MAKING SENSE OF THE INFORMATION IN THE REAL TIME WEB

Jurgen Van Gael (@jvangael)
Microsoft Research
Future of Social Networks - 2010

Microsoft®
FUSELABS

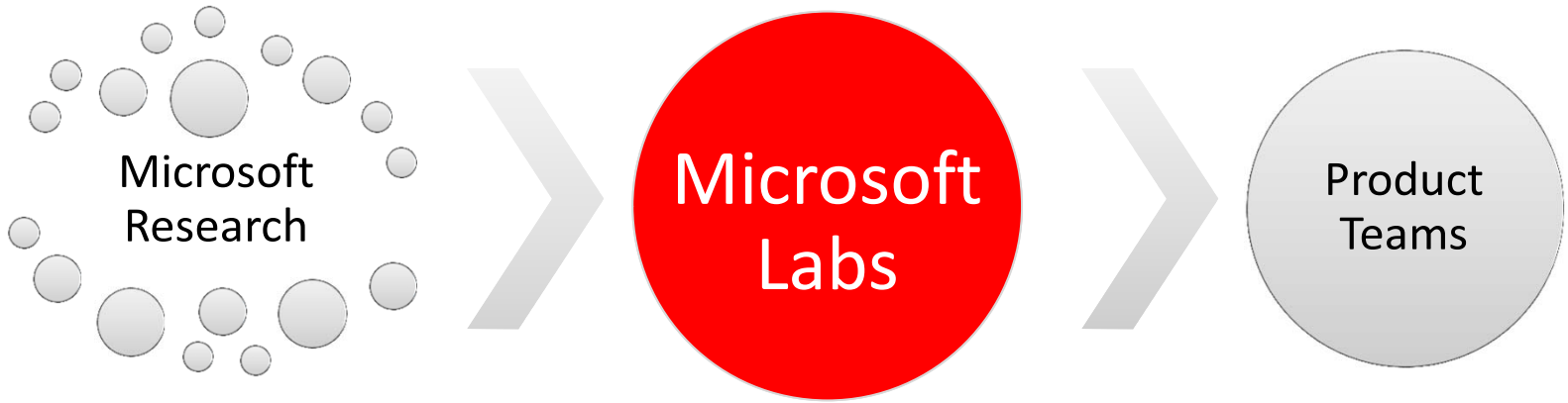
Overview

- Future Social Experiences Labs
- The Real-Time Web
- Making Sense of Information
- Project Emporia

Why are we here and who are we?

FUSE LABS

Labs at Microsoft



Pushing the limits of technology

Proving or dis-proving concepts

Execution

XBOX LIVE

bing

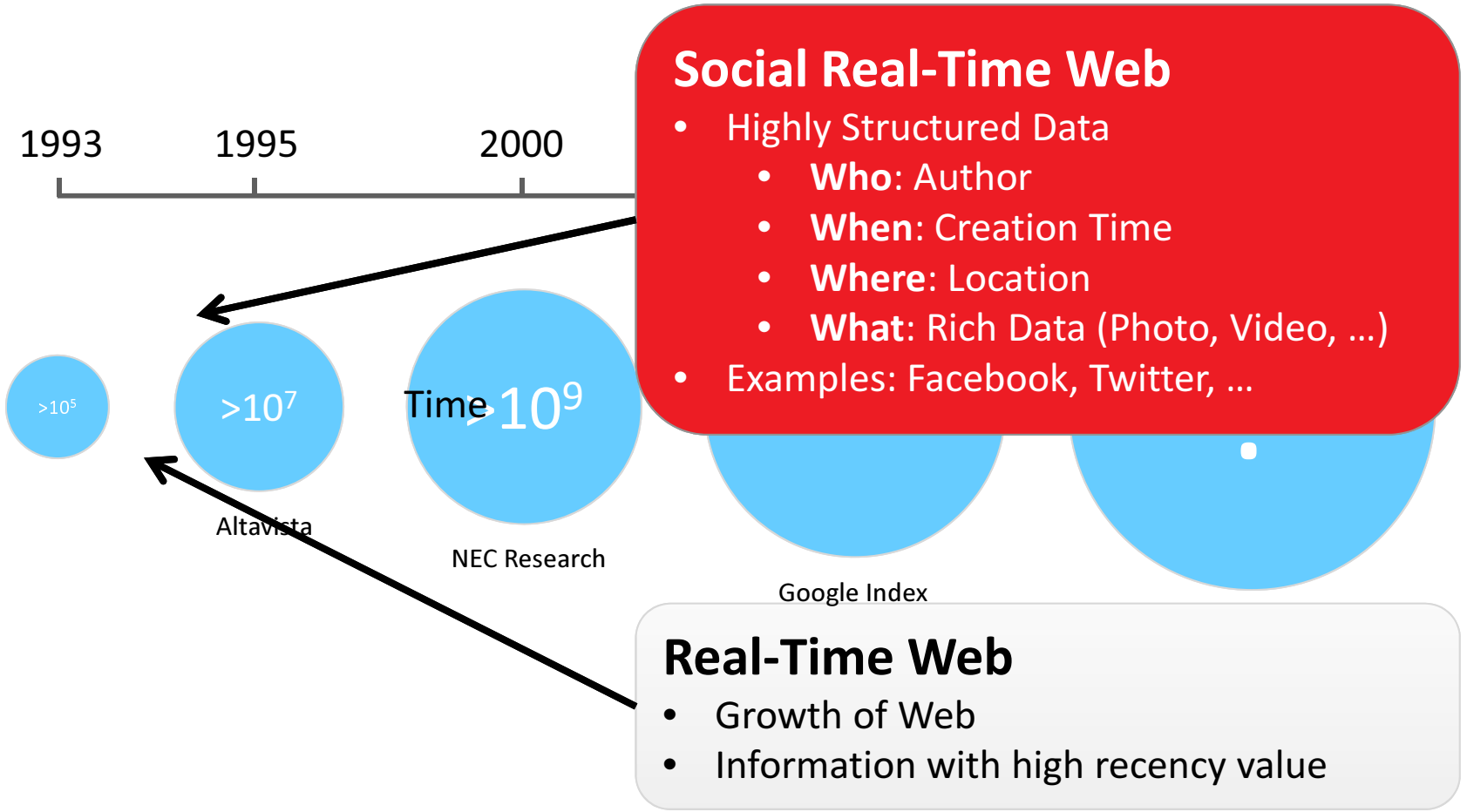
Docs BETA

spindex

What is the real-time web and what is the problem?

THE REAL-TIME WEB

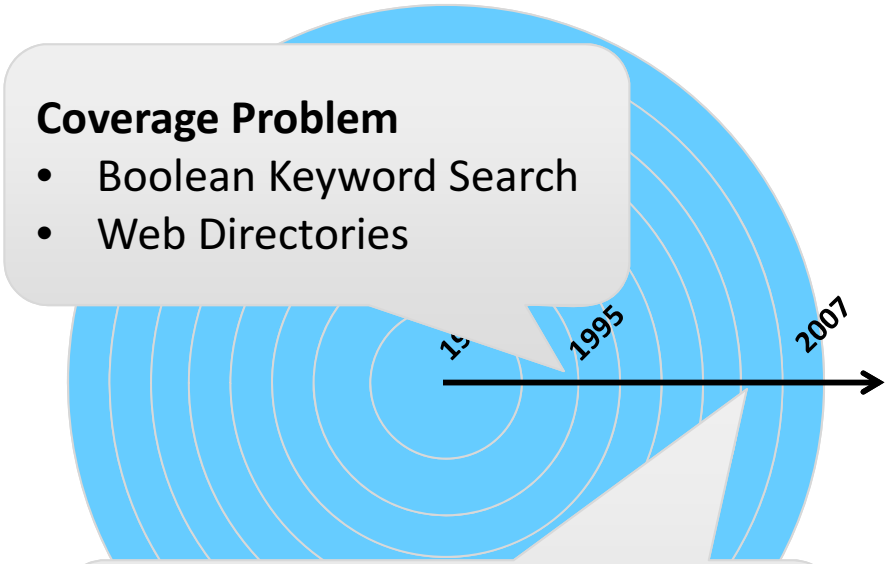
Information on the Web



From Coverage to Relevance

World Wide Web

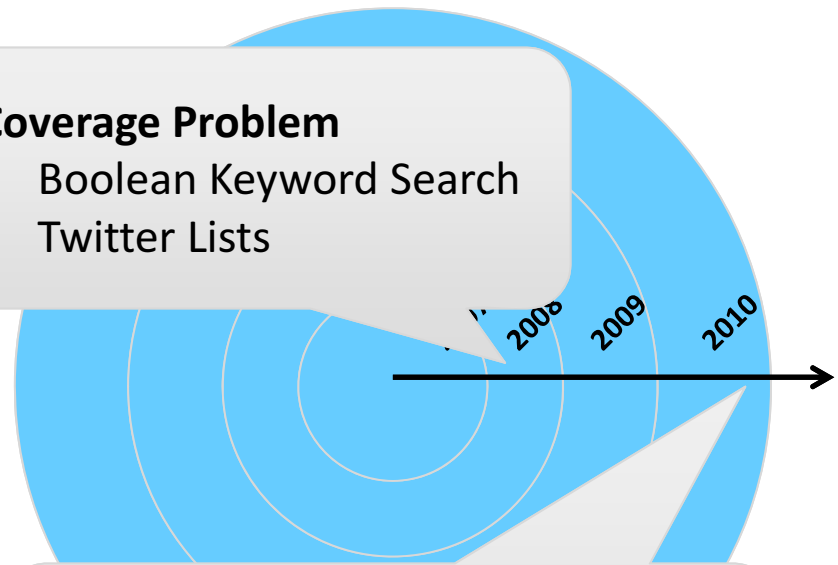
Coverage Problem
- Boolean Keyword Search
- Web Directories



Relevance Problem
- Relevance filtering & ranking
- Web Search

Twitter

Coverage Problem
- Boolean Keyword Search
- Twitter Lists



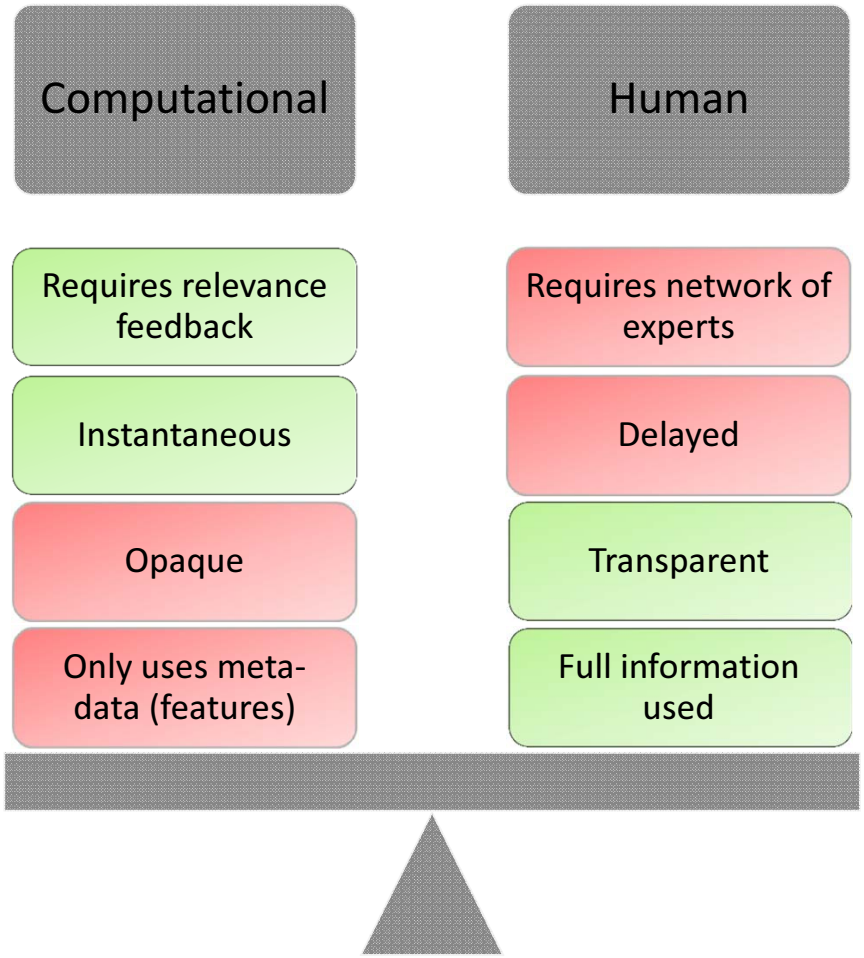
Relevance Problem
- Need for relevance filtering
- ???

DEMO

Twitter – Social Network or Microblog?

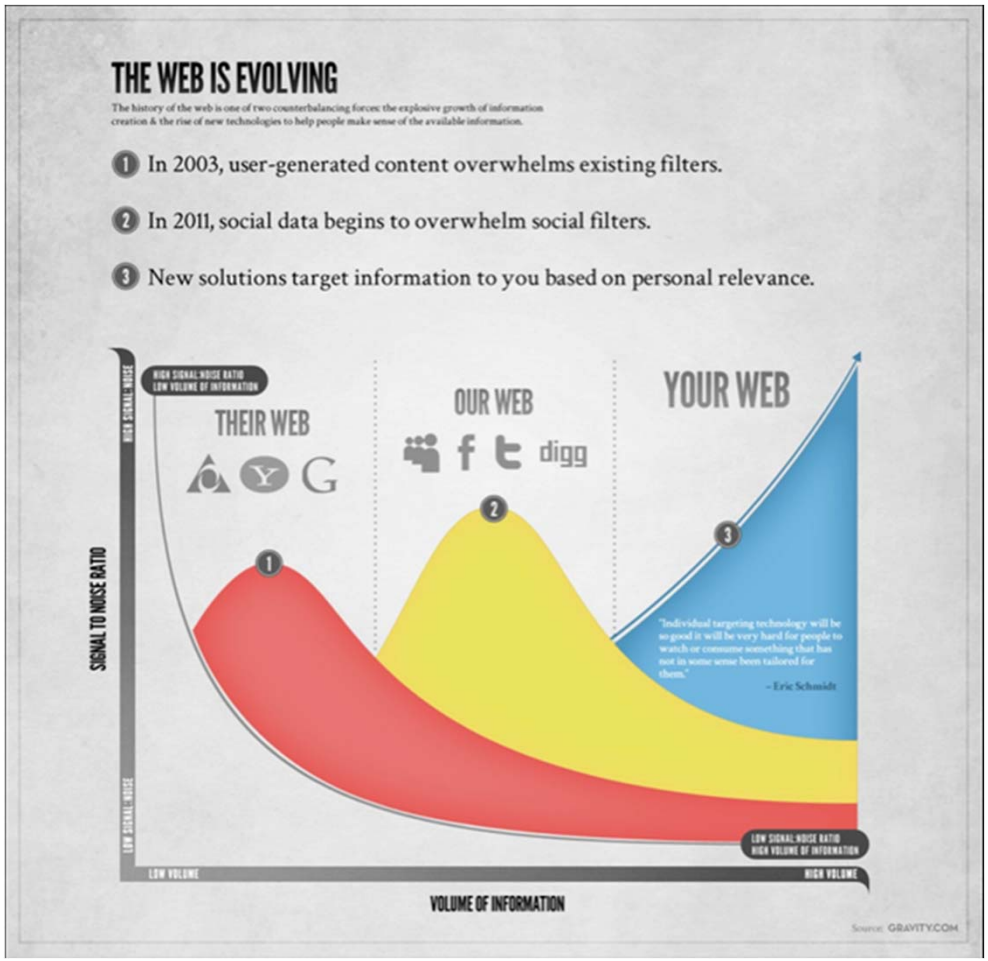
Social Network	Content Consumer	Content Growth
Asymmetric people relationship	> 190,000,000 monthly visitors	Twitter content doubles every 6 months
Public is default	80% of Twitter users posting less than 10 tweets overall	Existing users become more active
No rich media	40% of Twitter users have zero tweets!	Authoring user base is stagnant

Computational vs. Human Intelligence



The Future Will Be Personalized

TechCrunch - 17/11/2010 - Amit Kapur

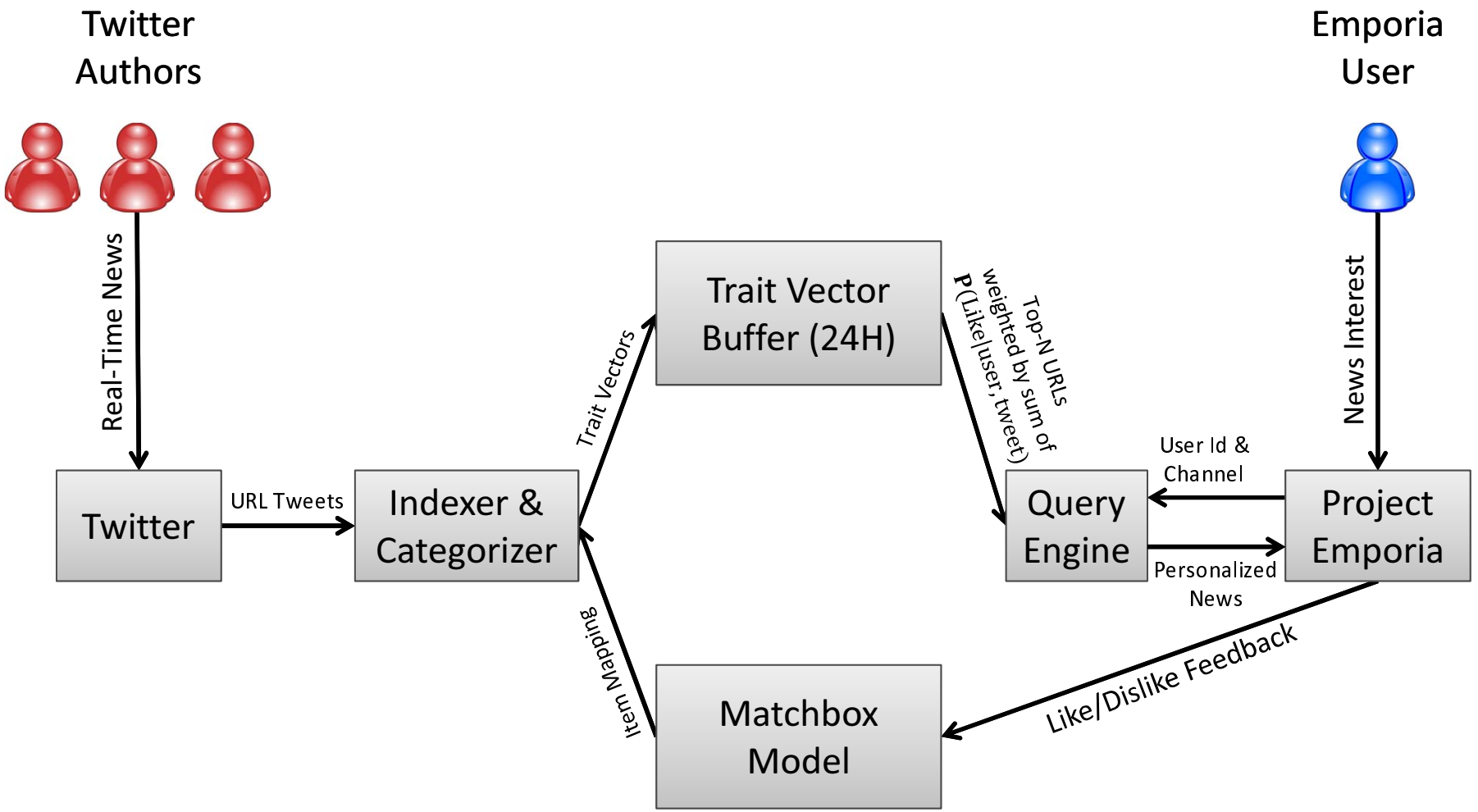


Open up the world of the social real time web.

PROJECT EMPORIA

DEMO

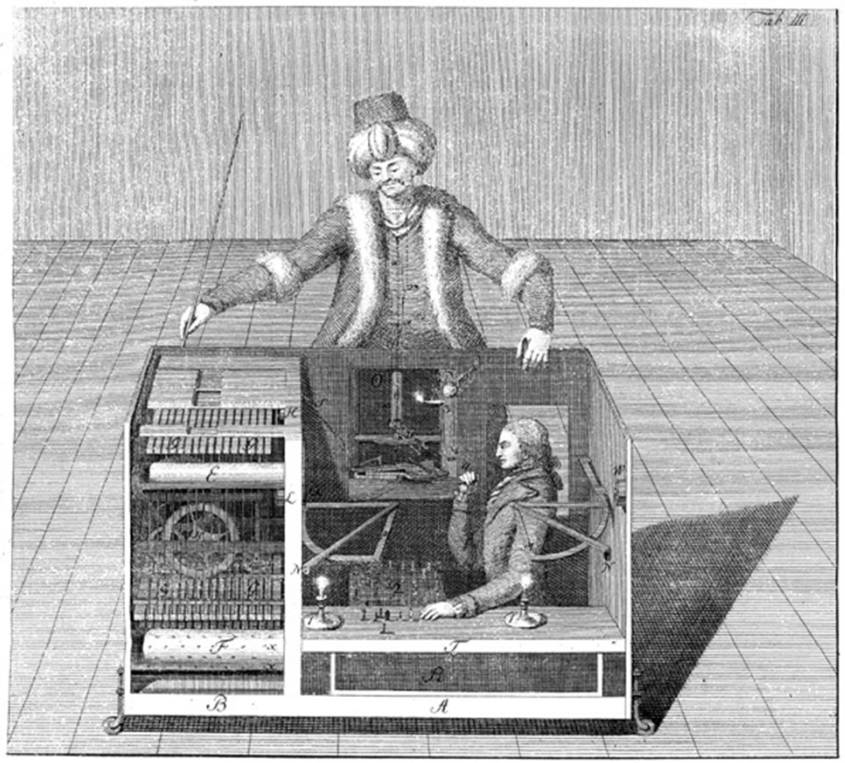
Project Emporia Pipeline



Leveraging the Wisdom of the Crowd

CLASSIFIER TRAINING

What is Mechanical Turk

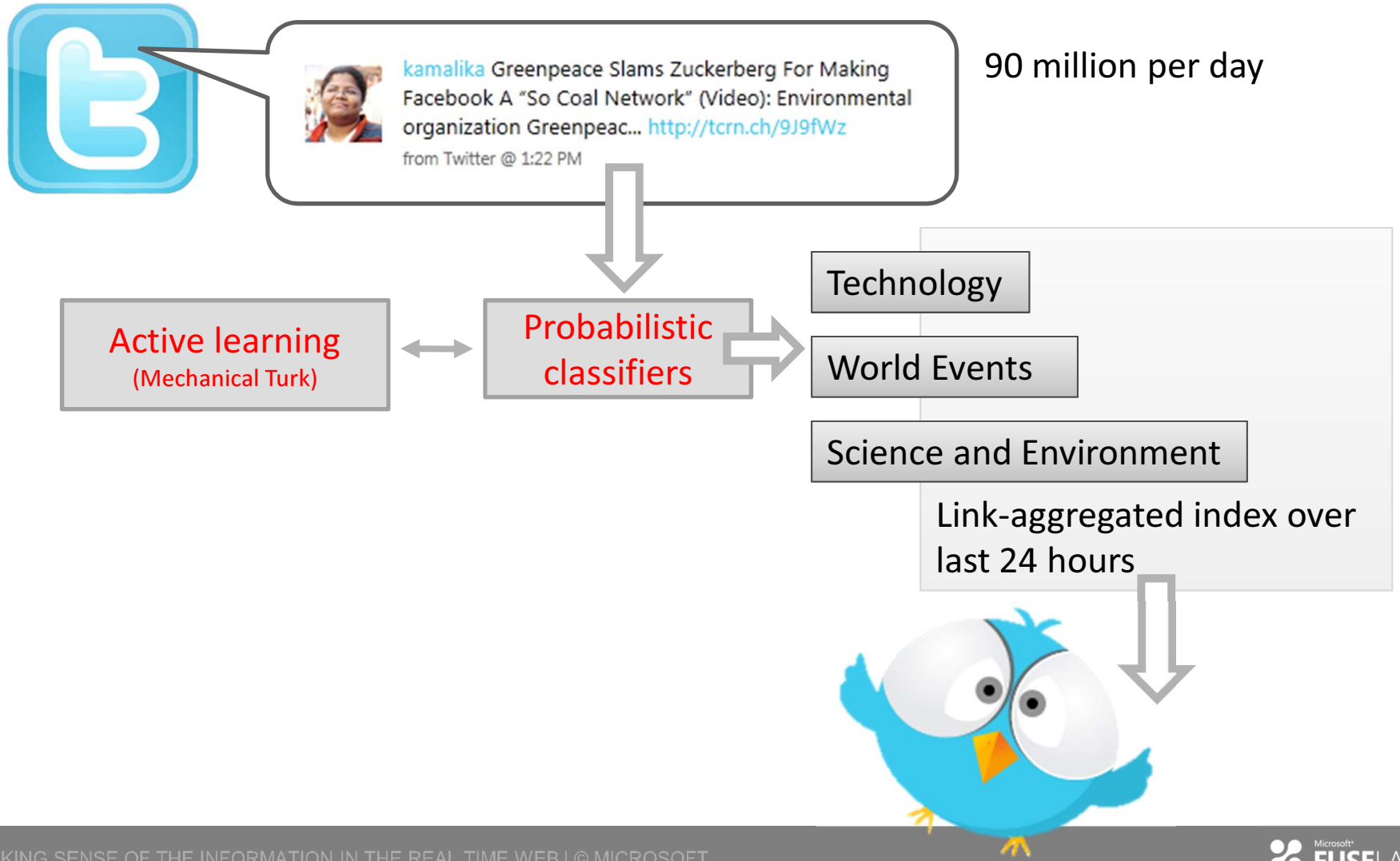


- Human Intelligence Task
- Pay per job (+bonus)

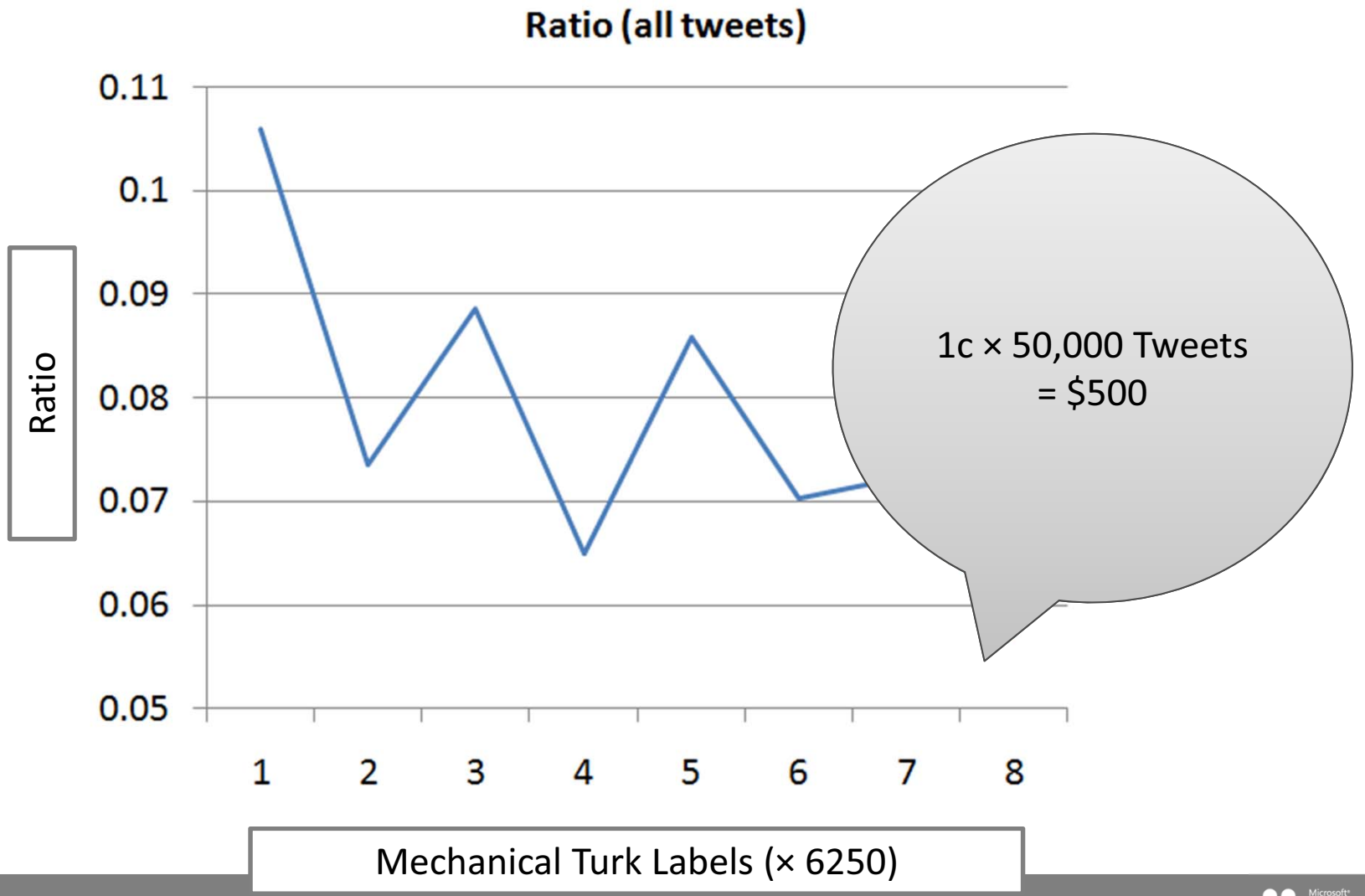
✓ Fast turnaround
✓ Large Marketplace

○ Noisy Results
○ Fraudsters

In Project Emporia



Learning Performance



Crowdsourcing Relevance

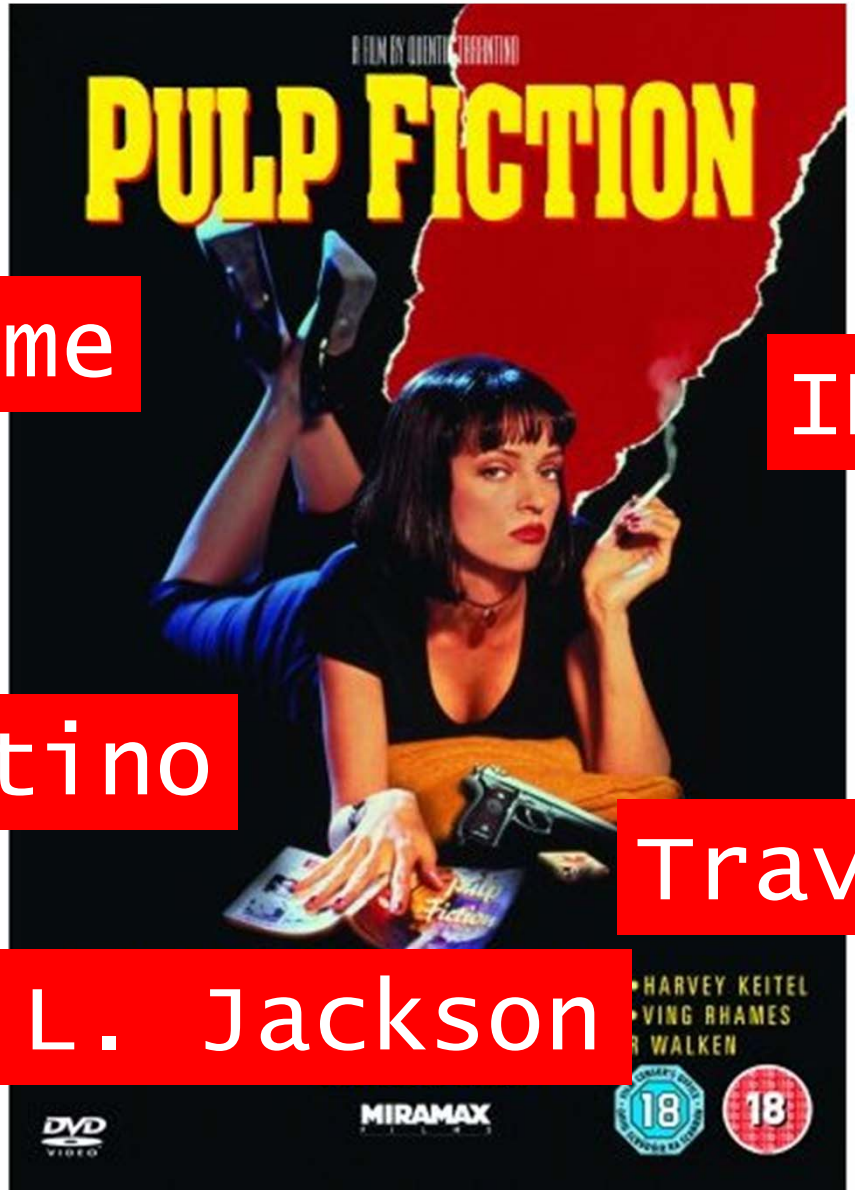
MATCHBOX RECOMMENDATIONS

ID=22549
Age<25
Student
Female

Drama

Crime

ID=2345

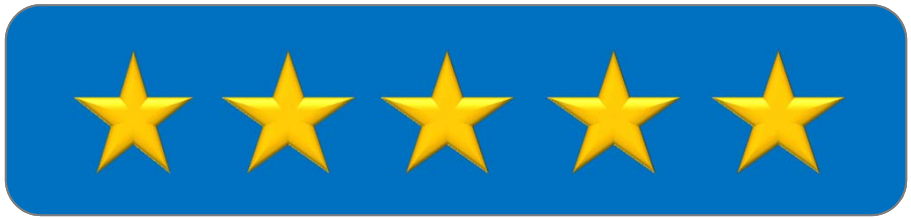


Tarantino

Travolta

Samuel L. Jackson

Training...



ID=22549

Age<25

Student

Female

Julia Roberts

Hugh Grant

ID=5432

Comedy

Romance

Prediction

?

ID=22549

Age<25

Student

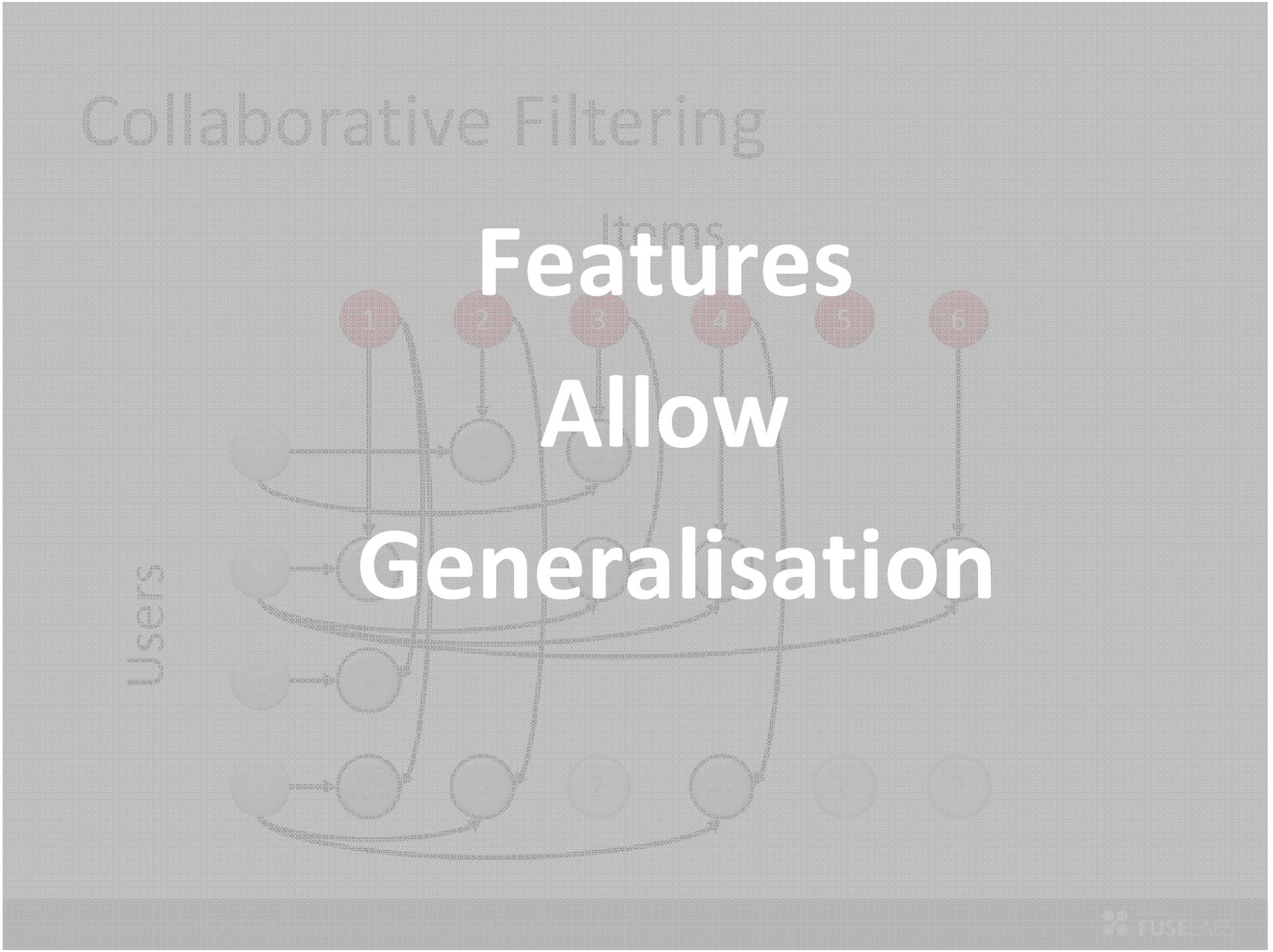
Female

ID=9765

Eva Green

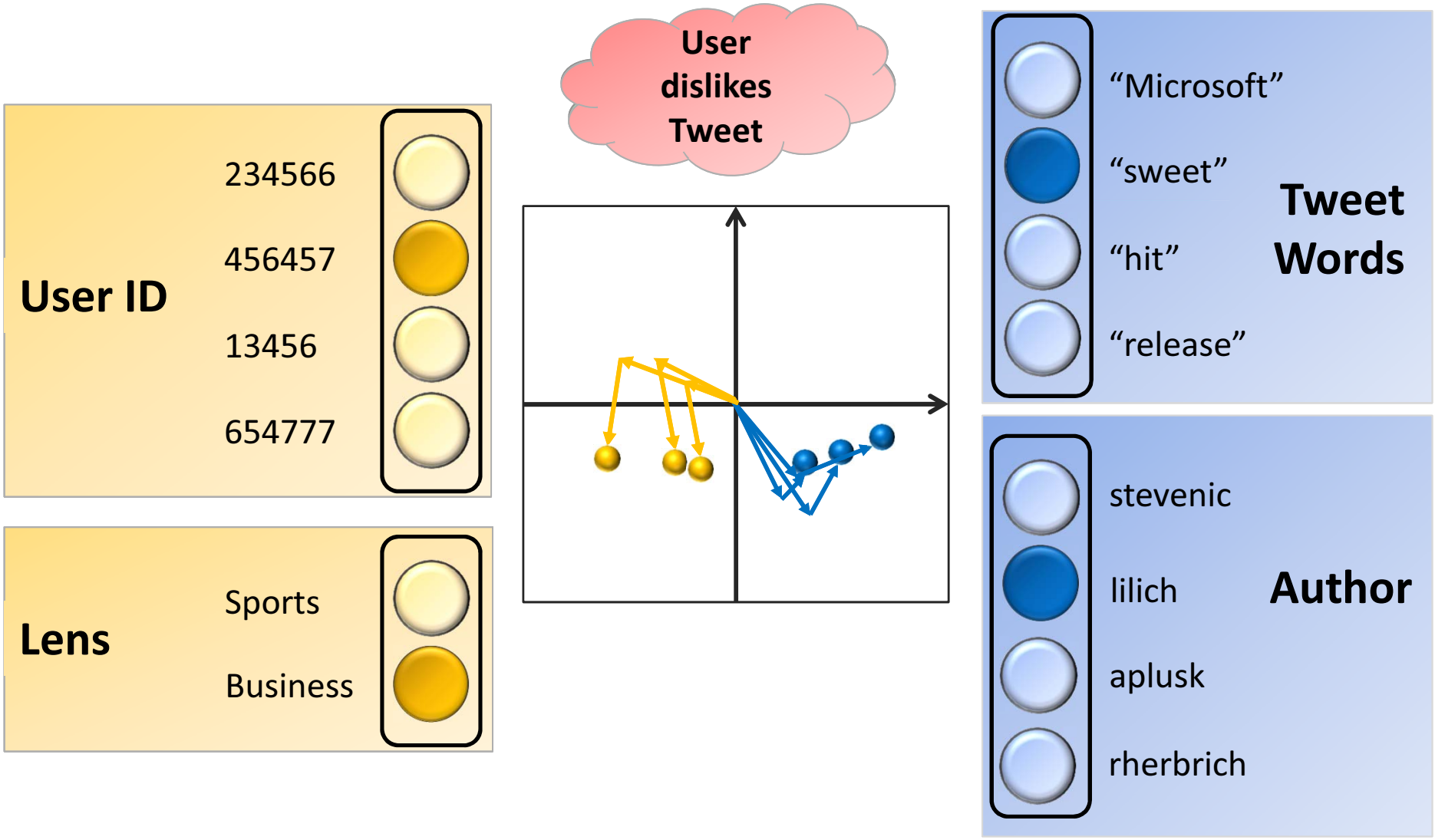
Daniel Craig

Action

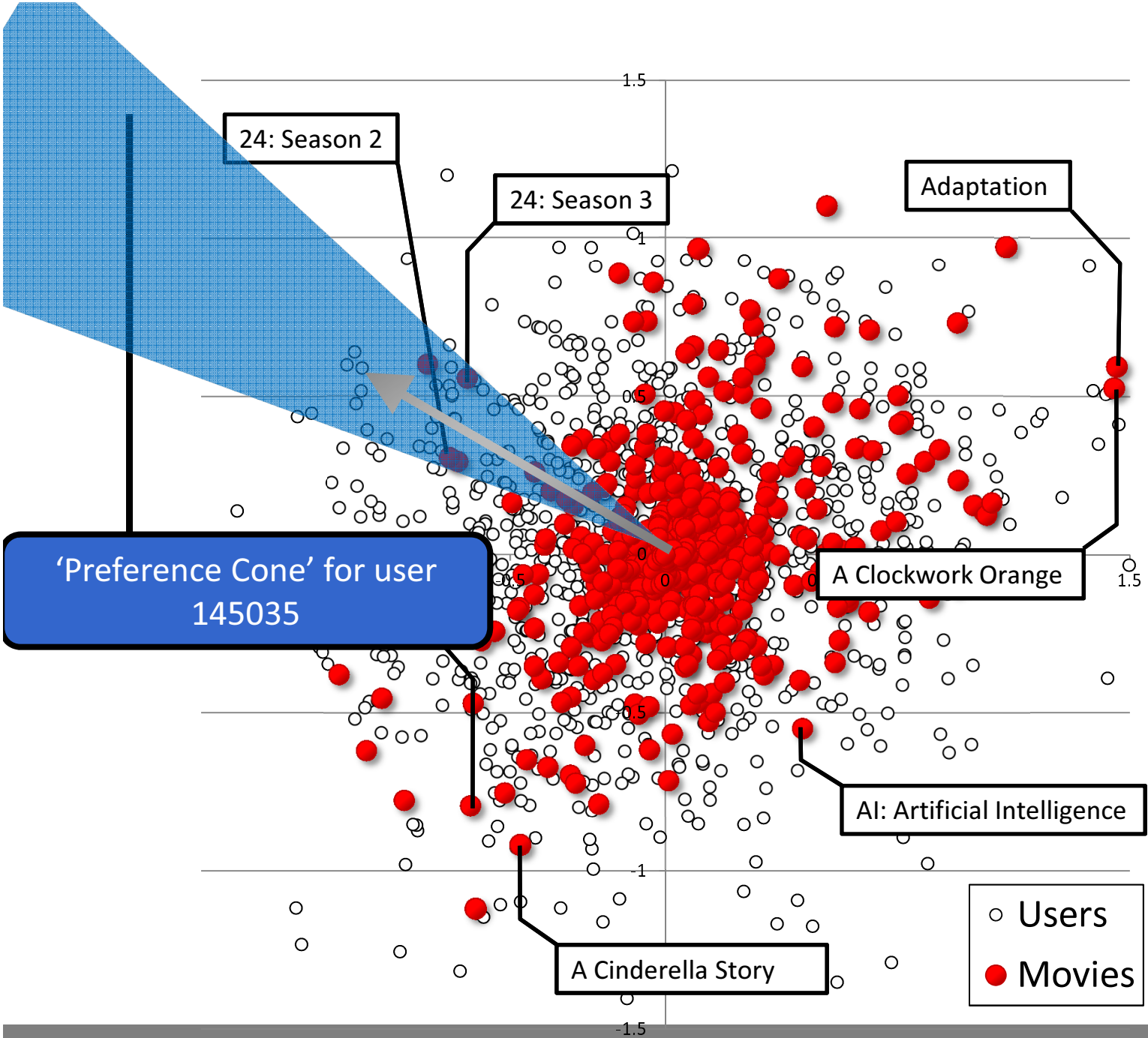


Features
Allow
Generalisation

Recommender System: MatchBox



User Item Trait Space



SUMMARY

Social Real-Time Web

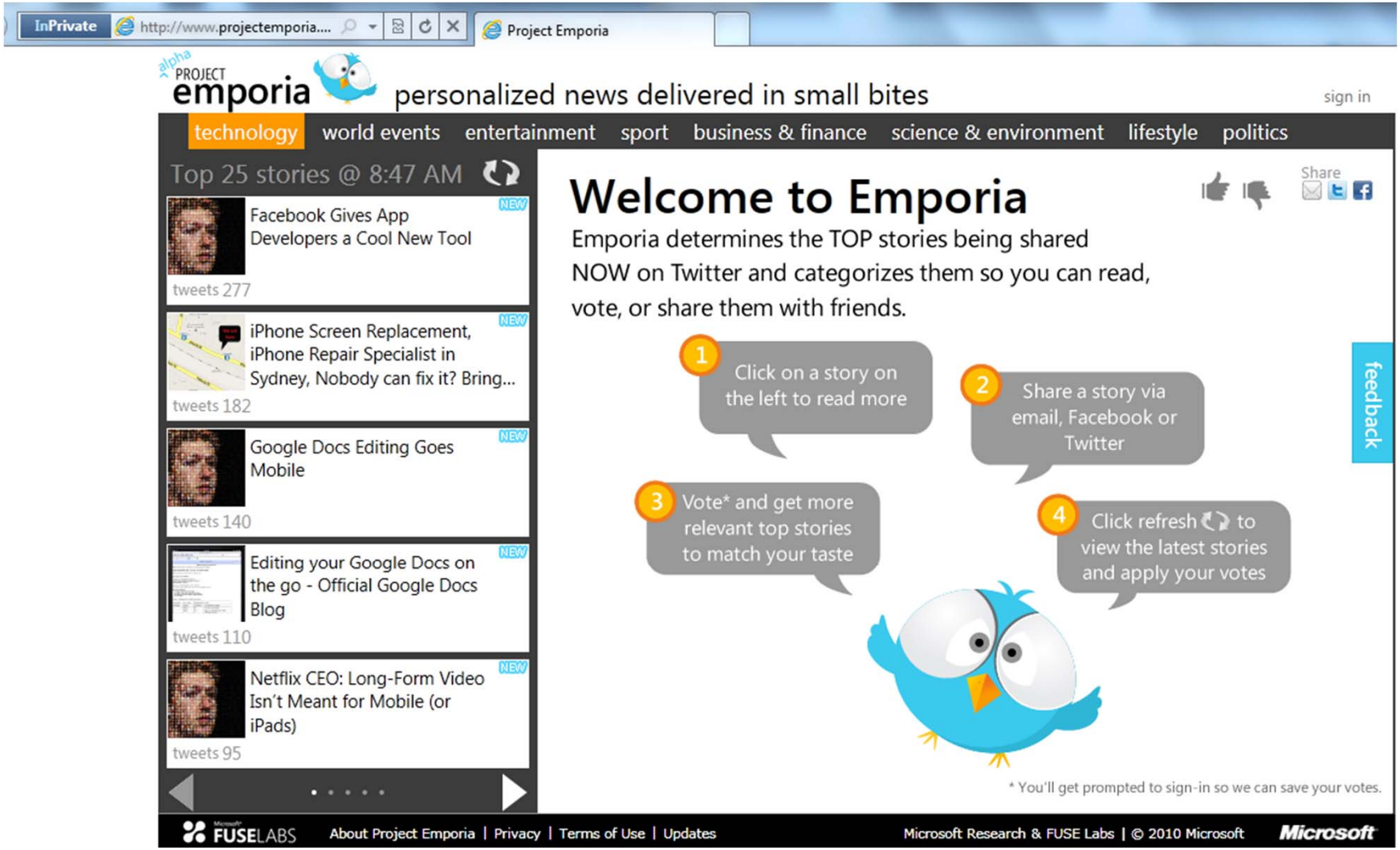
- Pivotal moment – relevance is the next big challenge
- Combining human and computational intelligence is key

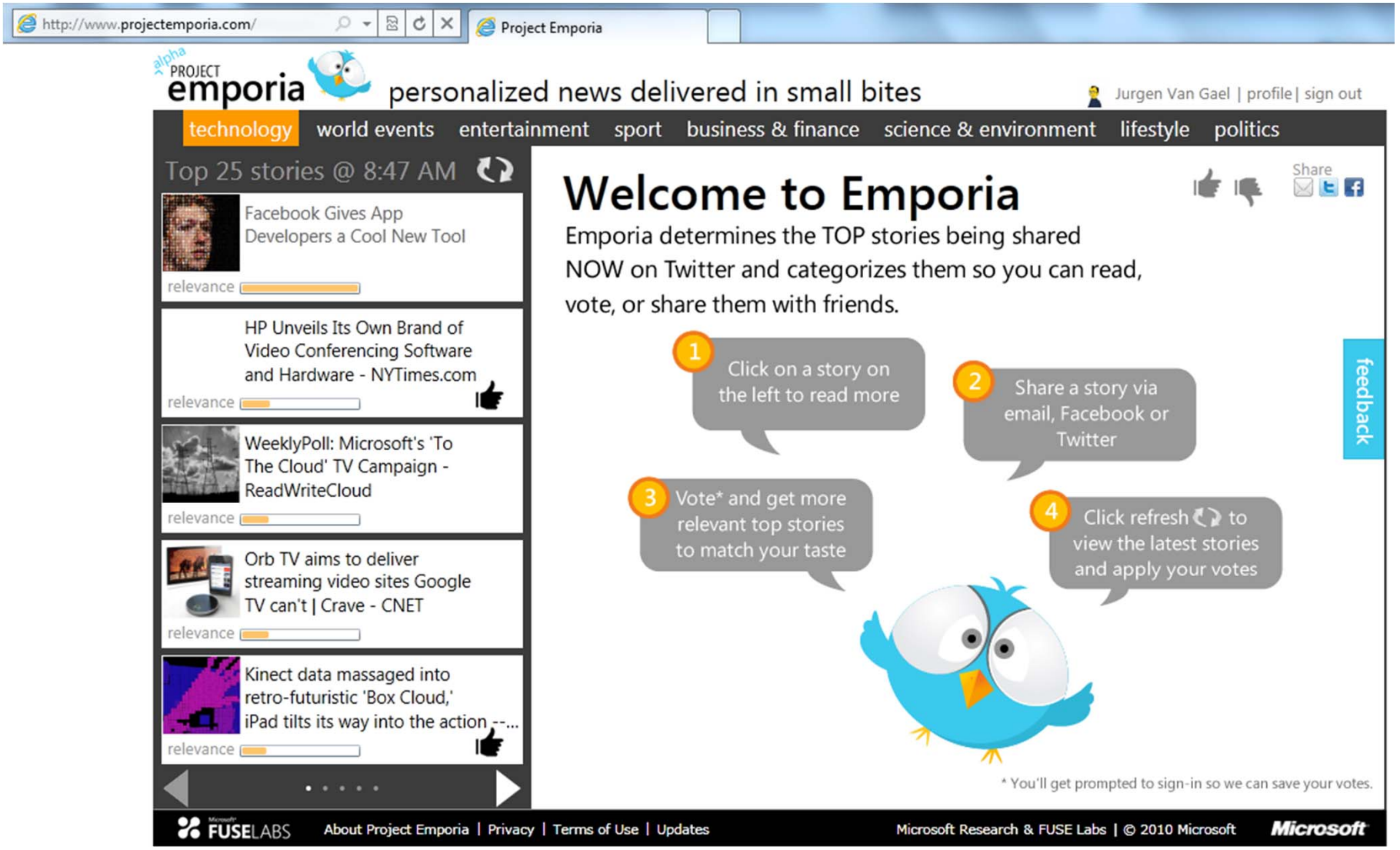
Project Emporia

- Alpha Release from Microsoft Research FUSE Labs
- Leverages the power of social networks:
  - Content Creation (Twitter, RSS)
  - Content Filtering (Mechanical Turk)
  - Content Recommendation (Emporia Users)

Go play at **www.projectemporia.com**

Microsoft


InPrivate
http://www.projectemporia....
Project Emporia
alpha
PROJECT
emporia
personalized news delivered in small bites
sign in
technology
world events
entertainment
sport
business & finance
science & environment
lifestyle
politics
Top 25 stories @ 8:47 AM
Facebook Gives App Developers a Cool New Tool
tweets 277
iPhone Screen Replacement, iPhone Repair Specialist in Sydney, Nobody can fix it? Bring...
tweets 182
Google Docs Editing Goes Mobile
tweets 140
Editing your Google Docs on the go - Official Google Docs Blog
tweets 110
Netflix CEO: Long-Form Video Isn't Meant for Mobile (or iPads)
tweets 95
NEW
Welcome to Emporia
Share
Emporia determines the TOP stories being shared NOW on Twitter and categorizes them so you can read, vote, or share them with friends.
1 Click on a story on the left to read more
2 Share a story via email, Facebook or Twitter
3 Vote* and get more relevant top stories to match your taste
4 Click refresh to view the latest stories and apply your votes
feedback
* You'll get prompted to sign-in so we can save your votes.
FUSELABS
About Project Emporia | Privacy | Terms of Use | Updates
Microsoft Research & FUSE Labs | © 2010 Microsoft
Microsoft


http://www.projectemporia.com/
Project Emporia
alpha
PROJECT
emporia
personalized news delivered in small bites
Jurgen Van Gael | profile | sign out
technology
world events
entertainment
sport
business & finance
science & environment
lifestyle
politics
Top 25 stories @ 8:47 AM
Facebook Gives App Developers a Cool New Tool
relevance
HP Unveils Its Own Brand of Video Conferencing Software and Hardware - NYTimes.com
relevance
WeeklyPoll: Microsoft's 'To The Cloud' TV Campaign - ReadWriteCloud
relevance
Orb TV aims to deliver streaming video sites Google TV can't | Crave - CNET
relevance
Kinect data massaged into retro-futuristic 'Box Cloud,' iPad tilts its way into the action --...
relevance
Welcome to Emporia
Share
Emporia determines the TOP stories being shared NOW on Twitter and categorizes them so you can read, vote, or share them with friends.
1 Click on a story on the left to read more
2 Share a story via email, Facebook or Twitter
3 Vote* and get more relevant top stories to match your taste
4 Click refresh to view the latest stories and apply your votes
feedback
* You'll get prompted to sign-in so we can save your votes.
FUSELABS
About Project Emporia | Privacy | Terms of Use | Updates
Microsoft Research & FUSE Labs | © 2010 Microsoft
Microsoft