

Entity Linking:

A Symbiosis between Knowledge Graphs and News, and Bringing the Knowledge Graph to Excel

Silviu Cucerzan
Microsoft Research



Entities and Entity Mentions

Entities: unique world bodies with (non-unique) names, such as people, organizations, locations

Unique KG Identifier

e.g. **Washington County**

Entity Types: categories of entities

(e.g. **administrative division #n**)

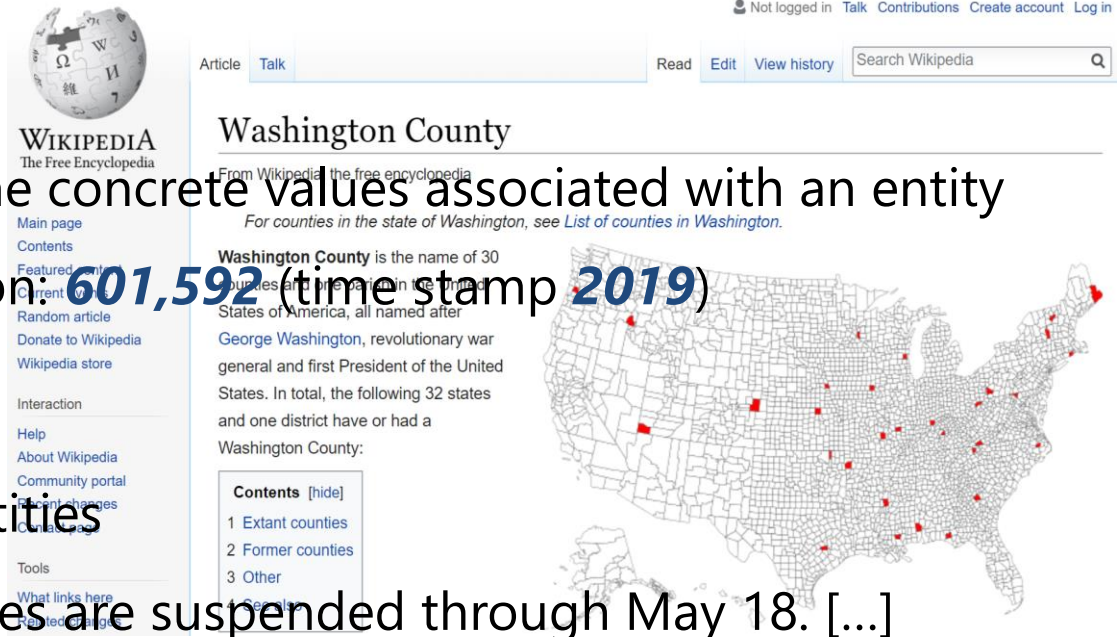
Entity Attributes and *Values*: properties and the concrete values associated with an entity

e.g. **population** of Washington County, Oregon: **601,592** (time stamp **2019**)

attribute
value

Entity Mentions: text fragments that refer to entities

e.g. "**Washington County** public facing services are suspended through May 18. [...]
(#StayHomeMN)"



Entity Linking

EDL (**E**ntity **D**etection and **L**inking)
ERD (**E**ntity **R**ecognition and **D**isambiguation)

Find *entity mentions* in text and assign them an ID from a knowledge repository.

e.g.

movie

year

director

- *2012 is a 2009 movie by Roland Emmerich.*

justice

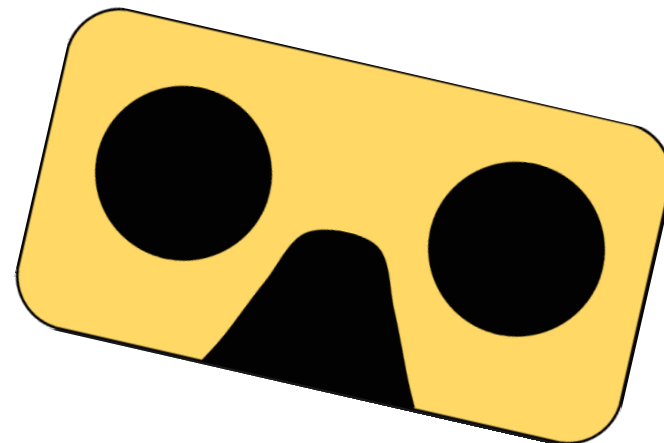
justice

justice

- *Kagan and Breyer disagreed with Roberts.*

?

- *Seattle is the best!*
 - *Seattle defeated the Red Bulls.*
 - *Seattle defeated the Red Stars.*



NEMO (Named Entities Made Obvious)

State-of-the-art EDL system

- built in MSR starting 2005
- best score in NIST's 2011, 2012, 2013, and 2014 TAC-KBP-EL evaluations and in Google-Microsoft-Yahoo 2014 ERD competition
- in production in Microsoft starting 2008

Framework:

The best evidence for entity disambiguation is provided by the set of co-occurring entities

- extract and disambiguate jointly all entities in a target document
- employ both *observable* and *latent* attributes

Syntax and local context are important; one-sense-per-discourse does not hold

- employ both whole-document and local-context features

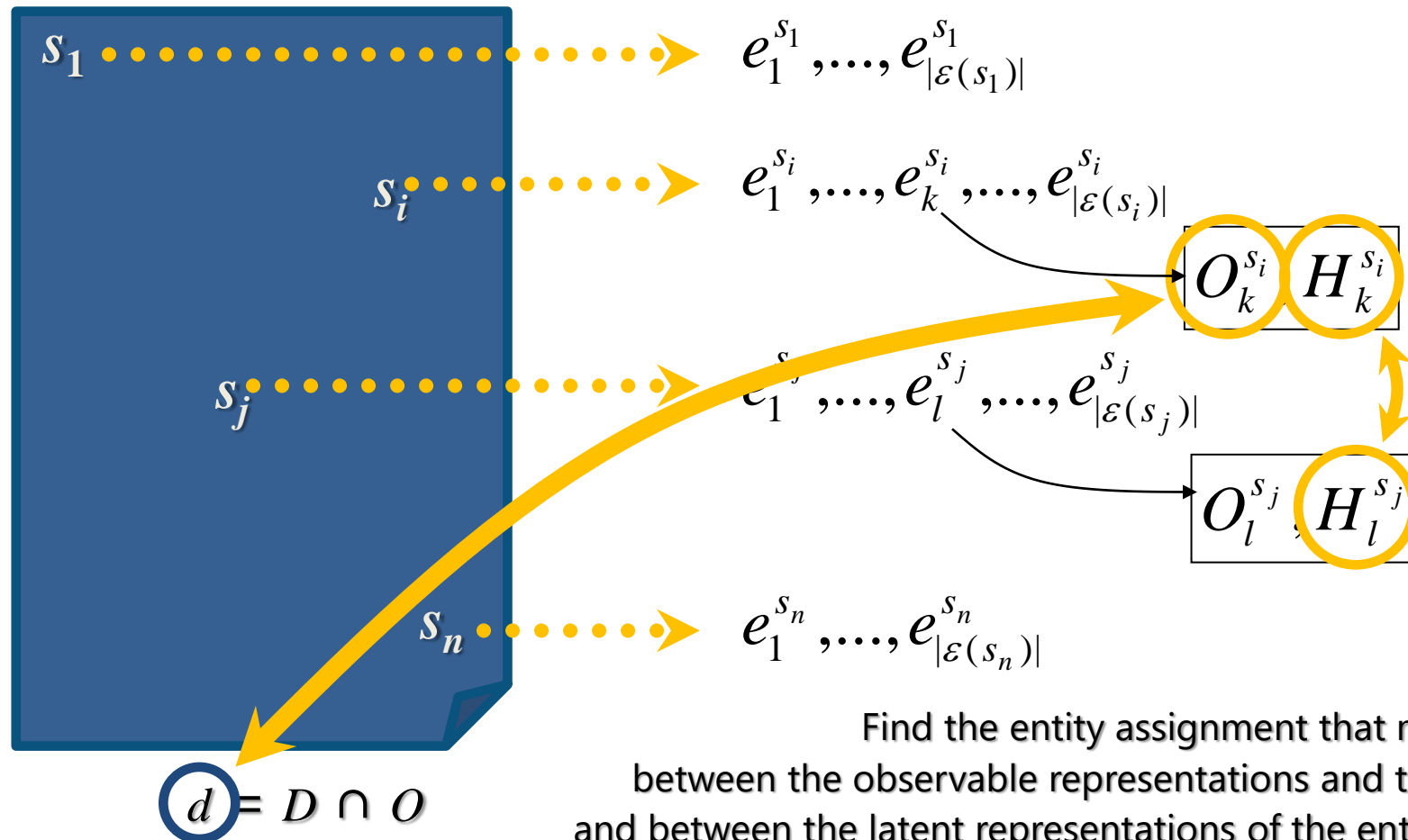


NEMO Disambiguation Paradigm

Each entity has multiple vectorial representations

observable
latent / hidden

Text document D



Find the entity assignment that maximizes the similarity between the observable representations and the document context d and between the latent representations of the entities in the assignment.

From Documents to Entity Graphs

REVIEW

Food Security: The Challenge of Feeding 9 Billion People

H. Charles J. Godfray,^{1,2} John R. Beddington,³ Ian R. Crute,³ Lawrence Haddad,⁴ David Lawrence,⁵ James F. Muir,⁶ Jules Pretty,⁷ Sherman Robinson,⁸ Sandy M. Thomas,⁹ Camilla

Continuing population and consumption growth will mean that the global demand increase for at least another 40 years. Growing competition for land, water, and energy, the overexploitation of fisheries, will affect our ability to produce food, as will the effects of climate change. But the world can produce more food and can ensure that it is used more equitably. A multifaceted and linked global strategy is needed to ensure sustainable security, different components of which are explored here.

The past half-century has seen marked growth in food production, allowing for a dramatic decrease in the proportion of the world's people that are hungry, despite a doubling of the total population (Fig. 1) (1, 2). Nevertheless, more than one in seven people today still do not have access to sufficient protein and energy from their diet, and even more suffer from some form of micronutrient malnourishment (3). The world is now facing a new set of intersecting challenges (4). The global population will continue to grow, yet it is likely to plateau at some 9 billion people by roughly the middle of this century. A major correlate of this deceleration in population growth is increased wealth, and with higher purchasing power comes higher consumption and a greater demand for processed food, meat, dairy, and fish, all of which add pressure to the food supply system. At the same time, food producers are experiencing greater competition for land, water, and energy, and the need to curb the many negative effects of food production on the environment is becoming increasingly clear (5, 6). Overarching all of these issues is the threat of the effects of substantial climate change and concerns about how mitigation and adaptation measures may affect the food system (7, 8).

A threefold challenge now faces the world (9): Match the rapidly changing demand for food

¹Department of Zoology and Institute of Biodiversity at the James Martin 21st Century School, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. ²U.K. Government Office for Science, 1 Victoria Street, London SW1H 0ET, UK. ³Agriculture and Horticulture Development Board, Stoneleigh Park, Kenilworth, Warwickshire CV8 2TL, UK. ⁴Institute of Development Studies, Falmes, Brighton BN1 9BE, UK. ⁵Syngenta AG, Post Office Box, CH-4002 Basel, Switzerland. ⁶Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, UK. ⁷Department of Biological Sciences, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK. ⁸Institute of Development Studies, Falmes, Brighton BN1 9BE, UK. ⁹Foreign Office, U.K. Government Office for Science, 1 Victoria Street, London SW1H 0ET, UK. ¹⁰International Institute for Environment and Development, 3 Endsleigh Street, London WC1H 0DD, UK.

*To whom correspondence should be addressed. E-mail: charles.godfray@ox.ac.uk

from a larger and more affluent supply; do so in ways that are socially sustainable; and ensure that the world's poorest people are not left behind. This challenge requires changes in production, storage, processing, and distribution that are as radical as

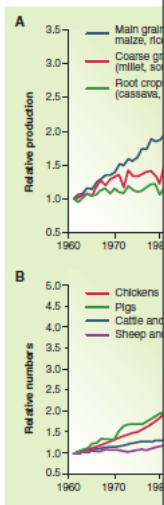


Fig. 1. Changes in the relative numbers of major crop plants and livestock since 1961 (when relative numbers = 1.0). (A) Major crop plants and (B) livestock. Data are from the Food and Agriculture Organization of the United Nations (10). [Source: (2)]

during the 18th and 19th-century Industrial and Agricultural Revolutions and the 20th-century Green Revolution. Increases in production will have an important part to play, but they will be constrained as never before by the finite resources provided by Earth's lands, oceans, and atmosphere (10).

Patterns in global food prices are indicators of trends in the availability of food, at least for those

Vol 463 | 11 February 2010 | doi:10.1038/nature08823

PERSPECTIVE

Measuring Food Insecurity

Christopher B. Barrett*

Food security is a growing concern worldwide. More than 1 billion people are estimated to lack sufficient dietary energy availability, and at least twice that number suffer micronutrient deficiencies. Because indicators inform action, much current research focuses on improving food insecurity measurement. Yet estimated prevalence rates and patterns remain tenuous because measuring food security, an elusive concept, remains difficult.

The 2008 global food price crisis, which sparked riots in more than two dozen countries, has focused attention on the need for economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life.

and nutritious food." Access is most closely related to social science concepts of individual or household well-being: What is the range of food choices open to the person(s), given their income, assets, and access to formal or informal safety net mechanisms, and the ways in which they can access them? (2) No doubt, as Amartya Sen wrote, "starvation is the characteristic of some people not having enough food to eat. It is not the characteristic of there being not enough food to eat. While the latter can be a cause of the former, it is but one of many possible causes" (7). Access reflects the demand side of food security, as manifest in uneven inter- and intra-household food distribution and in the sociocultural limits on what foods are consistent with prevailing tastes and values within a community. Access also encompasses problems in responding to adverse shocks such as unemployment spells, price increases in food-producing assets, and natural disasters, which's close

duce food, as linked global economy today still is likely to eat, dairy,

PERSPECTIVES

The next generation of scenarios for climate change research and assessment

Richard H. Moss¹, Jae A. Edmonds¹, Kathy A. Hibbard², Martin R. Manning³, Steven K. Rose⁴, Detlef P. van Vuuren⁵, Timothy R. Carter⁶, Seita Emori⁷, Mikiko Kainuma⁸, Tom Kram⁹, Nebojsa Nakicenovic^{10,11}, Keywan Riahi¹², Steven J. Smith¹, Ronald J. Stouffer¹³, John P. Weyant¹² & Thomas J. Wilbanks¹³

Advances in the science and observation of climate change are providing a better understanding of Earth's climate system and its likely response to human and natural environmental and societal changes. This understanding will depend not only on the response of the Earth system to changes in technology, economics, and society, but also on how humankind responds through changes in technology, economics, and society. Future forcings of and responses to climate change, necessitating the consequences of different response options. To date, such scenarios have been developed by climate change mitigation and adaptation, and have relied on research from physical, biological and social scientists. Here we describe a new generation of scenarios for climate change research and assessment.

To improve understanding of the complex interactions of the climate system, ecosystems, and human activities and conditions, the research community develops and uses scenarios. These scenarios provide plausible descriptions of how the future might unfold in several key areas—socioeconomic, technological and environmental conditions, emissions of greenhouse gases and aerosols, and climate. When applied in climate change research, scenarios help to evaluate uncertainty about human contributions to climate change, the response of the Earth system to human activities, the impacts of a range of future climates, and the implications of different approaches to mitigation (measures to reduce net emissions) and adaptation (actions that facilitate response to new climate conditions).

Traditionally, model-based scenarios used in climate change research have been developed using a sequential process focused on a step-by-step and time-consuming delivery of information between separated scientific disciplines. Now, climate change researchers from different disciplines have established a new coordinated parallel process for developing scenarios. This starts with four scenarios of future radiative forcings (the change in the balance between incoming and outgoing radiation to the atmosphere caused by changes in atmospheric constituents, such as carbon dioxide). Using this starting point, the parallel process will encourage research that will characterize a broad range of possible future climate conditions, taking into account recent climate observations and new information about climate system processes. Studies will give more attention to evaluating adaptation needs and strategies, exploring mitigation options, and improving understanding of potentially large feedbacks (that is, impacts of climate change such as melting of permafrost or dieback of forests that cause further changes in climate).

Cent radiative economic issues. "What a particular this adapt assess plex is choices will red their ul climate This climate with 'vide a research Altern The use in the e and oth a system decision scenario certain range c

PERSPECTIVE

Radically Rethinking Agriculture for the 21st Century

N. V. Fedoroff^{1,2}, D. S. Battisti³, R. N. Beachy⁴, P. J. M. Cooper⁴, D. A. Fischhoff⁵, C. N. Hodges⁶, V. C. Knauf⁷, D. Lobell⁸, B. J. Mazur⁹, D. Molden¹⁰, M. P. Reynolds¹¹, P. C. Ronald¹², M. W. Rosegrant¹³, P. A. Sanchez¹⁴, A. Vonshak¹⁵, J.-K. Zhu¹⁶

Population growth, arable land and fresh water limits, and climate change have profound implications for the ability of agriculture to meet this century's demands for food, feed, fiber, and fuel while reducing the environmental impact of their production. Success depends on the acceptance and use of contemporary molecular techniques, as well as the increasing development of farming systems that use saline water and integrate nutrient flows.

Population experts anticipate the addition of another roughly 3 billion people to the planet's population by the mid-21st century. However, the amount of arable land has not changed appreciably in more than half a century. It is unlikely to increase much in the future because we are losing it to urbanization, salinization, and desertification as fast as or faster than we are adding it (1). Water scarcity is already a critical concern in parts of the world (2).

Climate change also has important implications for agriculture. The European heat wave of 2003 killed some 30,000 to 50,000 people (3). The average temperature that summer was only about 3.5°C above the average for the last century. The 20 to 36% decrease in the yields of grains and fruits that summer drew little attention. But if the climate scientists are right, summers will be that hot on average by mid-

century, and by 2090 much of the world will be experiencing summers hotter than the hottest summer now on record. The yields of our most important food, feed, and fiber crops decline precipitously at temperatures much above 30°C (4). Among other reasons, this is because photosynthesis has a temperature optimum in the range of 20° to 25°C for our major temperate crops, and plants develop faster as temperature increases, leaving less time to accumulate the carbohydrates, fats, and proteins that constitute the bulk of fruits and grains (5). Widespread adoption of more effective and sustainable agronomic practices can help buffer crops against warmer and drier environments (6), but it will be increasingly difficult to maintain, much less increase, yields of our current major crops as temperatures rise and drylands expand (7). Climate change will further affect agriculture as the sea level rises, submerging low-lying cropland, and as glaciers melt, causing river systems to experience shorter and more intense seasonal flows, as well as more flooding (7).

Recent reports on food security emphasize the gains that can be made by bringing existing agronomic and food science technology and know-how to people who do not yet have it (8, 9), as well as by exploring the genetic variability in our existing food crops and developing more ecologically sound farming practices (10). This requires building local educational, technical, and research capacity; food processing capability; storage capacity; and other capabilities. Capabilities such as well-

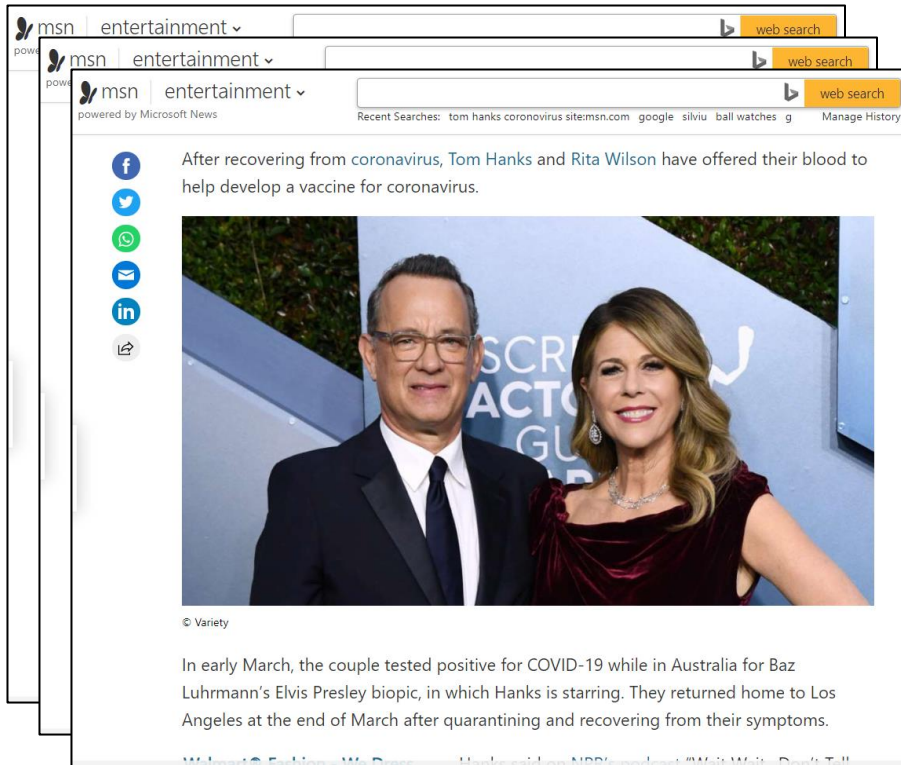


Downloaded from www.sciencemag.org on February 13, 2010

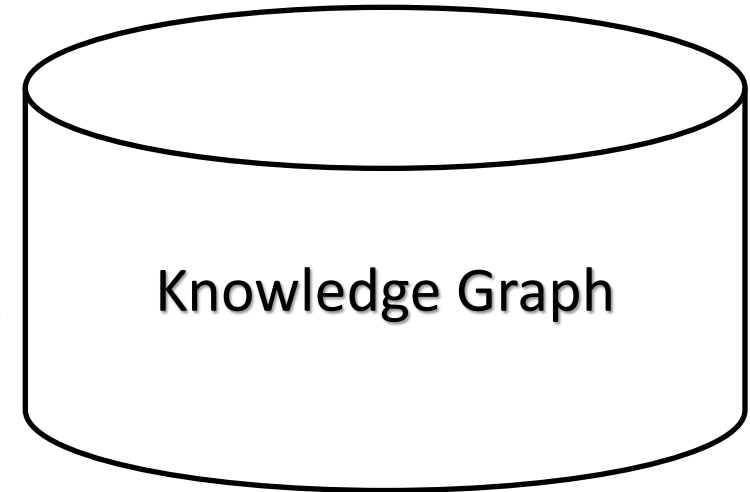
economy that are not reflected in the water shortages due to overextracted nonwidely recognized that food production nonrenewable inputs is unsustainable

Kathy A. Hibbard

News Mining for Knowledge Graphs



additions / changes
to be made



- Setting:

- a comprehensive KG
- reliable entity linking technology
- very high precision requirement (.99)

→ distance supervision paradigm

Distant Supervision Framework for Knowledge Mining



Full Name
William James Murray

How tall is Bill Murray?
1.88m or 6feet 2inches

Bill Murray Net Worth
\$140 Million

Bill Murray was born in Wilmette, Illinois, USA (September 21, 1950). Murray's mother is a post room clerk, Lucille and father is a salesman, Edward Joseph. He has eight siblings. On NBC 'Saturday Night Live' was his first TV experience in 1975-1976. In 1979, 'Meatballs' film was Murray's first role that he landed in. After 9 years break, he came back with 'Scrooged' film in 1988. Then in 1989, 'Ghostbusters II'. In 2014, he appeared in 'Dumb and Dumber To'.

The most known movies of Bill Murray;

- Ghost Busters (1984) with **Sigourney Weaver**, Dan Aykroyd
- Groundhog Day (1993) with Andie MacDowell
- Charlie's Angels (2000) with **Cameron Diaz**, **Lucy Liu**, Drew Barrymore
- Lost in Translation (2003) with **Scarlett Johansson**, Giovanni Ribisi
- Moonrise Kingdom (2012) with **Bruce Willis**, **Edward Norton**, Tilda Swinton



Bill Murray (height 6'2") with Melissa McCarthy (height 5'2") in St. Vincent (2014)

In 1977, he won Emmy Award with 'Saturday Night Live' in category of Outstanding Writing for a Variety, Music or Comedy Program.

In 2003, he won Golden Globe Award with 'Lost in Translation' in category of Best Actor – Motion Picture Musical or Comedy.

<http://www.celebheights.org/bill-murray-height>

Full Name

NEMO

William James Murray

targeted KG Id

How tall is Bill Murray?

1.88m or 6feet 2inches

Bill N

\$140

Train distant supervision models by
matching available attribute values

Knowledge Graph

Extract new arcs and attribute values
by employing the trained models

Bill Murray was born in Wilmette, Illinois, USA (**September 21, 1950**). Murray 's mother is a **post room** clerk, Lucille and father is a salesman, Edward Joseph. He has eight siblings. On NBC 'Saturday Night Live' was his first TV experience in **1975-1976**. In **1979**, 'Meatballs' film was Murray 's first role that he landed in. After **9 years** break, he came back with 'Scrooged' film in **1988**. Then in **1989**, 'Ghostbusters II'. In **2014**, he appeared in 'Dumb and Dumber To'.

The most known movies of Bill Murray;

- o Ghost Busters (**1984**) with Sigourney Weaver, Dan Aykroyd
- o Groundhog Day (**1993**) with Andie MacDowell
- o Charlie's Angels (**2000**) with Cameron Diaz, Lucy Liu, Drew Barrymore
- o Lost in Translation (**2003**) with Scarlett Johansson, Giovanni Ribisi
- o Moonrise Kingdom (**2012**) with Bruce Willis, Edward Norton, Tilda Swinton

Bill Murray (**height 6'2"**) with Melissa McCarthy (**height 5'2"**) in St. Vincent (**2014**)

In 1977, he won Emmy Award with 'Saturday Night Live' in **category** of Outstanding Writing for a Variety, Music or Comedy Program.

In 2003, he won Golden Globe Award with 'Lost in Translation' in **category** of Best Actor – Motion Picture Musical or Comedy.

Microsoft

Challenges

- Reliant on highly accurate entity detection and linking
- Very difficult to approach the targeted precision in any one document while maintain reasonable recall

(typical F-measure for fact extraction in low .70s)

- Difficult to deal with humorous discourse and opinions

Bill Murray's Height 6ft 1in (185 cm)

American actor best known for roles in Ghostbusters, GroundHog Day, Scrooged, Caddyshack, Stripes, The Life Aquatic with Steve Zissou, Rushmore and Lost in Translation. He once commented on his height, saying "I'm 6 foot 3 in my boots" when talking about the NBA in 1995. [Lucy Liu](#) made Bill out to be enormous, claiming in Maxim Magazine "He's tall, like, 6'4", and I'm 5'3."



<http://www.celebheights.com/s/Bill-Murray-32.html>

Bill Murray's Height

6ft 1in (185 cm)

American actor best known for roles in Ghostbusters, GroundHog Day, Scrooged, Caddyshack, Stripes, The Life Aquatic with Steve Zissou, Rushmore and Lost in Translation. He once commented on his height, saying "I'm 6 foot 3 in my boots" when talking about the NBA in 1995. Lucy Liu made Bill out to be enormous, claiming in Maxim Magazine "He's tall, like, 6'4", and I'm 5'3."

Challenges

- Reliant on highly accurate entity detection and linking
- Very difficult to approach the targeted precision in any one document while maintain reasonable recall
(typical F-measure for fact extraction in low .70s)
- Difficult to deal with humorous discourse and opinions

News-specific Challenges

- Thousands of news feeds every minute
- Need to decide on the importance and validity of new facts



Traffic-guided Fact Extraction from News

March 31, 2016: Ronnie Corbett

Over 3,000 news stories

News about **Ronnie Corbett**

bing.com/news



Ronnie Corbett, half of
[Yahoo News](#) · 17 min
LONDON (AP) — A dir
timing, Ronnie Corbet

Ronnie Corbett, one of
[CNN](#) · 19 minutes ago
(CNN)Ronnie Corbett,
exquisite timing and a

★ Get news updates

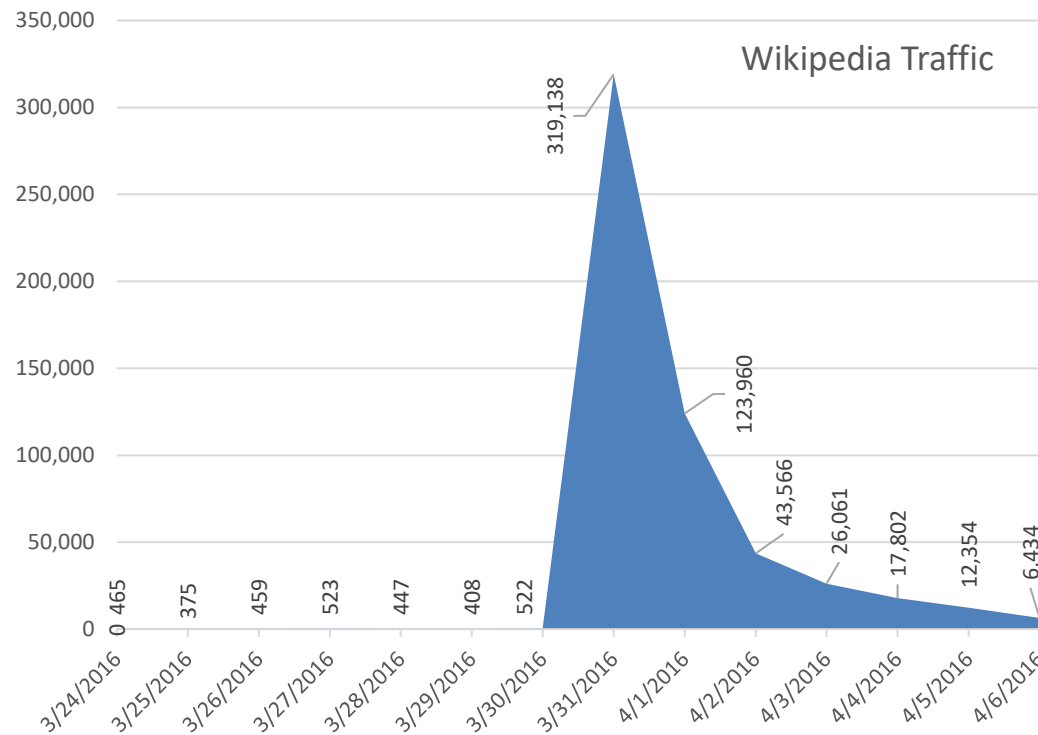
Ronnie Corbett

<https://en.wikipedia.org>
Ronald Balfour "Ronni
comedian who had a lo

Ronnie Corbett

www.bbc.co.uk/news/
Mar 31, 2016 · Comedi

British comedy great **Ronnie Corbett**



E is a Scottish actor and comedian
nie Barker in the television comedy
ed prominence in David Frost's
The Frost Report and later starred
s Me Over Here!

otland

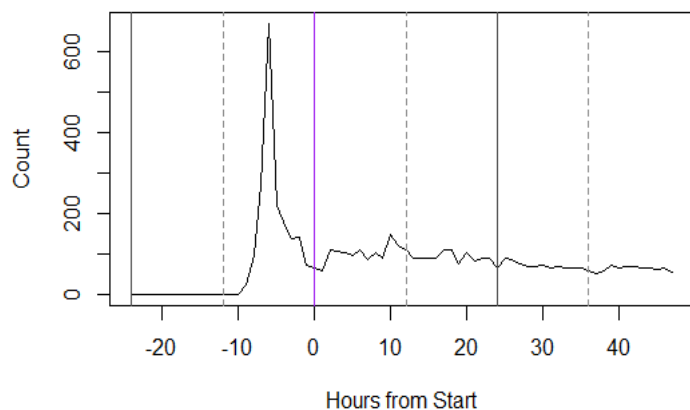
William Balfour Corbett

burgh

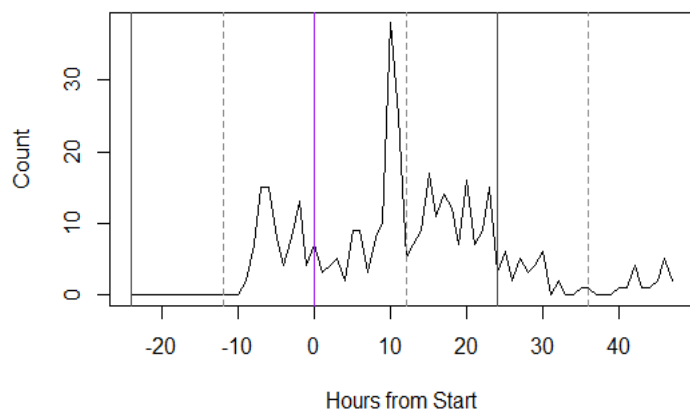
Traffic-guided Fact Extraction from News

News retrieved on the date of spike
for spiking entity in search queries:

Before Deduping

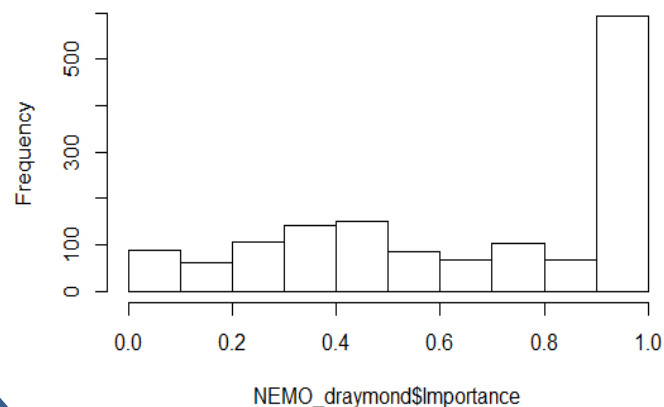


After Deduping

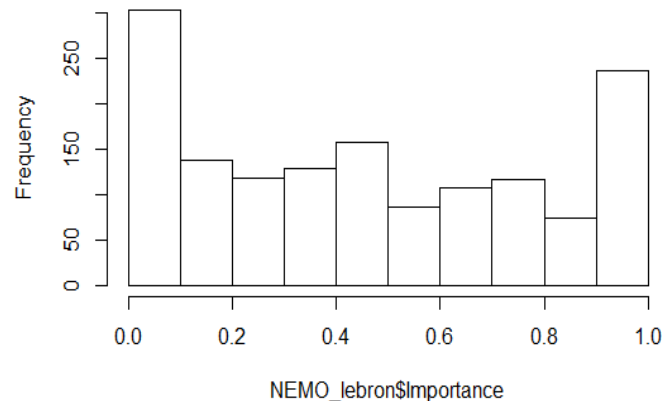


Importance of spiking entity
in the retrieved news:

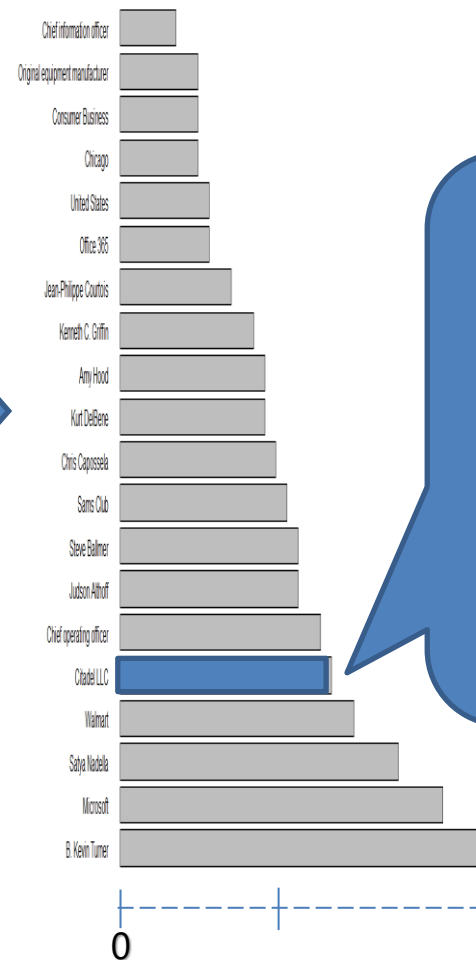
Importance of spiking entity



Importance of co-occurring entity



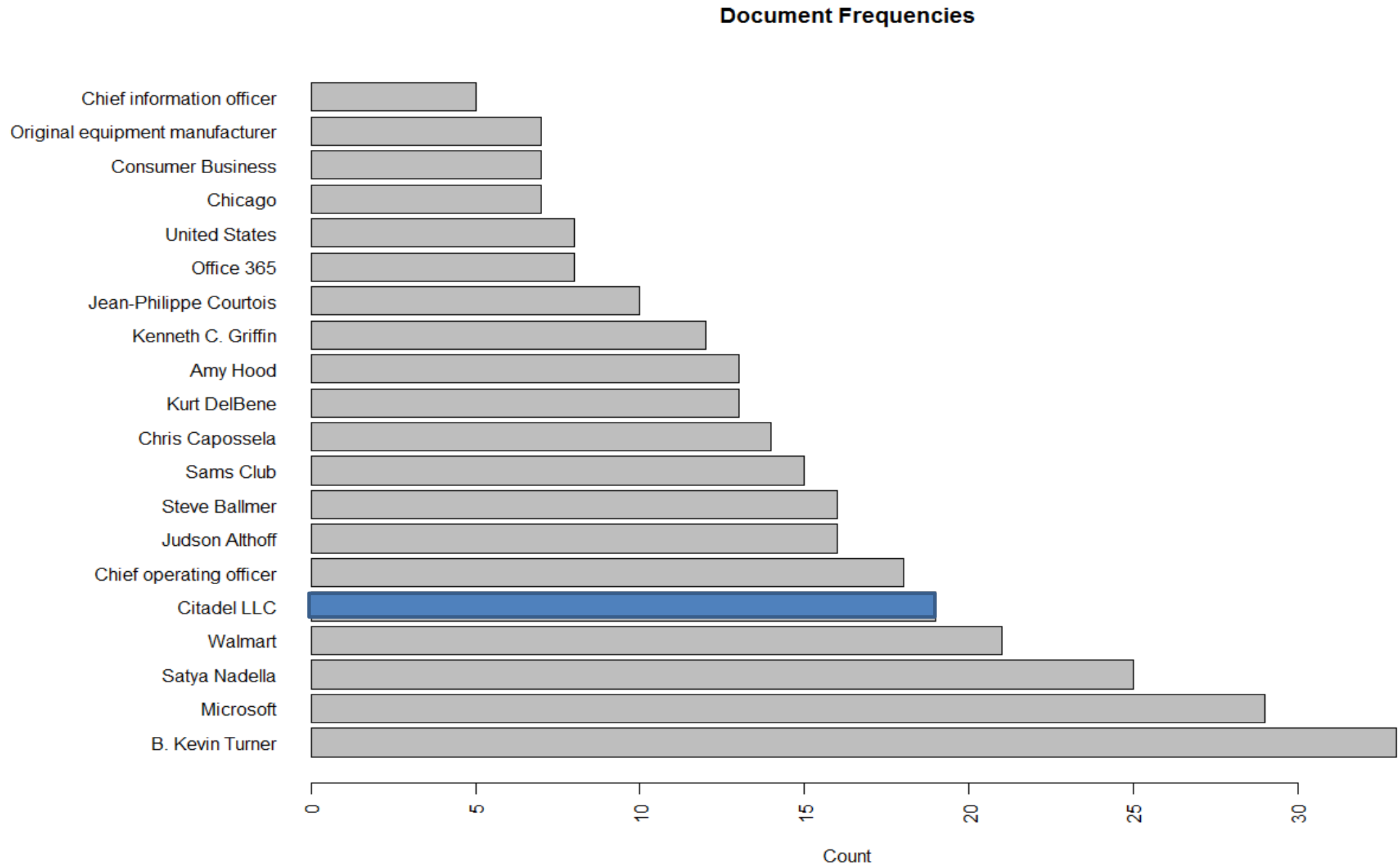
Normalized DF statistics for
co-occurring entities/attributes



Analyze all sentences that refer both the spiking entities and the selected co-occurring entity/attribute and compare with the data in the KG



e.g.: B. Kevin Turner (July 7, 2016)



e.g. B. Kevin Turner → Citadel LLC

The software giant announces changes to its executive lineup as its Chief Operating Officer [Kevin Turner](#) prepares to depart at the end of the month to run [Citadel](#) Securities. [Microsoft](#)'s chief operating officer, [Kevin Turner](#), is leaving the company to serve as **CEO** of [Citadel](#) Securities, a [Chicago-based](#) investment company, the [Redmond, Wash.](#), software maker announced July 7.

[Microsoft](#) chief operating officer [Kevin Turner](#) is leaving the software company to lead a unit of [Citadel](#), the [Chicago-based](#) investment firm.

[Microsoft](#) chief operating officer [Kevin Turner](#), who heads the company's partner channel and worldwide sales, is leaving to join [Citadel](#) Securities, according to a release by [Citadel](#).

[Microsoft Corp.](#) revamped its sales organization Thursday morning after its chief operating officer, [Kevin Turner](#), took a job at another company.

Global securities firm [Citadel](#) has hired a new key executive — from [Microsoft](#). [B. Kevin Turner](#), who was chief operating officer of [Microsoft](#) for 11 years, has joined [Citadel](#) as **CEO** of [Citadel](#) Securities, the hedge fund's market making business, and vice chairman of [Citadel](#).

[Microsoft Corp.](#) announced a revamp of its sales organization Thursday after [Kevin Turner](#), the company's chief operating officer, took a job as head of [Citadel LLC](#)'s securities unit. [Kevin Turner](#), [Microsoft](#)'s hard-charging and controversial **COO** whose folksy zingers at competitors were a highlight of the annual [Microsoft Worldwide Partner Conference \(WPC\)](#), is leaving [Microsoft](#) at the end of July to become **CEO** of [Citadel](#) Securities and vice chairman of [Citadel LLC](#). [Turner](#) has been [Microsoft](#)'s chief operating officer since 2005, after a 20-year career with [Walmart](#) that spanned roles from cashier to chief information officer and, ultimately, president and **CEO** of the [Sam's Club](#) division.

[Turner](#)'s new destination will be at leading financial firm [Citadel](#) Securities, where he will be replacing founder [Kenneth Griffin](#) as **CEO**, while also serving as Vice Chairman of the Firm. [Citadel](#) Securities and [Griffin](#) praises [Turner](#)'s leadership prowess and experience at major corporations, as well as his technology initiatives during the long career.

Hedge funds are poaching top executives from tech giants to manage their transformation into computer-driven firms full of engineers and mathematicians.

[Kevin Turner](#), [Microsoft](#) chief operating officer for the past 11 years, is moving to [Citadel](#) Securities, where he will be chief executive officer.

[Turner](#) is leaving [Microsoft](#) after more than a decade at the company to become **CEO** of the securities unit at financial-services firm [Citadel LLC](#).

[Citadel](#) Securities announced Thursday morning that it has hired [Kevin Turner](#) away from [Microsoft](#) to be **CEO**. [Turner](#) was **COO** running [Microsoft](#)'s enormous sales operations.

Chief Operating Officer [Kevin Turner](#) will become **CEO** of financial firm [Citadel](#) Securities, [Microsoft](#) **CEO** [Satya Nadella](#) said in an email to employees Thursday.

[SEATTLE](#) — A top [Microsoft](#) executive, [Kevin Turner](#), is leaving the company to join [Citadel](#) Securities as its new chief executive, continuing a wave of executive departures of technology industry leaders to financial firms.

[Turner](#) will remain at [Microsoft](#) to aid with the transition through this month, then leave to become chief executive officer at global financial firm [Citadel](#) Securities, according to [Nadella](#).

[B. Kevin Turner](#) was named **CEO** of [Citadel](#) Securities and vice chairman of its parent company, hedge fund manager [Citadel](#).

[Microsoft Corp.](#) disclosed a revamp of its sales organization on Thursday after [Kevin Turner](#), the company's chief operating officer, took a job as head of hedge-fund [Citadel LLC](#)'s securities unit.

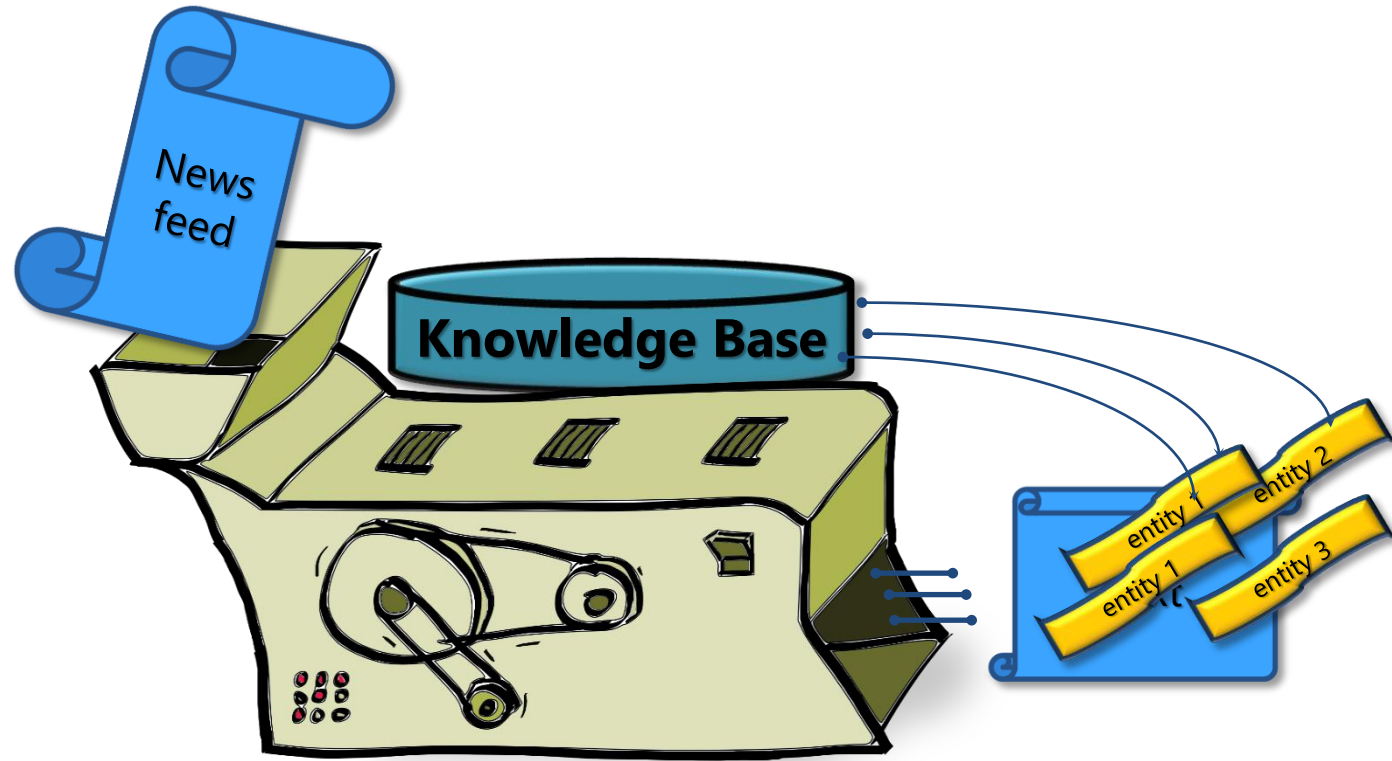
[Turner](#) will join [Citadel](#) as chief executive officer of the company's securities business, which is centered on a multifaceted trading platform that handles orders for everything from stocks to interest rate swaps.

([Bloomberg](#)) — [Citadel LLC](#) hired [Kevin Turner](#), the chief operating officer at [Microsoft](#), to become **CEO** of its securities unit as financial-services firms seek talent to help build their technology expertise.

The turnover of senior staff from [Microsoft](#) has continued with the news that its chief operating officer of 11 years, [Kevin Turner](#), is to leave to become **CEO** and vice chairman at financial services company [Citadel](#) Securities. [Turner](#) (pictured) was appointed **COO** of [Microsoft](#) in 2005, prior to which he held various positions at retailer [Wal-Mart](#), including **CIO** for [Wal-Mart Stores](#). [Turner](#) had joined the retailer as a cashier while studying for his undergraduate degree.

[Kevin Turner](#) will join [Citadel](#), a global financial firm, as chief executive officer of [Citadel](#) Securities and Vice Chairman of [Citadel](#).

Applications to News Processing and Presentation



Indexing and Annotation of News Video

MEET THE PRESS
Feb. 26
Presidential candidate Rick Santorum discusses the crisis in Afghanistan, the upcoming Michigan primaries and the role more

Find term in this clip:
religion
higher education
economy
Michigan
Rick Santorum

Washington, D.C.
Marquette, MI

Video
Video
Latest videos
Full broadcasts
Most viewed
Topics
Headliners
Roundtables
Web only: Take Two
MTP Minute
More David Gregory
MTP History
More NBC shows
Nightly News
Morning Joe

Will goals in Afghanistan be met?

now
Debating control and blame on Gulf oil spill

next
Takeaways from

MEET THE PRESS
Find term in this clip: U.S. House of Representatives, Carly Fiorina, Chuck Todd, oil spill, oil drilling

3 instances

BOEING

The Inauguration
Read Live votes
Topics White House transition Capitol Hill Obama
Collapse Move
Feedback

09/14 / 22:50

Script)
Still Waters (1)
America's Decline (1)
Scripture (1)
Gettysburg (1)
Normandy (1)
Battle of Khe Sanh (1)
Better Life (1)
Common purpose (1)
GDP (1)
Founding Fathers (1)

Show on timeline
☐ NBC analysis
☒ Keyword or phrase
☐ None

MSNBC: 2008 - 2010

Microsoft

MSNBC Topic Pages

More from msnbc.com

Categories

All

U.S.

World

Politics

Business

Entertainment

Topics by number of stories published in the last 24 hours

Latest headline

Trend this week

2 Lindsay Lohan

Lindsay Lohan Hopes To Postpone Rehab After Release

2 Barbra Streisand

Streisand named 2011 MusiCares Person of the Year

2 Microsoft

Microsoft and Salesforce settle patent tiff

1 Emmy Award

'Sons of Anarchy' star bashes Emmys

1 Anna Nicole Smith

Opening statements begin in Anna Nicole Smith trial



allDAY Blog

Kathie Lee & Hoda

TODAYMoms

Concert Series

About Us

Topics / Microsoft

Also see: Television program | Sony | Google

Microsoft and Salesforce settle patent tiff

Updated 2 hours, 49 minutes ago
Microsoft Corp said on Wednesday it reached a patent agreement with rival

Salesforce.com Inc, settling a recent flurry of lawsuits between the two software companies. [Full story](#)



Factbox: List of U.S. billionaires making charity pledge

Updated 5 hours, 2 minutes ago

A group of U.S. billionaires pledged on Wednesday to give away at least half of their wealth to charity as part of a campaign led by investor Warren Buffett and Microsoft founder Bill Gates.

[Full story](#)

Yahoo Japan to switch to Google's search engine

Updated 11:46 a.m. PT, Tues., July 27, 2010

Yahoo Japan, Japan's largest Internet portal operator, will adopt Google's search engine, refusing to follow Yahoo Inc in choosing Microsoft as a partner. [Full story](#)

Sponsored Links

Seattle-Tacoma: Local Mom Makes \$87/Hr Online!

Shocking Money-Making Secret Is Exposed By Local Mom. Read Her Story!

www.icq3news.com/jobs

Got a Computer? Make \$67/Hr From Home!

A Local Mom Shows Us Her Simple Secret To Making \$5,687 a Month Online

<http://today.msnbc.msn.com/id/38232697/>

Microsoft

MSN now: Social-Media-driven News

MSN: 2012

http://now.msn.com/

msnNOW: See what's trend... X

File Edit View Favorites Tools Help

msn NOW News Entertainment Sports Money Living

bing Web Search

Sign out

Make msnNOW my homepage

Like 183k Follow @msnNOW

msn now WHAT'S TRENDING

Wednesday, May 9, 11:10 am

Analyzing 144,186,249 Bing searches and 12,124,446 social updates.

biggest movers (now)

1. [john travolta](#)
2. [gavin smith missi...](#)
3. [brooke shields](#)
4. [melissa gilbert](#)
5. [hillary clinton p...](#)
6. [beastie boys laws...](#)
7. [christina aguilera](#)
8. [terrell owens](#)
9. [george zimmerman](#)
10. [dancing with the ...](#)

What is Biggest Movers?

CHOOSE YOUR VIEW: ALL FAME BALLS CASH SOUL WIRE

10 mins ago

This Texas-sized floating trash heap totally repulses us

24 mins ago

There are tree snakes invading Guam and they're really gross

39 mins ago

Receipt might prove Travolta's a tipper, not a groper

Bringing Entities to Office: Knowledge at Everyone's Fingertips

- Access to fresh information about entities that does not require app switching, and thus, without loss of context
 - Complete or verify facts about entities
 - Access attributes / values for entities
 - Track entities across document collections



Smart Lookup in Office

Word: 2014

Nordic Countries

CHAPTER 2: FINLAND

HELSINKI

Helsinki is the capital and largest city of Finland. It is in the region of Uusimaa, located in southern Finland, on the shore of the Gulf of Finland, an arm of the Baltic Sea. Helsinki has a population of 604,380 (31 December 2012), an urban population of 1,075,465 and a metropolitan population of 1,358,901, making it by far the most populous municipality and urban area in Finland. Helsinki is located some 80 kilometres (50 mi) north of Tallinn, Estonia, 400 kilometres (250 mi) east of Stockholm, Sweden, and 300 kilometres (190 mi) west of Saint Petersburg, Russia. Helsinki has close historical connections with these three cities.

The Helsinki urban area includes the city proper of Helsinki and large parts of other municipalities such as Espoo and Vantaa; it is the world's northernmost urban area of over one million people, and the city is the northernmost capital of an EU member state. Several well-known companies, such as Nokian Footwear, established as Finnish Rubber Works Ltd, were established in Helsinki.

TAMPERE

Tampere, part of the Pirkanmaa region, is a city in southern Finland. It is surrounded by the municipalities of Kangasala, Lempäälä, Nokia, Orivesi, Pirkkala, Ruovesi and Ylöjärvi. It is the most populous inland city in any of the Nordic



the Greater Helsinki municipalities of Helsinki and Espoo. The region had around 470,000 residents, of which 25% of 25 billion euros.

Tampere is located between two lakes, Näsijärvi and Pyhäjärvi, which differ in level by 18 metres (59 ft), the rapids linking them an important power source throughout history, making it. Tampere is dubbed the "Manchester of Finland" and is the center of Finnish industry, and this has given rise to terms such as "Manserock".

Tampere was founded as a market place on the site of the old village of Tammerkoski in 1775 by Gustav III of Sweden and four years later.



Nokia, Finland



en.wikipedia.org

Nokia, Finland is a town and a municipality on the banks of the Nokianvirta River in the region of Pirkanmaa, some 15 kilometres west of Tampere.

AREA 134.27 miles² 347.77 km²

POPULATION 32,353

HISTORY

- 1596 Club War
Nokia was the setting of one of the largest battles against Swedish feudal lords
- 1938 Pohjois-Pirkkala renamed Nokia
- 1977 Nokia was designated a city

RELATED TOPICS



Finland



Tampere



Helsinki

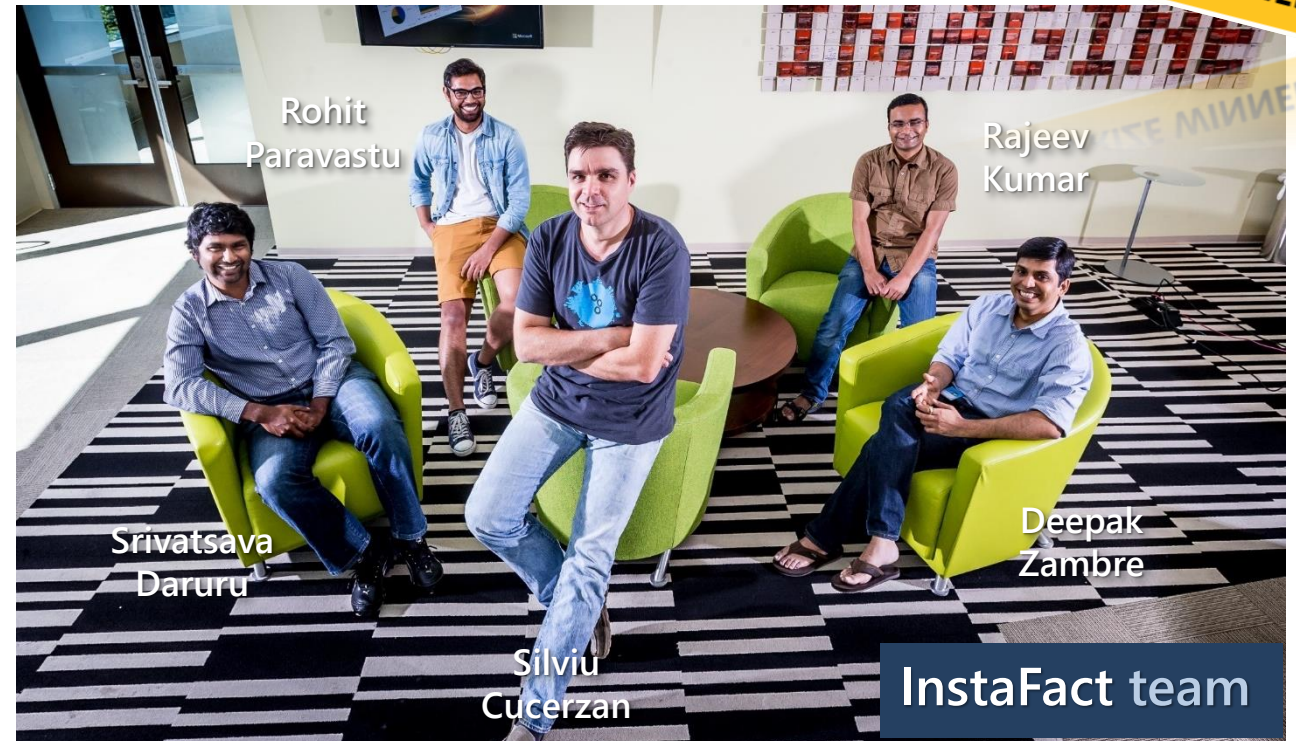
Entities and Attributes in Spreadsheets

Excel: 2018

<div> <div>HomeInsertPage LayoutFormulasDataReviewViewHelpSearch</div> <div> <div> <div>From text/CSV</div> <div>From Web</div> <div>From Table/Range</div> <div>Recent Sources</div> <div>Existing Connections</div> </div> <div> <div>Refresh All</div> <div>Queries & Connections</div> <div>Properties</div> <div>Edit Links</div> </div> <div> <div>Stocks</div> <div>Geography</div> </div> <div> <div>Sort</div> <div>Filter</div> <div>Clear</div> <div>Reapply</div> <div>Advanced</div> </div> <div> <div>Text to Columns</div> <div>Flash Fill</div> <div>Remove Duplicates</div> <div>Data Validation</div> <div>Consolidate</div> <div>Relationships</div> <div>Manage Data Model</div> <div>What-If Analysis</div> <div>Forecast Sheet</div> <div>Group</div> <div>Ungroup</div> <div>Subtotal</div> <div>Show Detail</div> <div>Hide Detail</div> </div> </div> <div>Get & Transform Data</div> <div>Queries & Connections</div> <div>Linked Records</div> <div>Sort & Filter</div> <div>Data Tools</div> <div>Forecast</div> <div>Outline</div> </div>								
Country	Area	Birth rate	Calling code	Capital	Forested area (%)	Abbreviation	National anthem	
Albania	28,748	11.88	355	Tirana	28.2%	AL	Himni i Flamurit	
Austria	83,871	9.80	43	Vienna	46.9%	AT	Land der Berge, Land am Strome	
Bulgaria	110,994	9.20	359	Sofia	35.2%	BG	Mila Rodino	
Hungary	93,028	9.40	36	Budapest	22.9%	HU	Himnusz	
Poland	312,685	9.70	48	Warsaw	30.8%	PL	Poland Is Not Yet Lost	
Romania	238,397	9.30	40	Bucharest	29.8%	RO	Deșteaptă-te, române!	
Slovakia	49,035	10.30	421	Bratislava	40.3%	SK	Nad Tatrou sa blýska	
Slovenia	20,273	10.00	386	Ljubljana	62.0%	SI	Anthem of the Slovene nation	
United States	9833517	12.4	1	Washington, D.C.	33.9%	US	The Star-Spangled Banner	
Instrument	Ticker symbol	P/E	Low	CEO	Price	Volume	Industry	
Apple Inc	AAPL	16.92	\$ 166.44	Mr. Timothy D. Cook	\$ 172.77	37,533,285	Consumer Electronics	
Microsoft Corp	MSFT	70.92	\$ 90.40	Satya Nadella	\$ 93.78	56,387,179	Software - Infrastructure	
Tesla Inc	TSLA	-80.65	\$ 291.36	Mr. Elon Musk	\$ 304.18	8,372,846	Auto Manufacturers	
Facebook Inc	FB	29.59	\$ 149.02	Mr. Mark Zuckerberg	\$ 160.06	126,080,767	Internet Content & Information	

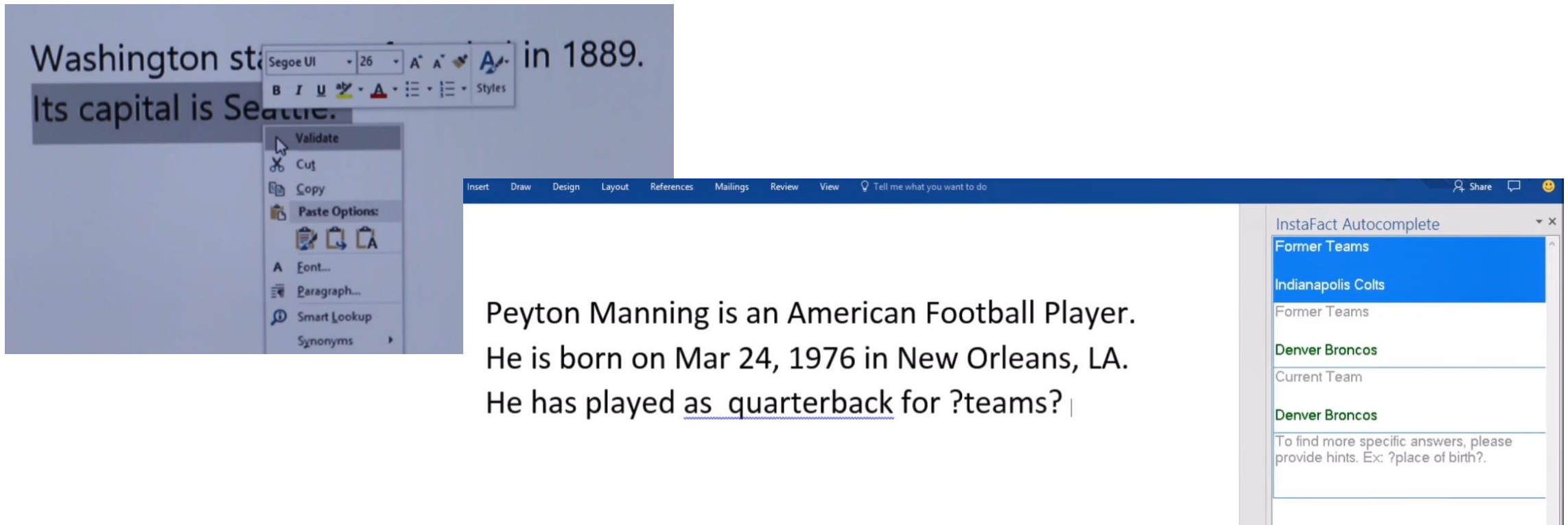
Starting Point: InstaFact – A Hackathon Project

- Office add-ins that contextually autocomplete facts using the Satori KG:
 - in **Word**: complete or verify short factual sentences
 - in **Excel**: link table cells to entities and populate tables with associated values based on:
 - analysis of column headers (InstaFact-retrieve)
 - prediction of relationships between the entities in different columns (InstaFact-complete)



Starting Point: InstaFact – A Hackathon Project

- Office add-ins that contextually autocomplete facts using the Satori KG:
 - in **Word**: complete or verify short factual sentences

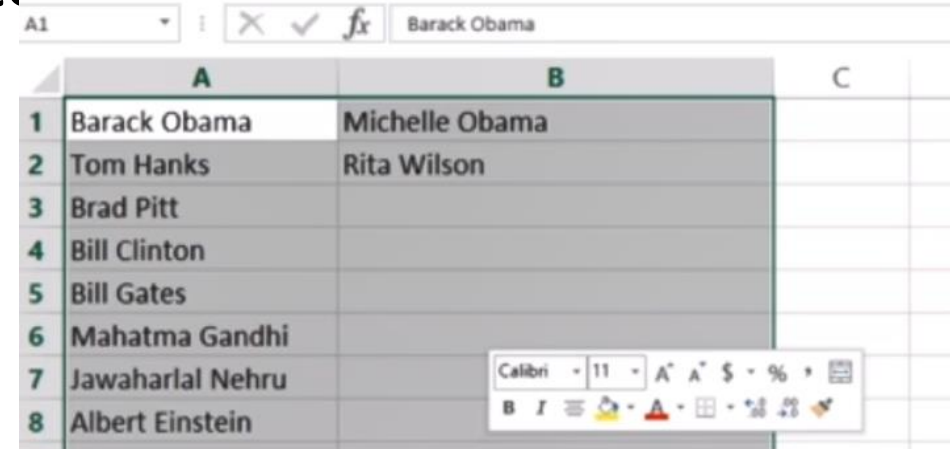


Starting Point: InstaFact – A Hackathon Project

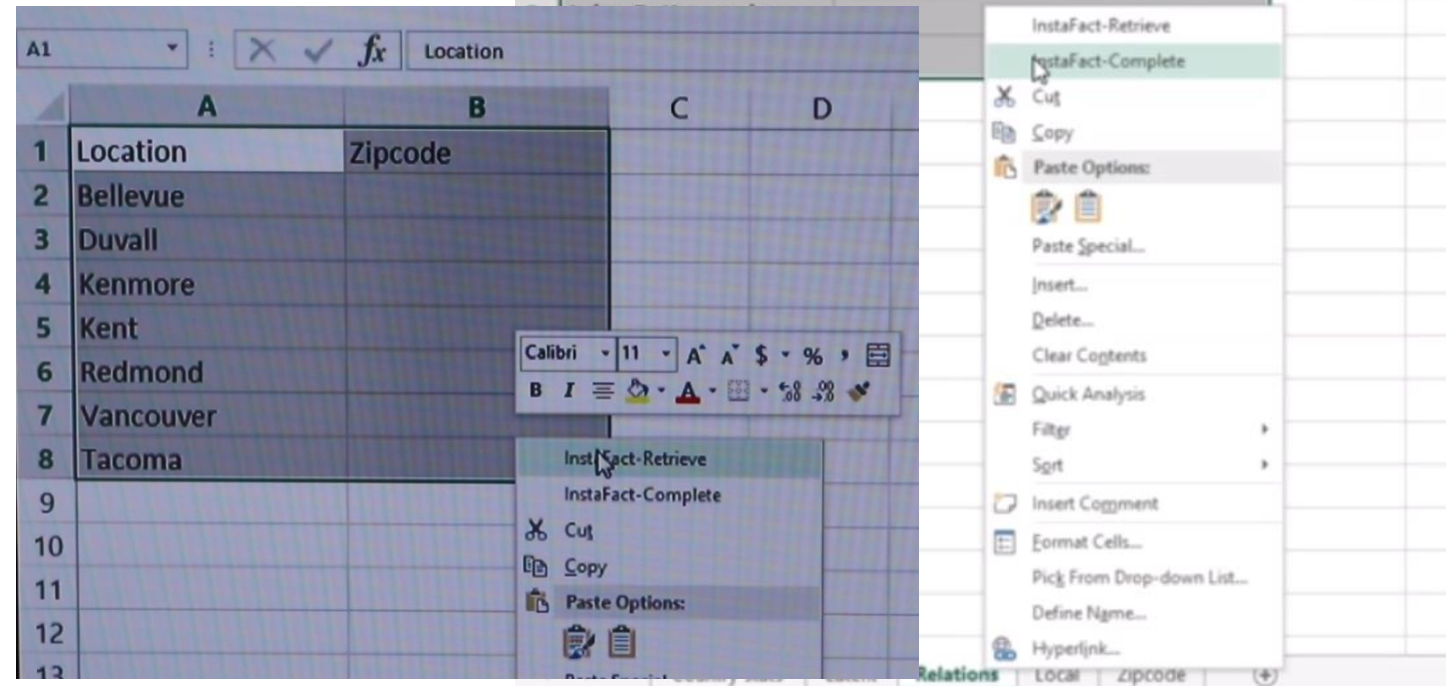
- Office add-ins that contextually autocomplete facts using the Satori KG:

- in **Excel**: link table cells to entities and populate tables with associated values based on:

- analysis of column headers (InstaFact-retrieve)
- prediction of relationships between the entities in different columns (InstaFact-complete)



	A	B	C
1	Barack Obama	Michelle Obama	
2	Tom Hanks	Rita Wilson	
3	Brad Pitt		
4	Bill Clinton		
5	Bill Gates		
6	Mahatma Gandhi		
7	Jawaharlal Nehru		
8	Albert Einstein		



	A	B	C	D
1	Location	Zipcode		
2	Bellevue			
3	Duvall			
4	Kenmore			
5	Kent			
6	Redmond			
7	Vancouver			
8	Tacoma			
9				
10				
11				
12				
13				

Main Challenge: Very Little Context

File Home Insert Page Layout Formulas Data Review View Add-ins LOAD TEST Foxit PDF Team Yellow Tell me what you want to do Share						
Clipboard Font Alignment Number Styles Cells Editing						
B20						
	A	B	C	D	E	F
1	main	founded	stadium	webpage	titles	
2	United	1/1/1995	Robert F. Kennedy Memorial Stadium	http://www.dcunited.com/	MLS Cup '96, MLS Cup '99, MLS Cup '97, MLS Cup 2004	
3	Real	1/1/2004	Rio Tinto Stadium, Rice-Eccles Stadium	http://www.rsl.com/	MLS Cup 2009	
4	Rapids	6/6/1995	Dick's Sporting Goods Park	http://www.coloradorapids.com/	MLS Cup 2010	
5	Sounders	11/13/2007	CenturyLink Field	http://www.soundersfc.com/	MLS Cup 2016	
			stadium	webpage	titles	
			78 Old Trafford	http://www.manutd.com/	2015–16 FA Cup, 1995–96 FA Cup, 2004 FA Cup Final, 1989–90 FA	
			ernabéu Stadium	http://www	Champions League Final, 2016–17 UEFA Champions Lea	
			n ArenaA, Olympic Stadium	http://www	Champions League Final, 1972 European Cup Final, 1973	
			adium Giuseppe Meazza Stad	http://www	Champions League Final, 1964 European Cup Final, 1965	

Microsoft

Google

Tesla

HP

United

Oracle

Dell

Twitter

Facebook

Sun

Amazon

Ajax

United

Real

Inter

United

Alaska

American

Air France

Delta

United

Real

Rapids

Sounders

Ajax

United

Continental

Acid Jazz

JOB

United

Outrage

Testament

Slayer

United

Revive

Newsboys

Beanbag

United

Dear Santa

Dimension

Page 8

National

United

Maori

ACT

Other Challenges for Deploying to Millions of Users

- Full compliance – preserve complete privacy for user data
- Speed – fast answers regardless of input size
- Support invariance properties in tables
- Balance between in-domain consistency for columns/rows in tables and semantic consistency in the graph
- Employ local context aggressively
- Allow incremental disambiguation
- Provide usable confidence measure / in-segment ranking
- Expand coverage beyond deployed domains; out-of-domain detection
- Case-insensitivity for entity mentions in table cells
- Integrate spelling correction with entity context
- Deploy a scalable platform, that can handle future domains of interest



Other Challenges (1)

Handle table format while preserving text analysis capabilities

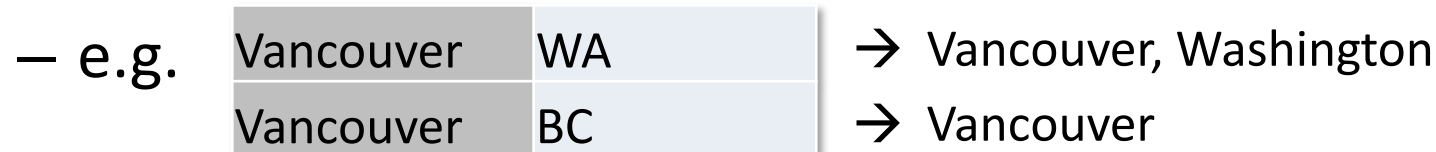
- invariance properties



- segment consistency (at the same time with semantic consistency)



- aggressive employment of local context



Other Challenges (2)

- Incremental disambiguation
 - symmetrical communication
 - consume entity context

e.g.

Portland
Bend

→


Ⓟ Portland, Oregon
Ⓟ Bend, Oregon
Redmond

→

→

Ⓟ Portland, Oregon
Ⓟ Bend, Oregon
Ⓟ Redmond, Oregon

- Usable confidence measure / in-segment ranking

– e.g.	Sammamish	7 disambiguations		p_1	usable confidence
	Seattle	50		p_2	
	Bellevue	85		p_3	
	Washington	490		p_4	

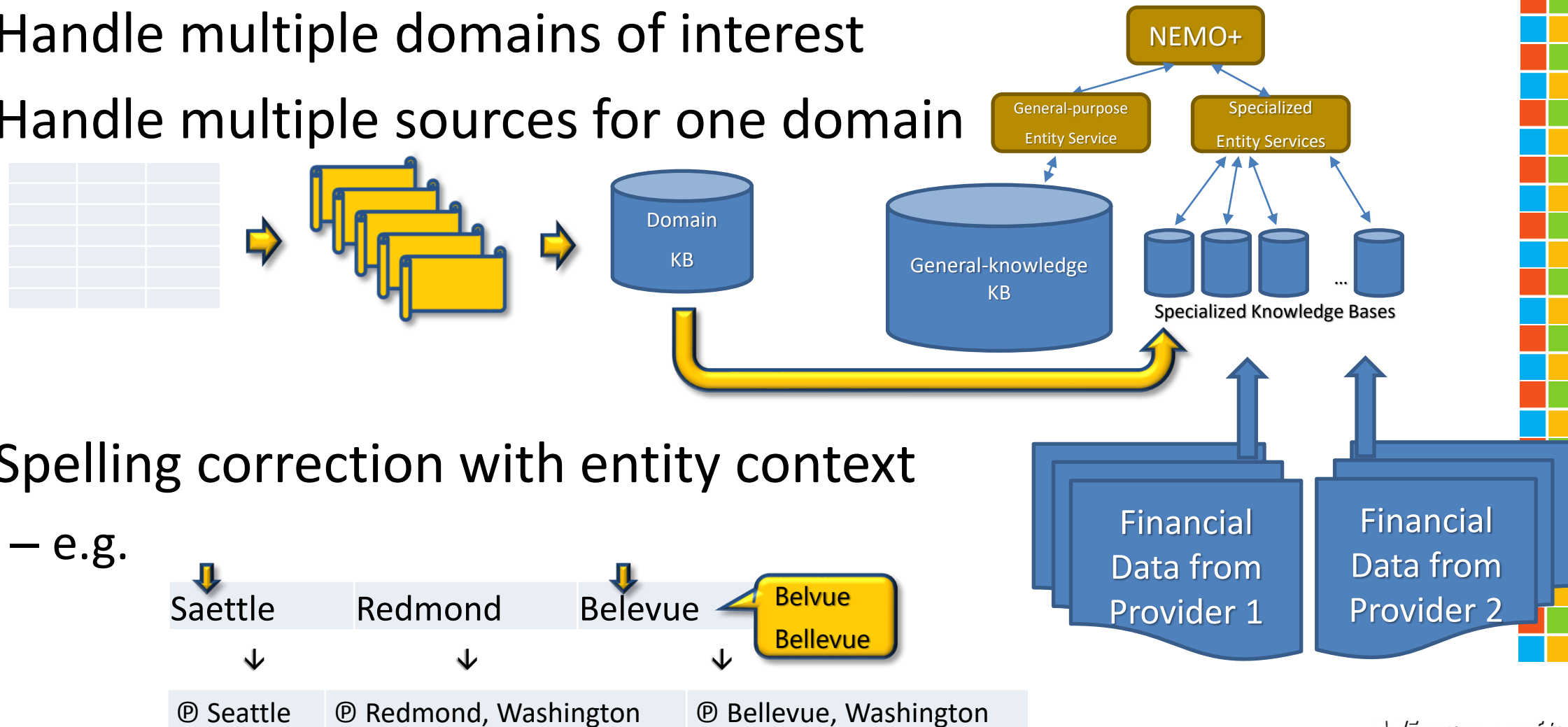
- Case-insensitive handling of entity mentions in tables

– e.g.	united		United	→ United Records
	continental		Continental	→ Continental Records
	acid JAZZ		Acid Jazz	→ Acid Jazz Records
	job		JOB	→ J.O.B. Records

- Speed (fast answers regardless of input type/size)
 - smart splitting of large tables to maintain consistent contexts

Other Challenges (3)

- Handle multiple domains of interest
- Handle multiple sources for one domain



Other Challenges (4)

no control over the user input

- Detect when input data does not match the target segment and provide an intelligent warning

Trump
Obama
Bush
Clinton

User selects to disambiguate
as "Geography"

→ Trump, Colorado
→ Obama, Fukui
→ Clinton, Nebraska
→ Bush, Kentucky

entities of the type
PRESIDENT

Labrador
Brittany
Yorkshire
Shetland
Chihuahua

User selects to disambiguate
as "Geography"

→ Labrador, Queensland
→ Brittany, France
→ Yorkshire, UK
→ Shetland Islands
→ Chihuahua (state)

entities of the type
DOG BREED

Appendix: Academic References

- S. Cucerzan. Large Scale Named Entity Disambiguation Based on Wikipedia Data. [EMNLP-CoNLL 2007](#)
- S. Cucerzan. The MSR System for Entity Linking at TAC. [TAC 2011](#). [TAC 2012](#). [TAC 2013](#). [TAC 2014](#)
- C.W. Leong and S. Cucerzan. Supporting Factual Statements with Evidence from the Web. [CIKM 2012](#)
- S. Cucerzan. Named Entities Made Obvious: Participation in the ERD 2014 Evaluation. [SIGIR-ERD 2014](#)
- S. Cucerzan. Linking to Web Knowledge Bases and Applications to Web Search. [ESAIR 2014 Keynote](#)
- N. Gao and S. Cucerzan. Entity Linking to One Thousand Knowledge Bases. [ECIR 2017](#)
- A. Sil, H. Ji, D. Roth, and S. Cucerzan. Multi-lingual Entity Discovery and Linking. [ACL 2018 Tutorial](#)