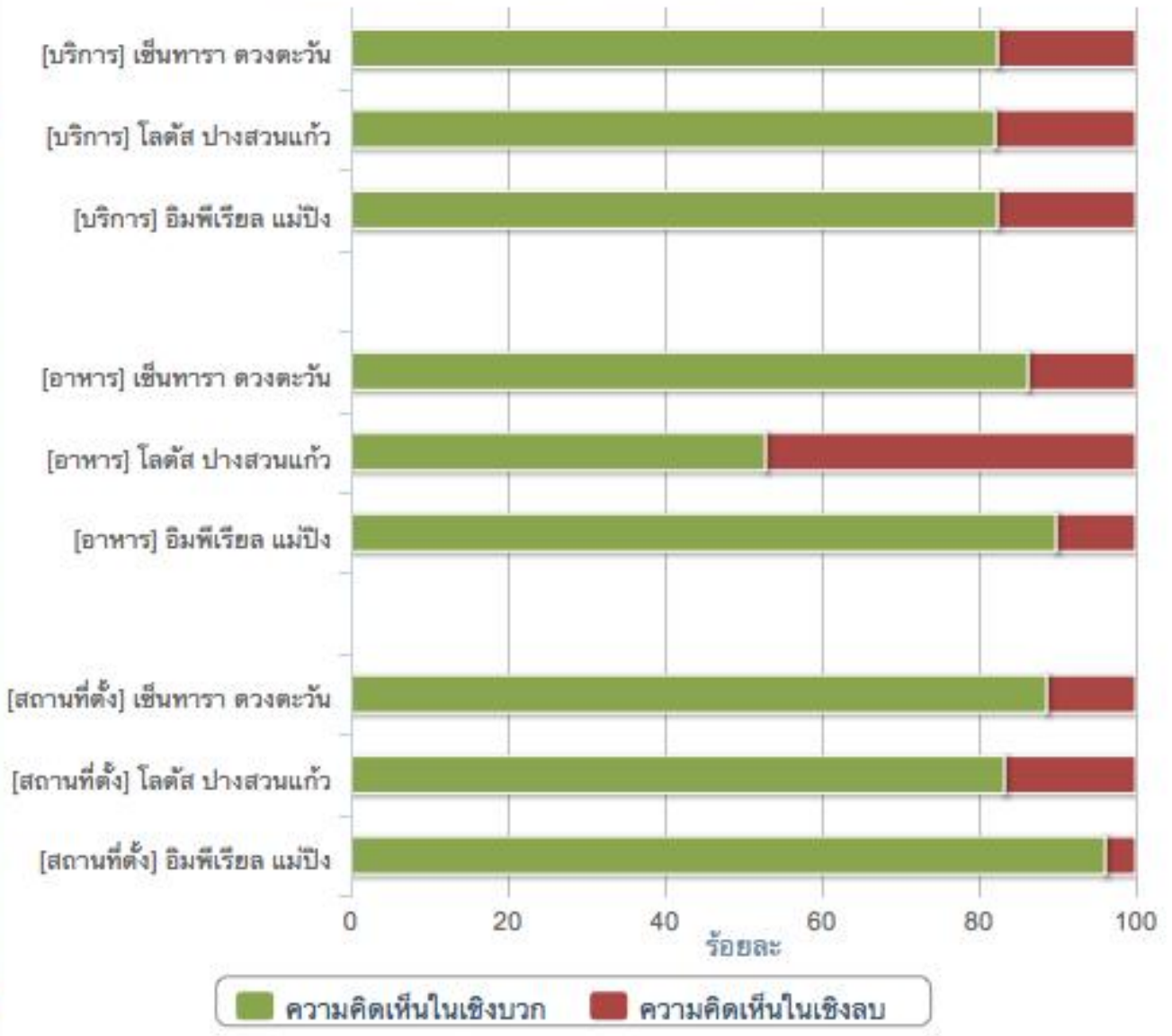
แสดงผลแบบข้อความ

เปรียบเทียบตามคุณลักษณะของโรงแรม (ครั้งละไม่เกิน 3 คุณลักษณะ)

- การบริการ
- อาหาร
- สภาพโรงแรม
- สถานที่ตั้ง
- ห้องพัก
- สิ่งอำนวยความสะดวก
- ราคา

เลือก

กราฟเปรียบเทียบภาพรวมคุณลักษณะของแต่ละโรงแรม





OPINION MINING ON MOBILE SERVICE

เครือข่าย

เปรียบเทียบเครือข่าย

กรุณาเลือกคุณสมบัติ:

- การบริการลูกค้า บริการเสริม โปรโมชั่น เครือข่าย ค่าบริการ อื่นๆ

กราฟเปรียบเทียบภาพรวมคุณลักษณะของแต่ละเครือข่าย



การบริการลูกค้า

ทำซิมทรูจตทะเบียนหายไปทำใหม่ที่shopมีค่าใช้จ่าย
ไหมครับ



CCTrueMoveแจ้งมีปัญหาไม่สามารถใช้BBService
ได้บริการช่วย



โทรขอพบบริการทำไมถึงช่วยแตกแบบนี้ล่ะครับ



ติดต่อพนักงานCallCenterของTrueยากมาก



บริการเสริม

โปรโมชั่น

เครือข่าย

ค่าบริการ

อื่นๆ

การบริการลูกค้า	
บริการเสริม	
โปรโมชั่น	
เครือข่าย	
ทรูมูฟชานา.....โทรออกไม่ได้ค่ะด่วนเลยคะ	-
โทรเข้า1331ทรูมูฟไม่มีคนรับสายนานมากคะ	-
สัญญาณมือถือขาดเป็นระยะตอนใช้บริการแอร์พอร์ตลิงค์	-
Trueonlinebillingระบบมีปัญหาไม่สามารถล็อกอินได้ครับ	-
ถนนเจริญกรุงแถวตลาดน้อยสัญญาณTrumoveห่วย	-
ค่าบริการ	
อื่นๆ	

Thank you for your attention

Q&A

References

- Haruechaiyasak C., A. Kongthon, P. Palingoon, and C. Sangkeettrakarn, "Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews," In Proc. of the 8th Workshop on Asian Language Resources, pp. 64–71, 2010.
- Kongthon, A., N. Angkawattanawit, C. Sangkeettrakarn, P. Palingoon, and C. Haruechaiyasak, "Using an opinion mining approach to exploit Web content in order to improve customer relationship management," *In Proc. of the 2010 Portland International Conference on Management of Engineering and Technology (PICMET 2010)*, pp. 1355-1360, 2010.
- Kongthon, A.; Haruechaiyasak, C.; Sangkeettrakarn, C.; Palingoon, P.; Wunnasri, W.; "HotelOpinion: An opinion mining system on hotel reviews in Thailand", *In Proc. of the 2011 Portland International Conference on Management of Engineering and Technology (PICMET 2011)*, 2011.