NUS - Extreme - Tsinghua

Challenges in Real-time Social Media Analytics:

What has changed over the last 5 years?

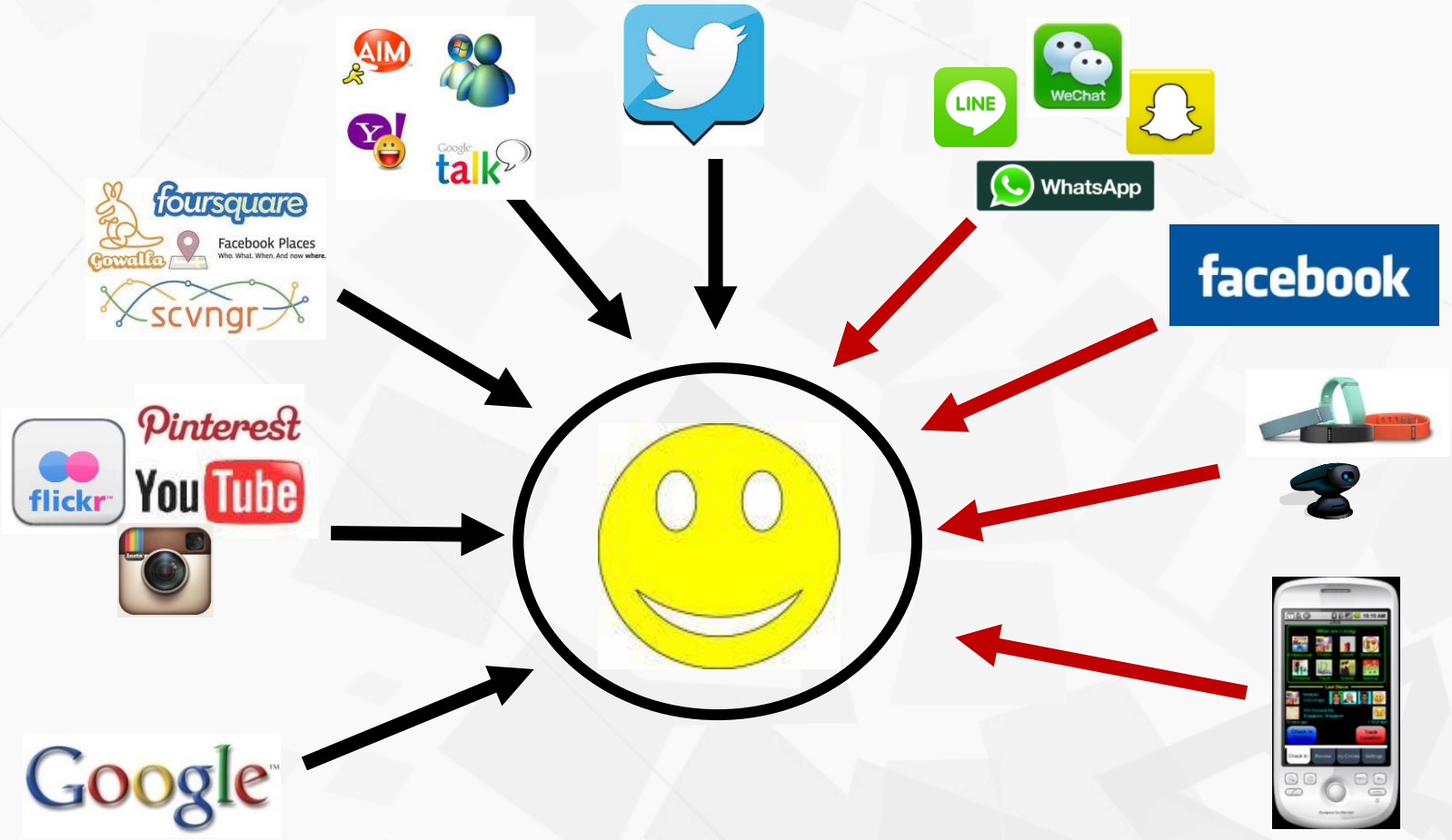
CHUA Tat-Seng

National University of Singapore

OUTLINE

- **The Live Social Observatory**
- **The Technologies behind the System**
- **Next Generation Systems and Challenges**
- **Other Related Research**
- **Summary**

Users at Center of Social Media Environment

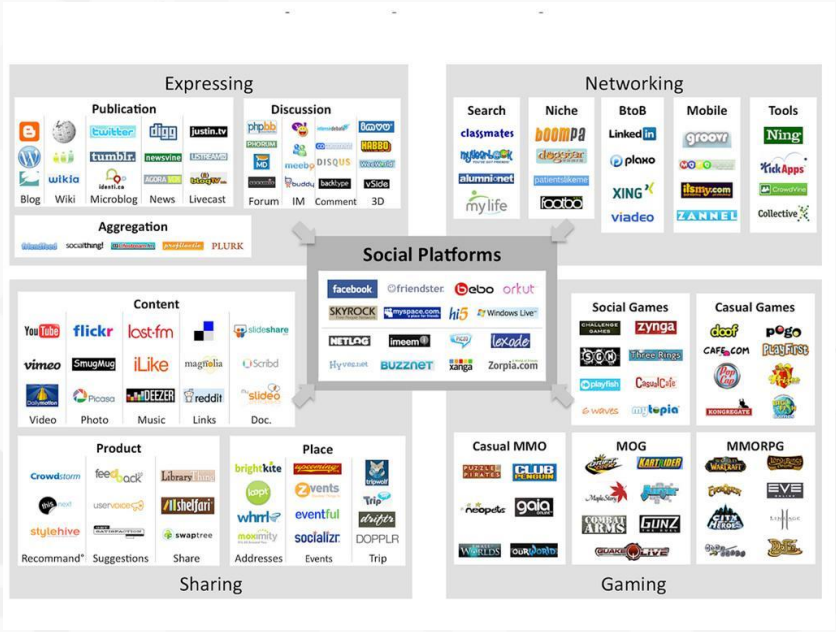


Functions of Social Media Platforms

- From users' perspective
  - Communications: sharing, interacting, keeping up-to-date.. with friends
  - Expression: air/project views
  - Self-preservation: Wellness, exercise, self-improvements
  - Local communities

- From platform providers' perspective
  - Attract and retain users; enrich contents; monetization
  - Offer innovative/fun services; improve user engagement
  - Co-creation of contents with users

Key Difference between Social Media Platform and Broadcast Media



Broadcast Media
Filter->Publish

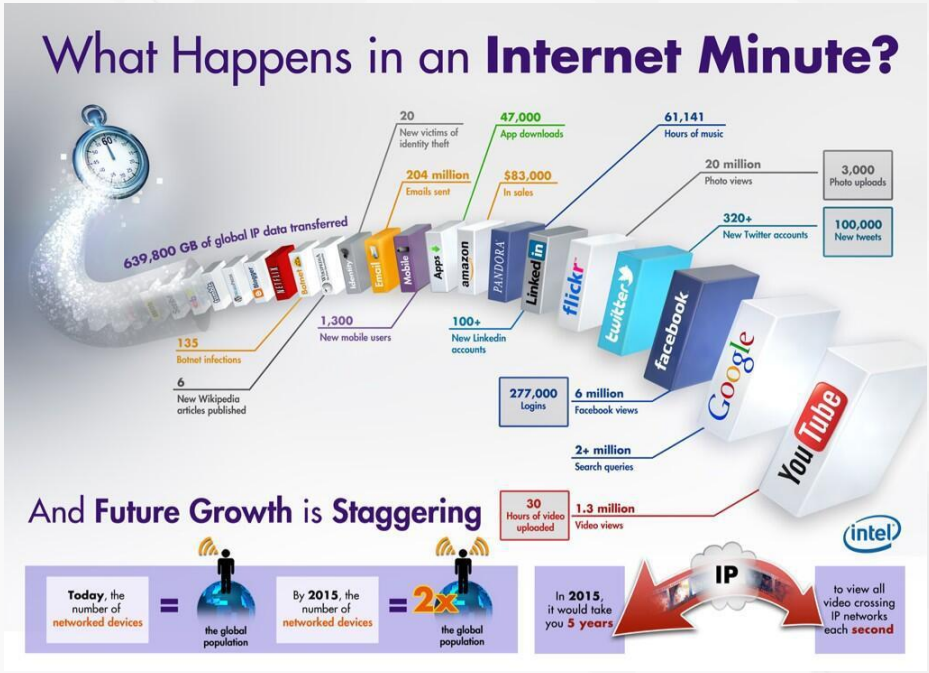
Social Media
Publish->Filter

Offers curated content

Offers no content;
Contents all created by Users

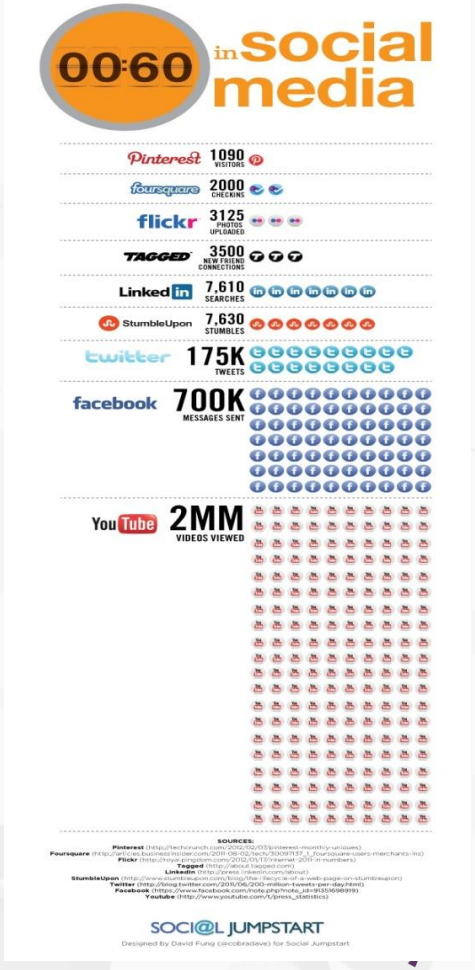
Functions of Social Media Platforms

- Lead to 3 key sources of live info streams:
  - Spontaneous User-Generated Contents (UGC)
  - Device-Generated Contents (DGC)
  - Structured Database Contents (SDC)



NExT Research Center
(http://next.comp.nus.edu.sg)

- NExT: NUS-Tsinghua Center on Extreme Search
  - Initiated in May 2010 with a 5-Year Grant of S$10 million from MDA/NRF of Singapore
  - Among the first to examine big data analytics
- Vision: To access live information that is not easily available on the current Web



Overview of Research

From Unstructured Info into Structured Knowledge

Gather Data from Multiple UGC Sources

Type 1: Text Contents: Tweet, cQA, User Comments

- Twitters
- cQA/ Forum
- Social News

Type 2: Images/ Videos

- Images/ Videos

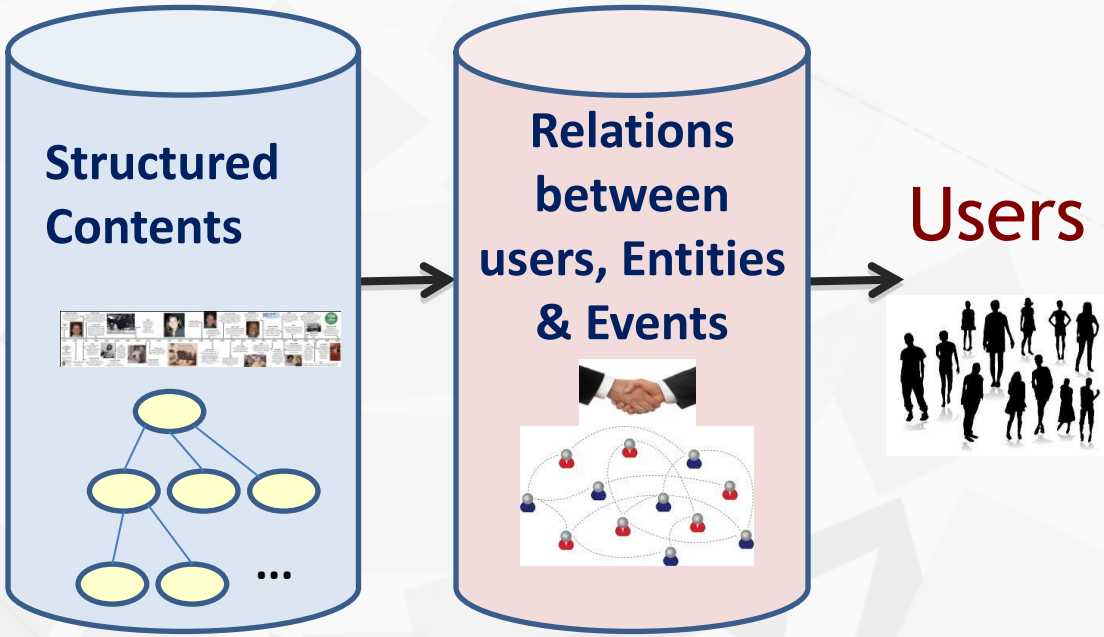
Type 3: Location Services/ Apps

- Check-in Venues/ Mobile Apps

Type 4: Structured Data

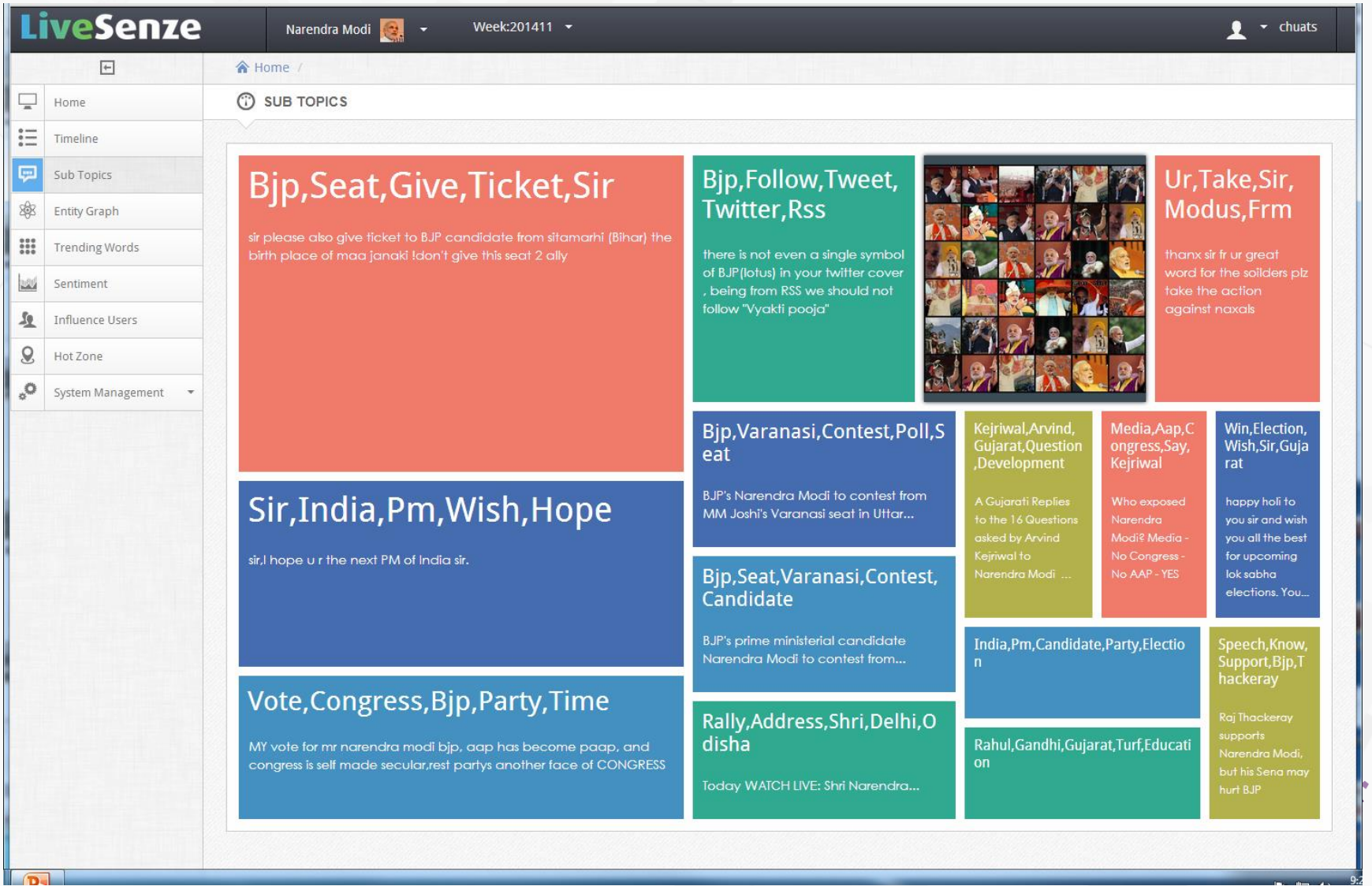
- Str. Data

Analysis to infer key entities, events, users and their relationships



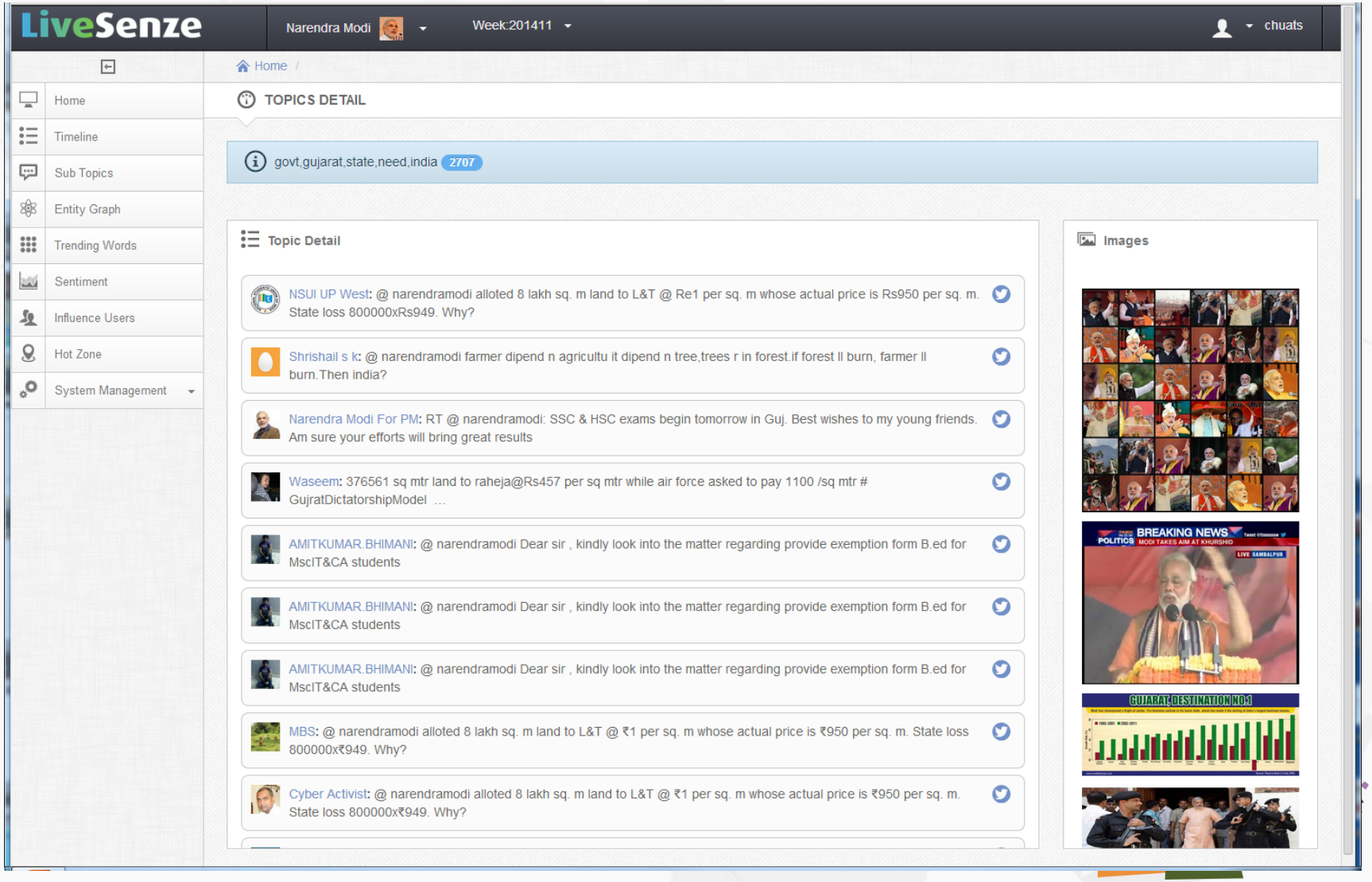
Platform as a Service: Indian Elections

Sub-Events Around an Entity (Modi)



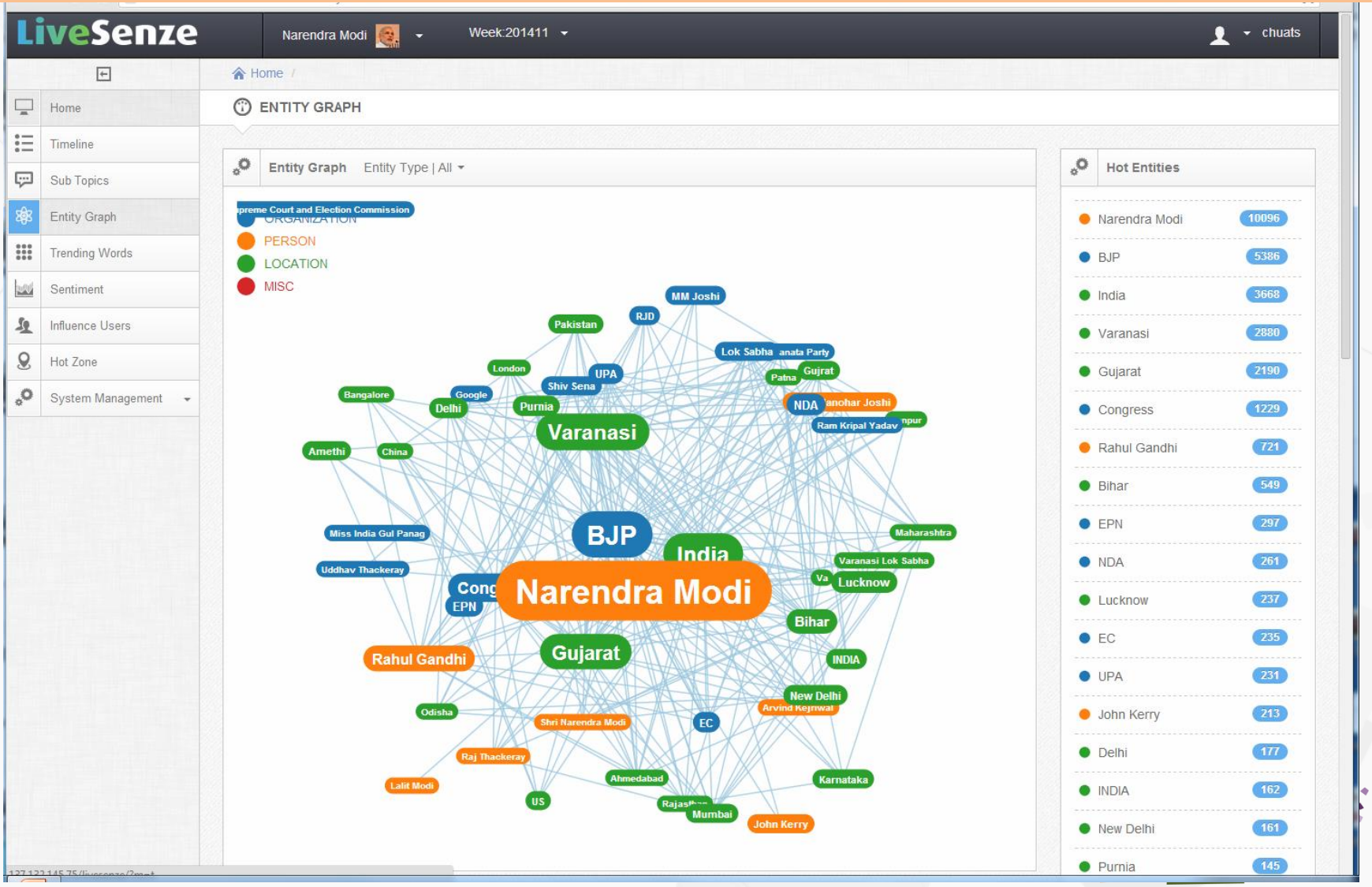
Platform as a Service: Indian Elections

Details of a Sub-Event (Modi)



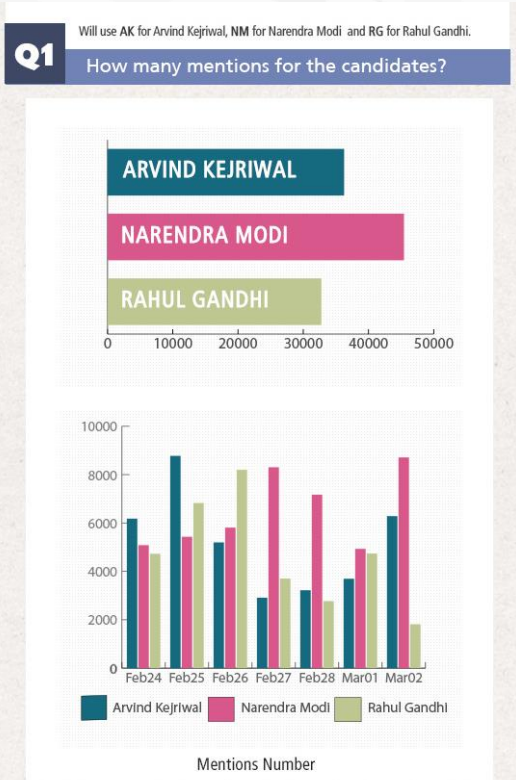
Platform as a Service: Indian Elections

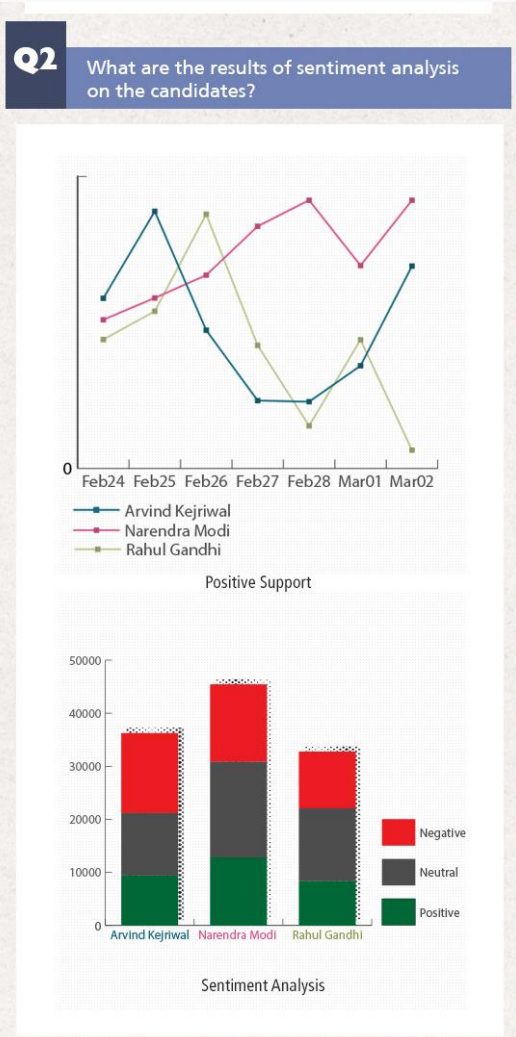
Relations between Entities (Modi)



Platform as a Service: Indian Elections

From Data to Event to Reports





Insights Gained from Social Media Analytics

Events, topics, Users

and

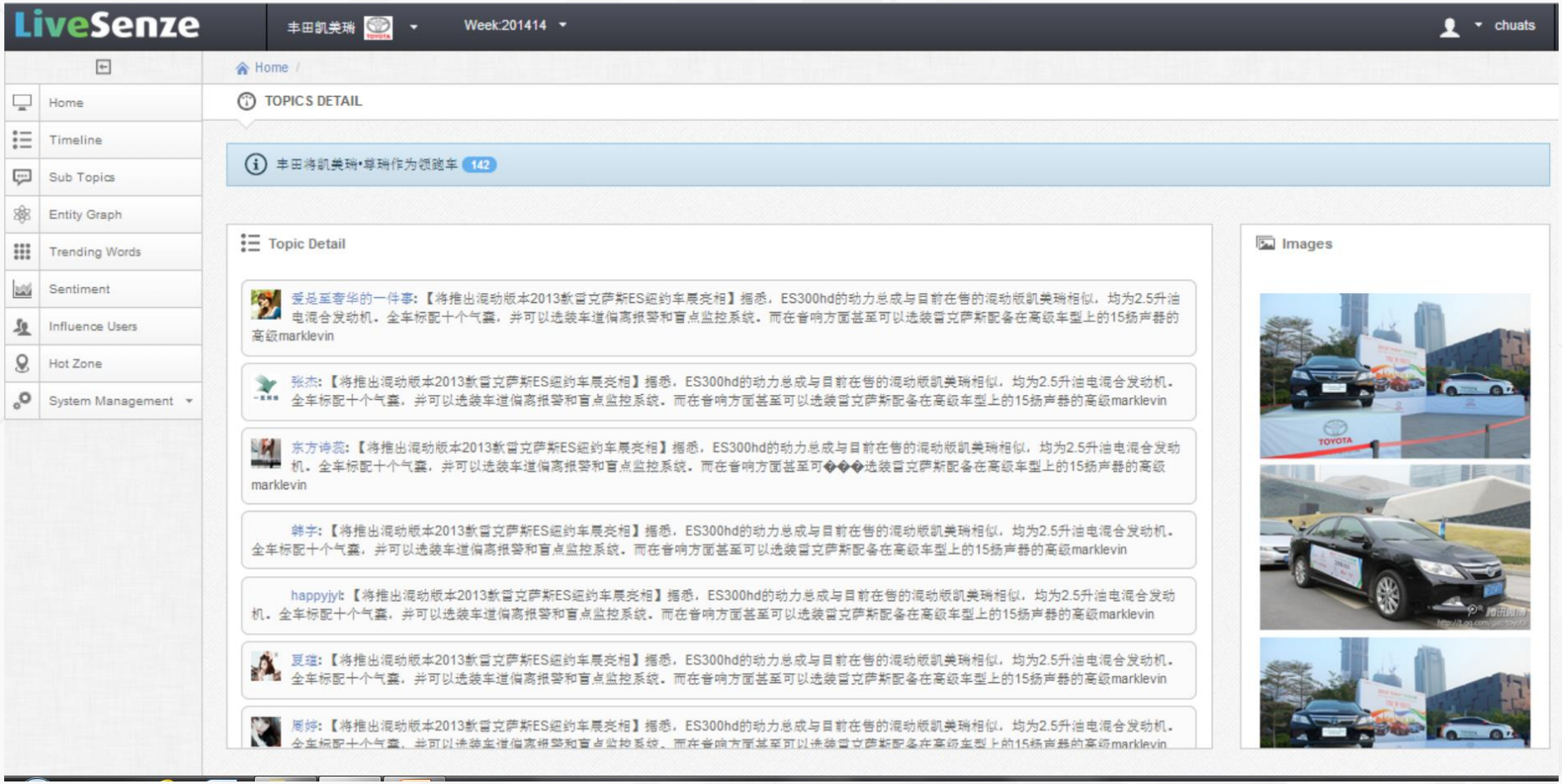
their relationships

Example 2: Automobile Sales in China

Sub-Events related to Camry

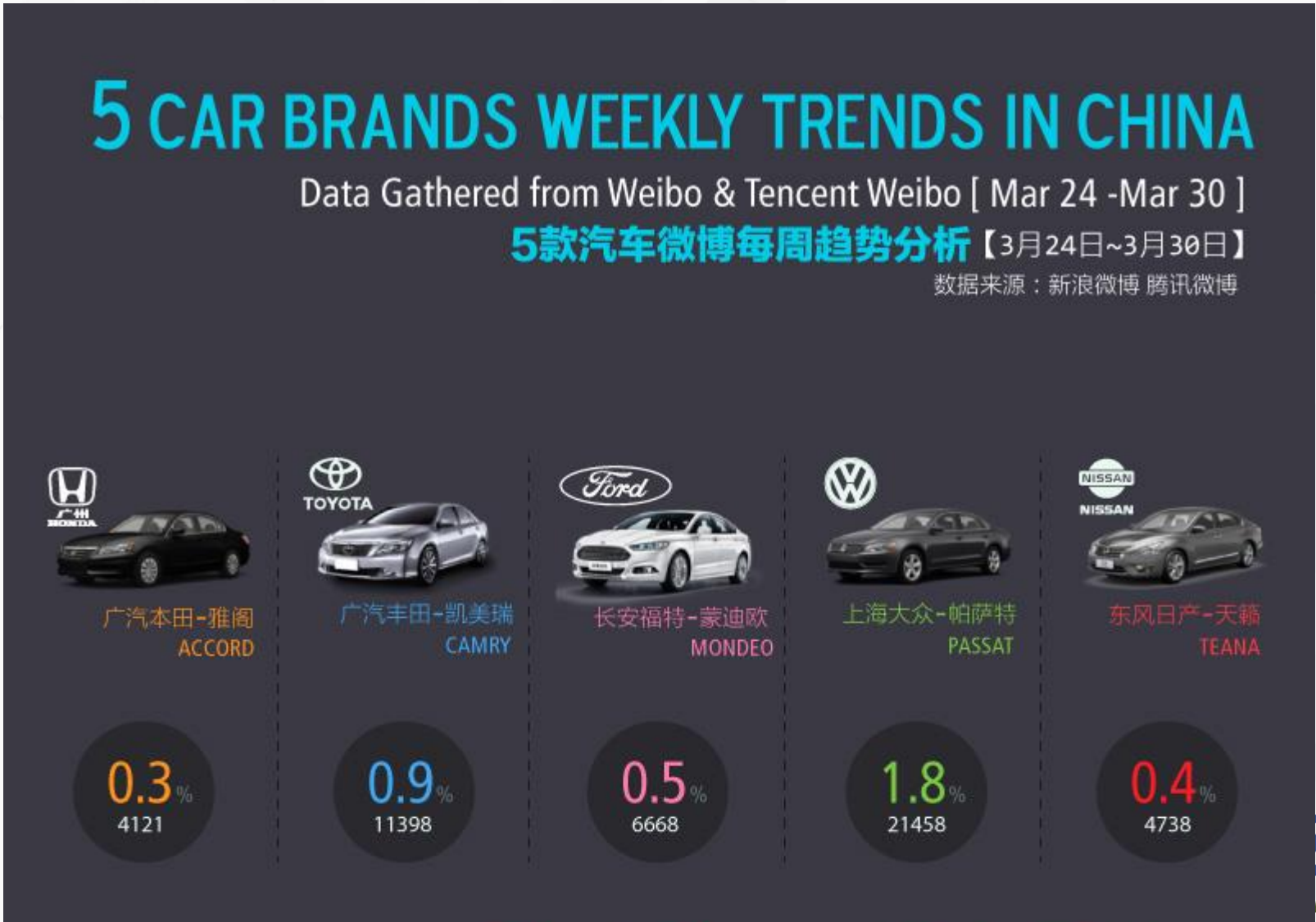
Example 2: Automobile Sales in China

Details of a Sub-Event



Example 2: Automobile Sales in China

From Data to Event to Reports



OUTLINE

- The Live Social Observatory
- The Technologies behind the System
- Next Generation Systems and Challenges
- Other Related Research
- Summary

Key Challenges

- As decided 5 years ago, the key challenges are Live, Big Data and 3M

What are they?

What is Big Data?

Volume	• From TB to PB or ZB • Fast growing
Variety	• Wide variety (text, image, video, docs) • Multi-Sources
Velocity	• Data stream vs batch • Real time analytics
Value	• Reporting and statistics • Trend and model prediction • Deep analytics

Big Data vs. Live

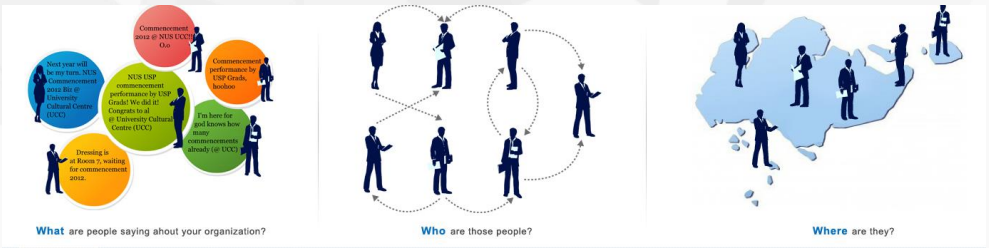
- Big Data is not about big, but about online and live
  - It is easy to handle (static) big data
  - But difficult to handle big live data streams
- Big Data Analytics
  - Data online and Live
  - Gather and Filter data (not Store, Sample and Search)
  - Feedback (close-loop) and React

Key Challenges in Live Social Media Analysis

ISSUES	KEY TECHNOLOGIES
Live Data	How to gather "representative" live data from multiple sources w.r.t. an entity
Multimodal	To analyze multimedia data (text, images, videos, locations and users)
Multilingual	Deep NLP for multiple languages (English, Chinese, …)
Events	To detect and track all events happening around any entity
Prediction	Predict future trends and detect hot posts / events
Prescription	Generate comparative reports, user demography, and prescribe actions
Automated	For scalability, the entire system should be fully automated

Live Event Detection & Tracking
from live social media streams

- Aims: monitor pulses of organization/ event/ location/ people:
  - What is going on; what is hot (Event detection)
  - Who is involved; who says what (User community detection)
  - Where are the users (Geo location detection)
  - What do users share in images/videos (Image Content Analysis)
  - Complain, admiration (Sentiment analysis)
- Main challenges
  - Huge amount of tweets – filtering problem
  - Dynamic & noisy vocabulary – conversational nature f social media
  - Dynamic user community
  - Media content analysis
  - Scalable distributed architecture



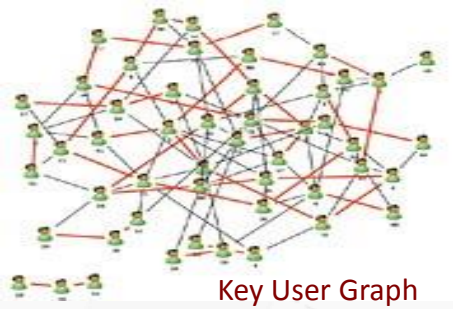
1. Gathering of Representative Data -1

- Key issue: How to deal with Live/Dynamic Data?
  - Some studies show that there will be 6 times more data if crawled in real-time
- 1st Source: Known accounts- Most organizations have substantial social media presence in building customer relations
  - Reliable but from Organization's point of views

- 2nd Source: Dynamic Keywords- Generate dynamic keywords to address dynamic vocabulary issue

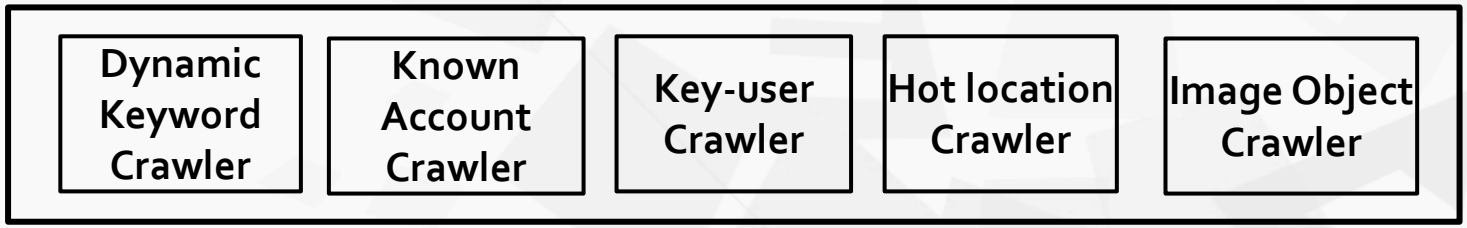
1. Gathering of Representative Data -2

- 3rd Source: Active Users- Identify dynamic user community
  - Needs to remove bot users



Key User Graph

- Overall: Reliable Data Collection Strategy
  - Employ multi-faceted approach to gather representative data
  - Simulate real user login behaviors
  - Our strategy ensures high relevance, coverage and diversity



- Results: High recall, but too much noise

2. Mutilingual Linguistic Analysis

(in English and Chinese)

- Key Phrase Extraction
- Named Entity (NE) Extraction
  - Identify name of people, organization, location and products etc.
  - May need finer grained NE's in vertical domains

One Approach to Phrase Extraction

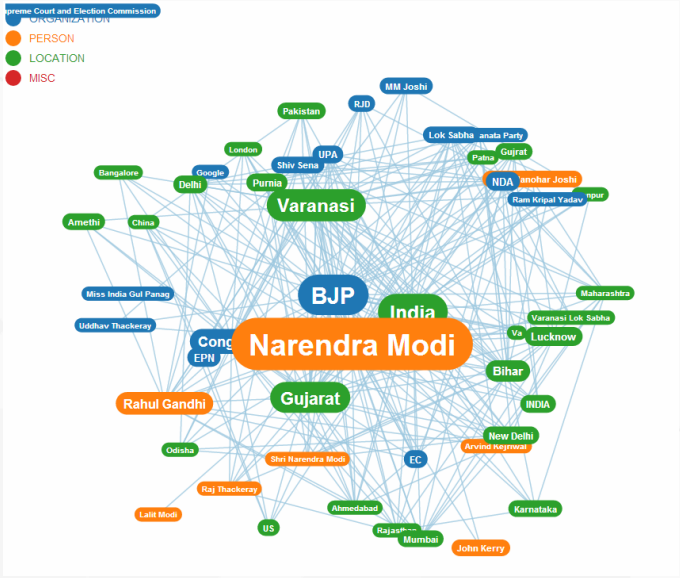
- Combine Short Text and Long Text
  - Short Text: Tweet, Weibo, Review from e-commerce website
  - Long Text: Wechat, News
- Procedure
  - Syntactic & semantic analysis
    - Filter potential bag of words (using LTP tools)
  - Keyword analysis
    - Generate word weights from different data source [Chen et al. ESWA2014]
  - Calculate Phrase
    - Keep top-score bag of words as phrase

Entity Extraction

- Prior Entity Database
  - Wiki, Baike (combine entity alias via link redirection)
  - Sogou Input lexicon database
  - Geographic entity database from Tsinghua University (Guoliang Li's group)
- Features
  - Word level: unigram, bigram
  - Part-of-Speech level: postag-unigram, postag-bigram
- Algorithms
  - Online learning algorithm
  - Semi-supervised learning algorithm

Multilingual Linguistic Analysis

- Entity Relations
  - Infer from co-occurrences of entities in messages
- Sentiment Analysis
  - Determine sentiment of posts at event, sub-events and entity level
- Classifier/ Filterer for noise
  - Employ text, social relations and temporal relations
  - To filter irrelevant contents
  - To filter noise



Filter Irrelevant Content

- Consider 4 Factors:
- Content: For each tweet, perform analysis to:
  - Informal language normalization
  - Irrelevant text tokens filtering
  - Identify evolving text features
- Location information
- User Social Relationships
- User Tweeting Tendencies Over Time

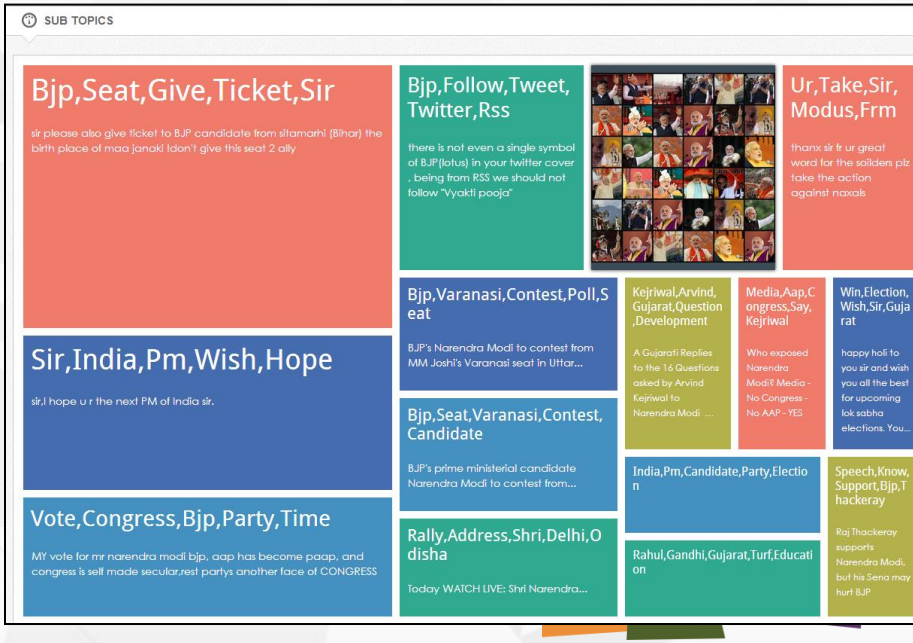
- Train a classifier based on training set using, say, kNN or SVM.
- For a new tweet $\underline{T}_{new}$, use the classifier to assign the class(es) of $\underline{T}_{new}$.

Noise Removal

- Identify Spam for Short Text
  - Filter out Tweet/Weibo which contain apparent commercial intent
  - Build a manually labeled dataset (10000+ labels)
- Features
  - Content-based: lexical pattern, part-of-speech pattern
  - Network-based: user follow information, retweet information
- Algorithms
  - KL divergence retrieval model [Qazvinian et al. EMNLP2011]
  - Random walk model [Li et al. IJCAI2015]

3. Event Detection (SIGIR' 2012)

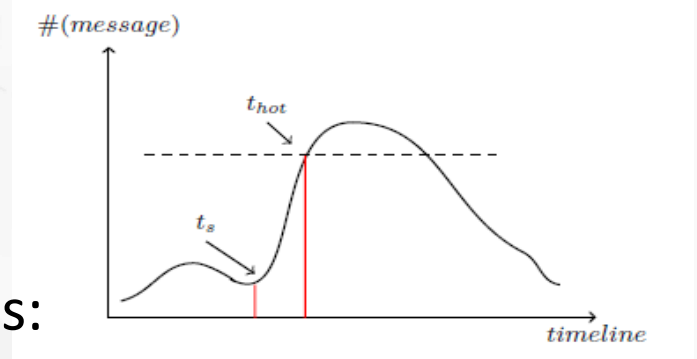
- Live Event Detection and Tracking
  - Extract named-entities in text messages, and visual concepts/ logos in associated images
  - Filter noise using a learned classifier
  - Cluster remaining (multimedia) message into events using an incremental clustering algorithm
  - Further divide the detected events into emerging and evolving events
  - Identify influence/active users and update evolving keywords
- Key research Issues:
  - How to do it Live??
  - How to extract relationships between events
  - and between events and users



4. Predictive Analytics -1

(SIGIR' 2012, 2013)

- Identify hot emerging events:
  - Aim: to detect such events before they become viral



- Key characteristics for hot emerging events:

✓ Has high rate of: User #, Message #, Re-tweet #, cumulated msg influence weight

✓ Has high level of overlap between: org key users and topic key users, org keywords and topic influence keywords

- 6 features for classifier with high accuracy:

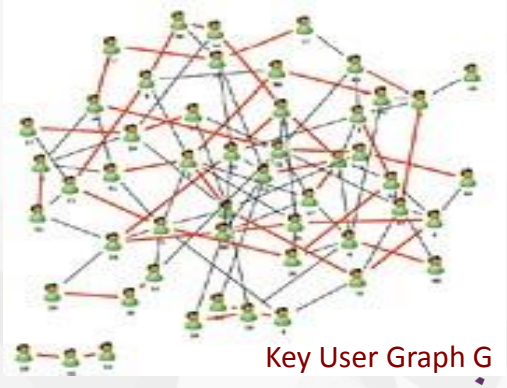
$$f_1 = \frac{|U^t|}{\sum_{x=0}^{t} \frac{1}{t-x+1}|U^x|} \qquad f_2 = \frac{|Tw^t|}{\sum_{x=0}^{t} \frac{1}{t-x+1}|Tw^x|} \qquad f_3 = \frac{|Rtw^t|}{\sum_{x=0}^{t} \frac{1}{t-x+1}|Rtw^x|}$$

$$f_4 = \frac{\#(ku_{tp} \cap ku)}{\#ku_{tp}} \qquad f_5 = \frac{\#(kw_{tp} \cap kw)}{\#kw_{tp}} \qquad f_6 = \frac{|A^t|}{\sum_{x=0}^{t} \frac{1}{t-x+1}|A^x|}$$

4. Predictive Analytics -2
(SIGIR' 2012, 2013)

- Framework for Viral Message and Event Detection:
  - Extract live social signals for early cues of trending events
  - Track multiple indicators: involvement of influential user, number and increase in msgs and re-tweets..
  - Identify also users likely to participate in diffusion process
  - Classifier for viral events/messages
  - Identify hot emerging events/ messages before they become viral

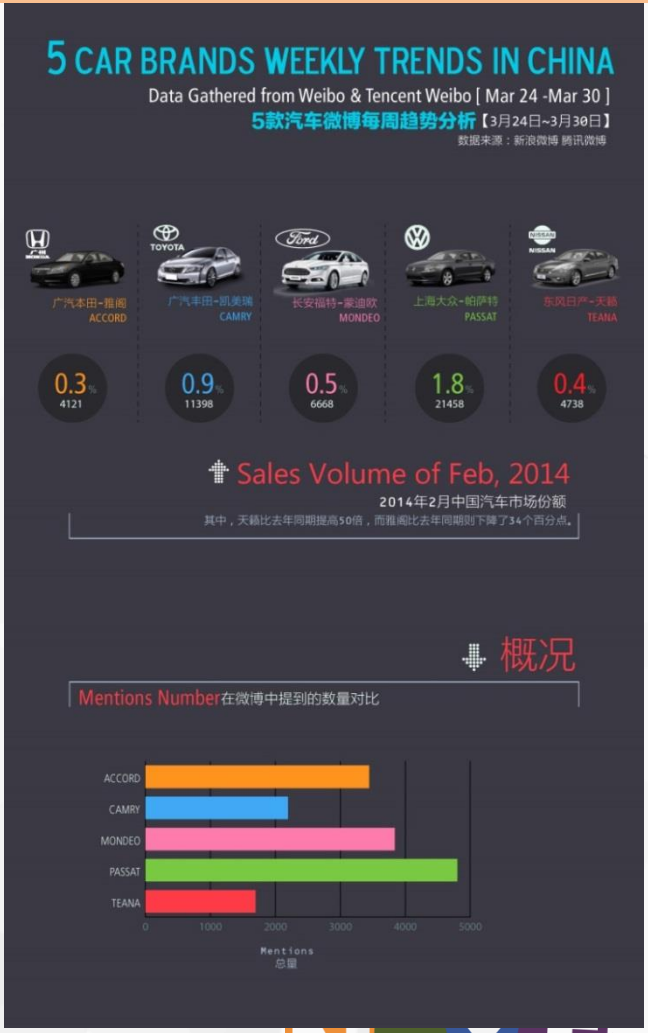
- Influential/ Active User Detection:
  - Characterization of influential and active users
  - Identification of mid-range influence users
  - Leading to prescriptive actions


Key User Graph G

5. Prescriptive Analytics

From Data to Review/ Report

- Converting Data into Actions:
  - Translating past, present, and future data into actionable items for better user engagements
- Reviews and Reports:
  - Smart linguistic program that aggregates social media sentiment and overall opinions about an entity
  - Generate comparative reports based on key indicators
  - Identify strengths and weaknesses for each entity
  - Towards prescribing (current and future) actions to improve image and user engagements



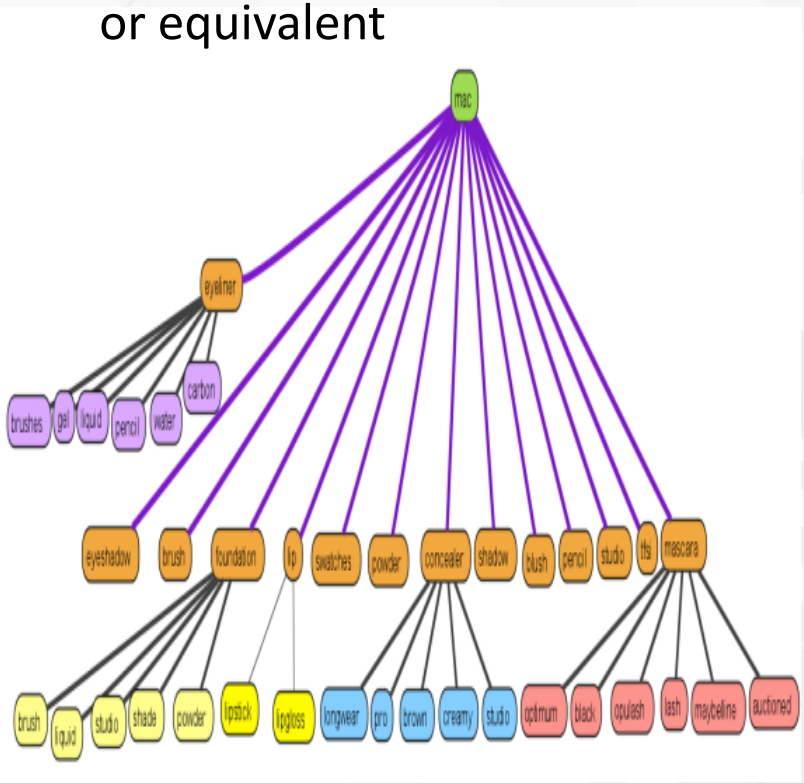
5. Prescriptive Analytics - 2

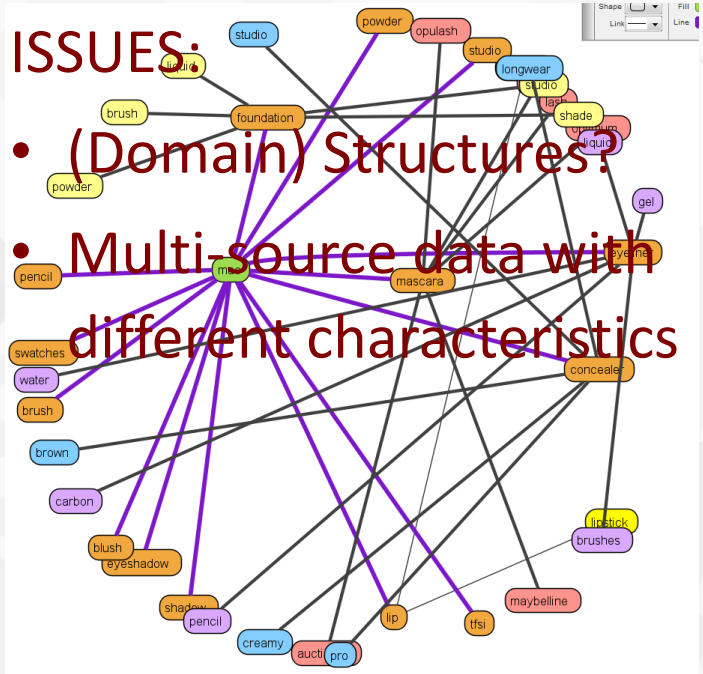
From Data to Review/ Report

- From Data -> Analysis -> Predictive -> Prescriptive Analytics
  - Towards analysis of “why” and “how”…
  - Work on characterization of problem, and solution to alleviate the problems

6. Information Organization:
from unstructured UGCs to Structured Knowledge

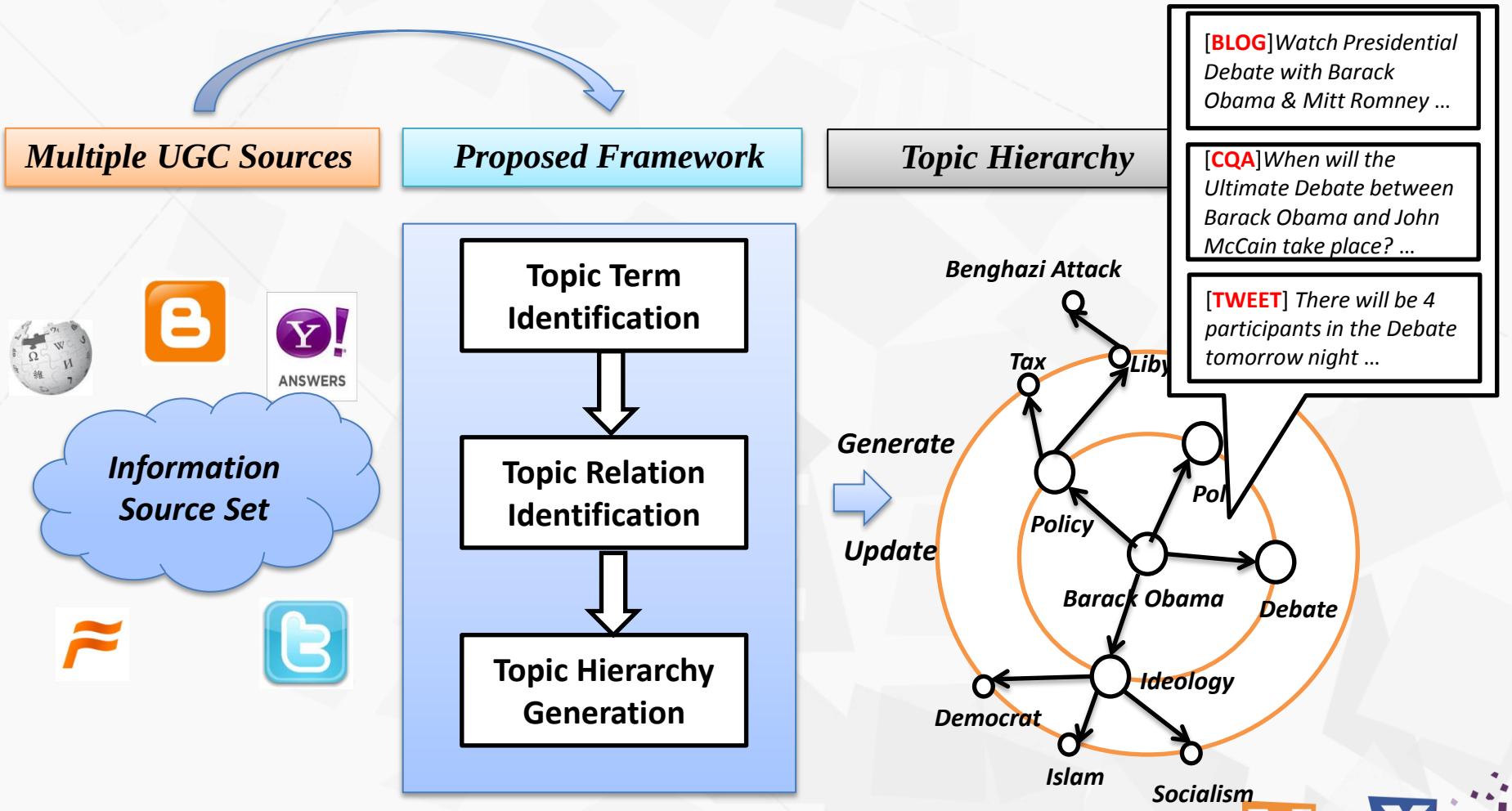
- Mining structure of data from multiple UGC sources
  - UGC sources includes: Wikipedia, Blog, cQA, Forum and Twitters, or equivalent





Information Organization

The Framework



- SIGIR 2013

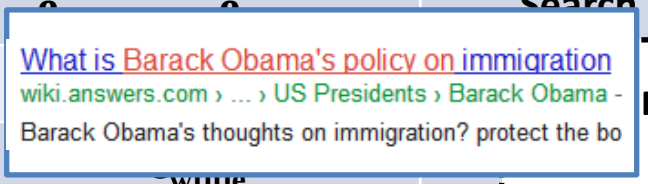
Key Research Focus:
Topic Relations Identification

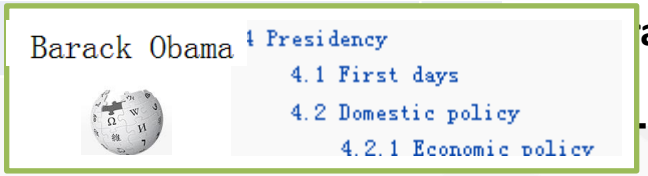
- Sub-topic relation: $R(t_a, t_b)$ indicates that t_b is a sub-topic of t_a
- We use multiple evidences to estimate probability $p\left(R(t_a, t_b)\right)$ of both directed and undirected relations

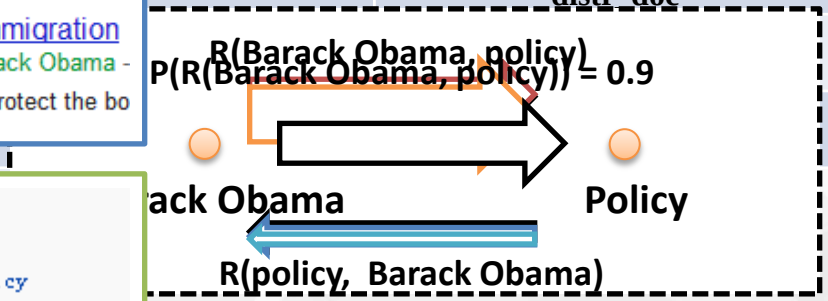
$$p(R(t_A,t_B)) = \left(\sum_{e_k \in E_{direct}} w_k \cdot e_k(t_A,t_B)\right) \cdot \prod_{e_s \in E_{undirect}} e_s(t_A,t_B)$$

directed-evidences	Source	undirected-e
	Search engine	e_{distr_doc}
		Information Source Set
		Wikipedia

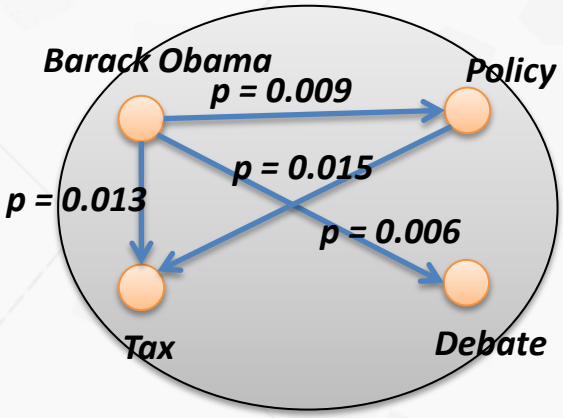
npmi(Barack Obama, policy) = 0.7

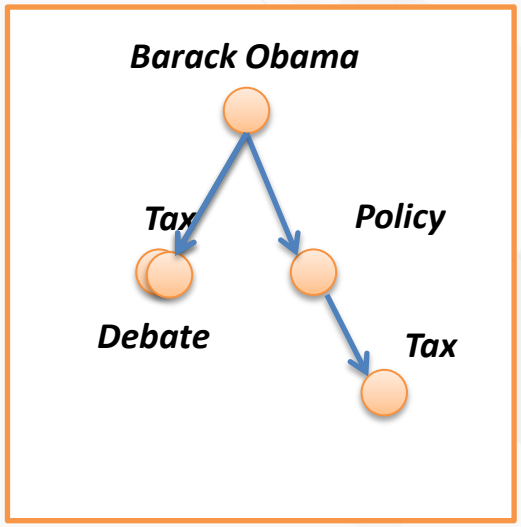






Information Organization:
Topic Hierarchical Generation





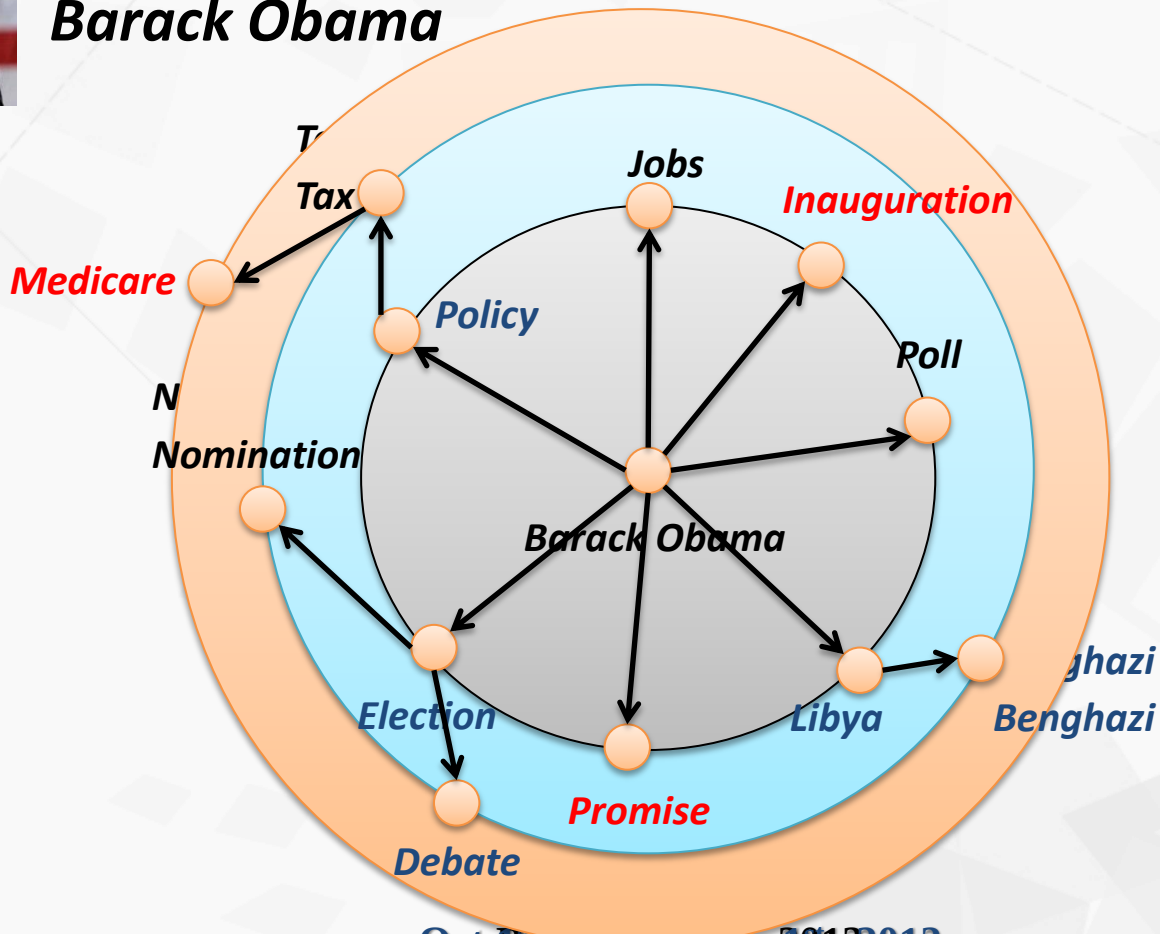
Resultant Hierarchy

Algorithm 1 Hierarchy Generation Algorithm

Input:
T: the candidate topic term set;
Output:
R_{ret}: the sub-topic relation set of the resultant hierarchy;
T_{ret}: the topic set of the resultant hierarchy;

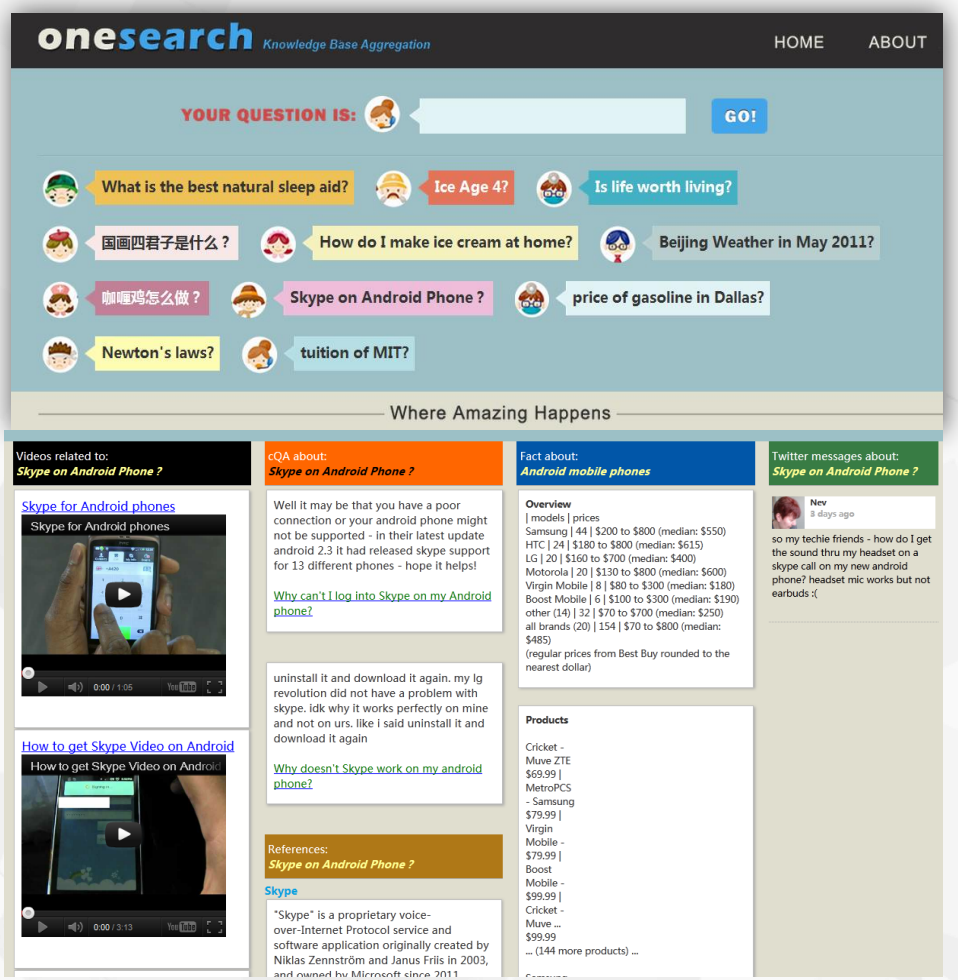
Initialize $T_0 = \{\mathcal{C}\}$, $R_0 = \phi$
for i = 1 TO ∞ **do**
  $t \leftarrow$ selectTermFrom$(T - T_{i-1})$
  if t is NIL **then**
    $R_{ret} \leftarrow R_{i-1}$
    $T_{ret} \leftarrow T_{i-1}$
    break
  end if
  $T_i \leftarrow t \cup T_{i-1}$
  $R_i \leftarrow R_{i-1}$
  for all t_k IN T_{i-1} **do**
    if edgeBetween(t_k, t) exists **then**
      $R_i \leftarrow R_i \cup$ edgeBetween(t_k, t)
    end if
  end for
  edgeWeighting(R_i)
  $R_i =$ hierarhcyPruning(R_i)
end for

Evolution of a Hot Topic
over time



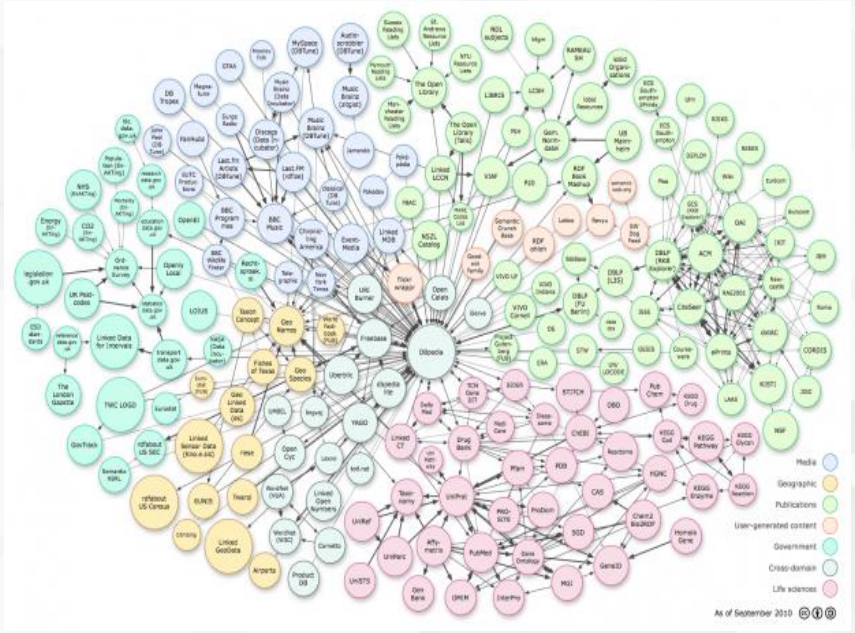
Application in Multimedia QA

- OneSearch: Multimedia/multilingual question-answering
  - Long term research: turning social QA into dynamic knowledge structures
  - Focus on returning timely multimedia answers

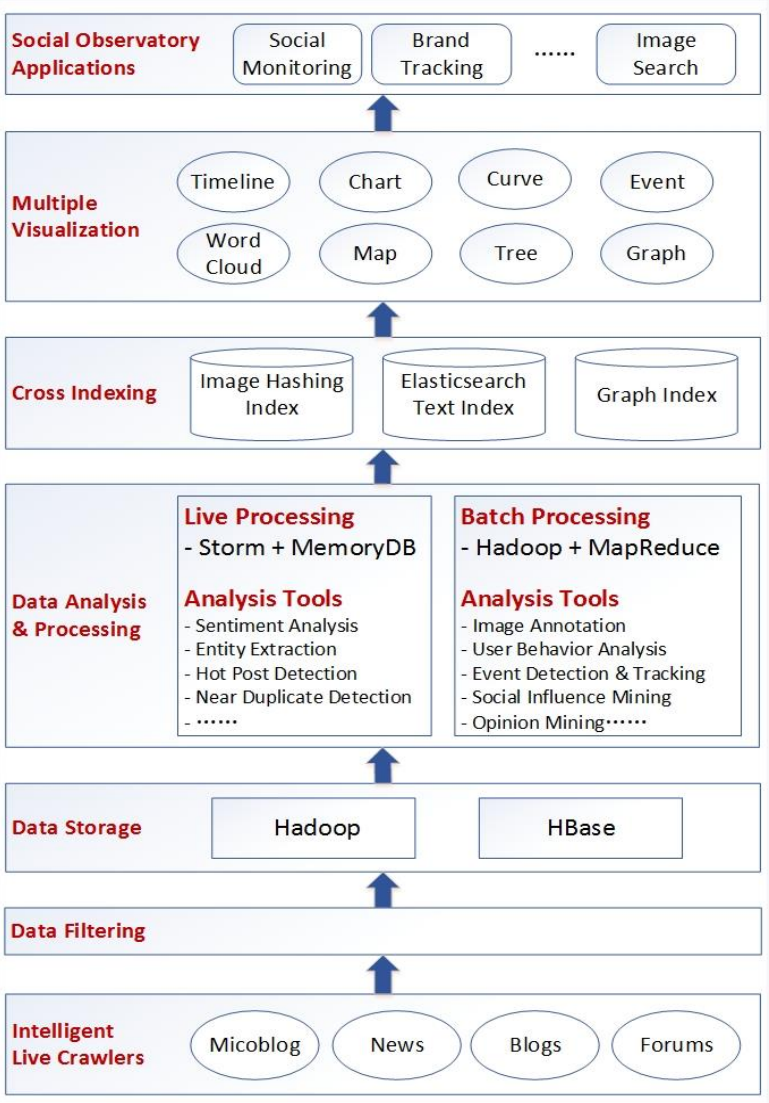


Overview of Information Organization
Unstructured to Structured Knowledge

- Towards knowledge structures/ graphs for different domains:
  - Information and relation extraction
  - Multi-source knowledge integration
  - Domain: Entertainment & Wellness (Critical Illnesses)



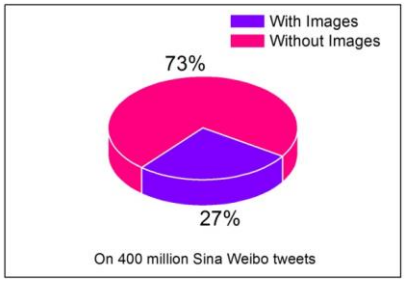
7. Scalable Infrastructures

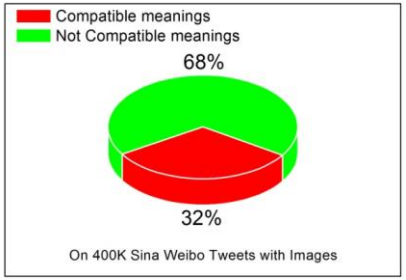


- **Back-End Tech Highlights**
  - Fully automated
  - Cloud-based
    - 80 VMs
  - Big data
    - >100 TB, 2.6 billion records
  - Live + Batch Processing
    - Storm + Hadoop + Hbase
  - Distributed
    - Crawlers, Database, File Storage, Processing
  - Cross indexing
    - Elasticsearch index for text
    - Own hash-based image index
    - Graph index

8. Large Scale MM Content Analysis (ICMR 2014, ACM MM 2014)

- Information is becoming increasingly multimedia
  - >30% of microblogs have images/videos
  - Over 40% of such microblogs do not have relevant text descriptions
  - Purely text based approach is inadequate

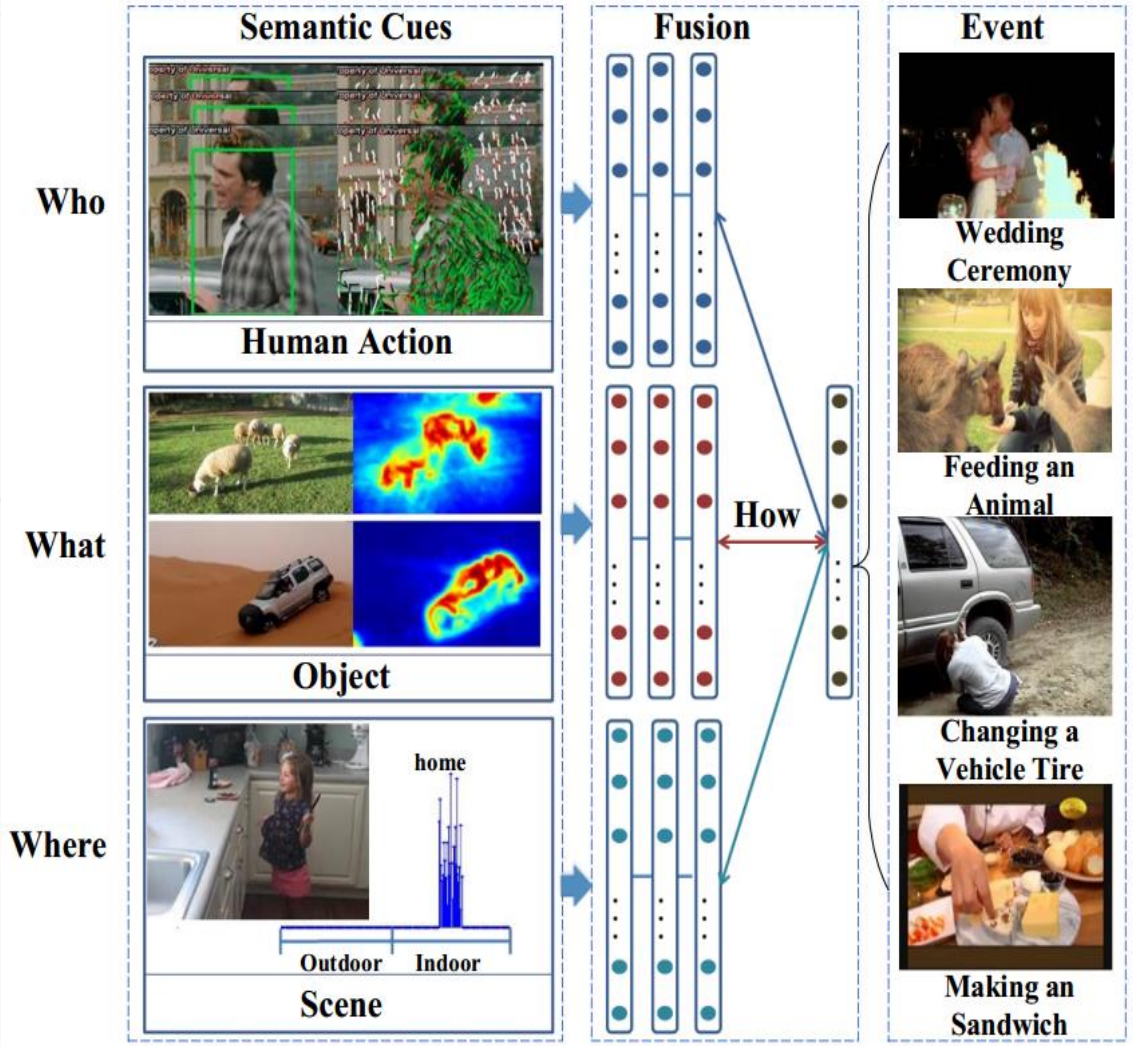




- Key problems: How to identify images relevant to entities/topics without relevant texts
  - ☺ Huge amount of social-tagged mm content available
  - ☹ Large semantic & intension gaps in mm contents
  - ☺ Classifiers to detection certain classes of visual concepts, logos are effective

Large-Scale Social Video Analysis

Deep Learning Architecture for Fusing Multiple Cues



Observation 1: ***Actions, objects and scenes compose the semantics of videos***

→ We devise a multi-channel deep NN.

Observation 2: ***Direct modeling of these semantics are not effective***

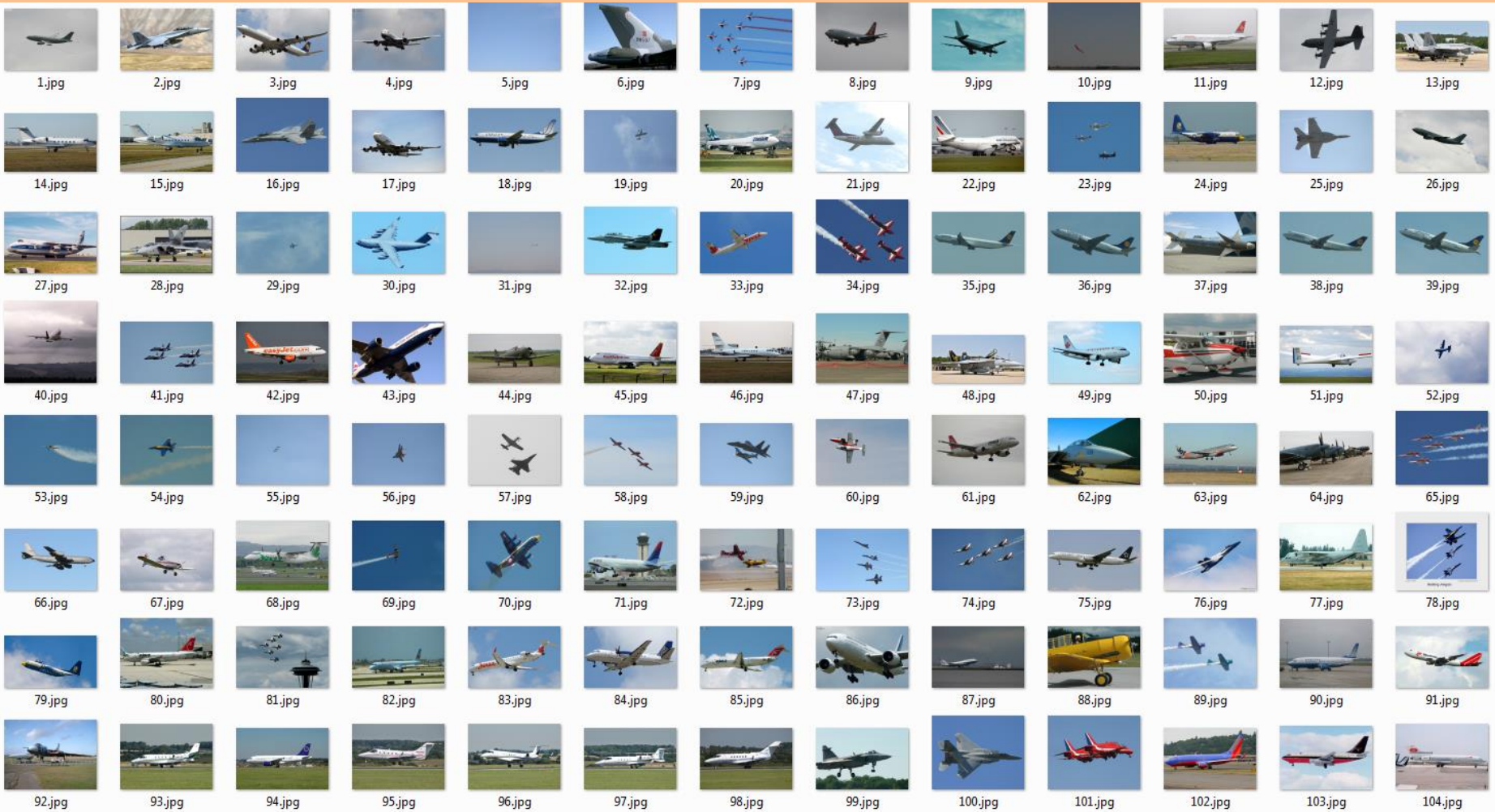
→ We extract more higher-level semantics that are more easier to be fused.

Large-Scale Social Video Analysis

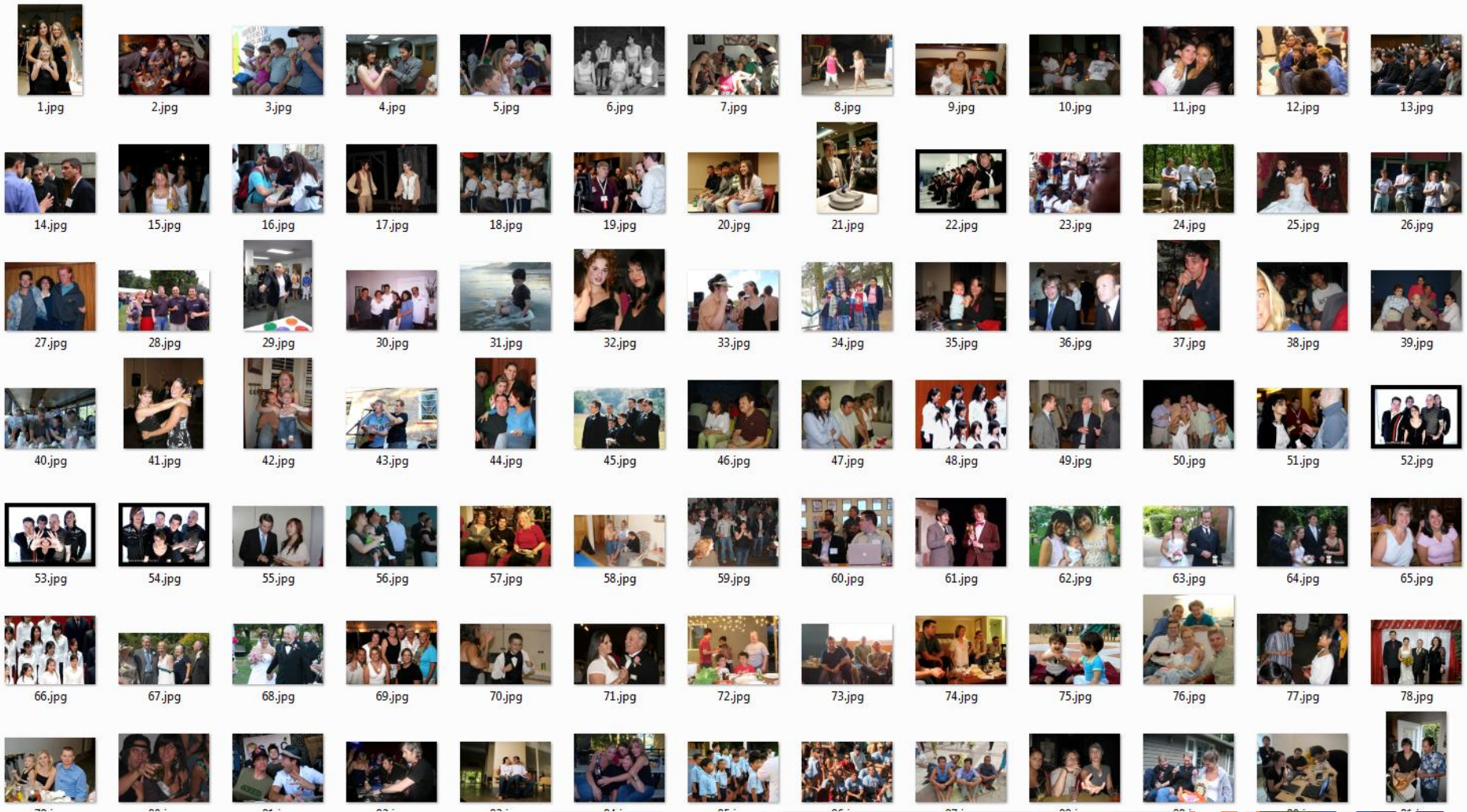
Deep Learning Architecture for Fusing Multiple Cues

- Experimenting with Multi-modal deep learning architecture
  - combining object, scene, motion and text features
- Automatically discover mid-level attributes
- Used in benchmark classification tasks
  - Out-performed existing state-of-the-arts features/systems on image annotation and video event detection
- Towards continuous learning system

Example of Attributes Discovered (Airplane)



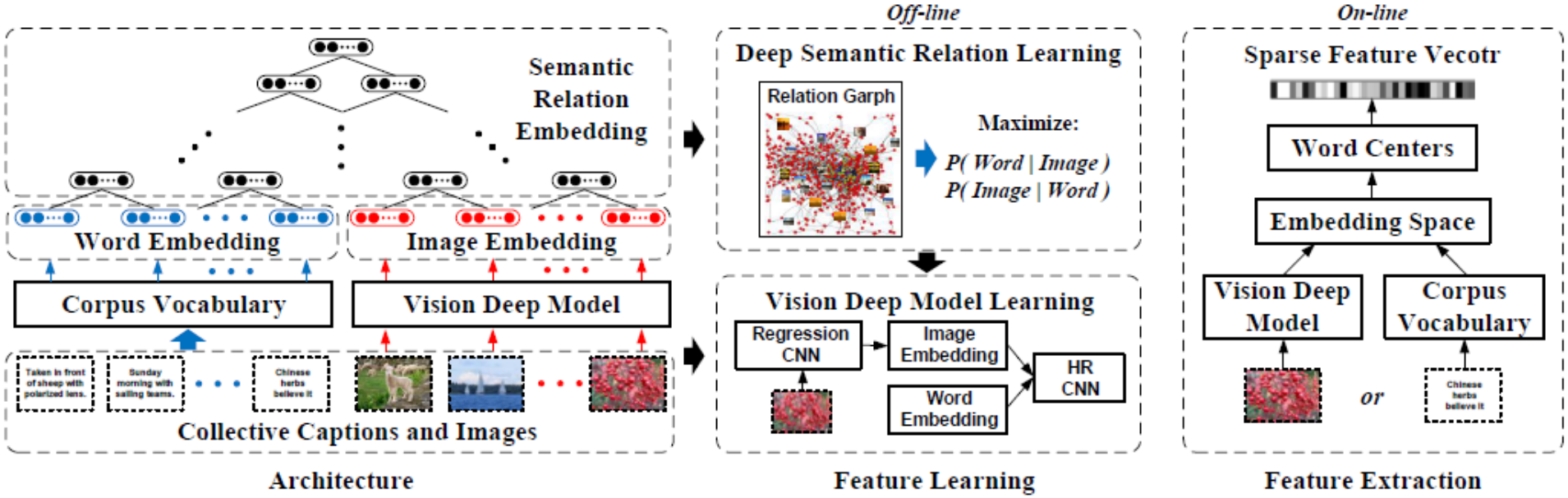
Example of Attributes Discovered (People)



Content Representation via Embedding

Learning Features from Images, Tags & Users

Deep Semantic Relation Embedding Architecture



mAP% of the proposed features (last two) as compared to state-of-the art on benchmarks

Dataset/Method	DeCAF	ICMAE	Ours-fc7	Ours-4kSc
NUSWIDE	38.6	32.7	41.3	**42.9**
CCV	62.5	58.4	67.2	**67.8**

Content Representation via Embedding

Examples: Image-Image Embedding

Content Representation via Embedding

Examples: Word-Image Embedding

Content Representation via Embedding

Examples: Image-Word Embedding

beach, girl, ocean, sand, water, playing, white, bikini, chair

soccer, field, game, players, zijn, groen, running, roelof, januari, ball

bike, girl, street, yellow, car, white, woman, boy, man, dog

wedding, wearing, dress, white, girl, red, black, pink, dressed, flower

Content Representation via Embedding

Examples: Word-Word Embedding

Dog	*foster, sleeping, service, cute, cat, puppy, stray, freshfields, spaniel, pug*
Girl	*dress, cute, hat, pink, boy, portrait, baby, shirt, birthday, blonde*
Morning + Drink	*drink, champagne, sky, juice, drinks, coffee, foggy, cold, soda, sun*
Evening + Drink	*evening, drink, wine, beer, bartender, alcohol, beverage, martini, liquor, bottle*

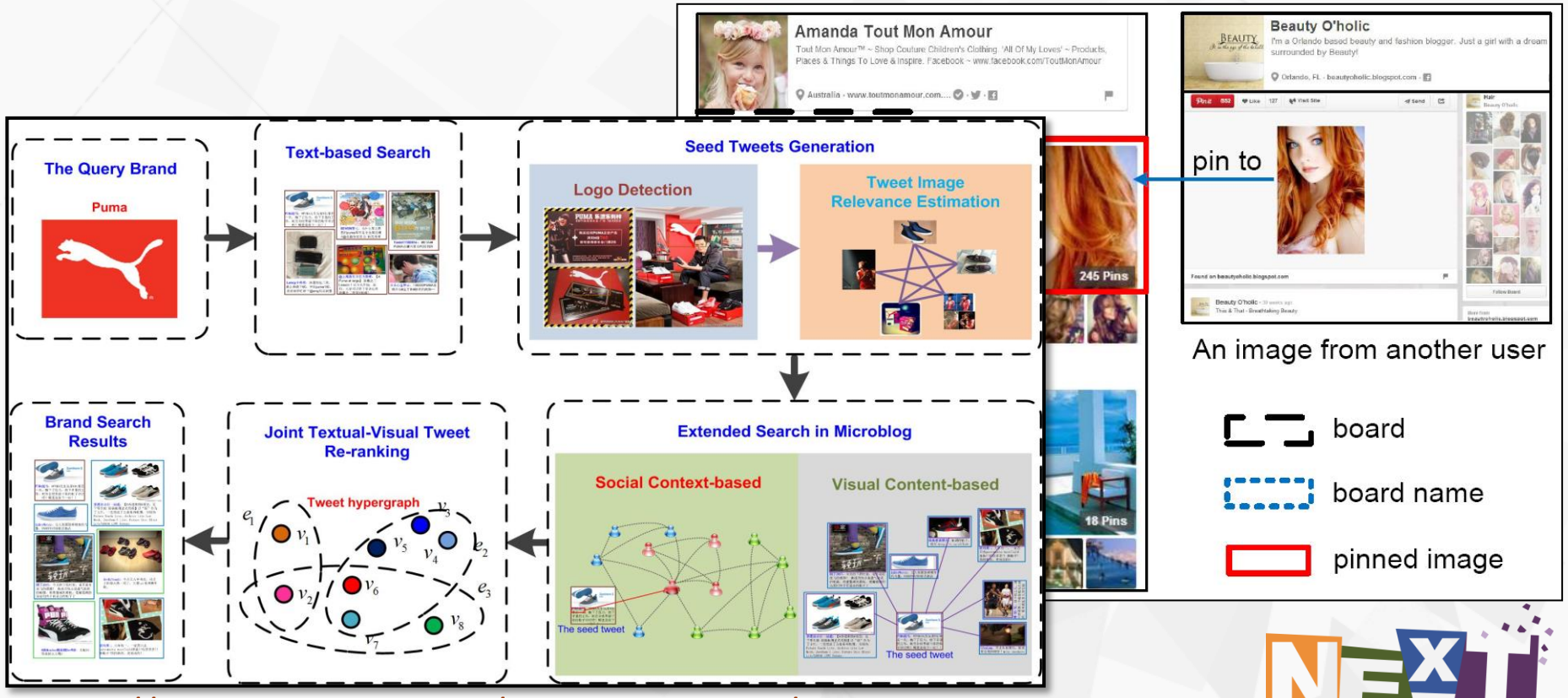
Table 1: Spearman correlation of 971 word-pair similarities computed by different methods and MEN human judgements. Our method learned from SBU is very close to Word2Vec learned from Google News.

Method	Word2Vec	S-CNN	Ours
MEN	**0.65**	0.36	0.64

Large-Scale Social Image Annotation

Mining Curated and Weibo Images

- Extend to social images from different types of social networks
  - Classifying curated images from Pinterests
  - Identify brand images from Weibo

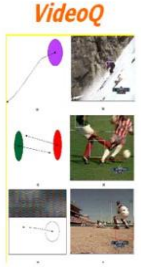


http://www.nextcenter.org/Brand-Social-Net/

Deployment of Visual Technologies -1

Variety of Visual Technologies

VideoGoogle



VisionGo

Mealsnap

Deployment of Visual Technologies -2

Some vertical domain visual search engines

Company	Appl^n	Technologies
Snapfashion	Fashion find Apps	• Similarity Search • App development
Superfish	Visual search API for image	• Similarity Search
Cortica	Contextual Ads Serving for image	• Matching-based recognition
Face++	Face, in finance & surveillance	• Similarity search • Object recognition
Dextro	Visual recognition for live video streams	• Visual analytics for images/videos
ViSenze	Visual Search & recognition API for image and video	• Similarity search • Domain ontology • Object recognition

ViSenze: A Startup Company under NExT

(www.visenze.com)

Size of Company:

- 30+ Employees
- 22 Technical (7 PhDs)
- Support over 30 million paid images
- Moving to USA and Europe

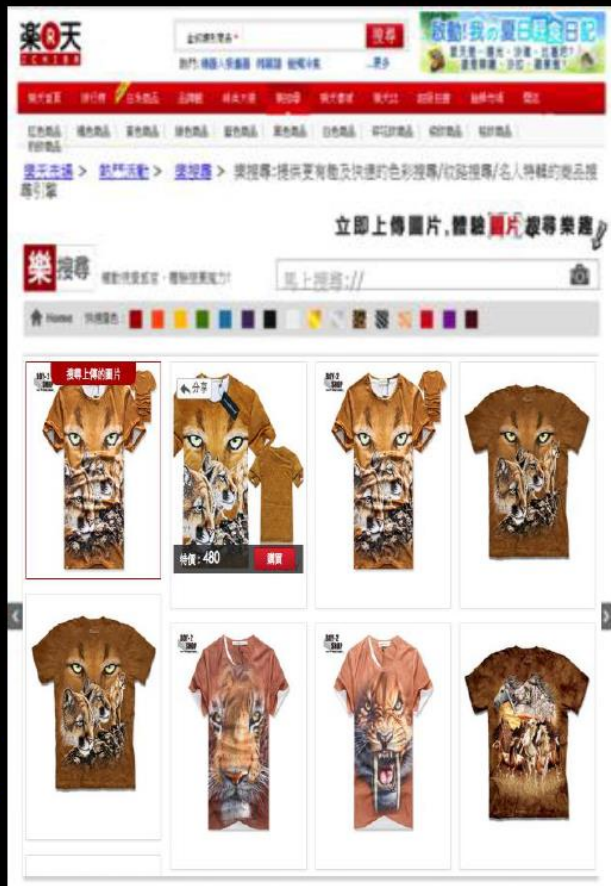
Key Customers:

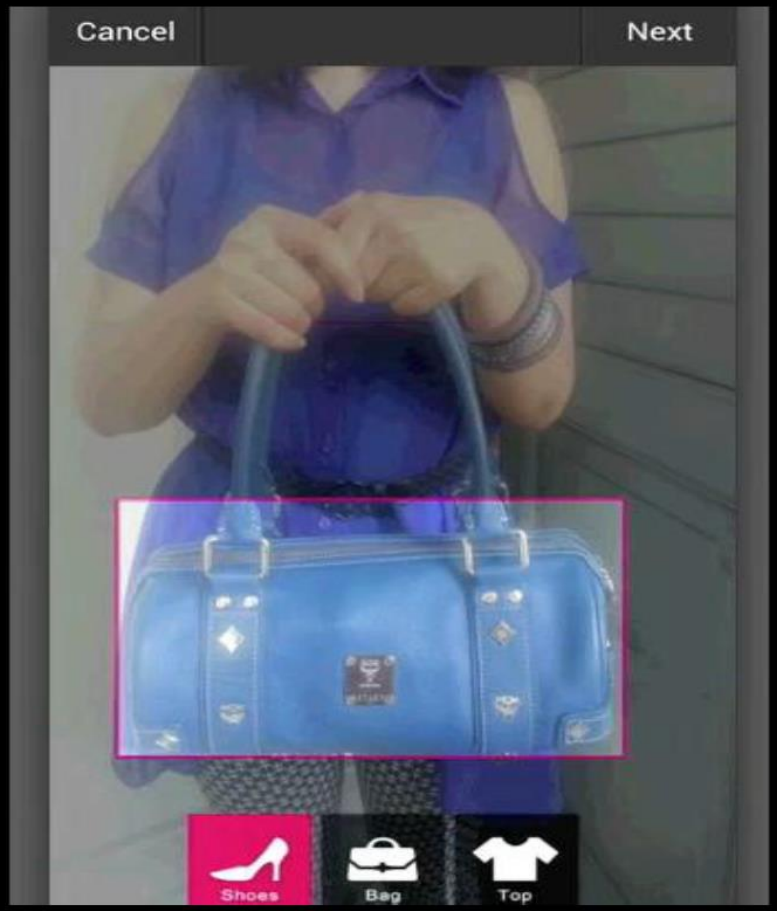
(over 50% of key sites in SE Asia and South Asia)

Flipkart, Rakuten, Zalora (Rocket), Lazala (Rocket), Reebonz, Patsnap

Download Flipkart App from Phone store

VISUAL POWERED E-COMMERCE SITES

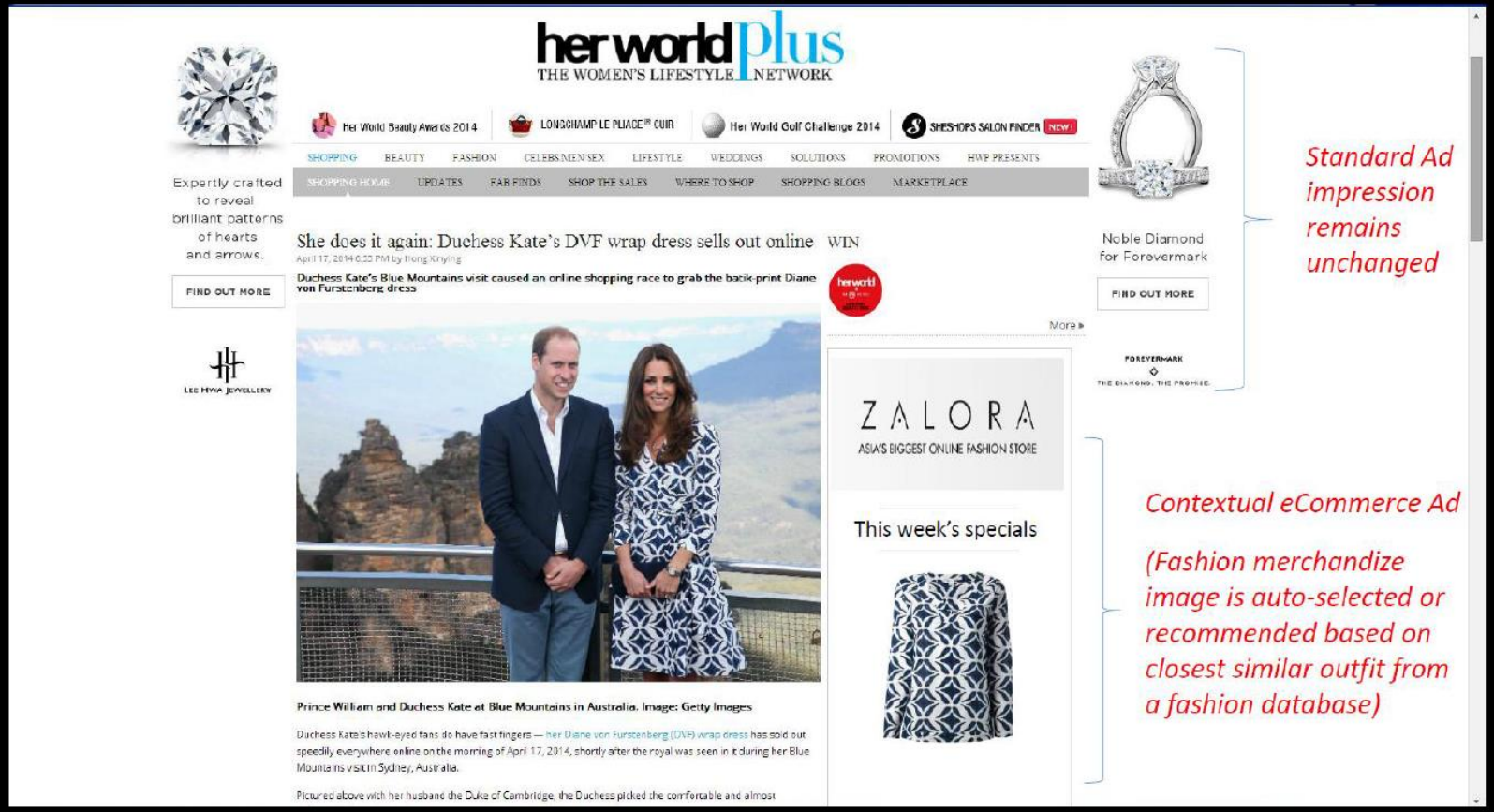




Snap Search Buy !

Shazam for Fashion Commerce !

Making Visual Content Targetable and Relevant to Advertisers



Contextual Advertising in Videos

Detect and recognize brands/logos in videos to associate with relevant Ads or Recommendations

In-Video Recognition

Use cases for Visual Search & Recognition

- Offline to Online
- Online to Offline
- Large-scale database retrieval
- Visual recommendation
- Supports multi-categories:
  - Fashion & related
  - Packaging
  - Home décor, furniture
  - Diagrams/drawings
  - And MANY more …….

Deployment of Visual Technologies -3

- Recent products are targeting at mostly 5 vertical domains:
  - Fashion
  - Home furnishing
  - Products
  - Surveillance
  - Finance

- Highlights of most current deployments:
  - Most are based on image matching & annotation technologies
  - Visual recognition has just started
  - Some impressive recognition systems are crowd-sourced-based
  - Hardly anyone offer real-time in-video products

OUTLINE

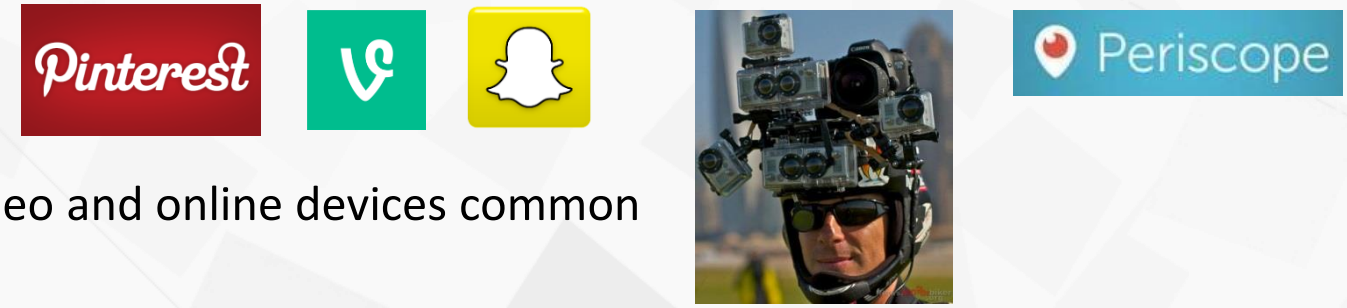
- The Live Social Observatory
- The Technologies behind the System
- Next Generation Systems and Challenges
- Other Related Research
- Summary

A Recap: Key Characteristics
of Current Generation of Social Media Systems

- Live, Big-Data and 3M (Multi-source, Multimedia & Multilingual)
- Support users' need for communication, sharing and interaction
- Support co-viewing and co-creation of contents (by users and systems)
- Develop social analytics aim to understand contents, events and users - targeting at recommendation

What Has Changed in Last 5 Years?

1. Image/Video handling
   - Top 3 recent social media platforms are all image/video centric



   - Live video and online devices common place

Visual-based Platform: SnapChat

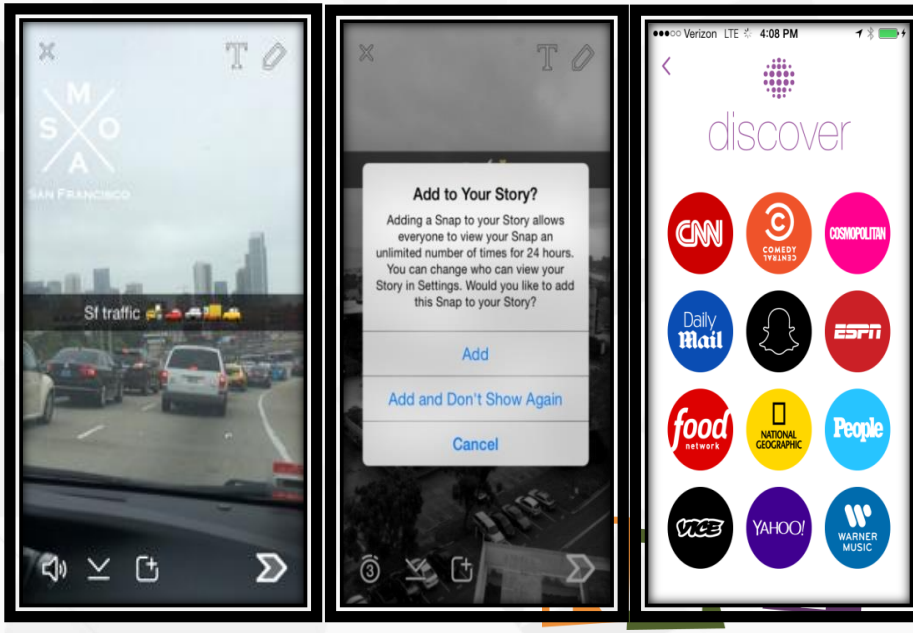
- Snapchat: founded in 2011; Sharing 10-sec video moments that erases itself after certain period;
  - 70% of users <24 yrs old (~100M active users);
  - user base will be bigger than Twitter soon

- Temporal Real Time Messages: people decide how long others view their photos/videos
- Share live stories with friends
- "Discover" news from top sources, eg CNN.
- See also Vine: 6-sec loop-videos

Share live videos

Creating a story for friends

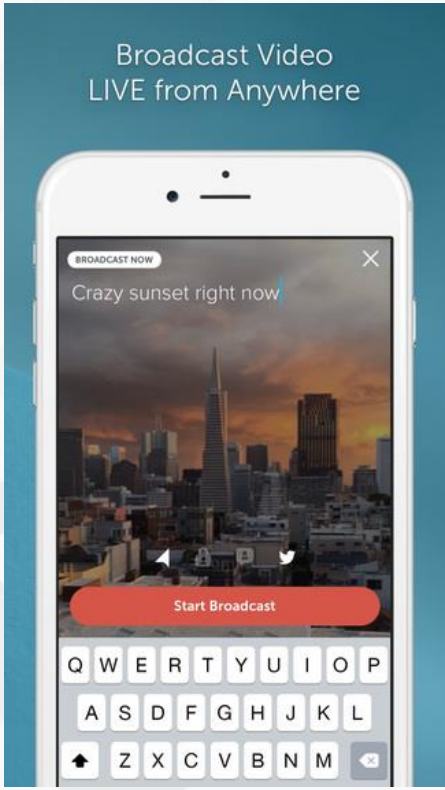
Consume content by brands -- news



Periscope

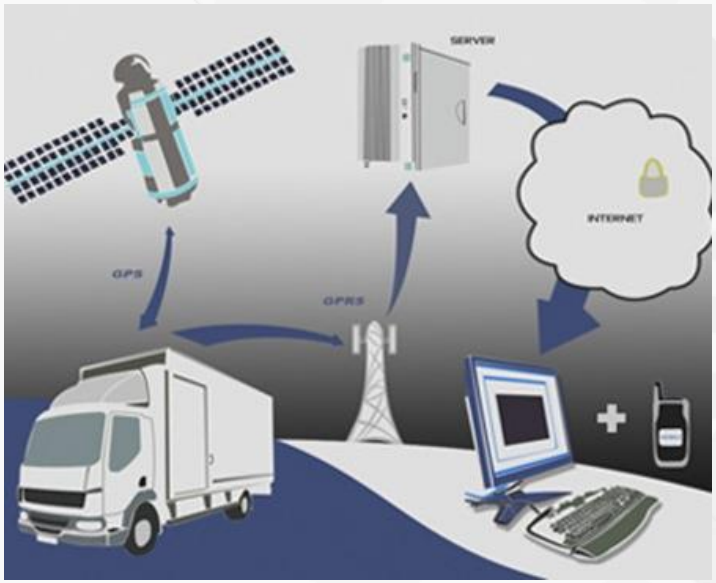
- Similar to SnapChat in supporting live broadcast from mobile phones
- Users can broadcast and watch any other live broadcasts

- Why it is hot??
  - Support mobile-based broadcast
  - Track events live
  - Interactive
  - Integrate with Twitter Social Network



What Has Changed in Last 5 Years?

2. Live location analytics
   - Sensor devices are everywhere – capable of multi-form of sensing
   - Multi-source information: location traces, POI and audio
   - Towards better location estimation and mobility analysis



What Has Changed in Last 5 Years?

3. Live aspect is central now
   - With Live, comes the ability for continuous sharing and feedback..
   - Users want instant feedback – from friends and systems

- Live videos are commonplace
  - Shoot, share, and feedback, all live
  - Many examples: home automation, SnapChat, Periscope
  - Problems in understanding and reconstructing the interesting places visited and/or events happened

What Has Changed in Last 5 Years?

4. Quality and structure of data
   - Deteriorating quality of data, with about 70% of UGC's belongs to noise/ spam/ rumors category
   - Representativeness of data is important, especially for large dynamic topics
   - Key approach to making data usable is to structure them – at both data and knowledge level

What Has Changed in Last 5 Years?

5. Co-Creation & Co-Invention
   - With live instant feedback, comes the possibility to co-create and co-invent
   - Not just the contents, but systems and design

Co-Inventing the Future -1

- GoPro points towards future of Consumer Electronic Products
  - Products online from Day 1
  - Create social communities
  - Feedback to product design
  - Bootstrap new products and applications
- Another example of intelligent appliances: Home Air Filters
  - Filters online from Day 1
  - Understand users' living habits
  - Auto order when filters break down
  - Extendable to any home appliances
  - Foundation for intelligent home

Co-Inventing the Future -2

- Product Design:
  - Building a dedicated user communities
  - Leverage opinion leaders in communities to feedback on deign of systems/products (hardware and software)
  - Offer new design through communities to influence more users
- Government policies
- Early Stage Examples

What Has Changed in Last 5 Years?

6. Online & Offline Integration
   - Online data represents certain segment of population
   - Offline data often represent the majority, if available
     (However, other than ecommerce, offline data is often limited)
   - Integration of both provides full picture of the world

- Examples:
  - E-commerce social media buzz; need offline data on actual sales and online sites that cater to buyers
  - Election of Singapore – social media and actual voting pattern is totally different

Big Data Revisited

- Big Data Analytics
  - Data online and Live
  - Gather and Filter data (not Store, Sample and Search)
  - Feedback (close-loop) and React
- Big Data is not about big, but about online and live
  - Through user and systems feedback, to realize better user experience
- Towards C2B
  - Alternatives to B2B and B2C, where business takes central roles
  - Customers now take central roles in influencing business

Key Technologies

- Predictive Analytics
  - Predict users preferences and dislikes
- Prescriptive Analytics
  - Analyze user communities and identify local communities
  - Analyze user feedbacks, identify needs/ concerns
  - Formulate actions

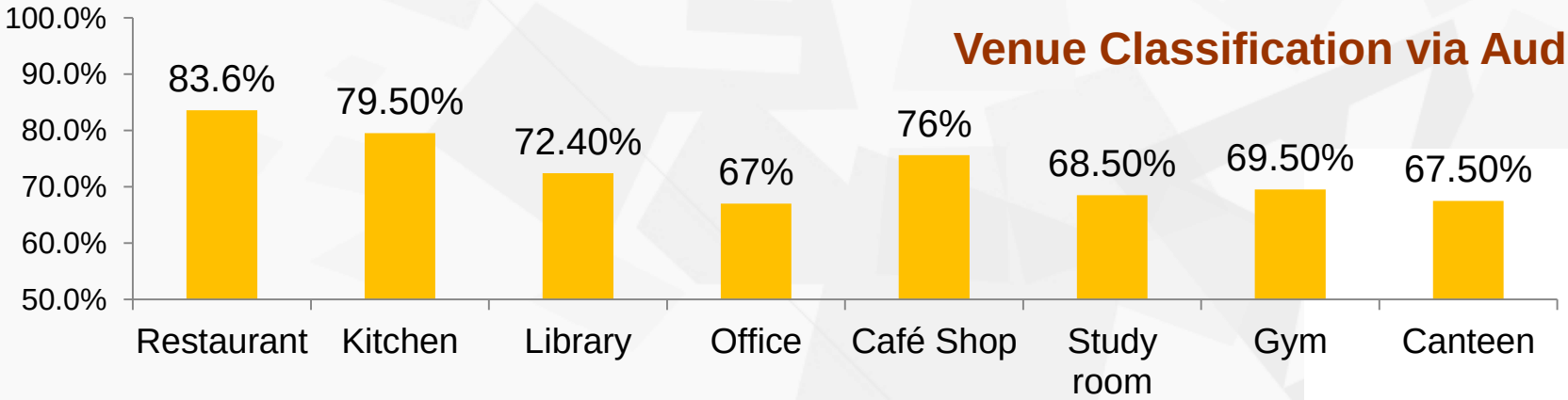
OUTLINE

- The Live Social Observatory
- The Technologies behind the System
- Next Generation Systems and Challenges
- Other Related Research
- Summary

Multimodal Location Estimation

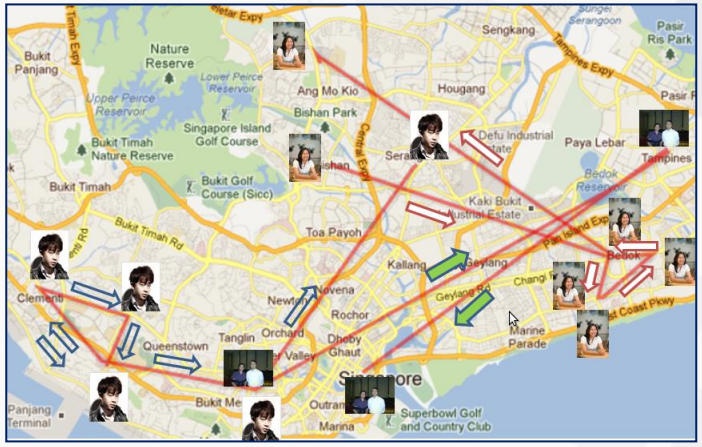
Combining Text, Visual, Location & Audio Analysis

- **Given a series geo-tagged tweets, how to map them to actual POI (Point of Interests)**
  - Each geo-tagged locations may be mapped to 100s of venues (as defined in, say, 4Square)
  - Sampled audio signals to estimate venue categories (for elder-care applications)
  - Results using Tri-modal Deep learning architecture (Timber, Rhythm, Pitch) > 67%
  - Classifier to estimate exact venue

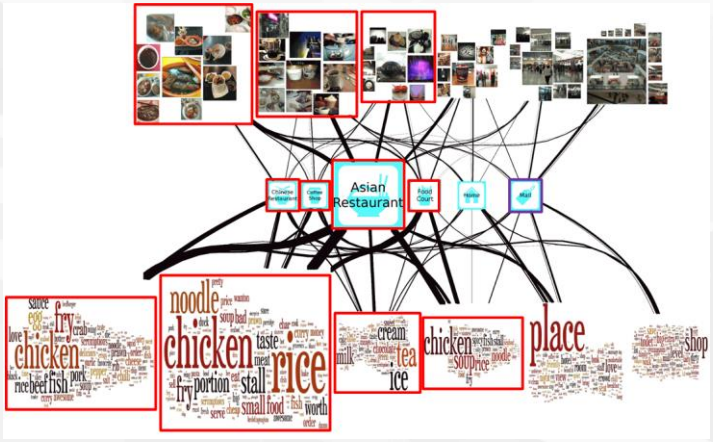


Location Analytics

- Mine relations between check-in venues
- Identify landmarks of local venues
- Identify popular trials (of individuals and their friends)
- Analyze user demographic and interests communities
- Generate multi-faceted user relation graph



Travel paths and flows of users



Interest community of food lovers in Singapore

Overview of Computational Wellness

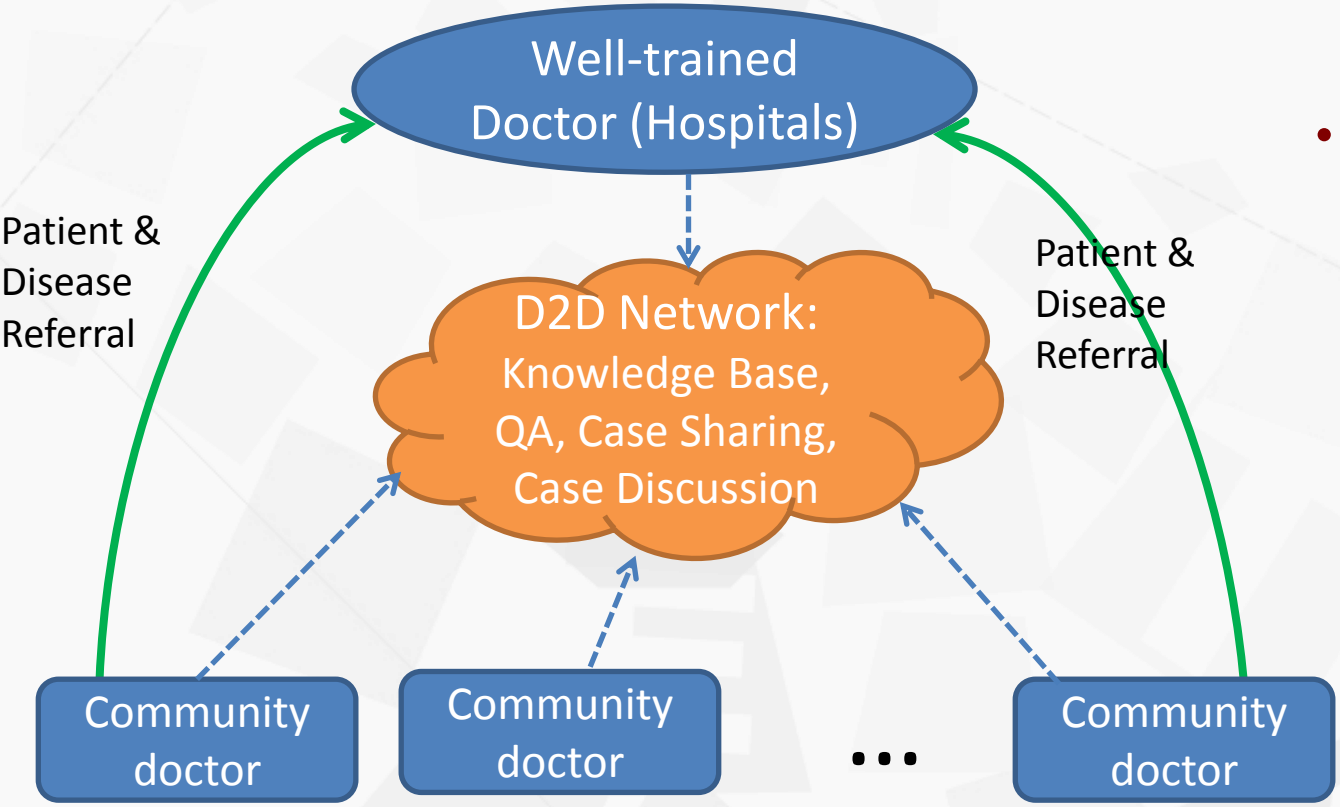
HealthSenze gathers & organizes big-data from multiple sources

- Personalized wellness with big data
- Focus on 6 Critical illnesses
- Predictive and prescriptive wellness

NEXT

Towards D2D (Doctor-to-Doctor) Network

A Form of Social Life Network



- With patients and disease referrals, D2D network also serves as national monitor for diseases and epidemic

- Able to reach hundreds of millions of patients at various Locations

Wellness Advisory System for End Users

HealthSenze

KB Sources

Crowd-Source

Forum/Blogs

Social Media Sources

Environment Source

Advisory System: based on 6 critical illnesses

End Users

- For end users
- Offer advises on appropriate activities and diets based on current conditions

OUTLINE

- The Live Social Observatory
- The Technologies behind the System
- Next Generation Systems and Challenges
- Other Related Research
- Summary

Summary

- We have built a live social observatory system:
  - Work on large-scale real-world problems of impact
  - Analyze large scale live UGCs and SDCs, and possibly MGCs
  - Monitor events and happenings in city to help users lead better life
- Research under NExT Research Center:
  - Over 40 researchers at all time
  - 2 Startup companies
- ViSenze Pte Ltd
  - 30+ staff
  - Mobile visual search
- 6Estates Pte Ltd
  - 10+ staff
  - From unstructured data to actionable knowledge

ViSENZE

6ESTATES

NExT

Summary

- Many Challenges:
  - Deeper analysis at entity and relation levels
  - Multimodal analysis at various levels
  - Real-time analysis and inference
  - Prescriptive analytics
  - Links to social understanding of events
  - Privacy and Transparency
  - ….
- Towards Innovative Applications
  - Social wellness
  - Smart city, E-government
  - Education..

THANKS

Visit our Web Observatory:
http:///WWW.NEXTCENTER.ORG/

We have openings:

- Interns, Research Staff, and Visiting Research Staff
- For NExT Research Center and Companies