WWW2015 FLORENCE, ITALY
24TH INTERNATIONAL WORLD WIDE WEB CONFERENCE
Florence 18 - 22 May 2015

An Introduction to Entity Recommendation and Understanding

Hao Ma and Yan Ke
Microsoft

The contents and opinions described in this tutorial do not necessarily reflect the opinions of Microsoft.

Technologies mentioned might or might not be in actual use.

Motivation

Database?

NLP?

Application?

Goals of this Tutorial

- Help identify many interesting applications in the field of Entity Recommendation and Understanding
- Present the current state of research on related topics
- Pinpoint challenging research problems

Outline

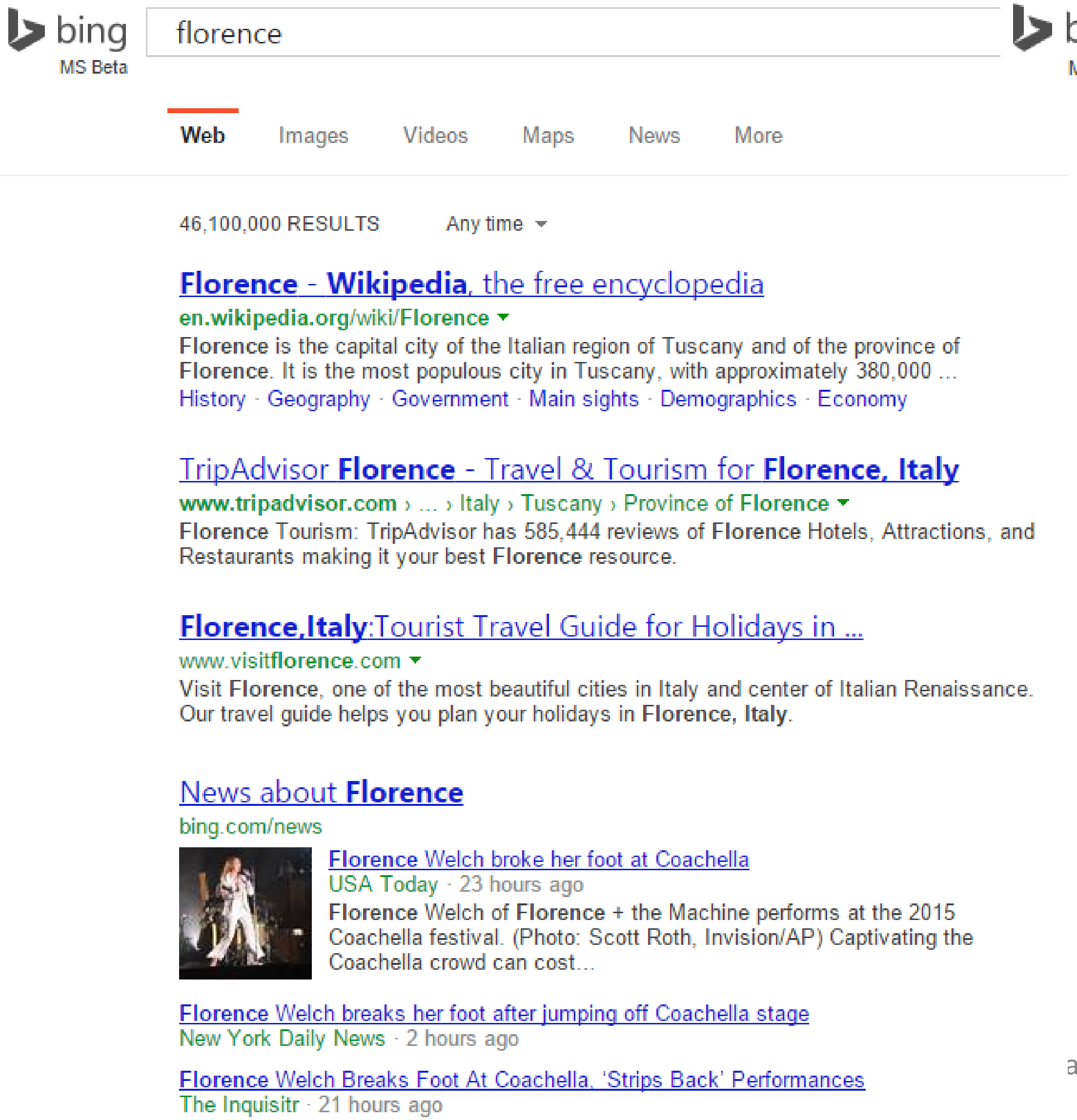
- Introduction to Entity and Knowledge
- Demonstration of Microsoft's Entity Experience
- Entity Recommendation and Understanding
  - $P(entity|entity)$
  - $P(entity|user)$
  - $P(entity|query)$
- Summary

Outline

- Introduction to Entity and Knowledge
- Demonstration of Microsoft’s Entity Experience
- Entity Recommendation and Understanding
  - $P(entity|entity)$
  - $P(entity|user)$
  - $P(entity|query)$
- Summary

Introduction to Entity and Knowledge

Why "Entities"



bing
MS Beta
florence

Web Images Videos Maps News More

46,100,000 RESULTS Any time

Florence - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Florence
Florence is the capital city of the Italian region of Tuscany and of the province of Florence. It is the most populous city in Tuscany, with approximately 380,000 …
History · Geography · Government · Main sights · Demographics · Economy

TripAdvisor Florence - Travel & Tourism for Florence, Italy
www.tripadvisor.com › … › Italy › Tuscany › Province of Florence
Florence Tourism: TripAdvisor has 585,444 reviews of Florence Hotels, Attractions, and Restaurants making it your best Florence resource.

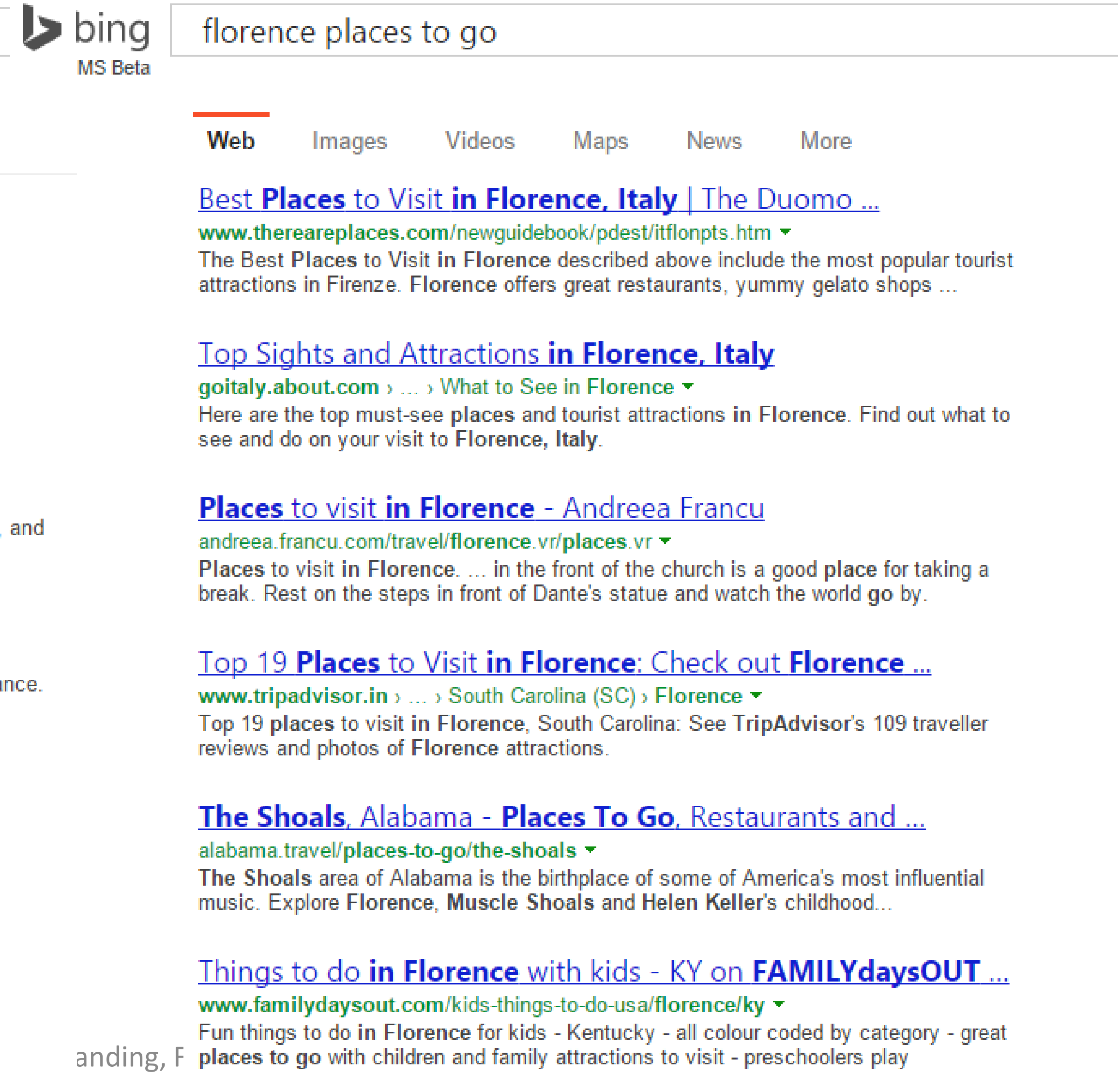
Florence,Italy:Tourist Travel Guide for Holidays in …
www.visitflorence.com
Visit Florence, one of the most beautiful cities in Italy and center of Italian Renaissance. Our travel guide helps you plan your holidays in Florence, Italy.

News about Florence
bing.com/news

Florence Welch broke her foot at Coachella
USA Today · 23 hours ago
Florence Welch of Florence + the Machine performs at the 2015 Coachella festival. (Photo: Scott Roth, Invision/AP) Captivating the Coachella crowd can cost…

Florence Welch breaks her foot after jumping off Coachella stage
New York Daily News · 2 hours ago

Florence Welch Breaks Foot At Coachella, 'Strips Back' Performances
The Inquisitr · 21 hours ago



bing
MS Beta
florence places to go

Web Images Videos Maps News More

Best Places to Visit in Florence, Italy | The Duomo …
www.thereareplaces.com/newguidebook/pdest/itflonpts.htm
The Best Places to Visit in Florence described above include the most popular tourist attractions in Firenze. Florence offers great restaurants, yummy gelato shops …

Top Sights and Attractions in Florence, Italy
goitaly.about.com › … › What to See in Florence
Here are the top must-see places and tourist attractions in Florence. Find out what to see and do on your visit to Florence, Italy.

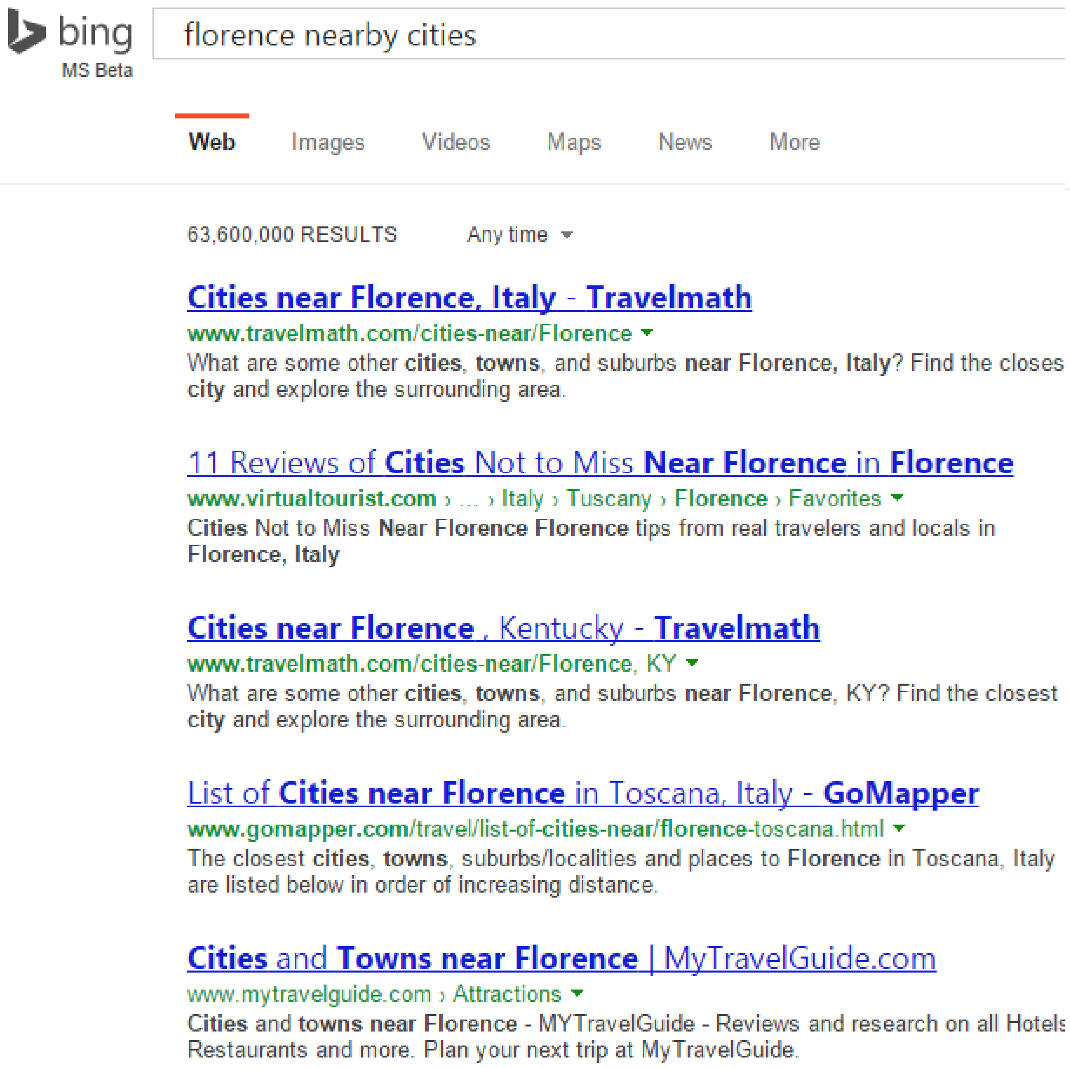
Places to visit in Florence - Andreea Francu
andreea.francu.com/travel/florence.vr/places.vr
Places to visit in Florence. … in the front of the church is a good place for taking a break. Rest on the steps in front of Dante's statue and watch the world go by.

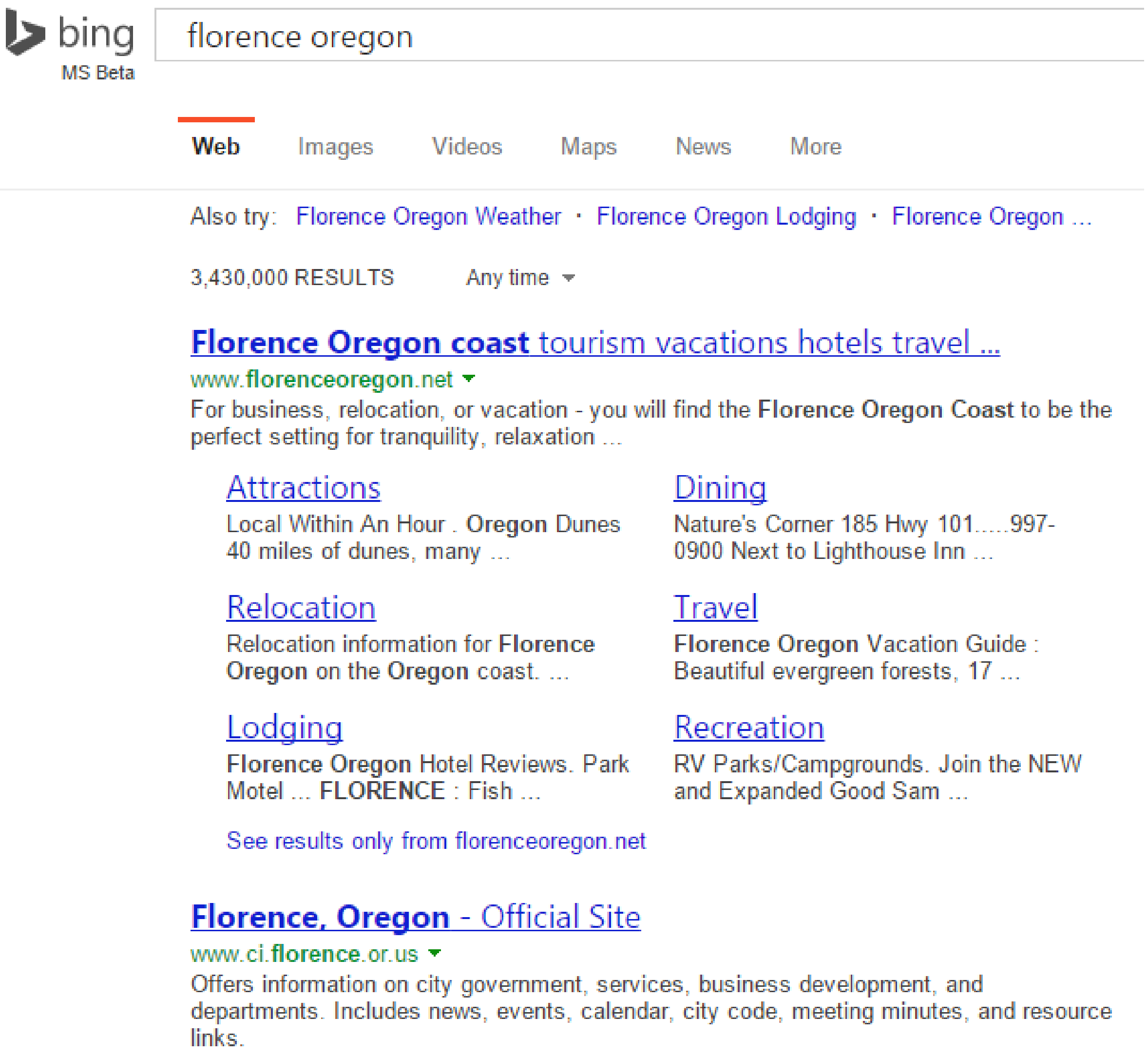
Top 19 Places to Visit in Florence: Check out Florence …
www.tripadvisor.in › … › South Carolina (SC) › Florence
Top 19 places to visit in Florence, South Carolina: See TripAdvisor's 109 traveller reviews and photos of Florence attractions.

The Shoals, Alabama - Places To Go, Restaurants and …
alabama.travel/places-to-go/the-shoals
The Shoals area of Alabama is the birthplace of some of America's most influential music. Explore Florence, Muscle Shoals and Helen Keller's childhood…

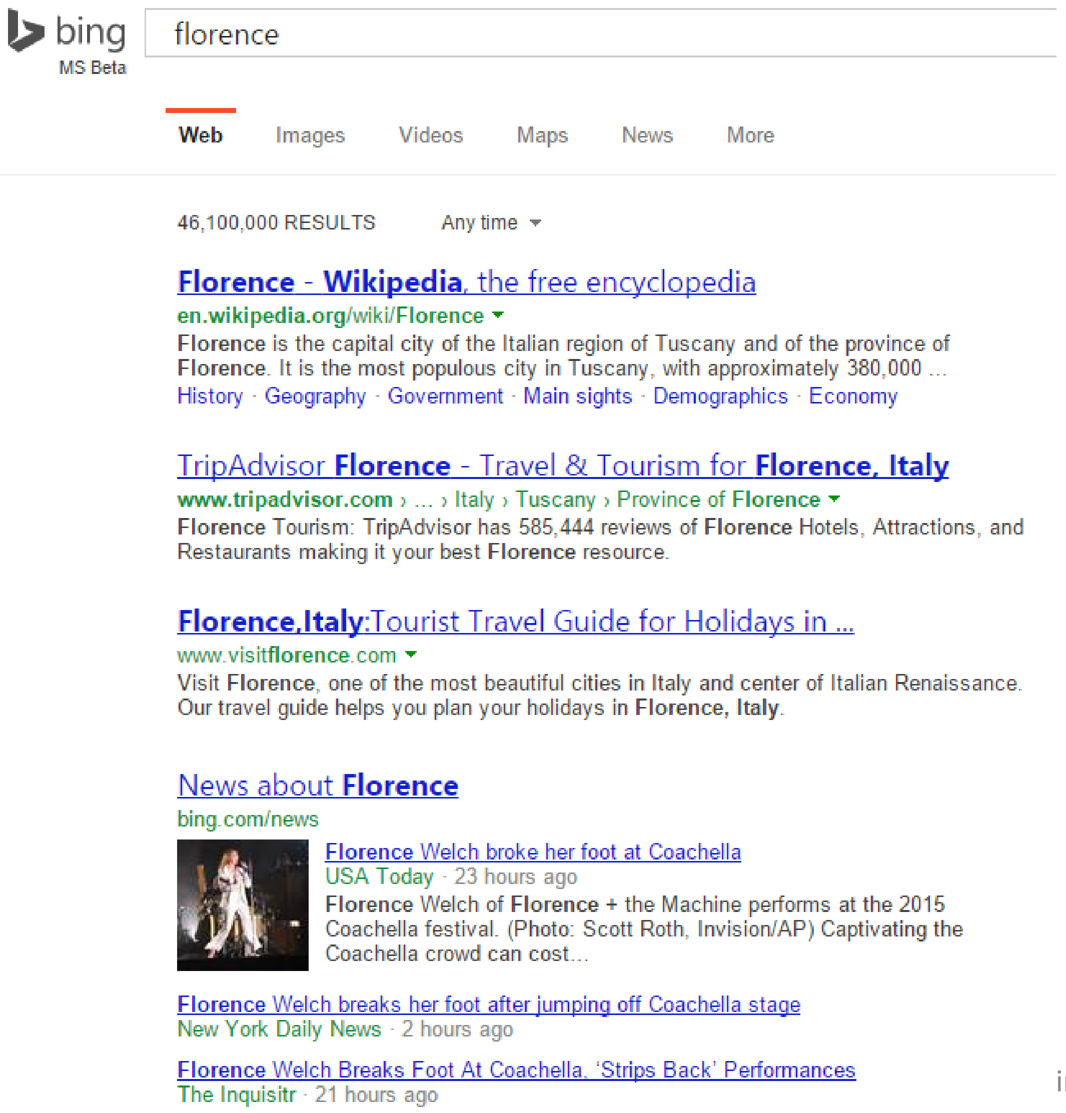
Things to do in Florence with kids - KY on FAMILYdaysOUT …
www.familydaysout.com/kids-things-to-do-usa/florence/ky
Fun things to do in Florence for kids - Kentucky - all colour coded by category - great places to go with children and family attractions to visit - preschoolers play

Why "Entities"

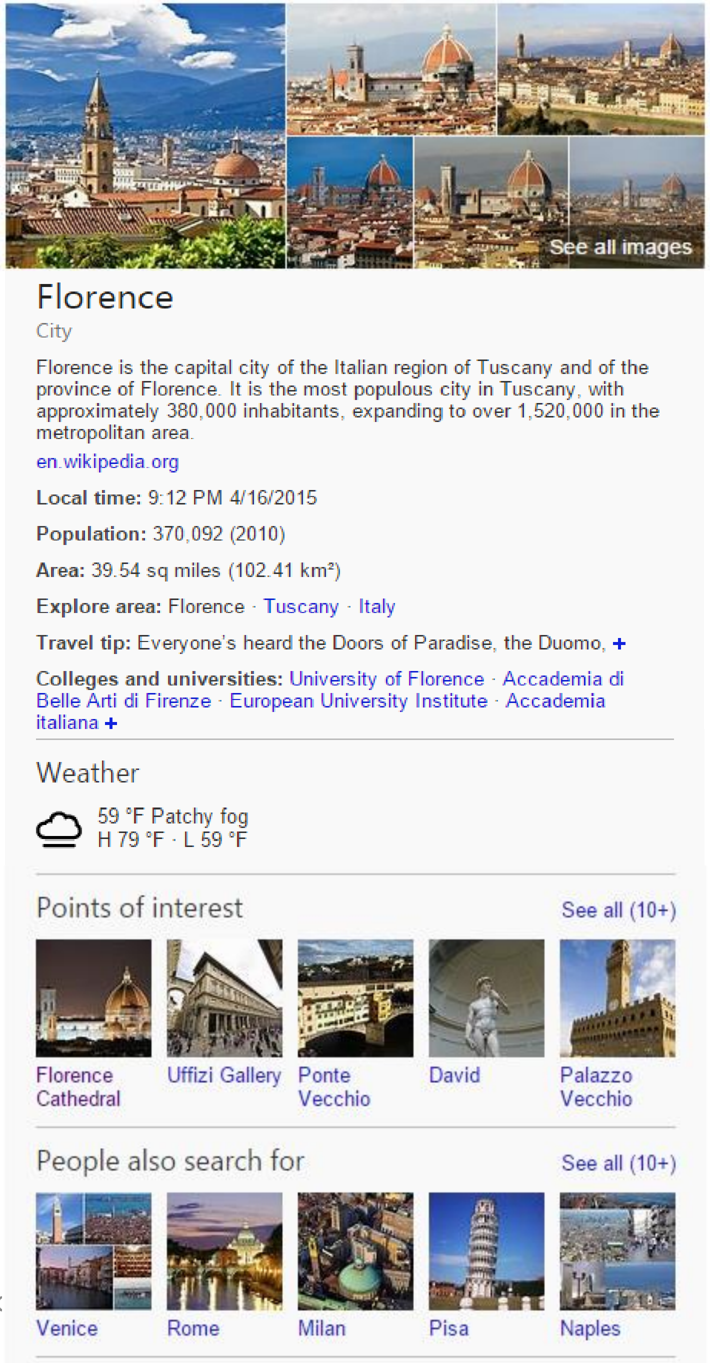




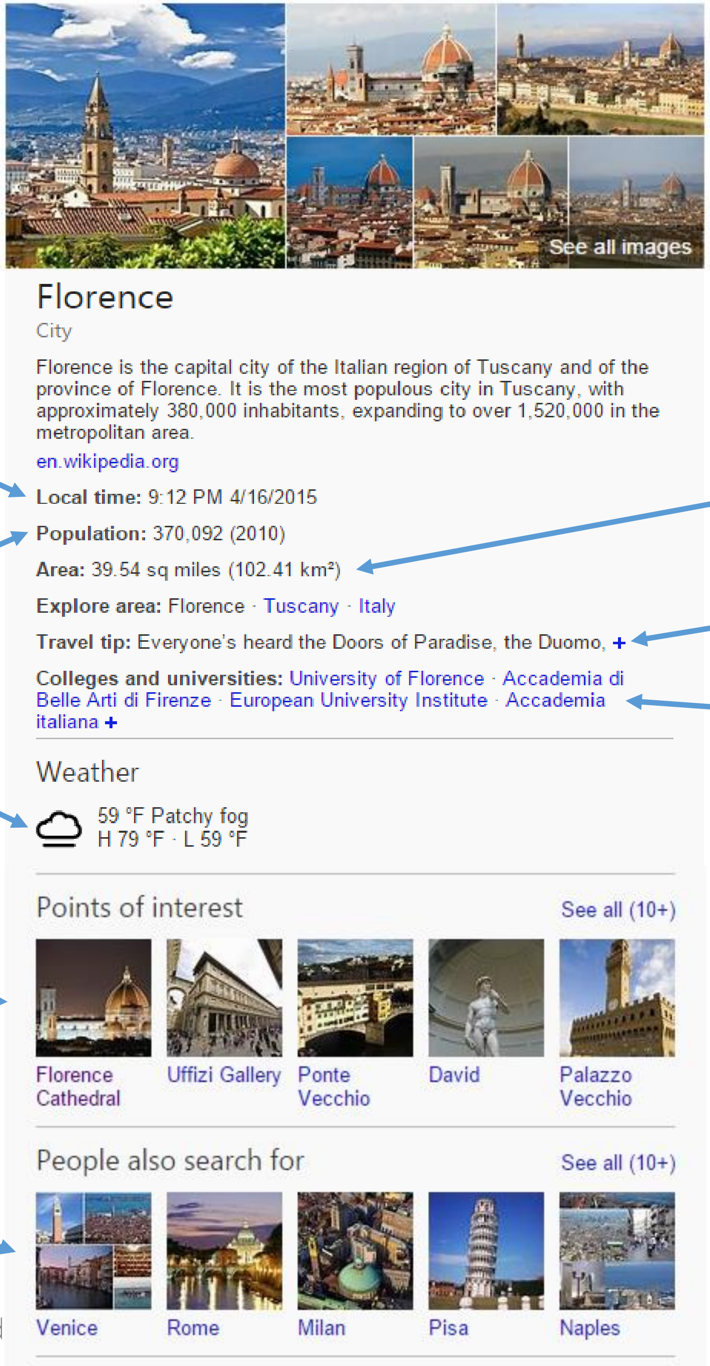
Why “Entities”



ing, Florenc



Why "Entities"



Florence photos

Time in florence

How big is florence

Florence population

Florence travel tips

Popular schools in florence

What's the weather in florence

Florence places to go

Cities similar to florence

One single Entity Pane can answer many user queries and satisfy users' diverse information needs

Why “Entities”

bing
MS Beta

florence

Web Images Videos Maps News More

46,100,000 RESULTS Any time

[**Florence** - **Wikipedia**, the free encyclopedia]
en.wikipedia.org/wiki/**Florence**
Florence is the capital city of the Italian region of Tuscany and of the province of **Florence**. It is the most populous city in Tuscany, with approximately 380,000 ...
History · Geography · Government · Main sights · Demographics · Economy

[TripAdvisor **Florence** - Travel & Tourism for **Florence, Italy**]
www.tripadvisor.com › ... › Italy › Tuscany › Province of **Florence**
Florence Tourism: TripAdvisor has 585,444 reviews of **Florence** Hotels, Attractions, and Restaurants making it your best **Florence** resource.

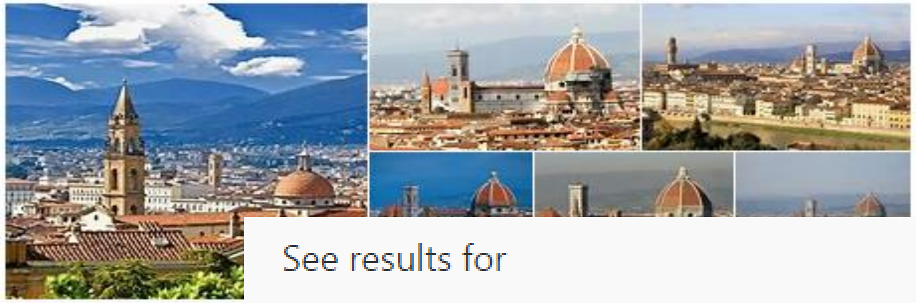
[**Florence,Italy**:Tourist Travel Guide for Holidays in ...]
www.visit**florence**.com
Visit **Florence**, one of the most beautiful cities in Italy and center of Italian Renaissance. Our travel guide helps you plan your holidays in **Florence, Italy**.

[News about **Florence**]
bing.com/news

[**Florence** Welch broke her foot at Coachella]
USA Today · 23 hours ago
Florence Welch of **Florence** + the Machine performs at the 2015 Coachella festival. (Photo: Scott Roth, Invision/AP) Captivating the Coachella crowd can cost...

[**Florence** Welch breaks her foot after jumping off Coachella stage]
New York Daily News · 2 hours ago

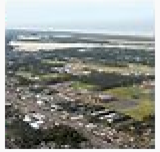
[**Florence** Welch Breaks Foot At Coachella, ‘Strips Back’ Performances]
The Inquisitr · 21 hours ago



Florence
City
Florence is the ca
province of Flore
approximately 38
metropolitan area
en.wikipedia.org
Local time: 9:12
Population: 370,
Area: 39.54 sq m
Explore area: Flo
Travel tip: Every
Colleges and un
Belle Arti di Firen
italiana +

See results for

Florence
Oregon
Florence is a city in Lane County, Oregon, United States. As of the 2010 census, the city had a total pop...



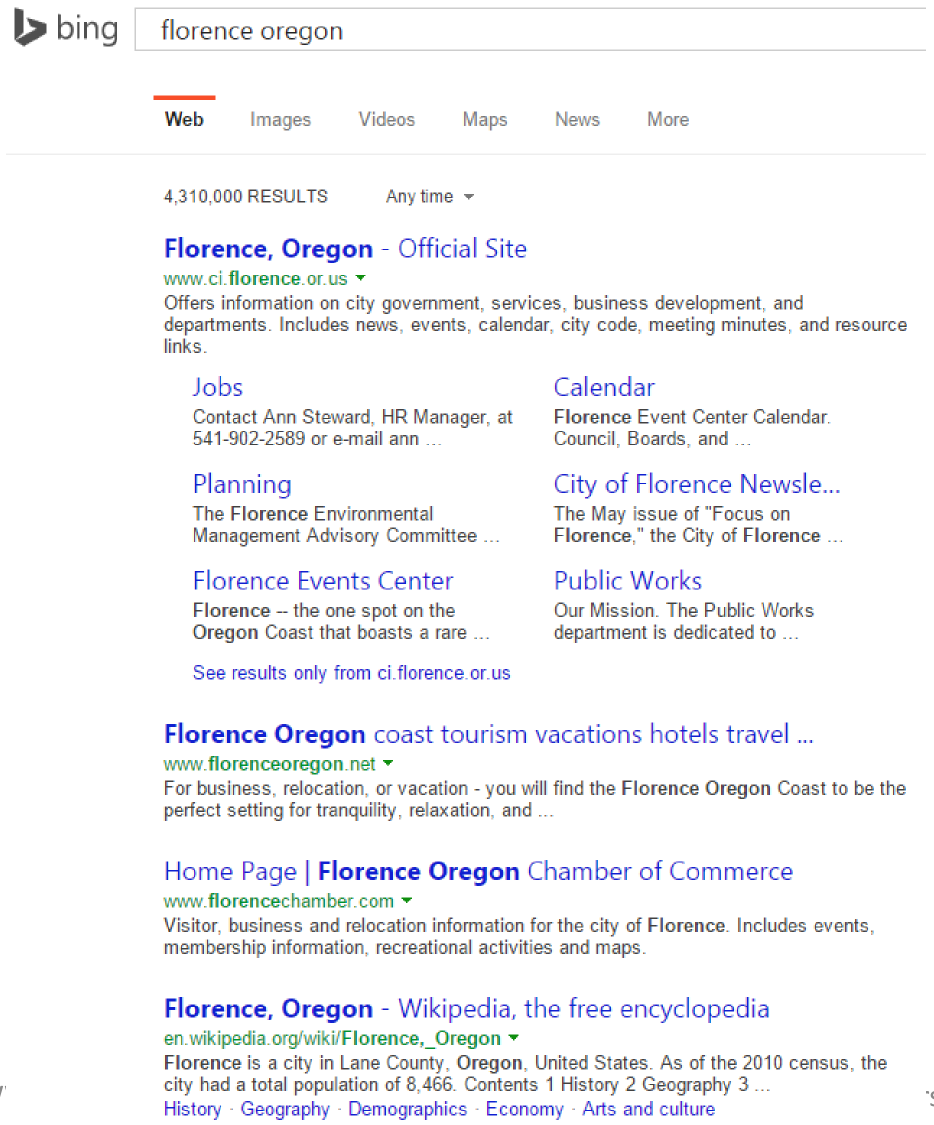
Florence
South Carolina
Florence is a city located in Florence County, South Carolina, United States. The city is the county seat of ...

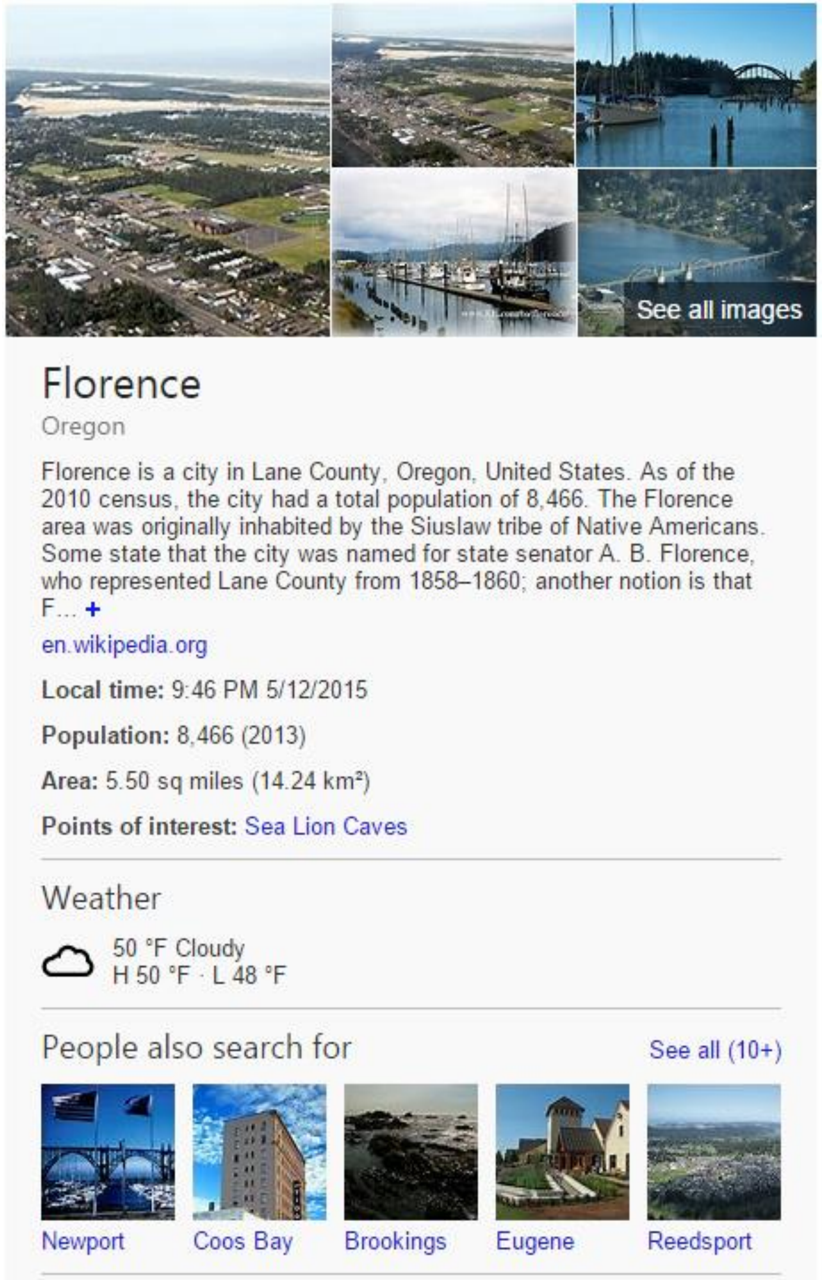
Florence
Kentucky
Florence is a home rule-class city in Boone County, Kentucky, in the United States. Florence is part of the ...

Florence
New Jersey
Florence Township is a township in Burlington County, New Jersey, United States. As of the 2010 United Stat...

Florence
Arizona
Florence is a town in and the county seat of Pinal County, Arizona, United States. The population was 25...

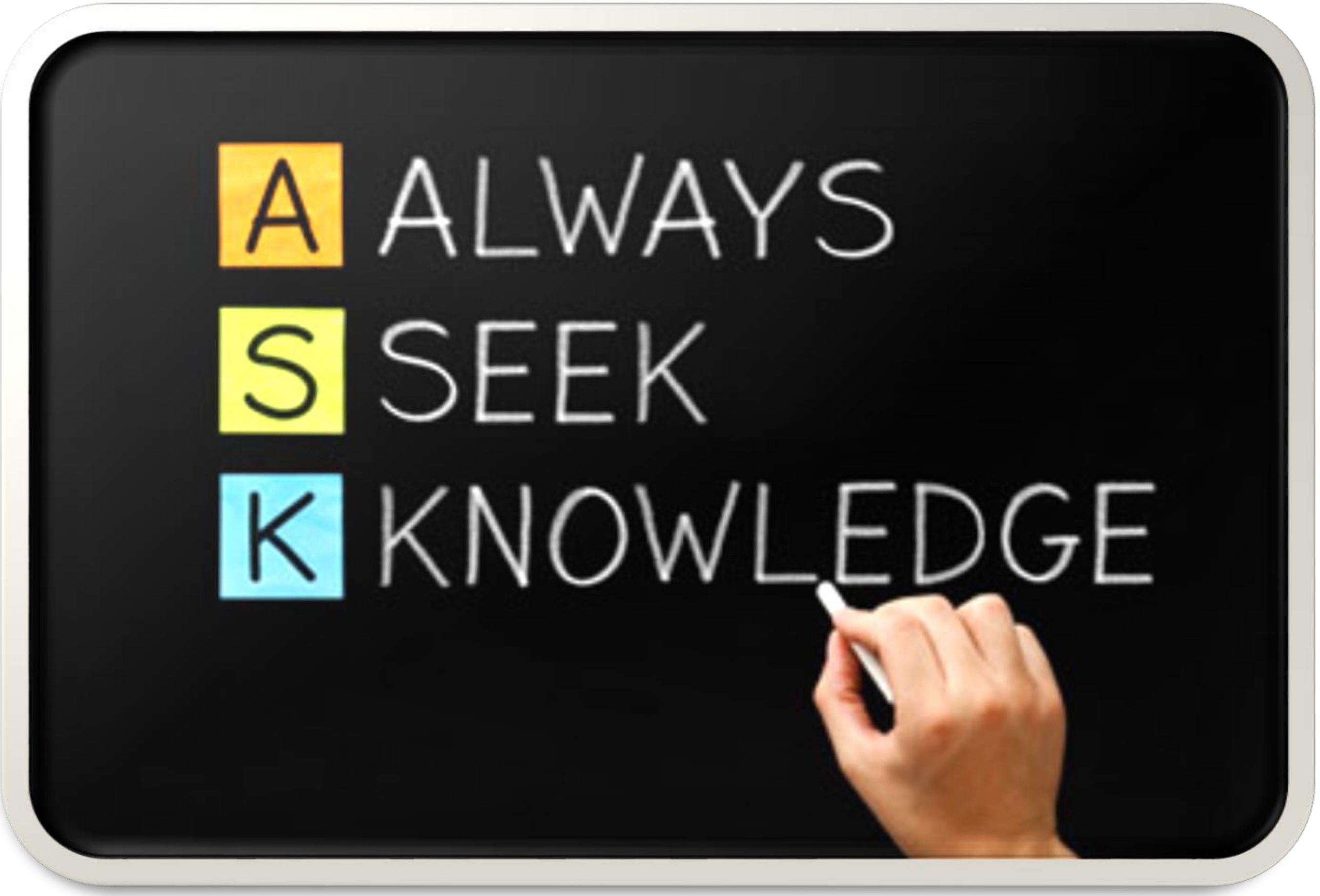
Why "Entities"





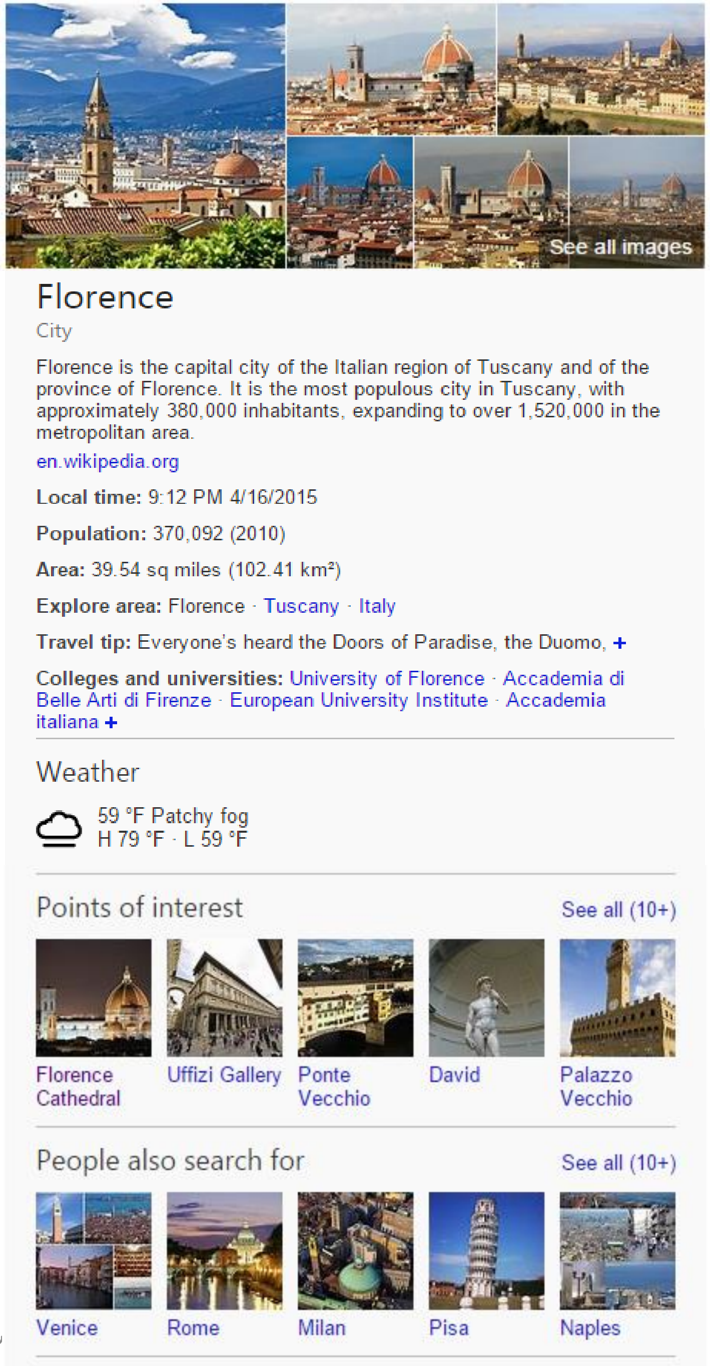
Because Entities are Surrounded by Knowledge

Why People Search?



Our Job


A ALWAYS
S SEEK
K KNOWLEDGE

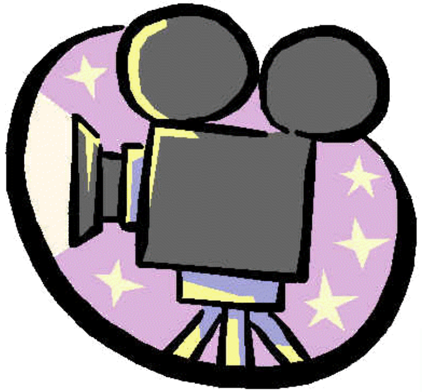

See all images
Florence
City
Florence is the capital city of the Italian region of Tuscany and of the province of Florence. It is the most populous city in Tuscany, with approximately 380,000 inhabitants, expanding to over 1,520,000 in the metropolitan area.
en.wikipedia.org
Local time: 9:12 PM 4/16/2015
Population: 370,092 (2010)
Area: 39.54 sq miles (102.41 km²)
Explore area: Florence · Tuscany · Italy
Travel tip: Everyone's heard the Doors of Paradise, the Duomo, +
Colleges and universities: University of Florence · Accademia di Belle Arti di Firenze · European University Institute · Accademia italiana +
Weather
59 °F Patchy fog
H 79 °F · L 59 °F
Points of interest
See all (10+)
Florence Cathedral
Uffizi Gallery
Ponte Vecchio
David
Palazzo Vecchio
People also search for
See all (10+)
Venice
Rome
Milan
Pisa
Naples

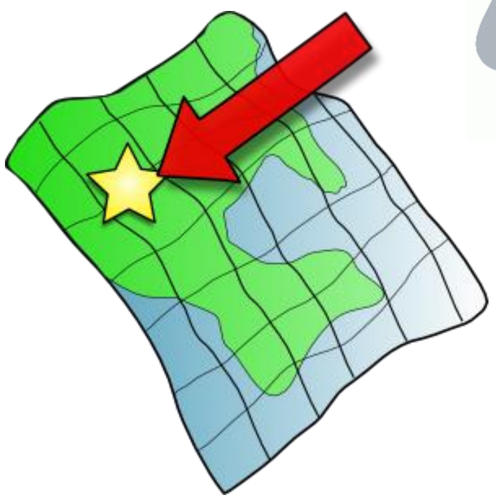
Entity Graphs

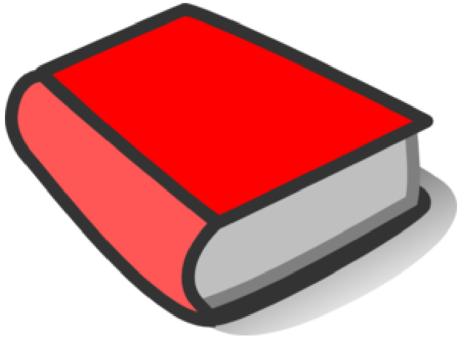
Knowledge Graph

Satori Knowledge Base

Entity Definition







Knowledge Definition

Address: 400 Broad St, Seattle, 98109
Phone: (800) 937-9582
Opened: Apr 21, 1962
Height: 605 feet (184.41 m)
Floors: 6

NFL championships: 2013
Head coach: Pete Carroll
Founded: 1976
Division: NFC West

Location

Home Field

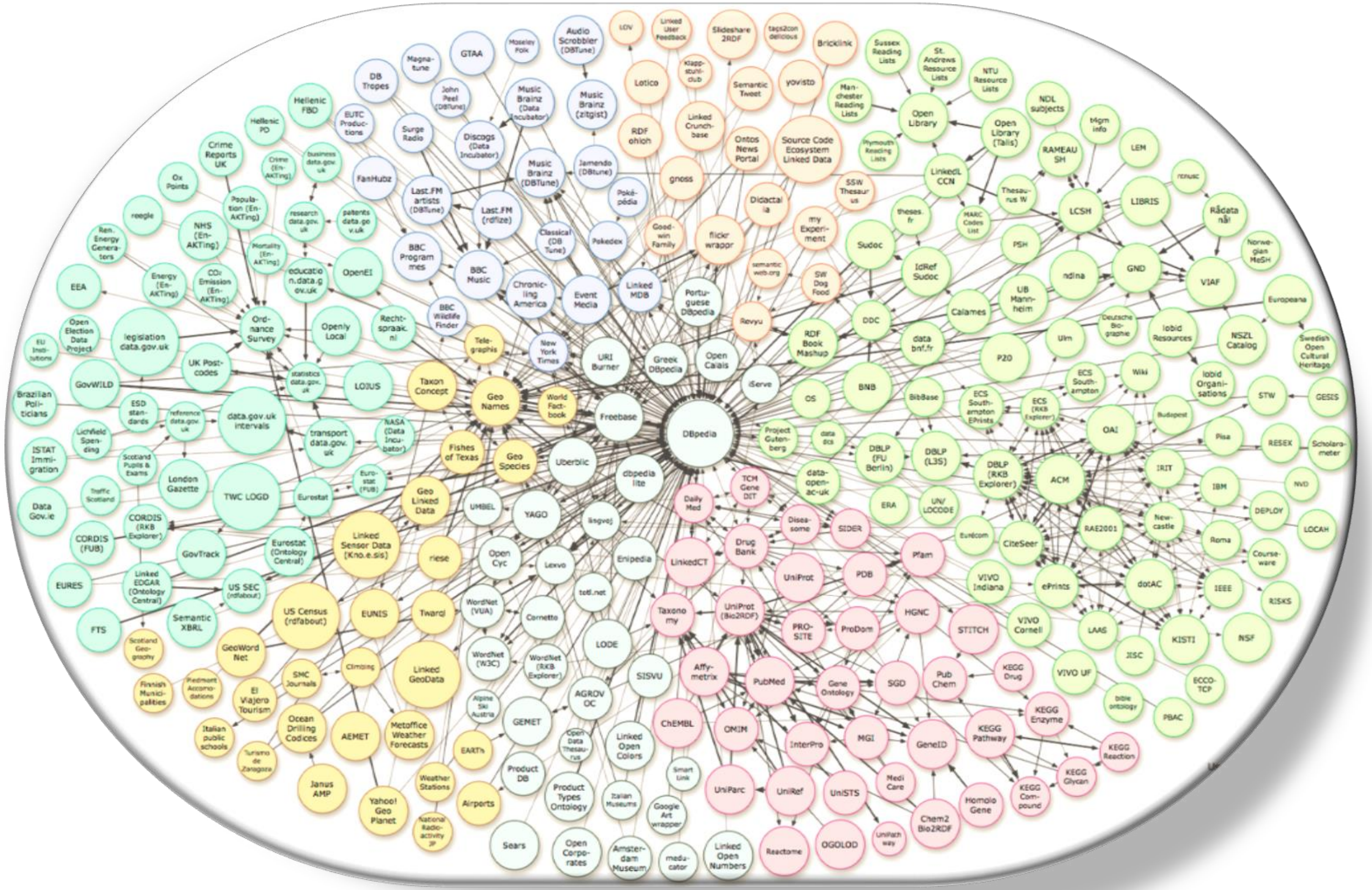
Seattle

Population: 652,405 (2013)
Area: 142.55 sq miles (369.20 km²)
Mayor: Ed Murray

Headquarters

Founded: Mar 30, 1971 · Pike Place Market
Customer service: +1 800-782-7282
CEO: Howard Schultz
Founders: Jerry Baldwin · Zev Siegl · Gordon Bowker

Entity Graphs



Why "Entity Recommendation"

- Information Explosion

Why “Entity Recommendation”

Freebase

47,558,128 Topics (and counting)

2,899,191,441 Facts (and counting)

DBpedia English

4,580,000 Topics (and counting)

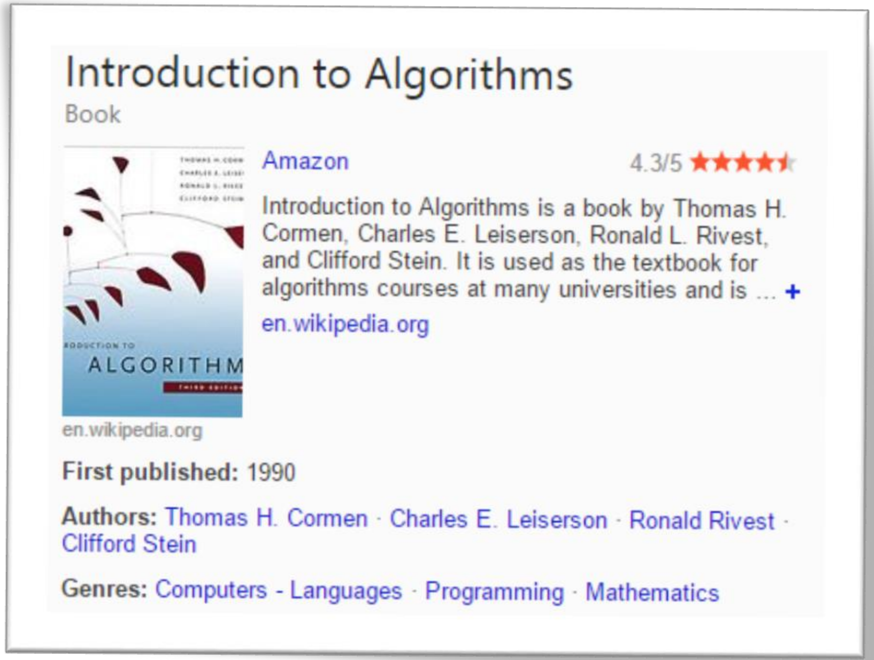
580,000,000 Facts (and counting)

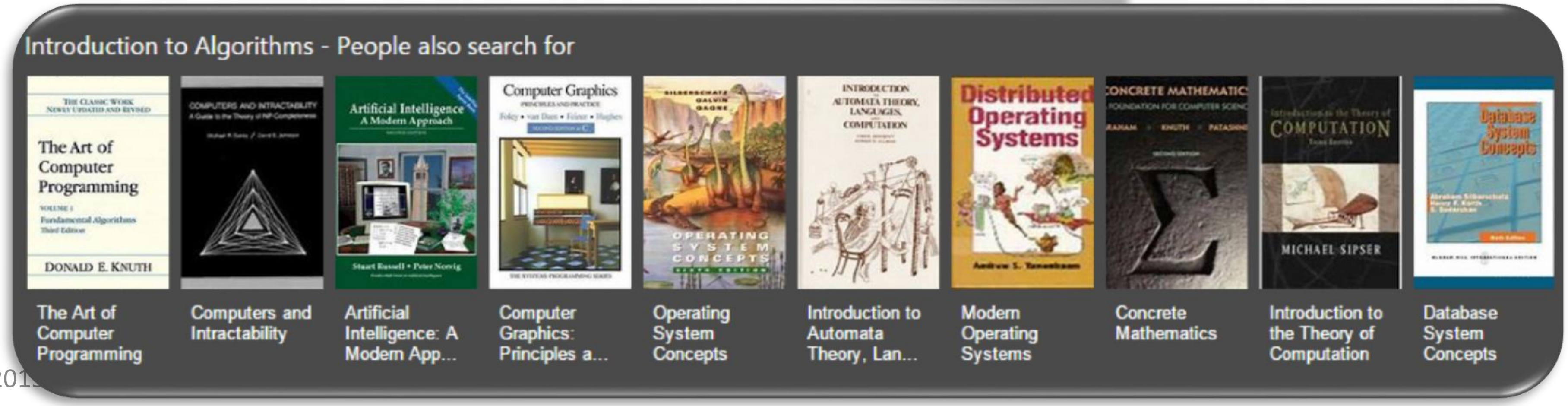
Why "Entity Recommendation"

- Information Overload

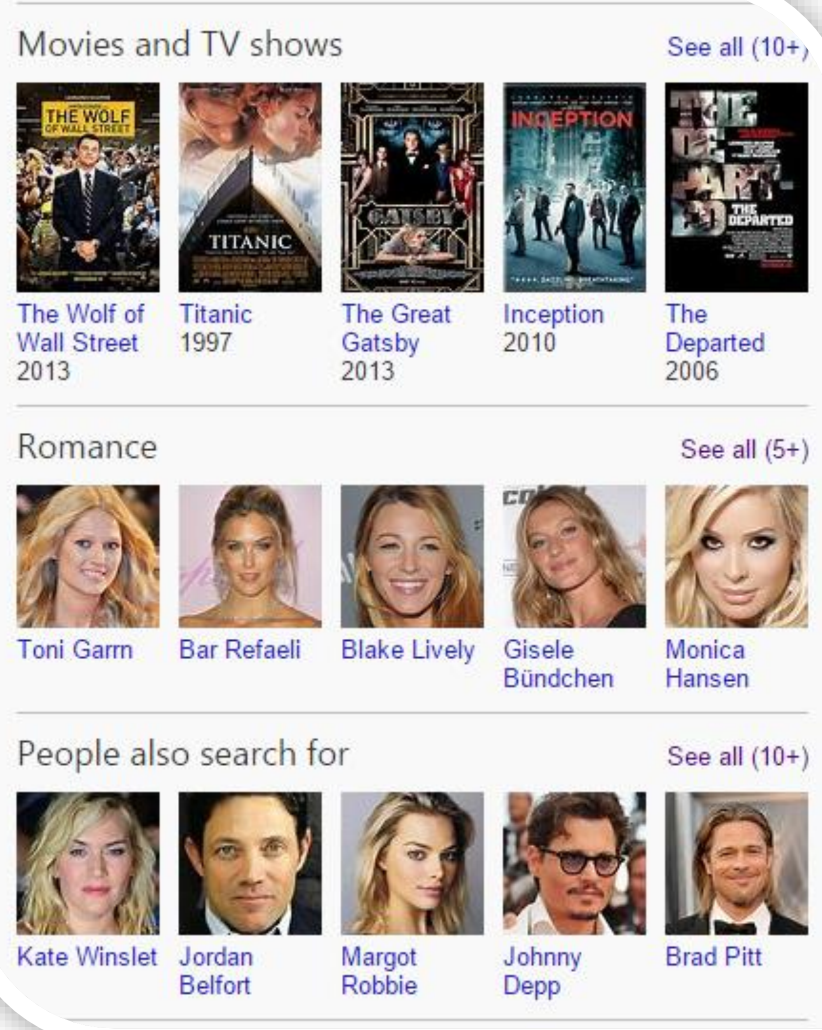
Help Explore the Knowledge Base

Why “Entity Recommendation”





Why “Entity Recommendation”



Why "Entity Recommendation"

bing machine learning

Machine learning

Machine learning is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the construction and study of algorithms that can ... +

en.wikipedia.org

Subdisciplines of: Artificial intelligence · Computer Science

Subdisciplines: Supervised learning · Deep learning

Academic conferences: AAAI 2016 · NIPS 2015 · ICML 2015 · IJCAI-15 · KDD 2015 · CVPR 2015 · ICASSP 2016 · ICDM 2015 +

Related people

See all (10+)

Andrew Ng · Arthur Samuel · Michael I. Jordan · Geoffrey Hinton · Tom M. Mitchell

People also search for

- Data mining
- Artificial neural network
- Artificial intelligence
- Supervised learning
- Pattern recognition
- Deep learning
- Unsupervised learning

bing www 2015

WWW 2015

Dates: May 20 - 22, 2015

Location: Florence

Website: WWW 2015

Abstracts due: Nov 03, 2014

Submissions due: Nov 10, 2014

Notification date: Jan 17, 2015

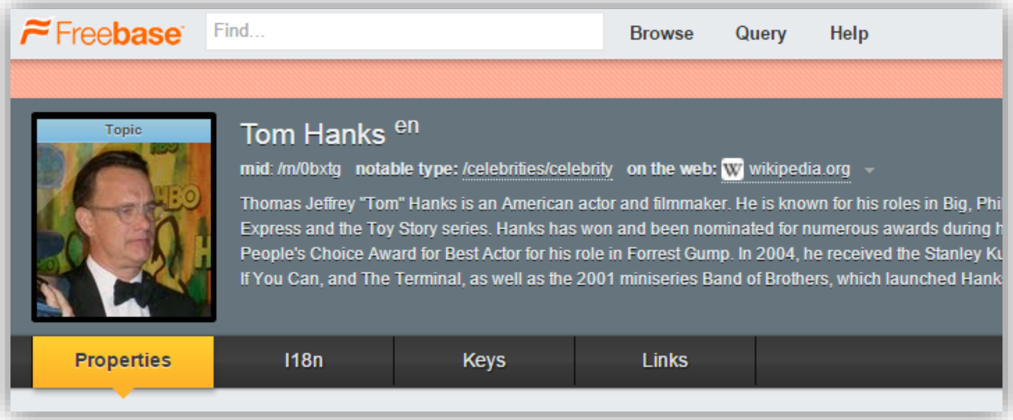
Final version due: Mar 08, 2015

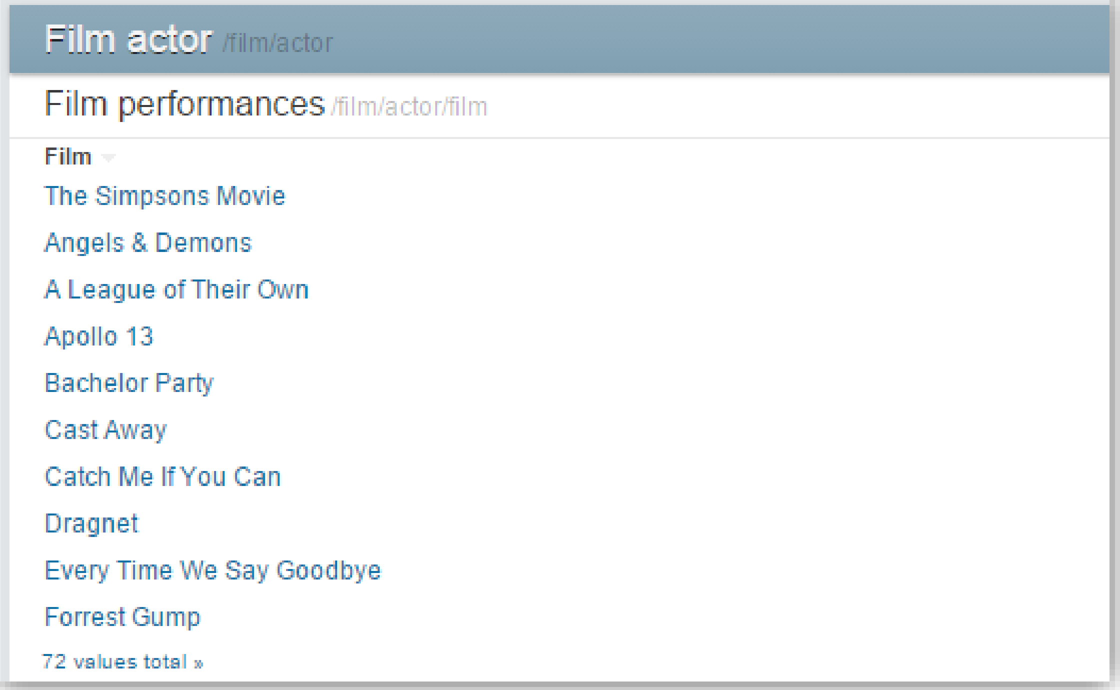
People also search for

- CIKM 2015 (Oct 19, 2015)
- AAAI 2016 (Feb 12, 2016)
- SIGIR 2015 (Aug 09, 2015)
- KDD 2015 (Aug 10, 2015)
- VLDB 2015 (Aug 31, 2015)
- IJCAI-15 (Jul 25, 2015)
- ICDE 2015 (Apr 13, 2015)

Why "Entity Understanding"

- Knowledge Bases are just an unordered list of facts





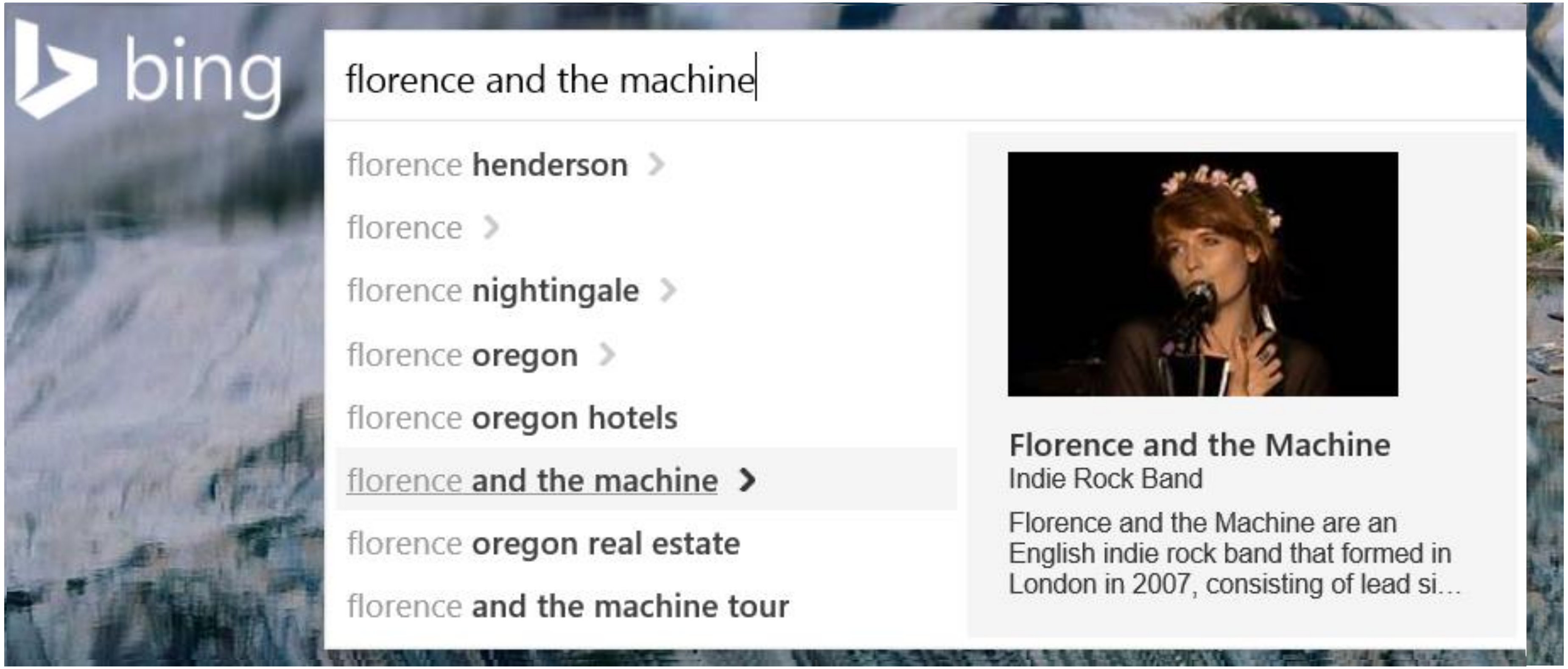
Why “Entity Understanding”

- Knowledge Bases are just an unordered list of facts

- Understanding is
  - Ranking facts
  - Creating connections between entities
  - Connecting entities and facts to queries and documents

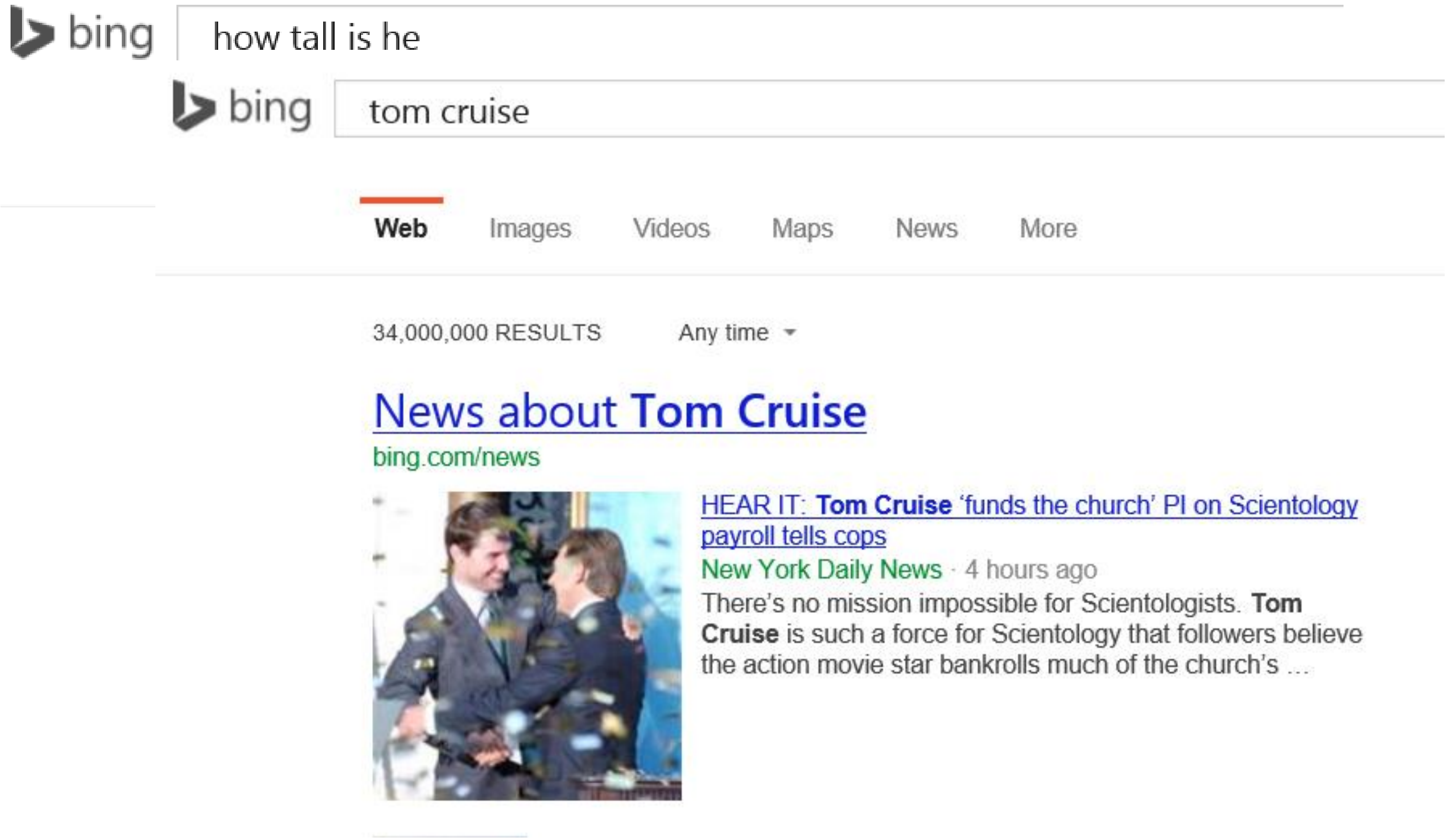
Why "Entity Understanding"

- When a user typed "Florence", how do you know which "Florence"?



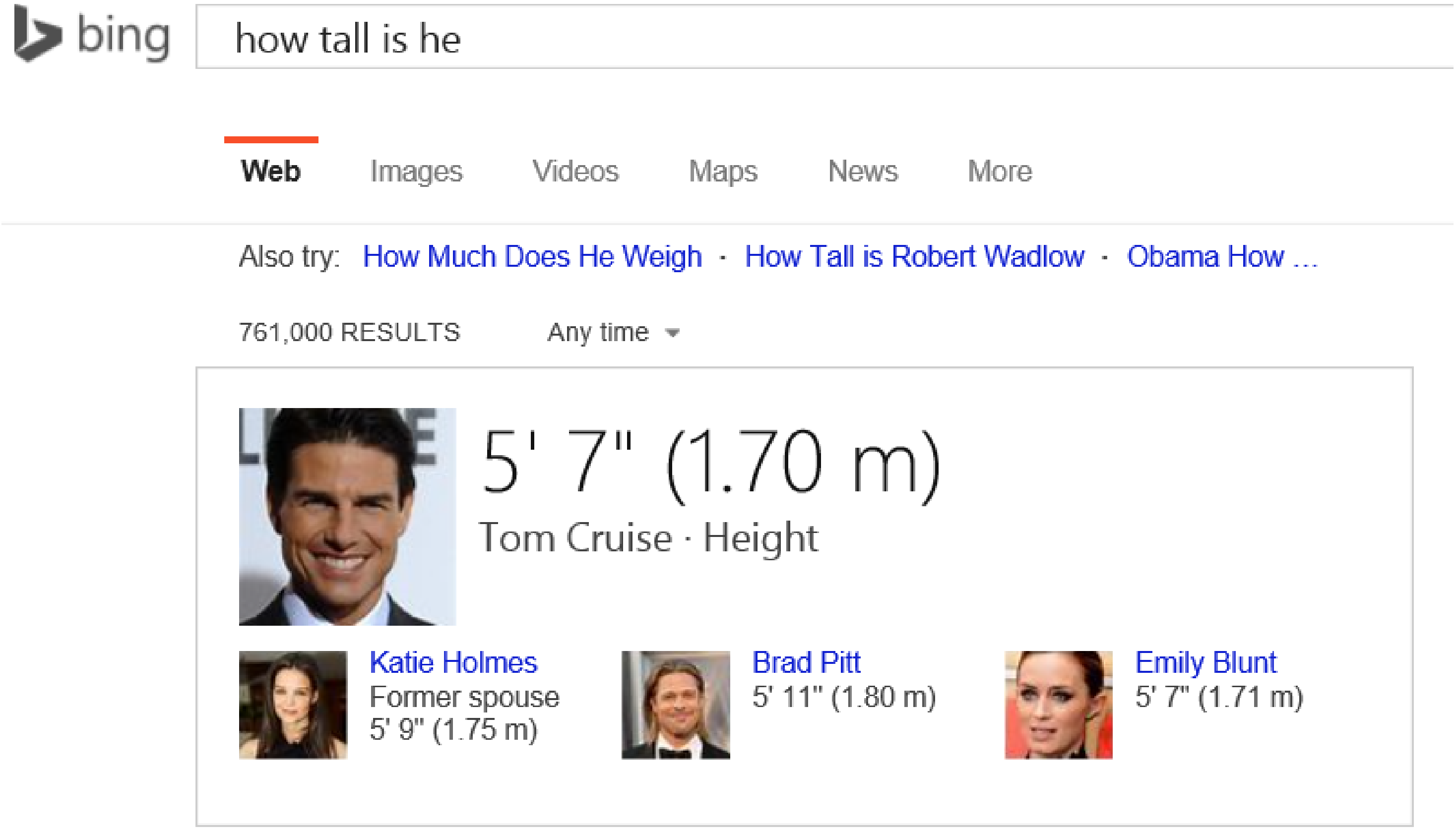
Why “Entity Understanding”

- When a user typed “how tall is he”, how do you know who is “he”?



Why "Entity Understanding"

- When a user typed "how tall is he", how do you know who is "he"?



Why "Entity Understanding"

- When a user clicked a few Web pages, how do you know what kind of entities this user is interested in?

Technologies

- Natural Language Processing
- Machine Learning
- Information Retrieval
- Recommender Systems
- Text and Log Mining
- ……

Data Sources

- Wikipedia
  - Semi-structured free Internet encyclopedia, contributed by community members
- Freebase
  - Structured data composed mainly by its community members
  - Acquired by Google on July 2010, and will be retired on June 2015
  - Data will be ported to WikiData
- WikiData
  - a collaboratively edited knowledge base
- DBPedia
  - Extracted structured information from Wikipedia
- Yago
  - a knowledge base automatically extracted from Wikipedia and other sources
  - Accuracy 95.02% based on manual evaulation
- Web Documents
- Queries and Search Click Log

Applications

- Entity Pane Experiences
- Entity Recommendation
  – Recommendation and Ranking
  – Interpretation
  – Exploration
  – Personalization
- Factoid Answers
- Graph Search
- Conversational Question and Answering
- Natural Language Question and Answering
- … …

Outline

- Introduction to Entity and Knowledge
- Demonstration of Microsoft's Entity Experience
- Entity Recommendation and Understanding
  - $P(entity|entity)$
  - $P(entity|user)$
  - $P(entity|query)$
- Summary

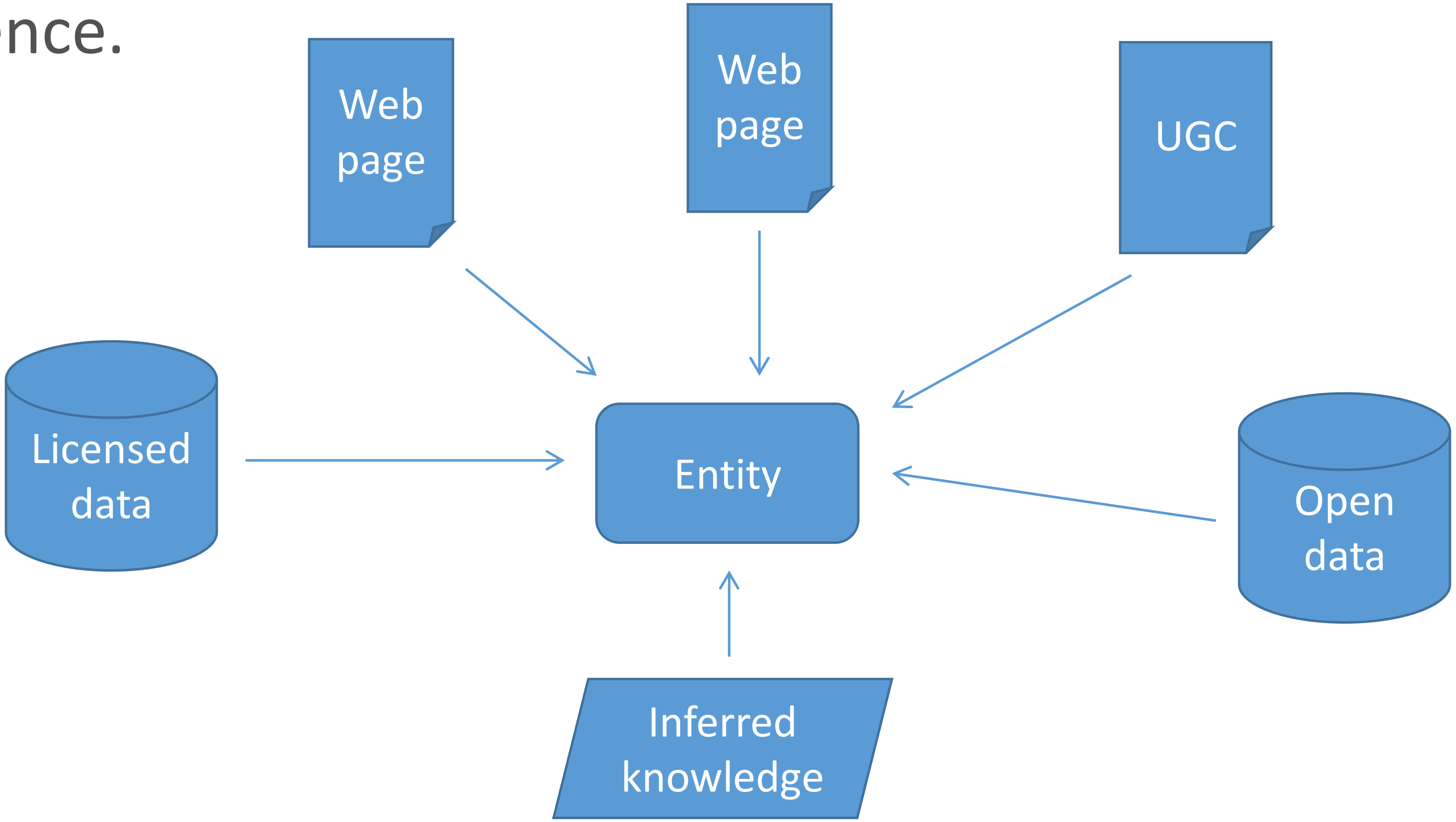
Demonstration of Microsoft's Entity Experience

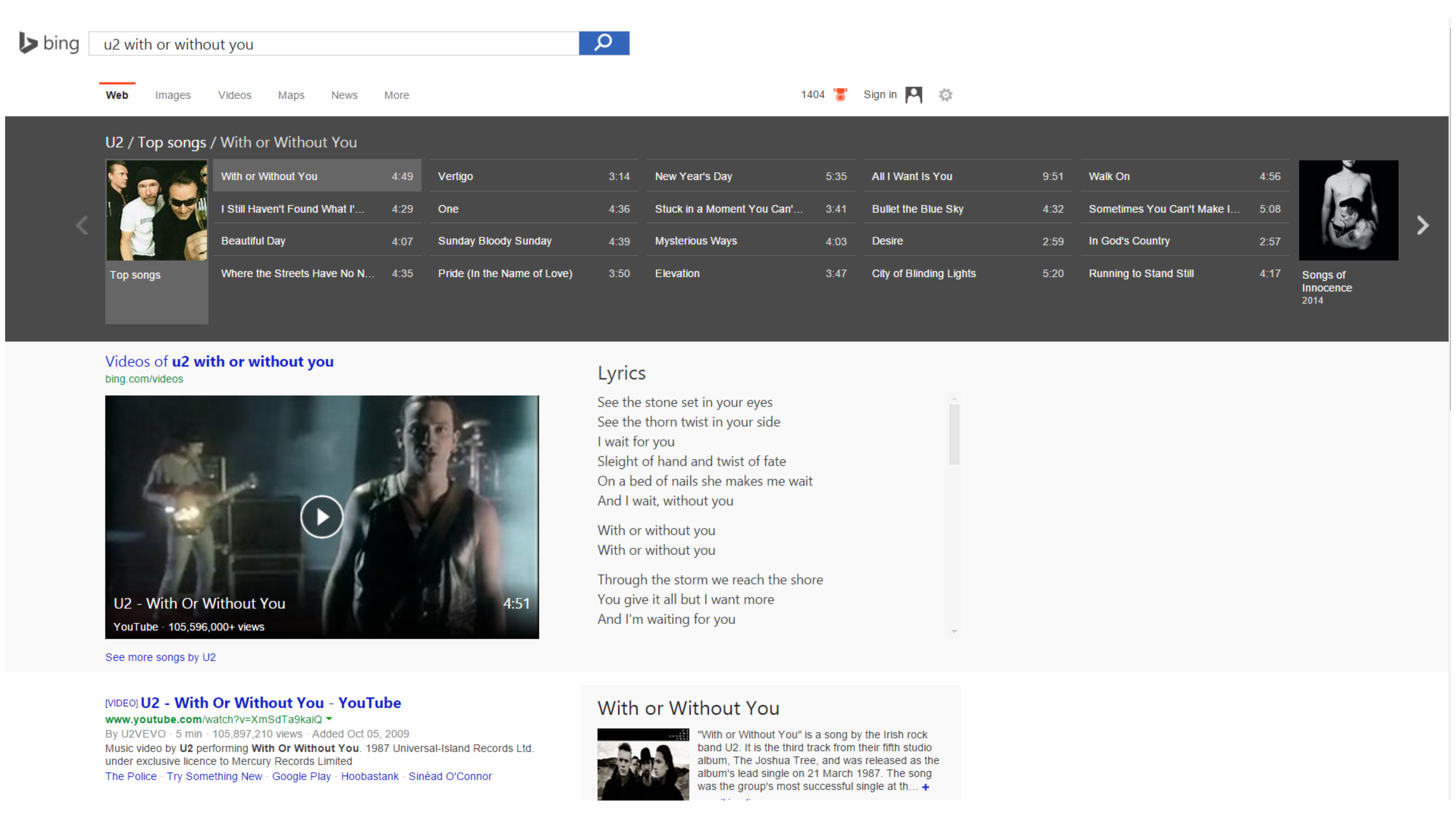
Entities are deeply integrated

- Bing
  - SERP, Images, Videos, Maps, …
- Office
- Windows
- Edge Browser
- Phone
- Xbox

Integrated Entity Experiences

- Combine data from many sources for an entity to build a rich user experience.





bing
u2 with or without you
Web
Images
Videos
Maps
News
More
1404
Sign in
U2 / Top songs / With or Without You
Top songs
With or Without You 4:49
I Still Haven't Found What I'... 4:29
Beautiful Day 4:07
Where the Streets Have No N... 4:35
Vertigo 3:14
One 4:36
Sunday Bloody Sunday 4:39
Pride (In the Name of Love) 3:50
New Year's Day 5:35
Stuck in a Moment You Can'... 3:41
Mysterious Ways 4:03
Elevation 3:47
All I Want Is You 9:51
Bullet the Blue Sky 4:32
Desire 2:59
City of Blinding Lights 5:20
Walk On 4:56
Sometimes You Can't Make I... 5:08
In God's Country 2:57
Running to Stand Still 4:17
Songs of Innocence
2014
Videos of u2 with or without you
bing.com/videos
U2 - With Or Without You 4:51
YouTube · 105,596,000+ views
See more songs by U2
Lyrics
See the stone set in your eyes
See the thorn twist in your side
I wait for you
Sleight of hand and twist of fate
On a bed of nails she makes me wait
And I wait, without you
With or without you
With or without you
Through the storm we reach the shore
You give it all but I want more
And I'm waiting for you
[VIDEO] U2 - With Or Without You - YouTube
www.youtube.com/watch?v=XmSdTa9kaiQ
By U2VEVO · 5 min · 105,897,210 views · Added Oct 05, 2009
Music video by U2 performing With Or Without You. 1987 Universal-Island Records Ltd. under exclusive licence to Mercury Records Limited
The Police · Try Something New · Google Play · Hoobastank · Sinéad O'Connor
With or Without You
"With or Without You" is a song by the Irish rock band U2. It is the third track from their fifth studio album, The Joshua Tree, and was released as the album's lead single on 21 March 1987. The song was the group's most successful single at th... +

You give it all but I want more
And I'm waiting for you

See more songs by U2

[VIDEO] **U2 - With Or Without You - YouTube**
www.youtube.com/watch?v=XmSdTa9kaiQ
By U2VEVO · 5 min · 104,667,422 views · Added Oct 05, 2009
Music video by **U2** performing **With Or Without You**. 1987 Universal-Island Records Ltd. under exclusive licence to Mercury Records Limited

With or Without You - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/With_or_Without_You
"**With or Without You**" is a song by the Irish rock band **U2**. It is the third track from their fifth studio album, **The Joshua Tree** (1987), and was released as the album ...
Writing and recording · Composition · Release and chart ... · Live performances

With Or Without You Lyrics - **U2** - **LyricsFreak.com**
www.lyricsfreak.com › U2
Lyrics to **With Or Without** by **U2**: See the stone set in your eyes / See the thorn twist in your side / I'll wait for **you** / Sleight of hand

U2 - With Or Without You Lyrics | **MetroLyrics**
www.metrolyrics.com/with-or-without-you-lyrics-u2.html
"**With Or Without You**" is track #4 on the album U218 Singles. It was written by Adam Clayton, Dave Evans, Paul Hewson, Larry Mullen. (No other information is available ...

Related searches for **u2 with or without you**

U2 All I Want Is **You**
With or Without You Lyrics
U2 Greatest Hits
U2 Sunday Bloody Sunday
With or Without You U2 YouTube
With or Without You Song

U2 - With or Without You Lyrics | **SongMeanings**
songmeanings.com › Artists - U

With or Without You

"With or Without You" is a song by the Irish rock band U2. It is the third track from their fifth studio album, The Joshua Tree, and was released as the album's lead single on 21 March 1987. The song was the group's most successful single at th... +
en.wikipedia.org

Album: The Joshua Tree

Artist: U2

Duration: 4:49

Genre: Rock, Mainstream Rock

Listen or buy

Xbox Music · iTunes
Amazon MP3

Also appears on

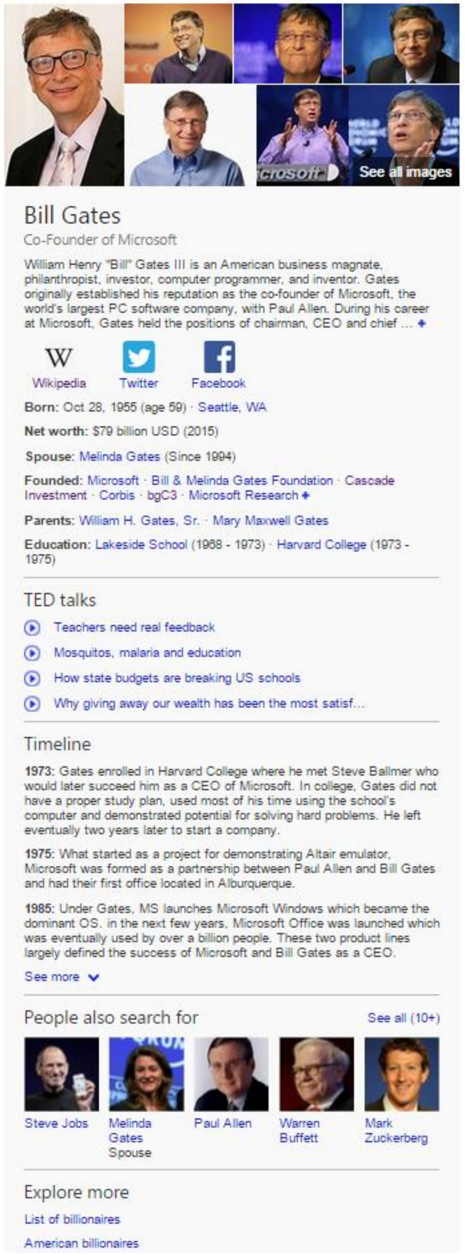
U218 Singles

The Best Of 1980 - 1990

Please

Data from: Xbox Music · LyricFind · Freebase

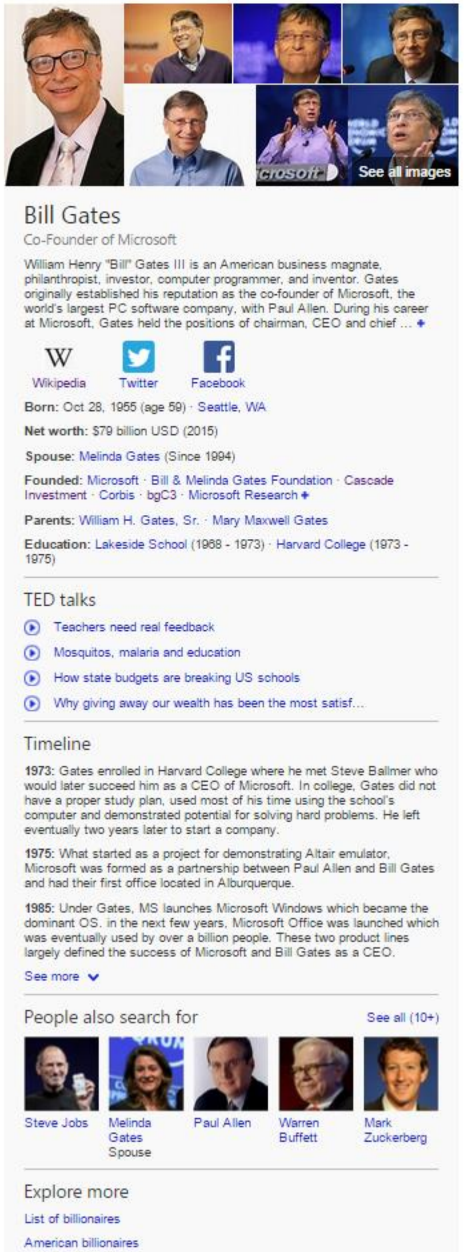
Feedback



TED talks

- Teachers need real feedback
- Mosquitos, malaria and education
- How state budgets are breaking US schools
- Why giving away our wealth has been the most satisf…

http://blogs.bing.com/search/2013/12/12/expand-your-understanding-with-bing/



Timeline

1973: Gates enrolled in Harvard College where he met Steve Ballmer who would later succeed him as a CEO of Microsoft. In college, Gates did not have a proper study plan, used most of his time using the school's computer and demonstrated potential for solving hard problems. He left eventually two years later to start a company.

1975: What started as a project for demonstrating Altair emulator, Microsoft was formed as a partnership between Paul Allen and Bill Gates and had their first office located in Alburquerque.

1985: Under Gates, MS launches Microsoft Windows which became the dominant OS. in the next few years, Microsoft Office was launched which was eventually used by over a billion people. These two product lines largely defined the success of Microsoft and Bill Gates as a CEO.

See more

http://blogs.bing.com/search/2014/02/21/timeline-understanding-important-events-in-peoples-lives/

Online Courses

Stanford University

University

Stanford University is a private research university in Stanford, California, and one of the world's most prestigious institutions, with the highest undergraduate selectivity and the top position in numerous surveys and measures in the Unite... +

en.wikipedia.org

Address: 450 Serra Mall, Stanford, CA 94305

Ranking: #4 National University (2015)

Undergraduates: 7,063 (2015)

Acceptance rate: 5.70% (2015)

Tuition: $43,245 USD (2015)

Founded: Oct 01, 1891

Popular online courses

- Machine Learning
- Natural Language Processing
- Child Nutrition and Cooking 2.0
- Automata
- Game Theory

See more

Physics

Physics is the natural science that involves the study of matter and its motion through space and time, along with related concepts such as energy and force. More broadly, it is the general analysis of nature, conducted in order to understand h... +

en.wikipedia.org

Subdiscipline of: Natural science

Popular online courses

- Calculating average velocity or speed
- Relationship between angular velocity and speed
- Displacement from time and velocity example
- Inclined plane force components
- Balanced and unbalanced forces

See more

Subdisciplines

See all (10+)

Astronomy

Quantum mechanics

Mechanics

Nuclear physics

Particle physics

Long tail of interconnected entities

Richard Qian

LinkedIn

Current: Corporate Vice President at Microsoft

Location: Greater Seattle Area

Experience

Corporate Vice President
Microsoft · 2015 - Present

Partner Development Manager - Web Indexing & Knowledge
Microsoft · 2008 - 2014

Principal Development Manager - Multimedia & News Search
Microsoft · 2006 - 2008

More on LinkedIn

Education

University of Illinois at Urbana-Champaign
Ph.D., Electrical & Computer Engineering · 1992 - 1996

University of Illinois at Urbana-Champaign
M.S., Electrical & Computer Engineering · 1991 - 1992

Tsinghua University
B.S., Computer Science · 1981 - 1986

More on LinkedIn

12160 168th Ct NE, Redmond, WA 98052

Beautiful Murray Franklyn home in Prescott. The Woodland floor plan has 4 bdrms/2.75 baths and bonus room. Upgraded throughout including slab granite, designer backsplash, Kitchen Aid appliances, & pendant lighting. Huge master … +

www.zillow.com

Directions | Zillow

Address: 12160 168th Ct NE,, Redmond, WA 98052

Floors: 2

Zestimate: $770,081

Year built: 2011

Price: $750,000

Rooms: 4 beds, 2.75 baths

Assigned schools

School	Distance	Rating
Albert Einstein Elementary School	1.20 miles	6/10
Evergreen Junior High School	4.20 miles	10/10
Redmond High School	1 miles	8/10

People also search for


17260 NE 123rd Way


17250 NE 123rd Way

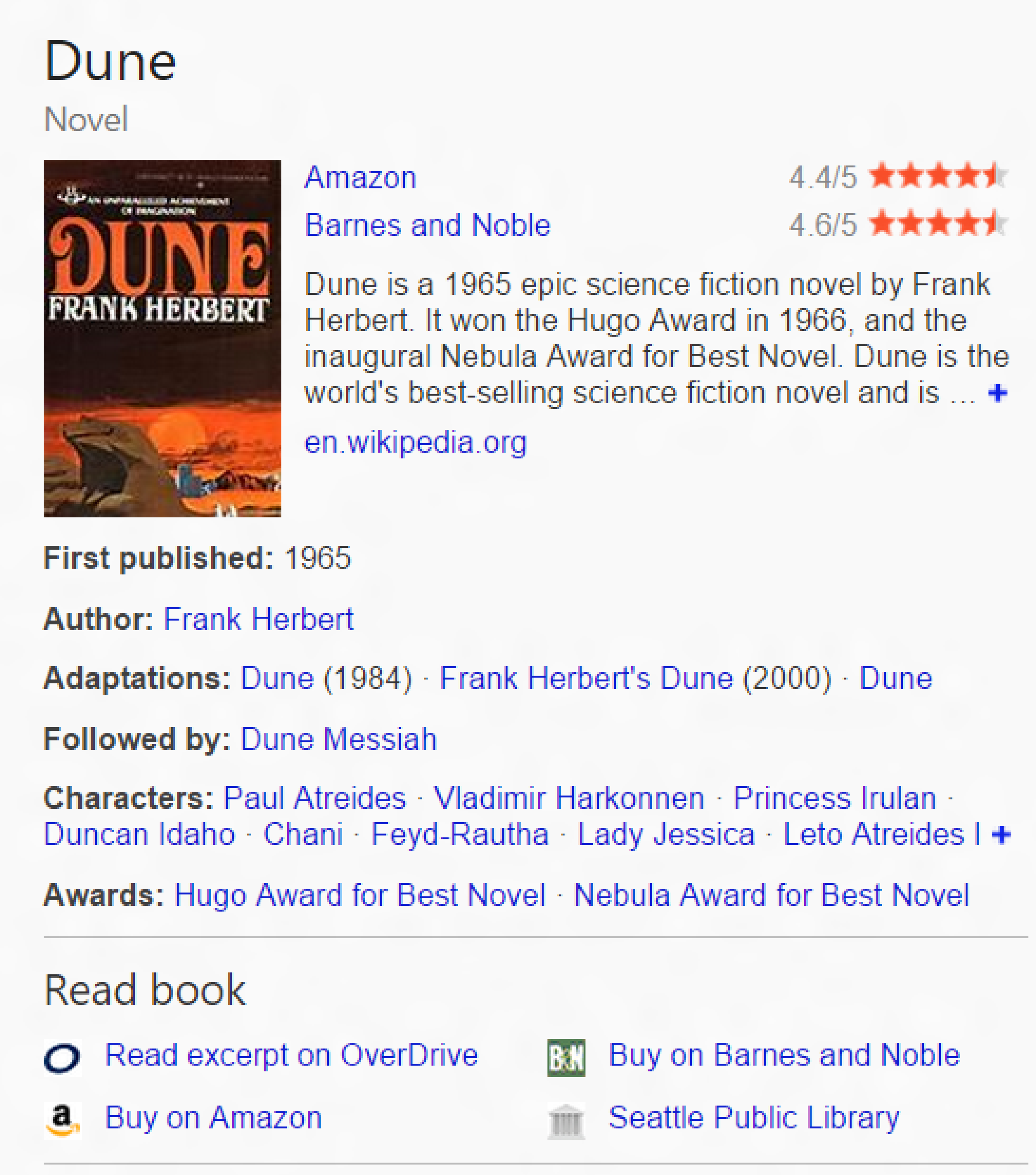

16829 NE 121st St


17215 NE 117th Way


17419 NE 123rd Way

http://blogs.bing.com/search/2014/03/31/150-million-more-reasons-to-love-bing-everyday/

Actions – facilitate task completion



The Avengers

PG-13 · 2hr 23min · Sci-Fi

IMDb 8.2/10

Rotten Tomatoes 92%

"Marvel's The Avengers"-the Super Hero team up of a lifetime, featuring iconic Marvel Super Heroes Iron Man, The Incredible Hulk, Thor, Captain America, Hawkeye and Black Widow. When an unexpected enemy emerges that threatens global safety and security, Nick Fury, Director of the international peacekeeping agency known as S.H.I.E.L.D., finds himself in need of a t... +

Estimated budget: $220 million USD

Release date: May 04, 2012

Director: Joss Whedon

Sequel: Avengers: Age of Ultron

Production company: Marvel Studios

Story by: Joss Whedon · Zak Penn

Watch now

Netflix

Amazon

Xbox Video

Watch trailer on IMDb

Songs of Innocence (2014)

Songs of Innocence is the thirteenth studio album by Irish rock band U2. Released on 9 September 2014, it was produced by Danger Mouse, with additional production from Paul Epworth, Ryan Tedder, Declan Gaffney and Flood. The albu... +

en.wikipedia.org

Genre: Rock, Mainstream Rock

Label: Interscope

Release year: 2014

Artist: U2

Listen or buy

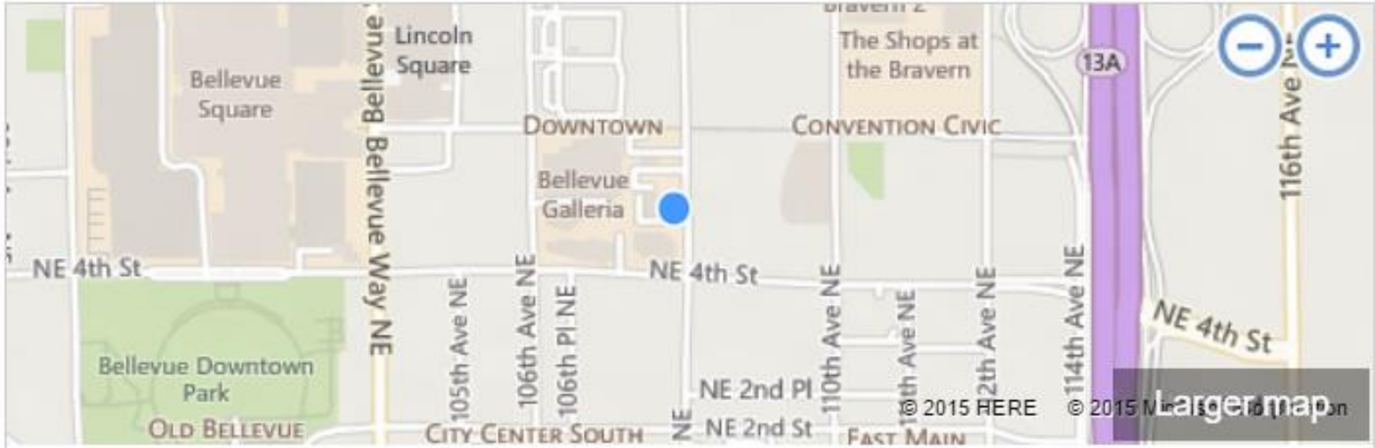
Xbox Music

iTunes

bing | el gaucho bellevue

Web Images Videos Maps News More

237,000 RESULTS Any time



El Gaucho

elgaucho.com

450 108th Avenue NE, Bellevue, WA 98004 · 6.31 mi

(425) 455-2715

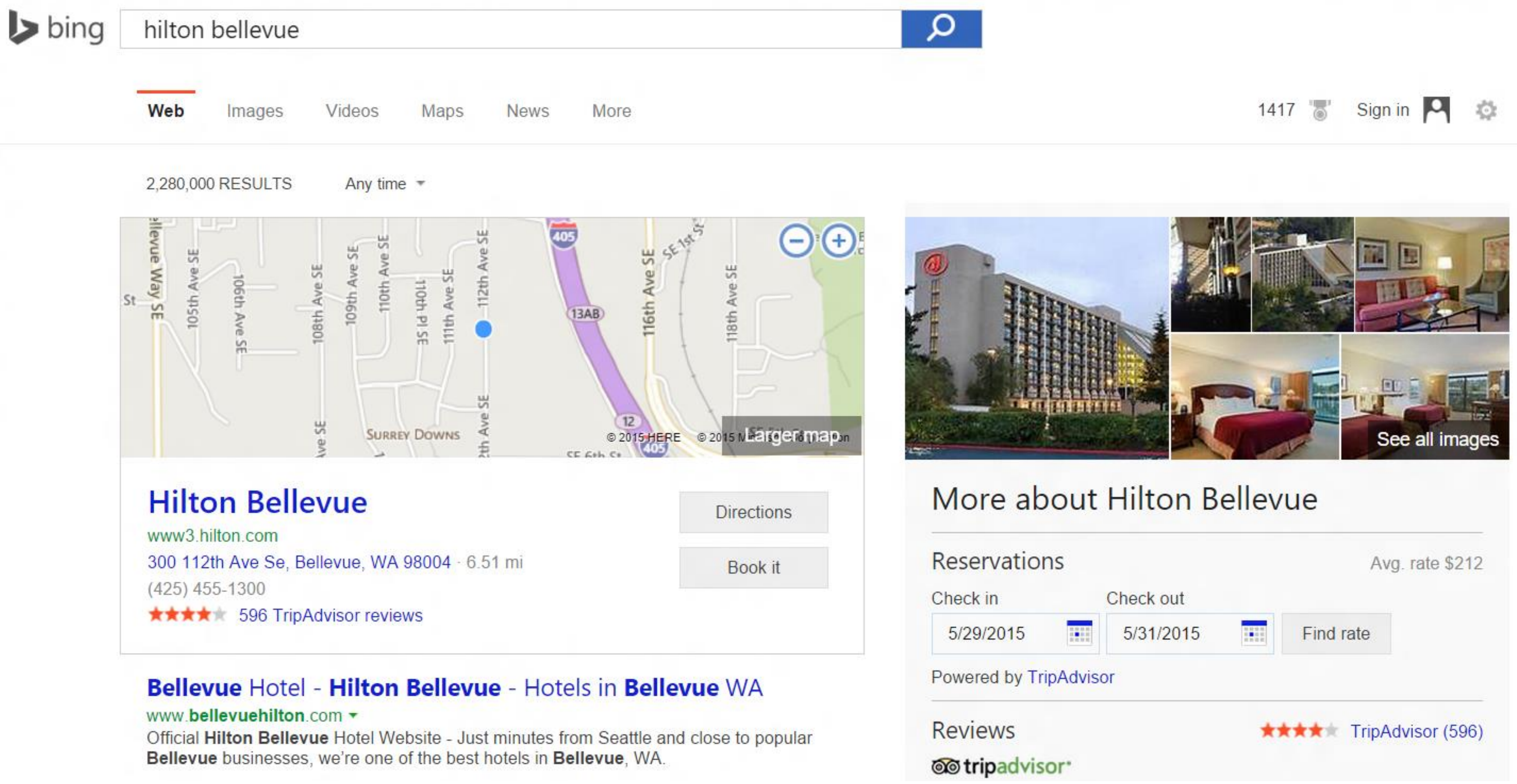
Open 11:30 AM - 10:00 PM

★★★★★ 347 Yelp reviews · $$$$$

Directions

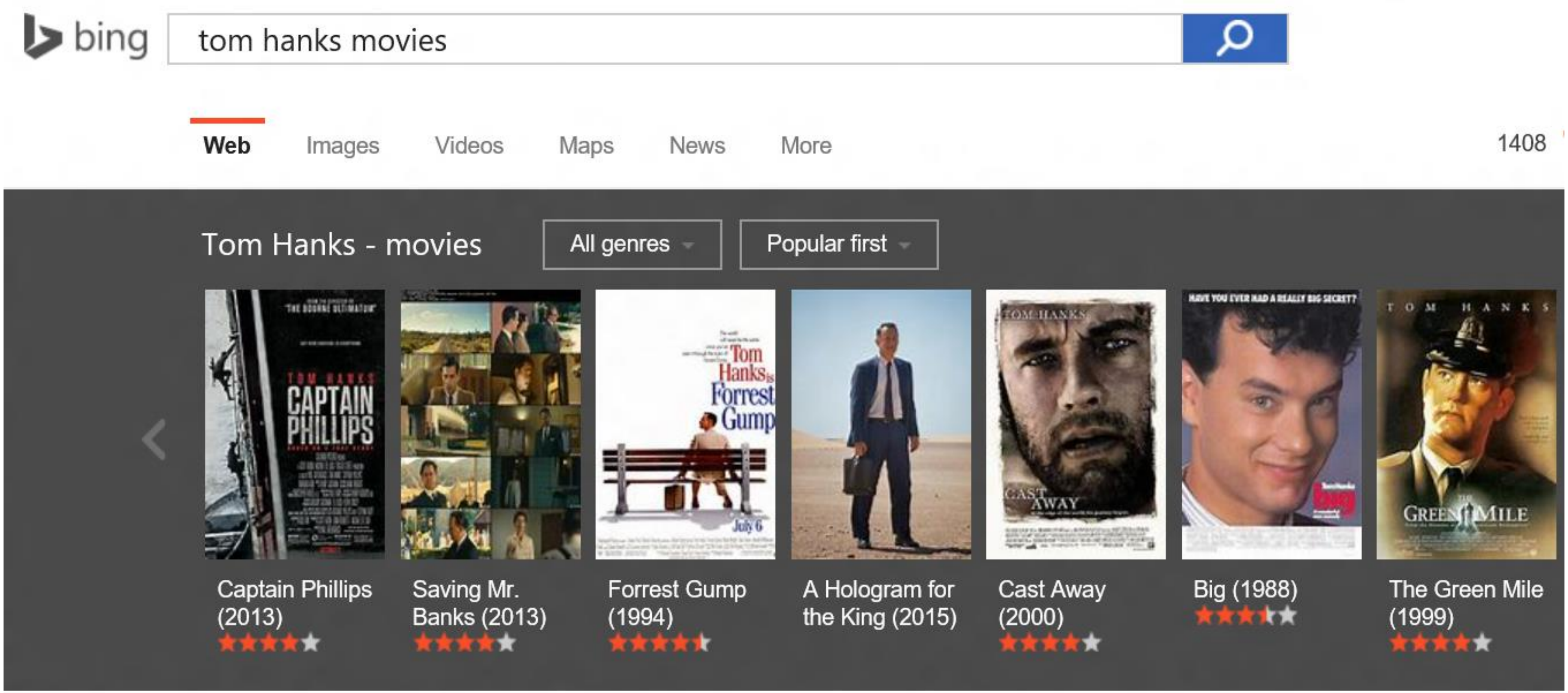
Menu

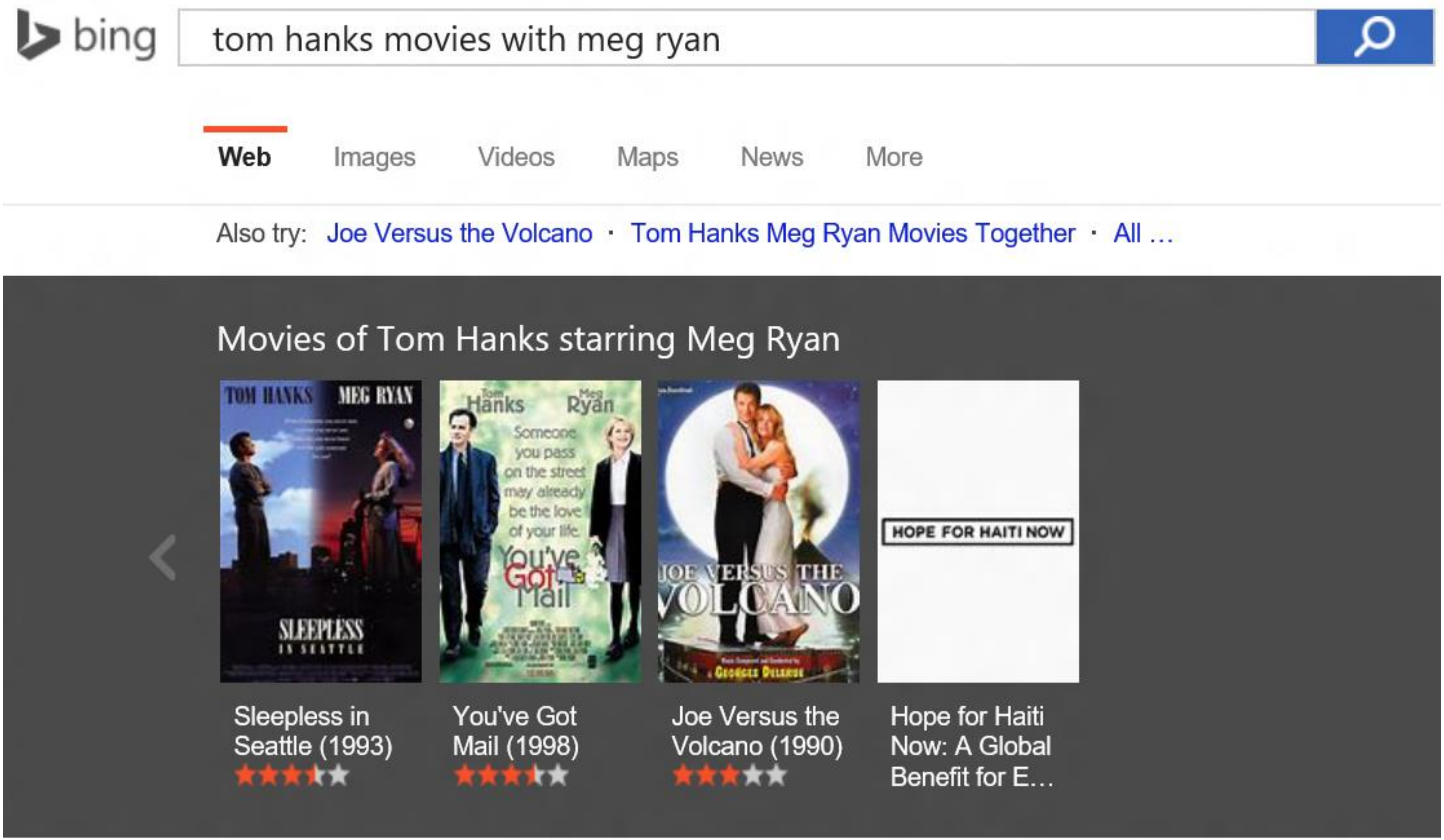
Book it

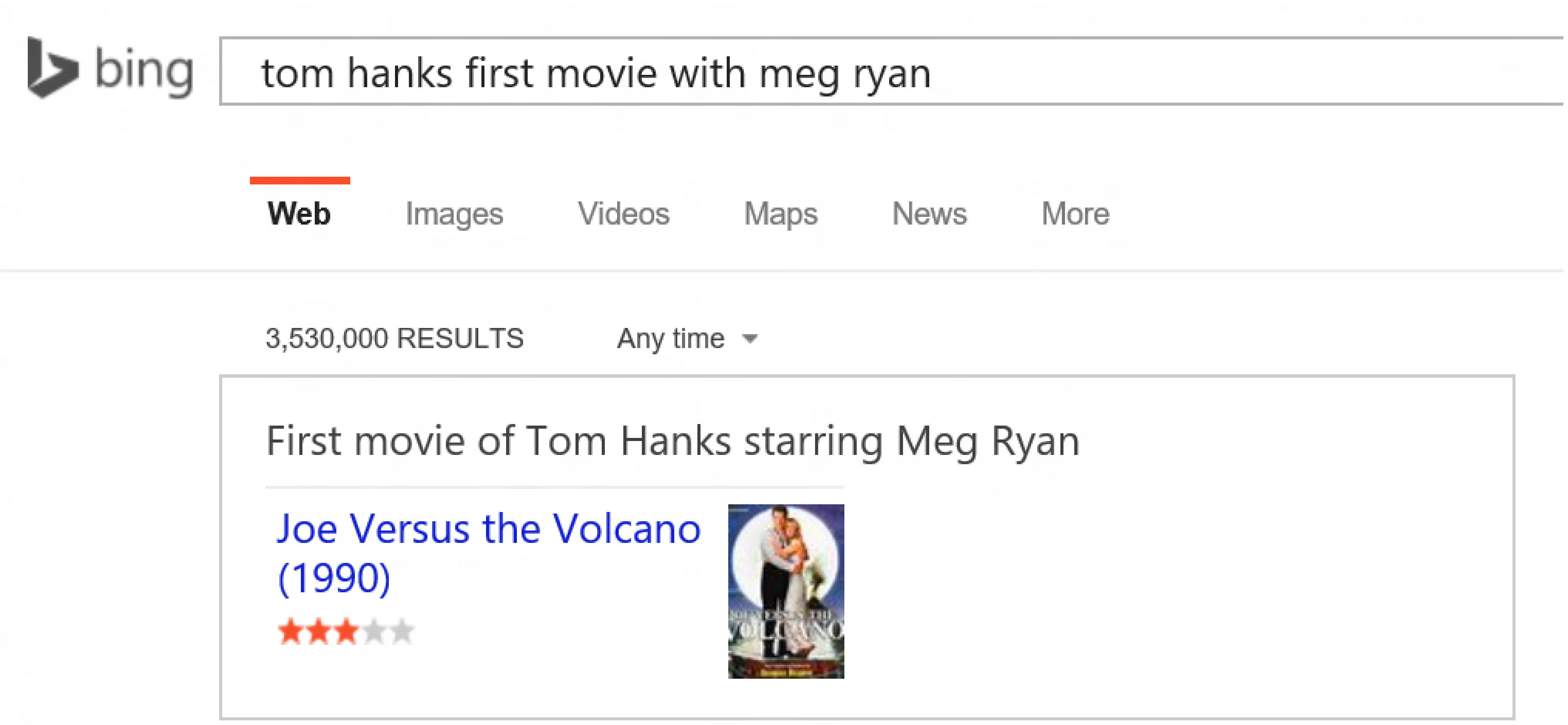

bing
hilton bellevue
Web
Images
Videos
Maps
News
More
1417
Sign in
2,280,000 RESULTS
Any time
Larger map
See all images
Hilton Bellevue
www3.hilton.com
300 112th Ave Se, Bellevue, WA 98004 · 6.51 mi
(425) 455-1300
596 TripAdvisor reviews
Directions
Book it
More about Hilton Bellevue
Reservations
Avg. rate $212
Check in
Check out
5/29/2015
5/31/2015
Find rate
Powered by TripAdvisor
Reviews
TripAdvisor (596)
tripadvisor
Bellevue Hotel - Hilton Bellevue - Hotels in Bellevue WA
www.bellevuehilton.com
Official Hilton Bellevue Hotel Website - Just minutes from Seattle and close to popular Bellevue businesses, we're one of the best hotels in Bellevue, WA.

Demonstration of Microsoft’s Entity Experience

Question Answering


bing
tom hanks movies
Web
Images
Videos
Maps
News
More
1408
Tom Hanks - movies
All genres
Popular first
Captain Phillips (2013)
Saving Mr. Banks (2013)
Forrest Gump (1994)
A Hologram for the King (2015)
Cast Away (2000)
Big (1988)
The Green Mile (1999)


bing
tom hanks movies with meg ryan
Web
Images
Videos
Maps
News
More
Also try: Joe Versus the Volcano · Tom Hanks Meg Ryan Movies Together · All …
Movies of Tom Hanks starring Meg Ryan
TOM HANKS MEG RYAN
SLEEPLESS IN SEATTLE
Hanks Ryan
Someone you pass on the street may already be the love of your life.
You've Got Mail
JOE VERSUS THE VOLCANO
HOPE FOR HAITI NOW
Sleepless in Seattle (1993)
You've Got Mail (1998)
Joe Versus the Volcano (1990)
Hope for Haiti Now: A Global Benefit for E…


bing
tom hanks first movie with meg ryan
Web
Images
Videos
Maps
News
More
3,530,000 RESULTS
Any time
First movie of Tom Hanks starring Meg Ryan
Joe Versus the Volcano (1990)

bing director of tom hanks first movie with meg ryan

Web Images Videos Maps News More

1,620,000 RESULTS Any time

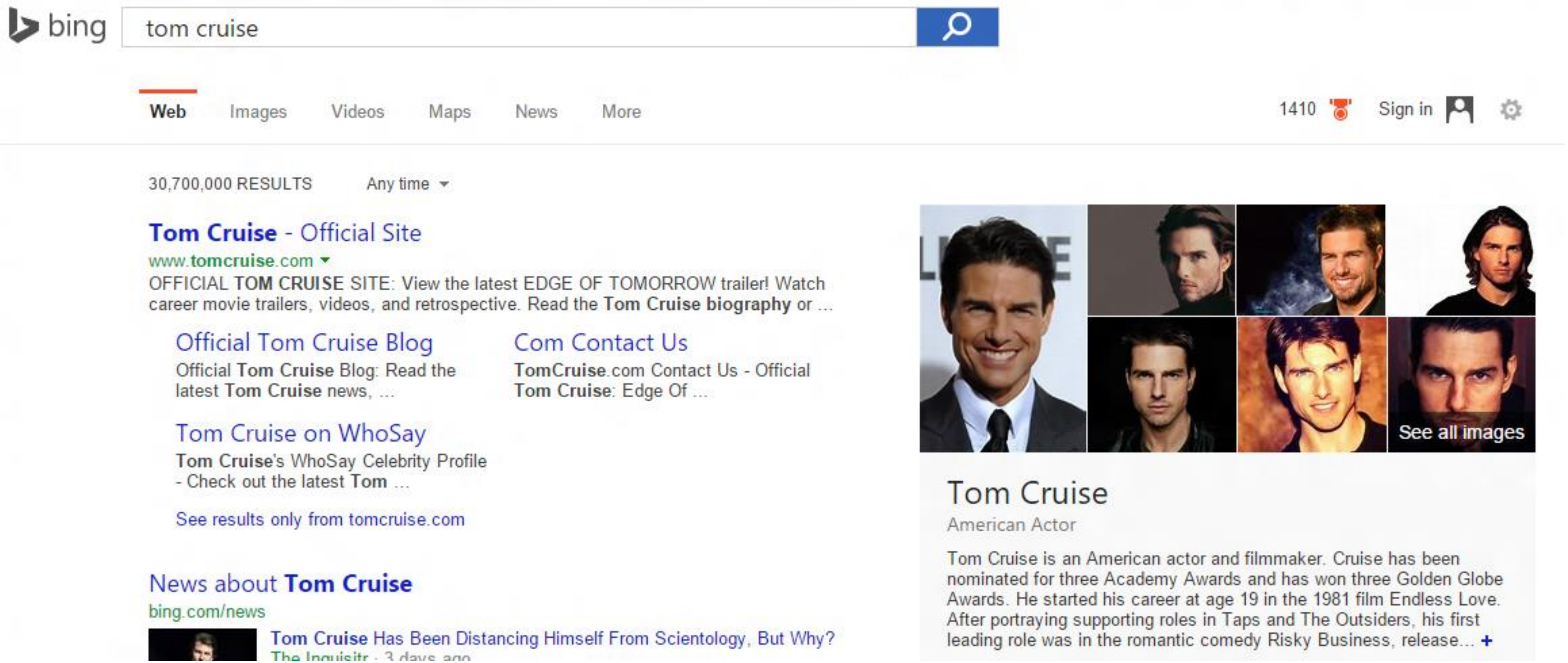
Director of first movie of Tom Hanks starring Meg Ryan

John Patrick Shanley

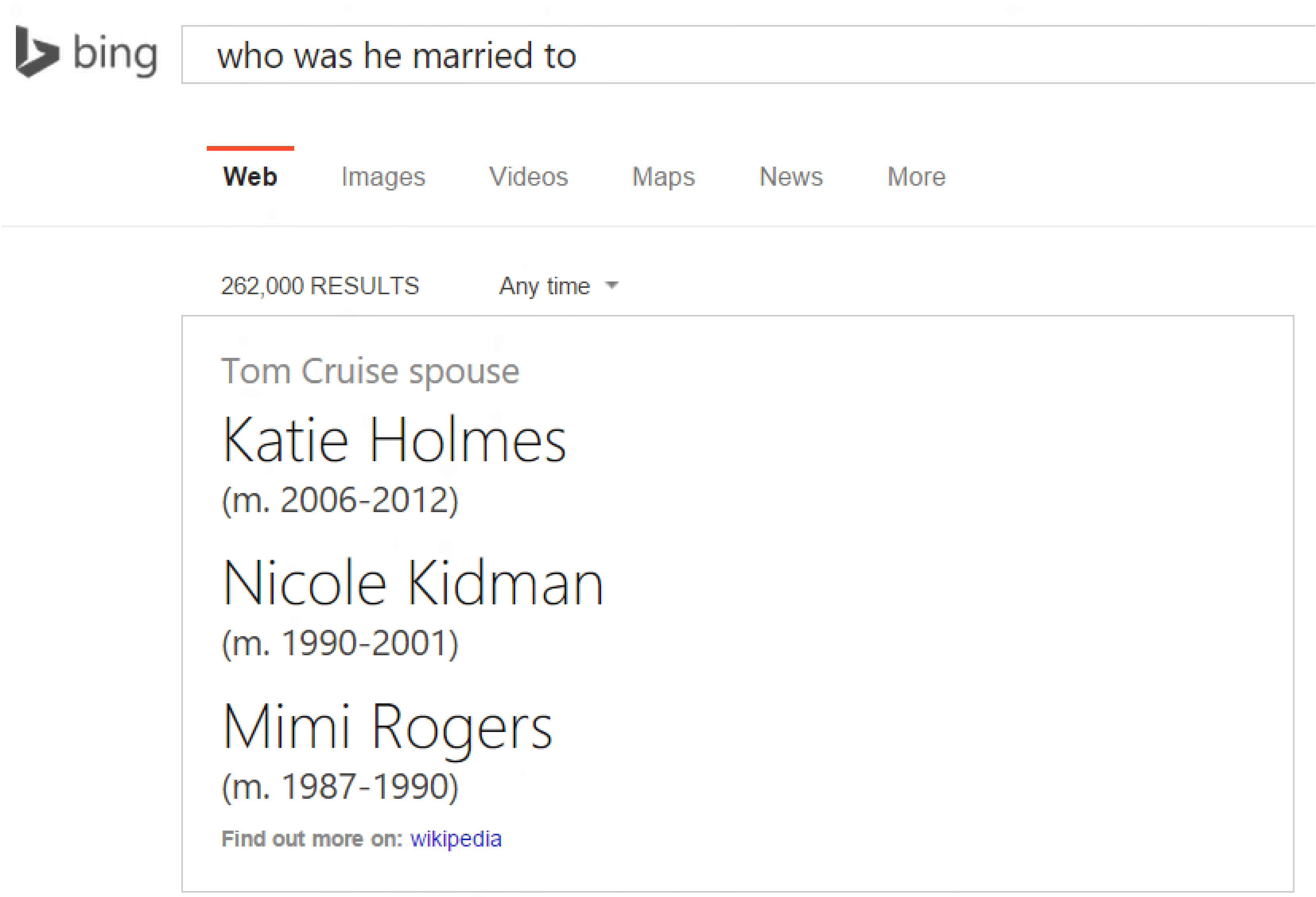
Joe Versus the Volcano (1990) - IMDb
www.imdb.com/title/tt0099892
★★★☆☆ Rating: 5.7/10 · 25,640 ratings · Comedy/Romance · PG · 102 min
Joe Versus the Volcano PG ... **Director**: **John Patrick Shanley**. Writer: **John Patrick Shanley**. Stars: **Tom Hanks**, **Meg Ryan**, Lloyd Bridges | See full cast and crew »

Meg Ryan Reteams With **Tom Hanks** for Ithaca , Actress Set ...
www.eonline.com/news/505216/**meg-ryan**-reteams-with-**tom-hanks**-for...
Jan 29, 2014 · **Meg Ryan** and **Tom Hanks** are teaming ... latest to step into the role of **director**. ... and was instrumental in making the **first film** such a ...

Conversational Question Answering



http://blogs.bing.com/search/2014/08/13/lets-have-a-conversation/


bing
who was he married to
Web
Images
Videos
Maps
News
More
262,000 RESULTS
Any time
Tom Cruise spouse
Katie Holmes
(m. 2006-2012)
Nicole Kidman
(m. 1990-2001)
Mimi Rogers
(m. 1987-1990)
Find out more on: wikipedia

how tall are they

Web Images Videos Maps News More

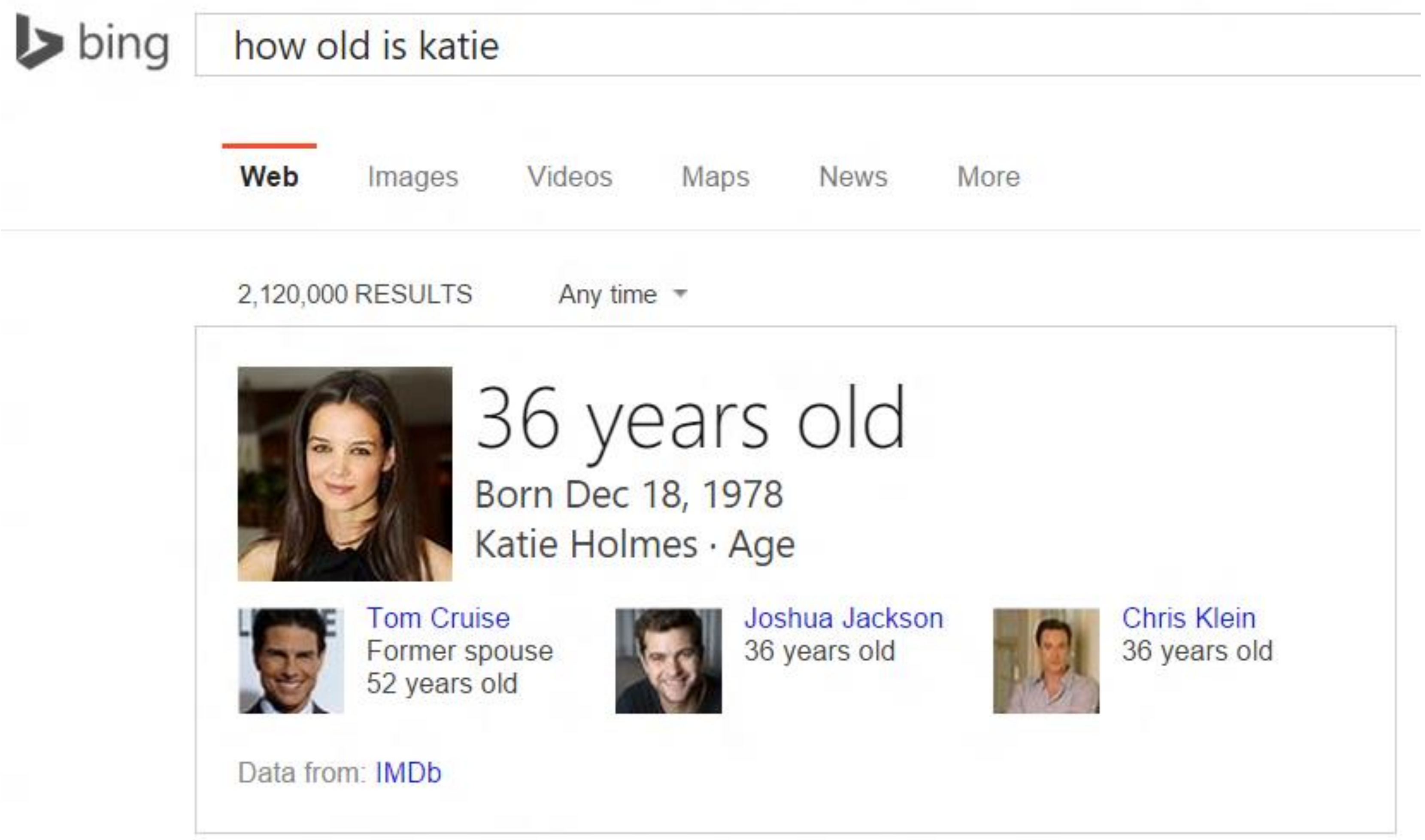
87,800 RESULTS Any time

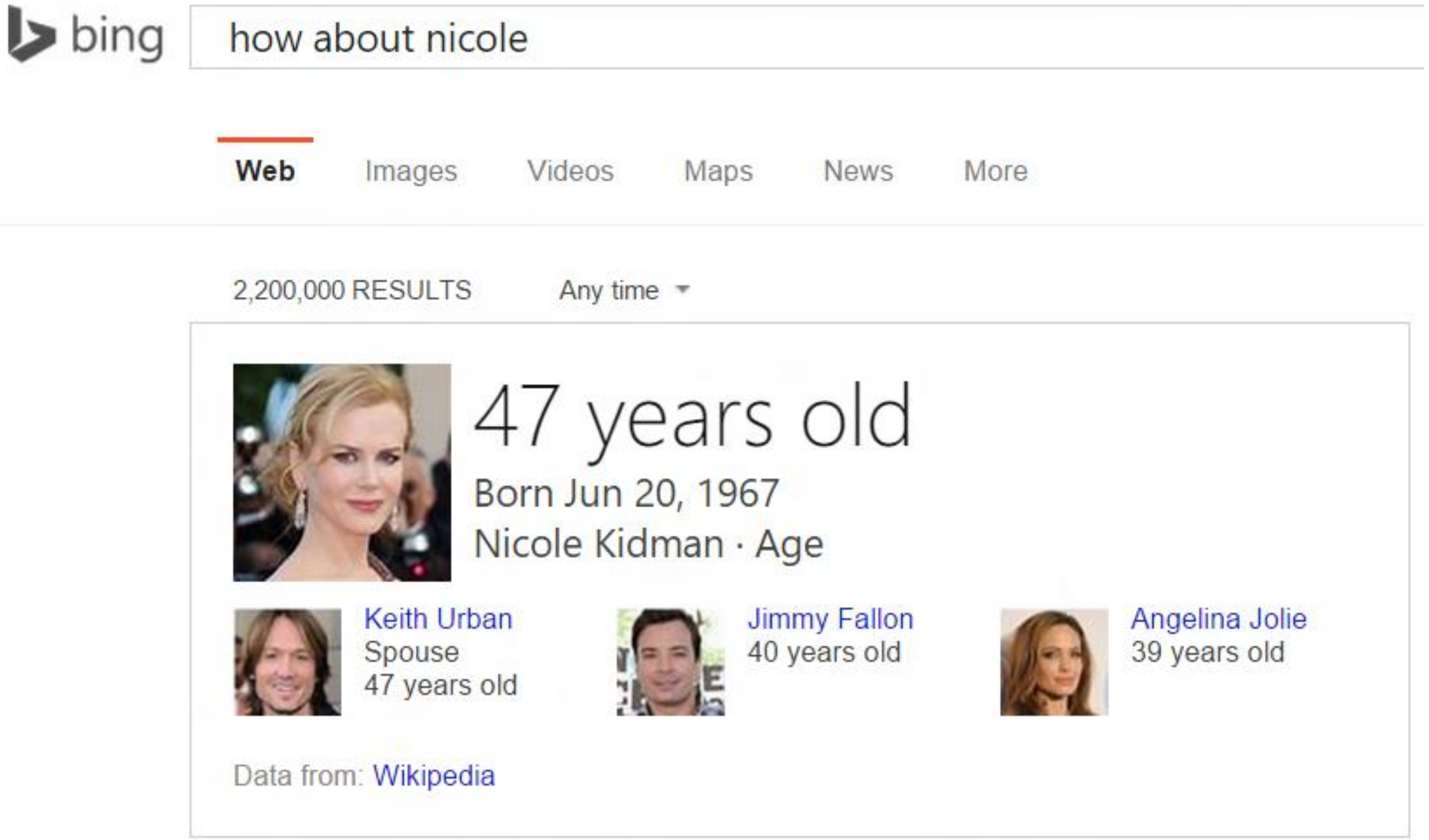
Katie Holmes height: 5 feet 9 inches (1.75 meters)

Nicole Kidman height: 5 feet 11 inches (1.80 meters)

Mimi Rogers height: 5 feet 9 inches (1.74 meters)

Find out more on: IMDb


bing
how old is katie
Web
Images
Videos
Maps
News
More
2,120,000 RESULTS
Any time
36 years old
Born Dec 18, 1978
Katie Holmes · Age
Tom Cruise
Former spouse
52 years old
Joshua Jackson
36 years old
Chris Klein
36 years old
Data from: IMDb


bing
how about nicole
Web
Images
Videos
Maps
News
More
2,200,000 RESULTS
Any time
47 years old
Born Jun 20, 1967
Nicole Kidman · Age
Keith Urban
Spouse
47 years old
Jimmy Fallon
40 years old
Angelina Jolie
39 years old
Data from: Wikipedia

bing where was she born

Web Images Videos Maps News More

544,000 RESULTS Any time

Honolulu, HI

Nicole Kidman · Birthplace

Data from: Wikipedia

Nicole Kidman - Biography - IMDb

www.imdb.com/name/nm0000173/bio

48 years old · News

Elegant redhead **Nicole Kidman**, known as one of Hollywood's top Australian imports, was actually **born** in **Honolulu, Hawaii**. Kidman is the daughter of ...

Awards · Films

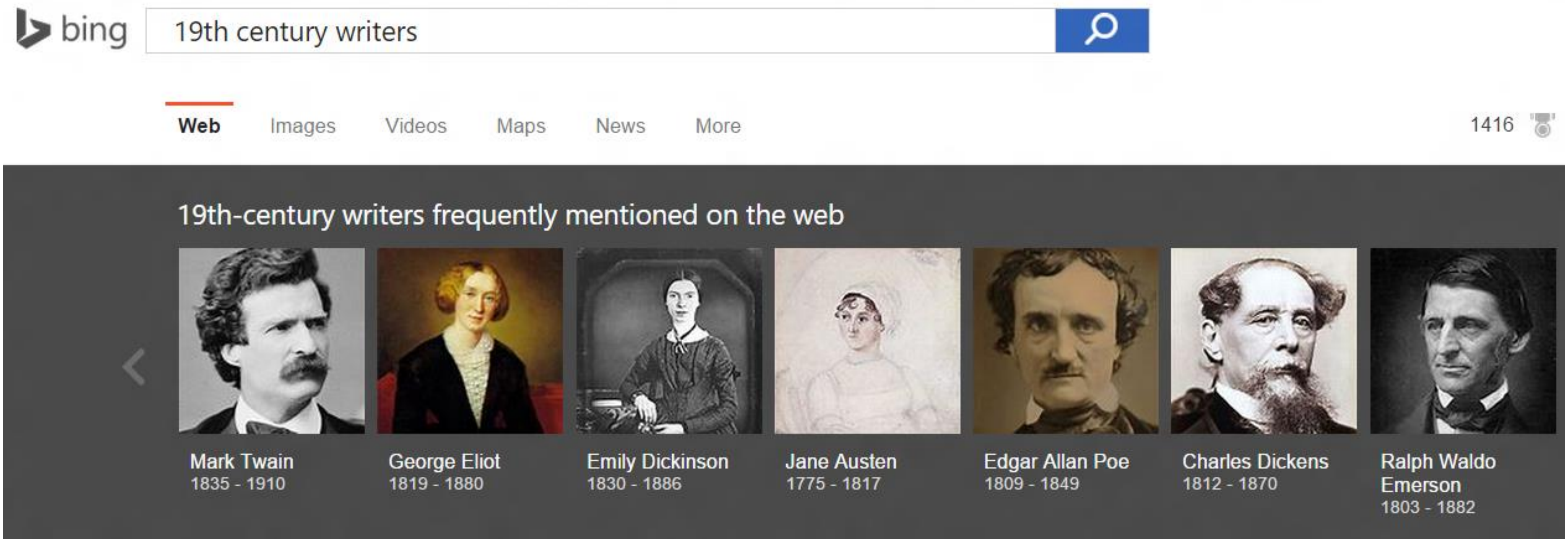
Nicole Kidman - Biography - Film Actress - **Biography.com**

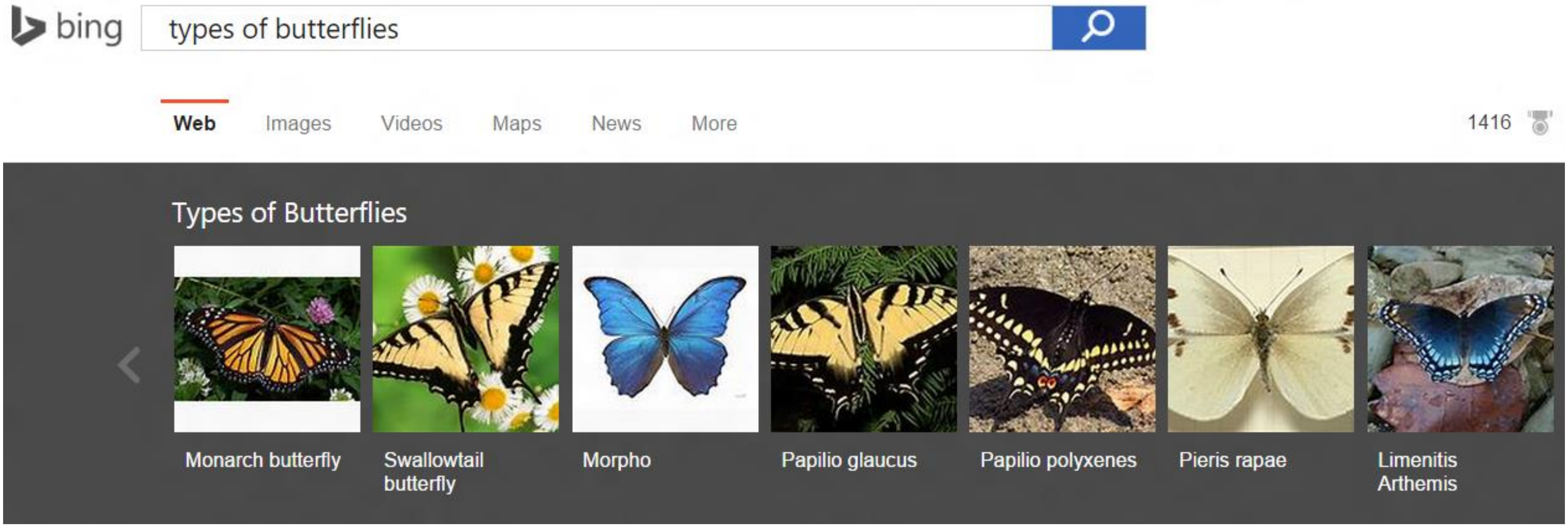
www.biography.com/people/nicole-kidman-9364474

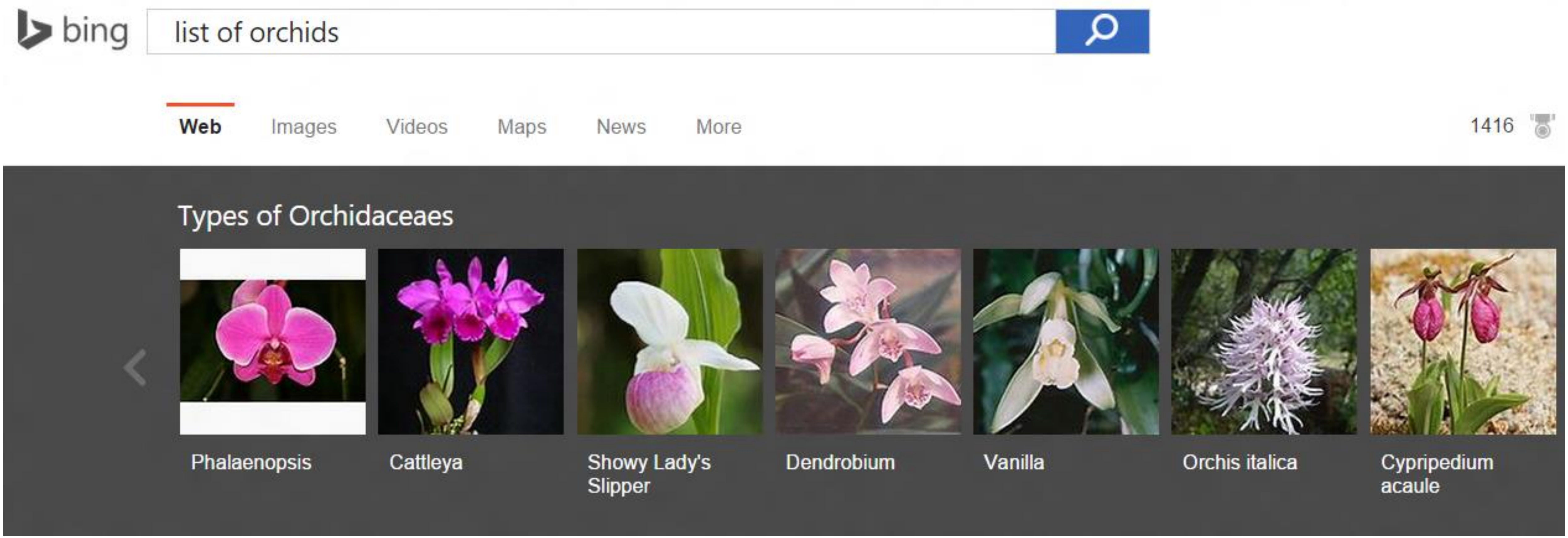
Meet Academy Award-winning actress **Nicole Kidman** at **Biography.com**. **She** is best known for movies like Moulin Rouge and her ten-year marriage to Tom Cruise.

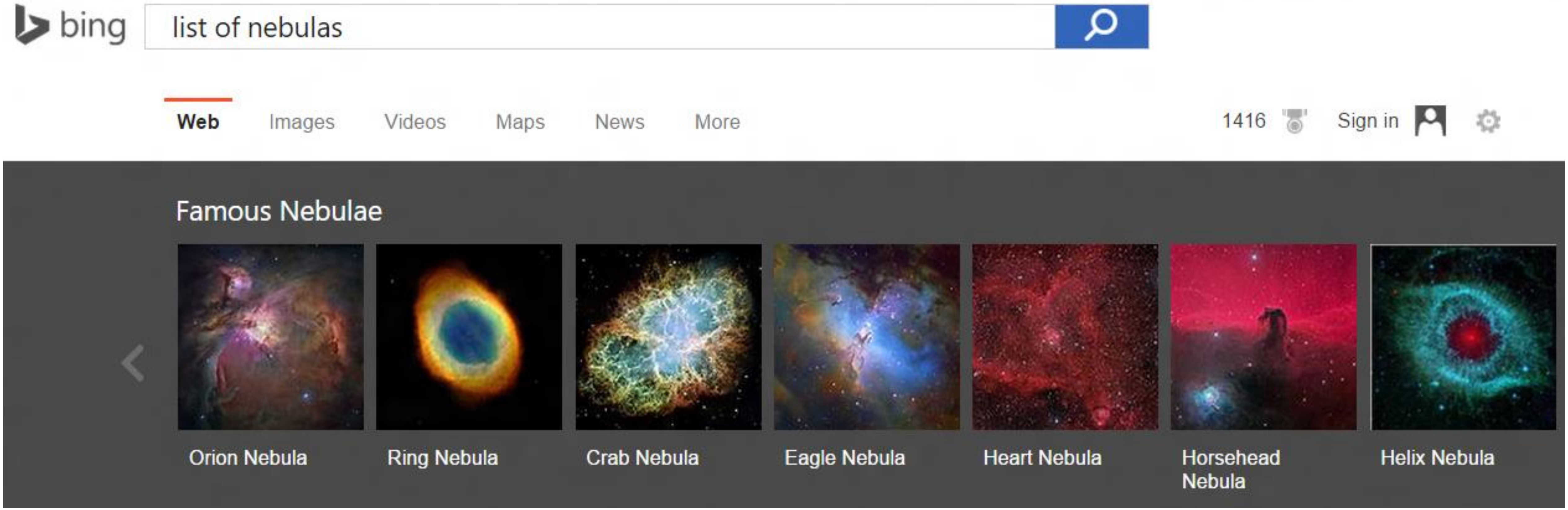
Demonstration of Microsoft's Entity Experience

Diversity of entity collections

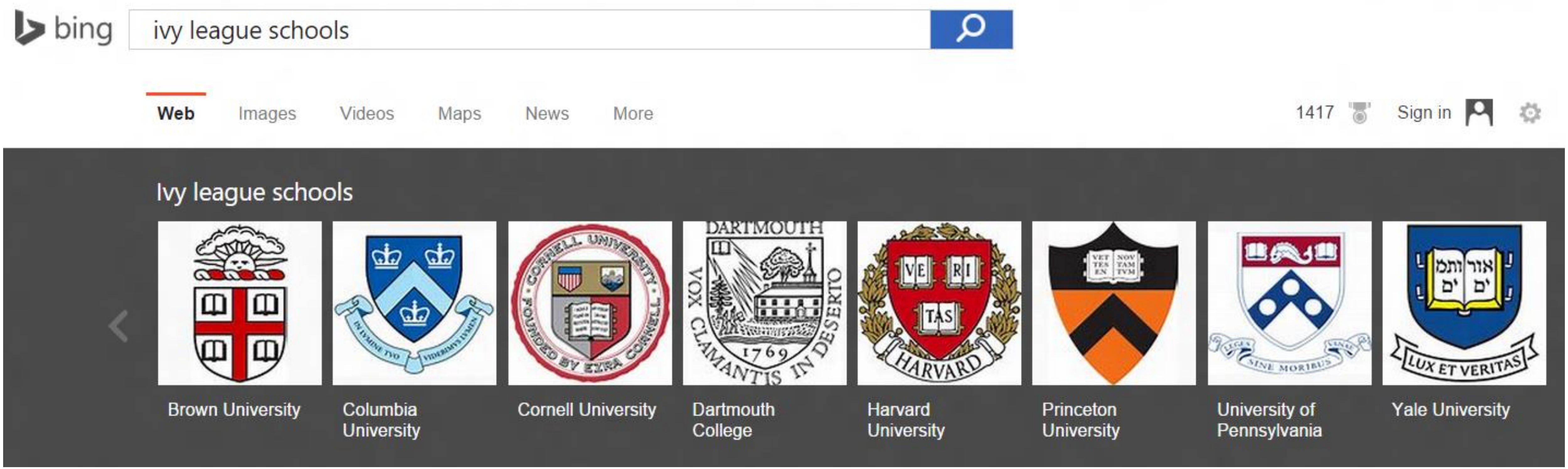

bing
19th century writers
Web
Images
Videos
Maps
News
More
1416
19th-century writers frequently mentioned on the web
Mark Twain
1835 - 1910
George Eliot
1819 - 1880
Emily Dickinson
1830 - 1886
Jane Austen
1775 - 1817
Edgar Allan Poe
1809 - 1849
Charles Dickens
1812 - 1870
Ralph Waldo Emerson
1803 - 1882


bing
types of butterflies
Web
Images
Videos
Maps
News
More
1416
Types of Butterflies
Monarch butterfly
Swallowtail butterfly
Morpho
Papilio glaucus
Papilio polyxenes
Pieris rapae
Limenitis Arthemis

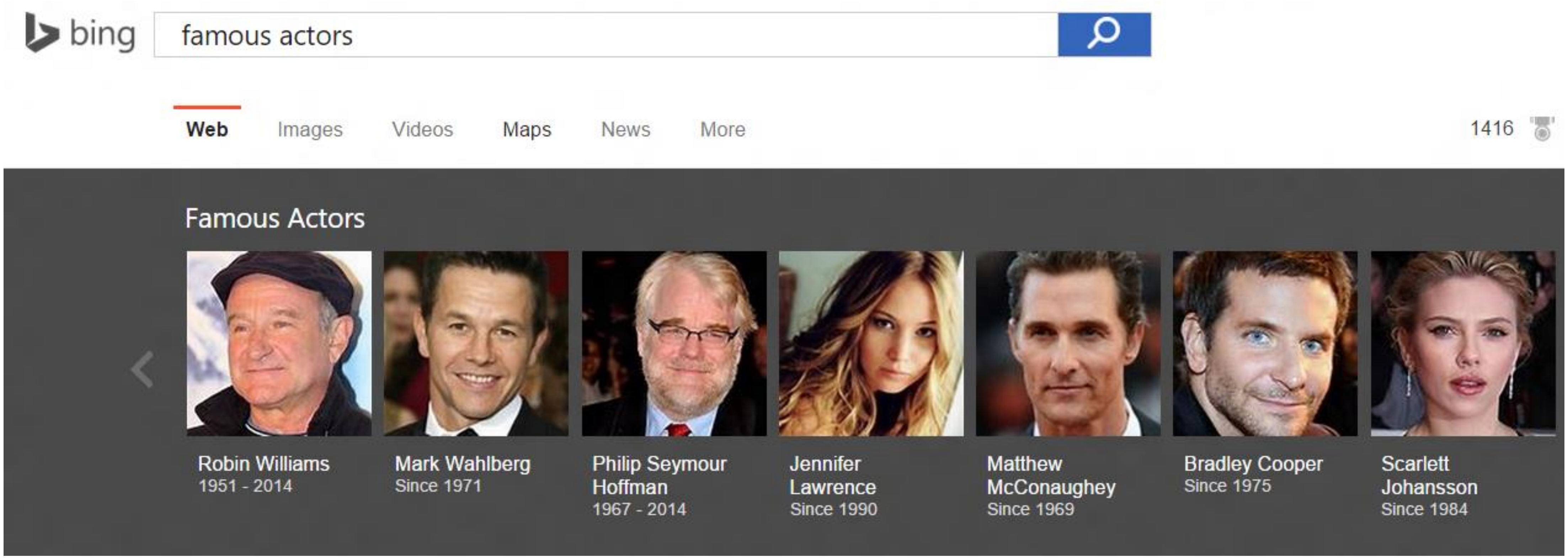

bing
list of orchids
Web
Images
Videos
Maps
News
More
1416
Types of Orchidaceaes
Phalaenopsis
Cattleya
Showy Lady's Slipper
Dendrobium
Vanilla
Orchis italica
Cypripedium acaule


bing
list of nebulas
Web
Images
Videos
Maps
News
More
1416
Sign in
Famous Nebulae
Orion Nebula
Ring Nebula
Crab Nebula
Eagle Nebula
Heart Nebula
Horsehead Nebula
Helix Nebula


bing
republican united states presidents
Web
Images
Videos
Maps
News
More
1416
Republican Party United States President
Abraham Lincoln
1861 - 1865
Theodore Roosevelt
1901 - 1909
George W. Bush
2001 - 2009
Dwight D. Eisenhower
1953 - 1961
Richard Nixon
1969 - 1974
George H. W. Bush
1989 - 1993
Ulysses S. Grant
1869 - 1877


bing
ivy league schools
Web
Images
Videos
Maps
News
More
1417
Sign in
Ivy league schools
DARTMOUTH
VOX CLAMANTIS IN DESERTO
1769
VE RI TAS
HARVARD
LUX ET VERITAS
Brown University
Columbia University
Cornell University
Dartmouth College
Harvard University
Princeton University
University of Pennsylvania
Yale University

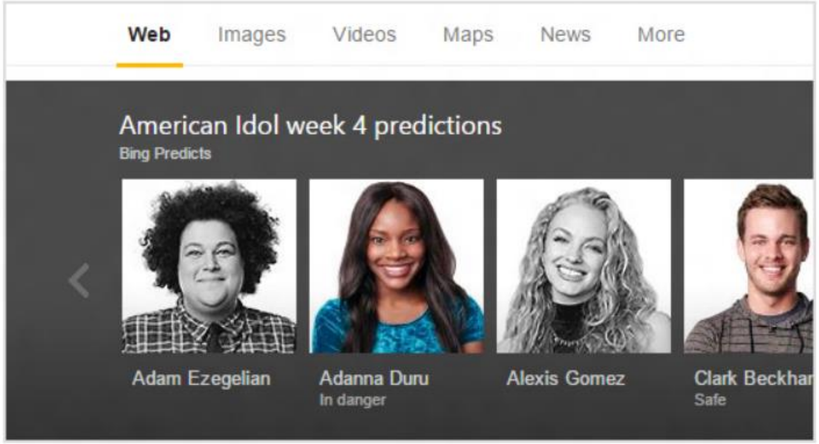
Segment-specific entity rankings



Demonstration of Microsoft's Entity Experience

Entities in the platform

Bing Predicts



American Idol

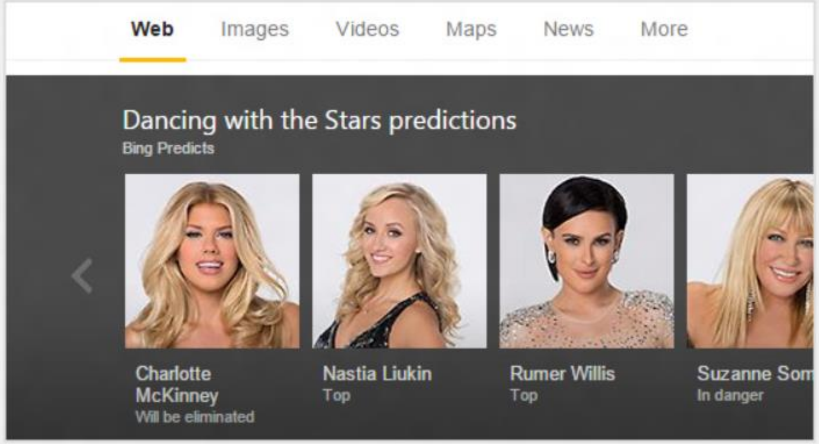
Search for "American Idol predictions" to find out who Bing predicts will be eliminated and who'll be safe.

Try it >

Dancing with the Stars

Search for "Dancing with the Stars predictions" to see who Bing predicts will make it to the next round and who won't.

Try it >



https://www.bing.com/explore/predicts

Bing Predicts

March Madness

Prediction accuracy: 73 percent

American Idol

Prediction accuracy: 90 percent

Dancing with the Stars

Prediction accuracy: 95 percent

The Voice

Prediction accuracy: 85 percent

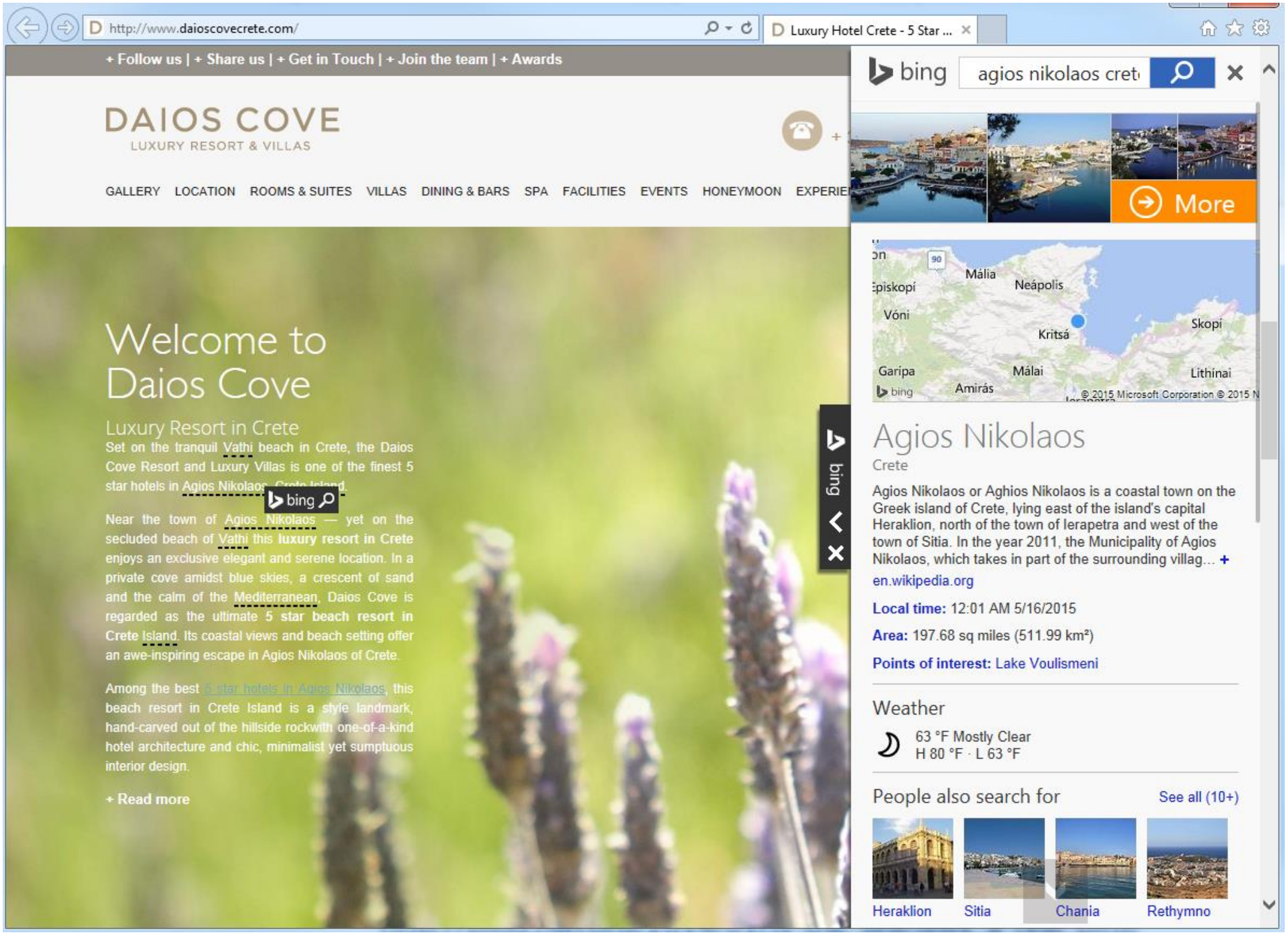
Academy Awards

Prediction accuracy: 84 percent

Golden Globe Awards

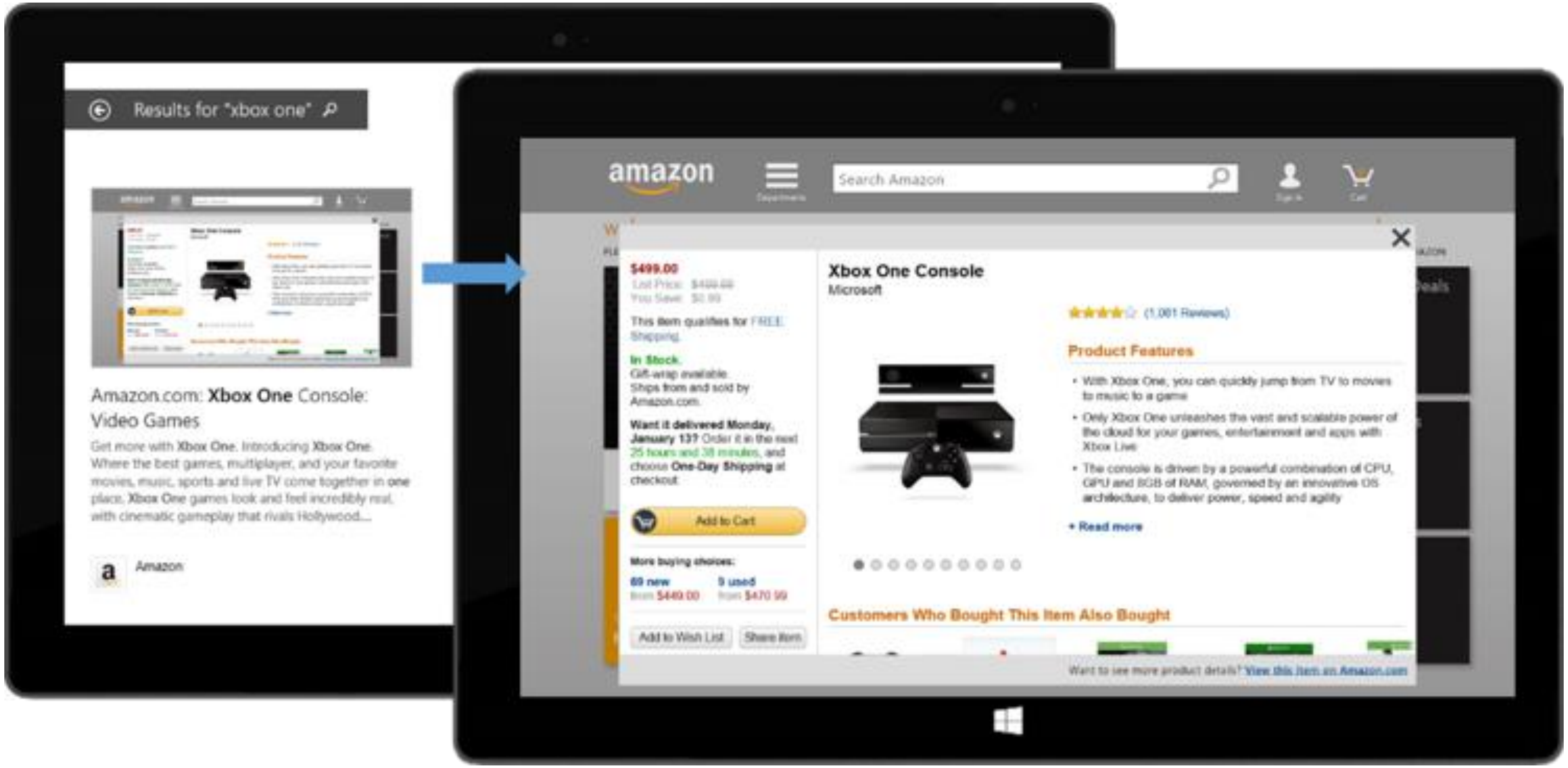
Prediction accuracy: 83 percent

Bing Widget API



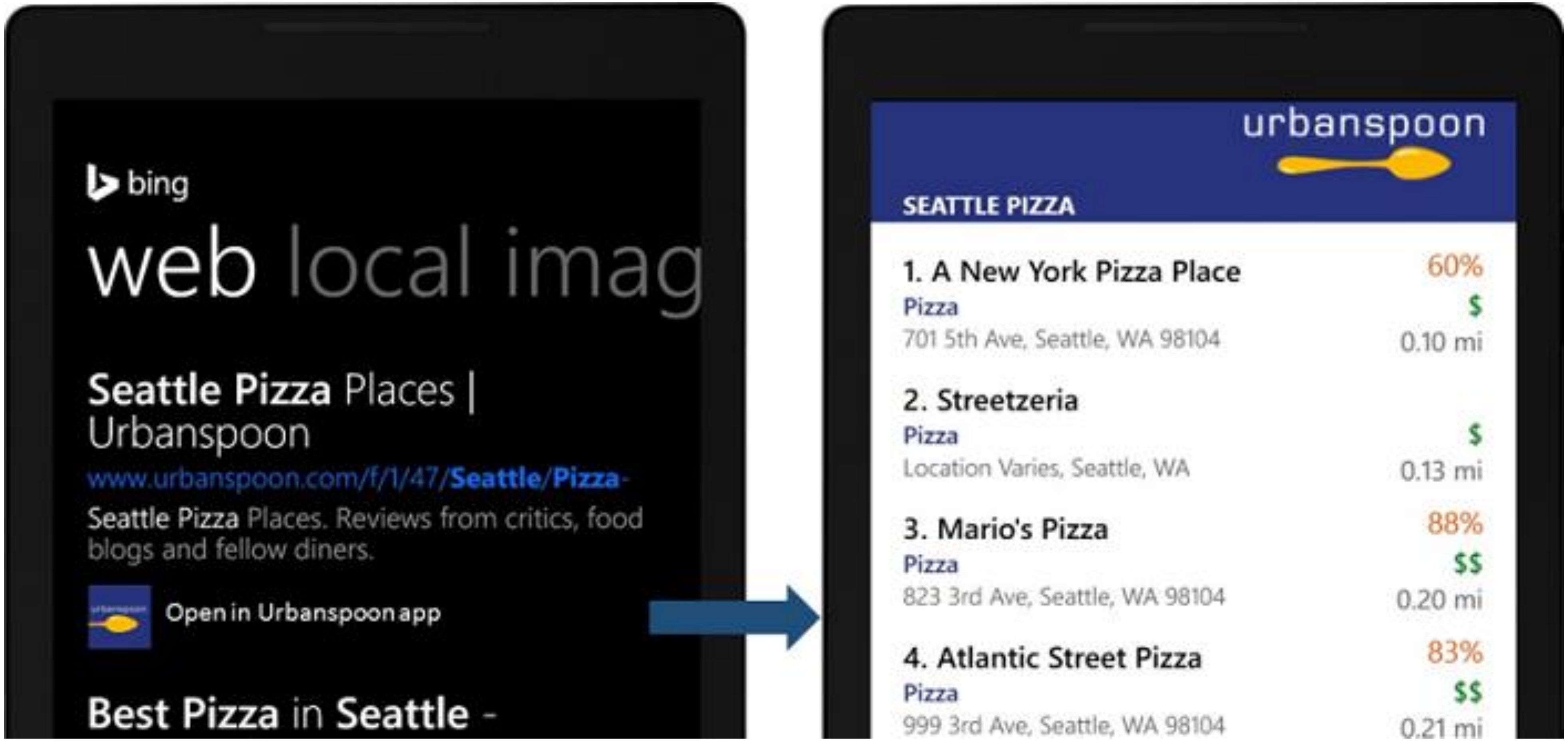
http://blogs.bing.com/webmaster/2014/01/03/bringing-the-power-of-bing-knowledge-to-webmasters/
http://blogs.bing.com/webmaster/2014/06/19/bing-knowledge-comes-to-webmaster-tools/
http://blogs.bing.com/webmaster/2014/05/15/mark-it-up/

App Linking

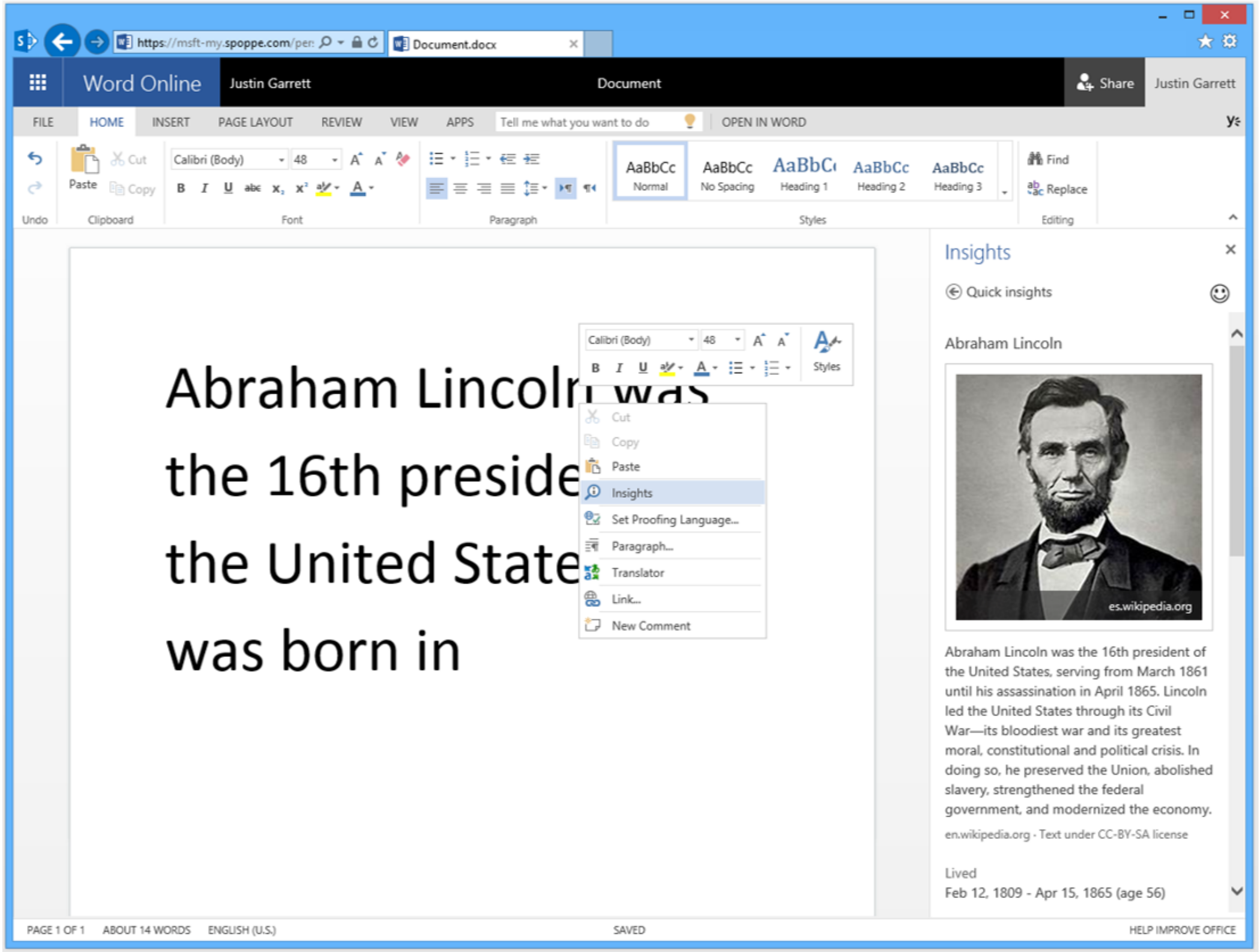


http://blogs.bing.com/webmaster/2014/04/02/announcing-bing-knowledge-widget-app-linking/
http://www.bing.com/webmaster/help/app-linking-09399b4b

App Linking

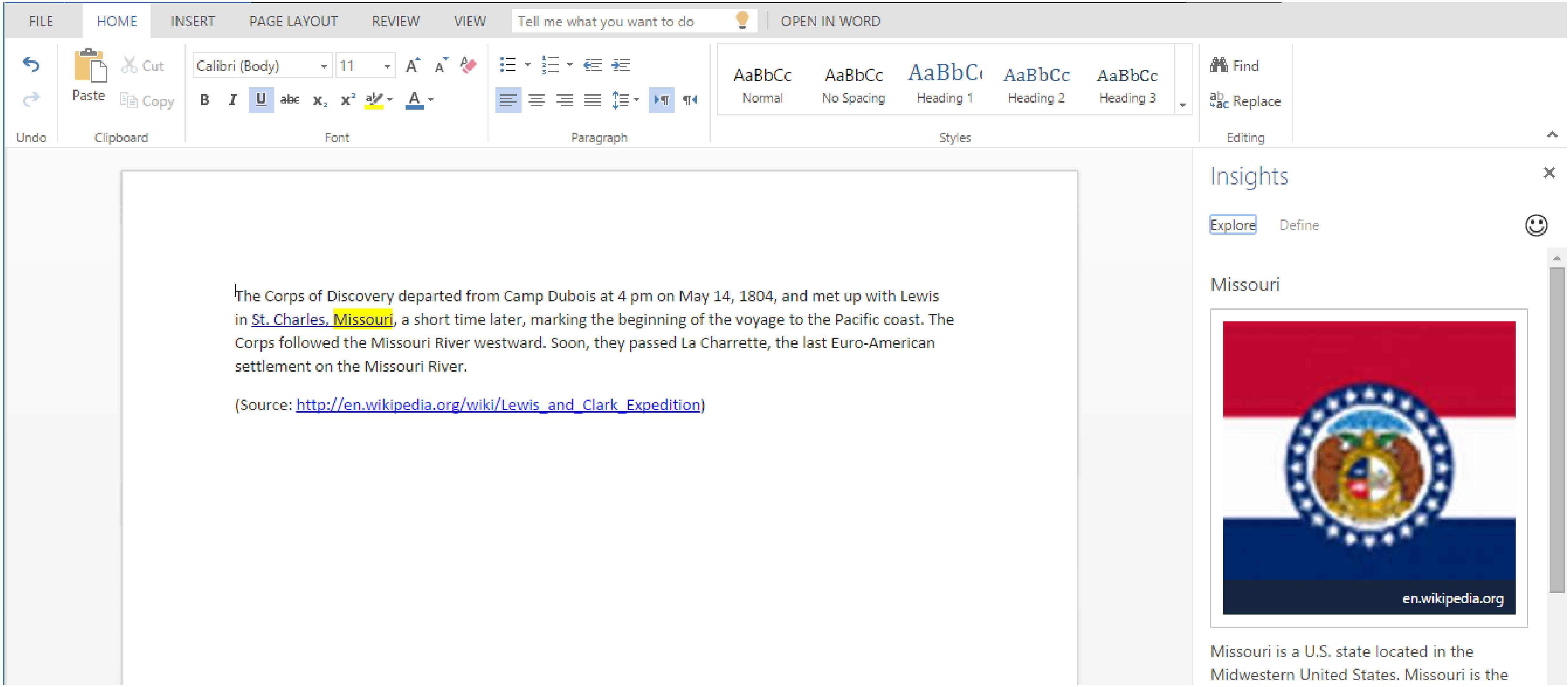

bing
web local imag
Seattle Pizza Places | Urbanspoon
www.urbanspoon.com/f/1/47/Seattle/Pizza-
Seattle Pizza Places. Reviews from critics, food blogs and fellow diners.
Open in Urbanspoon app
Best Pizza in Seattle -
urbanspoon
SEATTLE PIZZA
1. A New York Pizza Place
Pizza
701 5th Ave, Seattle, WA 98104
60%
$
0.10 mi
2. Streetzeria
Pizza
Location Varies, Seattle, WA
$
0.13 mi
3. Mario's Pizza
Pizza
823 3rd Ave, Seattle, WA 98104
88%
$$
0.20 mi
4. Atlantic Street Pizza
Pizza
999 3rd Ave, Seattle, WA 98104
83%
$$
0.21 mi

Office Insights

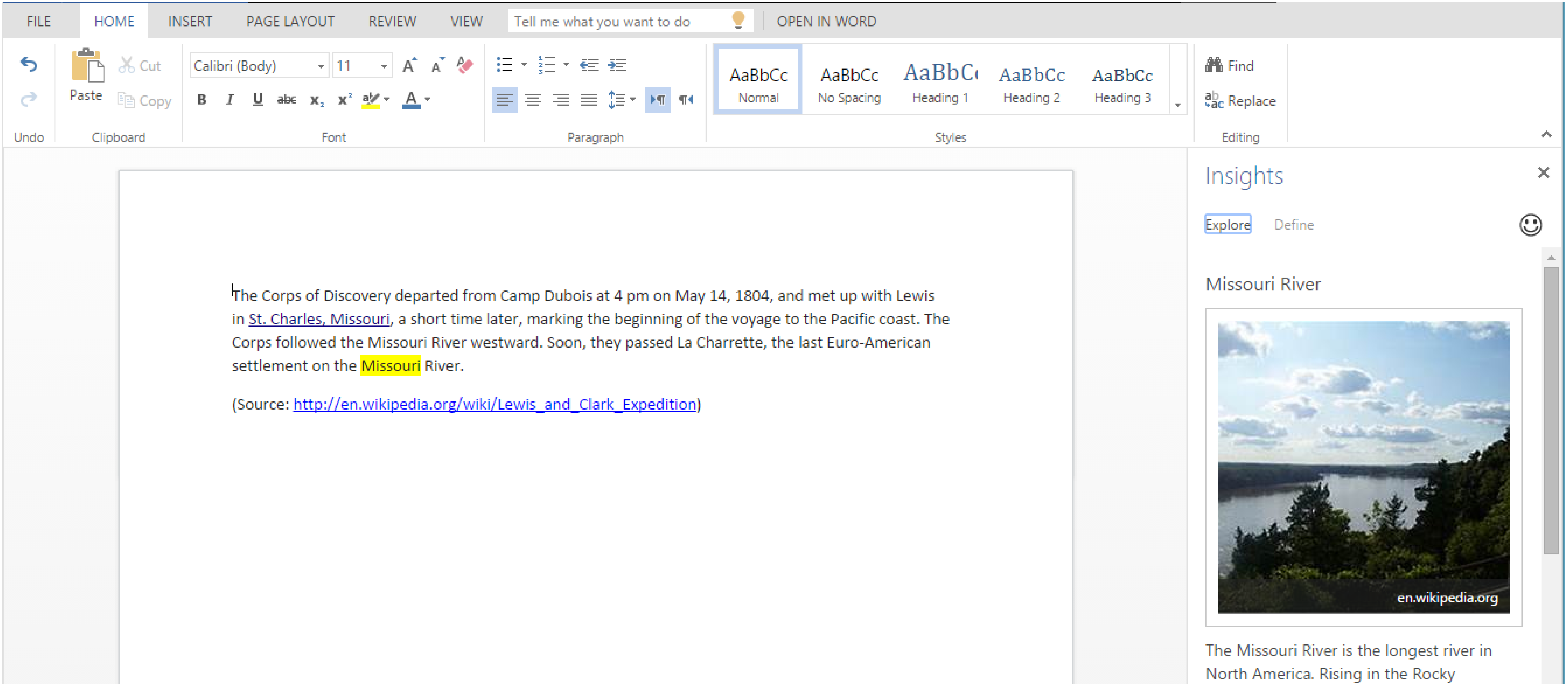


http://blogs.bing.com/search/2014/12/10/bing-brings-the-worlds-knowledge-straight-to-you-with-insights-for-office/
http://blogs.office.com/2014/12/11/first-look-insights-office-word-online-bringing-knowledge-web-document/
http://blogs.office.com/2014/12/10/whats-new-office-online/

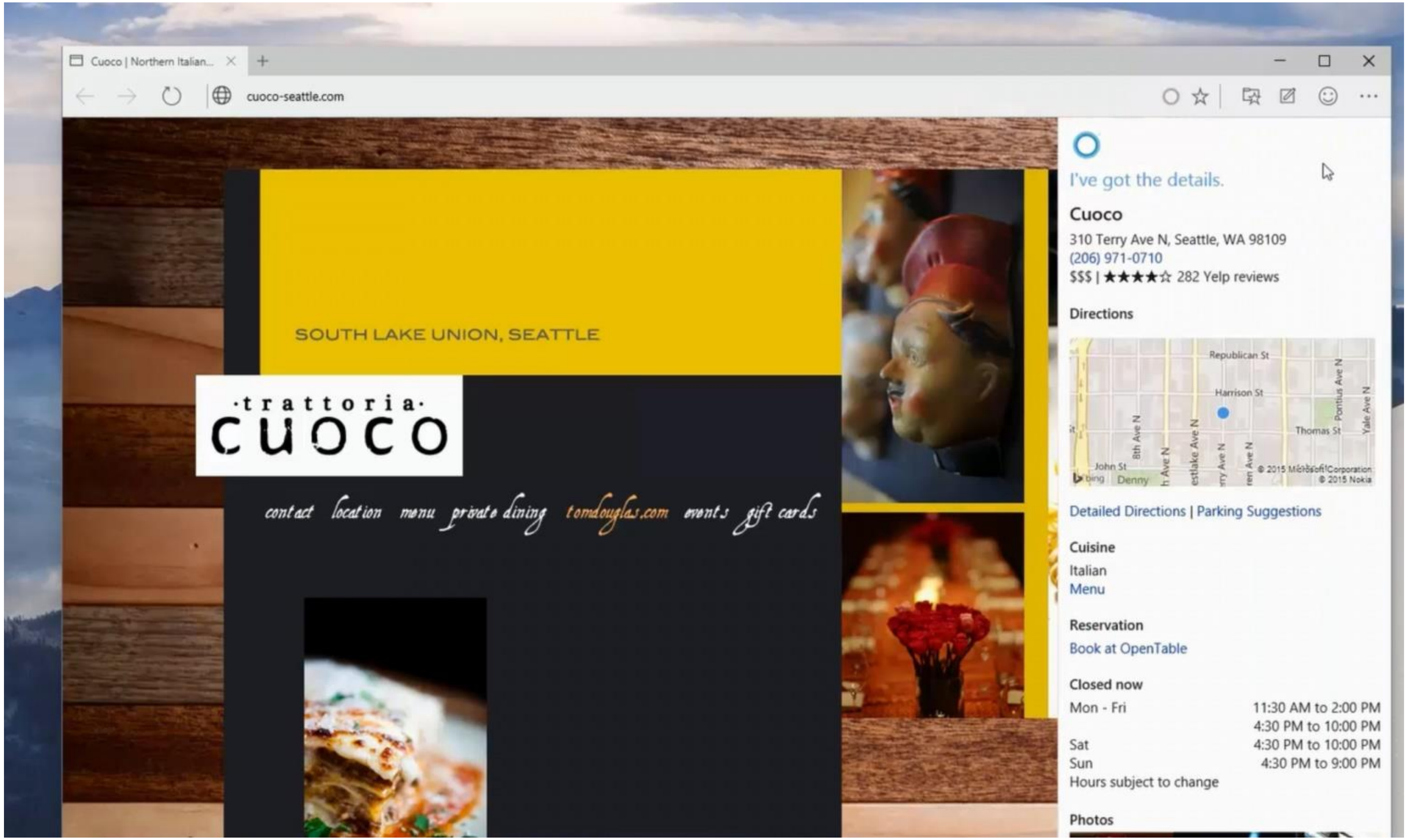
Office Insights



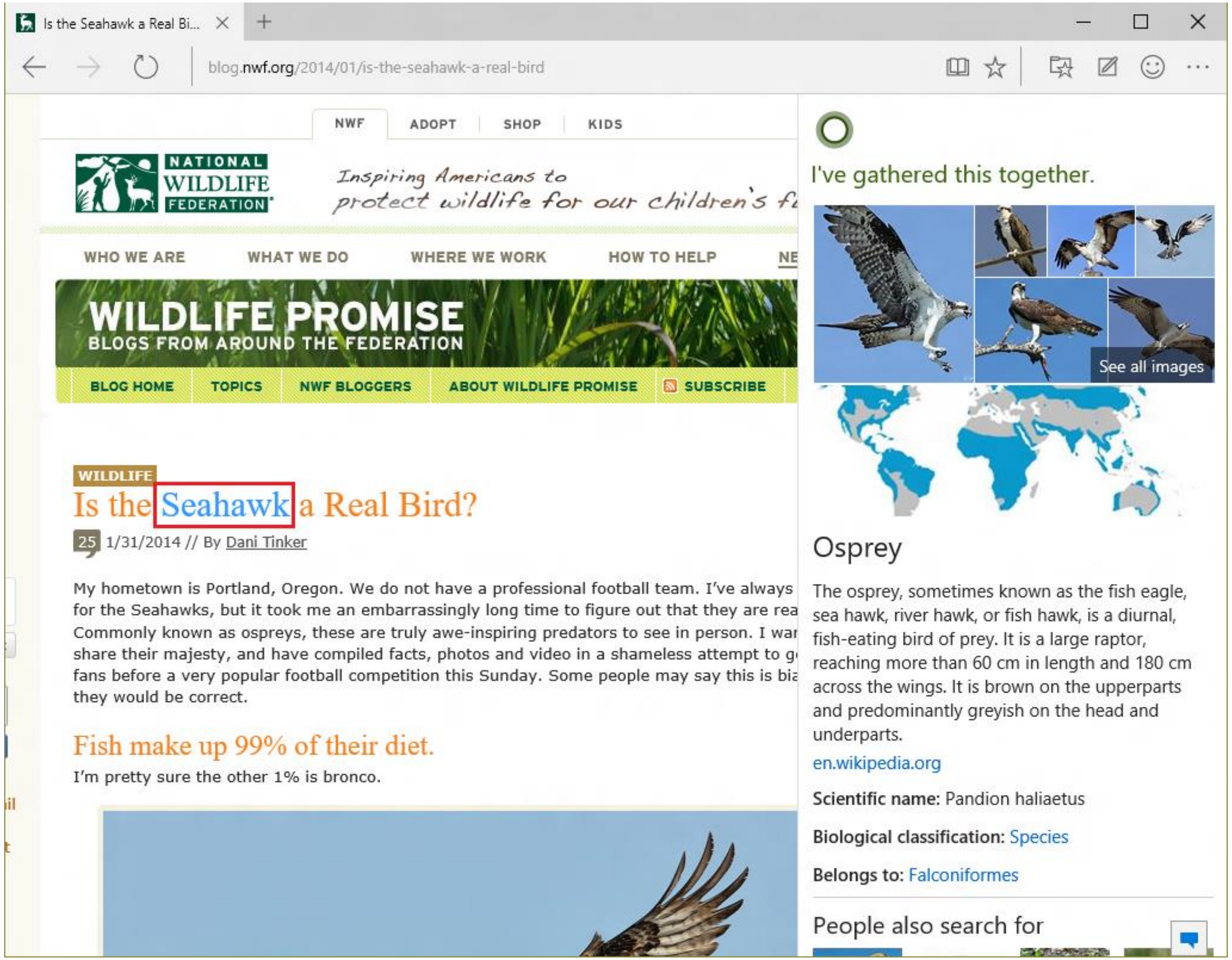
Office Insights



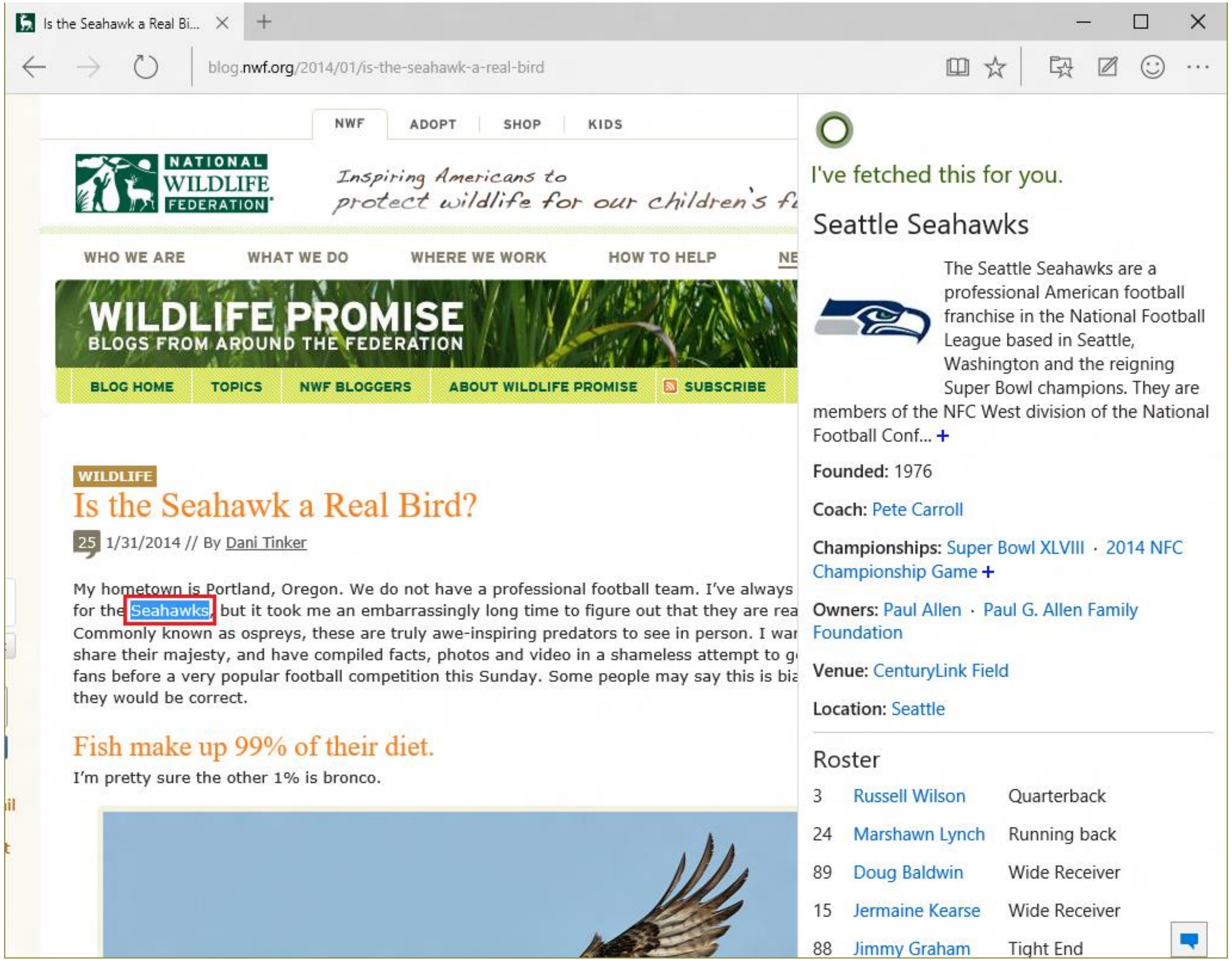
Edge Browser



http://windows.microsoft.com/en-us/windows/preview-microsoft-edge-pc

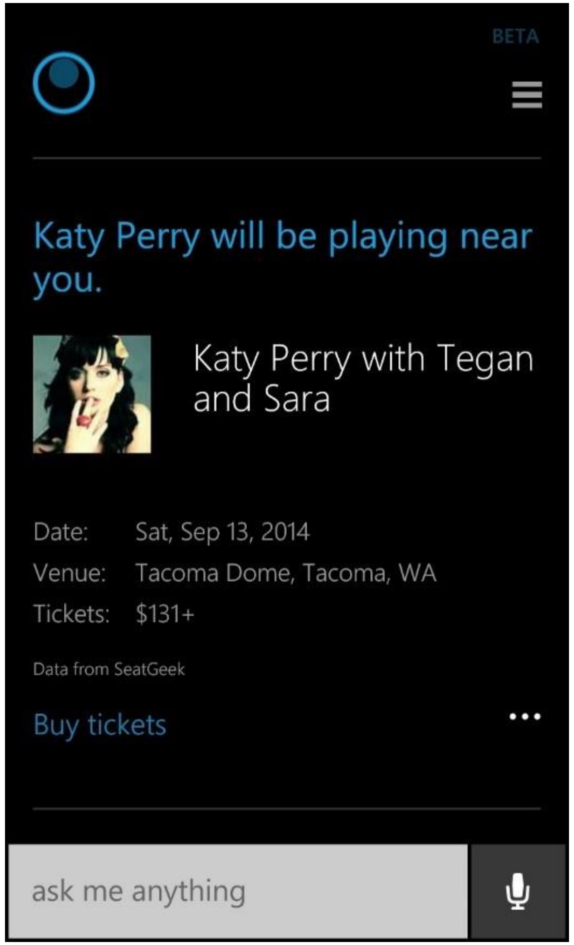


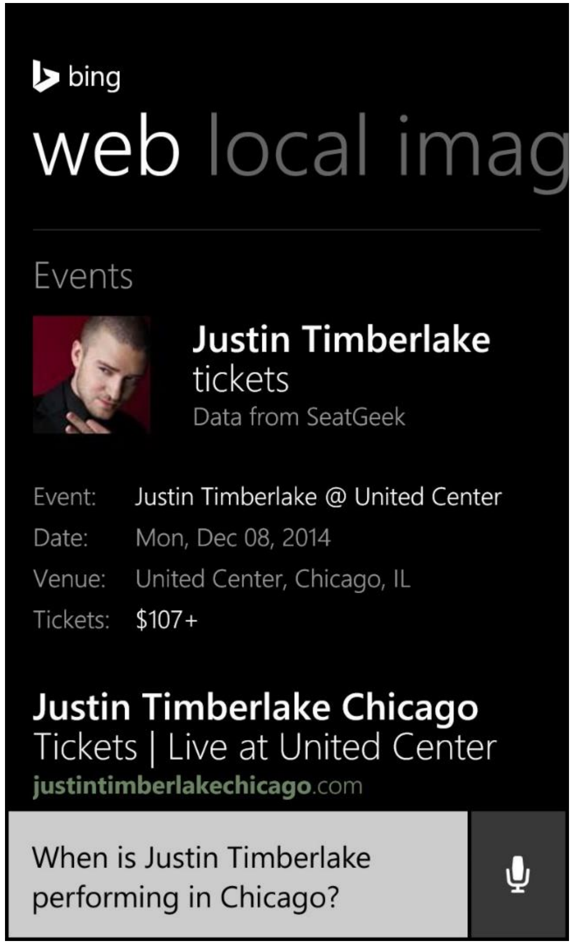
http://blog.nwf.org/2014/01/is-the-seahawk-a-real-bird/

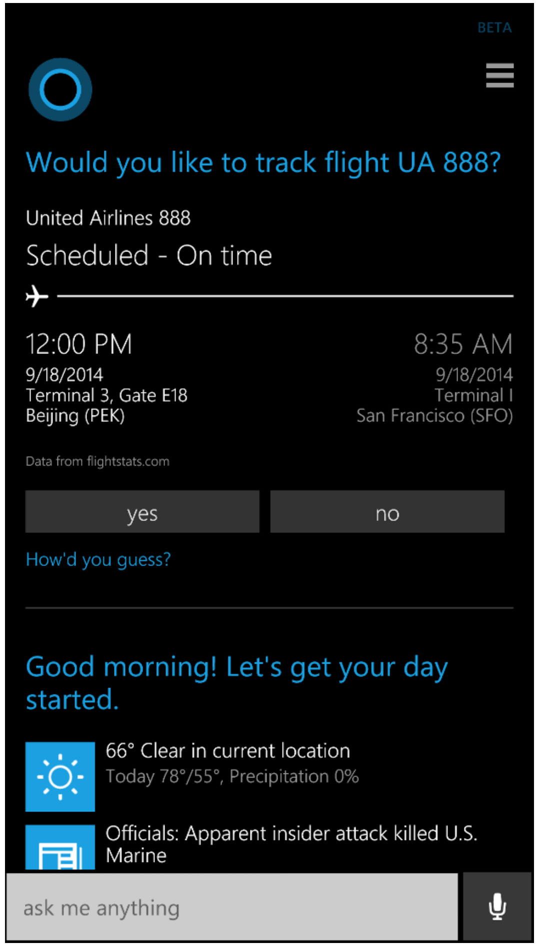


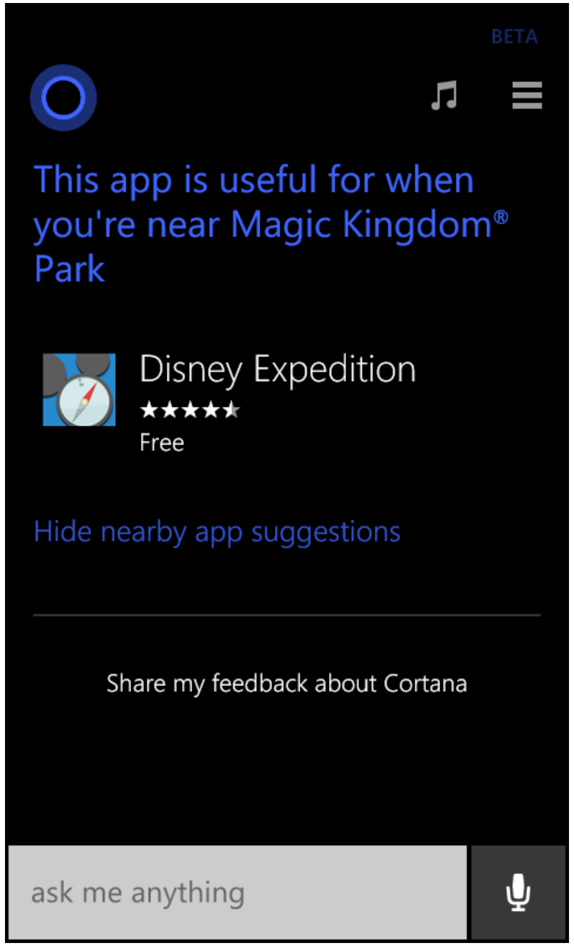
http://blog.nwf.org/2014/01/is-the-seahawk-a-real-bird/

Phone / Cortana







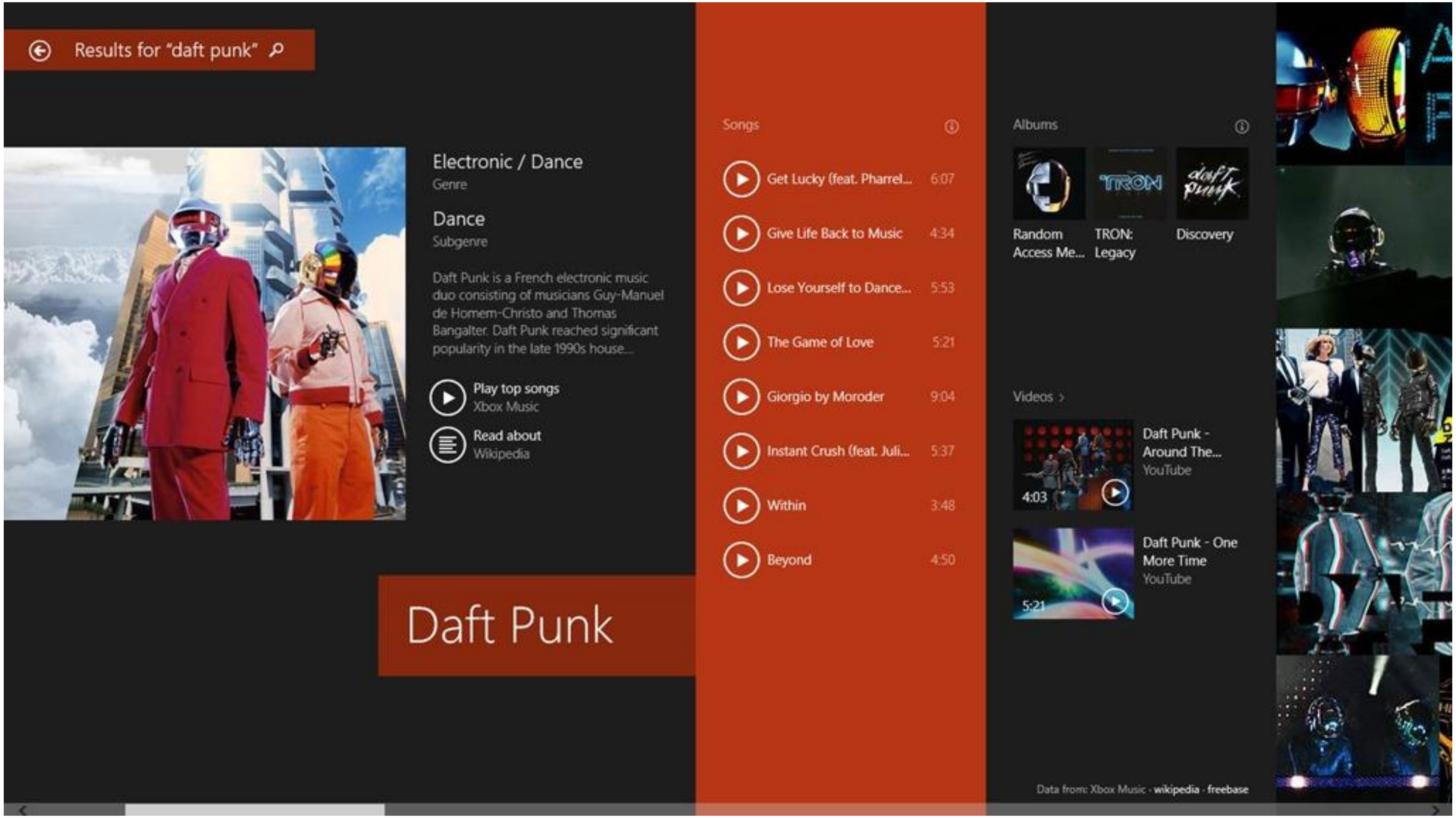


http://www.windowsphone.com/en-us/how-to/wp8/cortana/what-can-i-say-to-cortana
http://blogs.bing.com/search/2014/10/17/bing-adds-smarts-to-cortana/

Windows

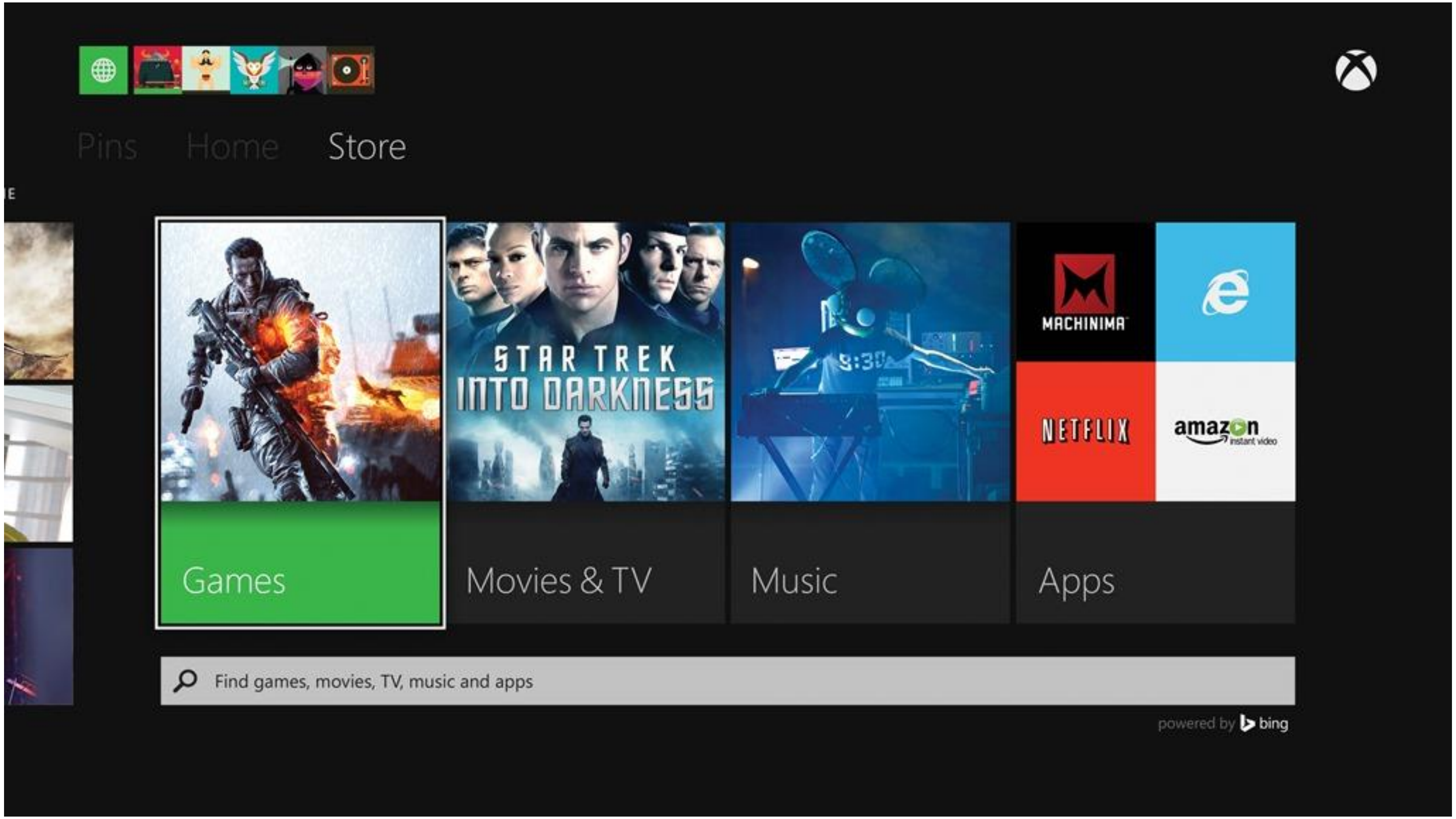
http://blogs.bing.com/search/2013/10/18/unwrapping-windows-8-1-bing-smart-search/

Windows



Results for "daft punk"
Electronic / Dance
Genre
Dance
Subgenre
Daft Punk is a French electronic music duo consisting of musicians Guy-Manuel de Homem-Christo and Thomas Bangalter. Daft Punk reached significant popularity in the late 1990s house...
Play top songs
Xbox Music
Read about
Wikipedia
Daft Punk
Songs
Get Lucky (feat. Pharrel... 6:07
Give Life Back to Music 4:34
Lose Yourself to Dance... 5:53
The Game of Love 5:21
Giorgio by Moroder 9:04
Instant Crush (feat. Juli... 5:37
Within 3:48
Beyond 4:50
Albums
Random Access Me...
TRON: Legacy
Discovery
Videos
Daft Punk - Around The...
YouTube
4:03
Daft Punk - One More Time
YouTube
5:21
Data from: Xbox Music · wikipedia · freebase

Xbox



http://blogs.bing.com/search/2013/11/19/xbox-bing-deliver-me-a-whole-new-way-to-search/

Challenges

- Speed – entity search can’t slow down web search.
- Size – serve hundreds of millions of entities online.
- Generalize to tail – how to retrieval and recommend tail entities when there are no popularity signals.
- Ambiguity – how to ask users to clarify intent.

Outline

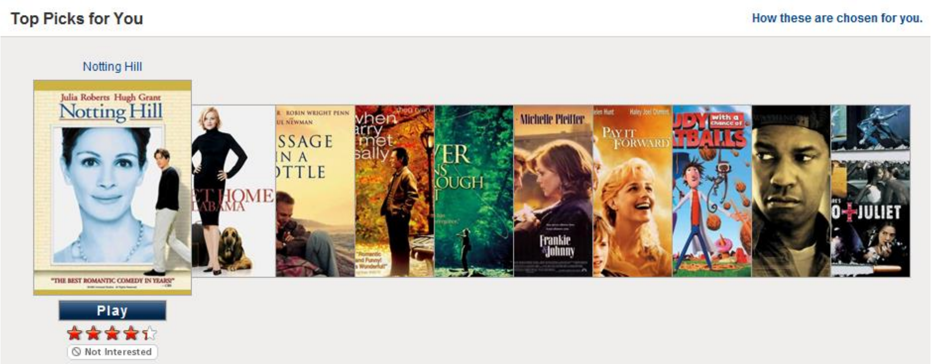
- Introduction to Entity and Knowledge
- Demonstration of Microsoft's Entity Experience
- Entity Recommendation and Understanding
  - $P(entity|entity)$
  - $P(entity|user)$
  - $P(entity|query)$
- Summary

Entity Recommendation and Understanding

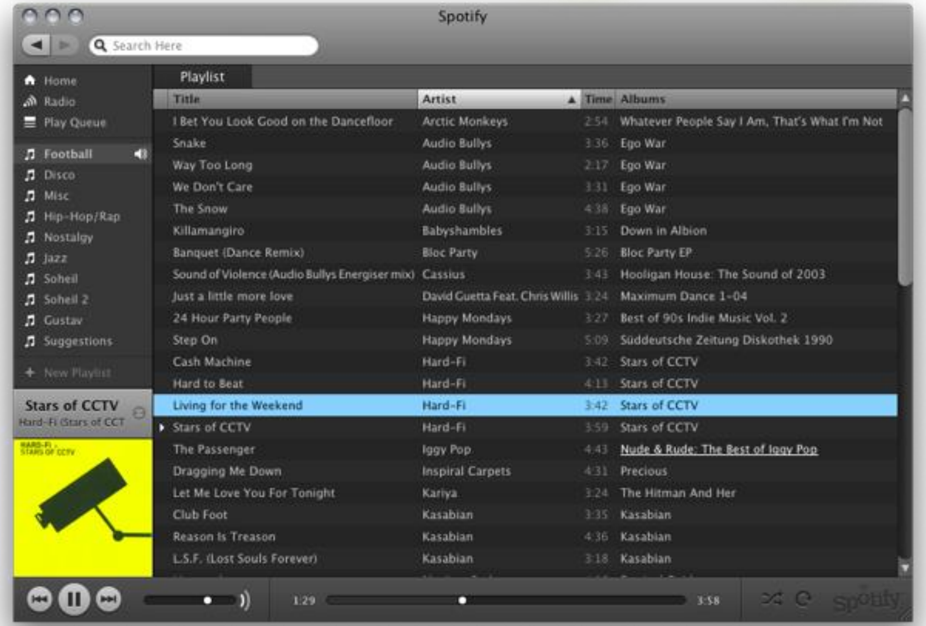
Traditional Recommender Systems

amazon.com

NETFLIX

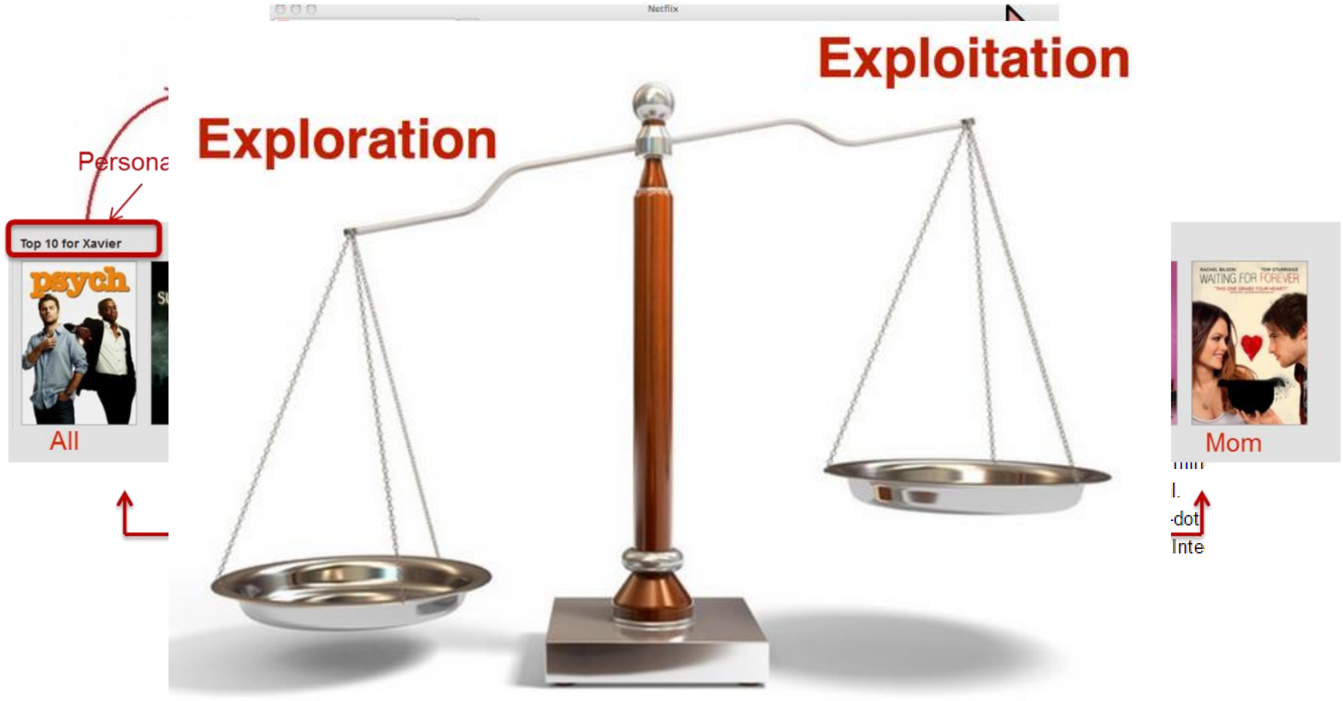


Spotify



Traditional Recommender Systems

- Recommender Systems have been well studied from many aspects
  - Collaborative filtering
  - Content-based
  - Context-aware
  - Rating-based
  - Learning to Rank
  - Diversity
  - Serendipity
  - Social-Aware
  - Temporal
  - Explore/Exploit
  - … …



Traditional Recommender Systems

- Majority of the algorithms is focusing on Personalization
  - $P(item|user)$
  - User-Item Matrix

	v_1	v_2	v_3	v_4	v_5
u_1	1		2	3	
u_2		3			1
u_3		4		5	
u_4	5			4	
u_5		2	5		4

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		1	1		1	
u_2	1			1		1
u_3			1			1
u_4	1			1		
u_5		1	1			1

Entity Recommender Systems

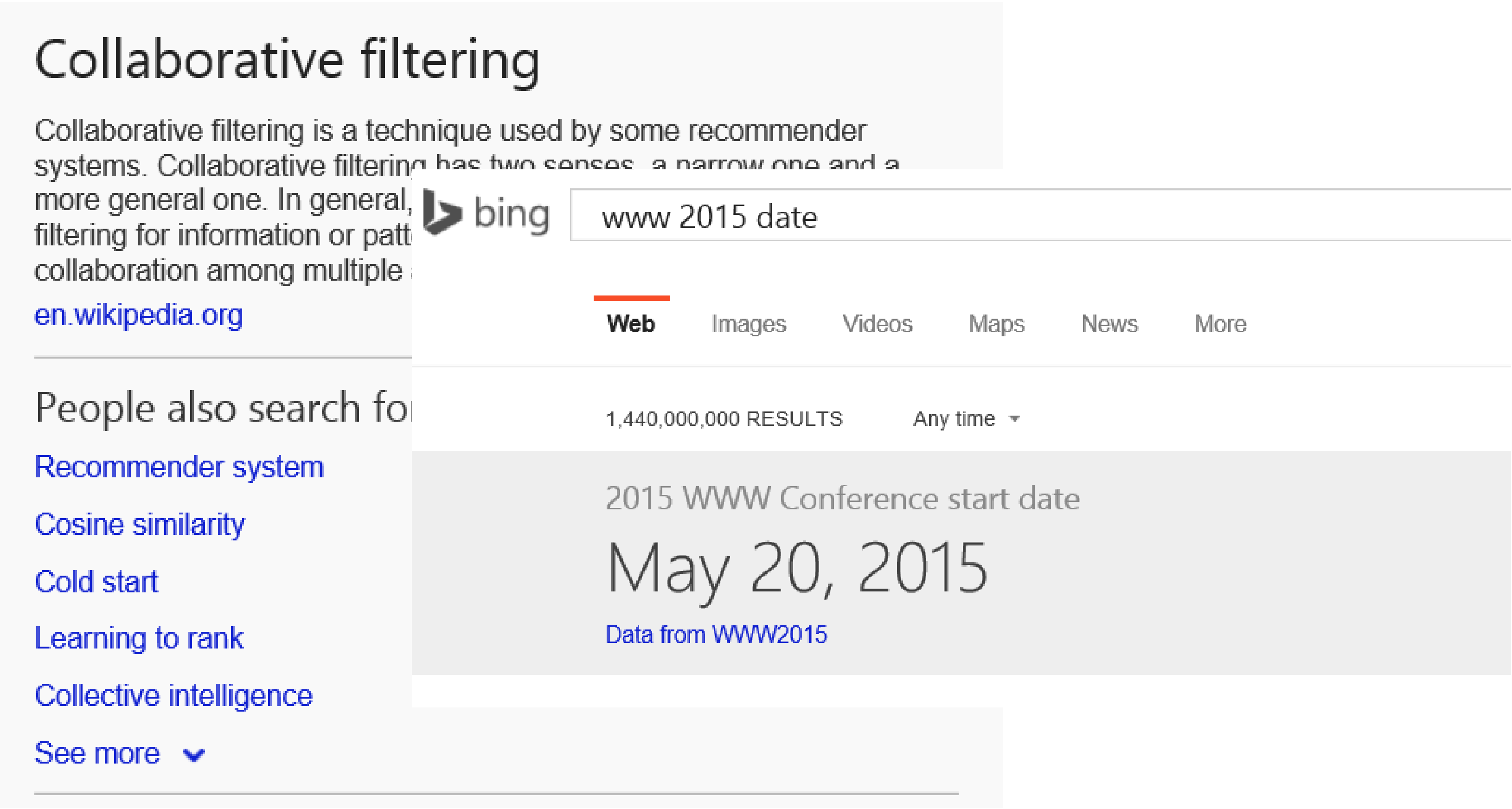
- Entity Graph
  - Heterogeneous Graph
    Freebase
    2K+ commonly used types
    30K+ commonly used properties

Entity Recommender Systems

- Entity Graph
  – Hugh Size
    Freebase
    47M+ topics
    2.9B+ facts

Entity Recommender Systems

- Non-Personalized
  - $P(item|item)$
  - $P(item|query)$



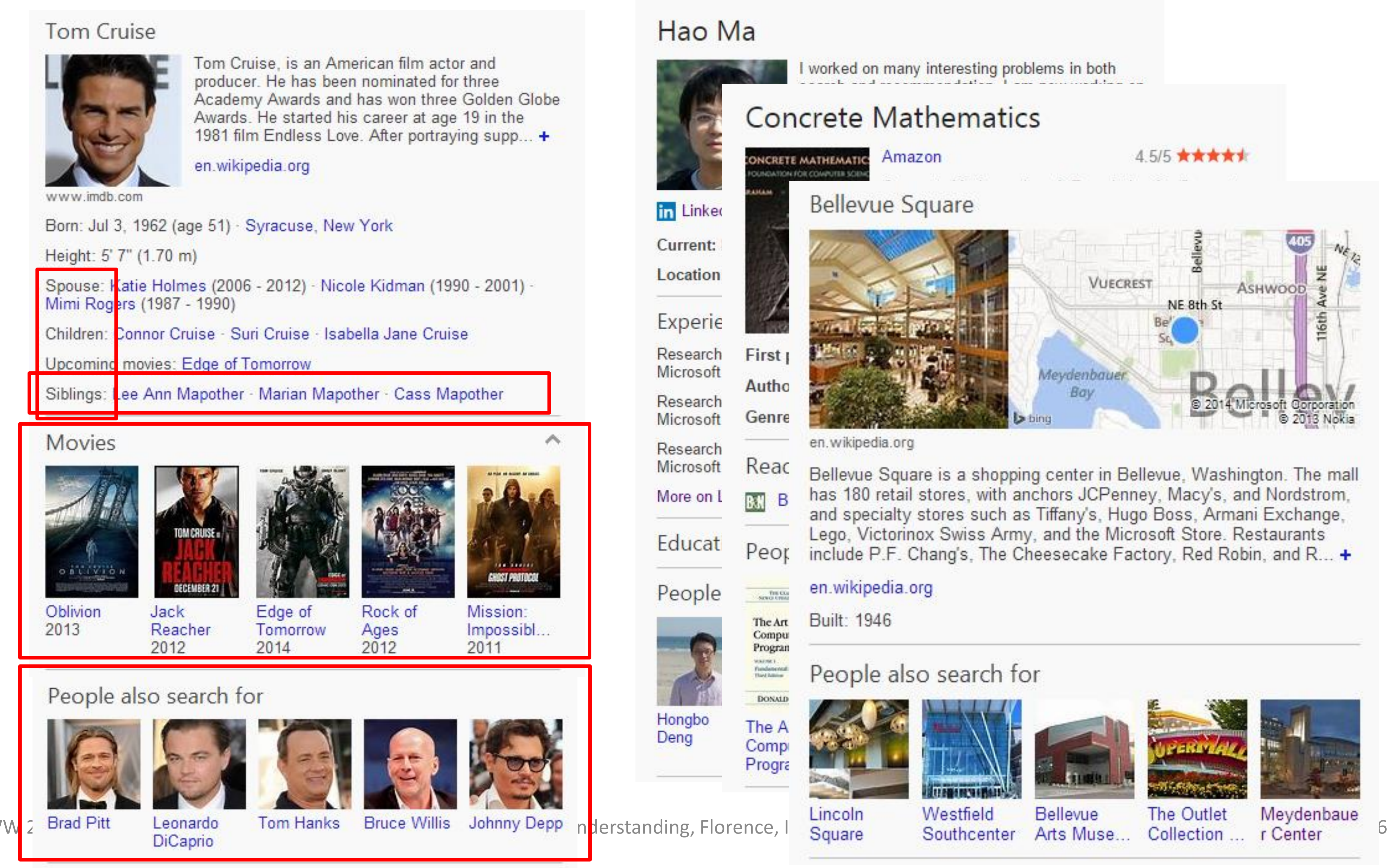
Entity Recommender Systems

- Personalization
  - $P(item|user)$

Entity Recommender Systems in Search

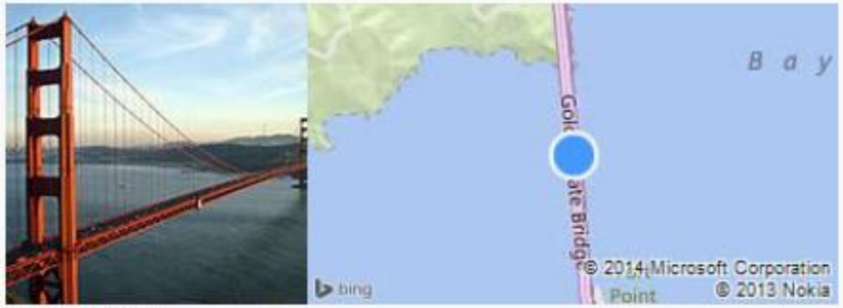
- Recommendation & Ranking
- Interpretation
- Entity Collection Recommendation
- Exploration
- Personalization

Recommendation & Ranking



Interpretation

Golden Gate Bridge



en.wikipedia.org

The Golden Gate Bridge is a suspension bridge spanning the Golden Gate strait, the mile-wide, three-mile-long channel between San Francisco Bay and the Pacific Ocean. The structure links the U.S. city of San Francisco, on the northern tip of the San Francisco Peninsula, to Marin County, bridging both U.S. Route 101 and California State R… +

en.wikipedia.org

Built: 1937

Length: 8,980 feet (2,737 m)

Height: 746 feet (227.40 m)

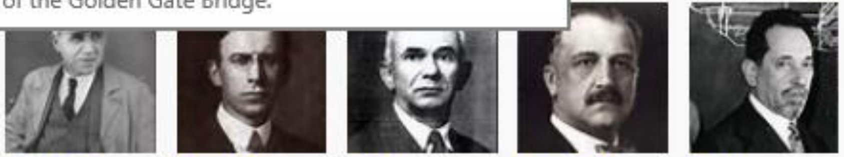
Architects: Joseph Strauss · Charles Alton Ellis · Irving Morrow

Location: San Francisco · Marin County

Crosses: Golden Gate

Bridge type: Suspension bridge · Truss bridge · Truss arch bridge

Charles Alton Ellis was a professor, structural engineer and mathematician who was chiefly responsible for the structural design of the Golden Gate Bridge.



Joseph Strauss
Architect

Charles Alton Ellis
Architect

Irving Morrow
Architect

Amadeo Giannini

Leon Moisseiff

Harvard University

www.harvard.edu

Harvard University is a private Ivy League research university in Cambridge, Massachusetts, whose history, influence and wealth have made it one of the most prestigious universities in the world. Established in 1636 by the Massachusetts legislature and soon thereafter named for John Harvard, Harvard is the United States' o… +

en.wikipedia.org

Address: 1350 Massachusetts Ave, Cambridge, MA, 02138 · Directions

Ranking: #2 National University (2014)

Undergraduates: 6,658 (2014)

Acceptance rate: 6% (2014)

Tuition: $42,292 USD (2014)

Founder: John Harvard

School color: Crimson

Mascot: Harvard University John Harvard

Born in Honolulu, Hawaii, Barack Obama is a graduate of Columbia University and Harvard Law School, where he was president of the Harvard Law Review.

Related people

John Harvard
Founder

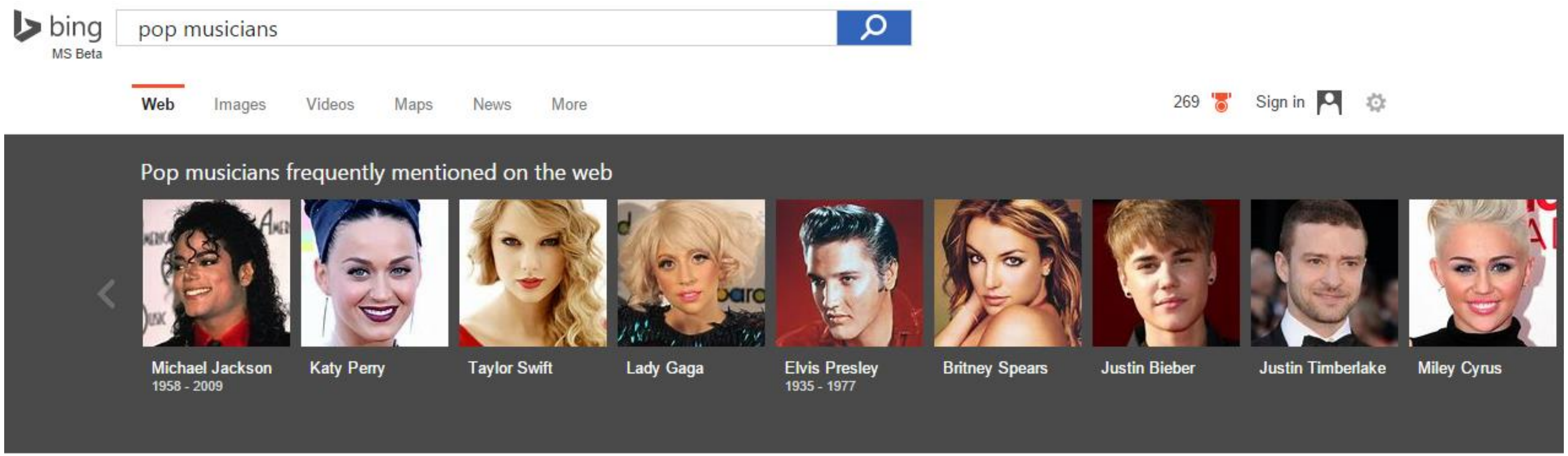
Drew Gilpin Faust

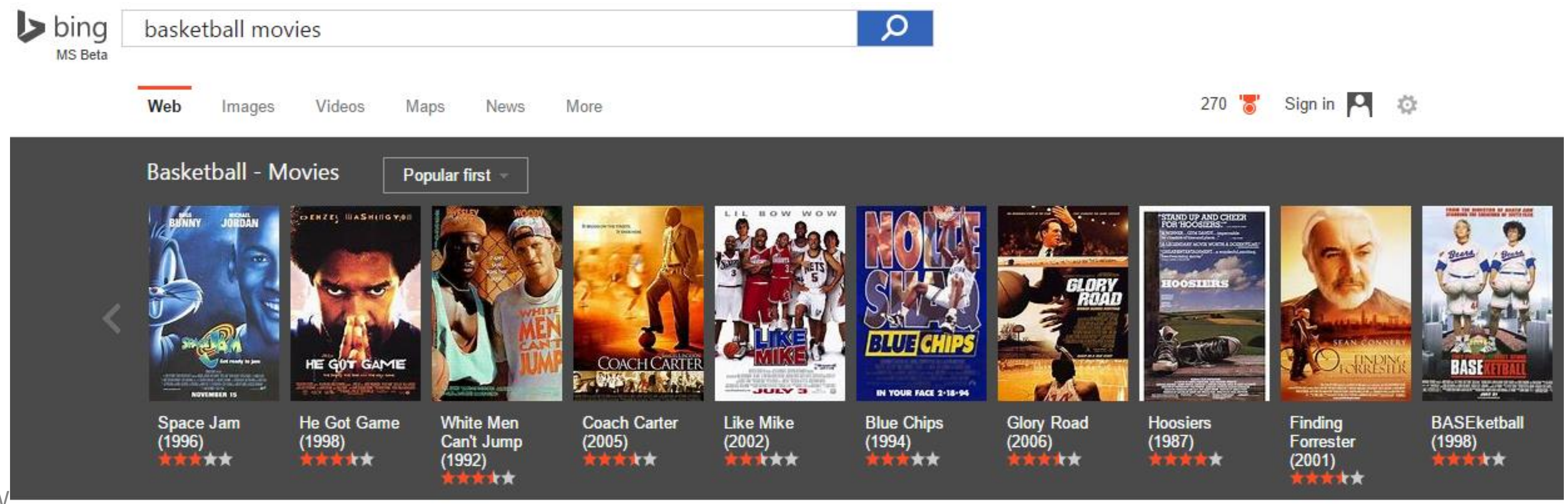
Mark Zuckerberg
Alumni

Barack Obama
Alumni

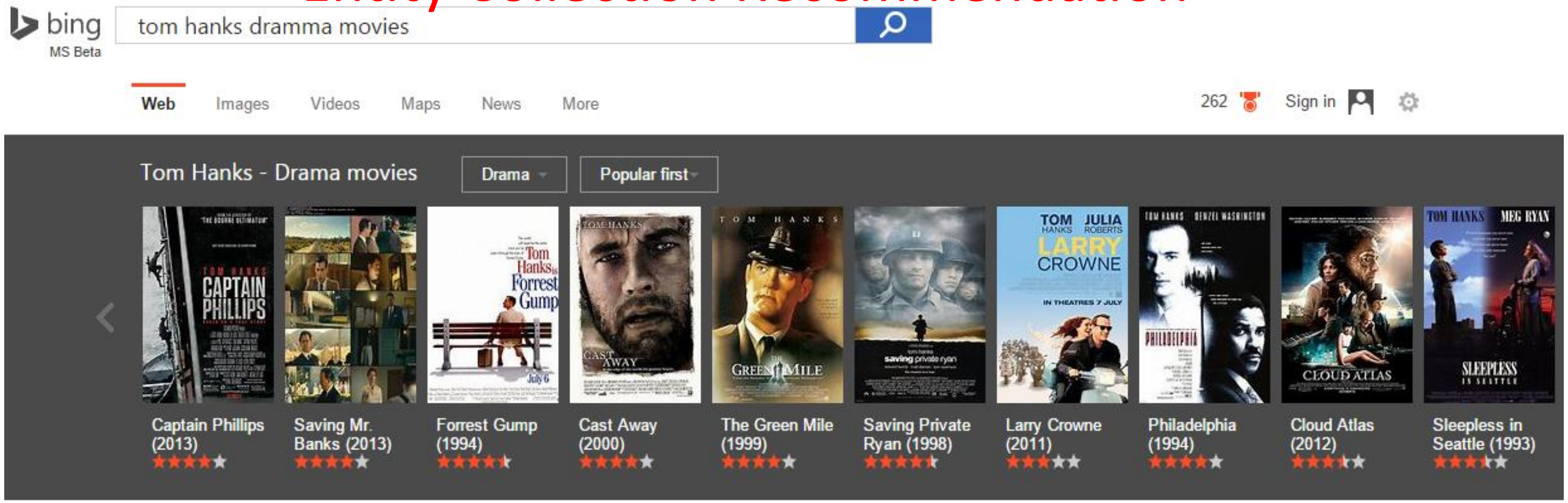
Bill Gates
Alumni

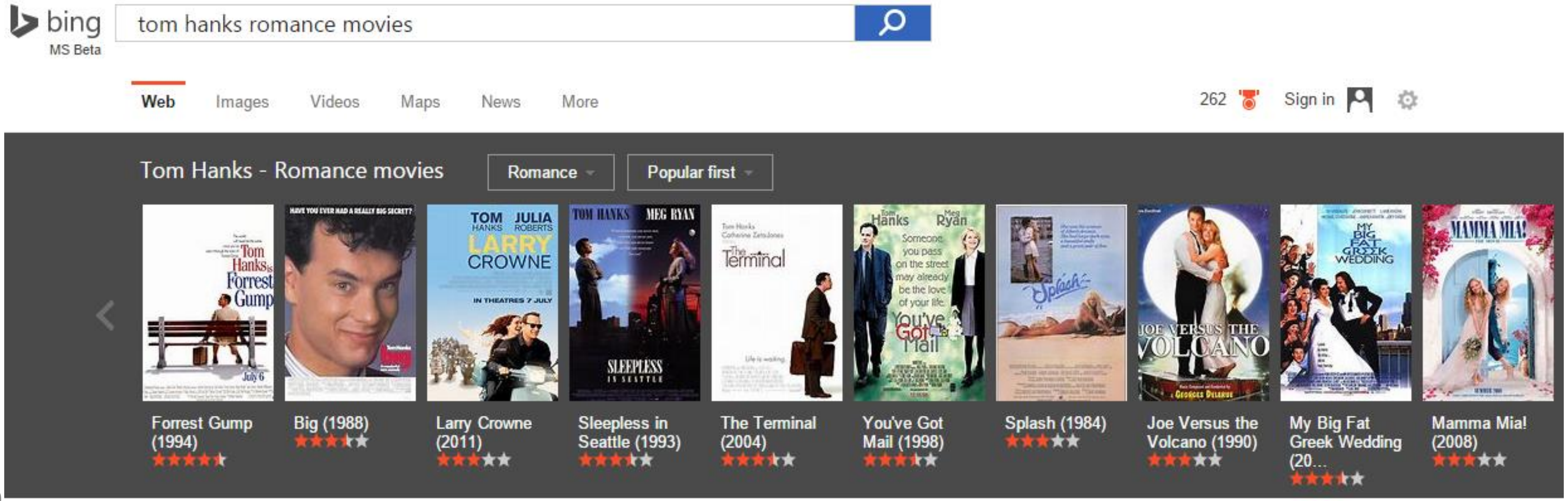
Entity Collection Recommendation



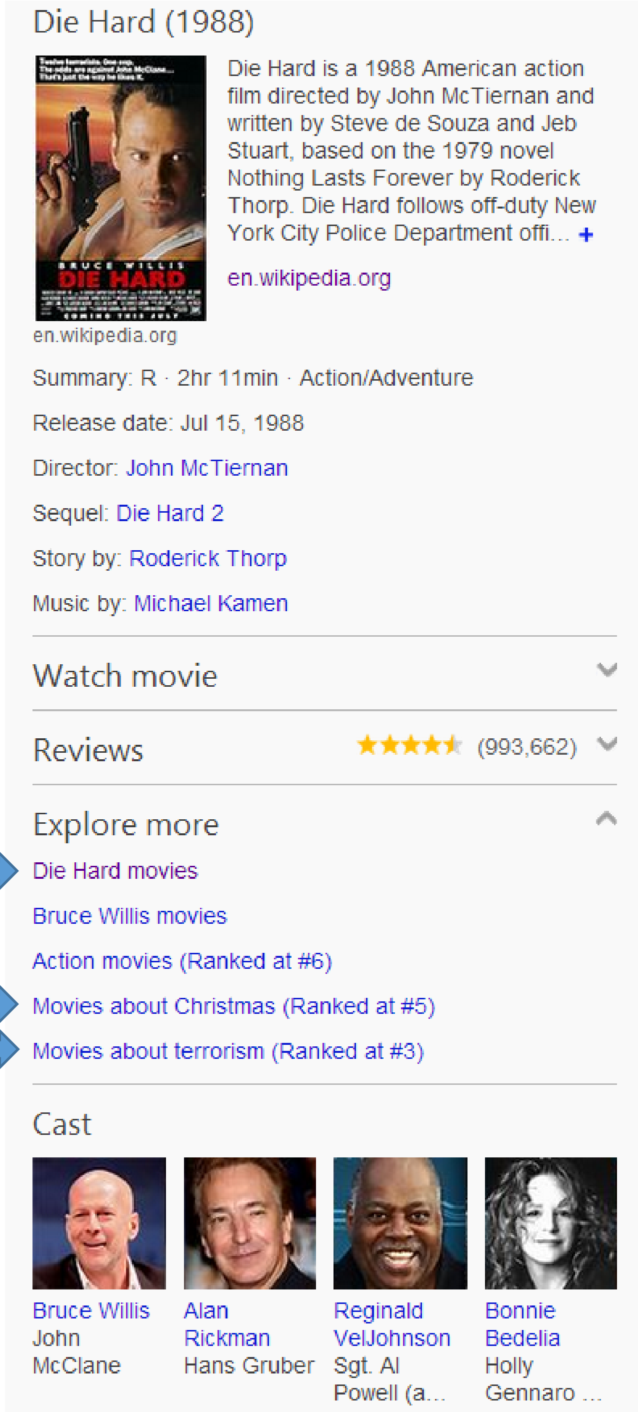


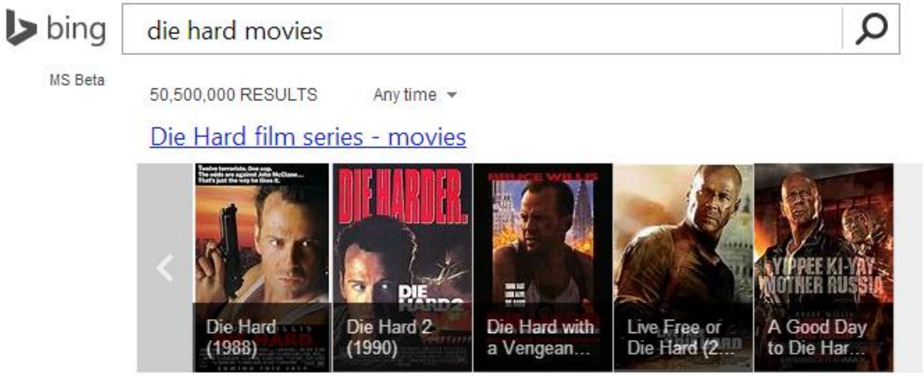
Entity Collection Recommendation

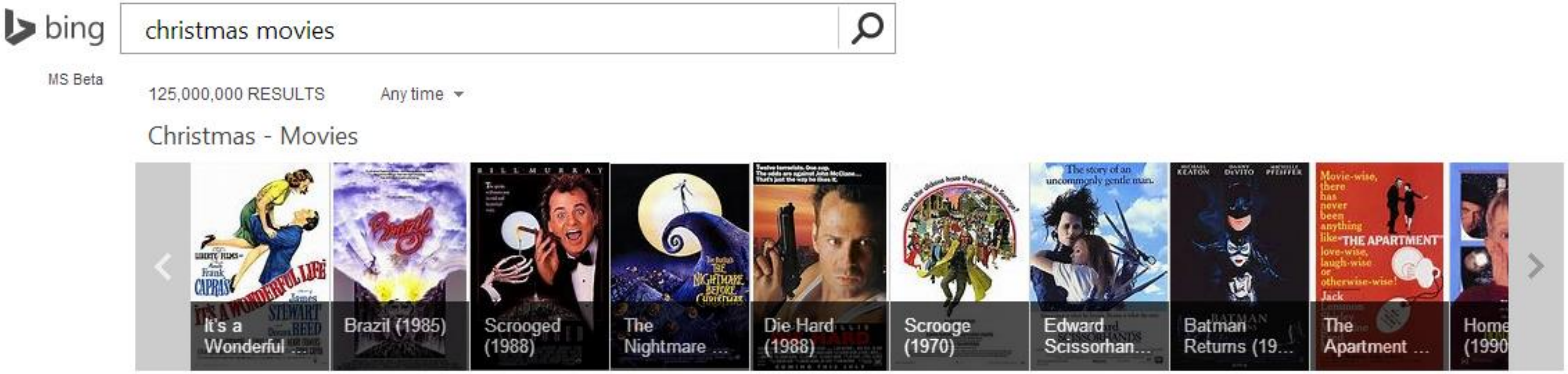


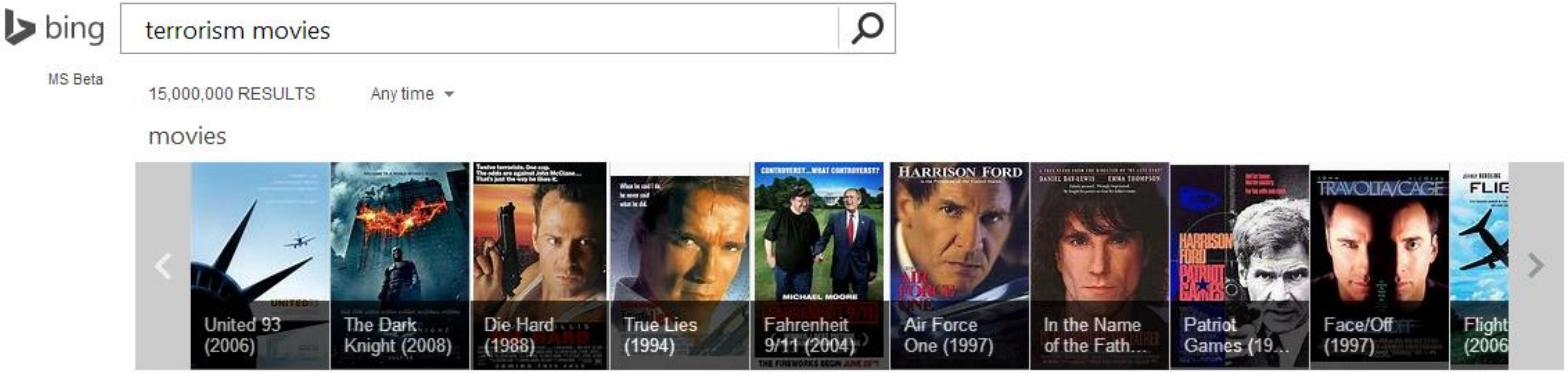


Exploration









Entity Recommendation & Understanding Taxonomy

- $P(item|item)$
  - Recommendations given an item
- $P(item|user)$
  - Recommendations given a user
- $P(item|query)$
  - Recommendations given a query

Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$
  - Recommendations given an entity
- $P(entity|user)$
  - Recommendations given a user
- $P(entity|query)$
  - Recommendations given a query

Item $\longleftrightarrow$ Entity

$P(entity|entity)$

- $P(Florence|Italy) = \frac{Freq(Florence, Italy)}{Freq(Italy)}$ Co-occurrence

- $P(Florence|Italy) = \frac{Sim(Florence,Italy)}{\sum Sim(*, Italy)}$ Cosine Similarity

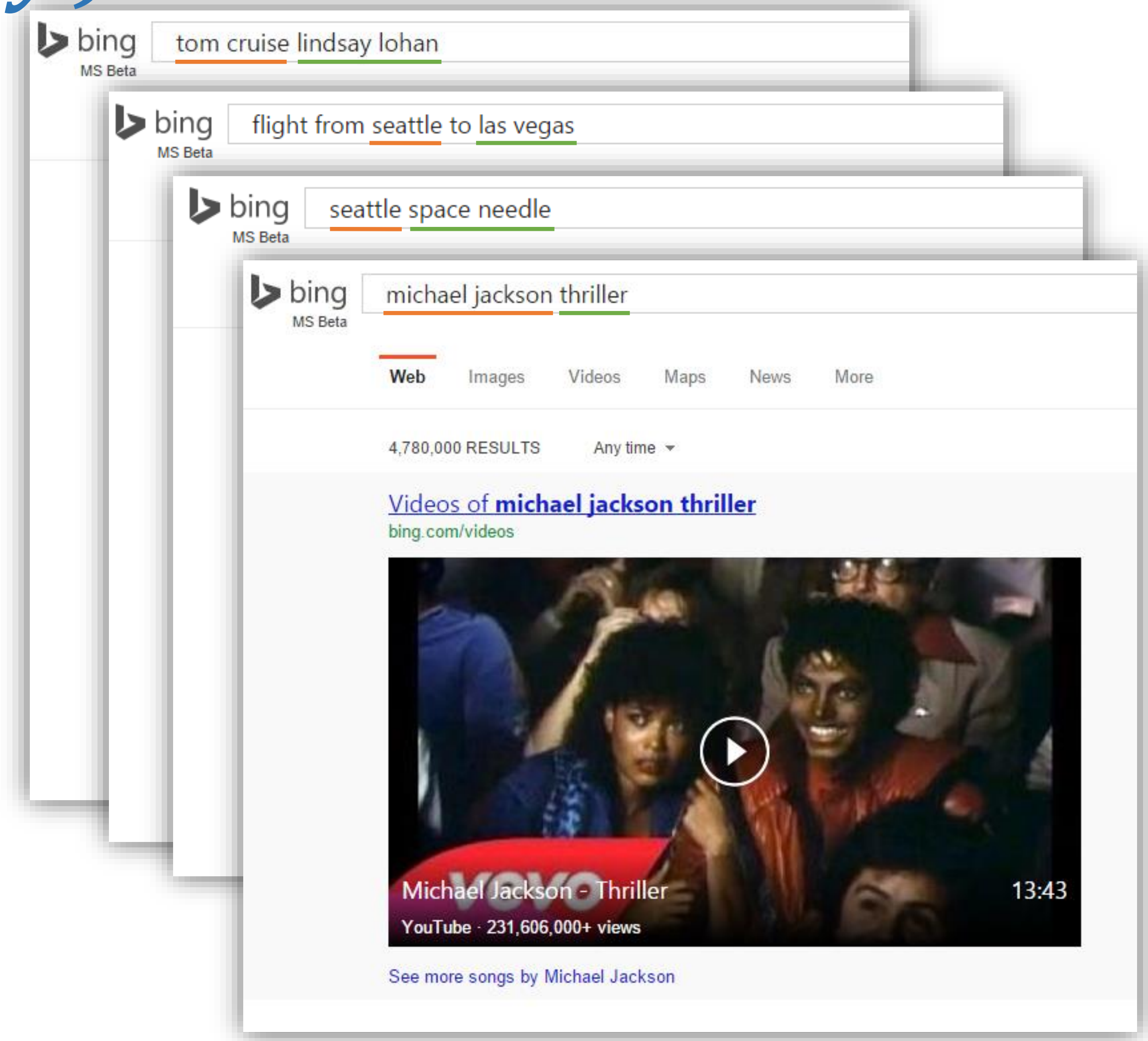
Item-Based Top-N Recommendation Algorithms [Mukund Deshpande, et al., ACM TOIS 2004]

$P(entity|entity)$ – Co-occurrence

- Sources
  - Within Queries
  - Across Queries
  - User Url Clicks
  - Wikipedia Pages
  - Wikipedia Categories/Templates
  - Wikipedia Revision Histories
  - Web documents
  - … …

$P(entity|entity)$ – Co-occurrence

- Within Queries

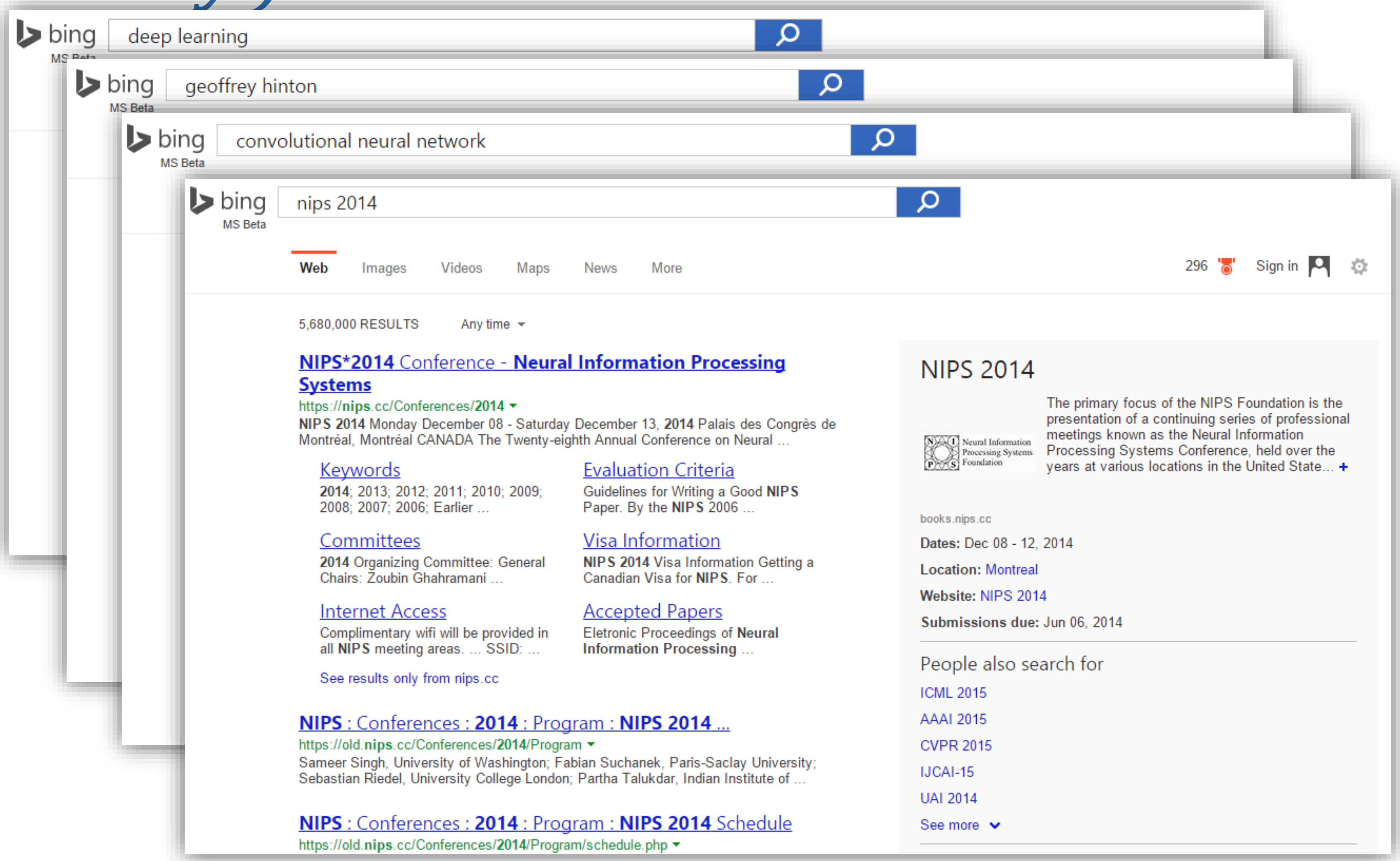


How to extract those entities?

Entity Recommendations in Web Search [Roi Blanco, et al., ISWC 2013]

$P(entity|entity)$ – Co-occurrence

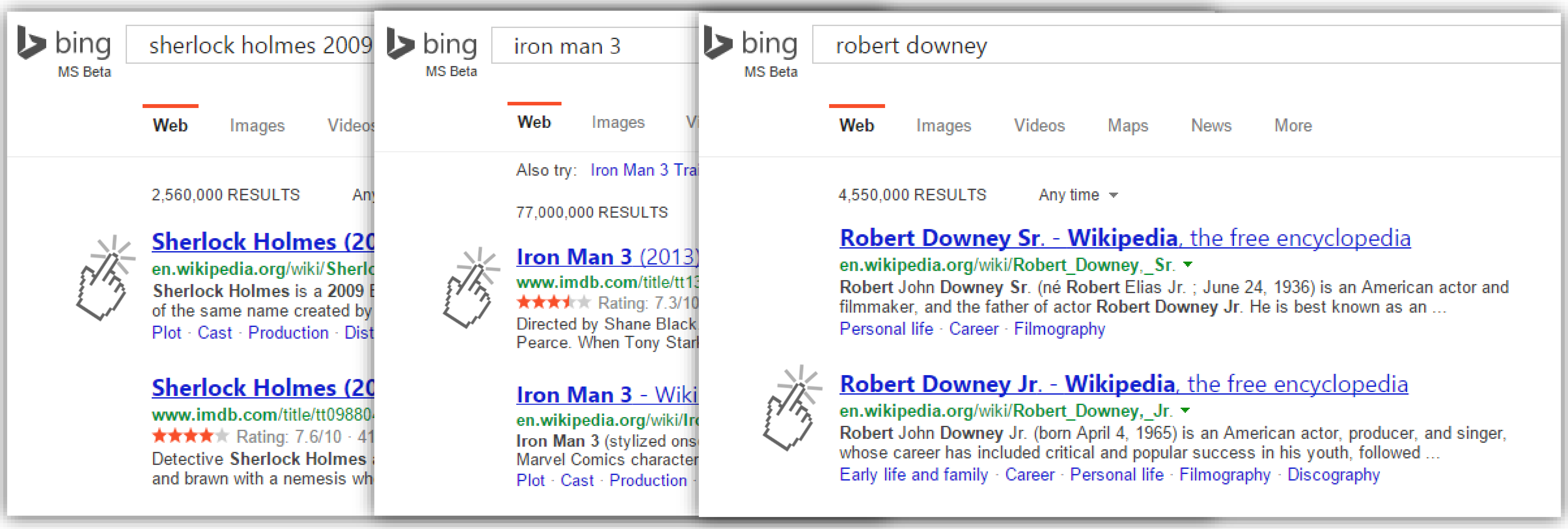
- Across Queries



Entity Recommendations in Web Search [Roi Blanco, et al., ISWC 2013]

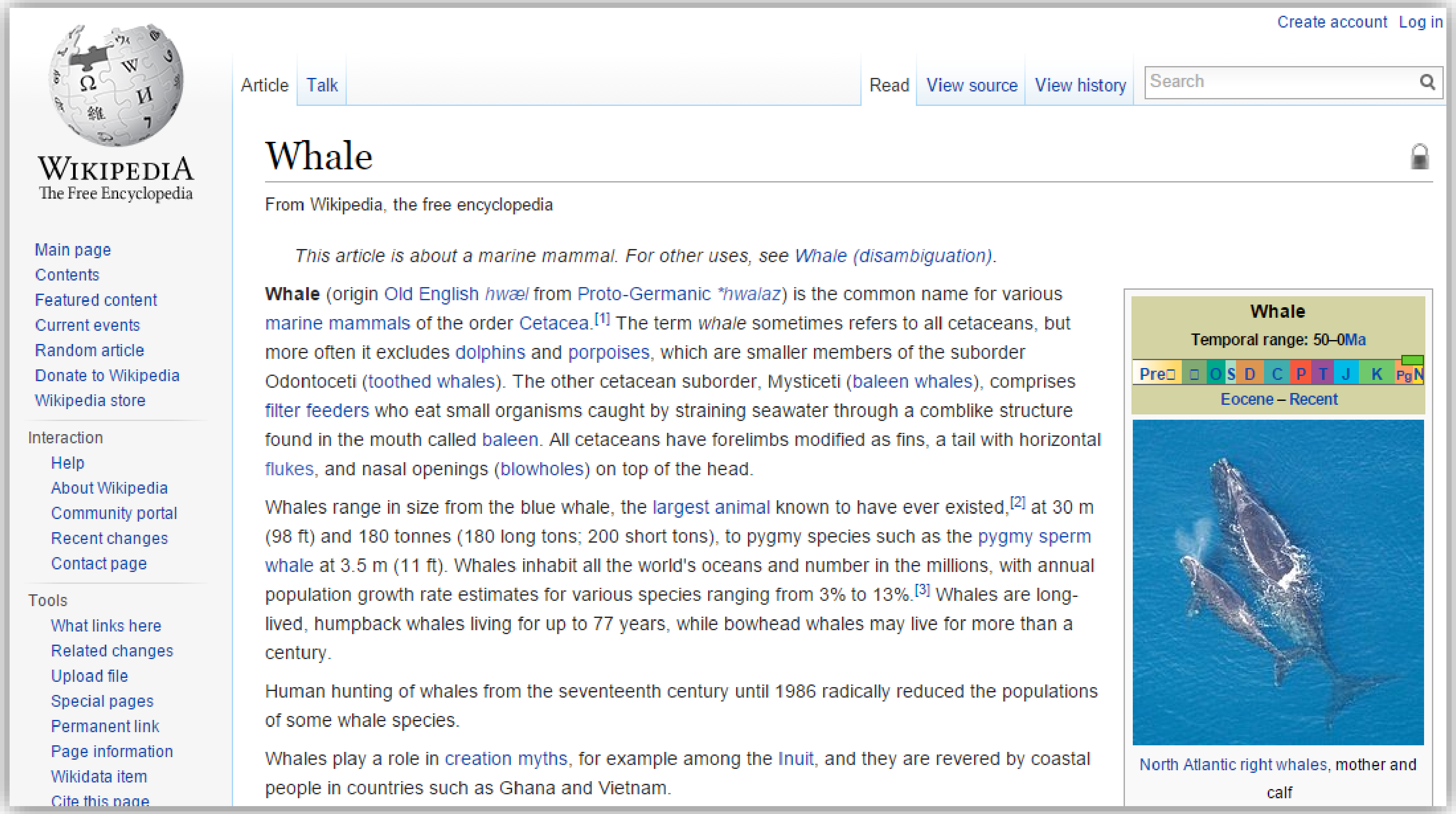
$P(entity|entity)$ – Co-occurrence

– User Url Clicks



$P(entity|entity)$ – Co-occurrence

– Wikipedia Pages



Create account Log in

Article Talk Read View source View history Search

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Whale

From Wikipedia, the free encyclopedia

This article is about a marine mammal. For other uses, see Whale (disambiguation).

Whale (origin Old English *hwæl* from Proto-Germanic **hwalaz*) is the common name for various marine mammals of the order Cetacea.[1] The term *whale* sometimes refers to all cetaceans, but more often it excludes dolphins and porpoises, which are smaller members of the suborder Odontoceti (toothed whales). The other cetacean suborder, Mysticeti (baleen whales), comprises filter feeders who eat small organisms caught by straining seawater through a comblike structure found in the mouth called baleen. All cetaceans have forelimbs modified as fins, a tail with horizontal flukes, and nasal openings (blowholes) on top of the head.

Whales range in size from the blue whale, the largest animal known to have ever existed,[2] at 30 m (98 ft) and 180 tonnes (180 long tons; 200 short tons), to pygmy species such as the pygmy sperm whale at 3.5 m (11 ft). Whales inhabit all the world's oceans and number in the millions, with annual population growth rate estimates for various species ranging from 3% to 13%.[3] Whales are long-lived, humpback whales living for up to 77 years, while bowhead whales may live for more than a century.

Human hunting of whales from the seventeenth century until 1986 radically reduced the populations of some whale species.

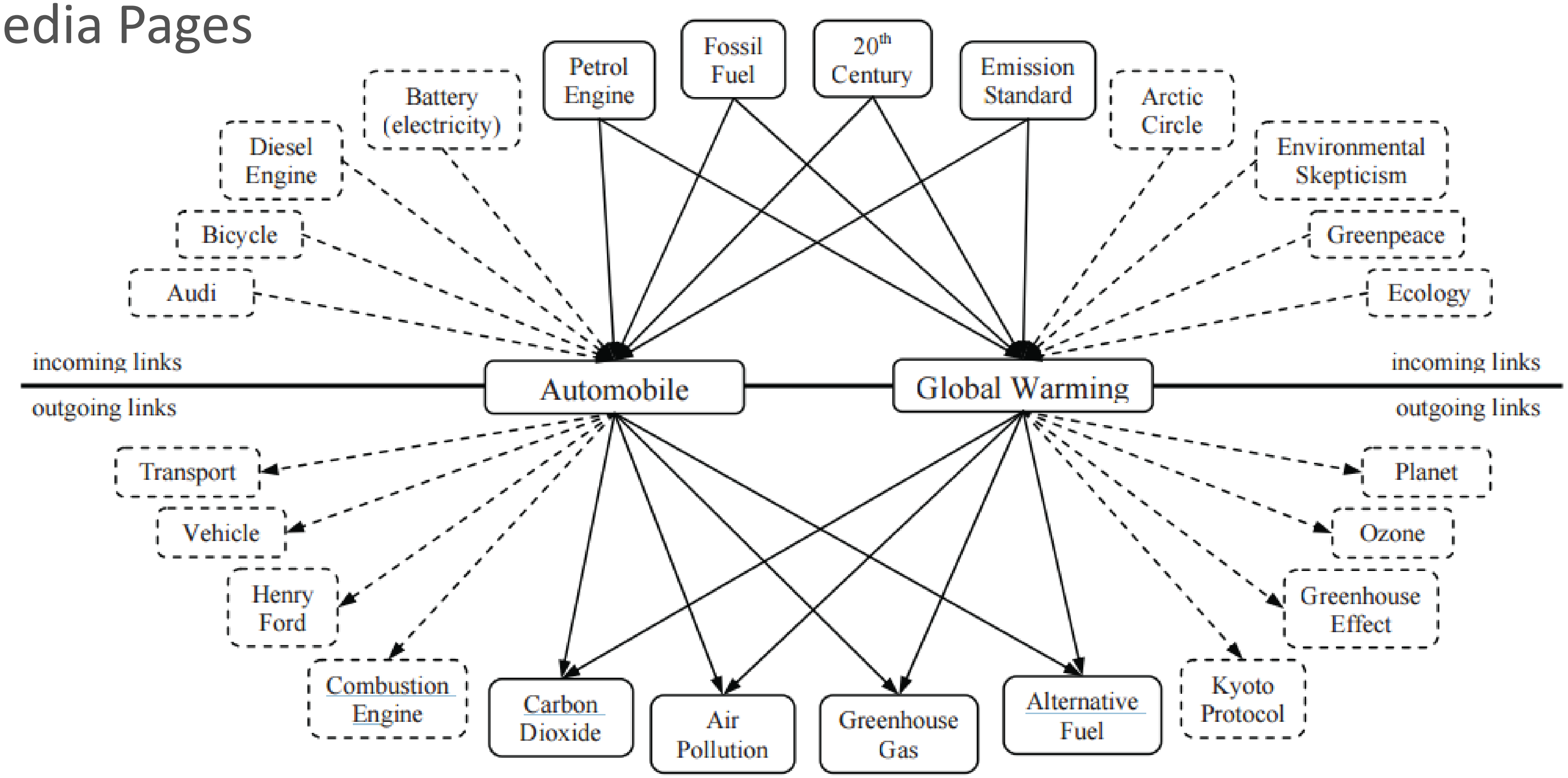
Whales play a role in creation myths, for example among the Inuit, and they are revered by coastal people in countries such as Ghana and Vietnam.

Whale
Temporal range: 50–0Ma
Pre O S D C P T J K Pg N
Eocene – Recent

North Atlantic right whales, mother and calf

$P(entity|entity)$ – Co-occurrence

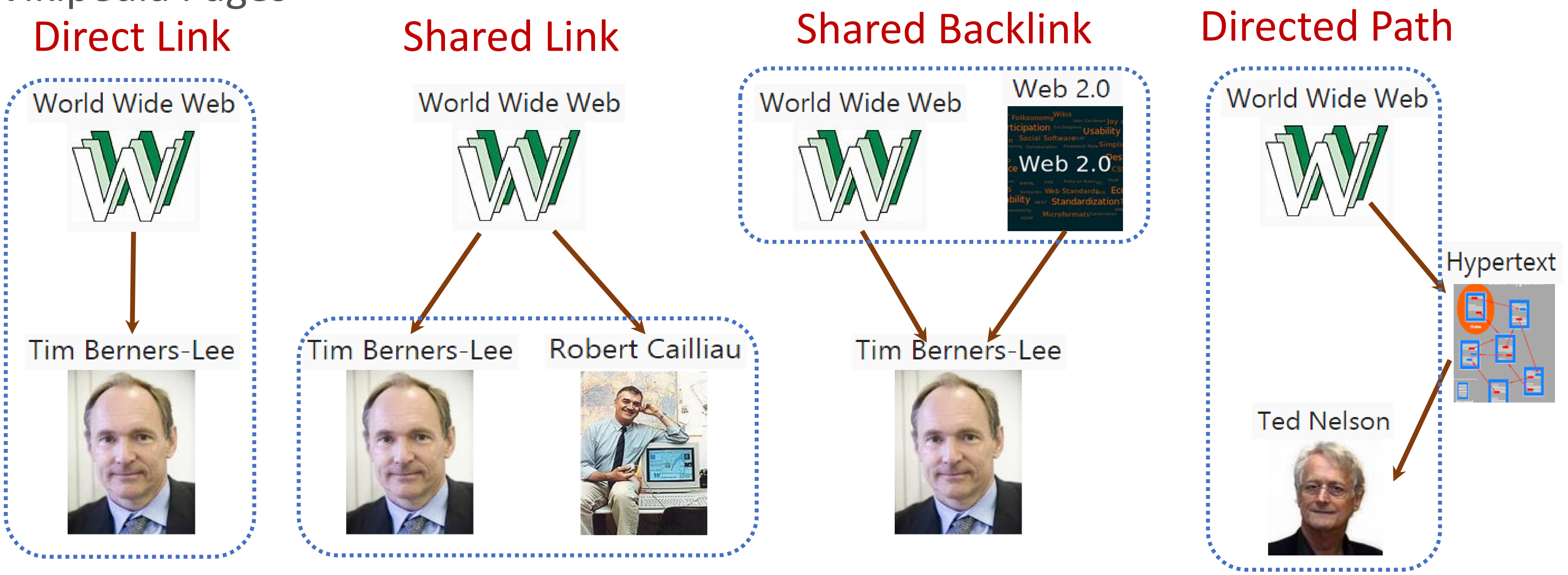
– Wikipedia Pages



An Effective, Low-Cost Measure of Semantic Relatedness Obtained from Wikipedia Links [David Milne, et al., AAAI 2008]

$P(entity|entity)$ – Co-occurrence

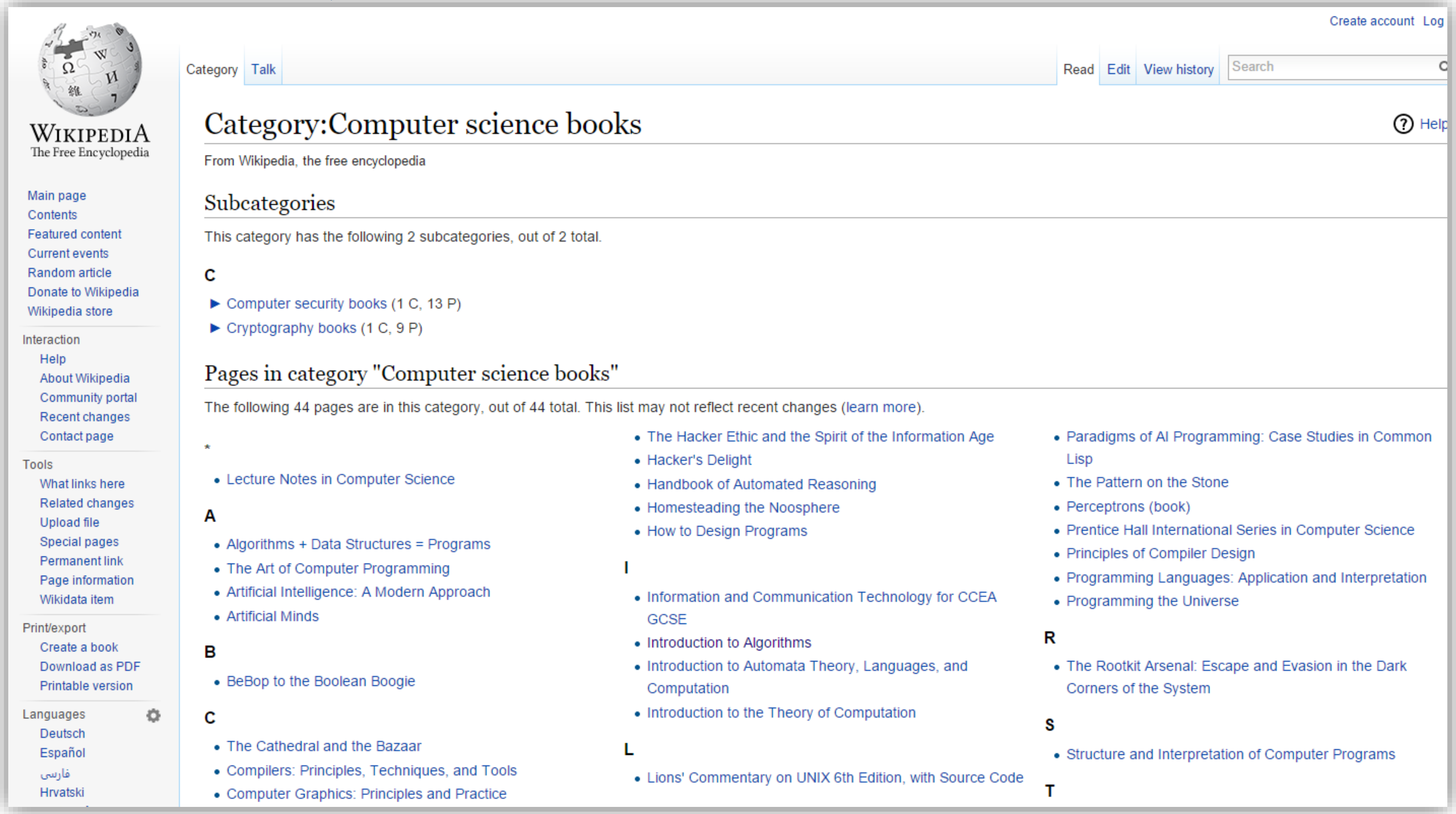
– Wikipedia Pages



Exploratory Recommendations Using Wikipedia's Linking Structure [Adrian M. Kentsch, et al., Benelearn 2011]

$P(entity|entity)$ – Co-occurrence

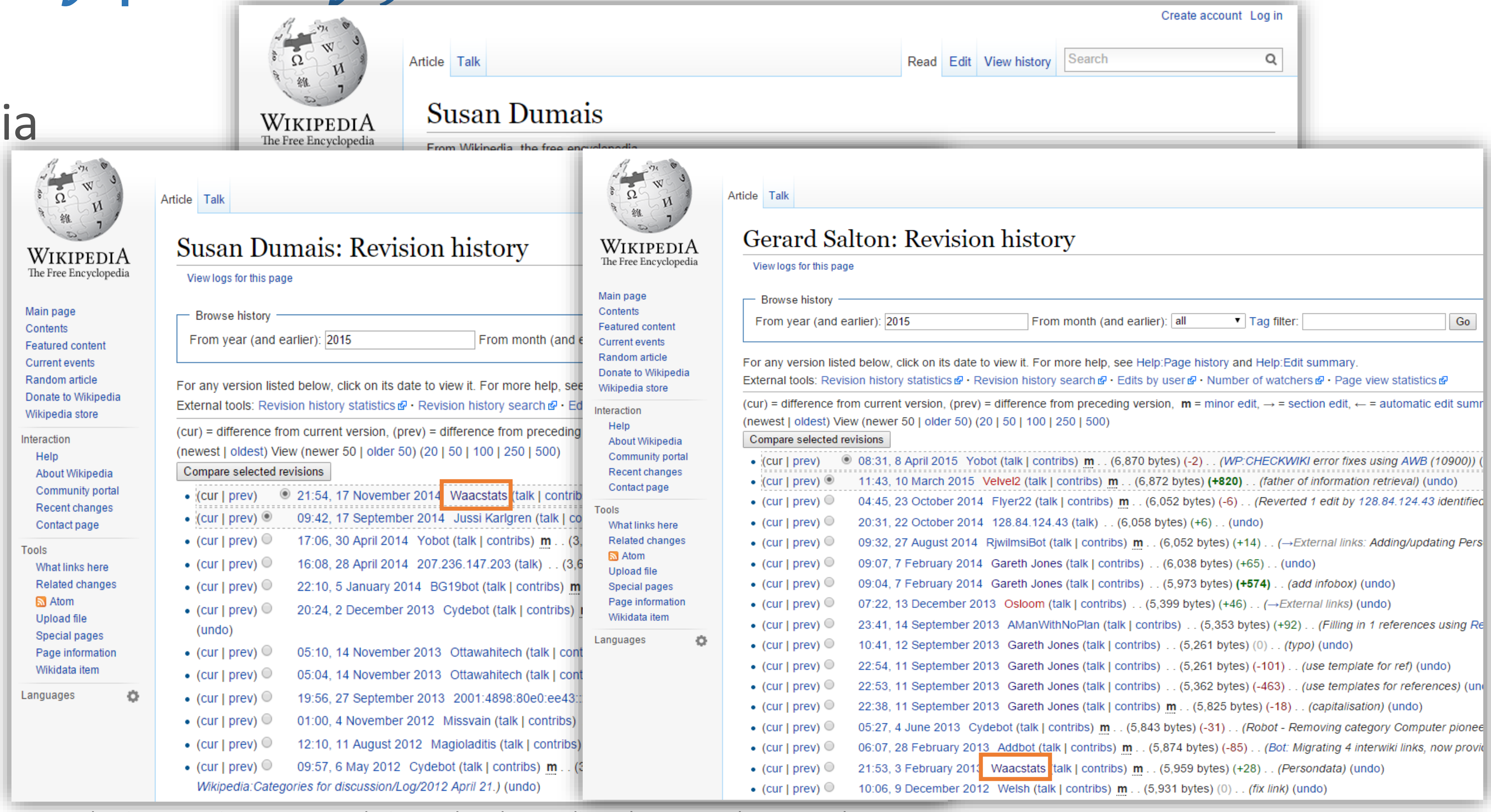
- Wikipedia Categories & Templates



WikiRelate! Computing Semantic Relatedness Using Wikipedia [Michael Strube, et al., AAAI 2006]

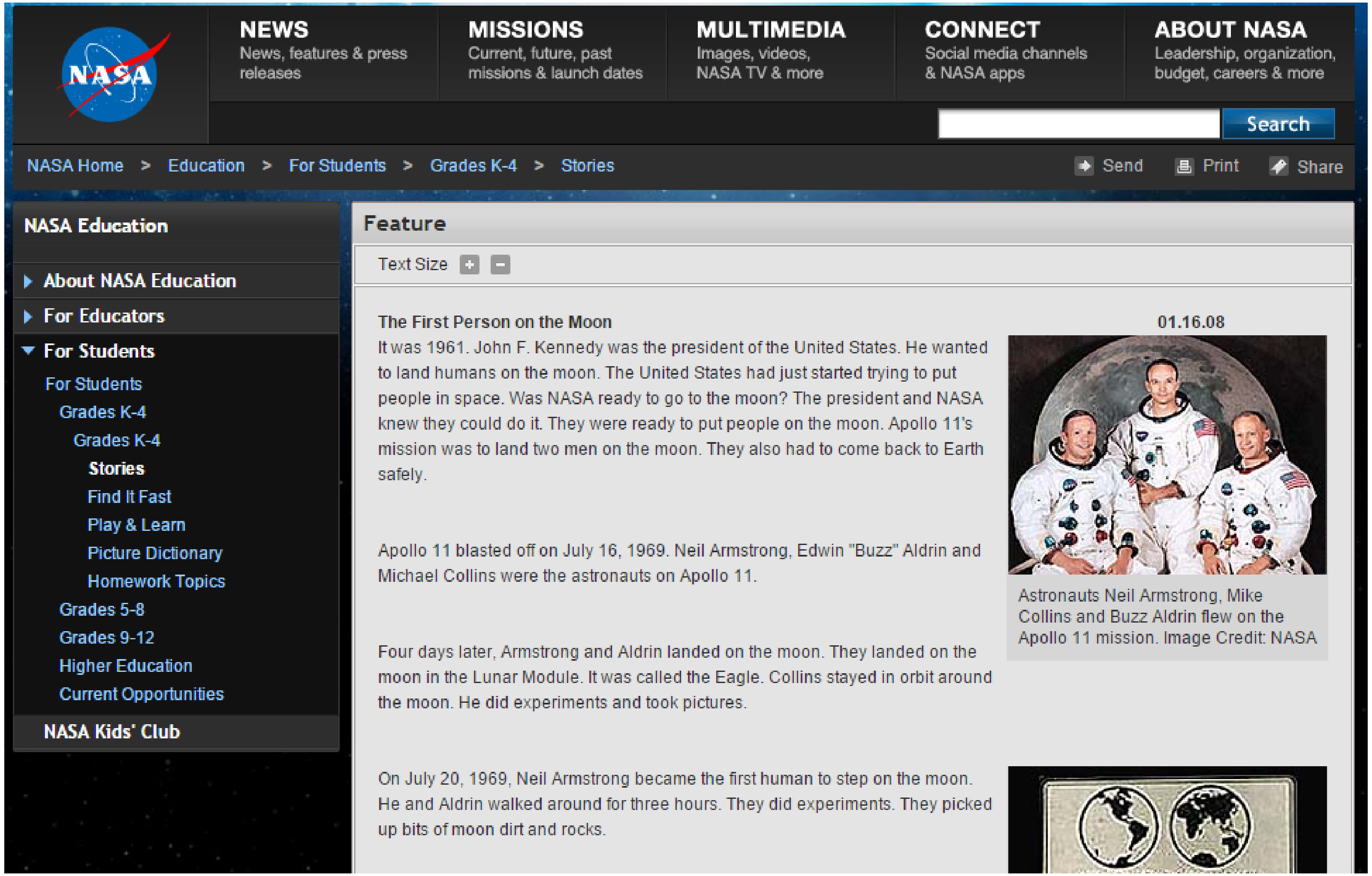
$P(entity|entity)$ – Co-occurrence

- Wikipedia Revision History



$P(entity|entity)$ – Co-occurrence

- Web documents
  - Moon
  - John F. Kennedy
  - United States
  - NASA
  - Apollo 11
  - Neil Armstrong
  - Edwin “Buzz” Aldrin
  - Michael Collins
  - Astronauts



TimeMachine:Timeline Generation for Knowledge-Base Entities [Tim Althoff, et al., arXiv:1502.04662, 2015]

$P(entity|entity)$ – Co-occurrence

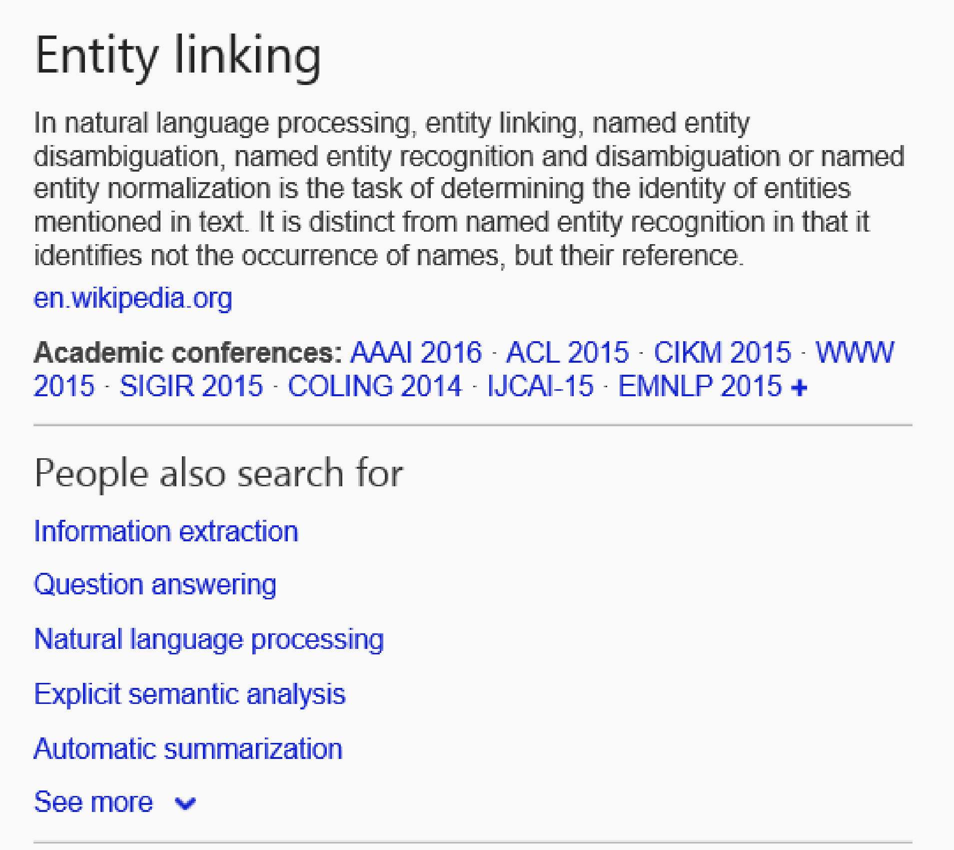
- Entity Recommendations based on Wikipedia Co-occurrence

Whale	Susan Dumais	Tom Cruise
Dolphin	C. J. van Rijsbergen	Nicole Kidman
Pinniped	W. Bruce Croft	Brad Pitt
Shark	Eric Horvitz	Steven Spielberg
Killer whale	George Furnas	Tom Hanks
Humpback whale	Thomas Landauer	John Travolta

Susan Dumais
Latent semantic analysis
Information retrieval
Gerard Salton Award
SIGIR
Singular value decomposition

Entity Linking

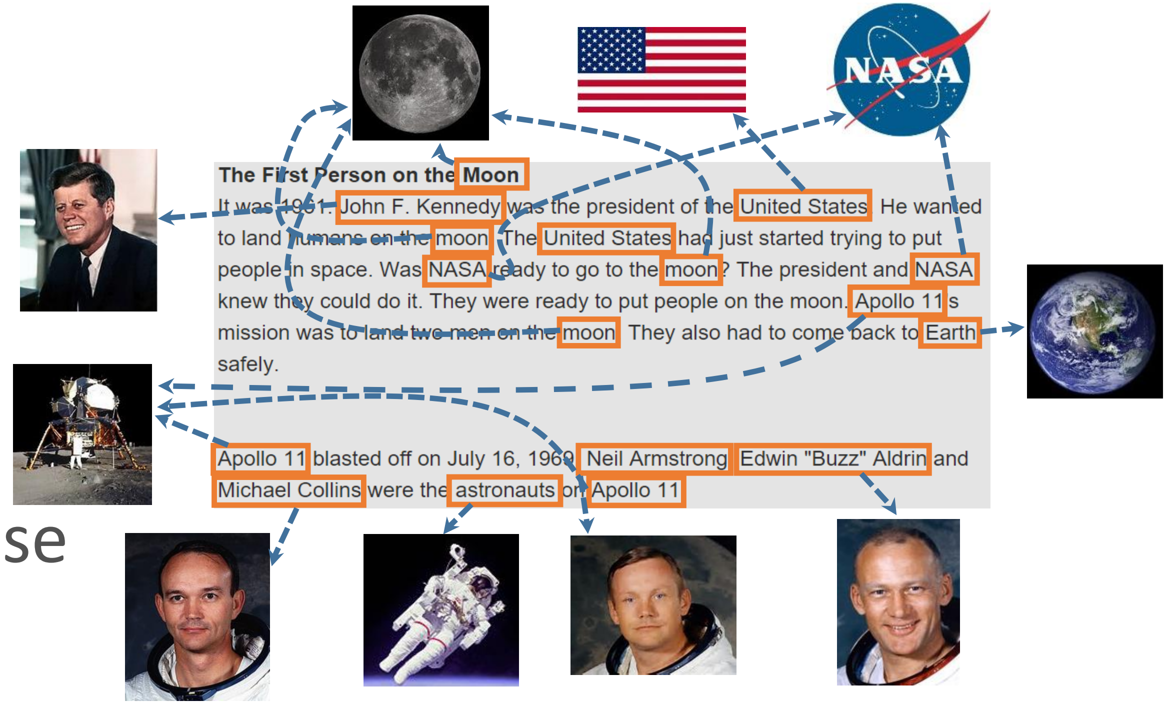
- How to extract entities from Queries and Documents?
  - Through Entity Linking!



Entity Linking

Entity Linking - Main Problem

- Linking free text to entities
  - Any piece of text
    - News document
    - Blog posts
    - Tweets
    - Queries
    - …
- Entities taken from a knowledge base
  - Freebase
  - Wikipedia
  - …



Entity Linking and Retrieval for Semantic Search [Edgar Meij, et al., WSDM 2014]

Entity Linking - Common Steps

- Determine “linkable” phrases
  - Mention detection
- Select candidate entity links
  - Link generation
  - May include NILs (null values, i.e., no target in KB)
- Use “context” to disambiguate/filter/improve
  - Disambiguation

Entity Linking and Retrieval for Semantic Search [Edgar Meij, et al., WSDM 2014]

Entity Linking

- An Example

Depth-first search

From Wikipedia, the free encyclopedia

Depth-first search (DFS) is an algorithm for traversing or searching a tree, tree structure or graph. One starts at the root (selecting some node as the root in the graph case) and explores as far as possible along each branch before backtracking.

Formally, DFS is an uninformed search that progresses by expanding the first child node of the search tree that appears and thus going deeper and deeper until a goal node is found, or until it hits a node that has no children. Then the search backtracks, returning to the most recent node it hadn't finished exploring. In a non-recursive implementation, all freshly expanded nodes are added to a LIFO stack for exploration.

sense	commonness	relatedness
Tree	92.82%	15.97%
Tree (graph theory)	2.94%	59.91%
Tree (data structure)	**2.57%**	**63.26%**
Tree (set theory)	0.15%	34.04%
Phylogenetic tree	0.07%	20.33%
Christmas tree	0.07%	0.0%
Binary tree	0.04%	62.43%
Family tree	0.04%	16.31%
…		

Learning to Link with Wikipedia [David Milne, et al., CIKM 2008]

Public Toolkits for Entity Linking

- Wikipedia Miner
- TagMe
- DBpedia Spotlight
- Illinios Wikifier
- AIDA
- RPI Entity Linking System

Entity Linking and Retrieval for Semantic Search [Edgar Meij, et al., WSDM 2014]

$P(entity|entity)$ - Recap

- Co-occurrence
  - Within Queries
  - Across Queries
  - User Url Clicks
  - Wikipedia Pages
  - Wikipedia Categories/Templates
  - Wikipedia Revision Histories
  - Web documents
- Entity Linking
- Similarity

$P(entity|entity)$ – Similarity

- TF*IDF scores based on Wikipedia Corpus

Florence

From Wikipedia, the free encyclopedia

"Firenze" and "Florentine" redirect here. For other uses, see Florence (disambiguation), Florentin (disambiguation) (disambiguation).

Florence (/ˈflɒrəns/; Italian: *Firenze* [fiˈrɛntse] (listen), alternative obsolete form: *Fiorenza*; Latin: *Florentia*) is the capital city of the Italian region of Tuscany and of the province of Florence. It is the most populous city in Tuscany, with approximately 380,000 inhabitants, expanding to over 1,520,000 in the metropolitan area.[2]

Florence is famous for its history: a centre of medieval European trade and finance and one of the wealthiest cities of the time,[3] it is considered the birthplace of the Renaissance, and has been called "the Athens of the Middle Ages".[4] A turbulent political history includes periods of rule by the powerful Medici family, and numerous religious and republican revolutions.[5] From 1865 to 1871 the city was the capital of the recently established Kingdom of Italy.

The Historic Centre of Florence attracts millions of tourists each year, and Euromonitor International ranked the city as the world's 89th most visited in 2012, with 1.8 million visitors.[6] It was declared a World Heritage Site by UNESCO in 1982. The city is noted for its culture, Renaissance art and architecture and monuments.[7] The city also contains numerous museums and art galleries, such as the Uffizi Gallery and the Palazzo Pitti, and still exerts an influence in the fields of art, culture and politics.[8] Due to Florence's artistic and architectural heritage, it has been ranked by *Forbes* as one of the most beautiful cities in the world.[9]

Rome

From Wikipedia, the free encyclopedia

This article is about the city in Italy. For the civilization of classical antiquity, see Ancient Rome. For other uses

Rome (/ˈroʊm/, Italian: *Roma* [ˈroːma] (listen), Latin: *Rōma*) is a city and special *comune* (named "Roma Capitale") in Italy. Rome is the capital of Italy and region of Lazio. With 2.9 million residents in 1,285 km^2 (496.1 sq mi), it is also the country's largest and most populated *comune* and fourth-most populous city in the European Union by population within city limits. The Metropolitan City of Rome has a population of 4.3 million residents.[2] The city is located in the central-western portion of the Italian Peninsula, within Lazio (Latium), along the shores of Tiber river. Vatican City is an independent country within the city boundaries of Rome, the only existing example of a country within a city: for this reason Rome has been often defined as capital of two states.[3][4]

Rome's history spans more than two and a half thousand years. While Roman mythology dates the founding of Rome at only around 753 BC, the site has been inhabited for much longer, making it one of the oldest continuously occupied cities in Europe.[5] The city's early population originated from a mix of Latins, Etruscans and Sabines. Eventually, the city successively became the capital of the Roman Kingdom, the Roman Republic and the Roman Empire, and is regarded as one of the birthplaces of Western civilization. It is referred to as "Roma Aeterna" (The Eternal City) [6] and "Caput Mundi" (Capital of the World), two central notions in ancient Roman culture.

$P(entity|entity)$ – Similarity

- Entity Recommendations based on Wikipedia textual similarity

Whale	Susan Dumais	Tom Cruise
Beluga whale	C. J. van Rijsbergen	Leonardo DiCaprio
Toothed whale	W. Bruce Croft	Nicole Kidman
Killer whale	Harry Shum	Clint Eastwood
Pygmy killer whale	Gerard Salton	Mark Rathbun
Humpback whale	Jaime Teevan	L. Ron Hubbard

$P(entity|entity)$ – Similarity

- Challenges
  - Textual Similarity suffers the vocabulary mismatch problem
    - “USA” and “United States of America” are semantically equivalent, yet share no terms in common
- Solution
  - Project entities into latent space that can semantically represent the entities

$P(entity|entity)$ – Word Embedding

- Word Embedding

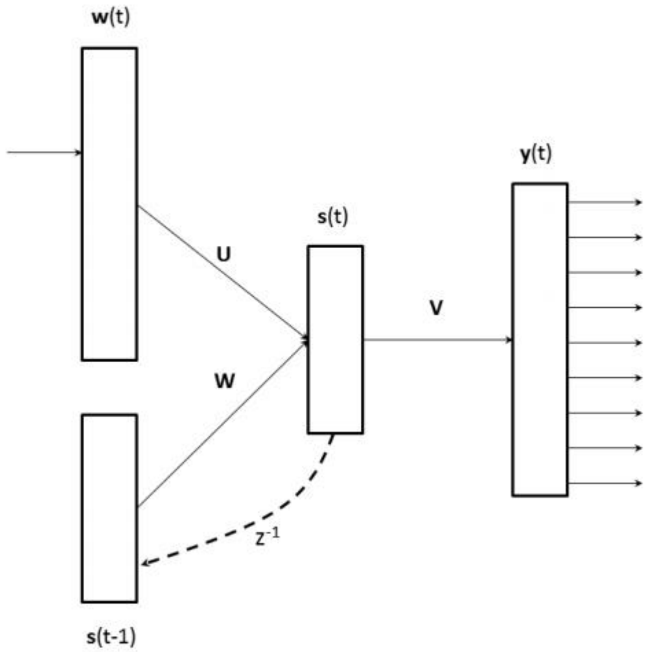


Figure 1: Recurrent Neural Network Language Model.

$$\mathbf{s}(t) = f\left(\mathbf{U}\mathbf{w}(t) + \mathbf{W}\mathbf{s}(t-1)\right)$$

$$\mathbf{y}(t) = g\left(\mathbf{V}\mathbf{s}(t)\right)$$

$$f(z) = \frac{1}{1+e^{-z}}, \quad g(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}}$$

Linguistic Regularities in Continuous Space Word Representations [Tomas Mikolov, et al., ACL 2013]

$P(entity|entity)$ – Word Embedding

- Word Embedding

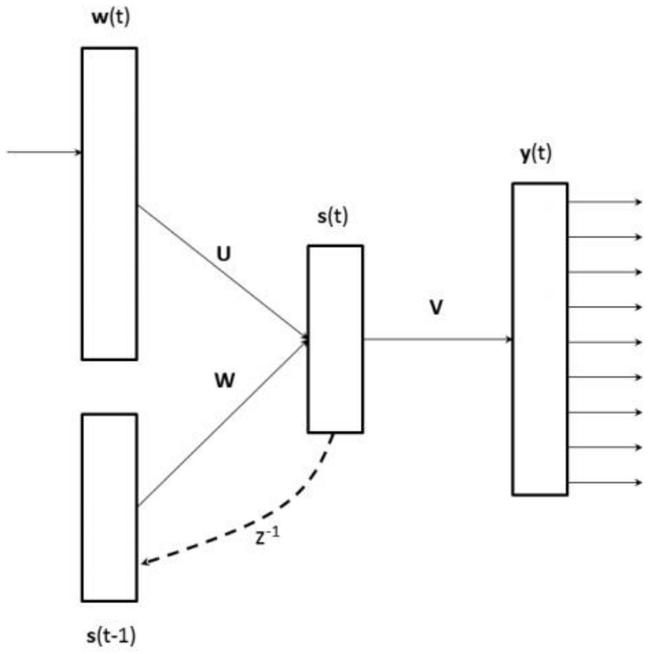
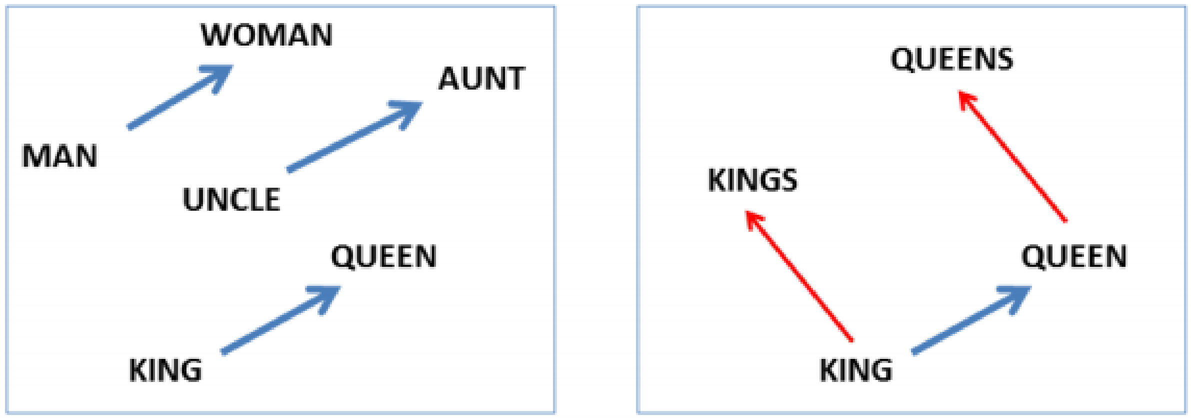


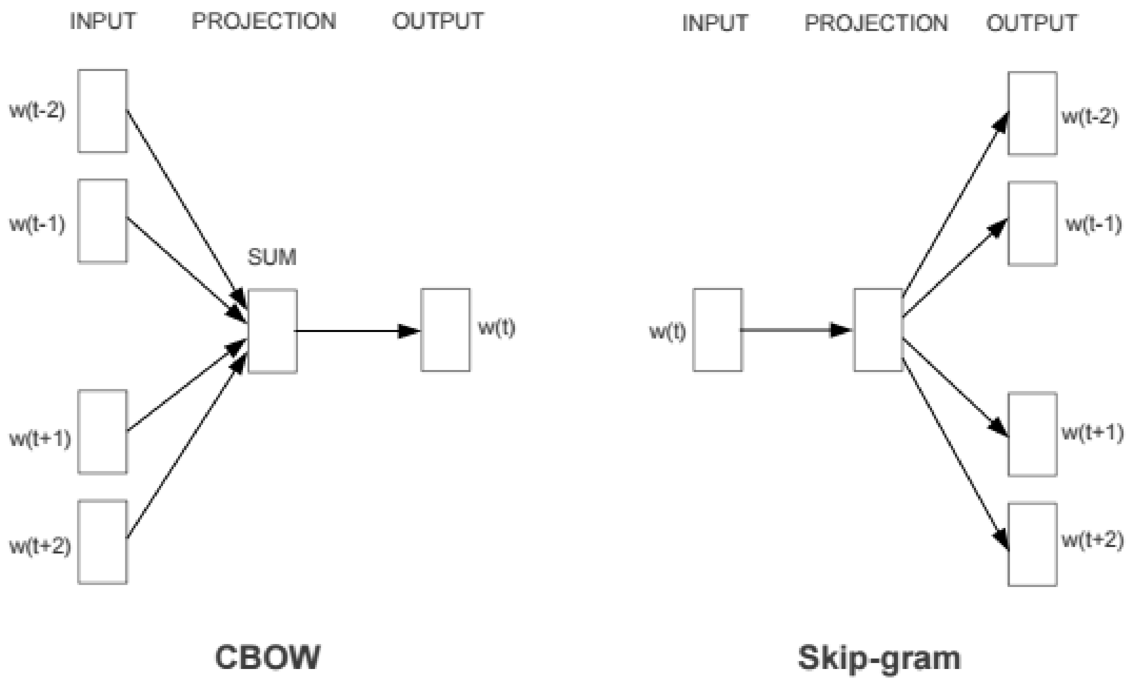
Figure 1: Recurrent Neural Network Language Model.



Linguistic Regularities in Continuous Space Word Representations [Tomas Mikolov, et al., ACL 2013]

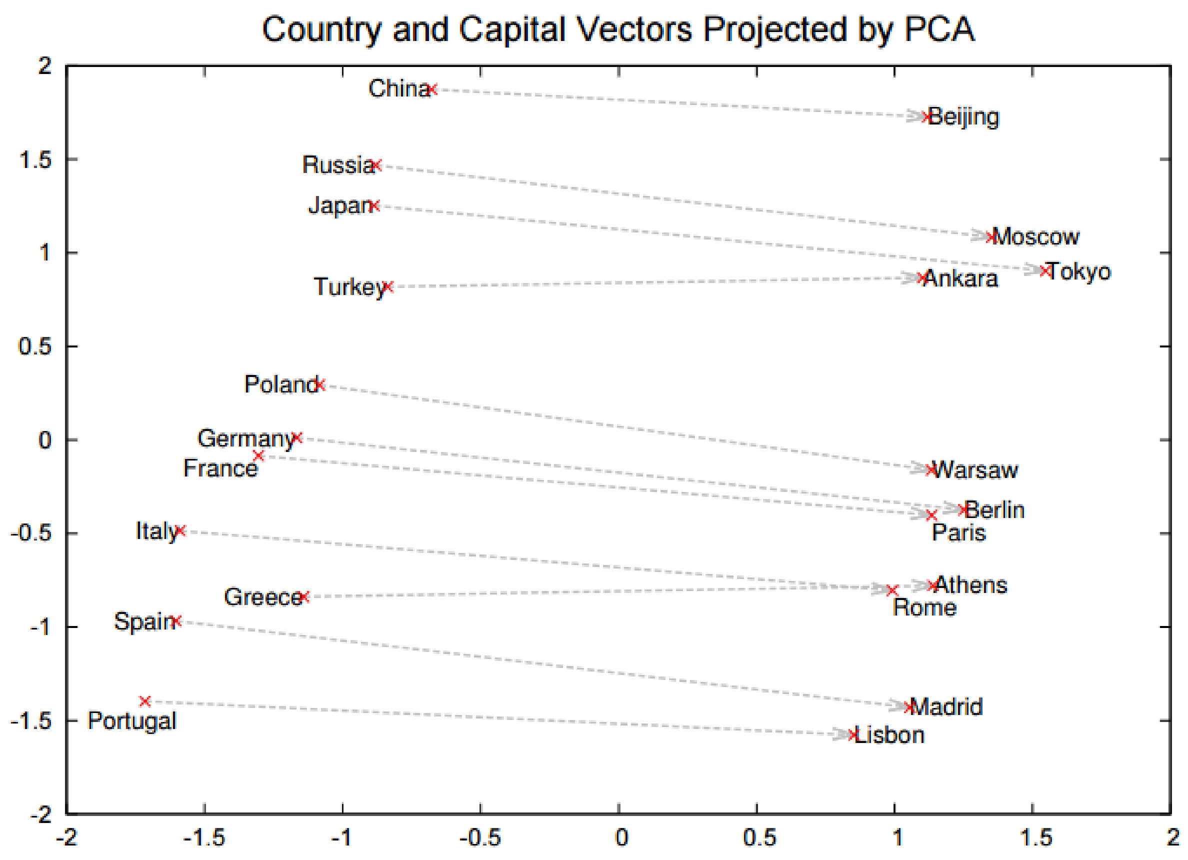
$P(entity|entity)$ – Word Embedding

- Word Embedding



Efficient Estimation of Word Representations in Vector Space [Tomas Mikolov, et al., ICLR 2013]

$P(entity|entity)$ – Word Embedding



Distributed Representations of Words and Phrases and their Compositionality [Tomas Mikolov, et al., NIPS 2013]

$P(entity|entity)$ – Word Embedding

Relationship	Example 1	Example 2	Example 3
France - Paris	Italy: Rome	Japan: Tokyo	Florida: Tallahassee
big - bigger	small: larger	cold: colder	quick: quicker
Miami - Florida	Baltimore: Maryland	Dallas: Texas	Kona: Hawaii
Einstein - scientist	Messi: midfielder	Mozart: violinist	Picasso: painter
Sarkozy - France	Berlusconi: Italy	Merkel: Germany	Koizumi: Japan
copper - Cu	zinc: Zn	gold: Au	uranium: plutonium
Berlusconi - Silvio	Sarkozy: Nicolas	Putin: Medvedev	Obama: Barack
Microsoft - Windows	Google: Android	IBM: Linux	Apple: iPhone
Microsoft - Ballmer	Google: Yahoo	IBM: McNealy	Apple: Jobs
Japan - sushi	Germany: bratwurst	France: tapas	USA: pizza

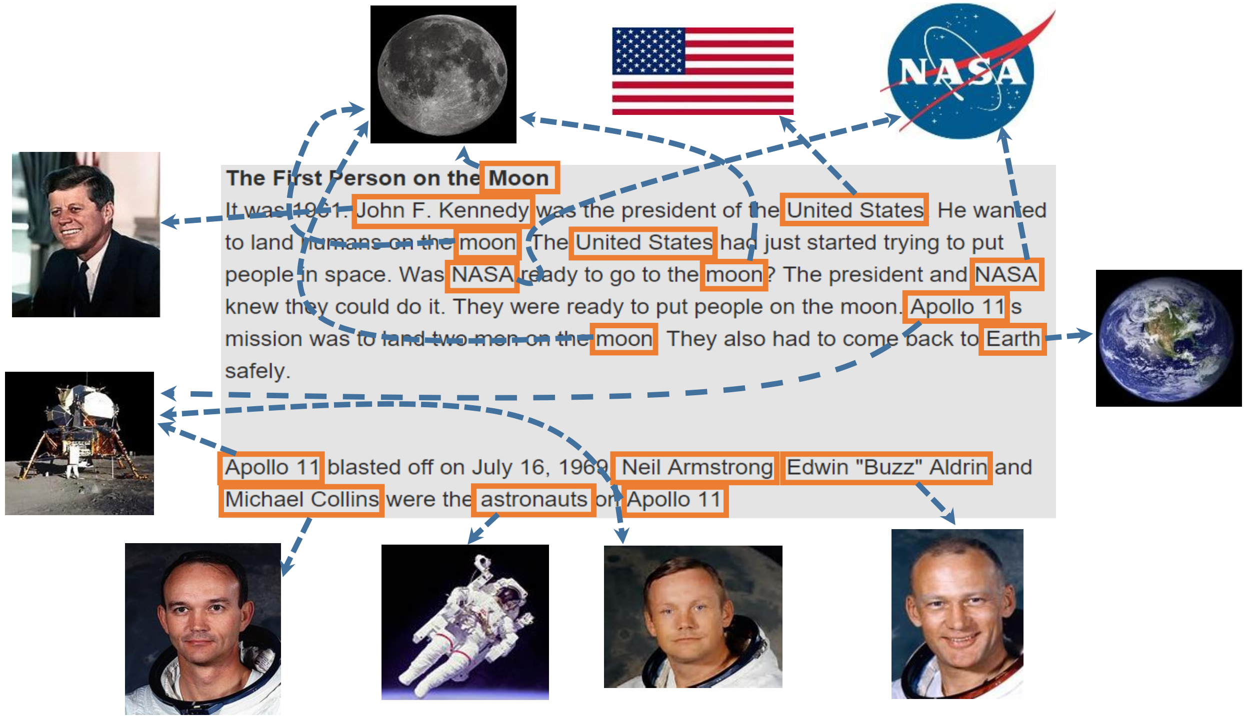
$$y = x_b - x_a + x_c$$

$$w^* = argmax_w \frac{x_w y}{||x_w|| ||y||}$$

Efficient Estimation of Word Representations in Vector Space [Tomas Mikolov, et al., ICLR 2013]

$P(entity|entity)$ – Word Embedding

- How to apply Word Embedding in Entities?
  - Perform entity linking on documents
  - Treat each entity as a single word
  - Learning the representation



Word2Vec [Tomas Mikolov, et al., ICLR 2013]

$P(entity|entity)$ – Word Embedding

- Entity Recommendations based on Word Embedding (Skip-gram)
- Trained on 100M Google News Articles

Whale	Susan Dumais	Tom Cruise
Penguin	Gary William Flake	Katie Holmes
Humpback whale	Andrei Broder	Suri Cruise
Killer whale	Bill Buxton	Nicole Kidman
Turtle	Harry Bruce	Cameron Diaz
Rat	Richard Rashid	Connor Antony

Word2Vec [Tomas Mikolov, et al., ICLR 2013]

$P(entity|entity)$

- Co-occurrence and Textual Similarity methods work well
- Textual Similarity method is very topic- or genre-related
- Word Embedding might not always work (depend on training data)

Co-occurrence

Whale	Susan Dumais	Tom Cruise
Dolphin	C. J. van Rijsbergen	Nicole Kidman
Pinniped	W. Bruce Croft	Brad Pitt
Shark	Eric Horvitz	Steven Spielberg
Killer whale	George Furnas	Tom Hanks
Humpback whale	Thomas Landauer	John Travolta

Textual Similarity

Whale	Susan Dumais	Tom Cruise
Beluga whale	C. J. van Rijsbergen	Leonardo DiCaprio
Toothed whale	W. Bruce Croft	Nicole Kidman
Killer whale	Harry Shum	Clint Eastwood
Pygmy killer whale	Gerard Salton	Mark Rathbun
Humpback whale	Jaime Teevan	L. Ron Hubbard

Word Embedding

Whale	Susan Dumais	Tom Cruise
Penguin	Gary William Flake	Katie Holmes
Humpback whale	Andrei Broder	Suri Cruise
Killer whale	Bill Buxton	Nicole Kidman
Turtle	Harry Bruce	Cameron Diaz
Rat	Richard Rashid	Connor Antony

$P(entity|entity)$ - Recap

- Co-occurrence
  - Within Queries
  - Across Queries
  - User Url Clicks
  - Wikipedia Pages
  - Wikipedia Categories/Templates
  - Wikipedia Revision Histories
  - Web documents
- Entity Linking
- Similarity
  - Textual Similarity
  - Word Embedding
- Interpretation

$P(entity|entity)$

Florence Cathedral

Church

The Cattedrale di Santa Maria del Fiore is the main church of Florence, Italy. Il Duomo di Firenze, as it is ordinarily called, was begun in 1296 in the Gothic style to the design of Arnolfo di Cambio and completed structurally in 1436 with the dome ... +

en.wikipedia.org

Opened: Mar 25, 1436

Height: 376 feet (114.50 m)

Architects: Filippo Brunelleschi · Giotto · Arnolfo di Cambio · Francesco Talenti

Architectural styles: Renaissance architecture · Gothic architecture · Italian Gothic architecture · Gothic Revival architecture

Categories: Basilica · Cathedral · Minor basilica · Church

Burials: Filippo Brunelleschi · Conrad II of Italy · Niccolò da Tolentino

Related people — See all (10+)

Filippo Brunelleschi
Architect

Michelang...

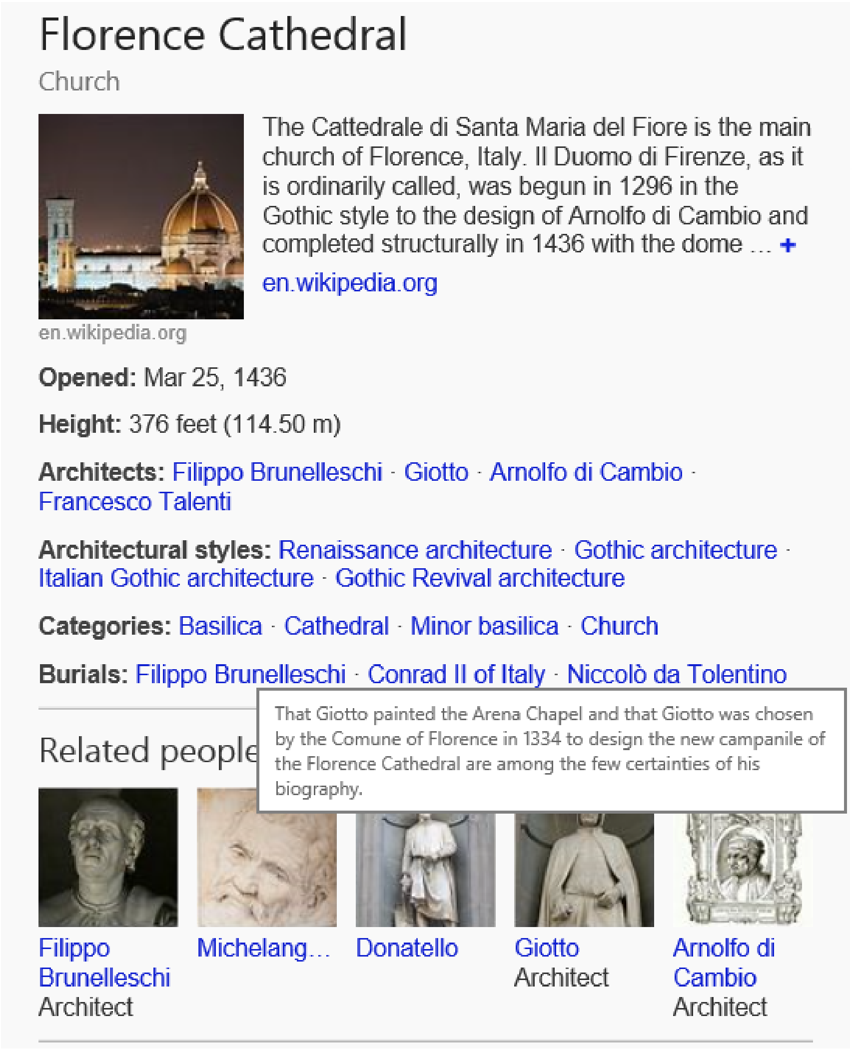
Donatello

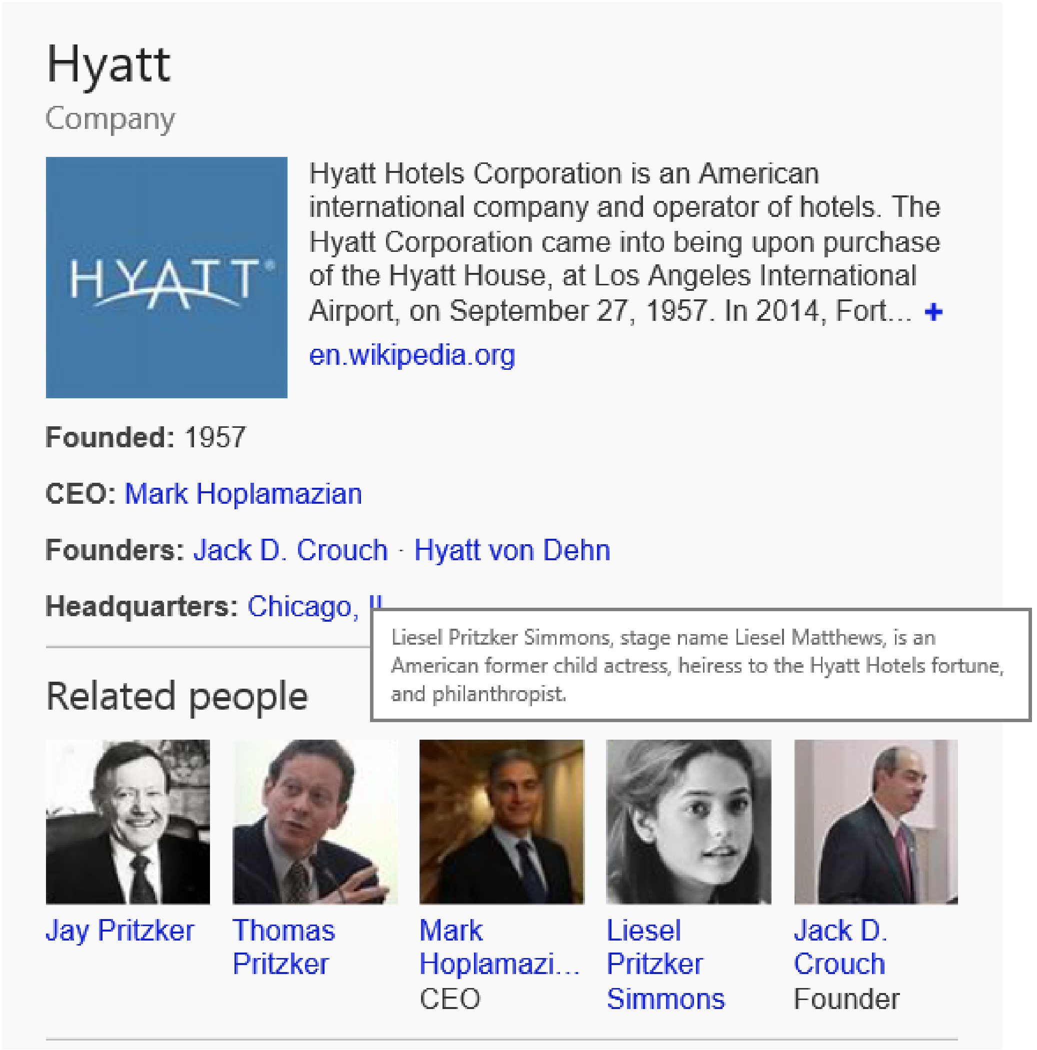
Giotto
Architect

Arnolfo di Cambio
Architect

Why are they related to “Florence Cathedral”?

$P(entity|entity)$ - Interpretation





$P(entity|entity)$ - Interpretation

- Problem definition

	Notation	Explanation
Input	e_a	the first entity of the entity pair.
	e_b	the second entity of the entity pair.
	r	the relation of interest between e_a and e_b.
Output	S	a set of candidate sentences possibly referring to e_a and e_b.

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Entity text representation

Wikipedia Anchor:
President obama

Wikipedia Title:
Barack Obama

Wikipedia Redirect:
Obama

Wikipedia Anchor:
Clinton

Wikipedia Title:
Bill Clinton

Wikipedia Redirect:
President clinton

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Candidate Sentences
  - On the wiki pages of “Barack Obama” and “Bill Clinton”
  - Keep those sentences that contains at least one entity’s text representations

Wikipedia Anchor: President obama

Wikipedia Title: Barack Obama

Wikipedia Redirect: Obama

Wikipedia Anchor: Clinton

Wikipedia Title: Bill Clinton

Wikipedia Redirect: President clinton

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Sentences enrichment
  - Perform Co-reference resolution
  - Replace detected strings with the entity text representations
  - Examples:
    - "He" → Barack Obama
    - "The company" → Toyota
    - "The film" → Titanic

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Sentences enrichment
  - Perform Entity Linking on Wikipedia articles

In response to the 2010 Haiti earthquake, U.S. President $\boxed{\text{Barack Obama}}$ announced that $\boxed{\text{Clinton}}$ and George W. Bush would coordinate efforts to raise funds for Haiti's recovery.

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Ranking sentences
  - Generate features and using Learning to Rank algorithms to rank sentences
- Features
  - **Text features:** Average IDF of terms of s in Wikipedia; Number of terms in s; Part of Speech distribution of s; etc.
  - **Entity features:** Number of entities in s; Whether s contains links to both e_a and e_b; Distance between e_a and e_b in s; Number of entities between e_a and e_b; etc.
  - **Relation features:** Whether s contains any term of r (binary); Average score of phrases in word2vec(r) that are matched in s; etc.
  - **Source features:** Position of s in document d; etc.

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

(#1) Ben Affleck - Bruce Willis (MovieActor_CoCastsWith_MovieActor)

Affleck starred in "Armageddon" (1998) opposite Bruce Willis.

(#2) Hugh Jackman - Kate Winslet (MovieActor_CoCastsWith_MovieActor)

Katie Finneran's most recent film was "Movie 43" in which she played Angie and also appeared alongside Hugh Jackman and Kate Winslet.

(#3) Bryan Singer - Tom Cruise (MovieDirector_Directs_MovieActor)

The film stars Tom Cruise and is directed by Bryan Singer.

(#4) Cameron Diaz - Tom Cruise (MovieActor_CoCastsWith_MovieActor)

The following year Cruise starred in the romantic thriller "Vanilla Sky" (2001) with Cameron Diaz and Penélope Cruz.

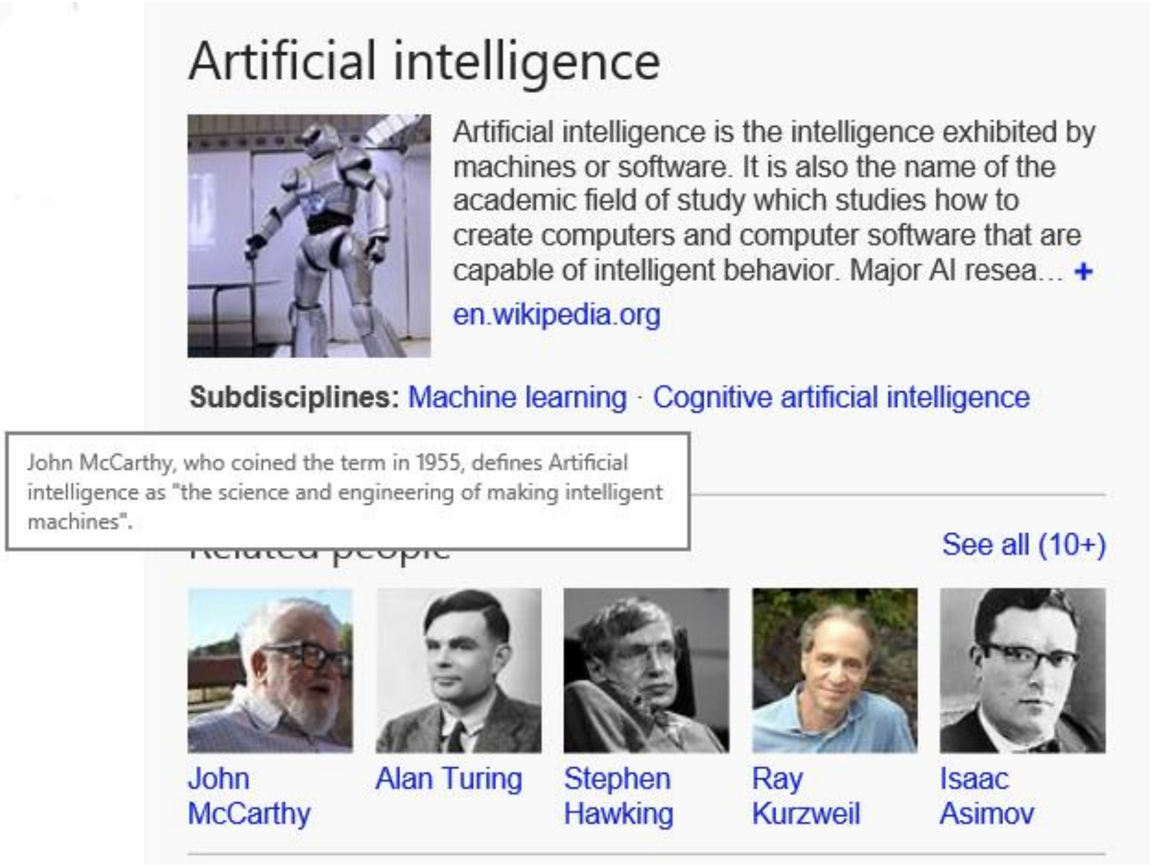
(#5) Cristiano Ronaldo - Karim Benzema (Athlete_PlaysSameSportTeamAs_Athlete)

Karim Benzema was also shortlisted by the French magazine France Football for the 2008 Ballon d'Or award, won by Cristiano Ronaldo.

Explaining Relationships Between Entities [Nikos Voskarides]

$P(entity|entity)$ - Interpretation

- Challenges
  - Relationships are missing or unknown in the real world scenarios



$P(entity|entity)$ - Interpretation

- Challenges
  - The selected sentences should be more "interesting" instead of just replicating the relationships

(#1) Ben Affleck - Bruce Willis (MovieActor_CoCastsWith_MovieActor)
Affleck starred in "Armageddon" (1998) opposite Bruce Willis.

Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$
  - Recommendations given an entity
    - Co-occurrence
    - Similarity
    - Entity Linking
    - Interpretation
- $P(entity|user)$
  - Recommendations given a user
- $P(entity|query)$
  - Recommendations given a query

Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$
  - Recommendations given an entity
    - Co-occurrence
    - Similarity
    - Entity Linking
    - Interpretation
- $P(entity|user)$
  - Recommendations given a user
- $P(entity|query)$
  - Recommendations given a query

Personalized Recommender Systems

- Problem definition
  - User-Item Matrix

Rating/Frequency

	v_1	v_2	v_3	v_4	v_5
u_1	1		2	3	
u_2		3			1
u_3		4		5	
u_4	5			4	
u_5		2	5		4

Implicit/One-class

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		1	1		1	
u_2	1			1		1
u_3			1			1
u_4	1			1		
u_5		1	1			1

Personalized Recommender Systems

- Memory-based methods
  - Pearson Correlation Coefficient
  - Vector Space Similarity/Cosine Similarity
- Model-based methods
  - Matrix Factorization
  - Probabilistic models
  - Clustering
  - Classification
  - …

Personalized Recommender Systems

- Memory-based methods
  - Pearson Correlation Coefficient
  - Vector Space Similarity/Cosine Similarity
- Model-based methods
  - Matrix Factorization
  - Probabilistic models
  - Clustering
  - Classification
  - …

Personalized Recommender Systems

- Memory-based methods

Users	Items												
u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Memory-based methods

Items

Users

u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Memory-based methods

Items

Users

u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Memory-based methods

Items

Users

u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Memory-based methods

Items

Users

u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Memory-based methods

Items

Users

u_1													
u_2	1	3		4		2		5			3	4	
u_3													
u_4		3		4			3	4		3	4		4
u_5													
u_6	1			3	5	2		4	1			3	

Personalized Recommender Systems

- Pearson Correlation Coefficient

$$Sim(a,u) = \frac{\sum_{i \in I(a) \cap I(u)} (r_{a,i} - \overline{r}_a) \cdot (r_{u,i} - \overline{r}_u)}{\sqrt{\sum_{i \in I(a) \cap I(u)} (r_{a,i} - \overline{r}_a)^2} \cdot \sqrt{\sum_{i \in I(a) \cap I(u)} (r_{u,i} - \overline{r}_u)^2}}$$

GroupLens: an open architecture for collaborative filtering of netnews [Paul Resnick, et al., CSCW 1994]

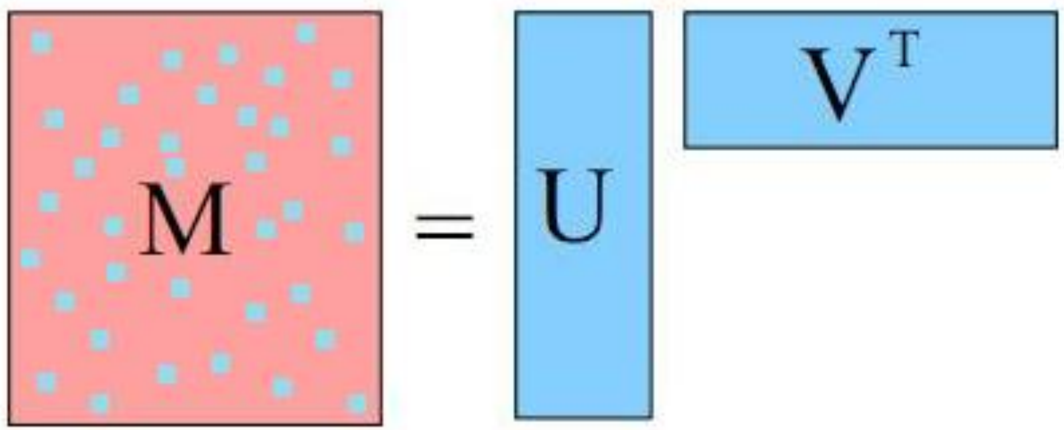
Personalized Recommender Systems

- Challenges on using Memory-based methods
  - Sparsity issue
    - User-item matrix is normally very sparse, and the density is normally under 1%
    - The sparsity issue will make estimating user similarity difficult and inaccurate
  - Scalability
    - Finding nearest neighbors require computation that grows with both the number of users and the number of items

Personalized Recommender Systems

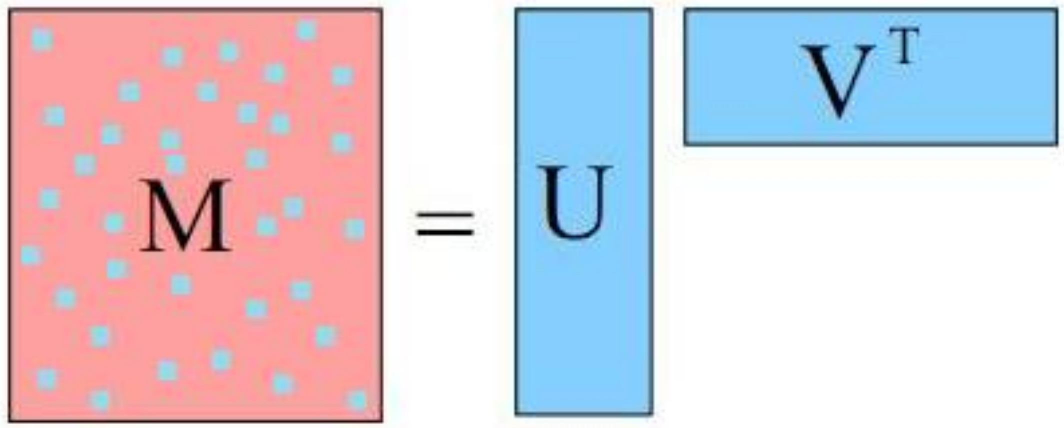
- Matrix Factorization

$$M_{[m\times n]} = U_{[m\times r]} V^T_{[n\times r]}$$



Personalized Recommender Systems

- Matrix Factorization
  - Adding bias
  - $b_{ij} = \mu + b_i + b_j$
  - Predicting rating as
  - $r_{ij} = b_{ij} + u_i v_j^T$

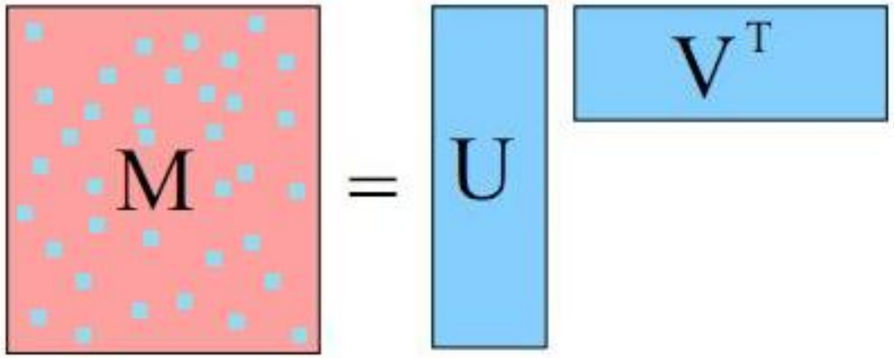


- Matrix factorization method is the single most effective method in the Netflix 1M prize challenge

The Recommender Problem Revisited [Xavier Amatriain, et al., KDD 2014]

Personalized Recommender Systems

- Matrix Factorization
  - What about implicit user-item matrix?
  - Negative Sampling



Implicit/One-class

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		1	1		1	
u_2	1			1		1
u_3			1			1
u_4	1			1		
u_5		1	1			1

Negative Sampling

	v_1	v_2	v_3	v_4	v_5	v_6
u_1	0	1	1		1	
u_2	1	0		1		1
u_3			1		0	1
u_4	1		0	1		
u_5	0	1	1	0		1

One-Class Collaborative Filtering [Rong Pan, et al., ICDM 2008]

An Incomplete List of Academic Papers on RS

- Breese et al. “Empirical analysis of predictive algorithms for collaborative filtering. ” UAI-1998.
- Deshpande et al. “Item-based top-n recommendation.” ACM TOIS-2004.
- Herlocker et al. “An algorithmic framework for performing collaborative filtering.” SIGIR-2009.
- Hofmann et al. “Collaborative filtering via gaussian probabilistic latent semantic analysis.” SIGIR-2003.
- Koren et al. “Factorization meets the neighborhood: a multifaceted collaborative filtering model.” KDD-2008.
- Koren et al. “Collaborative filtering with temporal dynamics.” KDD-2009.
- Koren et al. “Matrix factorization techniques for recommender systems.” IEEE Computer-2009.
- Linden et al. “Amazon.com recommendations: Item-to-item collaborative filtering.” Internet Computing-2003.
- Rennie et al. “Fast maximum margin matrix factorization for collaborative prediction.” ICML-2005.
- Salakhutdinov et al. “Probabilistic matrix factorization.” NIPS-2007.
- Salakhutdinov et al. “Bayesian probabilistic matrix factorization using markov chain monte carlo.” ICML-2008.
- Sarwar et al. “Item-based collaborative filtering recommendation algorithms.” WWW-2001.
- Si et al. “Flexible mixture model for collaborative filtering.” ICML-2003.
- Srebro et al. “Weighted low-rank approximations.” ICML-2003.

Why is Entity Recommender System different?

- The problem definition for traditional recommender systems is clearer than entity recommender systems since most of the time user-item matrix is given

Rating/Frequency

	v_1	v_2	v_3	v_4	v_5
u_1	1		2	3	
u_2		3			1
u_3		4		5	
u_4	5			4	
u_5		2	5		4

Implicit/One-class

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		1	1		1	
u_2	1			1		1
u_3			1			1
u_4	1			1		
u_5		1	1			1

Why is Entity Recommender System different?

- Entity recommender systems are embedded within searching process, users' preferences on entities are more difficult to observe

Users

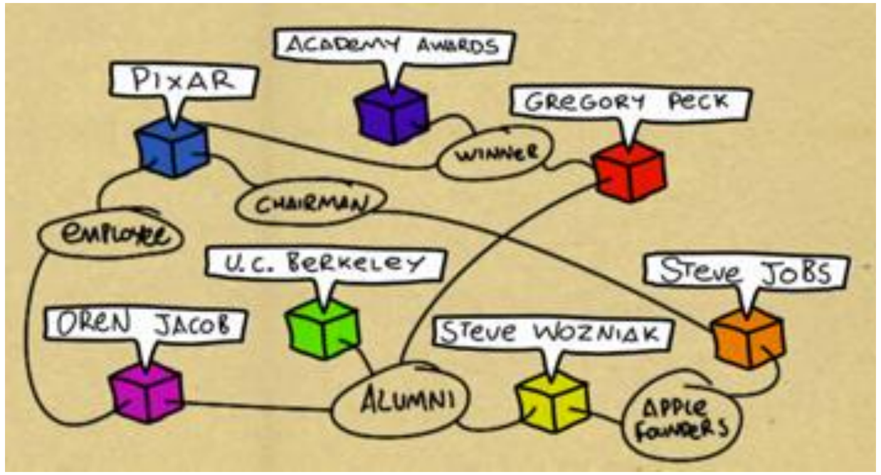
Queries

Clicks

Why is Entity Recommender System different?

- The huge Entity Graph with knowledge makes entity recommender systems more challenging and also appealing

Entities



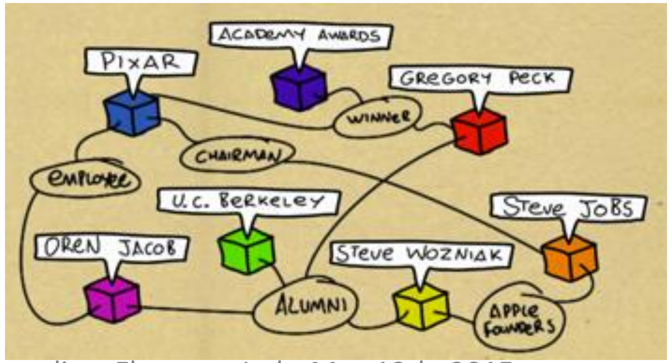
$P(entity|user)$

Queries

Clicks

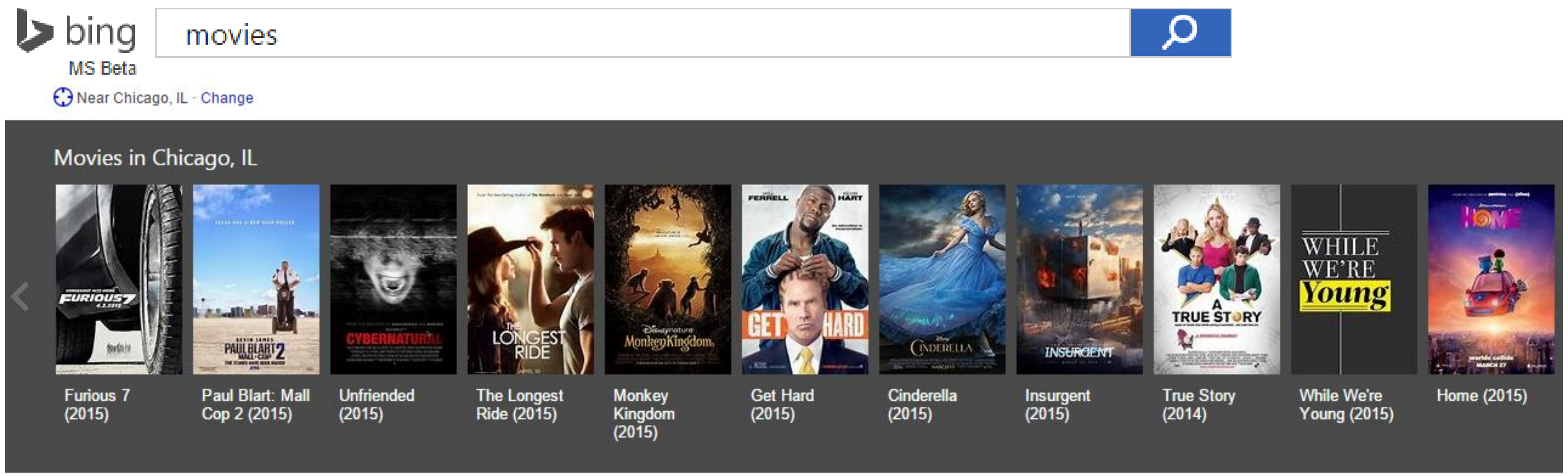
Users

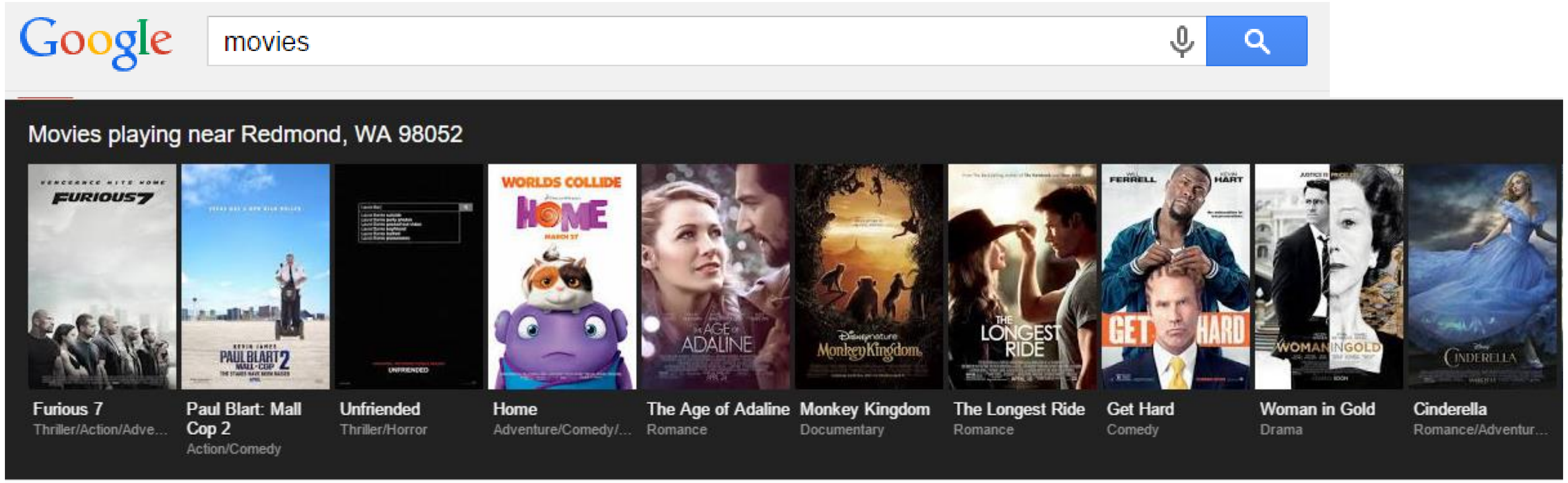
Entities



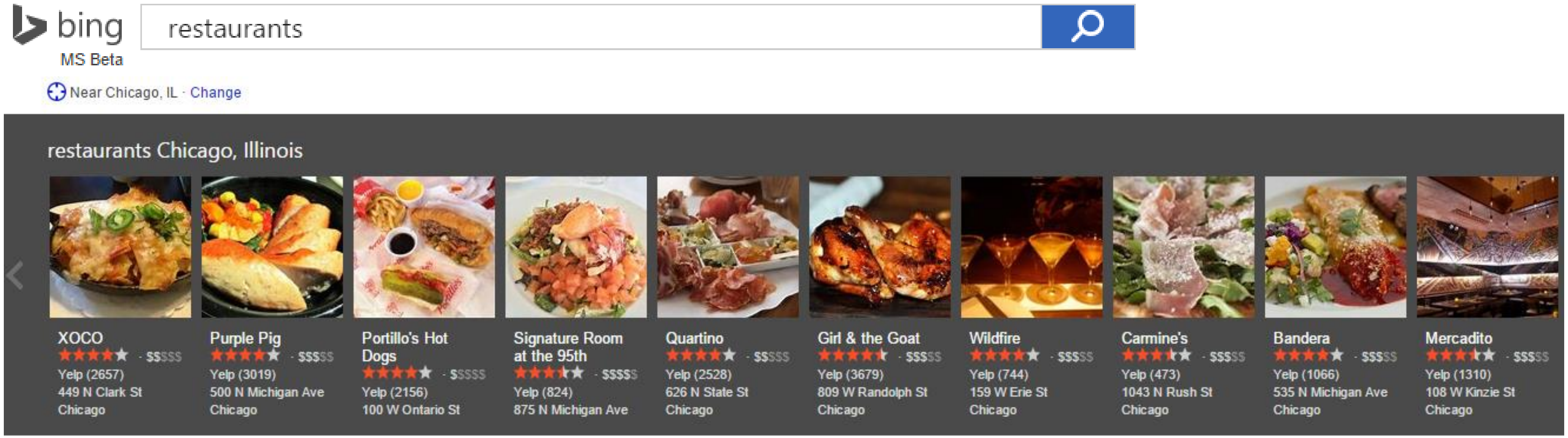
Universal
Recommender Systems

Current Entity Experience





Current Entity Experience



$P(entity|user)$

bing movies

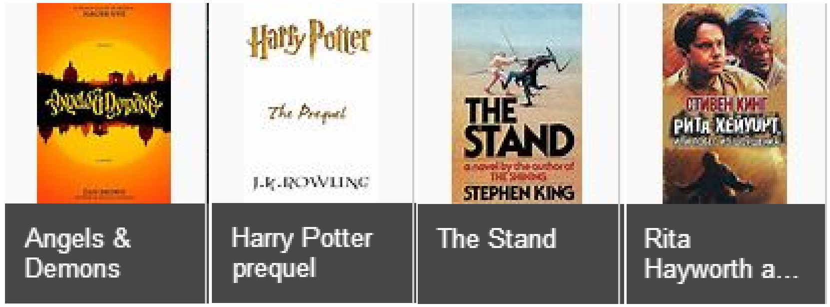
Movie recommendation for you

bing restaurants

Restaurant recommendation for you

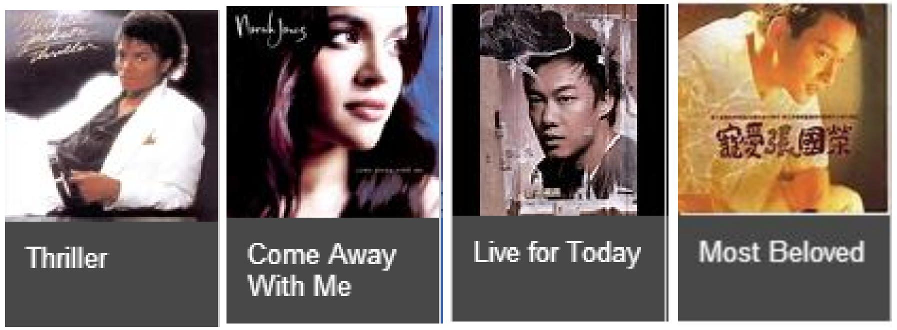
bing books

Book recommendation for you



bing music albums

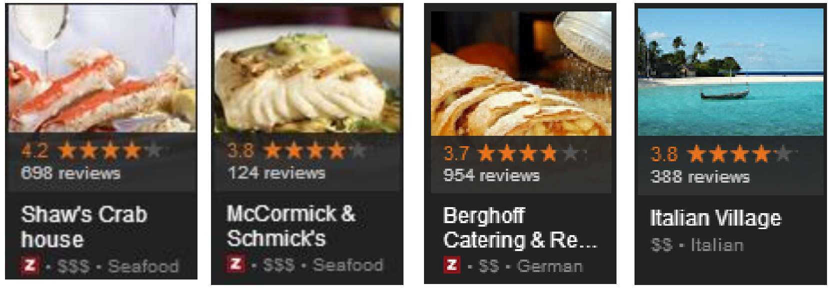
Music album recommendation for you



$P(entity|user)$

movie recommendation for you

restaurant recommendation for you

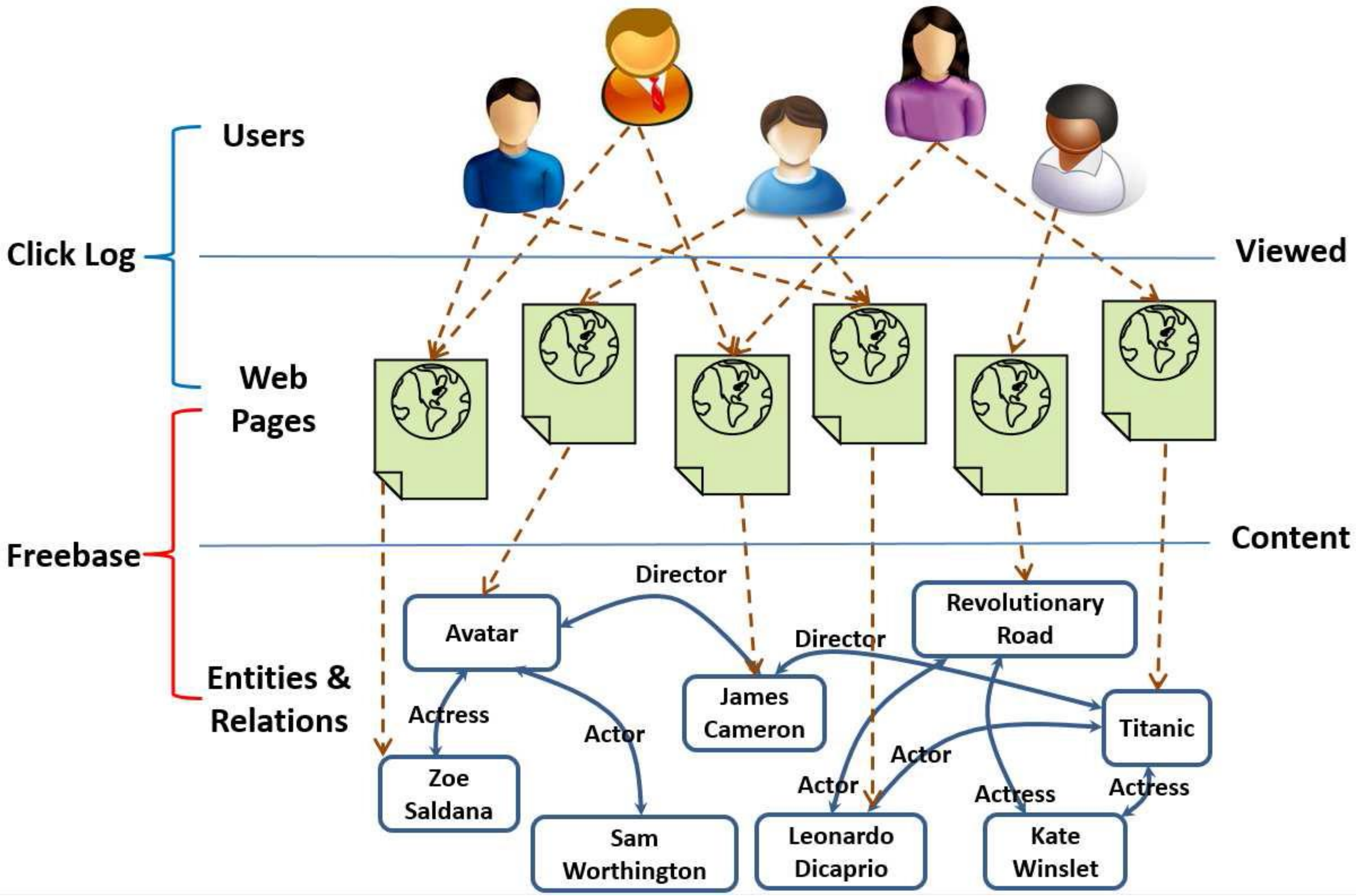


point of interest recommendation for you

event recommendation for you

dec 5^{th} dec 13^{th} dec 6^{th} dec 5^{th}

User Logs and Entity Graph



On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

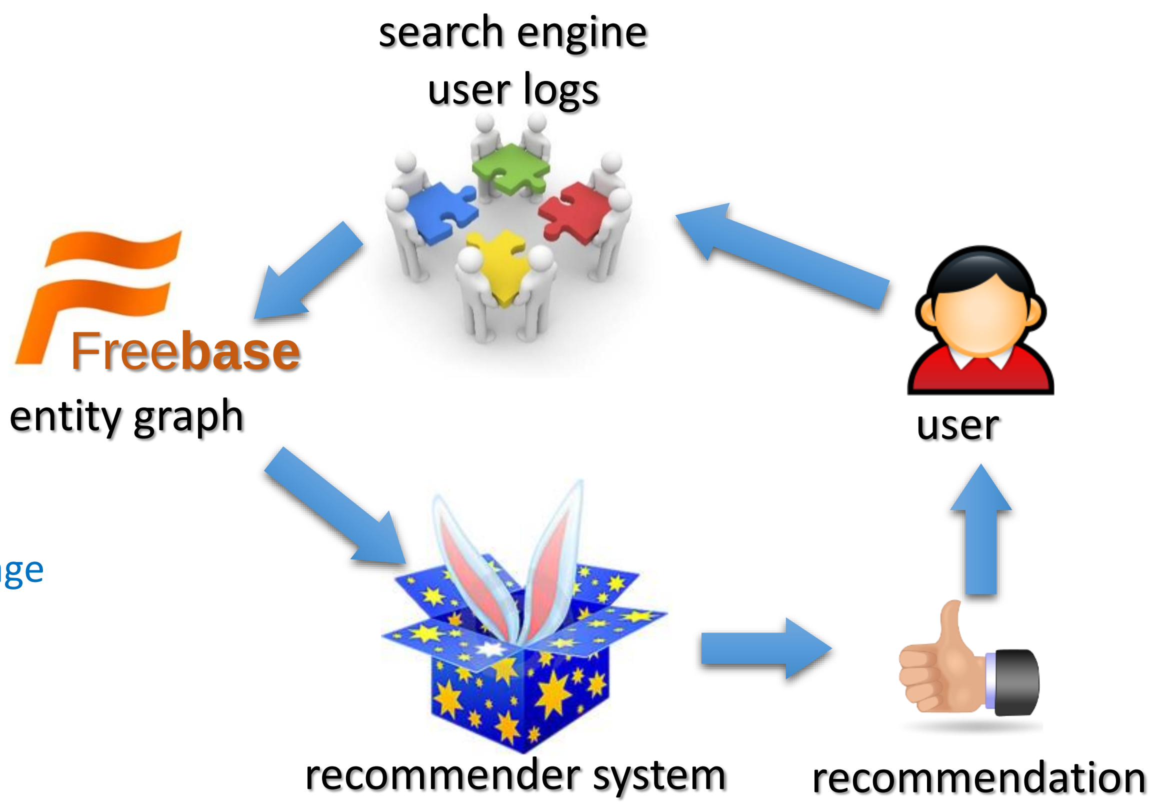
Problem Definition

- Consider entity related web pages
- User log sequence, sorted by timestamp

$$< e_1^u, e_2^u, \ldots, e_t^u, \ldots, e_{T-1}^u, e_T^u >$$ target entity page

user log sequence before T, denotes as L_T^u

- Use L_T^u to predict e_T^u



On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Benefits in Using User Logs and Entity Graph

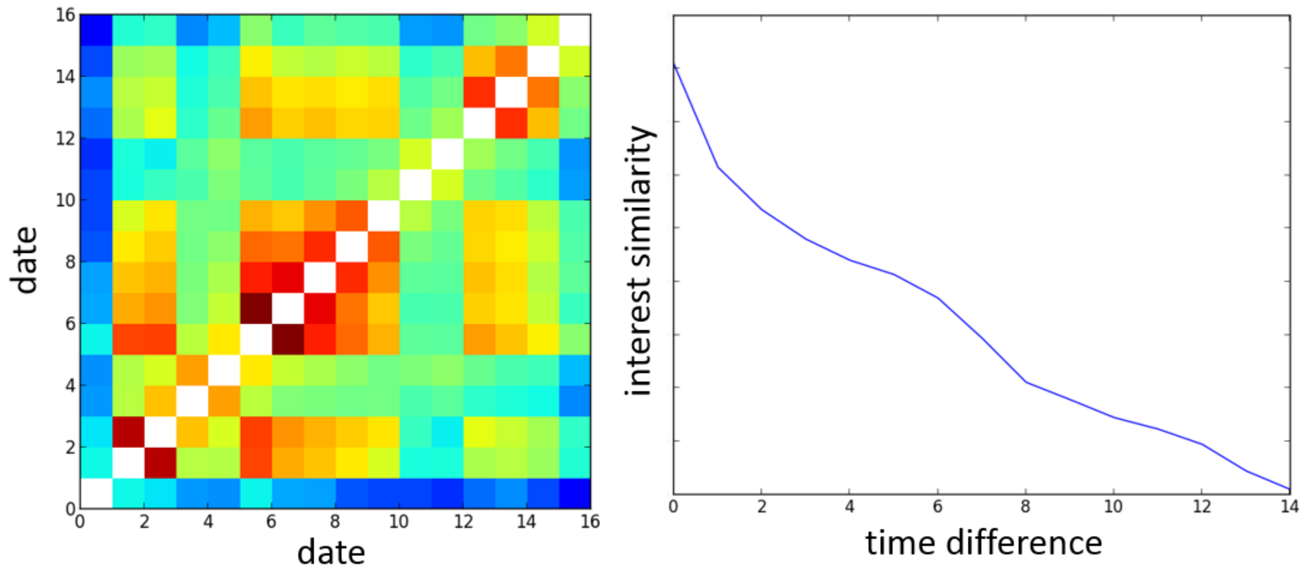
- Besides implicit feedback, user log also has
  - e.g., timestamp, dwell time, user country or region, time of day …
- Cross domain user log events
  - Which movies will foxnews.com readers like?
- Besides entity relationship, entity graph also has
  - e.g., movie release date, tagline, running time, gross revenue, budget, MPAA rating, text description, number of ratings, …

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Exploring data - User interests drift

User interests are consistent within a short time period but drift over time

The longer the time interval is, the less similar users' interests are



Left: accumulated user interest similarity for two weeks

Right: averaged user interest similarity with relative time difference (number of day)

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Exploring data - Cross Domain Correlation

comicbookresources.com (comic books)	ruelala.com (women shopping)
The Avengers	Magic Mike
Spider-Man	The Avengers
The Dark Knight Rises	Prometheus
Prometheus	Moonrise Kingdom
Men In Black 3	Ted
Iron Man 2	Snow White & Huntsman
Superman: The Man Of Steel	Savages
Thor	Hunger Games
Snow White & Huntsman	Rock of Ages
Battleship	The Best Exotic Marigold Hotel

Top 10 most viewed movies estimated using cross domain correlation

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Entity Pairwise Features

Meta-path based entity similarity from Freebase

$S(e_i, e_j)$

Other pairwise or non-pairwise feature (e.g., popularity)

Entity attribute similarity (e.g., movie description)

Entity relations derived from user log (e.g., conditional probability)

Cross domain user log events correlation

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Entity Pairwise Features

Representative Features

Entity Graph Path Features	
movie–actor–movie	movies with the same actors
movie–director–movie	movies with the same directors
movie–producer–movie	movies with the same producers
movie–star–movie	movies with the same stars
movie–writer–movie	movies with the same writers
movie–genre–movie	movies with the same genres
movie–language–movie	movies with the same language
Entity Graph Binary Features	
is_prequel	*movie*1 is a prequel of *movie*2
is_sequel	*movie*1 is a sequel of *movie*2
actor–movie	actor appears in the movie
director–movie	director directs the movie
producer–movie	producer produces the movie
Entity Graph Content Features	
release date	two movie with close release dates
description similarity	text similarity in movie descriptions
User Log Features	
co-click	conditional probability between entities
global popularity	movie popularity of all time
local popularity	movie popularity today
cross-domain	cross-domain correlation

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Recommendation Models

- Global recommendation model

entity at timestamp T

$$r(\boxed{\hat{e}_T^u}; L_T^u, \theta) = \boxed{\sum_{e_t^u \in L_T^u}} \boxed{w_t(\hat{e}_T^u, e_t^u)} \sum_k \theta_k \boxed{S_k(\hat{e}_T^u, e_t^u)}$$

events happened before T

$L_T^u = < e_1^u, e_2^u, \ldots, e_t^u, \ldots, e_{T-1}^u >$

time decay function

$w_t(e_T, e_t) = \beta e^{-\alpha(T-t)}$

pairwise features

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Recommendation Models

- Personalized recommendation model **(PRM)**

$$r(\hat{e}_T^u; L_T^u, \theta_T^u) = \sum_{e_t^u \in L_T^u} w_t(\hat{e}_T^u, e_t^u) \sum_k S_k(\hat{e}_T^u, e_t^u) \boxed{\theta_{T,k}^u}$$

for each user at target timestamp T

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Recommendation Models

- Personalized Recommendation with K-NN **(PRM+KNN)**

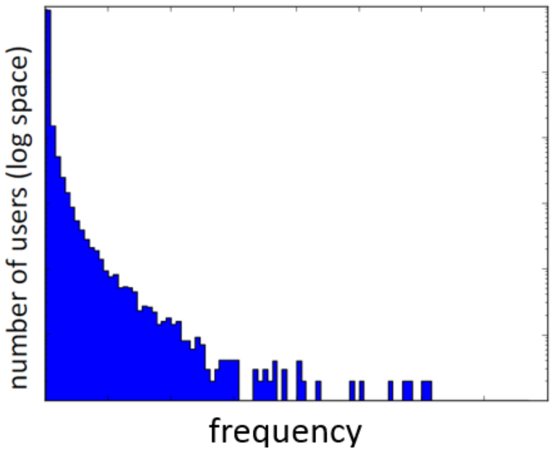
$$r_n(\hat{e}^u_t; L^u_T, \theta^u_T) = r(\hat{e}^u_t; L^u_T, \theta^u_T) + \lambda_1 \underbrace{\boxed{\sum_{L^{u\prime}_{T\prime} \in N(L^u_T)}}}_{\text{neighbor subsequences}} \underbrace{\boxed{w_u(L^{u\prime}_{T\prime}, L^u_T)}}_{\text{neighbor similarity}} r(\hat{e}^u_t; L^{u\prime}_{T\prime}, \theta^u_T)$$

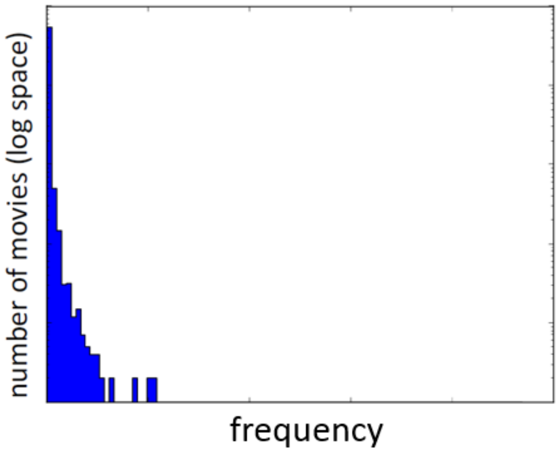
Collaborative filtering

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Experiment Setup

- Movie recommendation with search engine user log and movie related freebase knowledge graph
  - Sampled 1+ million users with at least one movie entity
  - 2+ million movie related entities with attributes and relationships, including movies, actors/actresses, directors, producers, etc.





On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Evaluation

- Top 10 Mean Reciprocal Rank (MRR) as evaluation metric

$$MRR = \frac{1}{|Test|} \sum_{i=1}^{|Test|} \frac{1}{rank_i}$$

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Comparison

Comparison Method	Description
Global Popularity	Frequently visited movies in 3 months
Local Popularity	Frequently visited movies in short time period
Domain Co-Click	Recommend based on non-movie related events
Matrix Factorization	Implicit feedback factorization
Co-Click with Time Decay	Conditional probability of events

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Performance

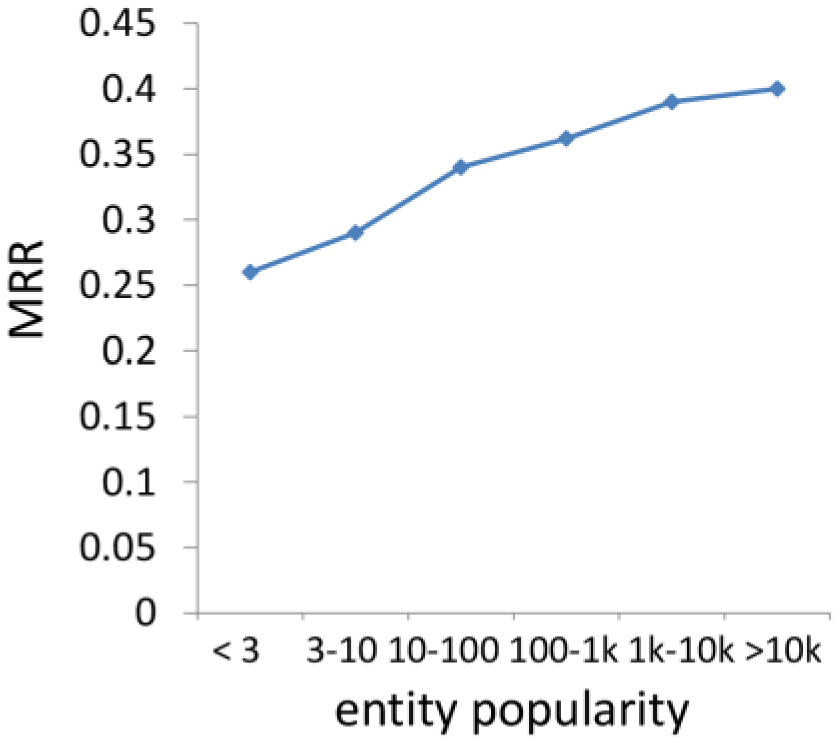
	Method	MRR
comparison methods	Global Popularity	0.024
	Local Popularity	0.043
	Domain Co-click	0.039
	Matrix Factorization	0.160
	Co-Click Time Decay	0.340
proposed methods	Global Model	0.354
	PRM	0.361
	PRM+KNN	0.451

- PRM+KNN utilizes neighbor information when recommending (CF)

On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Analysis – Entity Popularity

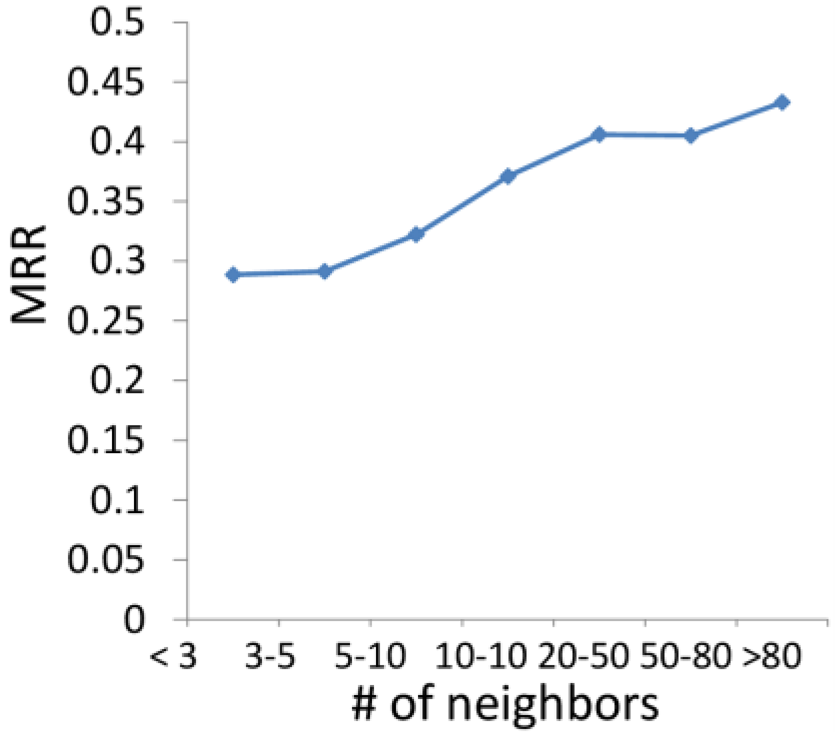
- Popular movies are easier to predict than other movies
  - Features in recommendation models favor popular movies, e.g., global and local popularity
  - With sufficient training data for popular movies, high quality recommendation models can be learned in such scenarios



On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Analysis – Number of Neighbors

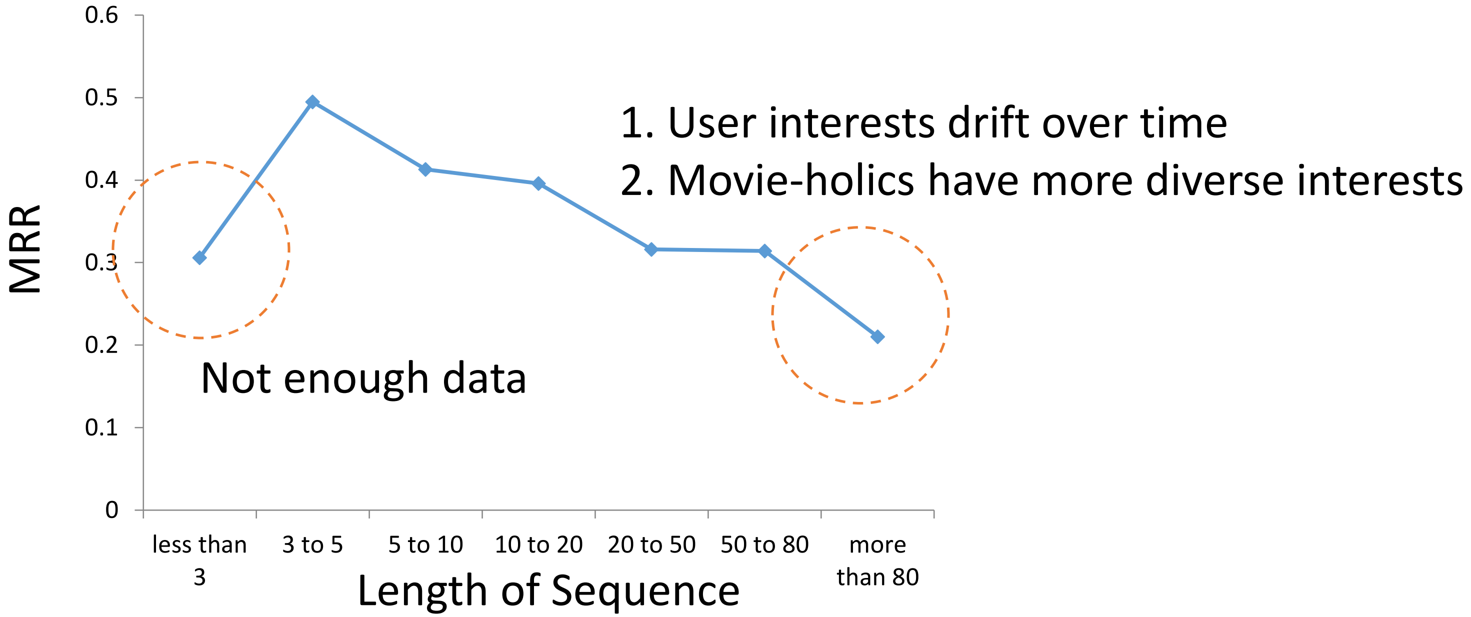
- The more neighbors each user log sequence has, the better the results
  - More neighbors indicate more data during parameter estimation
  - User log sequences with more neighbors, are usually associated with popular movie entities



On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

Analysis – Length of User Log Sequence

- Performance varies with sequence length



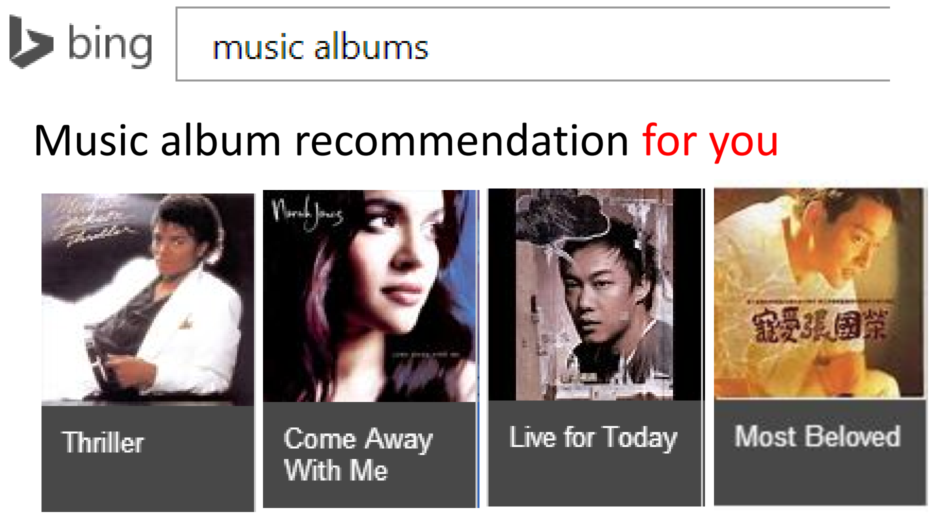
On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge [Xiao Yu, et al., WSDM 2014]

$P(entity|user)$ – Conclusion

- It is possible to build a universal recommender system on top of any search engines
- The heterogeneous information in the entity graph can be very helpful in improving the recommendation results

Entity Personalization

- Entity recommender systems - $P(entity|user)$



- What are other possible entity personalization experiences that are fundamentally different

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Definitions

A User

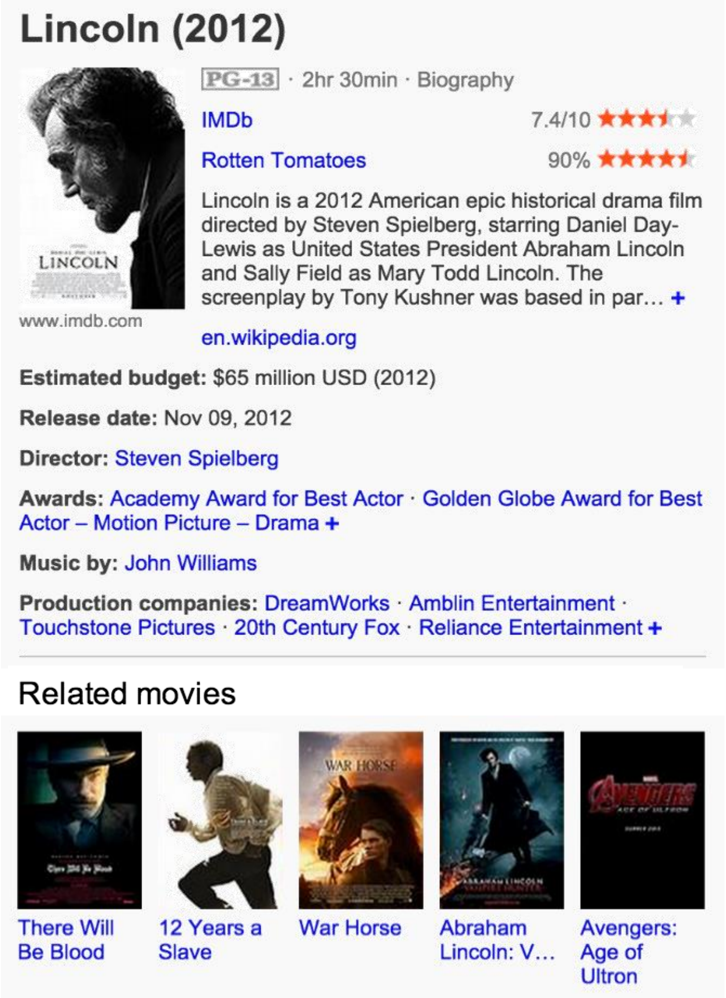
Main Entity

Related Entities

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Related Entities

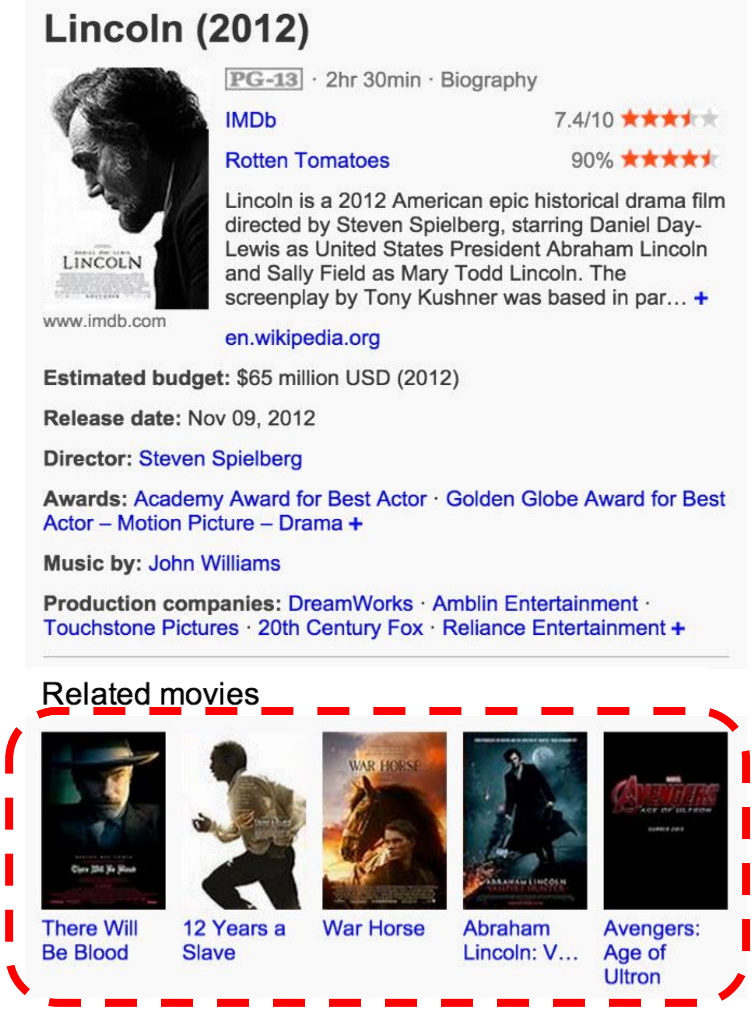
- Current recommender:
  - $P(entity|entity)$
  - $P(Other\ Movies|Lincoln)$
- User-specific information is completely ignored



Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

User's information needs are diverse

- Given "Lincoln" movie, a user may be interested in
  - Movies directed by "Steven Spielberg", or
  - Movies starred by "Daniel Day-Lewis", or
  - Movies related to "Abraham Lincoln ", or
  - Biographical movies, or
  - Civil War movies, or
  - ……

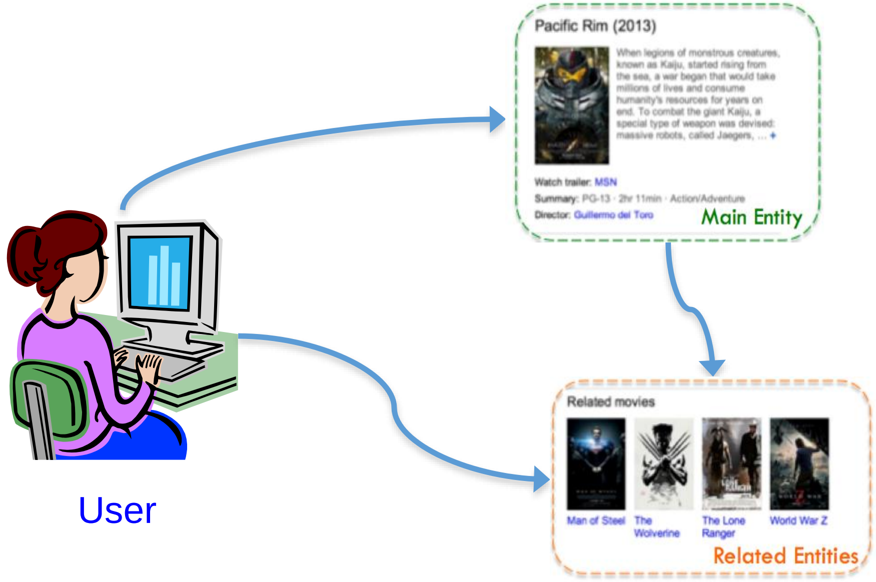


Personalizing Related Entities

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Related entity personalization

- **Goal**: Given a main entity, we aim to recommend a list of related entities based on the search user's interest.
- Three important dimensions are involved:



New Paradigm of Recommender Systems

$$P(entity|user, entity)$$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

User dimension

- User interest patterns can be mined through users' interactions with the search engine

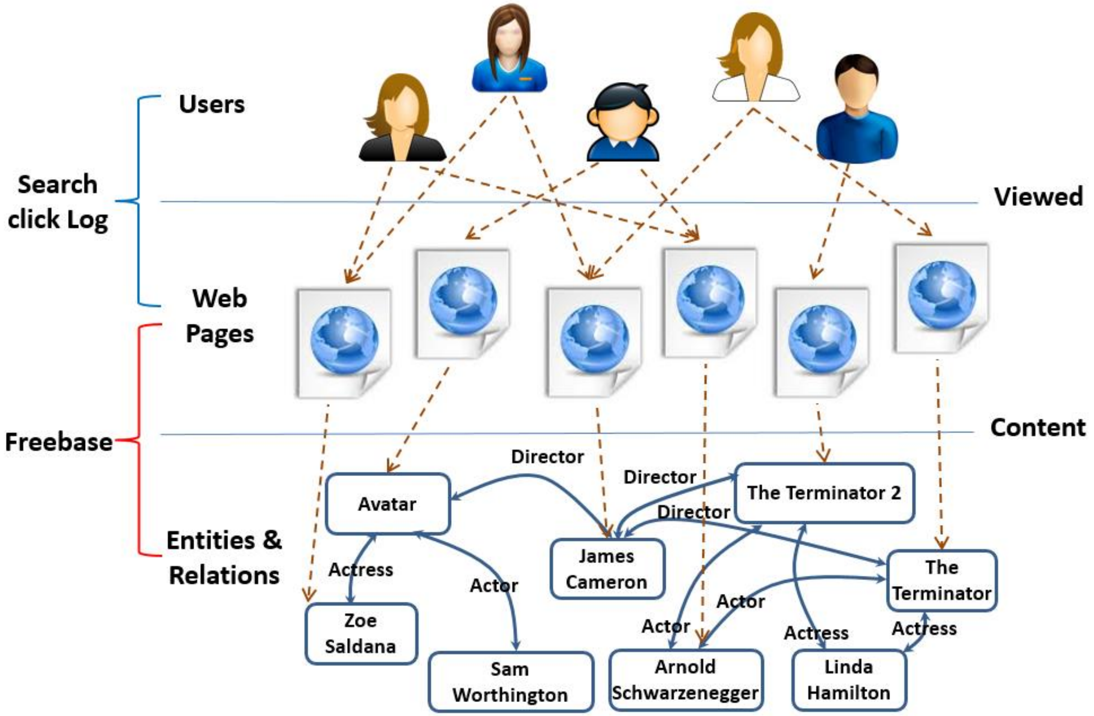
- Two sources:
  - Search click log
  - Entity pane log

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

User dimension

- User interest patterns can be mined through users' interactions with the search engine

- Two sources:
  - Search click log
  - Entity pane log



Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

User dimension

- User interest patterns can be mined through users' interactions with the search engine

- Two sources:
  - Search click log
  - Entity pane log

Entity pane log

User ID	Time	Main entity	Related entity	Rank	Click
32	7/9/2013 10:32:26	Pacific Rim	Man of Steel	1	0
32	7/9/2013 10:32:26	Pacific Rim	The Wolverine	2	0
32	7/9/2013 10:32:26	Pacific Rim	The Lone Ranger	3	1
498	6/16/2013 15:16:41	Leonardo DiCaprio	Kate Winslet	1	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Johnny Depp	2	1

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

User dimension

- User interest patterns can be mined through users' interactions with the search engine
- Two sources:
  - Search click log
  - Entity pane log
- Each user is represented as a vector of features: $\boldsymbol{x}$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Main entity

- Reflects the user's current search interest
- Ignoring main entities leads to inferior performance
  - If related entities are obtained based purely on the user's past interest, they will be completely independent of her information need
- Each main entity is represented as a feature vector: $\boldsymbol{y}$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Related entity

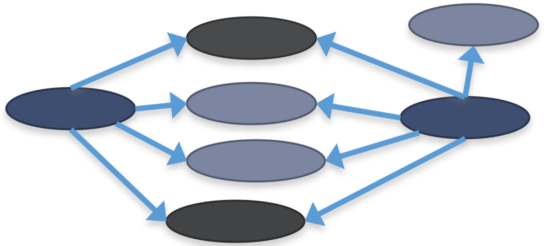
- A user may click a related entity, when it is aligned with both her interest pattern and current need

- Clicks on related entities specify user interests in certain facets of the main entities

- Each related entity is represented as a feature vector: z

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Goal

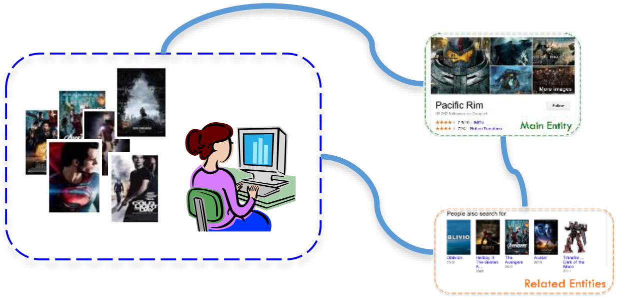


Knowledge base

Search click log

User ID	Time	Main entity	Related entity	Rank	Click
32	7/9/2013 10:32:26	Pacific Rim	Oblivion	1	0
32	7/9/2013 10:32:26	Pacific Rim	Hellboy II	2	0
32	7/9/2013 10:32:26	Pacific Rim	The Avengers	3	1
32	7/9/2013 10:32:26	Pacific Rim	Avatar	4	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Kate Winslet	1	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Baz Luhrmann	2	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Johnny Depp	3	1

Entity pane log



Three-way Entity Model

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Three-way Entity Model (TEM)

- Trilinear function:

$$\Phi_{umr}(\eta) = \sum_{i=0}^{I}\sum_{j=0}^{J}\sum_{k=0}^{K} \boxed{\eta_{ijk}} \cdot \boxed{x_{ui}} \cdot \boxed{y_{mj}} \cdot \boxed{z_{rk}}$$

i-th feature of user u / j-th feature of main entity m / k-th feature of related entity r

- Weights h capture the associations among users, main entities and related entities

- Feature vectors:

$$\begin{aligned} \mathbf{x}_u &= [1, x_{u1}, x_{u2}, \cdots, x_{uI}]^T, \\ \mathbf{y}_m &= [1, y_{m1}, y_{m2}, \cdots, y_{mJ}]^T, \\ \mathbf{z}_r &= [1, z_{r1}, z_{r2}, \cdots, z_{rK}]^T. \end{aligned}$$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

CTR incorporation

- Trilinear function contributes an important indicator to entity recommendation, especially for rare/new entities
- To further enhance recommendation on popular entities:
  - **CTR**(r): CTRs on related entities
  - **CTR**(m,r): CTRs on related entities specific to main entity m
  - **CTR**(u,m,r): CTRs on related entities specific to user u & main entity m
- Integration:

$$\Psi_{umr}(\eta, \beta) = \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \eta_{ijk} \cdot x_{ui} \cdot y_{mj} \cdot z_{rk} + \underbrace{\boxed{\beta}}_{\text{weight vector}}{}^{\top} \underbrace{\boxed{V_{\mathrm{CTR}}}}_{\text{CTR feature vector}}$$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Learning from entity pane log

- Clicks on entity pane
  - Positive feedback from users
- Negative feedback is missing
  - Users didn't click recommended entities for different reasons
- **Solution**: we use entity pairs as training data instead of individual entities

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Constructing training data

- Assumption
  - Users prefer the related entities they clicked over all the other suggestions

User ID	Time	Main entity	Related entity	Rank	Click
32	7/9/2013 10:32:26	Pacific Rim	Man of Steel	1	0
32	7/9/2013 10:32:26	Pacific Rim	The Wolverine	2	0
32	7/9/2013 10:32:26	Pacific Rim	The Lone Ranger	3	1
32	7/9/2013 10:32:26	Pacific Rim	World War Z	4	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Kate Winslet	1	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Baz Luhrmann	2	0
498	6/16/2013 15:16:41	Leonardo DiCaprio	Johnny Depp	3	1

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Likelihood function

- Likelihood function relating Ψ_{umr} values to pairwise preferences:

$$p(r_i \succ r_j | \Psi_{umr_i}, \Psi_{umr_j}) = \frac{1}{1 + e^{-g_{r_i r_j}(\Psi_{umr_i} - \Psi_{umr_j})}} \qquad g_{r_i r_j} \in \{-1, 1\}$$

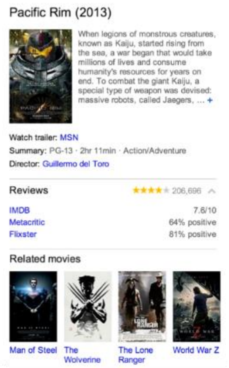
- Likelihood of all preference observations:

$$\begin{aligned} p(\mathcal{D}|\Psi) &= \prod_{(u,m,r_i,r_j)\in\mathcal{D}} p(r_i \succ r_j | \Psi_{umr_i}, \Psi_{umr_j}) \\ &= \prod_{(u,m,r_i,r_j)\in\mathcal{D}} \frac{1}{1 + e^{-g_{r_i r_j}(\Psi_{umr_i} - \Psi_{umr_j})}} \end{aligned}$$

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Experiments

- Two tasks: Movie recommendation Celebrity recommendation



- Data (3/2013 ~ 7/2013)

Dataset	#users	#entities	#instances
Movie	36,641	15,409	224,567
Celebrity	26,371	2,016	1,450,609

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Features used for movie & celebrity recommendation

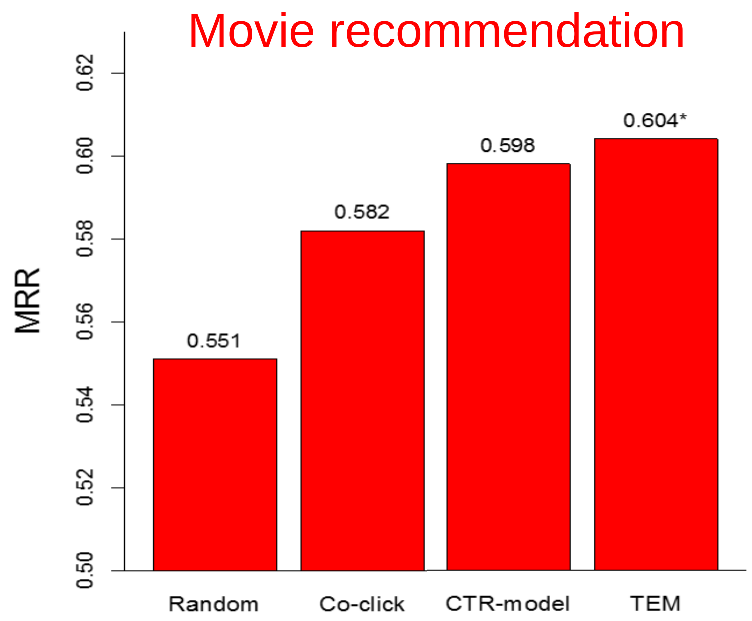
Movie recommendation	
User dimension	**Main & related movie**
Viewed entities Types of viewed entities Viewed movie's actors Viewed movie's directors Viewed movie's genres Viewed movie's country Viewed movie's language Viewed movie's producers Viewed movie's series Viewed movie's story Viewed movie's subject Viewed movie's music	Actors Directors Genres Country of origin Language Producers Series Story Subject Music ......

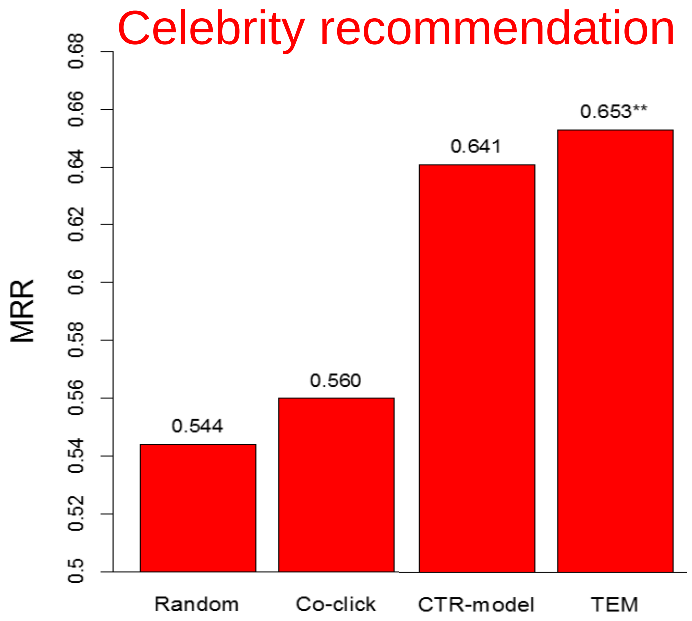
Celebrity recommendation	
User dimension	**Main & related movie**
Viewed entities Types of viewed entities Attributes of viewed entities Viewed pop singers Viewed business leaders Viewed writers Viewed musicians Viewed actors Viewed film directors	Profession Movie acted Movie directed Book written Music genre Organization Spouse Nationality Language Types ......

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Recommendation accuracy

- Metric: **MRR**
  - calculates the reciprocal of the rank of the first hit in the list
  - $MRR = \frac{1}{|Q|}\sum_{n=1}^{|Q|}\frac{1}{rank^{(n)}}$





Efficacy of personalization

- Case study

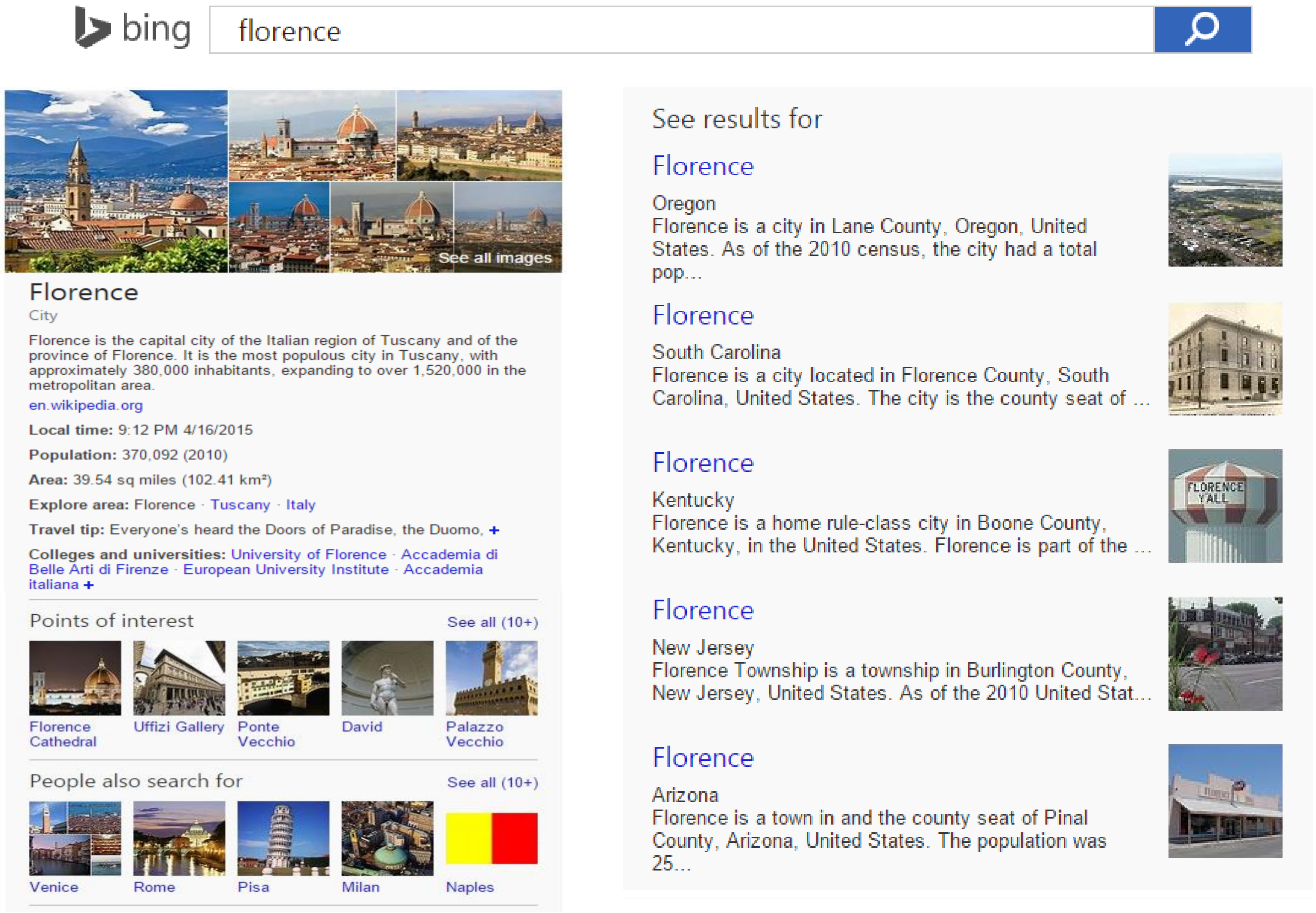
A fan of *Leonardo DiCaprio*

Main movie entity

Related movie entities		
Co-click	CTR-model	TEM
Iron Man 3 Man of Steel Star Trek (2013) Django Unchained	Iron Man 3 Star Trek (2013) Django Unchained Man of Steel	Django Unchained Iron Man 3 Star Trek (2013) Man of Steel

Learning to Recommend Related Entities to Search Users [Bin Bi, et al., WSDM 2015]

Other Entity Personalization Experience



Other Entity Personalization Experience

bing sigir

SIGIR 2015

SIGIR is the major international forum for the presentation of new research results and for the demonstration of new systems and techniques in information retrieval.

wikicfp.com

Dates: Aug 09 - 13, 2015

Location: Santiago

Website: SIGIR 2015

Submissions due: Jan 28, 2015

People also search for

- CIKM 2015 (Oct 19, 2015)
- WWW 2015 (May 20, 2015)
- AAAI 2016 (Feb 12, 2016)
- ECIR 2015 (Mar 29, 2015)
- KDD 2015 (Aug 10, 2015)

See more

Special Inspector General for Iraq Reconstruction

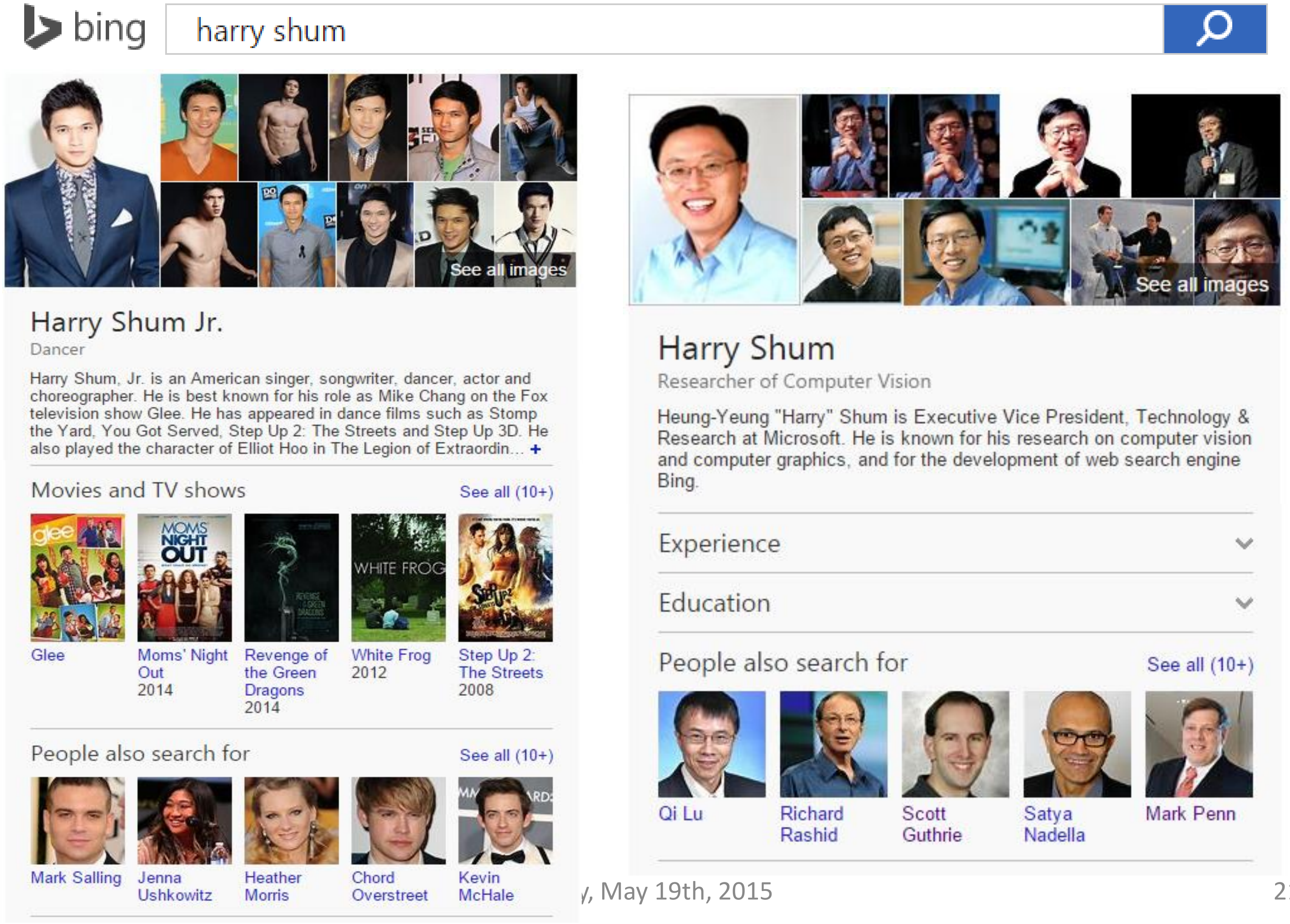
The Office of the Special Inspector General for Iraq Reconstruction was created as the successor to the Coalition Provisional Authority Office of Inspector General. SIGIR was an independent government agency created by the Congress to provide oversight of the use of the $52 billion U.S. reconstruction program in Iraq. Stuart W. Bowen, Jr. was app… +

en.wikipedia.org

Founded: 2004

Other Entity Personalization Experience

$P(entity|user, query)$



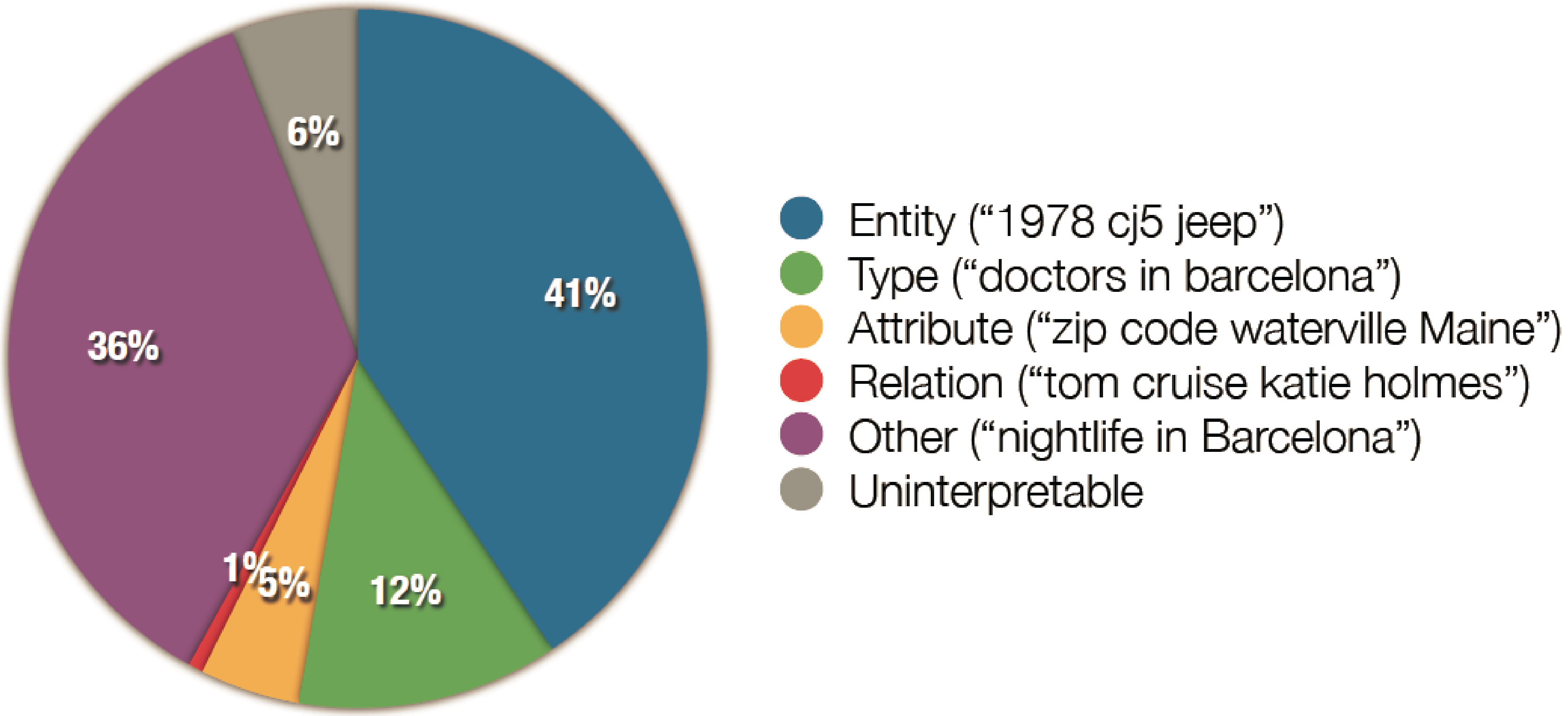
Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$
  - Recommendations given an entity
    - Co-occurrence
    - Similarity
    - Entity Linking
    - Interpretation
- $P(entity|user)$
  - Recommendations given a user
    - Universal recommender system
    - $P(entity|user, item)$
    - $P(entity|user, query)$
- $P(entity|query)$
  - Recommendations given a query

Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$
  - Recommendations given an entity
    - Co-occurrence
    - Similarity
    - Entity Linking
    - Interpretation
- $P(entity|user)$
  - Recommendations given a user
    - Universal recommender system
    - $P(entity|user, item)$
    - $P(entity|user, query)$
- $P(entity|query)$
  - Recommendations given a query

$P(entity|query)$



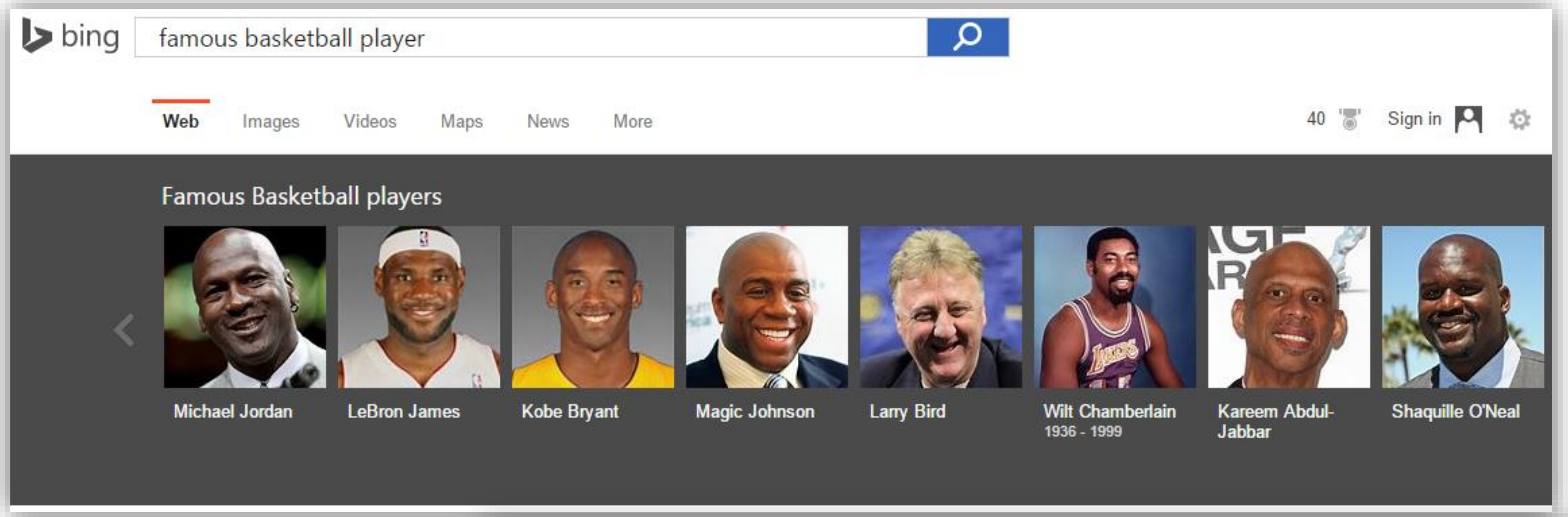
Ad-hoc Object Retrieval in the Web of Data [Jeffrey Pound, et al., WWW 2010]

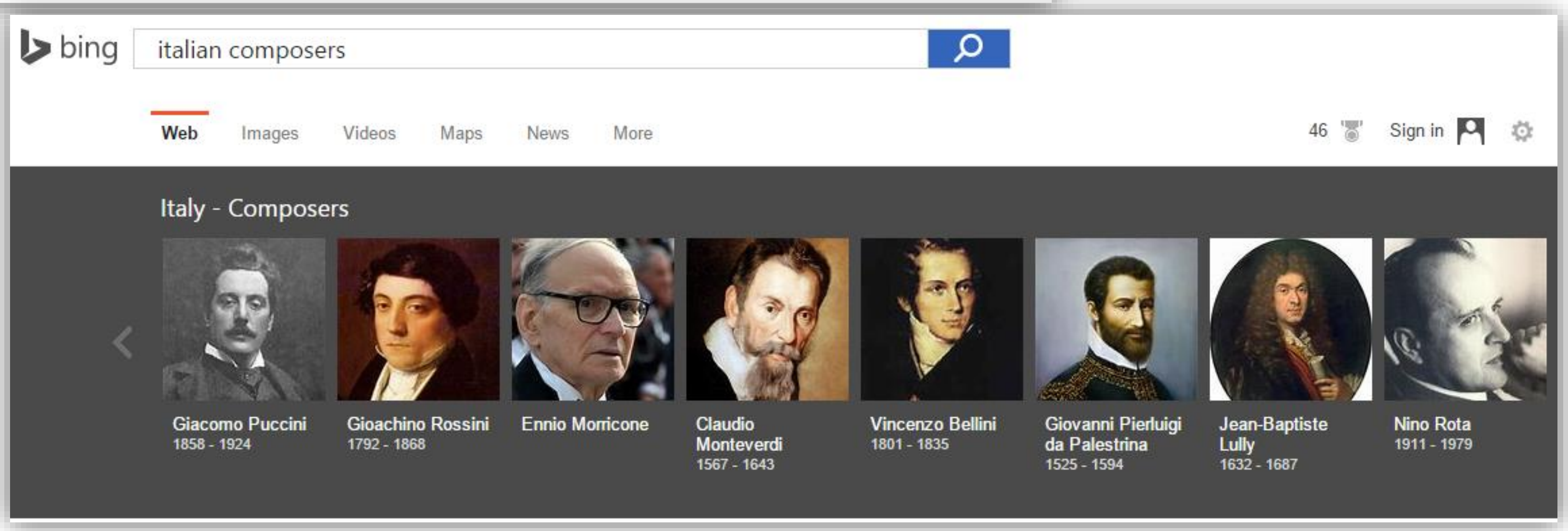
Entity Linking and Retrieval for Semantic Search [Edgar Meij, et al., WSDM 2014]

$P(entity|query)$

- Entity Retrieval/Finding
- Knowledge Base Question and Answering (KB QnA)
- Web-based Question and Answering (Web QnA)

$P(entity|query)$ – Entity Retrieval/Finding





$P(entity|query)$ – Entity Retrieval/Finding

- TREC Entity Track (2009 – 2011)
  - Related Entity Finding Task
  - Given
    - Input entity
    - Type of the target entity (PER/ORG/LOC)
    - Narrative (describing the nature of the relation in free text)
  - Return related entities

$P(entity|query)$ – Entity Retrieval/Finding

Input Entity: Boeing 747

Target Entity Type: Organization

Narrative: Airlines that currently use Boeing 747 planes

Input Entity: The food network

Target Entity Type: Person

Narrative: Chefs with a show on the food network

Input Entity: Eurail

Target Entity Type: Location

Narrative: What countries does Eurail operate in

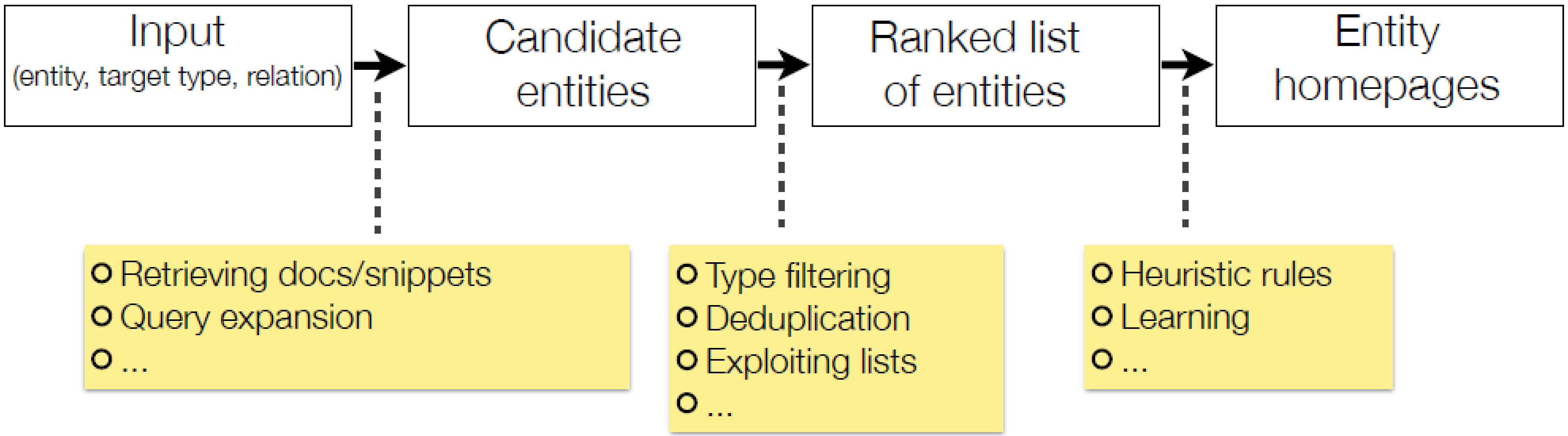
Input Entity: Dow Jones

Target Entity Type: Organization

Narrative: Find companies that are included in the Dow Jones industrial average

$P(entity|query)$ – Entity Retrieval/Finding

- A typical pipeline

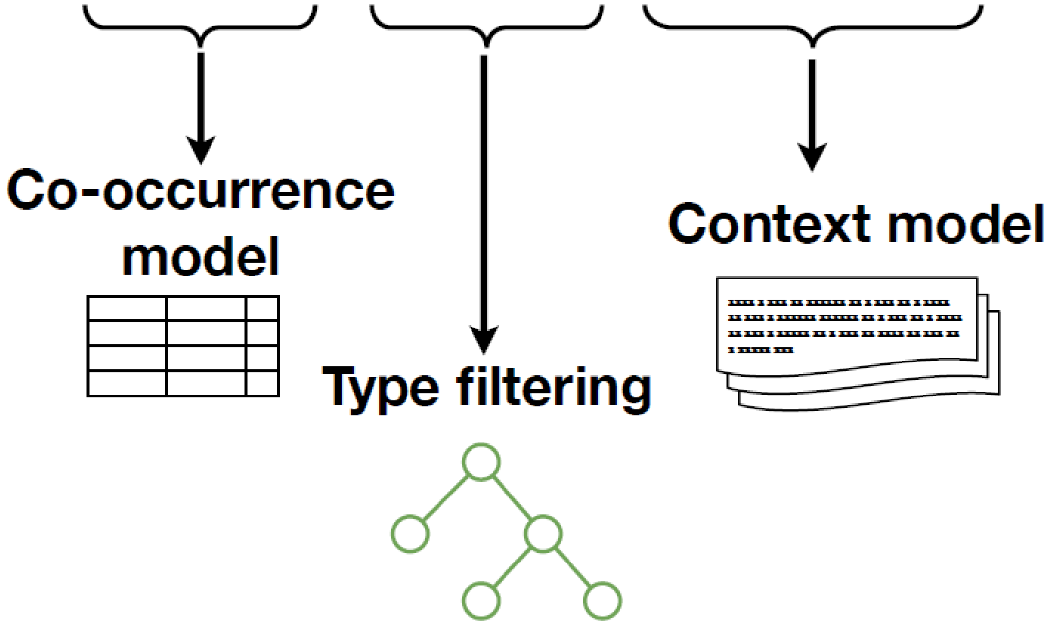


Entity Linking and Retrieval for Semantic Search [Edgar Meij, et al., WSDM 2014]

$P(entity|query)$ – Entity Retrieval/Finding

- Three component model

$$p(e|E,T,R) \propto p(e|E) \cdot p(T|e) \cdot p(R|E,e)$$



Related Entity Finding Based on Co-Occurrence [Marc Bron, et al., TREC 2009]

$P(entity|query)$ – Entity Retrieval/Finding

$$p(e|E,T,R) \propto \underbrace{p(e|E)}_{\text{Co-occurrence model}} \cdot \underbrace{p(T|e)}_{\text{Type filtering}} \cdot \underbrace{p(R|E,e)}_{\text{Context model}}$$

$$P(R|E,e) = P(R|\theta_{Ee}) = \prod_{t \in R} P(t|\theta_{Ee})^{n(t,R)}$$

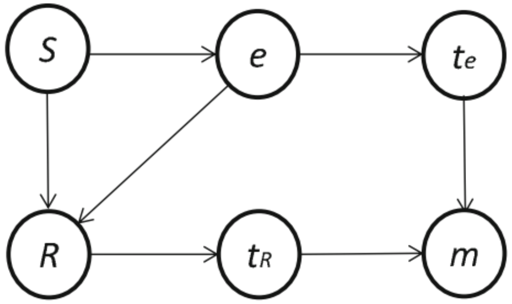
$$P(t|\theta_{Ee}) = \frac{1}{|D_{Ee}|} \sum_{d \in D_{Ee}} P(t|\theta_d)$$

$$P(t|\theta_d) = \frac{n(t,d) + \mu \cdot P(t)}{\sum_t' n(t',d) + \mu}$$

Related Entity Finding Based on Co-Occurrence [Marc Bron, et al., TREC 2009]

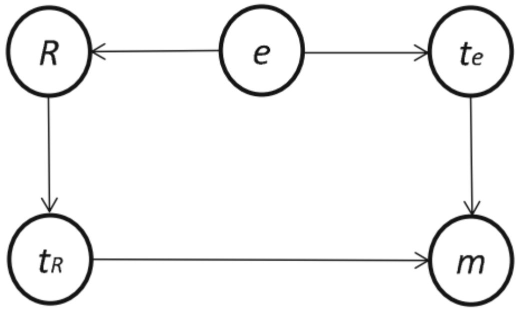
$P(entity|query)$ – Entity Retrieval/Finding

Model A



$$p(e, m = 1|R, S) \propto p(R|e, S)p(e|S) \sum_{t_R} \sum_{t_e} p(m = 1|t_e, t_R)p(t_e|e)p(t_R|R)$$

Model B



$$p(e, m = 1|R) \propto p(R|e)p(e) \sum_{t_R} \sum_{t_e} p(m = 1|t_e, t_R)p(t_e|e)p(t_R|R)$$

Related Entity Finding by Unified Probabilistic Models [Yi Fang, et al., World Wide Web 2013]

$P(entity|query)$ – Entity Retrieval/Finding

Input Entity: Dow Jones

Target Entity Type: Organization

Narrative: Find companies that are included in the Dow Jones industrial average

$p(m=1\|e,R)$	$p(R\|e)p(e)$	MA	$p(R\|e,S)p(e\|S)$	MB
nasdaq	**microsoft**	**boeing**	**coca cola**	**boeing**
bloomberg	**boeing**	**ibm**	**boeing**	**coca cola**
ibm	*federal reserve*	**pfizer**	*cnnmoney*	**microsoft**
news corporation	*european*	**coca cola**	*futures*	nasdaq
Yahoo	**coca cola**	**intel**	**microsoft**	**ibm**
atari	*uw*	**alcoa**	**pfizer**	**intel**
washington post	**ibm**	*cnnmoney*	**alcoa**	**merck**
boeing	**intel**	**mcdonald's**	**ibm**	**dupont**
stanford	*futures*	**merck**	*federal reserve*	**caterpillar**
enterprise media group	**merck**	**microsoft**	**mcdonald's**	*stanford*

Related Entity Finding by Unified Probabilistic Models [Yi Fang, et al., World Wide Web 2013]

$P(entity|query)$ – Entity Retrieval/Finding

- Knowledge base are largely incomplete

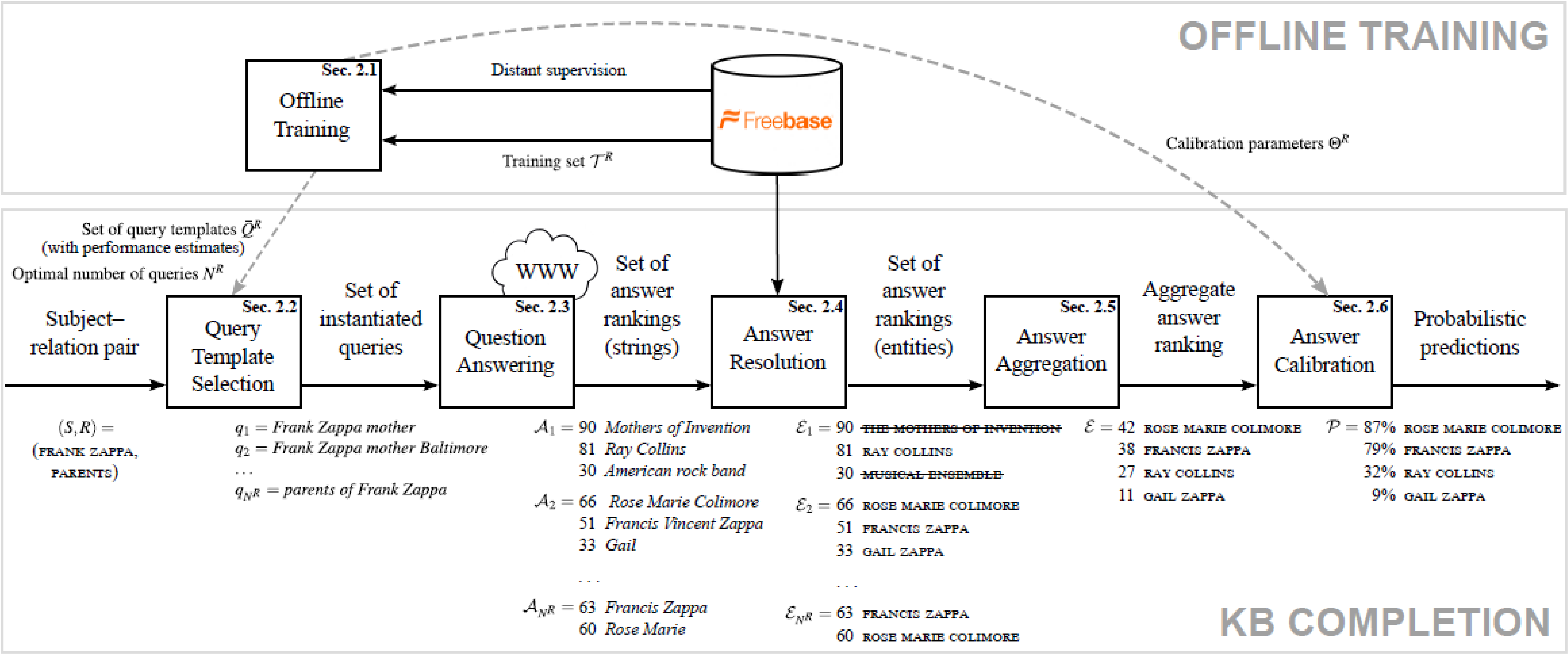
Freebase

Relation	Percentage unknown	
	All 3M	*Top 100K*
PROFESSION	68%	24%
PLACE OF BIRTH	71%	13%
NATIONALITY	75%	21%
EDUCATION	91%	63%
SPOUSES	92%	68%
PARENTS	94%	77%
CHILDREN	94%	80%
SIBLINGS	96%	83%
ETHNICITY	99%	86%

Entity Retrieval/Finding techniques can be used in Knowledge Base Completion

Knowledge Base Completion via Search-Based Question Answering [Robert West, et al., WWW 2014]

$P(entity|query)$ – Entity Retrieval/Finding



Knowledge Base Completion via Search-Based Question Answering [Robert West, et al., WWW 2014]

$P(entity|query)$ – Entity Retrieval/Finding

- Challenges
  - The TREC's related entity finding track is relatively easy since the "query intent" is known

Input Entity: Dow Jones

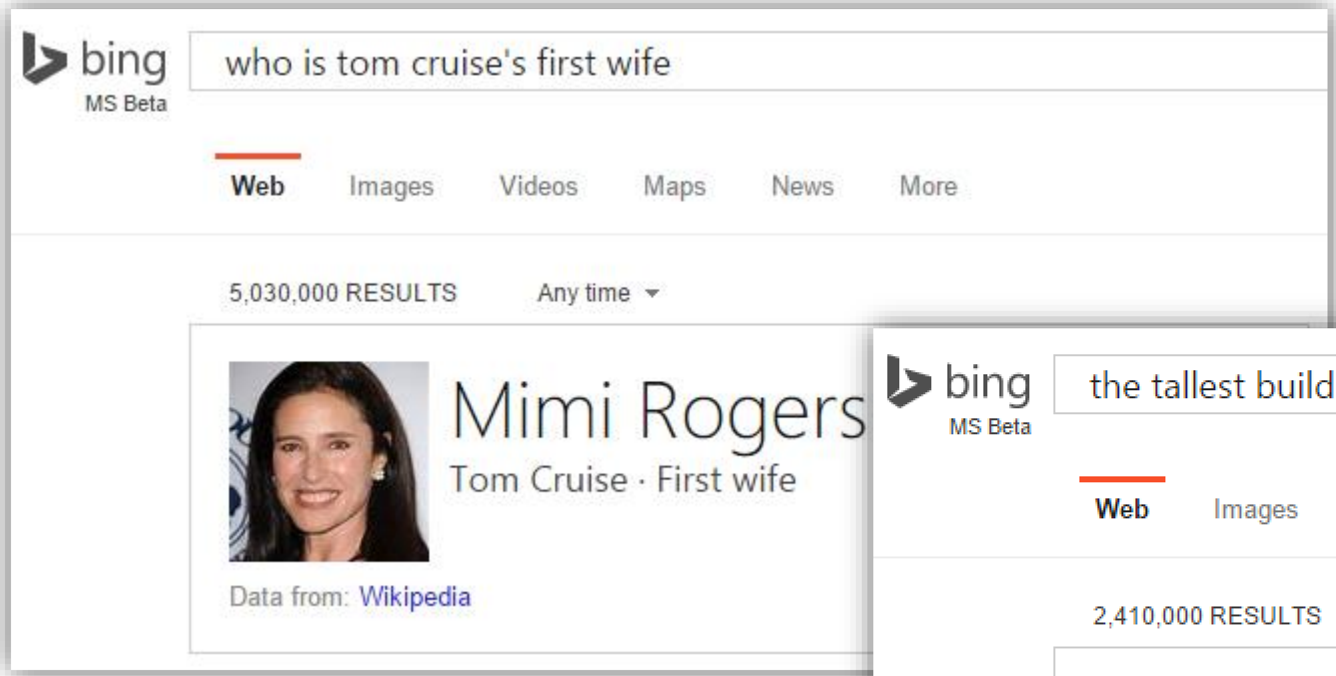
Target Entity Type: Organization

Narrative: Find companies that are included in the Dow Jones industrial average

  - In real world search engines, we need to understand the intent of queries

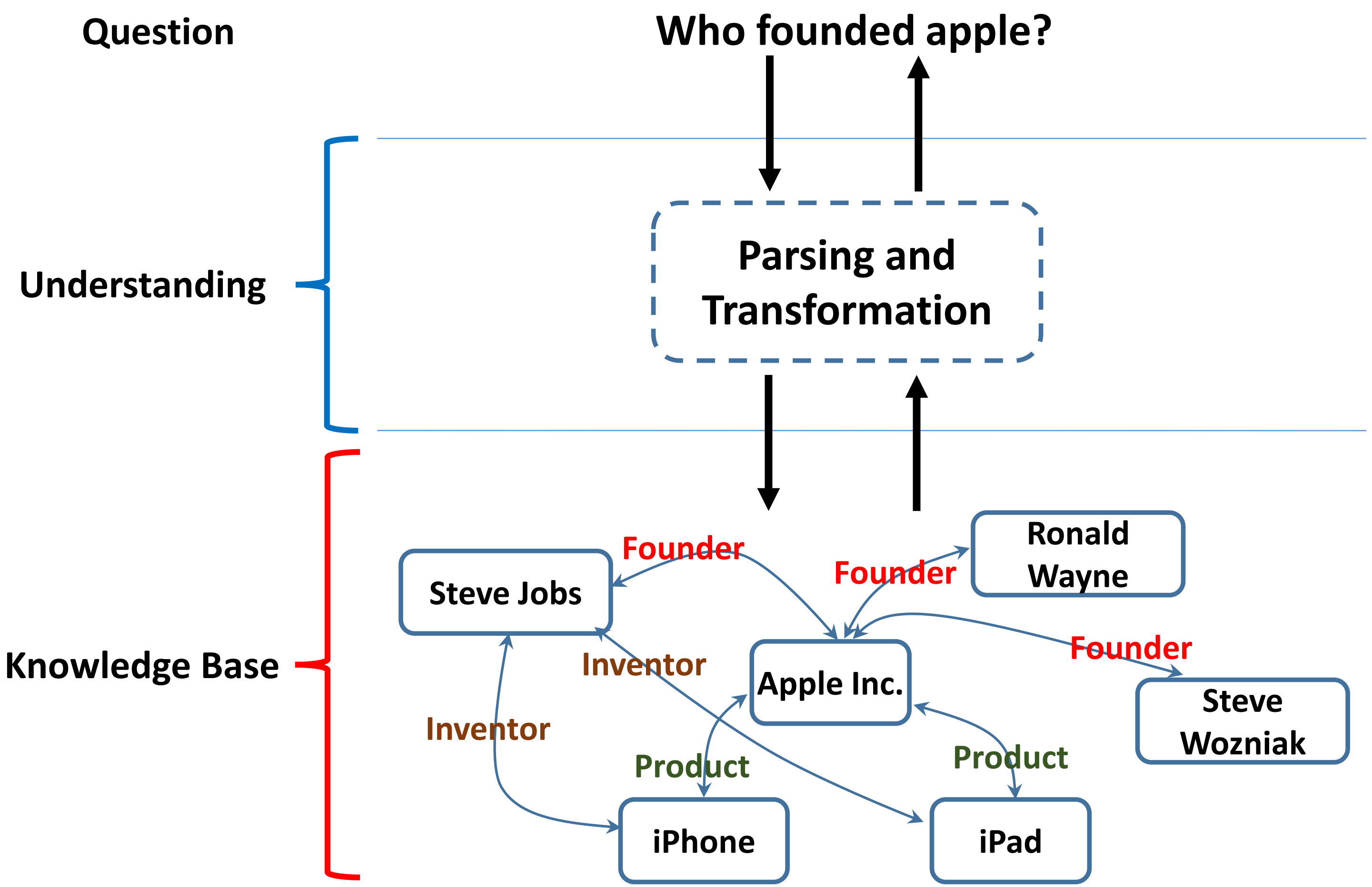
Companies in Dow Jones industrial

$P(entity|query)$ – KB QnA



$P(entity|query)$ – KB QnA

- Typical Architect of KB QnA



Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

An Incomplete List of Academic Papers on KB QnA

- Unger et al. “Template-based question answering over RDF data.” WWW-2012.
- Yahya et al. “Natural language questions for the web of data.” EMNLP-2012.
- Cai & Yates. “Large-scale semantic parsing via schema matching and lexicon extension.” ACL-2013.
- Kwiatkowski et al. “Scaling semantic parsers with on-the-fly ontology matching.” EMNLP-2013.
- Berant et al. “Semantic parsing on Freebase from question-answer pairs.” EMNLP-2013.
- Zou et al. “Natural language question answering over RDF: a graph data driven approach.” SIGMOD-2014.
- Yih et al. “Semantic parsing for single-relation question answering.” ACL-2014.
- Bao et al. “Knowledge-Based Question Answering as Machine Translation.” ACL-2014.
- Berant & Liang. “Semantic Parsing via Paraphrasing.” ACL-2014.
- Yao & Van Durme. “Information extraction over structured data: Question answering with freebase.” ACL-2014.
- Fader et al. “Open question answering over curated and extracted knowledge bases.” KDD-2014.
- Bordes et al. “Open question answering with weakly supervised embedding models.” ECML-PKDD-2014.
- Bordes et al. “Question answering with subgraph embeddings.” EMNLP-2014.
- Yang et al. “Joint relational embeddings for knowledge-based question answering.” EMNLP-2014.
- Reddy et al. “Large-scale Semantic Parsing without Question-Answer Pairs.” TACL, 2014.
- Yih et al. “Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base.” ACL-2015

$P(entity|query)$ – KB QnA

- Semantic Parsing

Who did Tom Cruise marry in 1987?

↓ semantic parsing

Type.Person ⊓ Marriage.(Spouse.TomCruise ⊓ StartDate.1987)

↓ execute logical form

Mimi Rogers

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

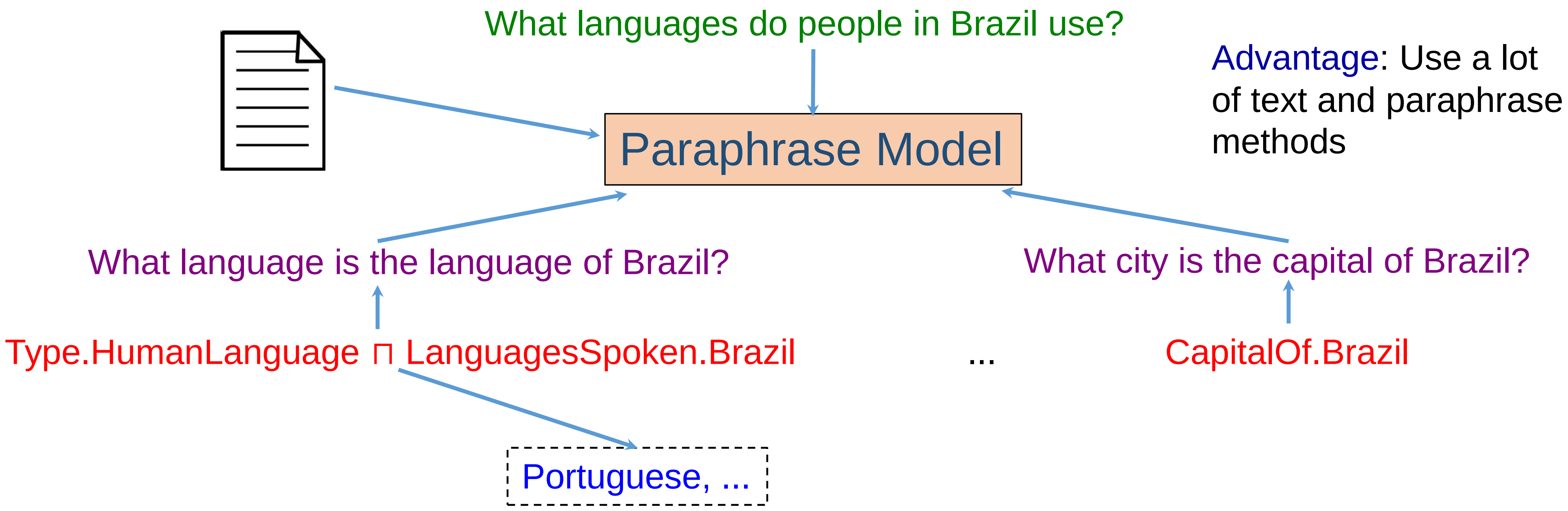
- Traditional statistical semantic parsing
  - Manually annotated logical forms

What's California's capital?	Capital.California
How long is the Mississippi river?	RiverLength.Mississippi
…	…

- Limitations
  - Requires experts | slow, expensive, does not scale!
  - Restricted to limited domains

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA



- Simple model suggests candidate logical forms
- Simple model generates canonical utterances
- Ranking of canonical utterances

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Input
  - Knowledge-base K
  - Training set of question-answer pairs $\{(x_i, y_i)\}_1^n$

    What are the main cities in California? SF, LA, ...
- Output
  - Semantic parser that maps questions x to answers y through logical forms z

Countries in Asia → Type.Country ⊓ ContainedBy.Asia

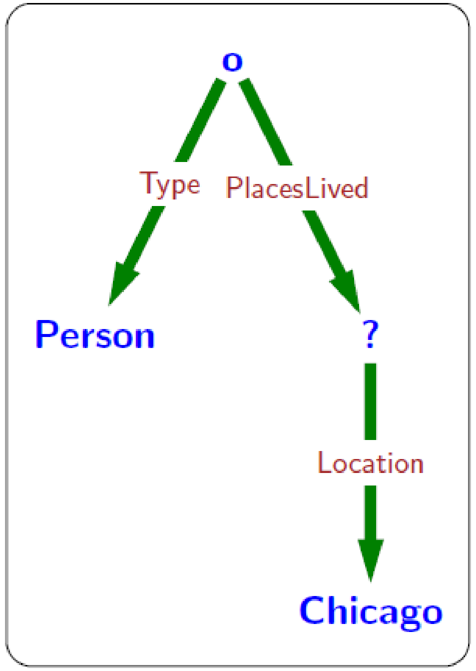
→ China, Japan, India, ...

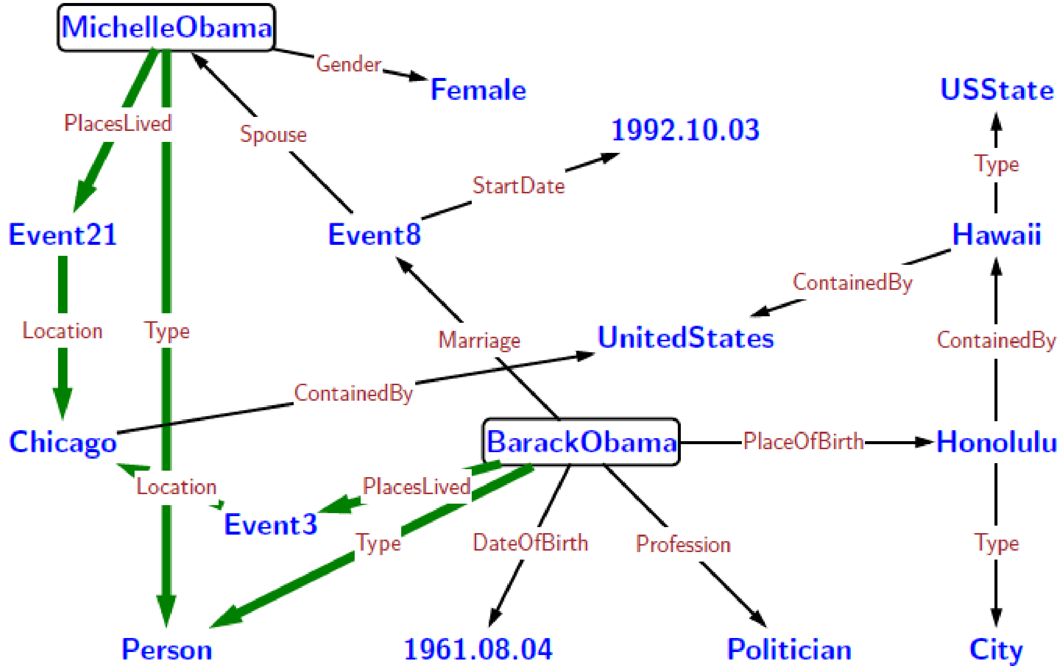
Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Logical forms are graph templates

Type.Person ⊓ PlacesLived.Location.Chicago

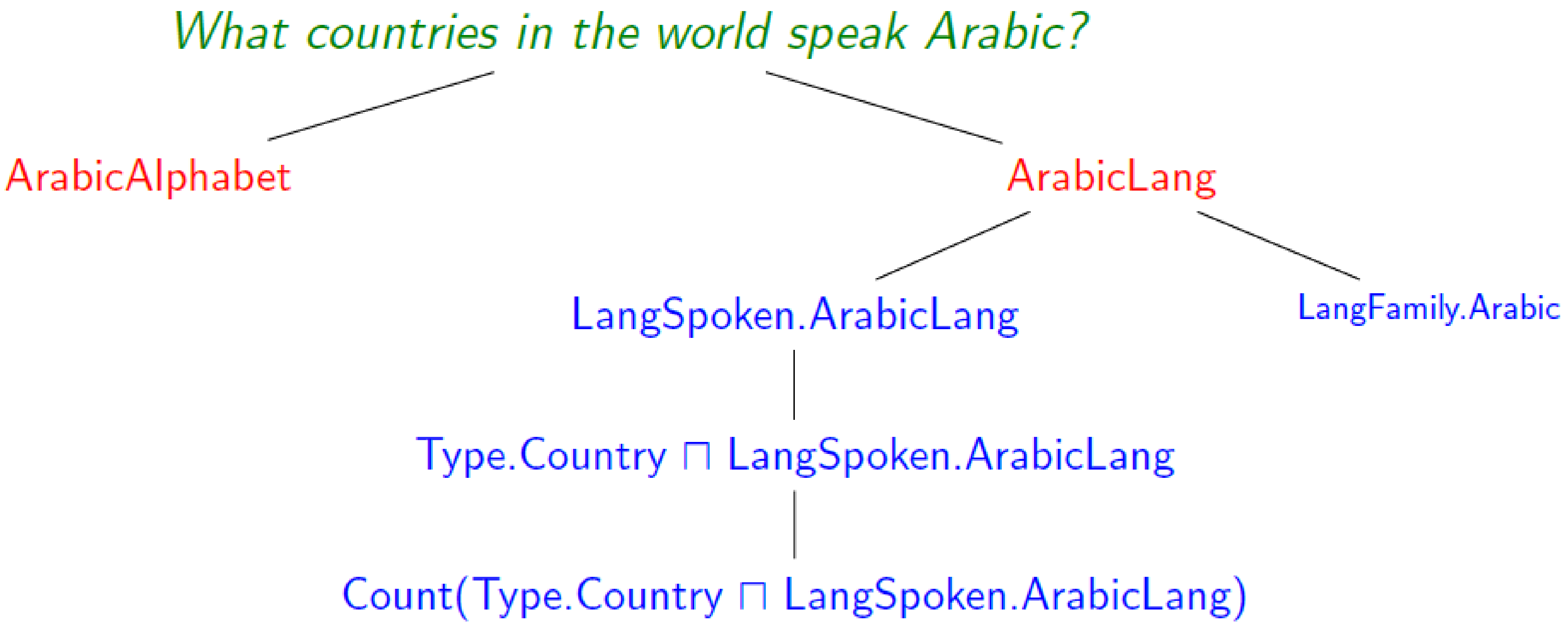




Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Candidate logical forms



Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Candidate logical forms

Template	Example	Question
$p.e$	Directed.TopGun	*who directed Top Gun*
$p_1.p_2.e$	Employment.EmployerOf.SteveBalmer	*Where does Steve Balmer work?*
$p.(p_1.e_1 \sqcap p_2.e_2)$	Character.(Actor.BradPitt⊓Film.Troy)	*Who did Brad Pitt play in Troy?*
Type.$t \sqcap z$	Type.Composer⊓SpeakerOf.French	*What composers spoke French?*
count(z)	count(BoatDesigner.NatHerreshoff)	*How many ships were designed by Nat Herreshoff?*

$p, p_1 p_2$– Freebase properties
e, e_1, e_2– Freebase entities

t – Freebase type
z – logical form

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Canonical utterance generation

$\text{Type.Country} \sqcap \text{LangSpoken.ArabicLang}$

country *spoken in* *languages spoken* *Arabic language*

syntactic analysis

What country is Arabic language spoken in?

What country spoken the languages Arabic language?

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Canonical utterance generation

	$d(p)$ **Categ.**	**Rule**	**Example**
$p.e$	NP	`WH` $d(t)$ `has` $d(e)$ `as` NP ?	*What* ***election contest*** *has* ***George Bush*** *as* ***winner?***
	VP	`WH` $d(t)$ `(AUX)` VP $d(e)$?	*What* ***radio station serves area New-York****?*
	PP	`WH` $d(t)$ PP $d(e)$?	*What* ***beer from region Argentina****?*
	NP VP	`WH` $d(t)$ VP `the` NP $d(e)$?	*What* ***mass transportation system served*** *the* ***area Berlin****?*
$\mathbf{R}(p).e$	NP	`WH` $d(t)$ `is the` NP `of` $d(e)$?	*What* ***location*** *is the* ***place of birth*** *of* ***Elvis Presley****?*
	VP	`WH` $d(t)$ `AUX` $d(e)$ VP ?	*What* ***film*** *is* ***Brazil featured in****?*
	PP	`WH` $d(t)$ $d(e)$ PP ?	*What* ***destination Spanish steps near travel destination****?*
	NP VP	`WH` NP `is` VP `by` $d(e)$?	*What* ***structure*** *is* ***designed*** *by* ***Herod****?*

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Paraphrase model

What countries in the world speak Arabic?

What country is Arabic language spoken in?

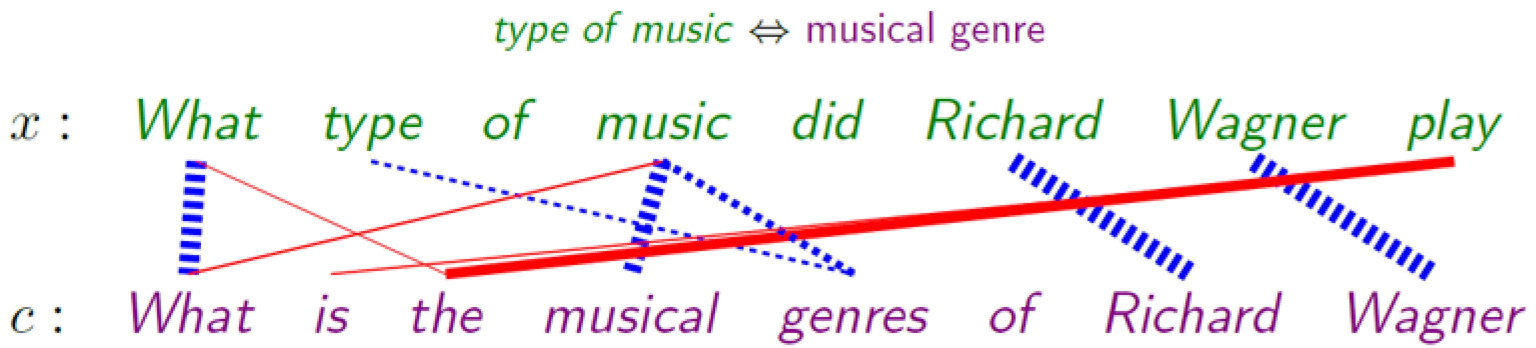
- Simple paraphrase model utilizing a lot of text
  - Association model - Paralex
  - Vector space model - Wikipedia

$$\phi_{\text{pr}}(x, c) = \phi_{\text{as}}(x, c) + \phi_{\text{vs}}(x, c)$$

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

Association: pair of spans $(x_{ij}, c_{i'j'})$



PARALEX corpus with 18 millions pairs of question paraphrases

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- Vector Space Model

Train word vectors $v(w)$:

C: content words in utterance x

$$v(x) = \frac{1}{|C|} \sum_{x_i \in C} v(x_i)$$

Learn a matrix W to estimate "similarity" score

$$s(x, c) = v(x)^\top W v(c)$$

Options for W

- Identity: dot product
- Diagonal: dot product with scaling
- Full matrix: interactions between dimensions

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

x : *What type of music did Richard Wagner play?*

as : *What is the musical genres of Richard Wagner?*

vs : *What composition has Richard Wagner as lyricist?*

x : *Where is made Kia car?*

as : *What place is founded by Kia motors?*

vs : *What city is Kia motors a headquarters of?*

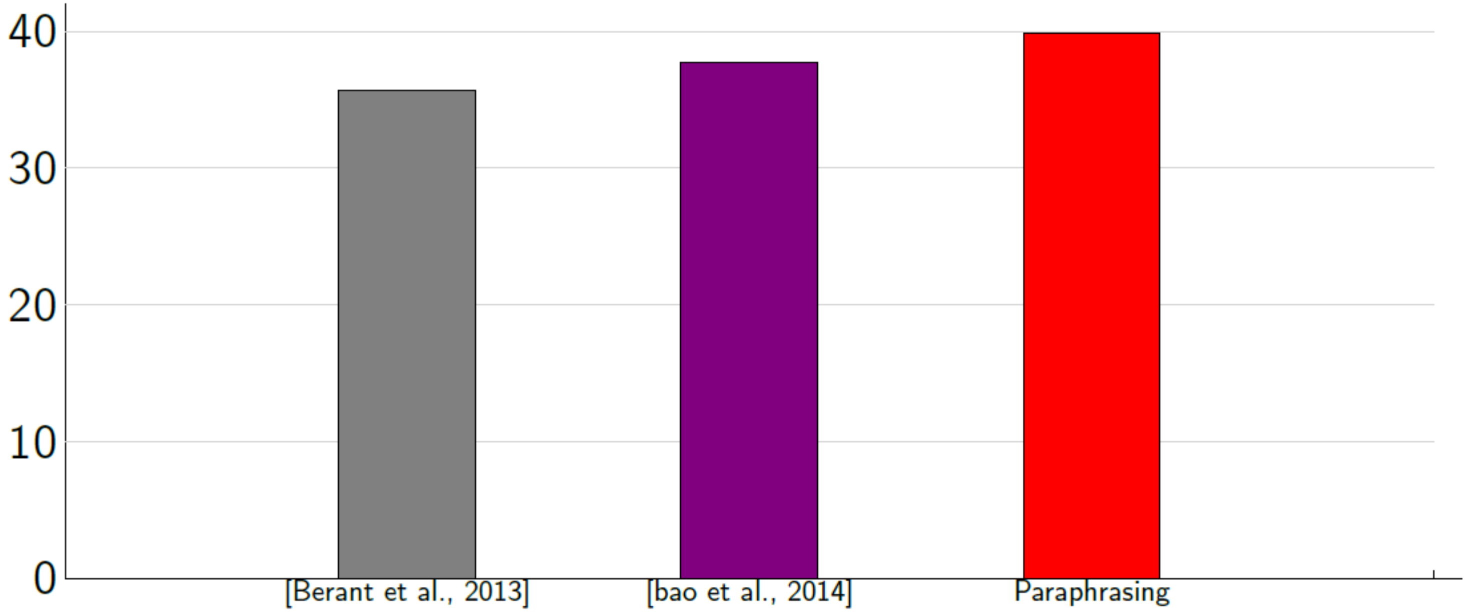
Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- WebQuestions dataset
  - What character did Natalie Portman play in Star Wars? → Padme Amidala
  - What kind of money to take to Bahamas? → Bahamian dollar
  - What currency do you use in Costa Rica? → Costa Rican colon
  - What did Obama study in school? → political science
  - What do Michelle Obama do for a living? → writer, lawyer
  - What killed Sammy Davis Jr? → throat cancer
- 5,810 questions crawled from Google Suggest and answered using AMT

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

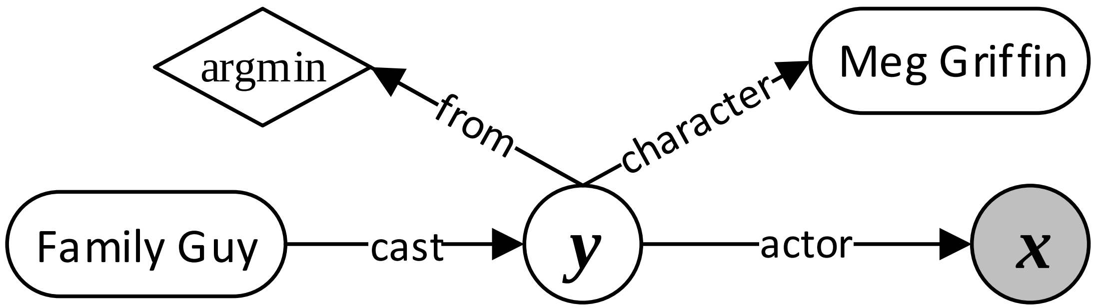


Outperforms previous state-of-the-art

Semantic Parsing via Paraphrasing [Jonathan Berant, et al., ACL 2014]

$P(entity|query)$ – KB QnA

- *"Who first voiced Meg on Family Guy?"*
- $\lambda x. \exists y. \mathrm{cast}(\mathrm{FamilyGuy}, y) \wedge \mathrm{actor}(y, x) \wedge \mathrm{character}(y, \mathrm{MegGriffin})$



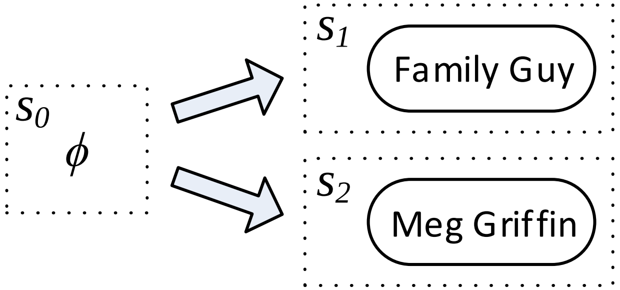
Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base [Scott Yih, et al., ACL 2015]

$P(entity|query)$ – KB QnA

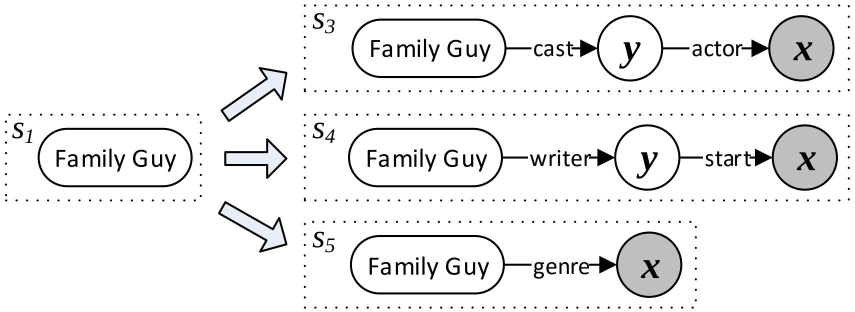
"Who first voiced Meg on Family Guy?"

- Leveraging KB more tightly when forming the parse (search pruning)
- The expressiveness of the query graphs controlled by search actions

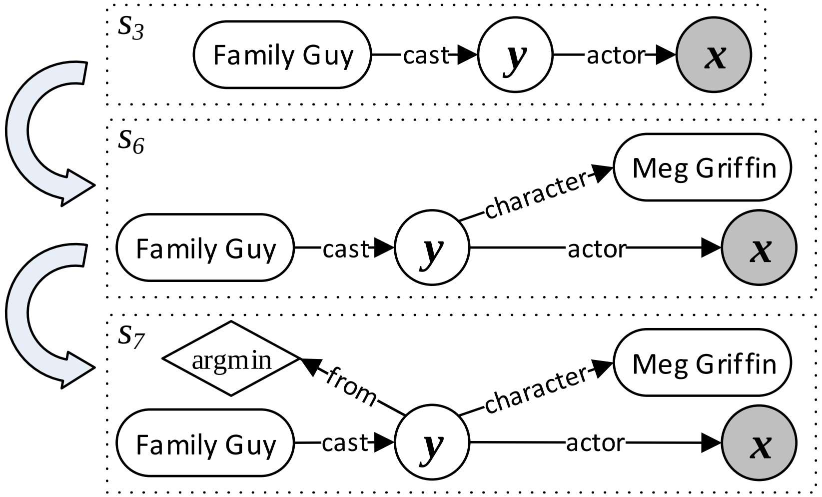
1. Topic Entity Linking (E2E tool)



2. Core Inferential Chain (DSSM)

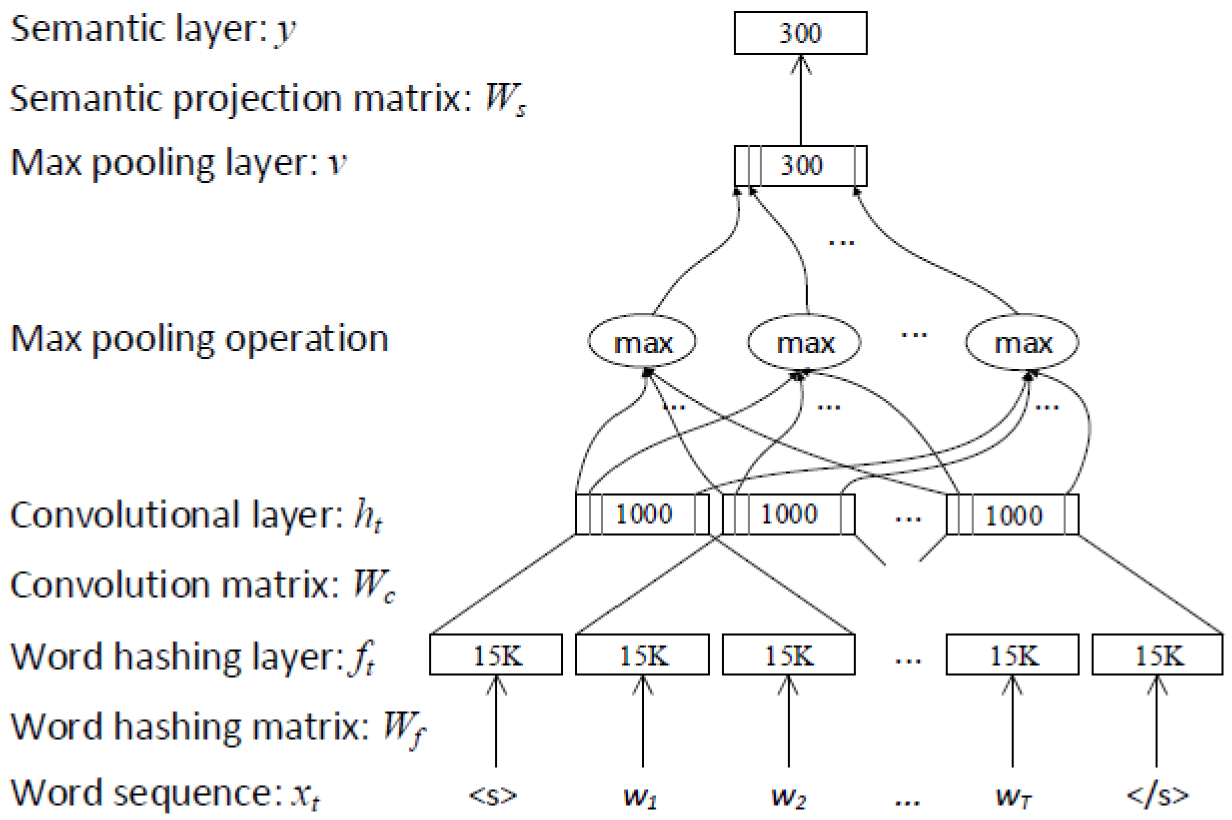


3. Augmenting Constraints



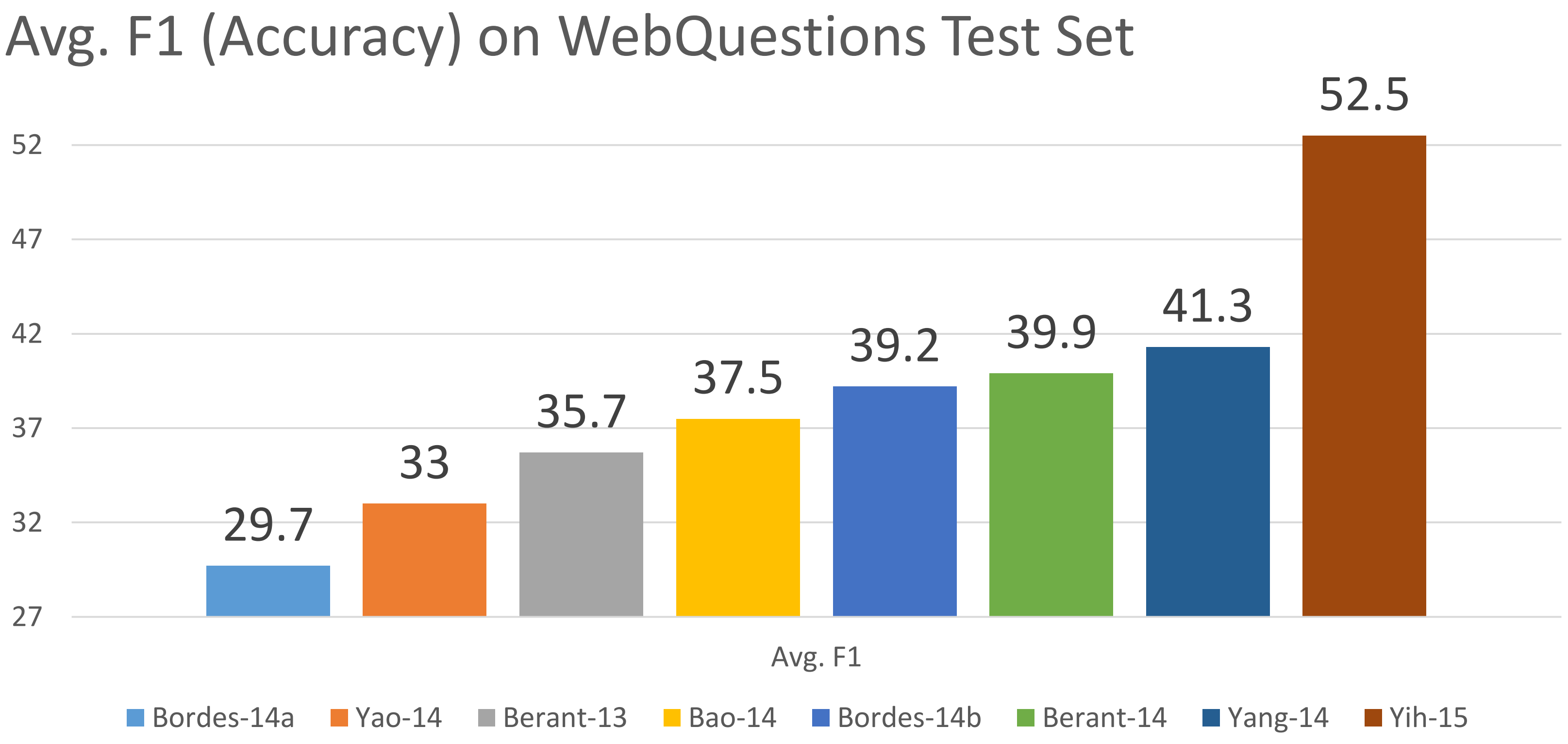
Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base [Scott Yih, et al., ACL 2015]

$P(entity|query)$ – KB QnA



Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base [Scott Yih, et al., ACL 2015]

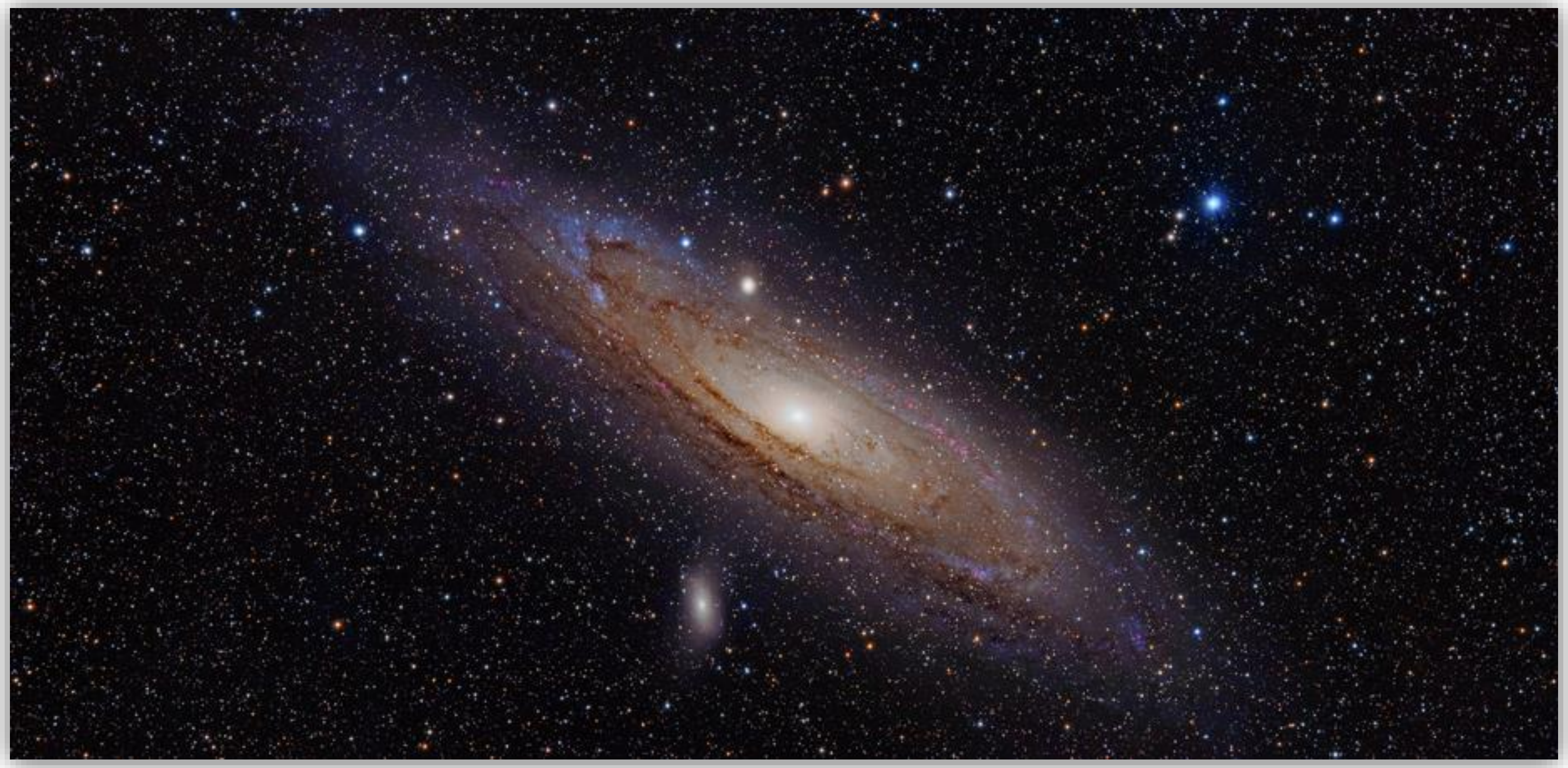
$P(entity|query)$ – KB QnA



Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base [Scott Yih, et al., ACL 2015]

$P(entity|query)$ – Web QnA

- Knowledge Base is largely incomplete



$P(entity|query)$ – Web QnA

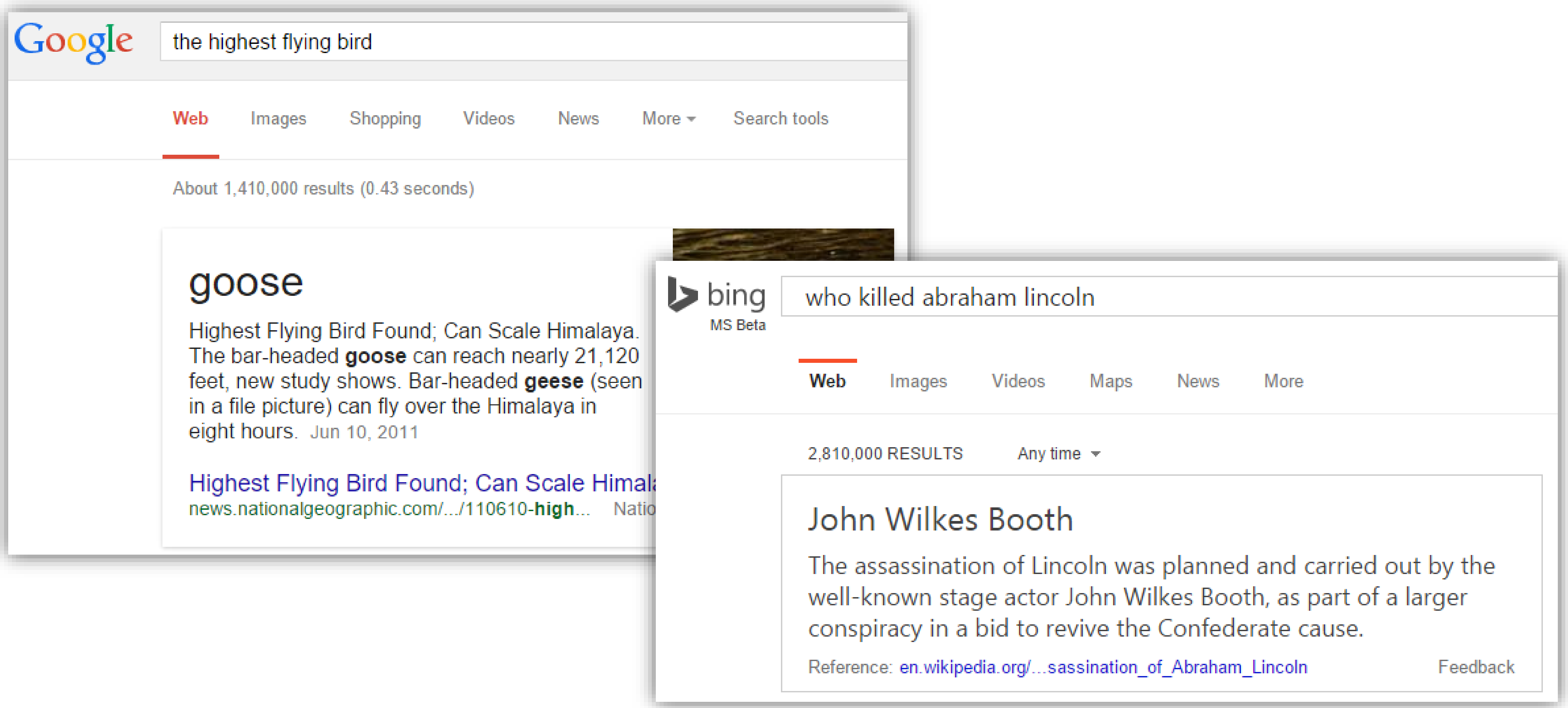
- Knowledge Base is largely incomplete

Freebase

Relation	Percentage unknown	
	All 3M	*Top 100K*
PROFESSION	68%	24%
PLACE OF BIRTH	71%	13%
NATIONALITY	75%	21%
EDUCATION	91%	63%
SPOUSES	92%	68%
PARENTS	94%	77%
CHILDREN	94%	80%
SIBLINGS	96%	83%
ETHNICITY	99%	86%

Knowledge Base Completion via Search-Based Question Answering [Robert West, et al., WWW 2014]

$P(entity|query)$ – Web QnA

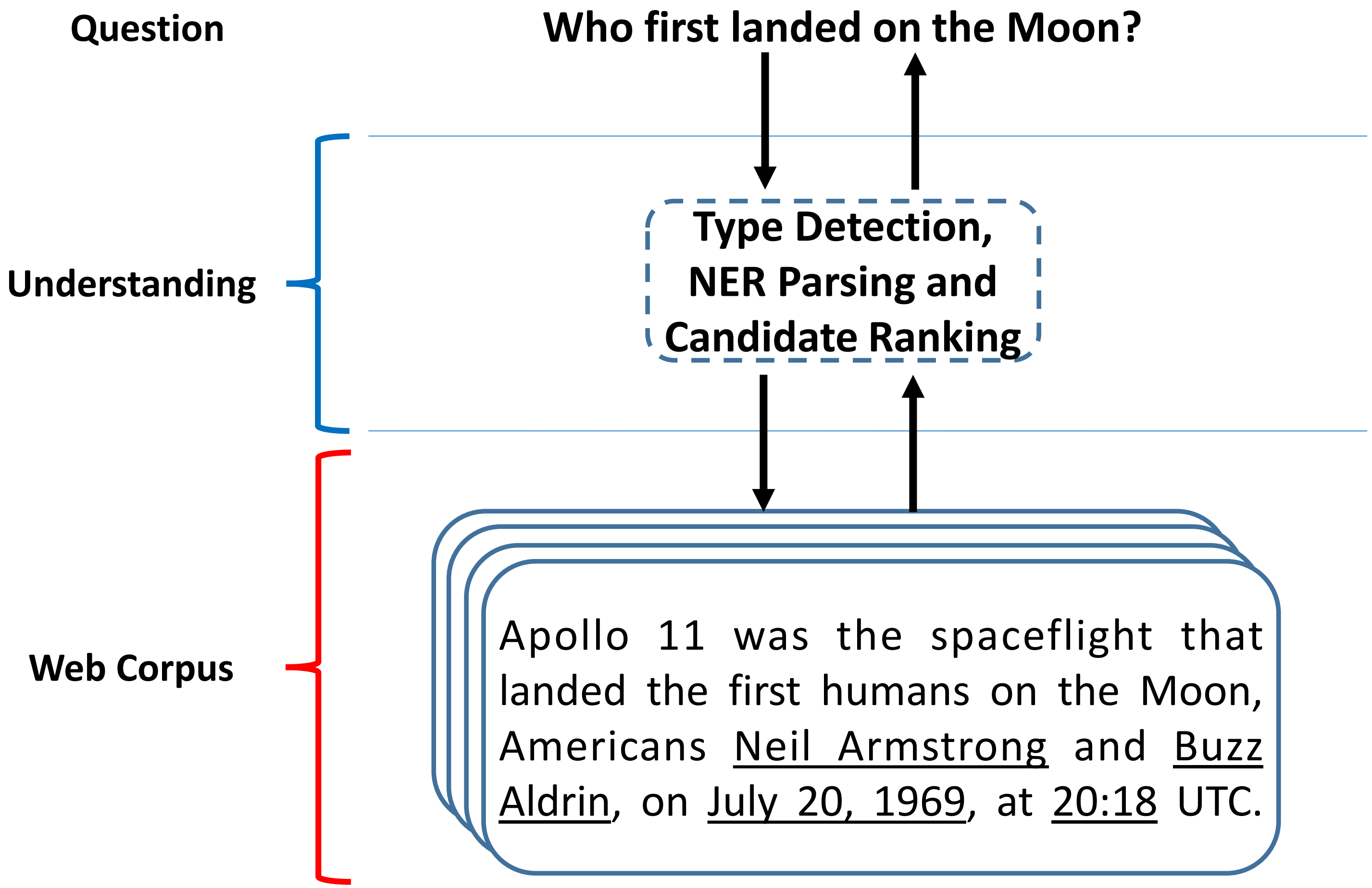


An Incomplete List of Academic Papers on Web QnA

- Dumais et al. “Web Question Answering: Is More Always Better? ” SIGIR-2002.
- Brill et al. “An analysis of the AskMSR question-answering system.” EMNLP-2002.
- Chu-Carroll et al. “A multi-strategy and multi-source approach to question answering.” Technical report-2006.
- Ko et al. “A probabilistic graphical model for joint answer ranking in question answering.” SIGIR-2007.
- Schlaefer et al. “A pattern learning approach to question answering within the ephyra framework.” TSD-2006.
- Azari et al. “Web-Based Question Answering: A Decision-Making Perspective.” UAI-2003.
- Ravichandran et al. “Learning surface text patterns for a Question Answering system.” ACL-2002.
- Kwok et al. “Scaling question answering to the web.” TOIS-2001.
- Brill et al. “Data-intensive question answering.” TREC-2001.
- Bian et al. “Finding the Right Facts in the Crowd: Factoid Question Answering over Social Media.” WWW-2008.
- Cheng et al. “EntityRank: Searching Entities Directly and Holistically.” VLDB-2007.
- Lin et al. “Question answering from the web using knowledge annotation and knowledge mining techniques.” CIKM-2003.
- Chaturvedi et al. “Joint question clustering and relevance prediction for open domain non-factoid question answering.” WWW-2014

$P(entity|query)$ – Web QnA

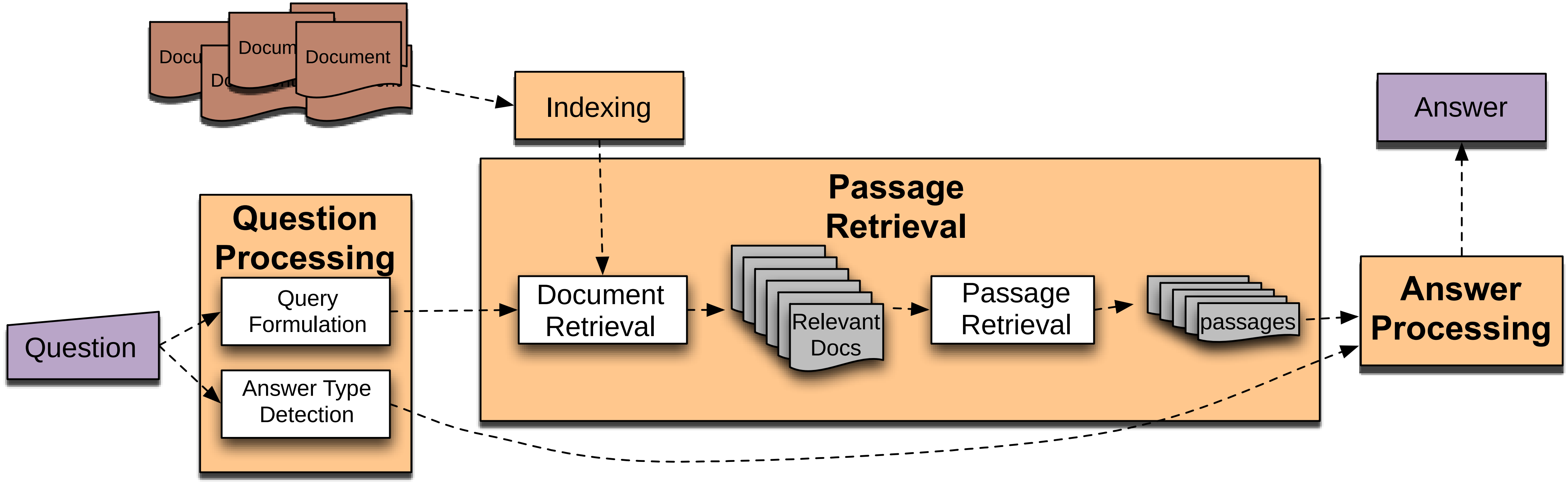
- Typical Architect of Web QnA



Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

$P(entity|query)$ – Web QnA

- Detailed Architect



Question Answering [Dan Jurafsky, Stanford]

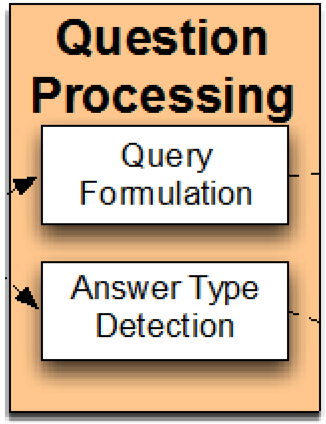
$P(entity|query)$ – Web QnA

- QUESTION PROCESSING
  - Detect question type, answer type
  - Formulate queries to send to a search engine
- PASSAGE RETRIEVAL
  - Retrieve ranked documents
  - Break into suitable passages and rerank
- ANSWER PROCESSING
  - Extract candidate answers
  - Rank candidates

Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$ – Web QnA

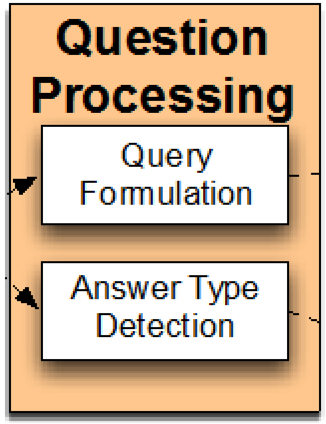
- Answer Type Detection: Name Entities
  - Who first landed on the moon?
    - Person
  - Where is the headquarters of Microsoft?
    - Location
  - What is the largest country in terms of population?
    - Country
  - Highest flying bird
    - Animal/Bird



Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$ – Web QnA

- 6 coarse classes
  - ABBEVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION, NUMERIC
- 50 finer classes
  - LOCATION: city, country, mountain…
  - HUMAN: group, individual, title, description
  - ENTITY: animal, body, color, currency…

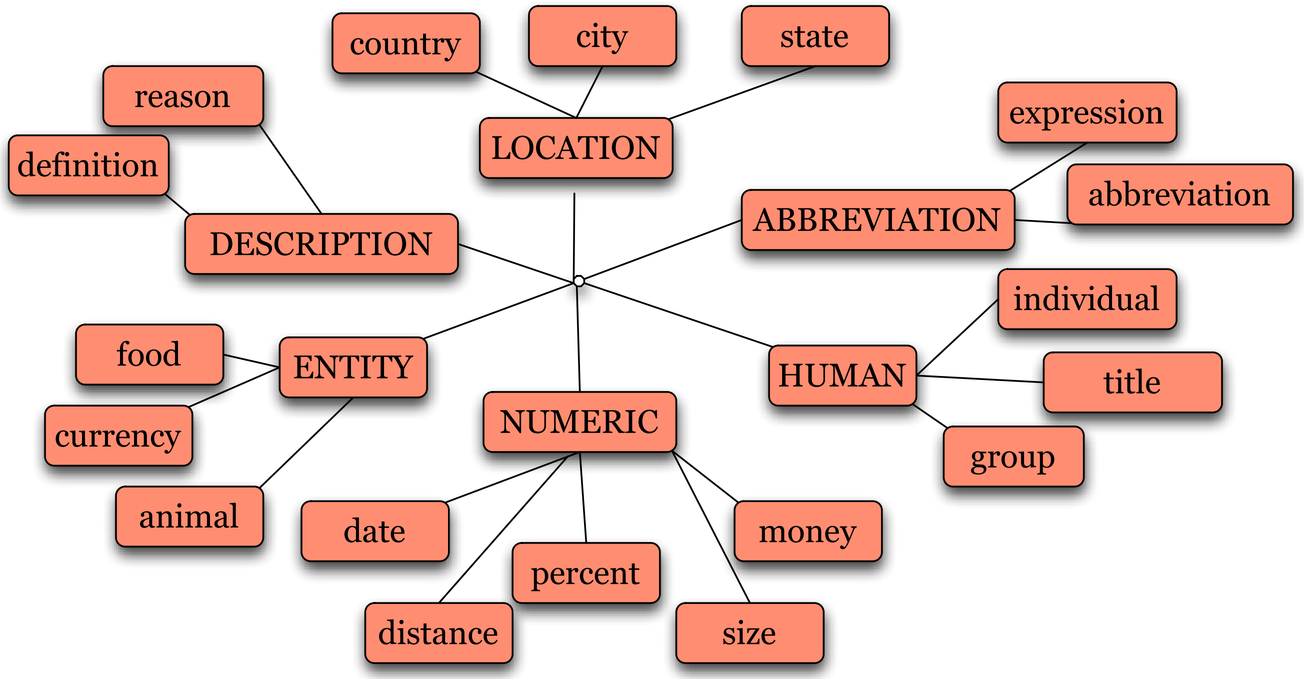


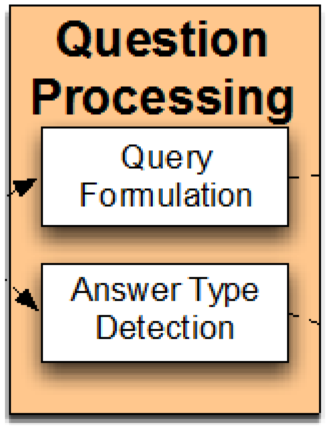
Learning Question Classifiers [Xin Li, et al., COLING 2002]

Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$ – Web QnA

- Part of the Answer Type Taxonomy



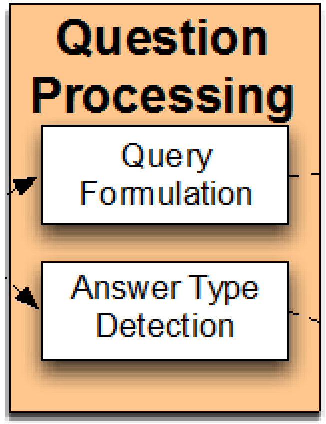


Learning Question Classifiers [Xin Li, et al., COLING 2012]

Question Answering [Dan Jurafsky, Stanford]

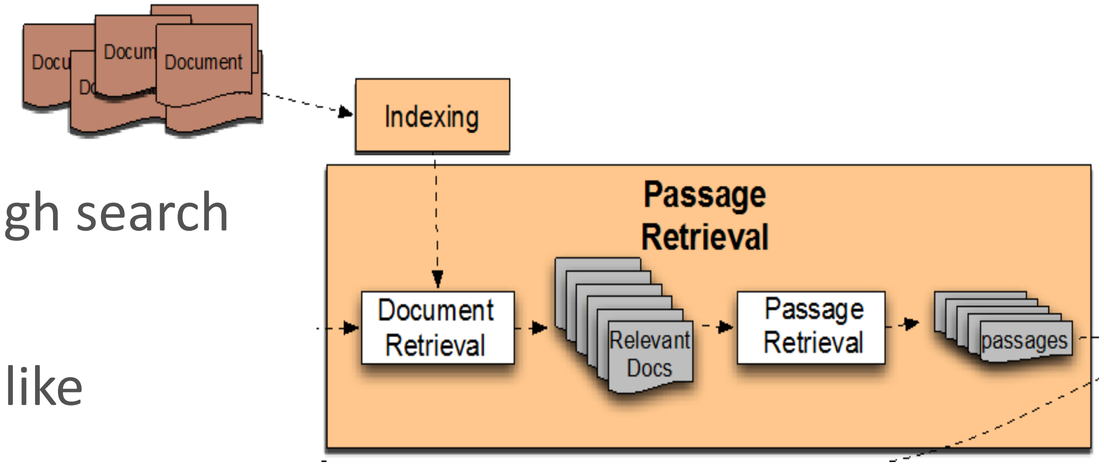
$P(entity|query)$ – Web QnA

- Answer Type Detection
  - Rules
    - Regular expression based rules
      - Who {is|was|are|were} PERSON
    - Question headword
      - Which **city** in China has the largest number of foreign financial companies?
      - What is the state **flower** of California?
  - Machine Learning
    - **Define** a taxonomy of question types
    - **Annotate** training data for each question type
    - **Train** classifiers for each question class using a rich set of features: Question words and phrases; Part-of-speech tags; Parse features (headwords); Named Entities; Related words



Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$ – Web QnA

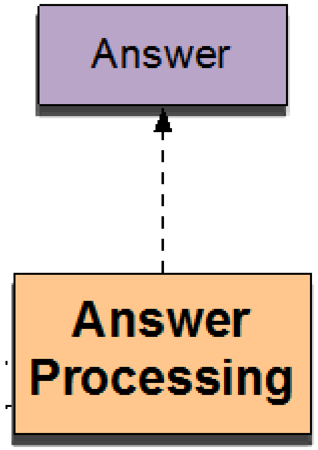


- Passage Retrieval
  - Retrieve documents using query terms through search engines
  - Segment the documents into shorter unites, like paragraphs.
  - Passage ranking, features
    - Number of Named Entities of the right type in passage
    - Number of query words in passage
    - Number of question N-grams also in passage
    - Proximity of query keywords to passage
    - Longest sequence of question words
    - Rank of the document containing passage
    - …

Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$ – Web QnA

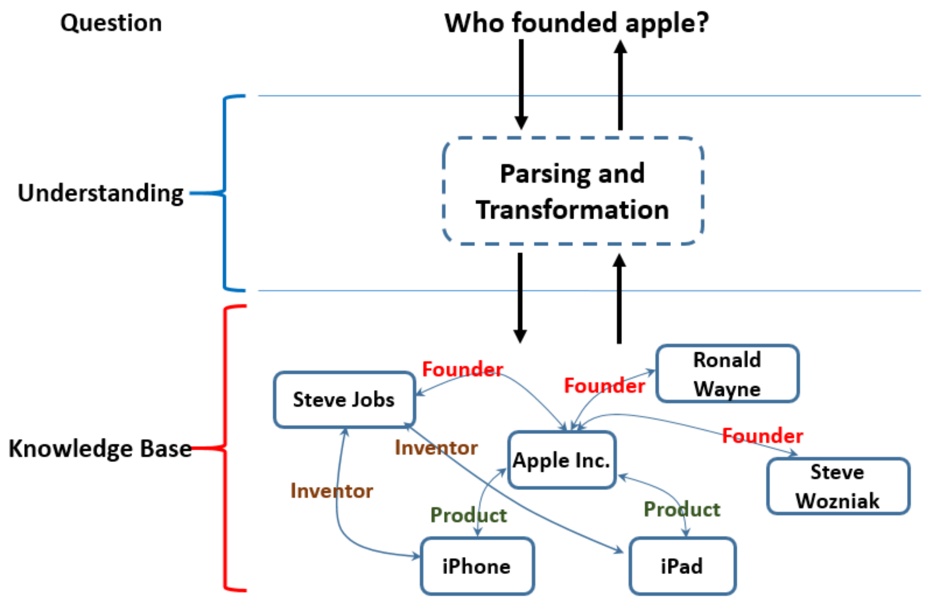
- Run an answer-type named-entity tagger on the passages
  - Each answer type requires a named-entity tagger that detects it
  - If answer type is CITY, tagger has to tag CITY
- Return the string with the right type:
  - How many bones in a human body? (Number)
    - The human skeleton is the internal framework of the body. It is composed of 270 bones at birth – this total decreases to 206 bones by adulthood after some bones have fused together.



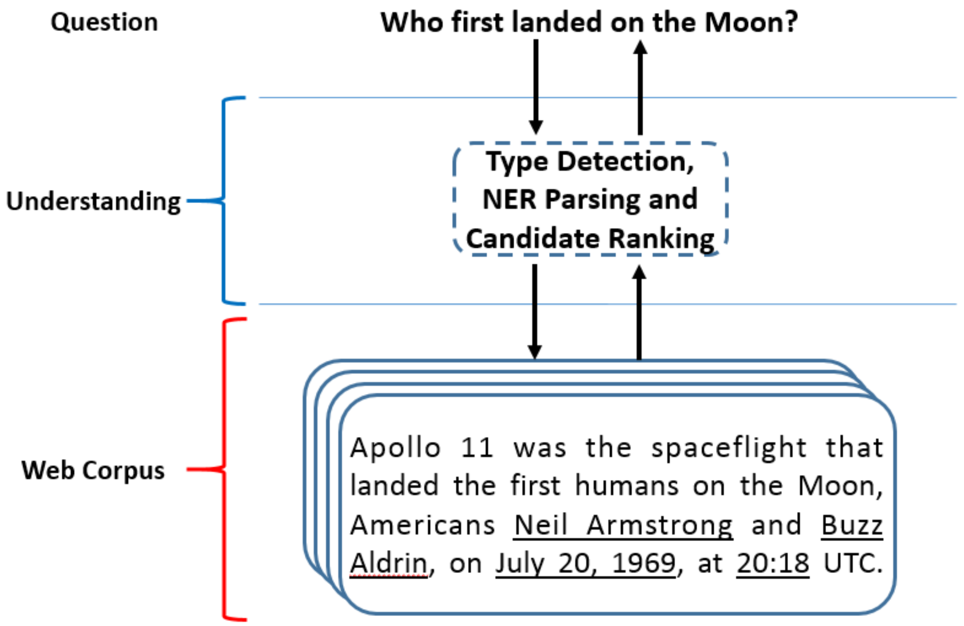
Question Answering [Dan Jurafsky, Stanford]

$P(entity|query)$

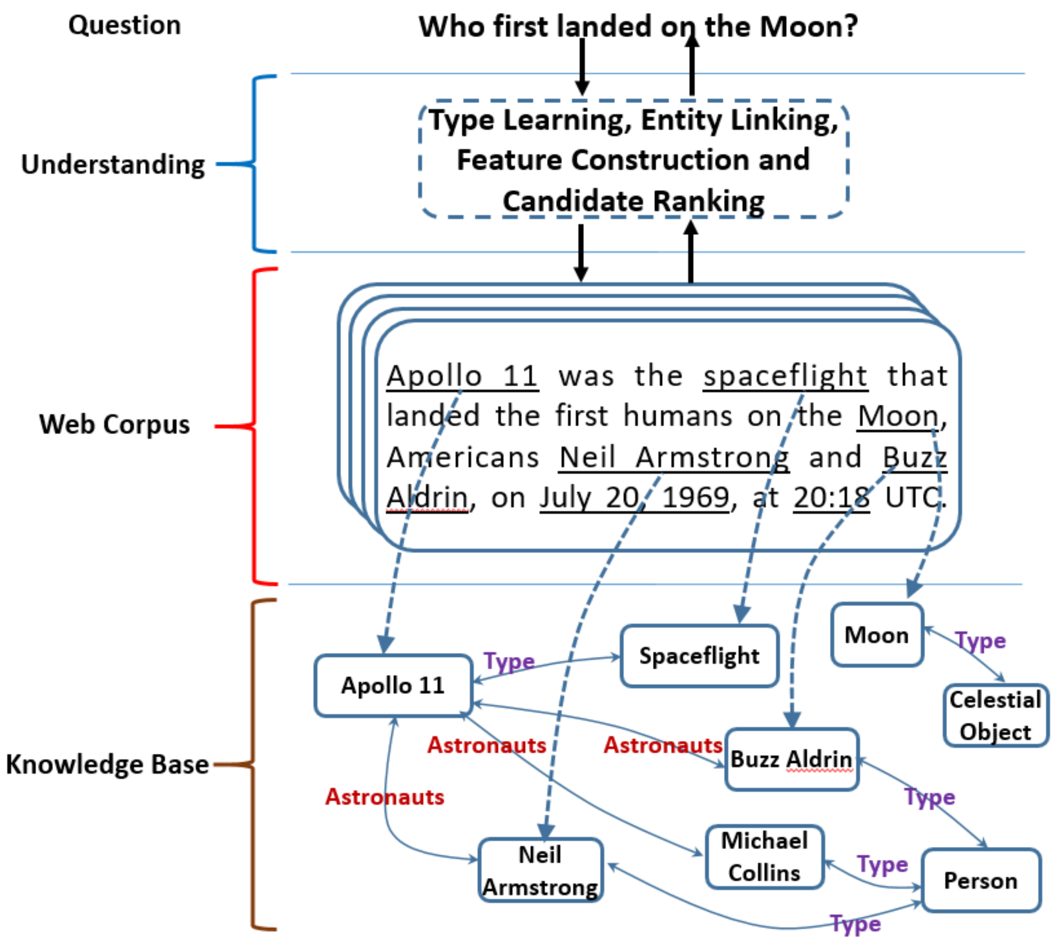
KB QnA



Web QnA



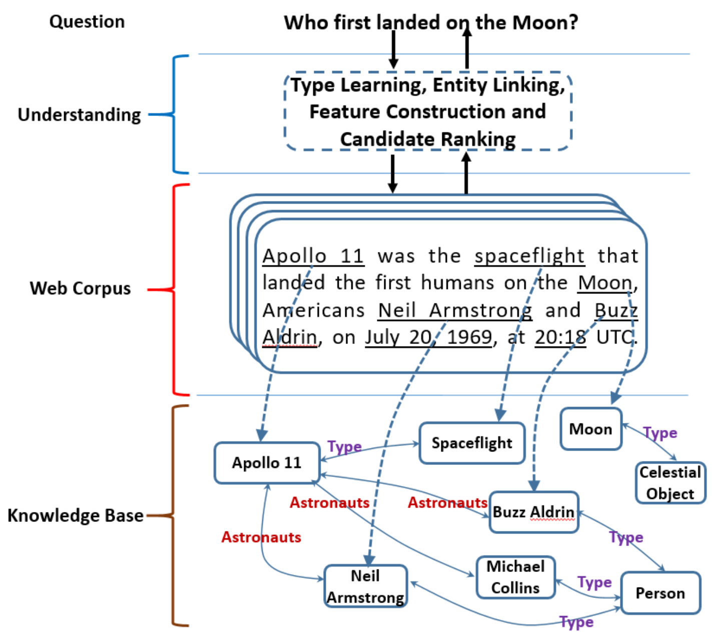
$P(entity|query)$ – Web-KB QnA



Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

$P(entity|query)$ – Web-KB QnA

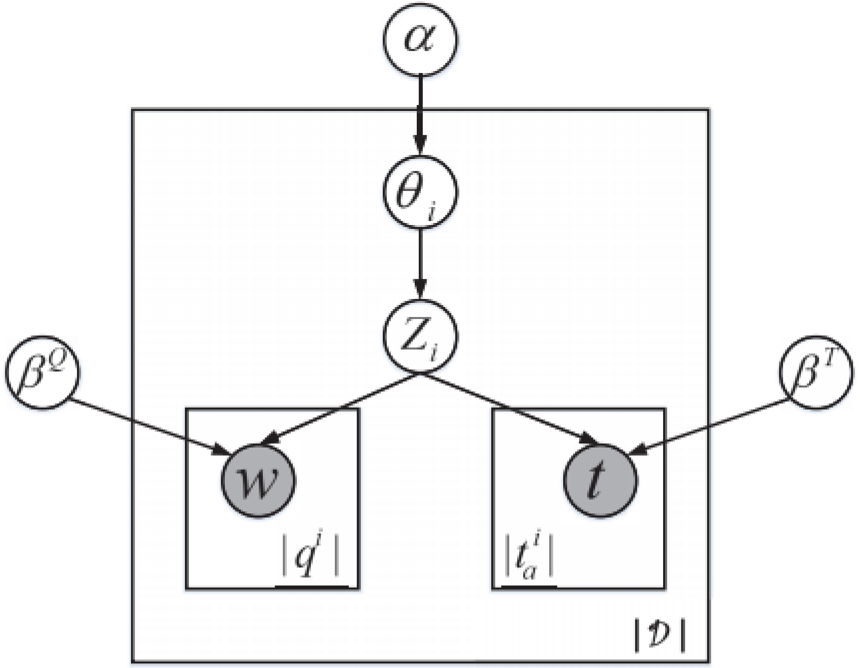
- Advantages
  - Entity Linking: Reduce redundancy among answer candidates
  - Answer candidates ⟶ Freebase types
  - Freebase information ⟶ semantic features for ranking



Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

$P(entity|query)$ – Web-KB QnA

- Type detection is modeled latently



The likelihood of observing a question $\boldsymbol{q}$ and its answer types $\boldsymbol{t}_a$ as:

$$\mathcal{L} = \prod_{i \in \mathcal{D}} P(q^i, t_a^i | \alpha, \beta^Q, \beta^T)$$

$$= \prod_{i \in \mathcal{D}} \int_{\theta_i} P(\theta_i|\alpha) P(q^i, t_a^i | \theta_i, \beta^Q, \beta^T) \, d\theta_i$$

Variational EM to optimize

$$\lfloor \log \mathcal{L} \rfloor = \sum_{i \in \mathcal{D}} E_Q \log P(\theta_i|\alpha) + \sum_{i \in \mathcal{D}} E_Q \log P(Z_i|\theta_i)$$
$$+ \sum_{i \in \mathcal{D}} E_Q \log P(w \in q^i, t \in t_a^i | Z_i, \beta^Q, \beta^T)$$
$$+ H(Q(\theta, Z))$$

Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

$P(entity|query)$ – Web-KB QnA

- Experiments
  - Search Queries

	Systems	MRR	Precision	Recall	F_1
	QuASE	**0.6402**	**0.5962**	**0.5691**	**0.5823**
Web QnA	AskMSR+	0.5337	0.3782	0.3760	0.3771
KB QnA	SEMPRE	0.2372	0.2646	0.1940	0.2239

Open Domain Question and Answering via Semantic Enrichment [Huan Sun, et al., WWW 2015]

Outline

- Introduction to Entity and Knowledge
- Demonstration of Microsoft’s Entity Experience
- Entity Recommendation and Understanding
  - $P(entity|entity)$
  - $P(entity|user)$
  - $P(entity|query)$
- Summary

Summary

Entity Recommendation & Understanding Taxonomy

- $P(entity|entity)$ - Recommendations given an entity
  - Co-occurrence
  - Similarity
  - Entity Linking
  - Interpretation
- $P(entity|user)$ - Recommendations given a user
  - Universal Recommender System
  - $P(entity|user, item)$
  - $P(entity|user, query)$
- $P(entity|query)$ - Recommendations given a query
  - Entity Retrieval/Finding
  - Knowledge Base Question and Answering
  - Web Question and Answering

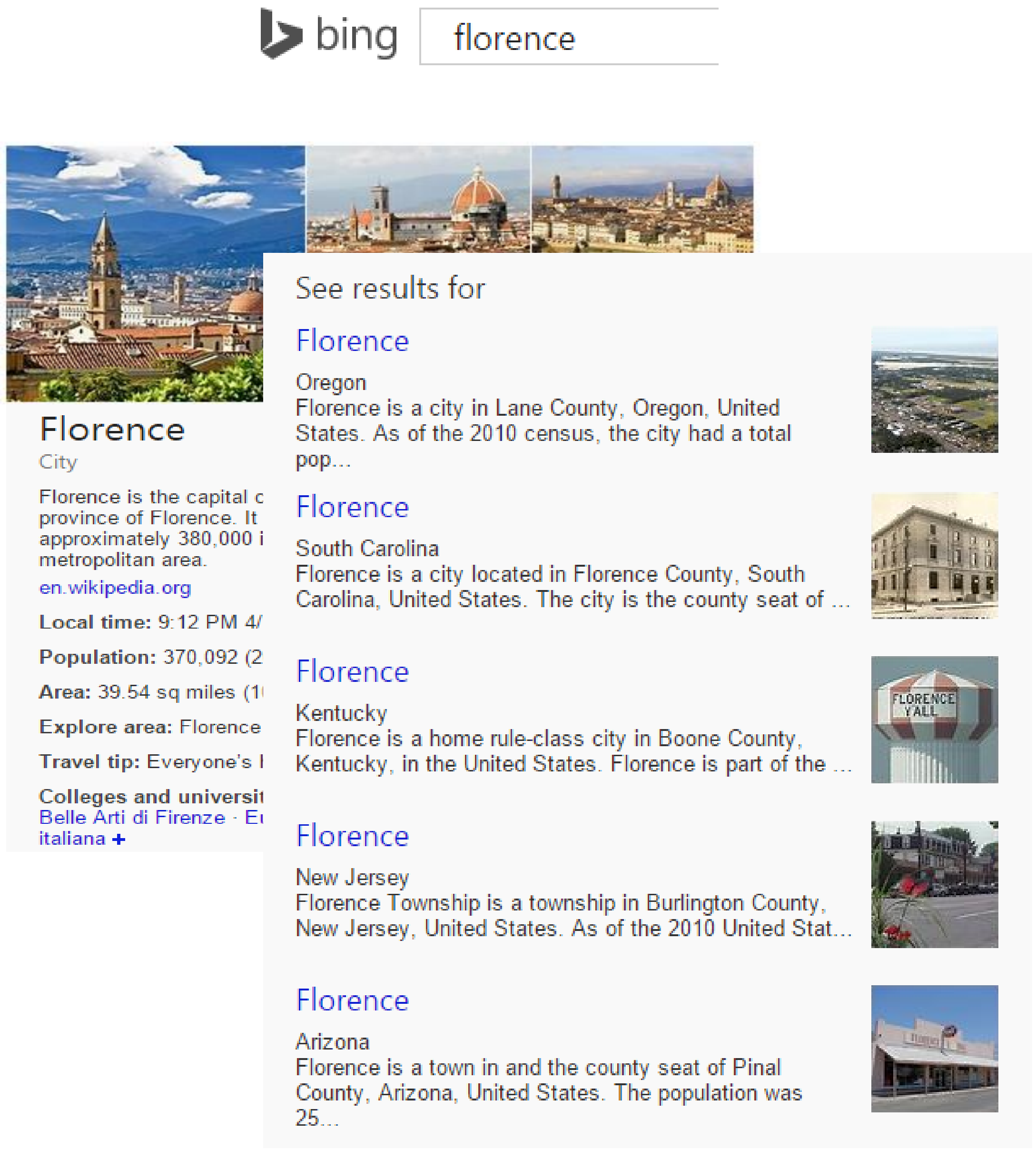
Challenges

- Entity Understanding
  - Ranking, KB completion, Entity Triggering ($P(query|entity)$), …
- User Understanding
  - Users' Entity Preference, Interest Drift, Multiple Sources (query, url click, entity pane click), …
- Query Understanding
  - Query Intent (definition, list, factoid, question, etc.), Question Type, …
- Document Understanding
  - Entity Linking, NER, Event Detection, …

Challenges

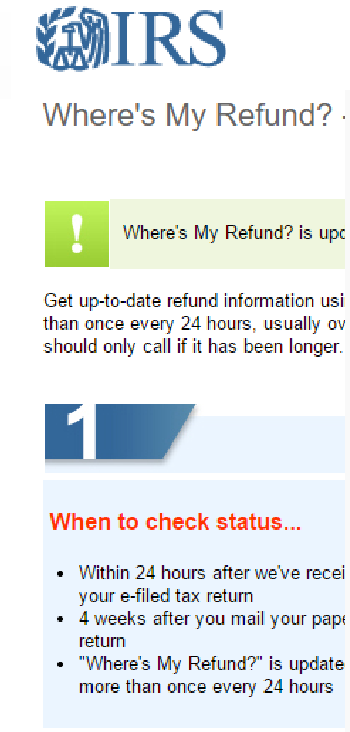
- Query Entity Linking
  - Short and noisy
  - When a user types “Florence”, which one to link?
  - Utilize user location
  - Utilize previous queries in the same session

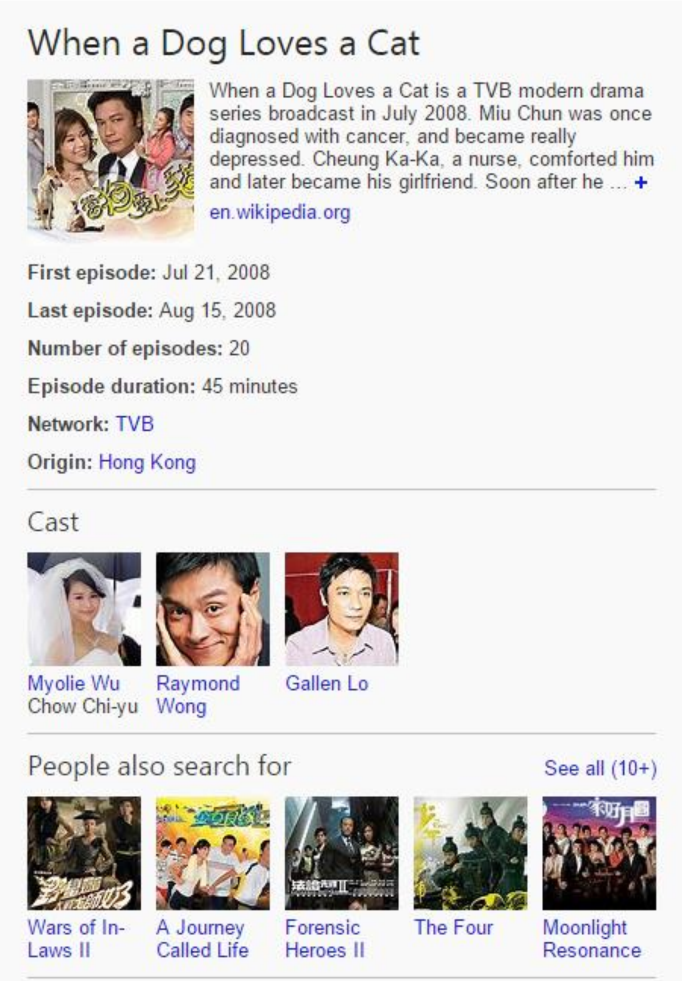
Personalized Entity Linking System



Challenges

- Question Understanding
  - Rules are not always correct
  - “where is my refund”
    - location?
    - When and how to get refund
  - “when a cat loves a dog”
    - Date Time?
    - TV series





Challenges

- Question and Answering
  - TREC data - Web QnA
  - WebQuestions data – KB QnA
  - All the question in these research datasets are real and valid questions
    - Who first landed on the moon
    - Who killed Abraham Lincoln
  - Real world scenario
    - When is the end of the world
    - Who won the world cup 2017
  - A data set contains both valid and invalid questions
    - Make sure the algorithms won’t return answers for invalid questions

> The ancient Mayans predicted the world would end on **December 21, 2012**. Or rather, those interpreting the Mayan calendar – which ran out on **December 21** – guessed it ran out then as there was no world beyond this last day. Aug 18, 2014
>
> Our days are numbered: 7 end of world predictions you've ...
> www.mirror.co.uk/news/weird.../days-numbered-7-**end**-**world**-4069965

Related Tutorials

- Entity Linking and Retrieval (Meij, Balog and Odijk)
  - http://ejmeij.github.io/entity-linking-and-retrieval-tutorial/
- Entity Resolution (Getoor and Machanavajjhala)
  - http://www.umiacs.umd.edu/~getoor/Tutorials/ER_VLDB2012.pdf
- Constructing and Mining Web-scale Knowledge Graphs tutorial (Bordes, Gabrilovich)
  - http://www.cs.technion.ac.il/~gabr/publications/papers/KDD14-T2-Bordes-Gabrilovich.pdf
- The Recommender Problem Revisited (Amatriain, Mobasher)
  - http://www.slideshare.net/xamat/kdd-2014-tutorial-the-recommender-problem-revisited
- Question Answering Lecture (Jurafsky)
  - https://web.stanford.edu/class/cs124/lec/qa.pdf

Thanks!

Hao Ma: haoma at Microsoft.com

Yan Ke: yanke at Microsoft.com