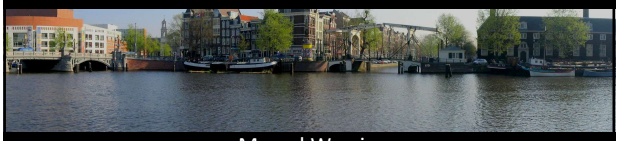
Amsterdam Data Science

Analyzing large multimedia collections in an urban context



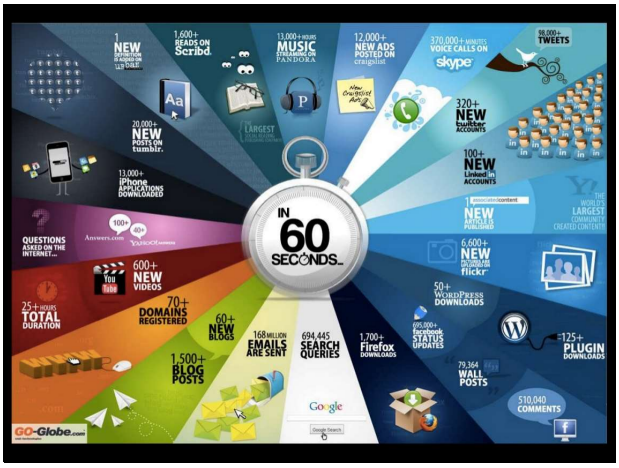
Marcel Worring

Stevan Rudinac, Jan Zahalka, Dennis Koelma

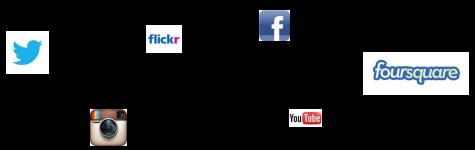
Joost Boonzajer Flaes, Jorrit van den Berg

Informatics Institute, Amsterdam Data Science

Amsterdam Data Science



Objective and Subjective data



Open Data

Structured data

Numeric data

Temporal data

Geographic data

Image data

Textual data

Unstructured data

Data Sources

Geo location

- Amsterdam, Netherlands

Exif

- Camera: Nikon N60
- Focal length: 55 mm
- Exposure time: 1/200
- Flash: off

Author

- josemanuelerre (Flickr)
- José Manuel Ríos Valiente

Tags

- cyclist
- bike
- street

Comments

- "I love Amsterdam! great photo!"
- "Great compostion, beautiful B&W!!"
- "Estupendo B&N, bella imagen."

...

Data Sources

- *"Koningsdag, or 'King's Day' is one of the principal holidays of the Netherlands. . . "*

- In this case, the **image** says more than the **text**

Photo: quantz @ Flickr

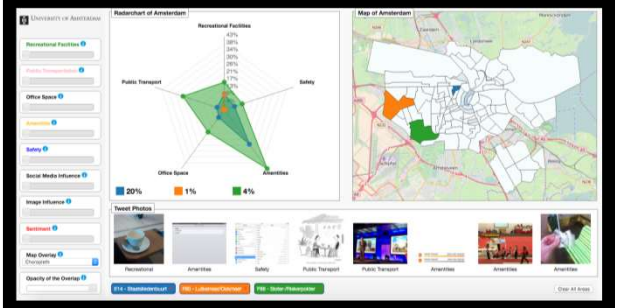
Objective and Subjective data

Open Data

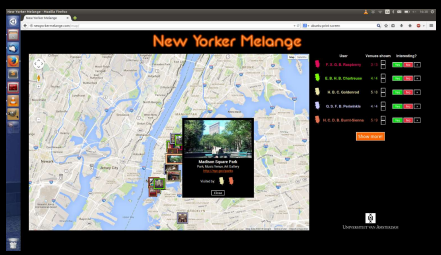
+ Content Analysis

WHAT DOES IS BRING?

Professional Recommender Systems

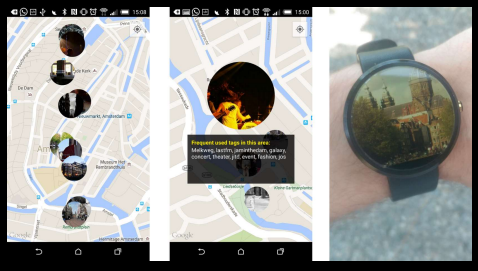


Recommender system for tourists

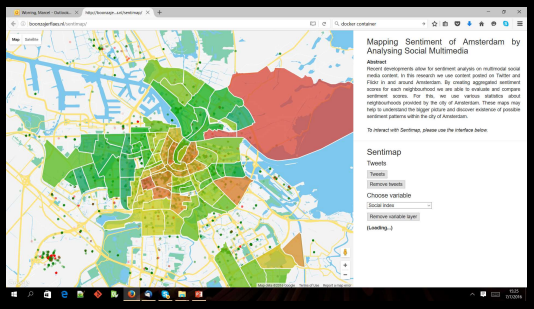


11

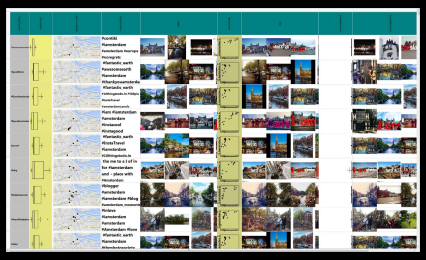
Touristic Routing



City Sentiment

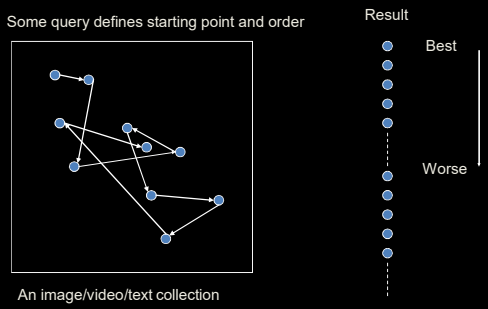


City Marketing Analytics



ALGORITHMS

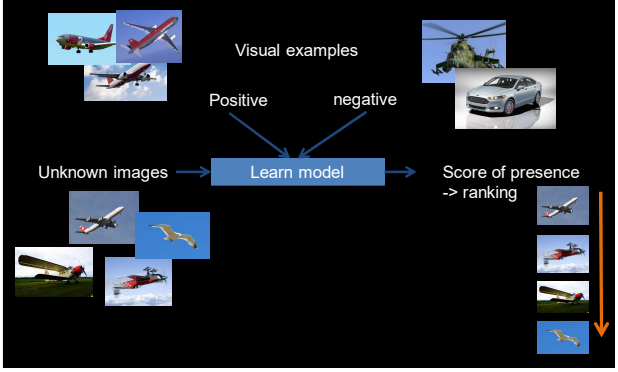
Ranking of data



For Social Media

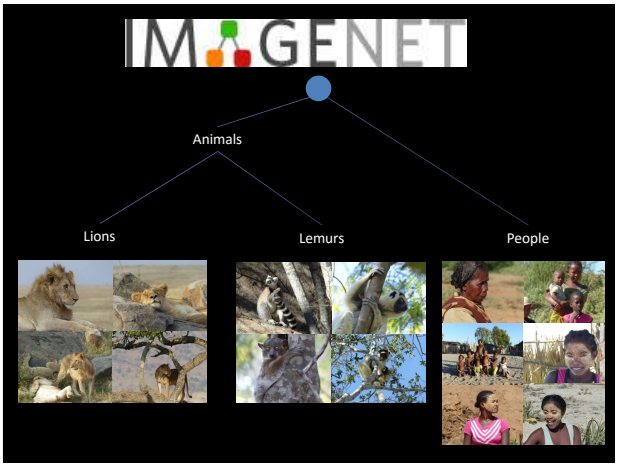
- The Ranking can be based on
  - The objective content of the comments
  - The subjective content of the comments
  - The objective visual content
  - The subjective visual content
  - ……
- Or any combination of the above

Concept detection



Requires annotation to learn

Zebu



What do we learn?

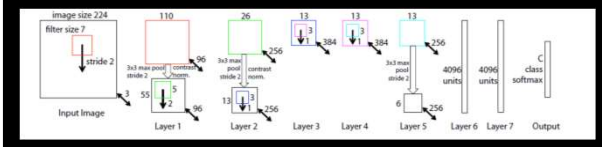
IMAGENET

14,197,122 images, 21841 synsets indexed

Visual Sentiment Ontology

1200 trained visual concept detectors for adjective-noun pairs

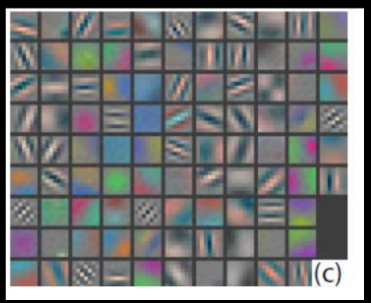
The new trend: Deep learning



Start with raw pixels, learn all parameters

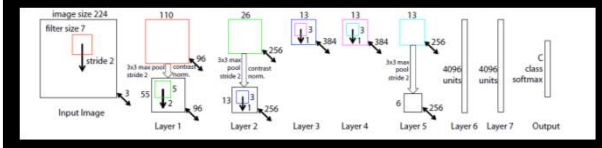
Krishevsky NIPS 2012

The learned filters



Zeiler and Fergus

The layered network



Convolution + pooling + fully connected layers + output layers

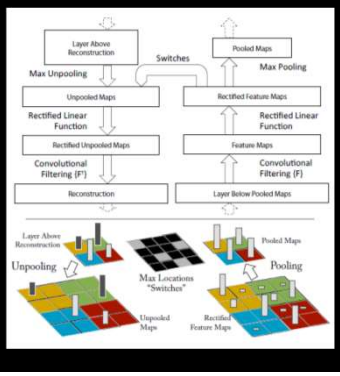
60.000.000 parameters to learn

But what do all these layers do?

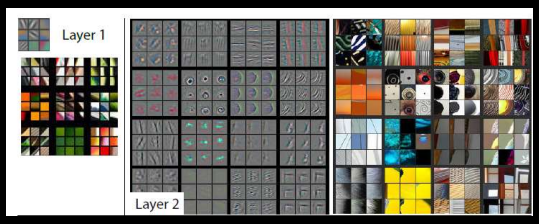
Krishevsky NIPS 2012

Zeiler and Fergus

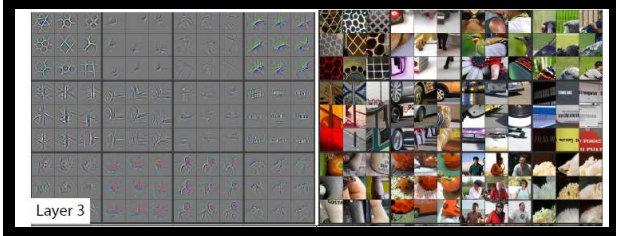
Visualizing deep networks



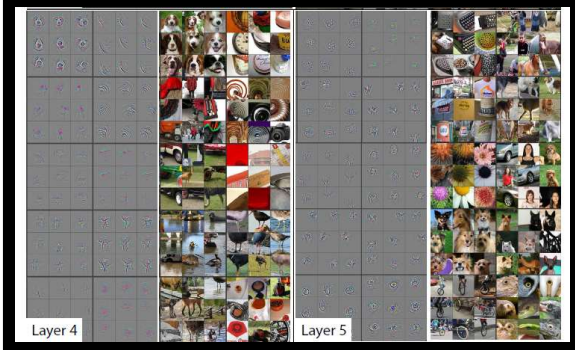
Visualizing deep networks



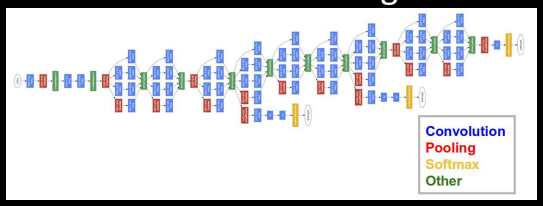
Visualizing deep networks



Visualizing deep networks



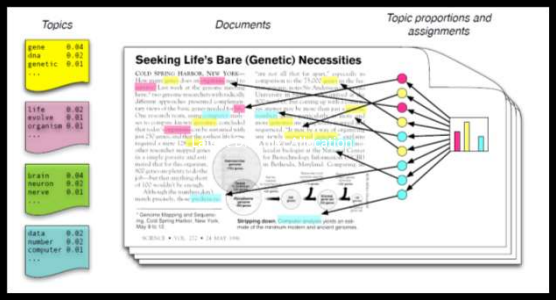
State-of-the-art: GoogleNet



and growing ……

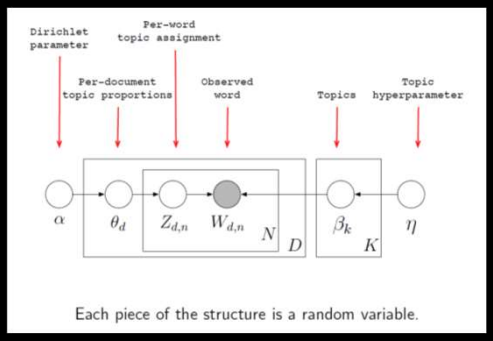
Makes image search keyword driven

Text Analysis



Latent Dirichlet Allocation D. Blei, 2003

Latent Dirichlet Allocation



D. Blei, 2003

Latent Dirichlet Allocation

- **Generative model**, discovers topics and scores them
- **100 topics** are enough to sufficiently cover **entire Wikipedia**
- **Input**: Raw text
- **Output**: Topic scores per document

```
0.054*mexico + 0.049*forest + 0.024*argentina
     + 0.022*islands + ...+ 0.014*aires
```

We treat comments or sets tags as documents

VENUE RECOMMENDER

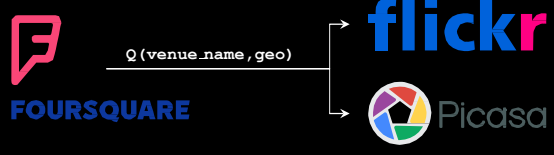
The task

- **Venue recommendation** — suggesting places of interest (**venues**) based on user preferences
- The classic approach is **collaborative filtering** utilizing the **user-item matrix**

City Melange Characteristics

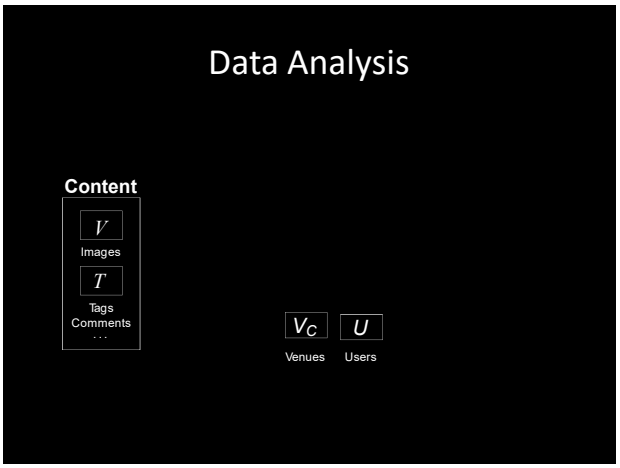
- **City Melange** — a venue explorer utilizing **multimedia analytics** techniques
- **Content-based** — based solely on the content of venue-related social media
- **Multimodal** — combining content from images *and* the associated text
- **Interactive** — user preferences are modelled on the fly as you explore the city
- **Cross Platform** — integrates data from diverse social platforms

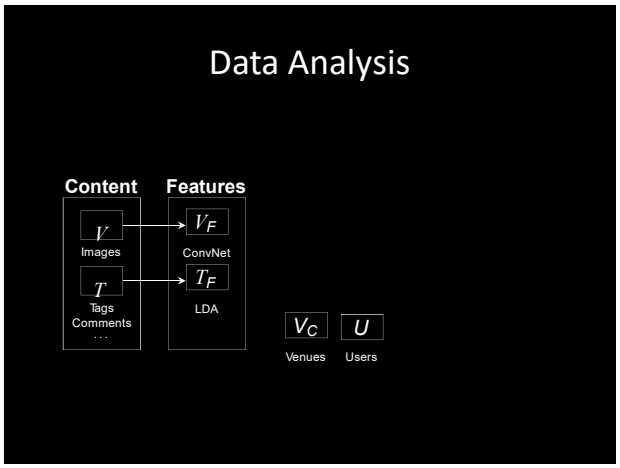
Data Gathering

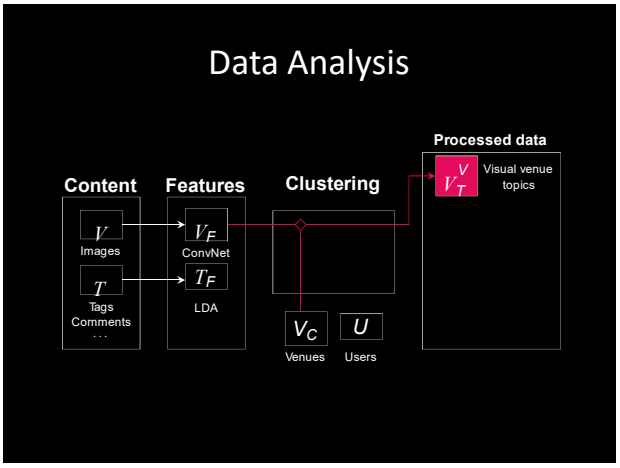


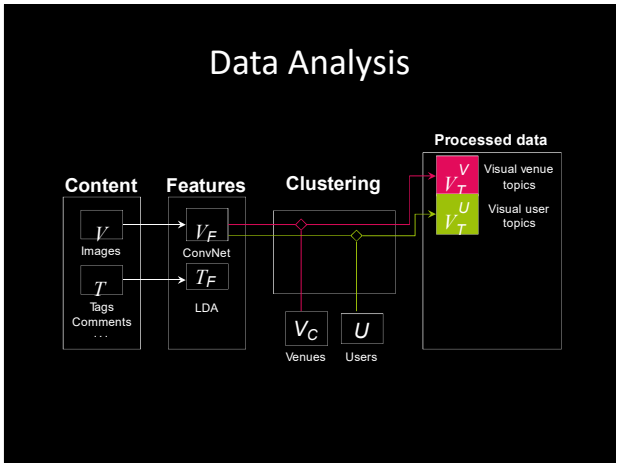
Venue information
Venue images

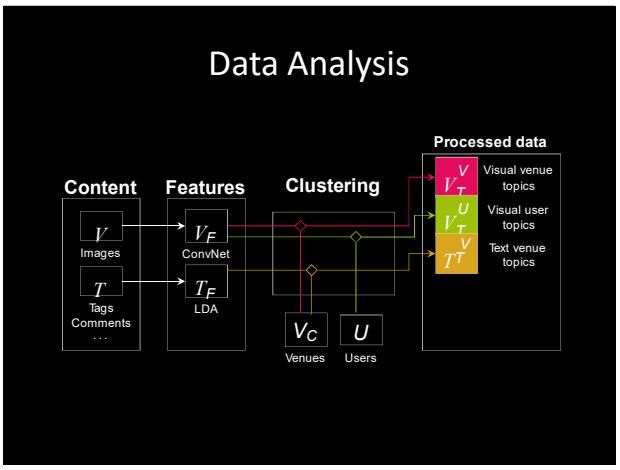
Images, metadata
User data

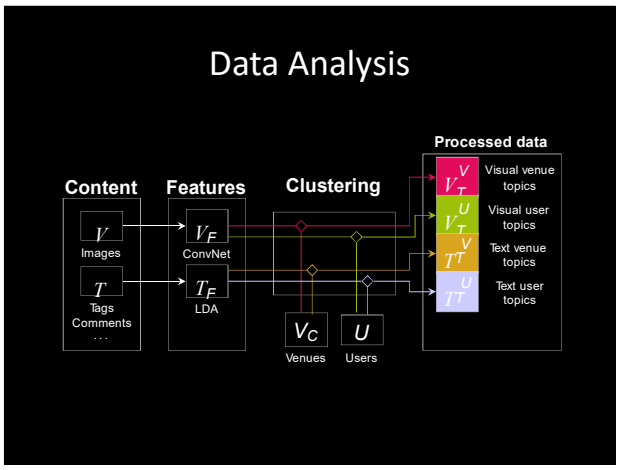

Data Analysis
Content
V
Images
T
Tags
Comments
Vc
Venues
U
Users


Data Analysis
Content
Features
V
Images
T
Tags
Comments
VF
ConvNet
TF
LDA
Vc
Venues
U
Users

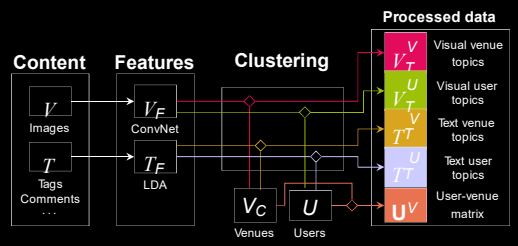

Data Analysis
Content
Features
Clustering
Processed data
V
Images
T
Tags
Comments
VF
ConvNet
TF
LDA
Vc
Venues
U
Users
V T
Visual venue topics


Data Analysis
Content
Features
Clustering
Processed data
V
Images
T
Tags
Comments
VF
ConvNet
TF
LDA
Vc
Venues
U
Users
V T
Visual venue topics
U T
Visual user topics


Data Analysis
Content
Features
Clustering
Processed data
V
Images
T
Tags
Comments
VF
ConvNet
TF
LDA
Vc
Venues
U
Users
V T
Visual venue topics
U T
Visual user topics
V T T
Text venue topics


Data Analysis
Content
Features
Clustering
Processed data
V
Images
T
Tags
Comments
VF
ConvNet
TF
LDA
Vc
Venues
U
Users
V T
Visual venue topics
U T
Visual user topics
V T T
Text venue topics
U T T
Text user topics

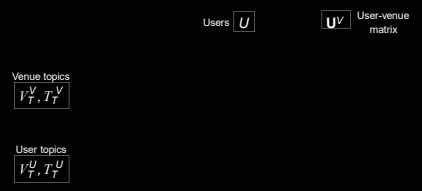
Data Analysis



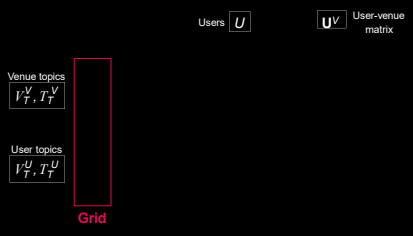
New Yorker Melange

- ACM Multimedia Grand Challenge 2014 **1st Prize**
- `newyorkermelange.com`

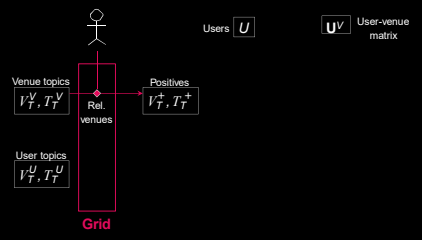
Interactive Recommendation



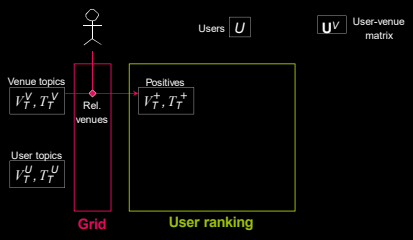
Interactive Recommendation



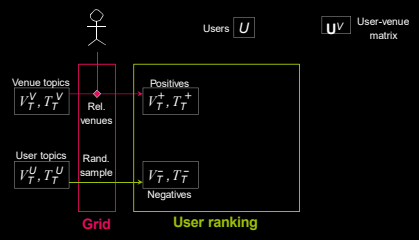
Interactive Recommendation



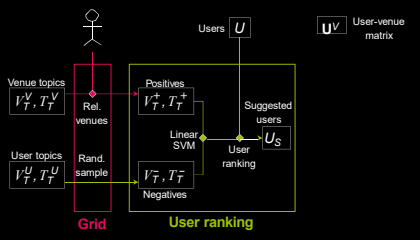
Interactive Recommendation



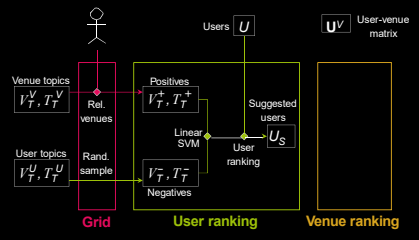
Interactive Recommendation



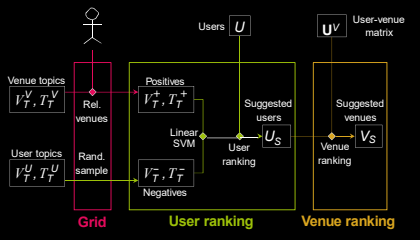
Interactive Recommendation



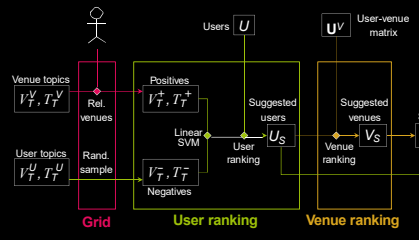
Interactive Recommendation



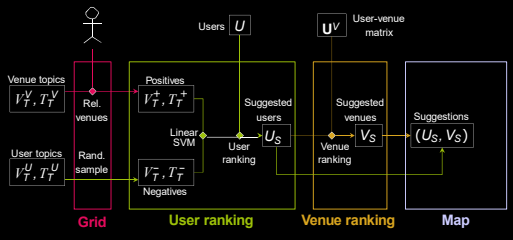
Interactive Recommendation



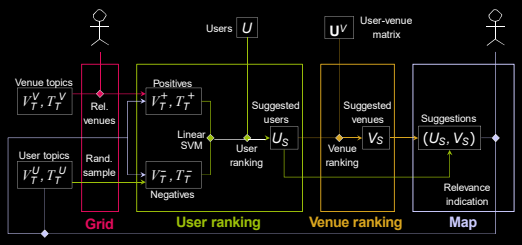
Interactive Recommendation



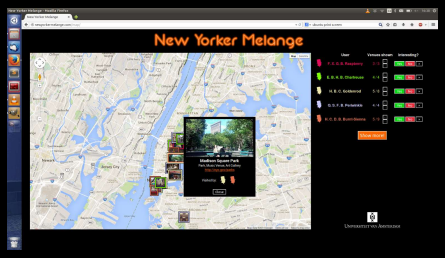
Interactive Recommendation



Interactive Recommendation



Recommender system for tourists



56

Evaluation

1. Can we recommend the **right type of venue**?
2. Can we recommend **mainstream venues to mainstream tourists** and **specialized venues to afficionados**?

Evaluation

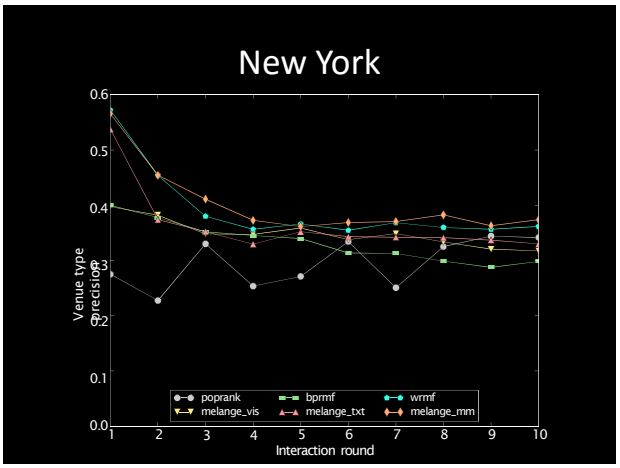
- **621 fine-grained venue types** (Japanese restaurant, skate park. . .)
- **100 artificial actors**, use 75% of the data to seed Melange
- Perform **10 interaction rounds**

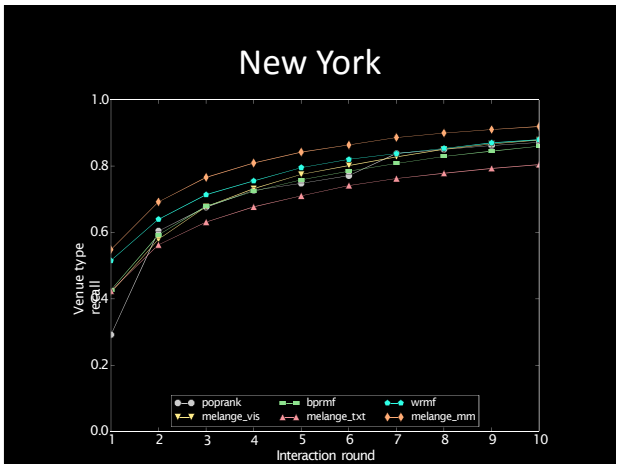
Methods Compared

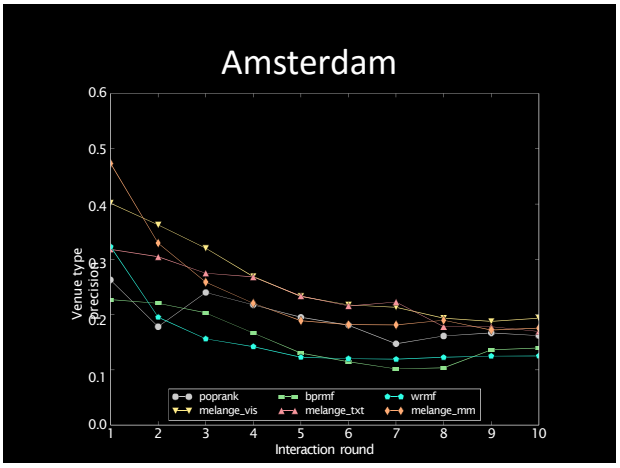
- City Melange
  - Visual modality only
  - Text modality only
  - Multimedia (vis + txt)
- Recommender baselines
  - WRMF — Weighted regularized matrix factorization
  - BPRMF — Bayesian personalized ranking matrix factorization
- Popularity ranking (PopRank) — most visited venues according to Foursquare

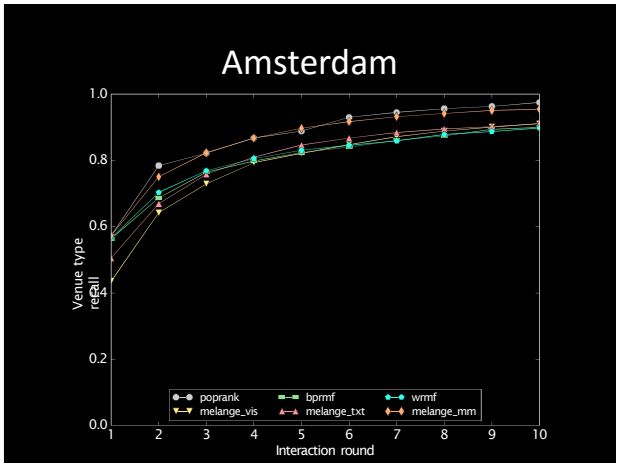
Data Collection

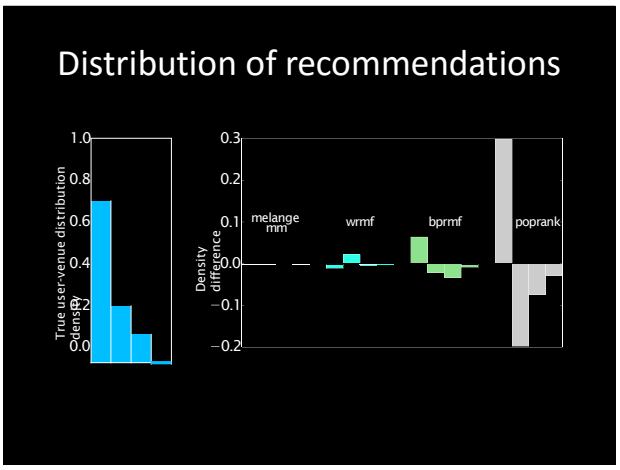
- **New York** — 1.07M images and associated text from Foursquare, Flickr, and Picasa
- **Amsterdam** — 56K images and associated text from Foursquare and Flickr









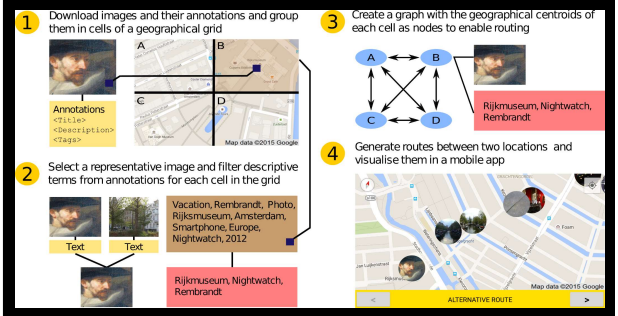


TOURIST ROUTING

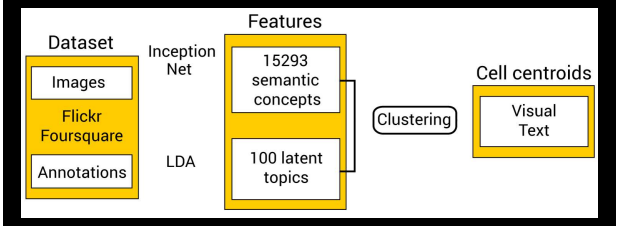
SceneMash

- **Data collection**
  - 150,000 geotagged Flickr and Foursquare images from the region of Amsterdam
  - Metadata associated with the images:
    - image title
    - description
    - tags
    - geotags

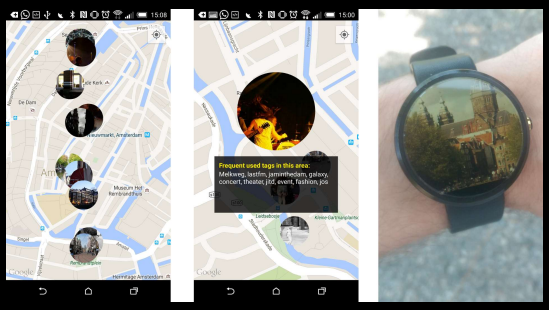
SceneMash



SceneMash



SceneMash



Demo

CITY SENTIMENT

Data Collection

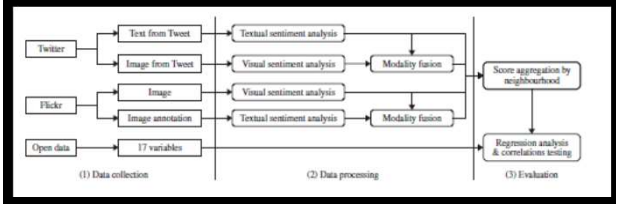
Amsterdam Neighborhoods

64K GeoTagged Tweets with Images

64K GeoTagged Images and comments

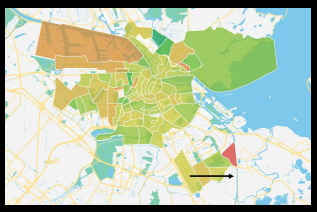
various neighborhood statistics (17 variables)

Methodology



Sentiment Maps

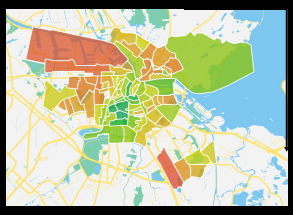
Sentiment analysis



Sentiment Maps

flickr

Sentiment analysis



Finding correlations

textual and visual content

flickr

textual and visual content

City of Amsterdam

various statistics

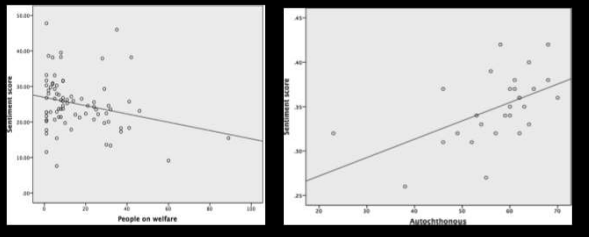
Sentiment analysis

Correlation Analysis

Correlations

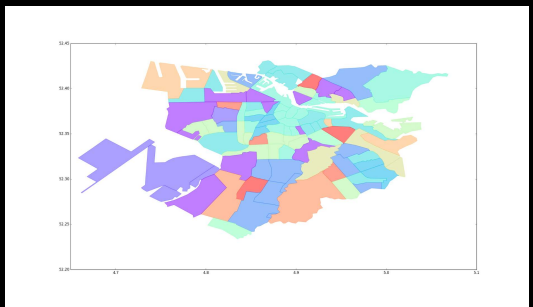
Flickr

Twitter

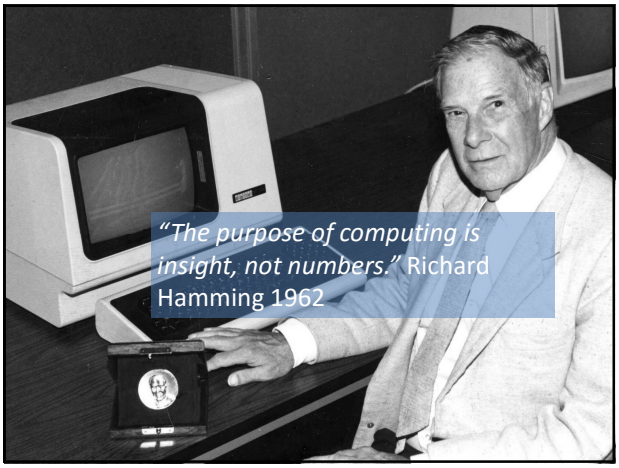


Correlations are only found with multimodal sentiment

Redefined Neighborhoods



MARKETING ANALYTICS



So what we want?

Insight

What is insight?

omplex
Insight is complex, involving all or large amounts of the given data in a synergistic way, not simply individual data values.

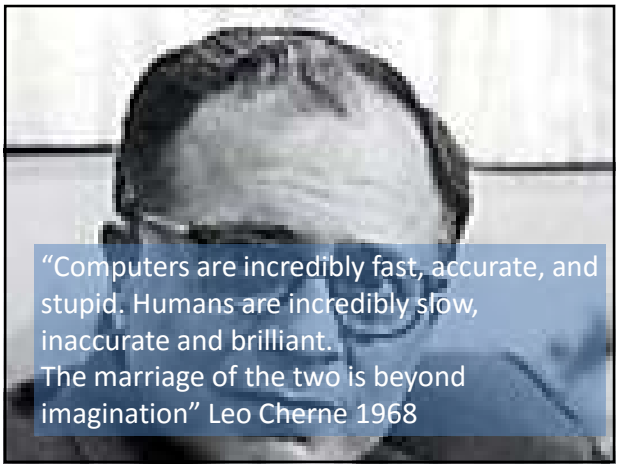
Deep
Insight builds up over time, accumulating and building on itself to create depth often generating further questions and, hence, further insight.

Insight

Qualitative
Insight is not exact, can be uncertain and subjective, and can have multiple levels of resolution.

Relevant
Insight is deeply embedded in the data domain, connecting the data to existing domain knowledge and giving it relevant meaning going beyond dry data analysis, to relevant domain impact.

Unexpected
Insight is often unpredictable, serendipitous, and creative.



Visual Analytics

- Combine the power of computer and human

- Compute power
- Storage capacity

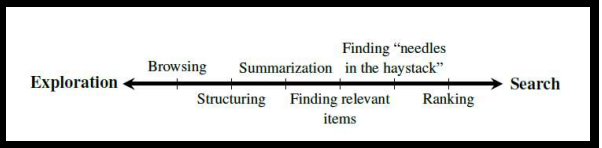
- Flexibility
- Creativity
- Expert knowledge

Ref:Chinchor2010

Definition

Multimedia Analytics
=
Multimedia Analysis
+
Visual Analytics

Multimedia Analytics



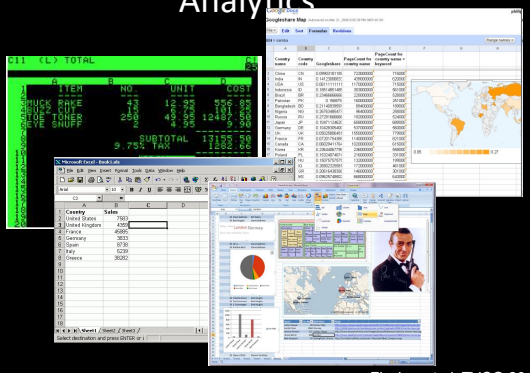
INSIGHT

Analytics

- What is the best known Analytic tool?

Yes the Spreadsheet

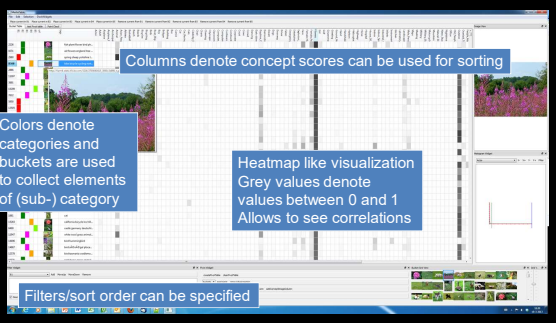
Analytics



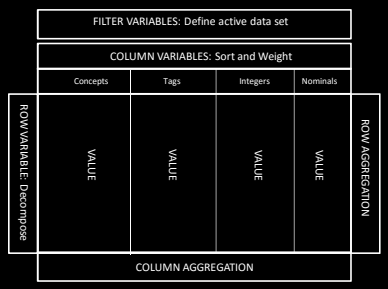
Fischer et.al, TVCG 2010.

Refs: deRooij2010b, deRooij2013

MediaTable

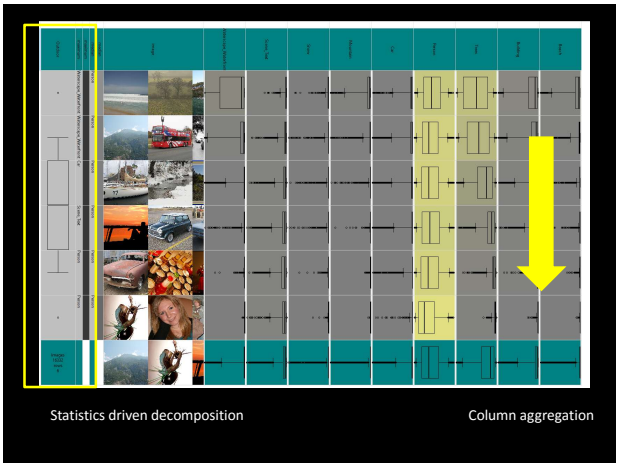


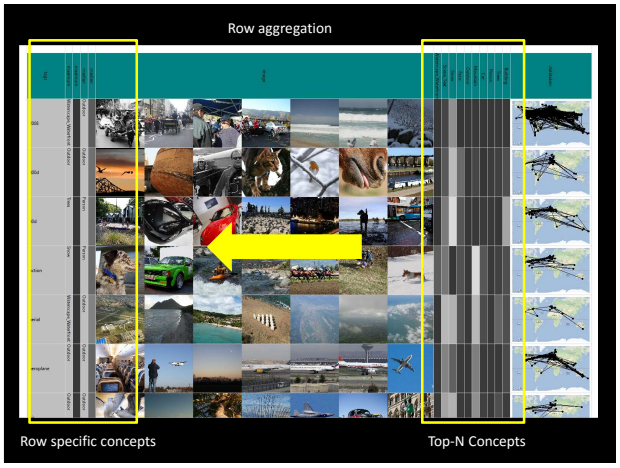
Multimedia Pivot Tables

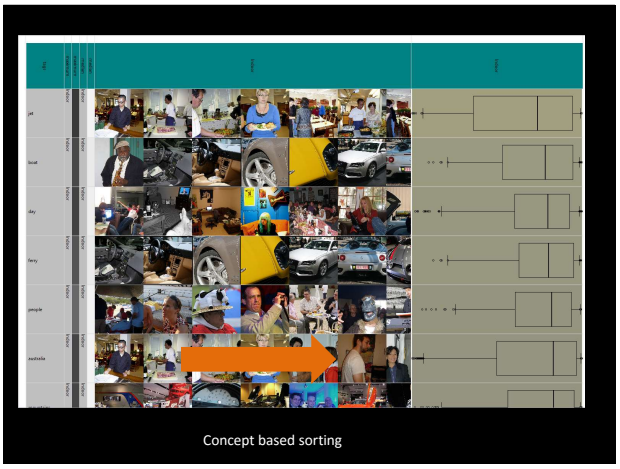


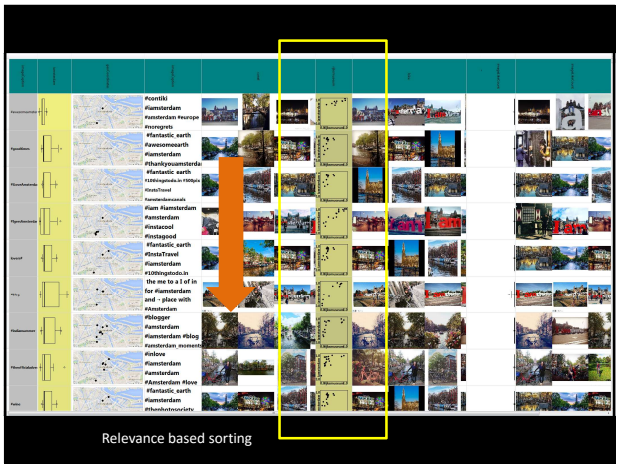
Visualizations

Type	Filter	Column	Row	Value	Visualization
Images	Selection to bucket	x	Individual images	Sorted list of images	
Nominal	Label selection	x	Individual labels	Sorted and weighted text histogram	
Buckets	Bucket selection	x	Individual buckets	Weighted histogram	
Geo	Selection to bucket	x	x	Map with weighted elements	
Numeric	Range selection	Weights	7-point summary	Sum, max, avg, weighted distribution	
Concepts	Range selection	Weights	7-point summary	Weighted distribution	
Tags	Tag selection	Weights	Individual tags	Sorted and weighted tag histogram	

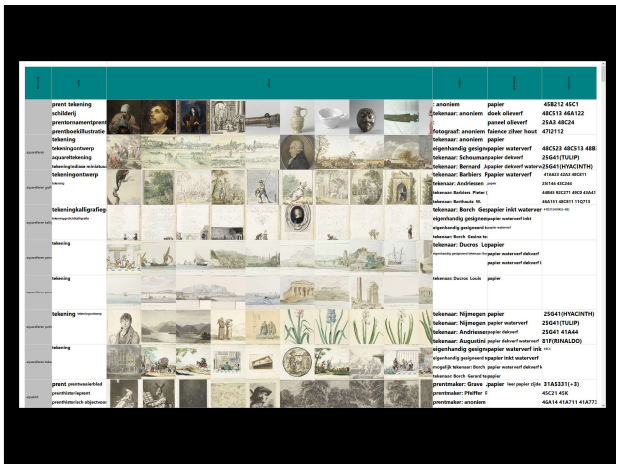


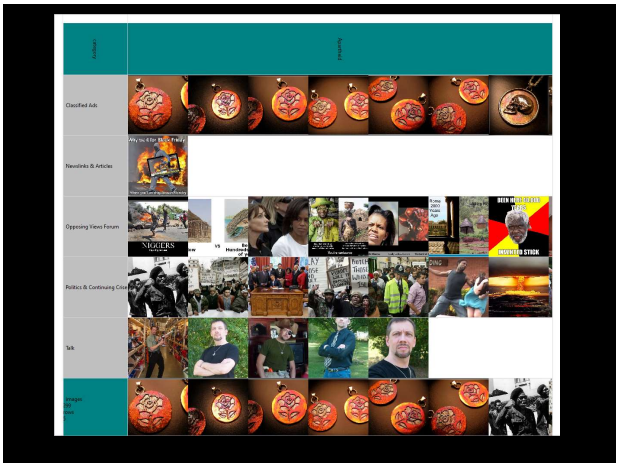






BM-25 BASED RANKING





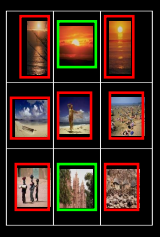
Demo

https://staff.fnwi.uva.nl/m.worring/pivot-tables.html

Learning from interaction

Employing user interaction

Selection of pos/neg examples

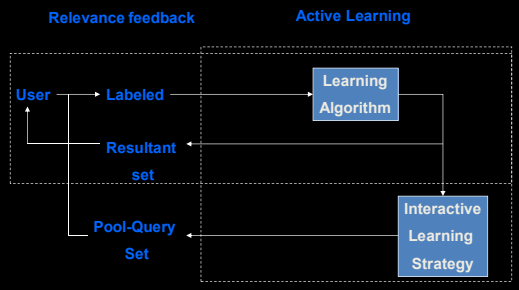


Some elements in the collection are labeled

Many are not

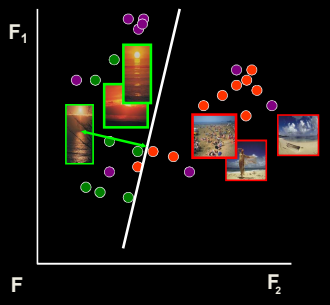
Ref: Huang2008

Employing user interaction



Chen in 2005 was the first to explore this for Video Retrieval

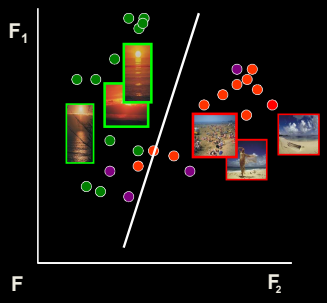
Relevance feedback



Try to find boundary in feature space best separating positive from negative examples

Measure of class membership probability

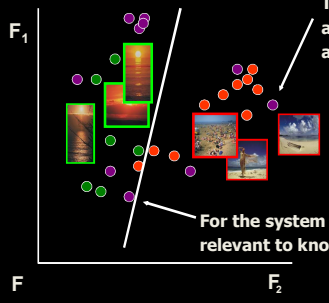
Relevance feedback



In the next iteration I will have more samples hence a better estimate of the boundary

This process is usually known as relevance feedback

Active Learning



The system can safely assume this sample is also negative

In active learning the system decides which elements to show for feedback and which not.

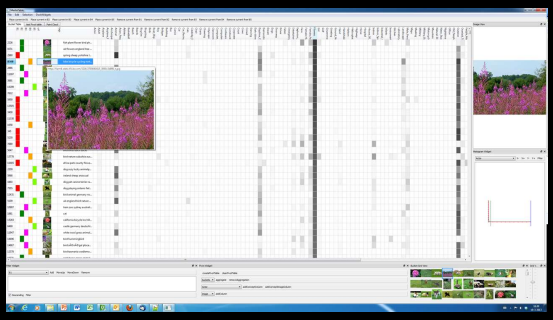
For the system it is relevant to know this label

Automatic AND interactive

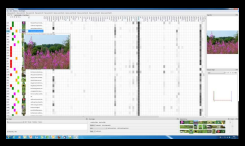
Three interactive strategies

- Fully interactive
  - User is interactively performing the sort/select/categorize process
- Manual relevance feedback
  - In addition to the above the user can perform relevance feedback on any of the categories
- Unobtrusive relevance feedback
  - In addition to the above the system automatically indicates new potentially relevant elements

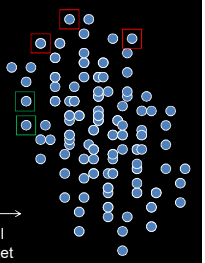
Fully interactive



On demand suggestions



Uncategorized images

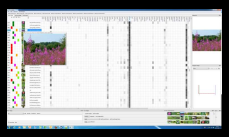


After categorizing some elements

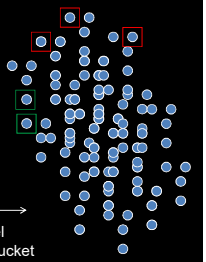
Learn and apply model for **user** selected bucket

Category suggestions

Unobtrusive assistance



Uncategorized images



Continously observe what happens

Learn and apply model for **system** selected bucket

Category suggestions

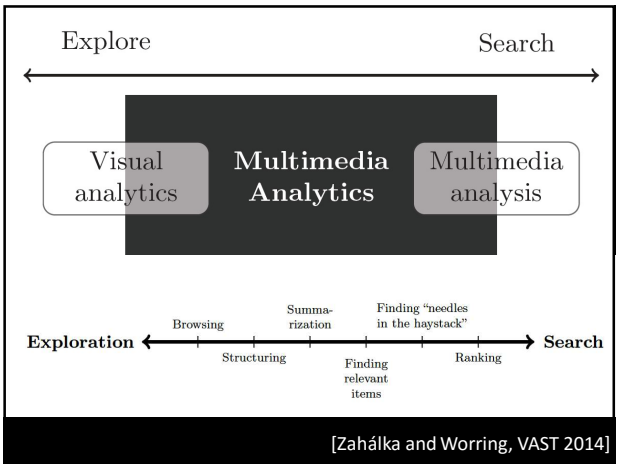
Results: elements found

Task	Baseline	Manual	Automatic
1	13 ± 7	36 ± 19 *	46 ± 8 * °
2	24 ± 14	23 ± 19	32 ± 8 * °
3	27 ± 24	31 ± 18	53 ± 18 * °
4	13 ± 8	19 ± 7 *	28 ± 5 * °
Average	20 ± 16	27 ± 18 *	40 ± 15 * °
Participants	8	7	6

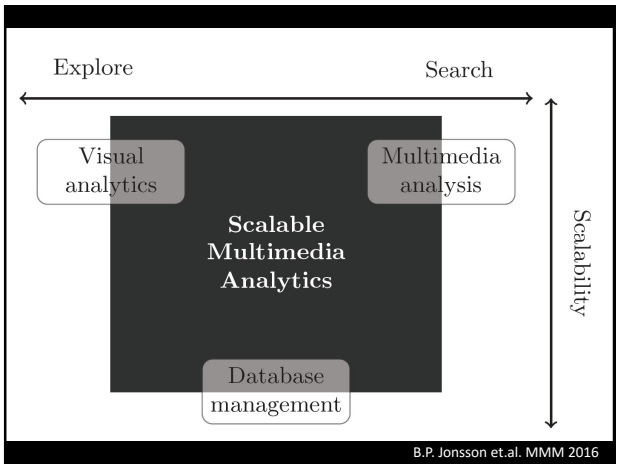
• significant at the p=0.01 level compared to baseline
o significant at the p=0.01 level compared to manual

Task 1: specific, high visual similarity
Task 2: generic visually diverse, concept available
Task 3: generic visually diverse, concept available
Task 4: generic visually diverse, no concept available

SCALABILITY



[Zahálka and Worring, VAST 2014]



B.P. Jonsson et.al. MMM 2016

WRAP-UP

Objective and Subjective data

Open Data

Structured data

Numeric data

Temporal data

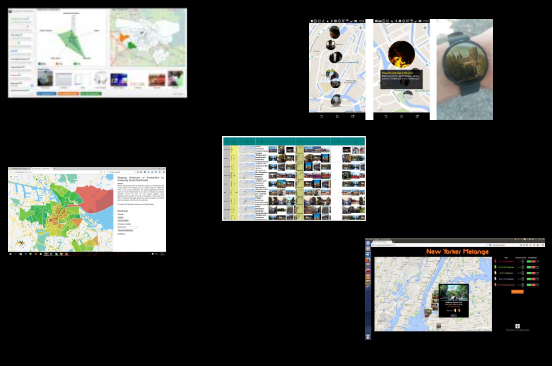
Geographic data

Image data

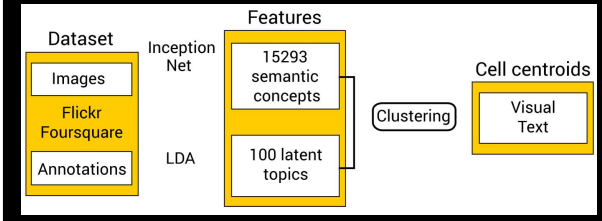
Textual data

Unstructured data

The applications

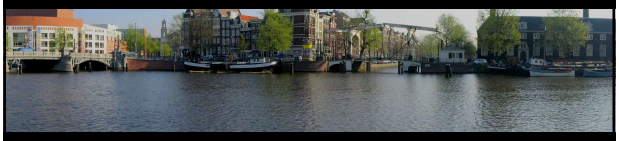


The Algorithms



And its variations

www.amsterdamdatascience.nl



m.worring@uva.nl