Multimedia Retrieval:
Are We Doing it Right?

CHUA Tat-Seng

National University of Singapore

OUTLINE

- Retrospective on Image Search Since 2010
- Why Not Much Research in Video
- Some Current MM Research Efforts
- Summary

Disclaimer: The coverage of this talk is very broad, and I have time to cover only topics and works close to me, and do not have time to highlight interesting related works from you.

Two Basic Questions?

What is Multimedia Retrieval?

Why from 2010?

Invited Talk at CIVR 2010

Outline of Talk

- Information Rich World
- Contributions of MM Research
- Achievements of MM Research
- Bridge the Semantic Gaps
- Summary

Contributions of MM Research?

- Has years of MM /vision research contributed towards success of these ventures??
- One (or majority) point of views:
  - Very little
  - Most big ideas are simple !!
  - We are benefiting form their success, rather than helping them

    We do research on their data (Flickr, YouTube , Twitter etc)..
- Reasons:
  1) They are not technology oriented – put up brave face
  2) Other CS research has not contributed much either – lay blames on all
  3) They are consumer oriented where accuracy is not important, and context info is more than sufficient – is only context sufficient?
  4) No real data for large-scale academic research– only toy problems

Contributions of MM Research

- Context-only approach has its limitation
  - As such companies evolved, media content analysis becomes more important
- Not all glooms and dooms, some success stories
  - Duplicates removal in YouTube
  - Large-scale visual matching in Bing/Google Image Search
  - Landmark matching
  - Snap-Tell applications.
  - Others..

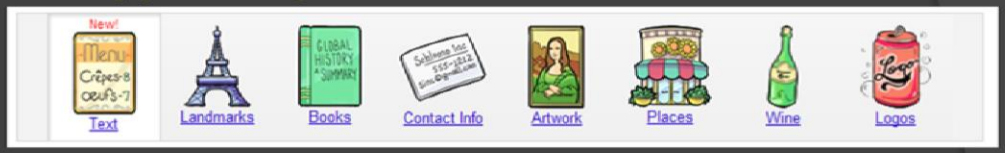
Introduced Several Systems

kooaba
point - snap - find

IQ Engines

Google Goggles
Use pictures to search the web.

- List of objects searchable via Goggle
  - Snap pictures and get answers



- More details of this Landmarks

- Give me the English translation of this menu?

- Key technologies: OCR, logo detection, image matching…

LMSearch

LMSearch

Technologies Behind These Early Commercial Systems?

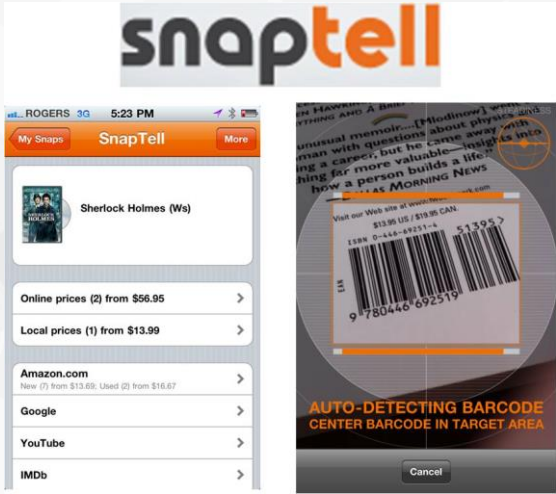
- Robust image matching technologies
- Hashing for large scale image indexing
- Concept Annotation for fixed number of common categories

What happen to these early systems??

SnapTell

- Founded in 2006 in Silicon Valley;
  obtained $4 million USD fund in Round 1
- It supports “snap and find” of covers of CD, DVD, book, or video game;
  returns product with ratings and pricing info from Google, Amazon, eBay and more.

- Acquired by Amazon on June 16, 2009
  To add technology to its mobile offering



Plink

- Founded in April 2009 @ London with $100K seed
- It stood out for its ability to quickly identify art with the snap of a photo
- Acquired by Google in April 2010

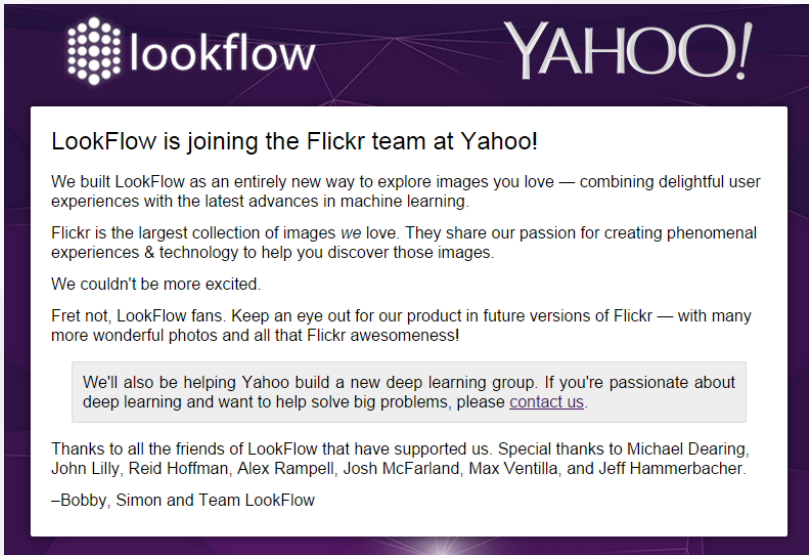
IQ Engine

- Founded in June 2008 @ Berkeley, CA
  With $3.8 million in Round 1
- If offers an image recognition platform called Glow that recognizes scenes, objects, landmarks, text and people in photos.
  Used to automatically tags and organizes users' photos.
- Acquired by Yahoo! in Aug 2013;
  Used to improve Flickr photo organization and search

LookFlow

- Founded in 2009 in Silicon Valley
- Build new ways for people to find, explore, collect, and share content they're interested in;
  using "an online search and discovery" technology

- Acquired by Yahoo! on Aug 2013;
  To leverage its ML algorithms for Flickr



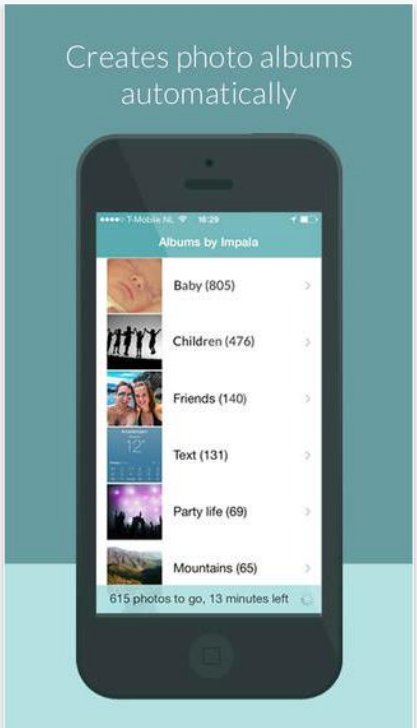
Kooaba

- Founded in Nov 2006 from ETH, Zurich;
  $2.9 million in 1 round;
- Image matching: mobile app that connects static printed newspaper format to dynamic online social sharing utilities
- Recently developed cloud-based image recognition solutions that integrate state-of-the-art visual recognition in applications
- Acquired by Qualcomm in Jan 2014;
  used to boost its Vuforia augmented reality platform

EuVision Technologies

- Founded in U of Amsterdam in 2010
- Technology: Image annotation
  Launched mobile app named Impala that helps user organize their photos automatically into 20+ categories.

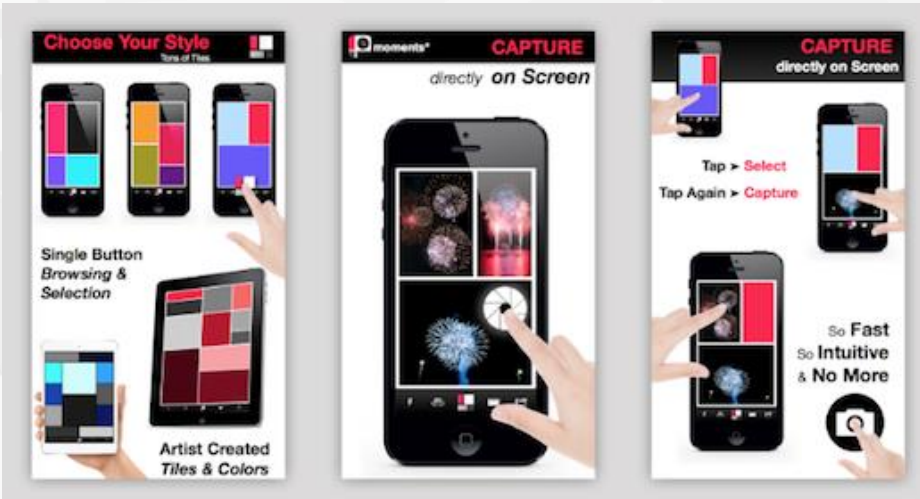
- Acquired by Qualcomm in Sept 2014
  Doing innovative research to integrate image/video technologies into mobile space



Madbits

- Founded in New York
- It deploys deep learning technology to assign relevant information to raw images.
  to automatically organize large databases of images.

- Acquired by Twitter in July 2014.
  To help bolster its image search feature and catalogue the firehose of photos uploaded to the network



VisualGraph

- Visual analysis techniques to connect images into a visual graph – towards visual Web
- It deploys state-of-the-art machine vision tools, such as object recognition (e.g. shoes, faces), with large-scale distributed search and machine learning infrastructures

- Acquired by Pinterest in Jan 2014.
  Recently announced products to offer recommendations to Pinterest users in fashion domain

Cloud-based, Big-data
Image Recognition / Visual Search platform

Examples of commercial visual-search system
[Google Goggles, Search-by-image] [Amazon A9]

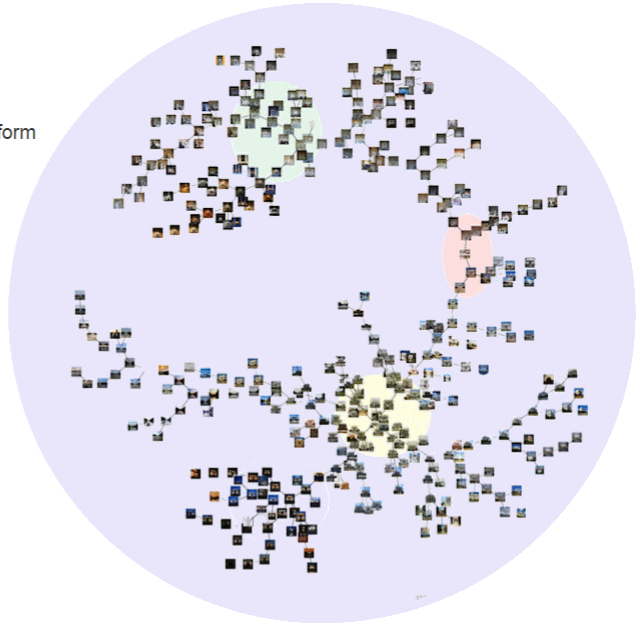
Our previous works on image graphs
[VisualRank For Product Images] [Google Image Swirl]

Technology
[Object Recognition]
- Face, People Detection (accuracy +98% of face.com)
- Cars, fashion objects, body, textured objects,

[Large-Scale Visual Search System]
- Product Recommendation
  (Distributed, in-memory, visual-search system)

[Image Annotation]
- Deep learning, coming soon.

Team / Mission
Ex-Google machine vision scientist and engineers.
Connecting world's visual inspirations.



Deployment of Visual Technologies -1

Variety of Visual Technologies: Lots of Opportunities

informedia
digital video understanding

Autonomy
VIRAGE

MediaMill
Semantic Video Search Engine

Sixth Sense

QBIC™

VideoGoogle

CuZero

bing

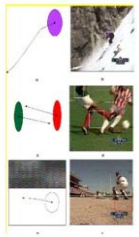
VisualSEEk

IBM
Multimedia Analysis and Retrieval System

snaptell

TALKMINER

VideoQ



blinkx

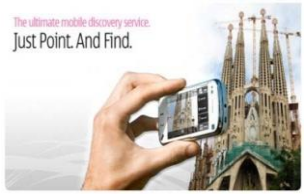
2001

Google goggles
labs

Google™
Video BETA

IQ Engines

VisionGo



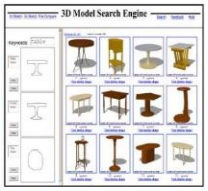
TinEye

kooaba
IMAGE RECOGNITION

SINGING FISH
the audio/video search engine

like.com
visual search

leafsnap



VIDEOSURF

ALIPR™
Automatic Photo Tagging and Visual Image Search

Stanford Mobile Visual Search

Mealsnap

NEXT

Deployment of Visual Technologies -2

In General Commercial Visual Search Engines

SERVICE	TECHNOLOGY	TARGET USERS	BUSSINESS MODEL
Kooaba	Image recognition	Entertainment	SaaS-based
IQ Engines (oMoby)	Image recognition Crowdsourcing	Shopping	SaaS-based
Mobile Acuity	Image recognition	Consumers, Marketers	Contracting with marketers
LinkMe Mobile	Image and audio recognition	Consumers, Marketers	Contracting with marketers
Snaptell	Image recognition	Consumers, Marketers	Contracting with marketers
Point&Find (NOKIA)	Image recognition	Mobile users, Marketers	Engage more users in NOKIA experience
Gazopa	Image recognition	Shopping, Mobile users	Improving on-line shopping experience
Google Goggles	Image recognition	Mobile users	Advertising-based
Clic2c	Watermarking	Entertainment	Contracting with marketers
WeKnowIt IMG REC	Image recognition	Tourism	Touristic promotion actions
Wizup	Image and audio recognition	Consumers, Marketers	Contracting with marketers
TinEye Mobile (Snooth)	Image recognition	Wine industry	Advertising-based

Deployment of Visual Technologies -3

- Highlights of most current deployments (mostly in B2B settings):
  - Most products are based on image search, matching & annotation technologies
  - Some of the impressive systems are actually crowd-sourced-based
  - Visual recognition has just started
  - Hardly anyone with real-time in-video products

- Recent products are based on vertical domains, in four popular vertical domains:
  - Fashion
  - Home furnishing
  - Products
  - Surveillance

Deployment of Visual Technologies -4

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
Facec++	Face, then to 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

Highlights of ViSenze

ViSENZE

Simplifying the Visual Web

AWARDS

SiTF AWARDS 2013
Awarding the finest in ICT innovations

EMERGING TECHNOLOGY

RED HERRING ASIA 100 WINNER

2014

RED HERRING GLOBAL 100 WINNER

2014

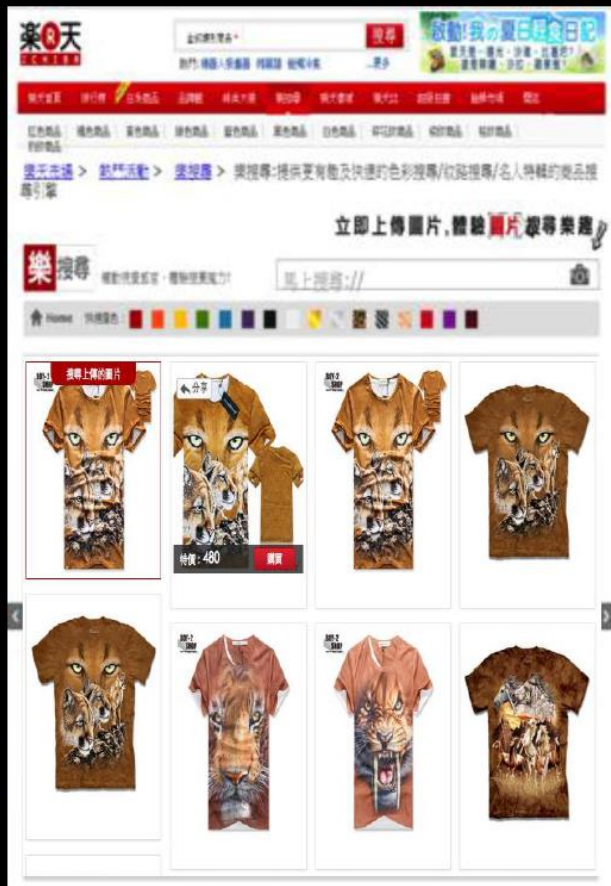
National Infocomm Awards 2014

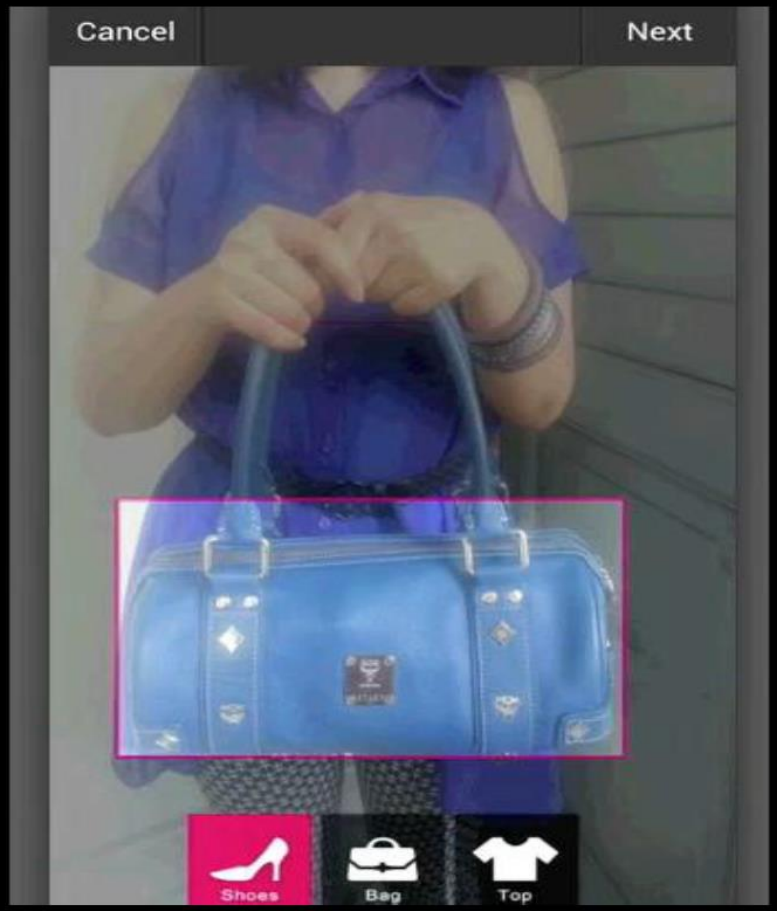
MOST INNOVATIVE SOLUTION

APICTA
Asia Pacific ICT Alliance
AWARDS

Application Tools & Platforms

VISUAL POWERED E-COMMERCE SITES

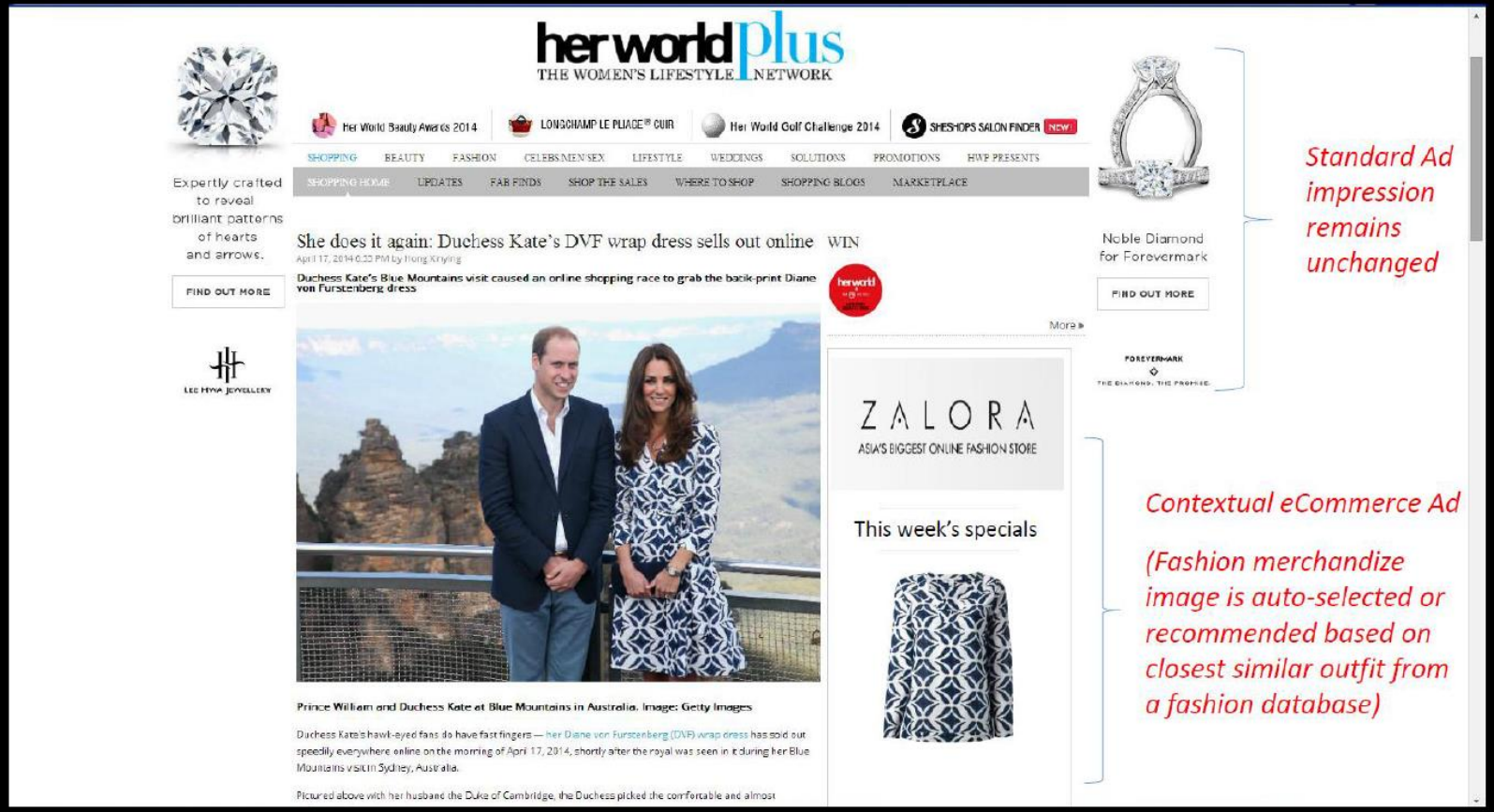




Snap Search Buy !

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 Current Visual Technologies

- Hardly anyone has deployable real-time in-video recognition technologies
- Video Recognition is still not widely used and talked about, even in research: *WHY?*

OUTLINE

- Retrospective on Image Search Since 2010
- Why Not Much Research in Video
- Some Current MM Research Efforts
- Summary

Current Needs of Industry

- Several urgent needs are all video based:
  - Visual recognition to in-video advertising
  - Live (video) event detection
  - Live Surveillance applications
- This can be seen in business/ industrial trends in social media..

Review of Social Media Platforms

- Social Network platforms
  - Three major platforms: Private, Professional, Public

- Media Sharing platforms
  - YouTube, Instagram, Flickr, .. , Periscope, Vine


- Social Messaging Platforms
  - Whatsapp, LINE, Wechat, SnapChat, …

★ Recent popular apps are all image/ video based..

- Social Curation Platforms: Pinterest …

Major Changes in Last few Years!

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



   - Live video from GoPro, health sensors & appliances, etc

2. Live aspect is central now
   - With Live, comes the ability for continuous sharing and feedback..
   - Users now want to get instant feedback: from friends and systems
   - Problems in understanding and reconstructing contents in live contents, including videos



SnapChat -1

- A Multimedia Machine?
- ***Why it is so popular?***

- "Temporal Real Time Messages": people decide how long others view their photos/videos

- Promotes creativity: Drawing, geo-filters & other 'add ons'.

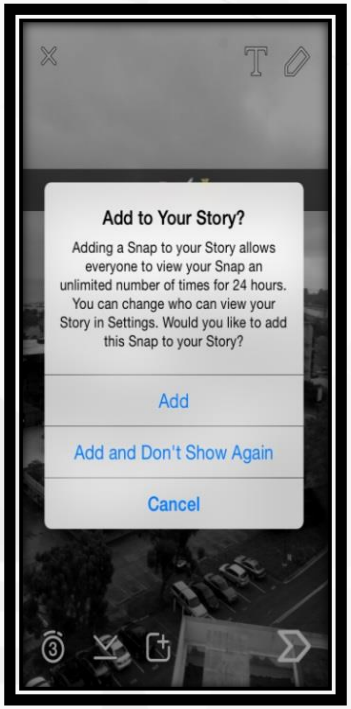
With one tap you can decide who sees the message.

Depending on where you are, it provides geo-related filters; this adds another layer for communicating.

SnapChat -2

- ***Why it is so popular?***
- People can view stories from friends, contribute to nearby stories eg “NUS Graduation”, view live stories, e.g. Oscars, Sporting Events, and Fashion Shows.

- The discover feature allows people to view news from top sources, eg CNN, ESPN…



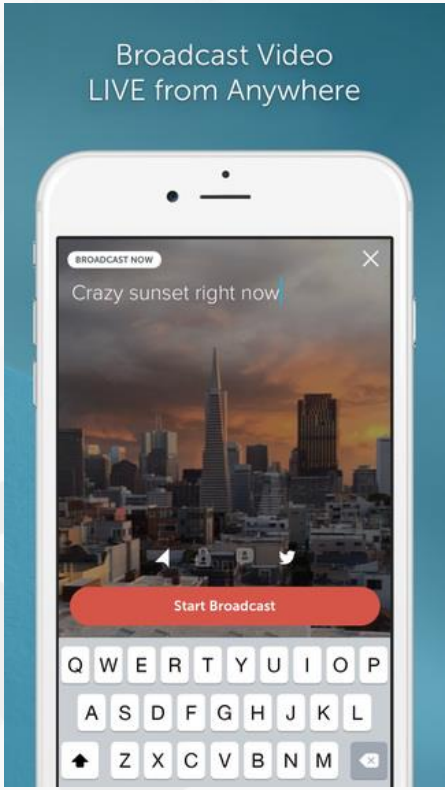
Creating a story, enables your friends to see what’s happening in your life through photos and videos. Paves the way for constant user engagement

Users are able to consume content by brands. Opening the door for a way to consume 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



Very Little Research in Video -1

- Given the needs and popularity of video in industry, there are insufficient research efforts on video
- Conference statistics in MM-related conferences:

Conference	Total #	# that is visual-related	# related to Video
MM2014	59	46	15 (32%)
MM2013	53	38	13 (34%)
MM2012	67	50	14 (28%)
ICMR2014	50	46	24 (52%)
ICMR2013	37	31	14 (45%)
ICMR2012	56	47	15 (32%)
CVPR2014	540	378	113 (30%)
CVPR2013	471	267	91 (57%)
CVPR2012	465	275	79 (29%)

Very Little Research in Video -2

- Major Research topics (though many might not actually be in video)
  - Event Recognition/Detection
  - Action Recognition
  - Object Tracking
  - Saliency Prediction
  - Video Segmentation
  - Video Classification
  - Geographic Location Tagging
  - Video Copy Detection
  - Summarization

Very Little Research in Video -3

- What are the major obstacles?
  - **Lack of large-scale datasets?**
  - **Ability to process large-scale videos**
  - **Too obsessed with visual analysis?**

An overview of existing Video datasets

(complement from Yigang Jiang)

Dataset	# Videos	# Classes	Year	Manually Labeled ?
Kodak	1,358	25	2007	✓
MCG-WEBV	234,414	15	2009	✓
CCV	9,317	20	2011	✓
UCF-101	13,320	101	2012	✓
THUMOS-2014	18,394	101	2014	✓
MED-2014	≈28,000	20	2014	✓
Sports-1M	1M	487	2014	✗
.....				

Recent Datasets

Dataset	Size	Source	Labels	Def. of Label	Publicly Accessible
MED	**~0.3 M**	**Youtube**	**40 Events**	**Self-Defined**	**~18%** (released 2014)
FCVID	**91,223**	**Youtube**	**239 Categories**	**Self-Defined**	**100%** (released 2015)
NUS/ CMU/ Yahoo!	**106,000 (to extend to 0.7 M)**	**Flickr** (consumer videos)	**420 Concepts + 100 Scenes**	**ImageNet + SUN**	**100%** (to be released soon)

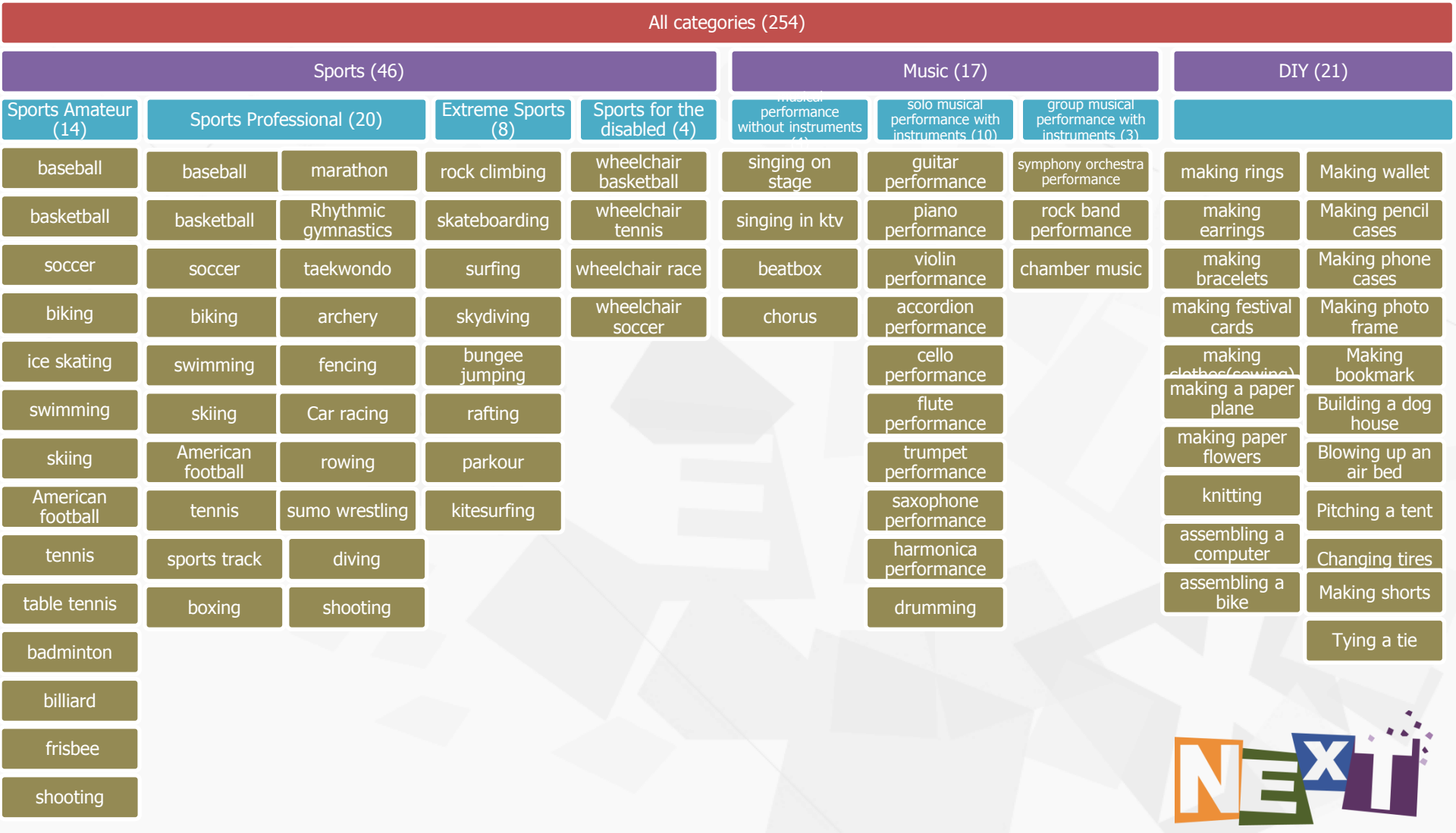
FCVID: Fudan-Columbia Video Dataset

- One of the largest public benchmark Internet videos with manual annotations (**239** categories)
- Covering many categories organized in a hierarchical structure
- Released in Feb. 2015!

  http://bigvid.fudan.edu.cn/FCVID/

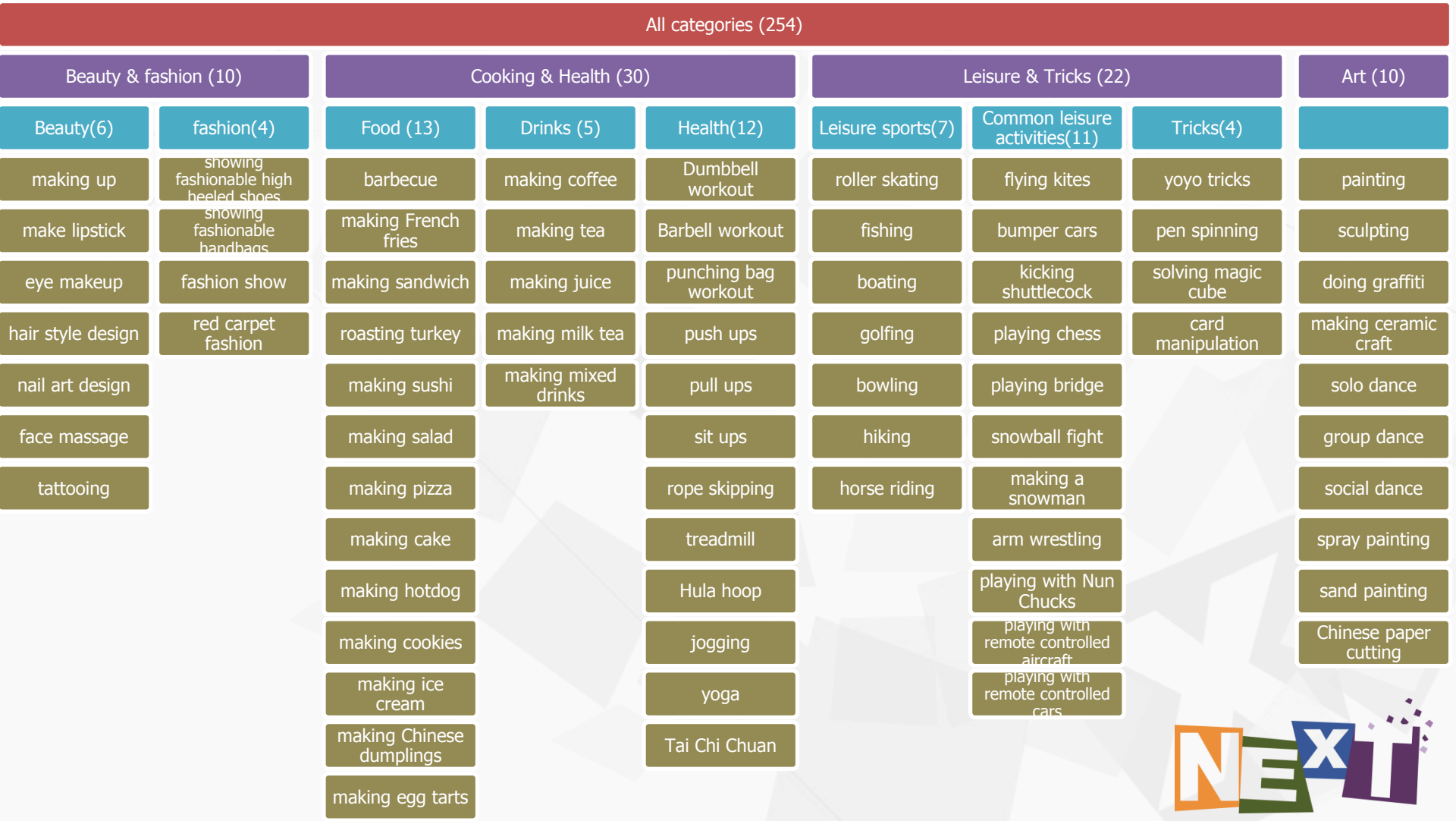
Fudan-Columbia Video Dataset (FCVID)

(239 categories; 11 higher level groups; 32 second level groups)



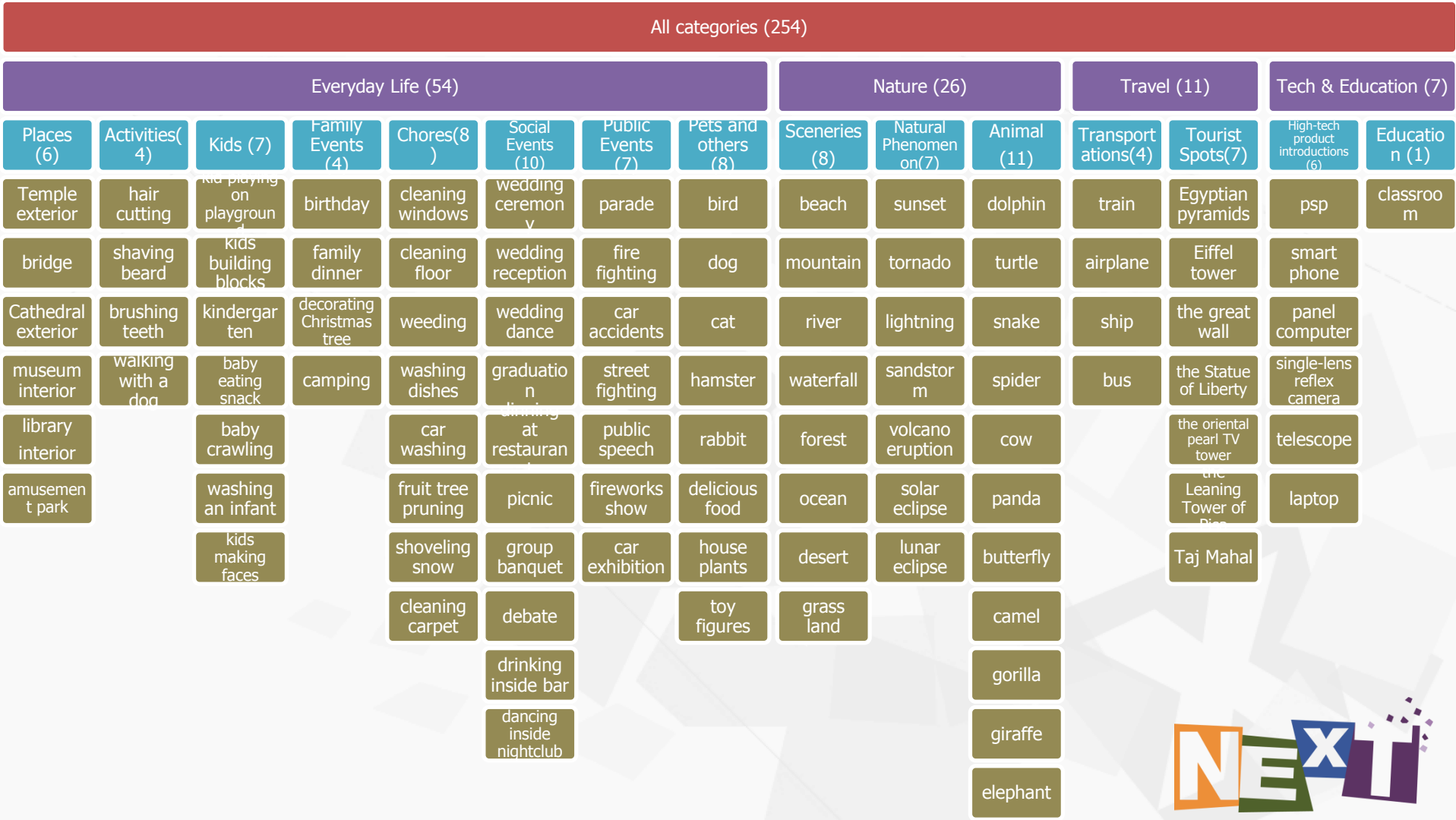
Fudan-Columbia Video Dataset (FCVID)

(239 categories; 11 higher level groups; 32 second level groups)



Fudan-Columbia Video Dataset (FCVID)

(239 categories; 11 higher level groups; 32 second level groups)



FCVID: Example 1

Everyday life

↓

Kids

↓

Kids playing blocks

FCVID: Example 2

Sports

↓

Sports Professional

↓

Boxing

NUS-CMU-Yahoo! Dataset

- Targets on consumer videos, mostly recorded by mobile devices
- Source and Scale: Yahoo Creative Commons 100M
  - 0.7 million videos from Flickr
- Targeted Annotations
  - 420 entry-level concepts from ImageNet, 100 scenes from SUN
  - To release tagging for 0.1M videos soon, the rest later

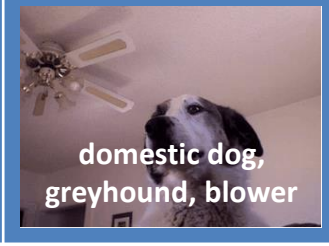
Concept Definitions

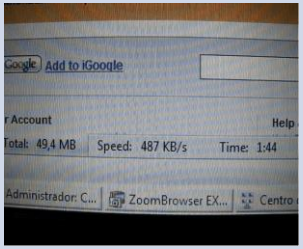
dog	flower	ball	wheel vehicle	vessel, water-craft	fish	insect	kitchen applia-nce	Compu-ter	gun
boxer, bull mastiff, great dane, husky, saint bernard, collie, kelpie, shetland, toy terrier, japanese spaniel,chihuahua, gordon setter, english setter	daisy, yellow lady‘s Slipper	ping-pong ball, volleyball, rugby ball, punching ball, soccer ball, tennis ball	Tricycle, motor scooter, unicycle	trimaran, catamaran, gondola, fireboat, submarine, canoe, yawl, speedboat, lifeboat, liner	coho, lionfish, garfish, barracouta, pufferfish, blowfish, eel, rock beauty, tench, goldfish, anemone fish	cricket, grasshopper, bee, ant, mantis, cockroach, walking stick, lacewing, fly, leafhopper, cicada, dung beetle, rhinoceros beetle, carabid beetle	microwave, toaster, waffle iron	web site, notebook, laptop, hand-held computer, desktop computer	six-shooter, assault rifle, cannon, rifle

Scene Definitions

INDOOR:				OUTDOOR NATURAL:			OUTDOOR MAN-MADE:		
shopping and dining	workplace	home or hotel	transportation	water, ice, snow	mountains, hills, desert, sky	forest, field, jungle	sports fields, parks, leisure spaces	houses, cabins, gardens, and farms	commercial buildings
banquet_hall, beauty_salon, Bookstore, butchers_shop, candy_store, coffee_shop, jewelry_shop, music_store, Pharmacy, barrel_storage	anechoic_chamber, control_room, Corridor, dentists_office, engine_room, hospital_room, veterinarians_office	Attic, Bedroom, home_office, living_room, Nursery, Pantry, utility_room, barrel_storage	bus_interior, engine_room, limousine_interior, subway_station/platform, van_interior,	Natural, Dock, ice_floe, Iceberg, Pond, Sandbar, ski_slope, Swamp, underwater/coral_reef	Badlands, desert/vegetation	field/cultivated, field/wild, golf_course, Pasture, putting_green, Rainforest, rice_paddy, Swamp	golf_course, hot_tub/outdoor, labyrinth/outdoor, Pavilion, putting_green, Racecourse, rope_bridge, ski_slope, stadium/baseball,	balcony/exterior, doorway/outdoor, field/cultivated, greenhouse/outdoor, House, hunting_lodge/outdoor, outhouse/outdoor, Patio, rice_paddy	Alley, apartment_building/outdoor, balcony/exterior, Crosswalk, doorway/outdoor, market/outdoor, office_building, phone_booth, Skyscraper

NUS-CMU-Yahoo! Dataset: Examples





NUS-CMU-Yahoo! Dataset: cont.

- Source and Scale: Yahoo Creative Commons 100M
  - 0.7 million videos from Flickr
  - To tag 420 entry-level concepts from ImageNet, & 100 scenes from SUN
  - To release tagging for 0.1M videos soon, the rest later
  - Support annotation, retrieval & video description tasks
- Plan to launch in MMM and/or ACM ICMR:
  - For interactive tassks and grand challenge tasks
- Encourage collaborations:
  - To tag other entities, and define/lead other tasks, like event detection

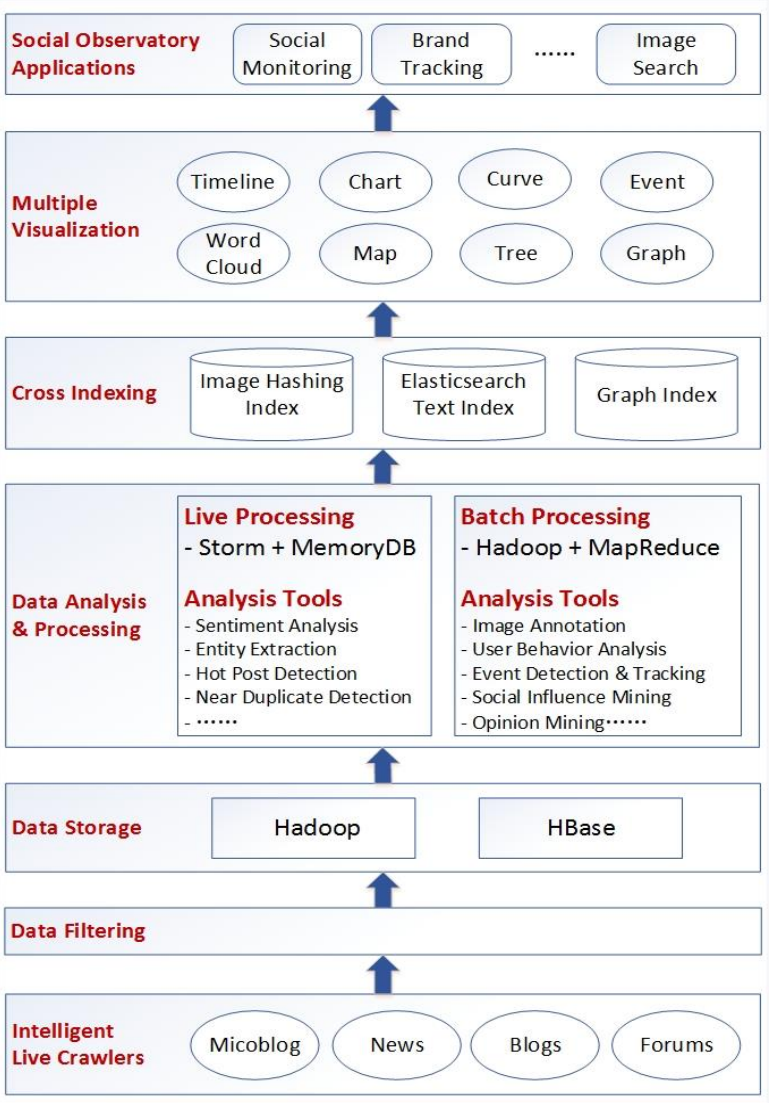
NOTE:

1. **Size**: largest when fully tagged
2. **Consumer-level**: first 100% consumer-level dataset
3. **Definition of Labels**: based on current popular ontologies; can directly use off-the-shelf concept or scene detectors pre-trained on ImageNet and SUN.

2. Ability to Process Large-Scale Videos

- Availability of Dataset is one of the main academic problem, but not the only problem
- The ability to process large-scale video is another:
  - Infrastructures for large-scale stream data
  - Technologies to process video in real-time
  - Mostly engineering tasks done in industry
- Works in such robust academic systems are beginning
  - For example, work by CMU in this conference: “Content-based video search over 1 million videos with 1 core in 1 second”
  - Need industrial scale infrastructures and expertise
  - Need technologies for visual audio and text analysis

NExT Architecture for Big Data Processing



- 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

3. Needs Multimedia Approach to Solve Real Problems

- All B2C apps use simple but reliable contextual info to drive applications
  - Mostly are ideas to help users communicate and have fun
  - For example, Geo-filter, Live Story and Discover features in SnapChat rely on manual curation and geo-tags
  - User profiling in social media products relies on check-in POI, user-defined profile, messages and images shared etc.
  - What visual techniques are needed?
- In contrast, B2B apps require to offer solutions to customers' problems
  - Most need technologies as basis/ barrier to entry

3. Needs Multimedia Approach to Solve Real Problems

- Because of the above, not many B2C ideas would come from academic community
- Our evaluation system values validated ideas with thorough evaluations, which could be the obstacles
  - This is especially the case when acceptance ratio is low
  - Hence hard for novel ideas to flourish..
  - We need to be conservative in pursing ideas – in name of publications
- How to address this problem?
  - Well recognized in ACM
  - We tried various means, including the introduction of “Brave New Idea” track

OUTLINE

- Retrospective on Image Search Since 2010
- Why Not Much Research in Video
- Some Current MM Research Efforts
- Summary

Need to Tackle Real Multimedia Problems

- Many real practical problems require multimedia, not visual, solutions
  - We need to make judicious use of text, image/video, location data etc.
- As illustration, I look at solutions to several problems using multimodal data:
  - MED Event Detection
  - Social TV
  - User Mobility
  - Event-Shop (Ramesh Jain)
  - Wellness

1) Event Detection in Video

- Given a video, determine if an event has occurred
- One key problem is to identify what the event-related concepts
  - Many current solutions employ complex visual analysis and corpus correlations to determine important related concepts

- For example,

  event – Changing a vehicle tire

  semantic concepts -

  vehicle,
  bicycle wheel,
  car wheel …

Why not use domain knowledge: From FrameNet Project

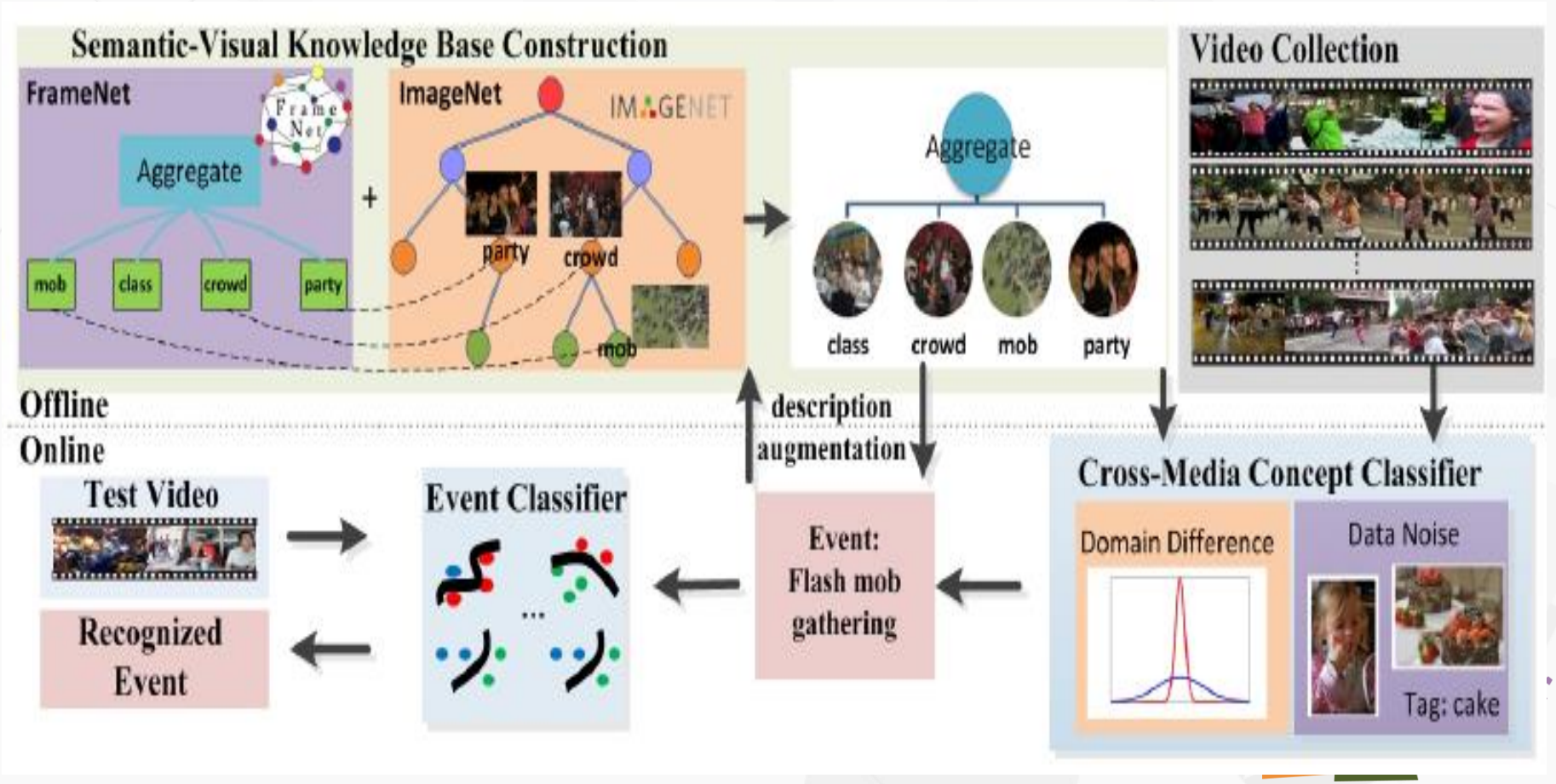
- What is Frame?
  A *frame* is a data-structure for representing a stereotyped situation, like going to a child's birthday party
  - **top levels** of frame:
    situation (e.g. going to a child's birthday party)
  - **lower levels** of frame:
    specific instances or data *(e.g. song birthday song, games, cake, party-meal ...)*
- Widely used in NLP research

[Marvin Minsky, A Framework for Representing Knowledge, 1974]

Frame	FrameNet LUs
Social_Event	banquet.n dance.n dinner.n party.n celebration.n ...
Forming_Relationships	betrothal.n divorce.n engagement.n wedding.n ...
Personal_Relationship	romance.n spouse.n marriage.n husband.n engagement.n...
Vehicle	bicycle.n bus.n cab.n car.n convertible.n limousine.n ...
Vehicle_Subpart	engine.n seatbelt.n tire.n wheel.n windshield.n ...
Leadership	premier.n prince.n president.n lawmaker.n...
Grooming	comb.v shampoo.v wash.v shower.v soap.v...

System Overview

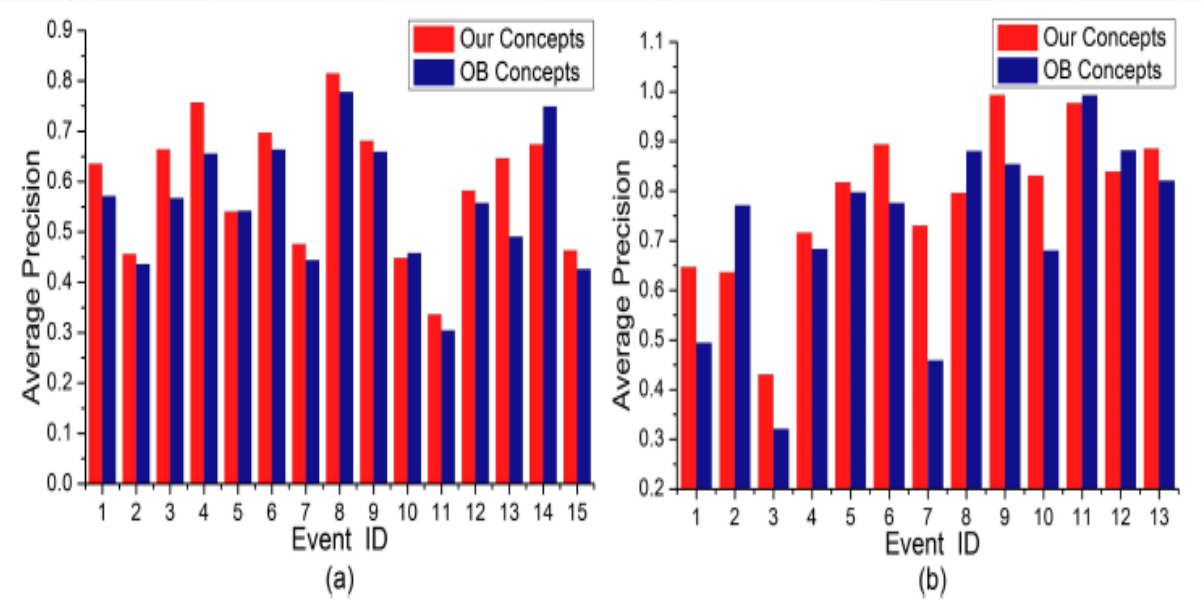
- Event: Flash mob gathering



Effects of Event-related Concept Set

- Video Event Recognition

  Comparison with OB concepts [1] consisting of 177 most frequent objects on: (a) MED11, and (b) EVVE

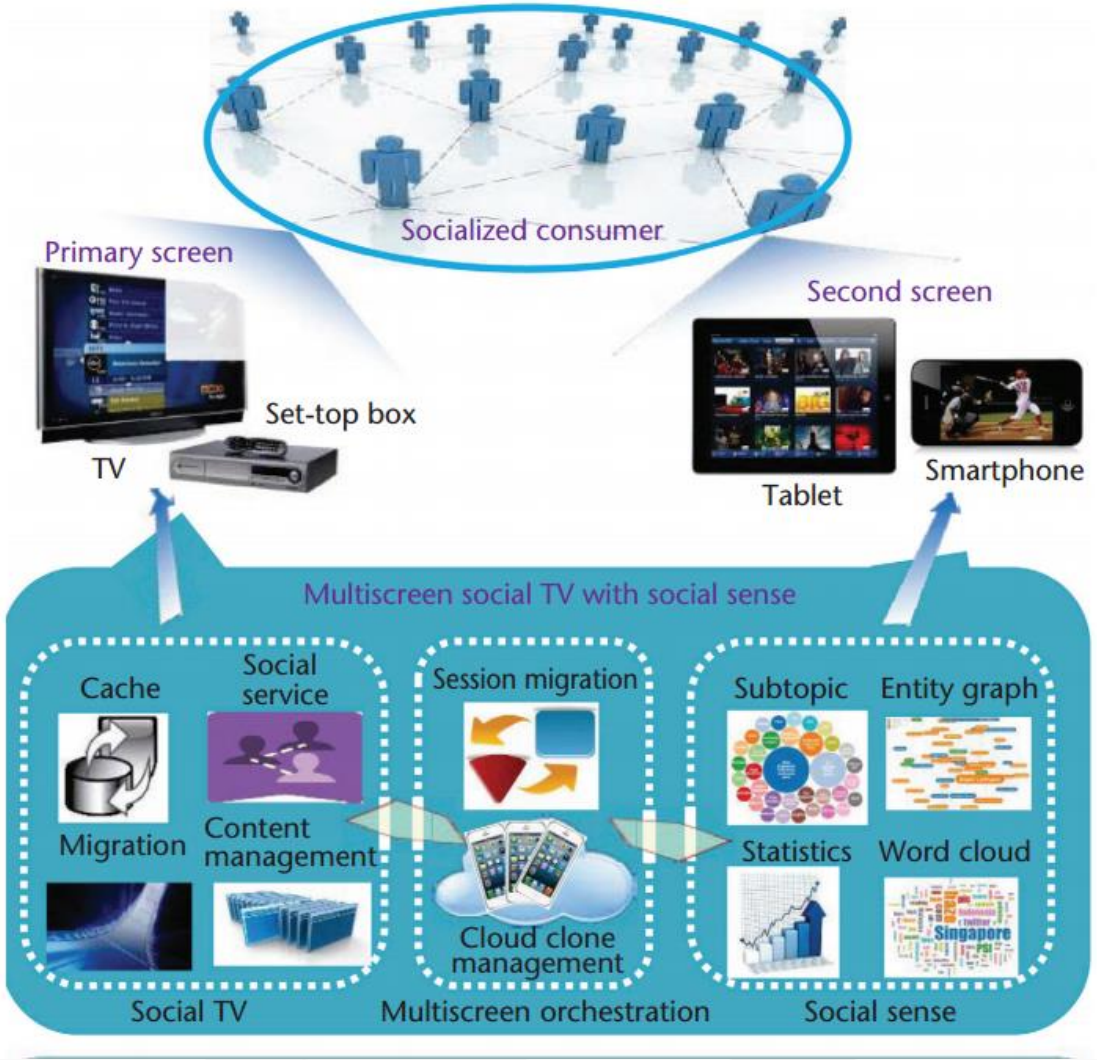


To appear in IEEE ToM

[1] L.-J. Li, H. Su, L. Fei-Fei, and E. P. Xing. Object bank. NIPS, 2010.
[2] J. Revaud, and H. Jegou. Event retrieval in large video collections with circulant temporal encoding. CVPR, 2013.

2) Social TV

- Utilize second screen to offer simultaneous social media discussions on live broadcast TV shows:
  - to provide social and interactive features simultaneously with the TV viewing experience.



Live Social Media Analytics

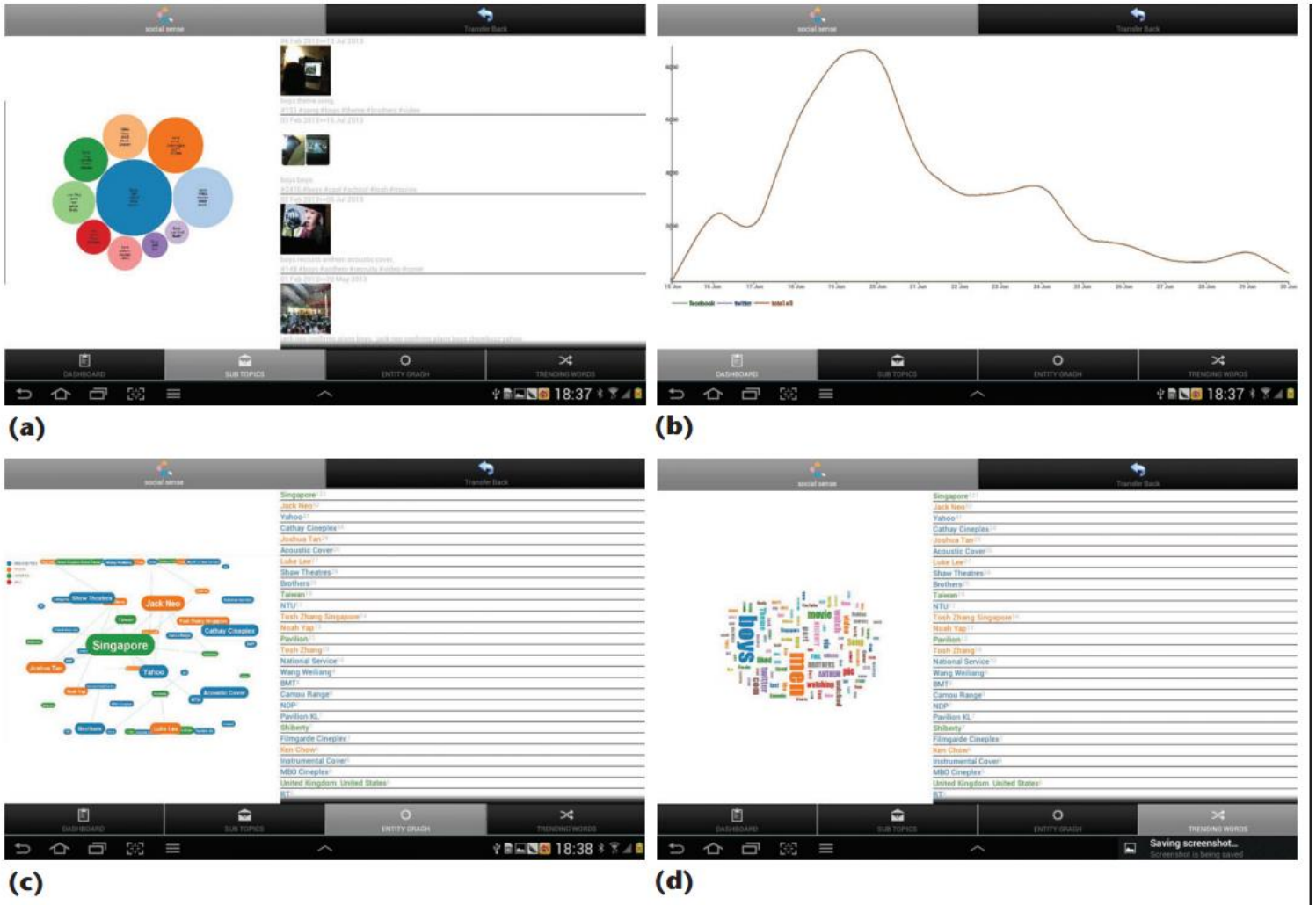
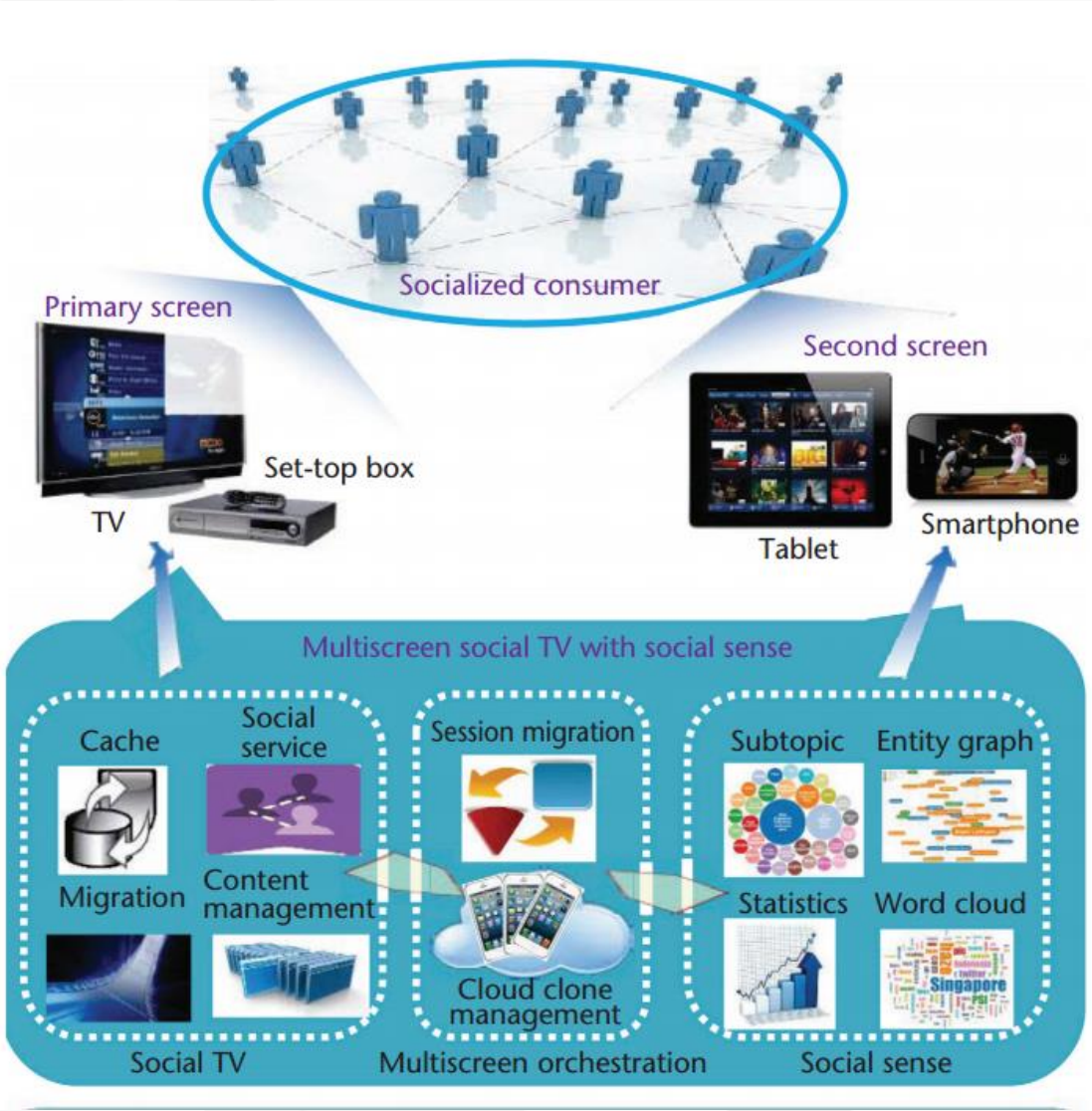


Figure 5. Social sense features. (a) Topics, (b) degree of interest, (c) entity graph, and (d) keyword cloud.

2) Social TV

- Utilize second screen to offer simultaneous social media discussions on live broadcast TV shows:
  - to provide social and interactive features simultaneously with the TV viewing experience.

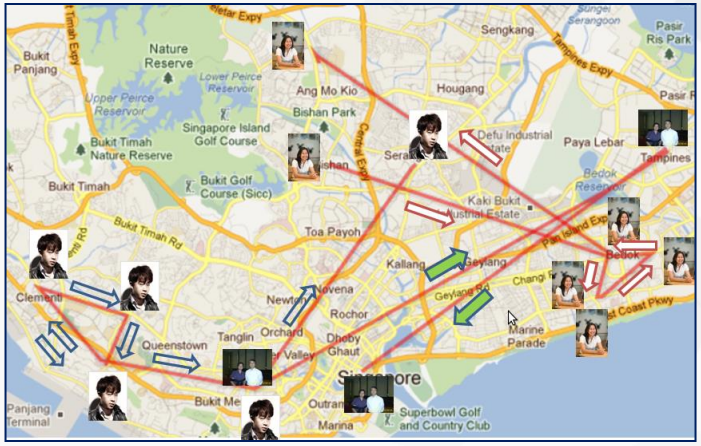


Selected as Best Paper for IEEE MM 2015

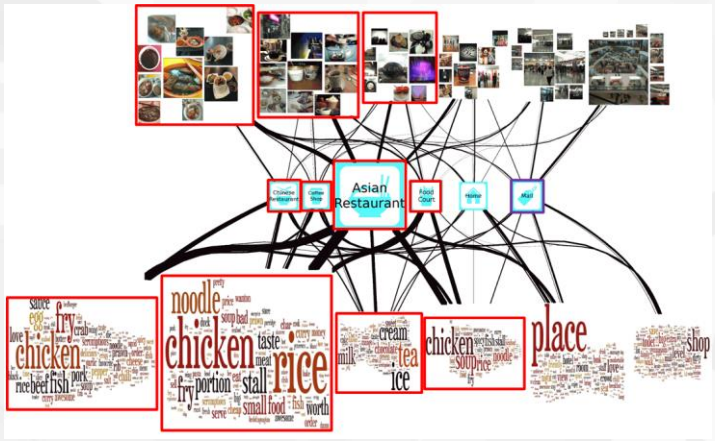
H Hu, et al. (2014): Toward Multiscreen Social TV with Geolocation-Aware Social Sense. IEEE MultiMedia 21(3): 10-19

3. User Mobility

- Utilize data from multiple social media sources
- Mine relations between check-in venues and UGCs
- 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



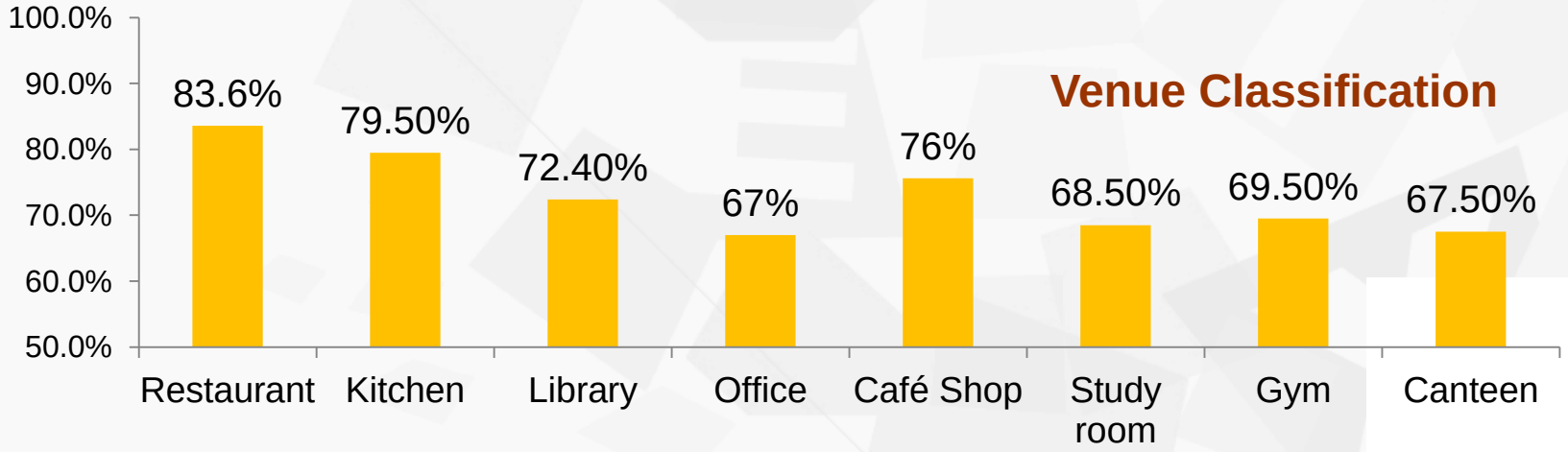
Travel paths and flows of users



Interest community of food lovers in Singapore

Multimodal Location Estimation

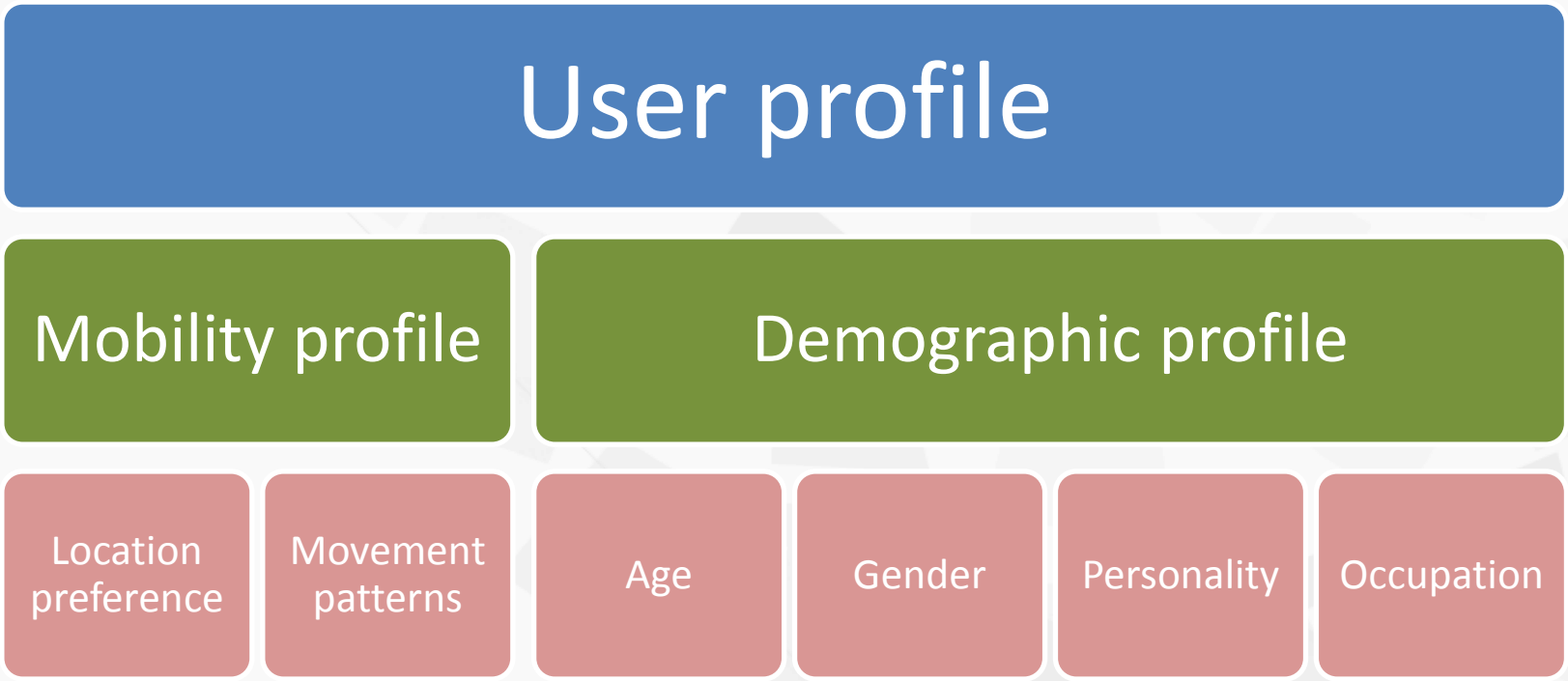
- **POI Estimation**
  - GPS trace of users, and messages if any
  - Estimate the POI's that users have visited
- **Mapping to POI (Point of Interests)**
  - Use nearby venues found in 4Square, and other info
  - Utilized audio signals to estimate venue categories
  - Essential for micro-videos in social media sites



- Classifier: POI of messages

User profile: Mobility + Demography

- **Estimate User Demography**
  - Utilize multi-source social media data, & if available, mobile phone traces
  - For Age and Gender – to over 80% accuracy



- **To present in this conference on Thu afternoon**

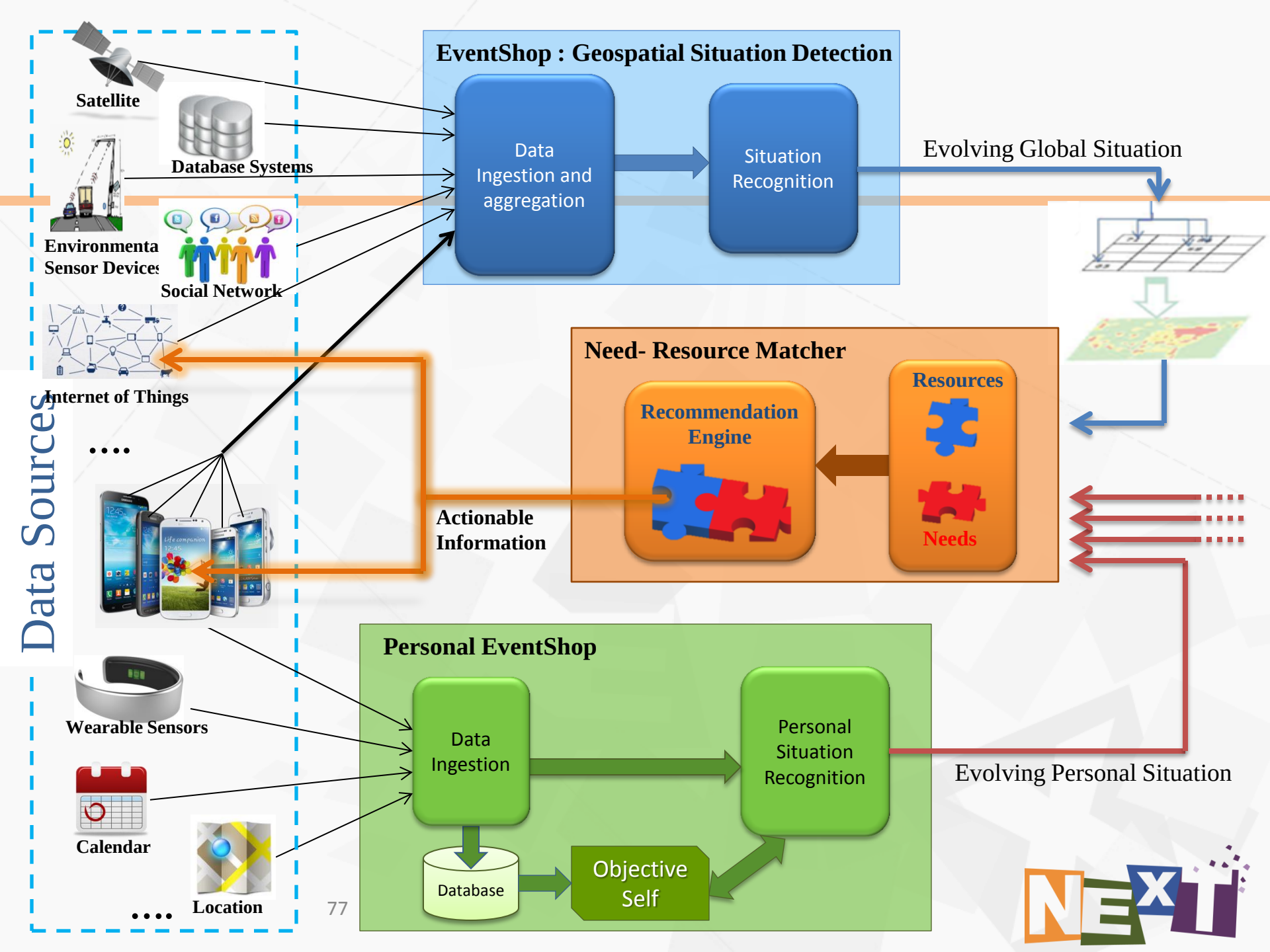
4) Geospatial Interpolation Analytics

“Social Life Network”

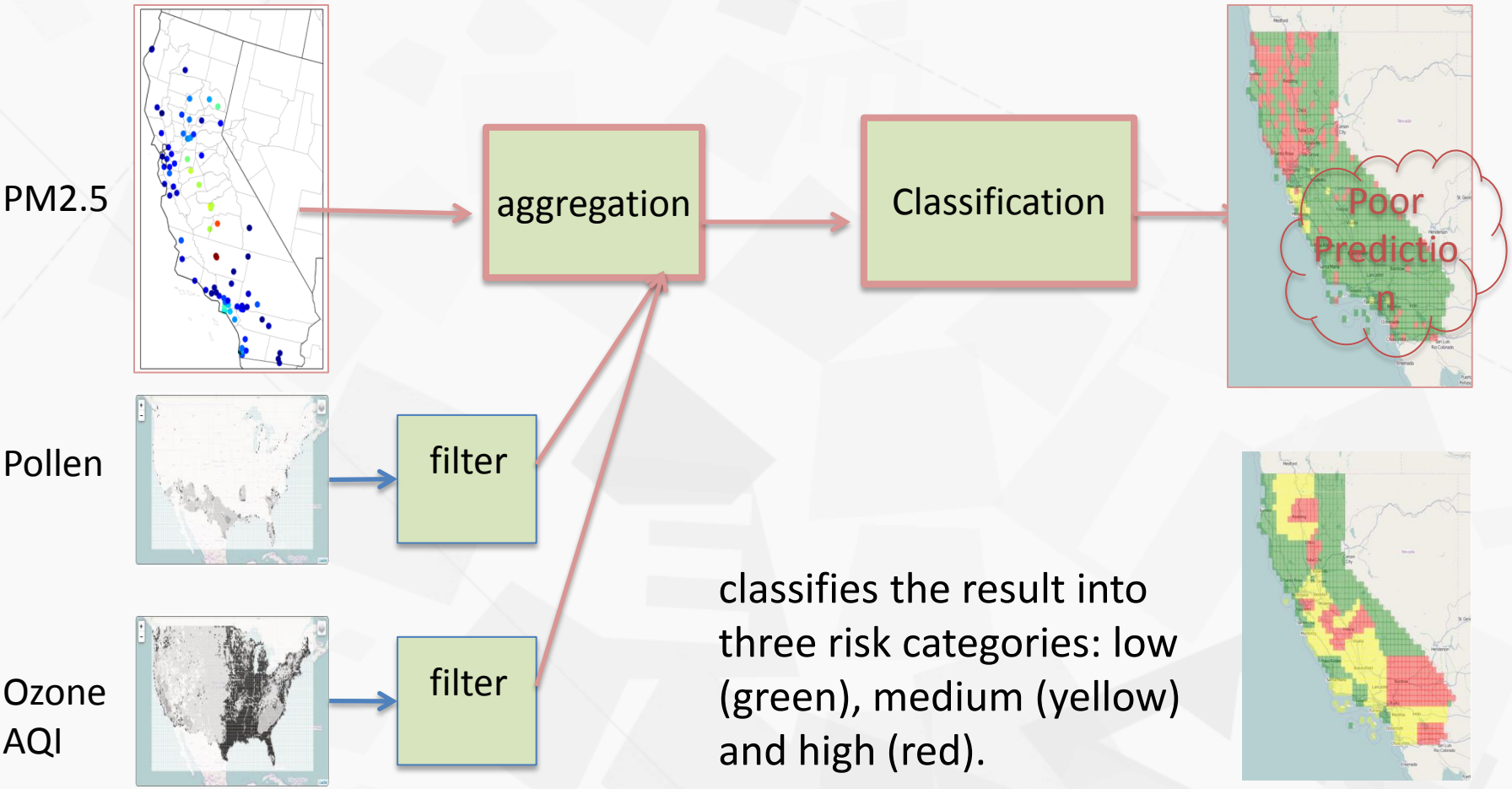
Providing right RESOURCES

at the right TIME

in the right PLACE


Data Sources
Satellite
Database Systems
Environmenta
Sensor Devices
Social Network
Internet of Things
....
Wearable Sensors
Calendar
Location
....
EventShop : Geospatial Situation Detection
Data Ingestion and aggregation
Situation Recognition
Evolving Global Situation
Need- Resource Matcher
Recommendation Engine
Resources
Needs
Actionable Information
Personal EventShop
Data Ingestion
Personal Situation Recognition
Database
Objective Self
Evolving Personal Situation

Asthma Risk Management Application



Geospatial Interpolation Analytics for Data Streams in EventShop

Nominated as Best Paper Candidate in ICME 2015

Asthma Hospitalization

5) Computational Wellness

HealthSenze gathers & organizes big-data from multiple sources

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

NEXT

OUTLINE

- Retrospective on Image Search Since 2010
- Why Not Much Research in Video
- Some Current MM Research Efforts
- Summary

Summary

- Current states of MM applications:
  - Technologies in media search have evolved and matured to tackle applications in certain vertical domains
  - Wide range of data, knowledge sources and social network data available online – for various wide range of applications
- Why little research in video?
  - Live online videos feature in most recent popular social apps
  - Lack of datasets and infrastructures?
  - Need to look into multi-modal approach to tackle problems
- Why lack of good industry ideas:
  - Technical research supports good B2B app
  - But not B2C which based on ideas leveraging robust technologies
  - Academic research not gears towards that
  - Need fundamental change of how to promote novel ideas

Summary

- To address problems raised:
  - Work on truly multimedia data, not just visual
  - Work on some risky new ideas, not just matured formula
  - Work on large-scale real-life (end-to-end) problem

THANKS

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

ACM ICMR Conferences

- ICMR2016: NYC, USA
  - John Smith (IBM Research), John Kender (Columbia Univ.)
- ICMR2017: Bucharest, Romania
  - Nicu Sebe (Univ. Trento), Bogdan Ionescu (Univ. Politehnica Bucharest)
- ICMR 2018 will be in Asia:
  - Call for organization out soon, please prepare your bid to support the conference