

Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources

Xin Luna Dong, Evgeniy Gabrilovich, Kevin Murphy,
Van Dang, Wilko Horn, Camillo Lugaresi,
Shaohua Sun, Wei Zhang

Google Inc.
@VLDB'2015

Motivation for Knowledge-Based Trust (KBT)

.....

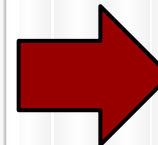
- Providing a new perspective to evaluate Web source quality
- What we have now--*Exogenous signals*
 - Link-based
 - Search log and click-through rate
 - Web spam
- Key idea: Evaluate trustworthiness of sources by *the correctness of its factual information--Endogenous signals*

Correctness of Factual Information

The screenshot shows the Wikipedia article for 'United States'. The page title is 'United States' and it is noted as being redirected from 'USA'. The main text states: 'The **United States of America** (**USA**), commonly referred to as the **United States** (**US**), **America** or simply **the States**, is a federal republic^{[10][11]} consisting of 50 states and a federal district. The 48 contiguous states and the federal district of Washington, D.C., are in central North America between Canada and Mexico. The state of Alaska is the northwestern part of North America and the state of Hawaii is an archipelago in the mid-Pacific. The country also has five populated and nine unpopulated territories in the Pacific and the Caribbean. At 3.79 million square miles (9.83 million km²) in total and with around 317 million people, it is the world's third largest country by area and third most populous. It is a transcontinental country, with the majority of its land area in North America and a small portion in the Pacific Ocean. The United States is a federal republic, with 50 states and a federal district. The 48 contiguous states and the federal district of Washington, D.C., are in central North America between Canada and Mexico. The state of Alaska is the northwestern part of North America and the state of Hawaii is an archipelago in the mid-Pacific. The country also has five populated and nine unpopulated territories in the Pacific and the Caribbean. At 3.79 million square miles (9.83 million km²) in total and with around 317 million people, it is the world's third largest country by area and third most populous. It is a transcontinental country, with the majority of its land area in North America and a small portion in the Pacific Ocean.

The right sidebar contains a summary box for the 'United States of America' with the following information:

- Flag**: Image of the United States flag.
- Great Seal**: Image of the United States Great Seal.
- Motto**: "In God we trust" (official)^{[1][2][3]}, "E pluribus unum" (Latin) (traditional), "Out of many, one"
- Anthem**: "The Star-Spangled Banner" (with a play button and volume icon)
- Map**: A globe showing the United States highlighted in green.

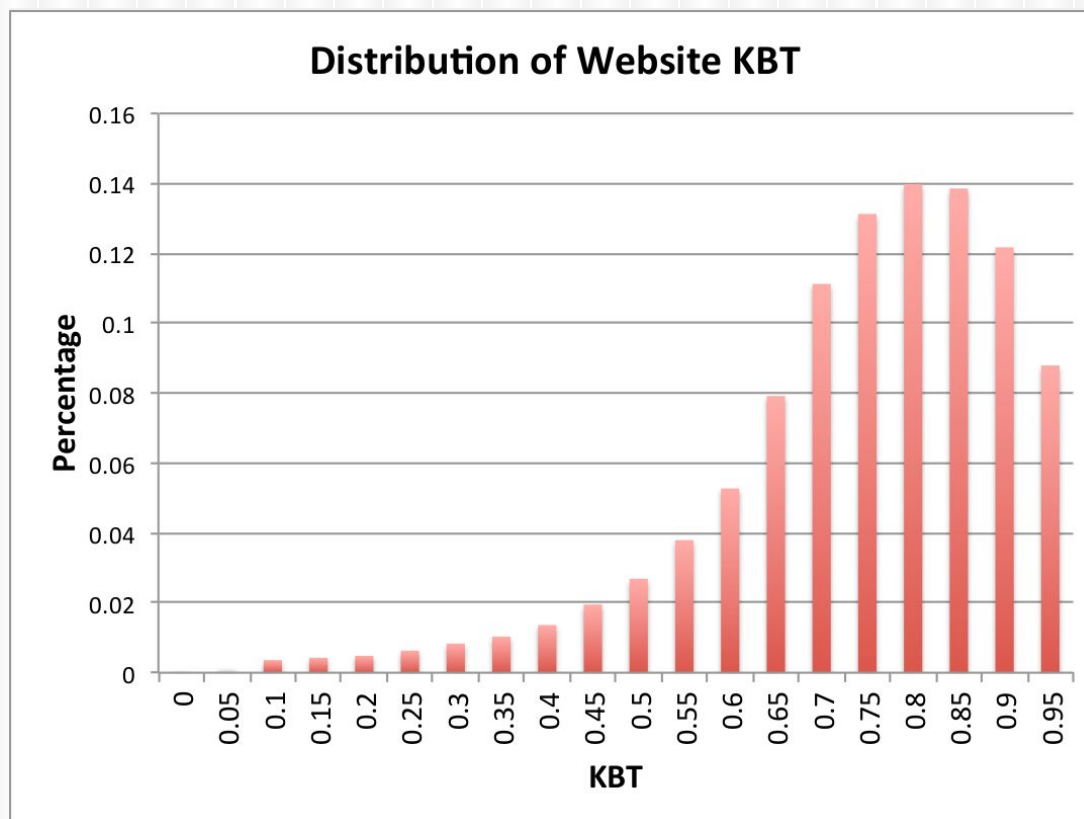


Fact 1	✓
Fact 2	✓
Fact 3	✗
Fact 4	✓
Fact 5	✗
Fact 6	✓
Fact 7	✓
Fact 8	✓
Fact 9	✓
Fact 10	✗
...	...
Accu	0.7

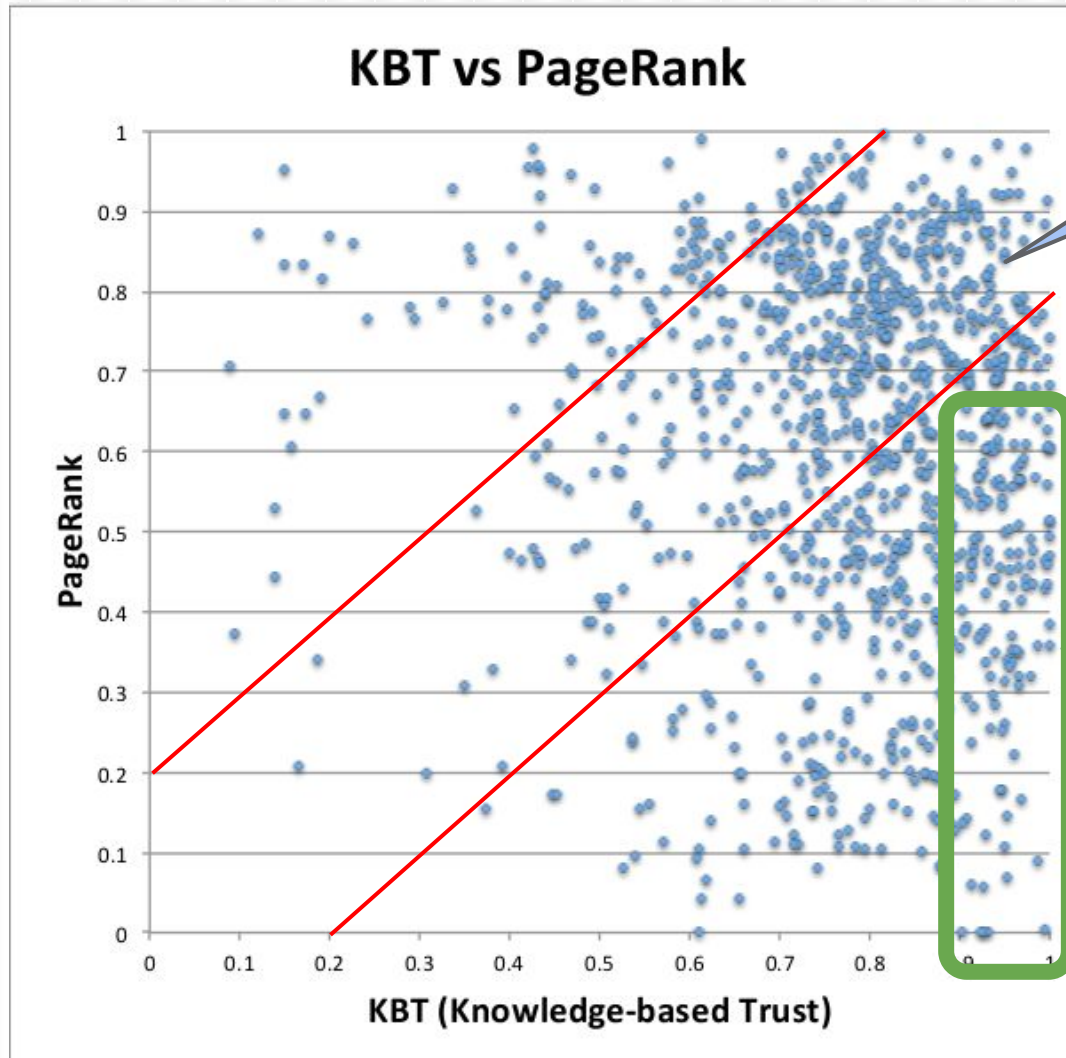
How Can Trustworthiness Help?

Knowledge-Based Trust (KBT)

Trustworthiness in $[0, 1]$ for 5.6M websites and 119M webpages



Knowledge-Based Trust vs. PageRank



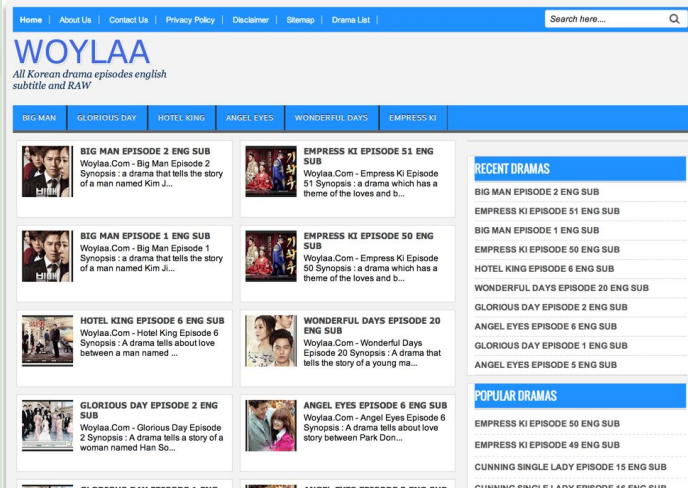
Correlated scores

Often tail sources w. high trustworthiness

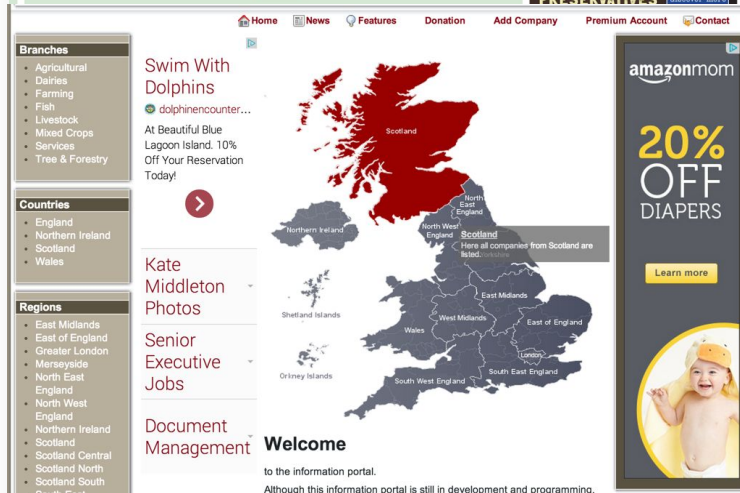
I. Tale Sources w. Low PageRank May Provide Valuable Info



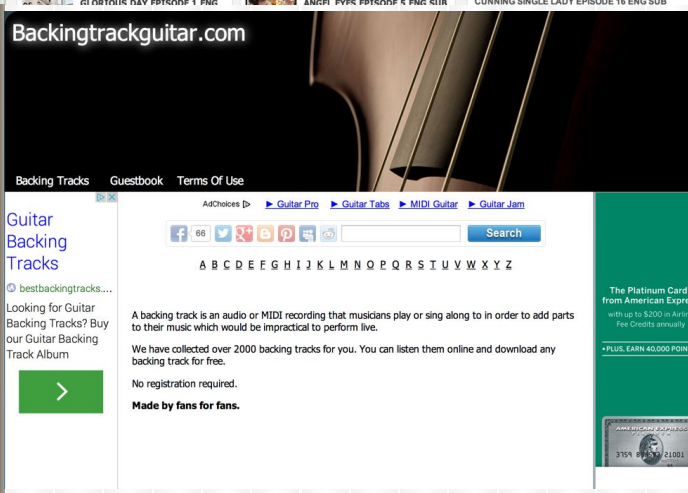
The screenshot shows the salary.com homepage. At the top, there is a search bar for salaries and jobs, with options to browse salaries and follow on social media. Below the search bar is a navigation menu with categories like Salary, Job Search, Education, Career Development, Work & Life, Features, and Business Products. The main content area is divided into several sections: 'I am an Employee or Individual' and 'I am an Employer or Business', both offering services like 'Help Me Negotiate' and 'Job or Salary Reports'. There are also sections for 'Contemplating A Move', 'Job Search', 'Salary Information', 'Job-Specific Competencies', and 'Compensation Subscriptions'. A prominent offer states 'Save 44% on 2 Job Postings' with a 'Save now' button. At the bottom, there is a survey question: 'Over the past 12 months, how frequently have you been bullied at work?' with radio button options for 'Never', 'Once or twice over the past 12 months', 'Once or twice a month', 'Once or twice a week', 'Once or twice every few months', and 'Almost every day at work'. A 'Submit' button is at the bottom of the survey.



The screenshot shows the WOYLAA website, which is a platform for Korean drama episodes. The header includes navigation links like Home, About Us, Contact Us, Privacy Policy, Disclaimer, Sitemap, and Drama List, along with a search bar. The main content area is titled 'WOYLAA All Korean drama episodes english subtitle and RAW'. Below this, there are several episode listings with thumbnails and titles, such as 'BIG MAN EPISODE 2 ENG SUB', 'EMPRESS KI EPISODE 51 ENG SUB', 'BIG MAN EPISODE 1 ENG SUB', 'EMPRESS KI EPISODE 50 ENG SUB', 'HOTEL KING EPISODE 6 ENG SUB', 'WONDERFUL DAYS EPISODE 20 ENG SUB', 'GLORIOUS DAY EPISODE 2 ENG SUB', 'WONDERFUL DAYS EPISODE 20 ENG SUB', 'GLORIOUS DAY EPISODE 1 ENG SUB', 'ANGEL EYES EPISODE 6 ENG SUB', 'WoylAA.Com - Angel Eyes Episode 6 Synopsis: A drama tells about love story between Park Don...', 'GLORIOUS DAY EPISODE 1 ENG SUB', and 'ANGEL EYES EPISODE 5 ENG SUB'. On the right side, there are sections for 'RECENT DRAMAS' and 'POPULAR DRAMAS', each listing several drama titles.



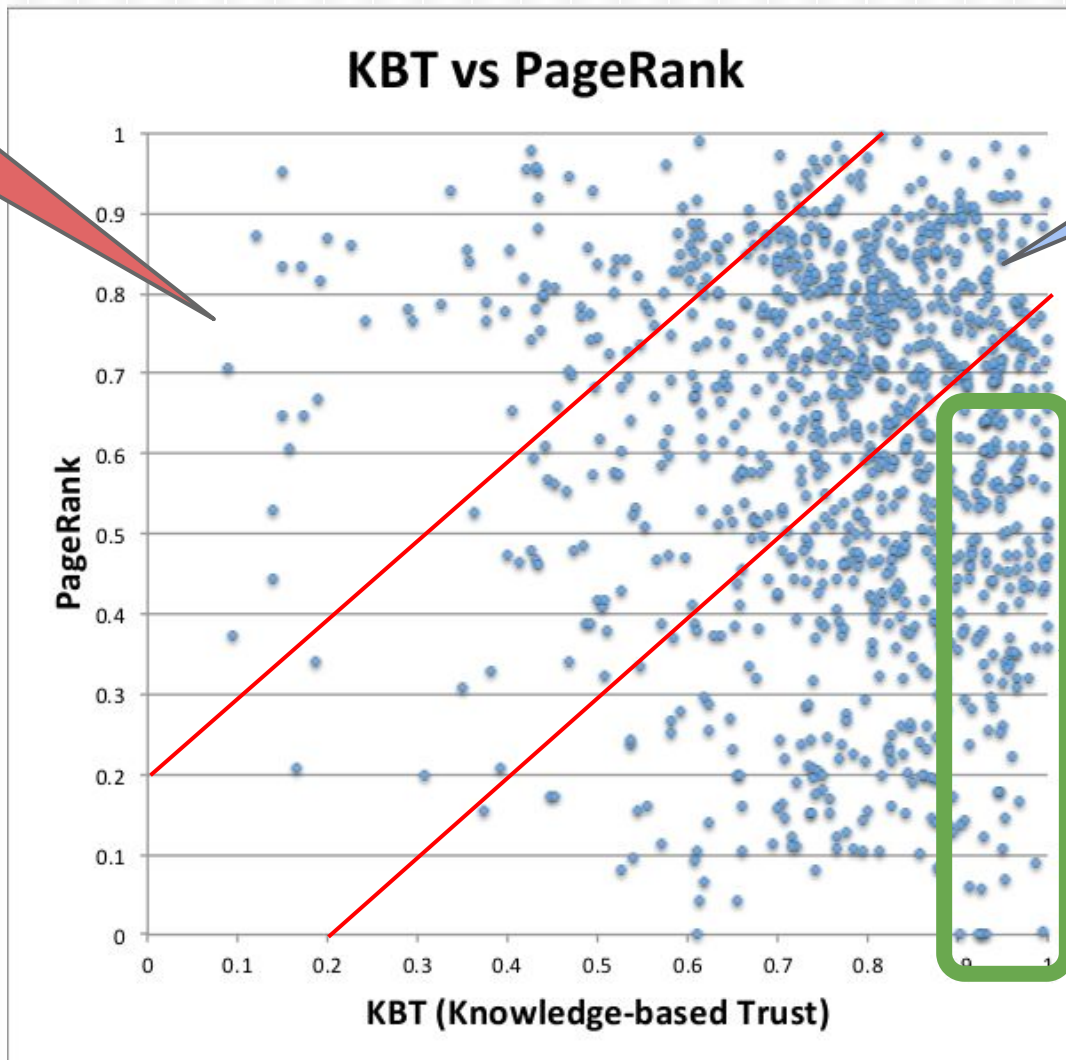
The screenshot shows a website with a map of the United Kingdom. The map is color-coded by region, with Scotland in red and the rest of the UK in blue. The website has a navigation menu with links for Home, News, Features, Donation, Add Company, Premium Account, and Contact. On the left side, there is a sidebar with 'Branches' (Agricultural, Dairies, Farming, Fish, Livestock, Mixed Crops, Services, Tree & Forestry), 'Countries' (England, Northern Ireland, Scotland, Wales), and 'Regions' (East Midlands, East of England, Greater London, Merseyside, North East England, North West England, Northern Ireland, Scotland, Scotland Central, Scotland North, Scotland South, South East). The main content area features a 'Swim With Dolphins' advertisement for 'dolphincounter...' at 'Beautiful Blue Lagoon Island, 10% Off Your Reservation Today!'. Below this is a 'Kate Middleton Photos' section with a 'Senior Executive Jobs' link and a 'Document Management' link. A 'Welcome' message is at the bottom, stating 'to the information portal. Although this information portal is still in development and programming.' An Amazon Mom advertisement for '20% OFF DIAPERS' is on the right side.



The screenshot shows the Backingtrackguitar.com website. The header includes navigation links for Backing Tracks, Guestbook, and Terms Of Use. The main content area is titled 'Backingtrackguitar.com' and features a search bar with a 'Search' button. Below the search bar is a list of categories: 'Guitar Pro', 'Guitar Tabs', 'MIDI Guitar', and 'Guitar Jam'. A 'Search' button is also present. The main content area is titled 'Guitar Backing Tracks' and includes a 'bestbackingtracks...' link. Below this is a 'Looking for Guitar Backing Tracks? Buy our Guitar Backing Track Album' section with a green arrow button. The text below the button states: 'A backing track is an audio or MIDI recording that musicians play or sing along to in order to add parts to their music which would be impractical to perform live. We have collected over 2000 backing tracks for you. You can listen them online and download any backing track for free. No registration required. Made by fans for fans.' On the right side, there is a 'The Platinum Card® from American Express' advertisement with a 'PLUS, EARN 40,000 POINTS*' offer and an image of the card.

Among 100 sampled websites, 85 are indeed trustworthy.

Knowledge-Based Trust vs. PageRank



Often sources w. low accuracy

Correlated scores

Often tail sources w. high trustworthiness

II. Popular Websites May Not Be Trustworthy

Gossip Websites

<http://www.ebizmba.com/articles/gossip-websites>

Domain
www.eonline.com
perezhilton.com
radaronline.com
www.zimbio.com
mediatakeout.com
gawker.com
www.popsugar.com
www.people.com
www.tmz.com
www.fishwrapper.com
celebrity.yahoo.com
wonderwall.msn.com
hollywoodlife.com
www.wetpaint.com

14 out of 15 have a PageRank among top 15% of the websites

All have knowledge-based trust in bottom 50%

II. Popular Websites May Not Be Trustworthy

YAHOO!
ANSWERS

Entertainment & Music > Celebrities



Why are British women so unattractive?

Seriously, what's with that? There are very few English chicks I would say are attractive yet so many countries who are drop-dead gorgeous.

Update: Catherine Zeta-Jones is from NEW ZEALAND!!!!!! Dummy!

Update 2: The SPICE GIRLS! Surely, you jest. Put them all together, all their best points, and then...

Update 3: crazy_lad wins the "moron" award for this question. He says all Americans are fat and being narrowminded! LMFAO at that IDIOT! LOL

☆ Follow 37 answers

[Are you getting your biggest tax refund?](#)

Get your taxes done right, and your biggest refund, guaranteed. Start for free today!

TurboTax Sponsored

[California Programs Contribute to STEM Careers](#)

California's public school system contributes to STEM careers by offering science-centric activities.

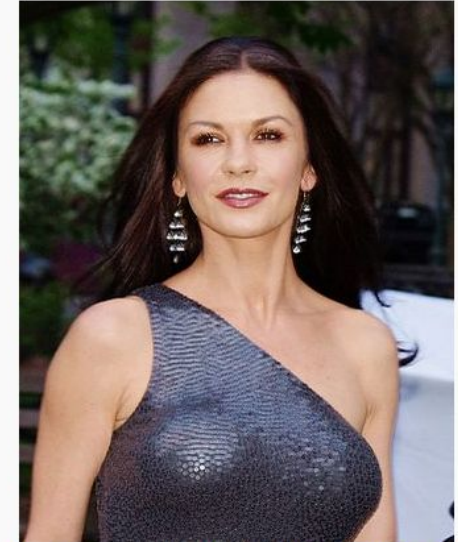
CBS Local Sponsored



WIKIPEDIA
The Free Encyclopedia

Catherine Zeta-Jones

CBE




Zeta-Jones at the 2012 Tribeca Film Festival

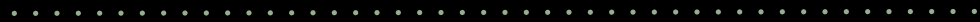
Born	Catherine Zeta Jones 25 September 1969 (age 45) Swansea, Glamorgan, Wales
Nationality	Welsh
Citizenship	Britain
Alma mater	Arts Educational Schools, London
Occupation	Actress
Years active	1981–present
Spouse(s)	Michael Douglas (m. 2000)
Children	2

III. Website Recommendation by Vertical

Sula

Site	Accuracy	# Triples	Score 
en.wikipedia.org	0.76	1,184	5.36
www.cheesewiki.com	0.84	230	4.55
www.ranker.com	0.93	105	4.34
scratchpad.wikia.com	0.92	101	4.26
www.marcellathecheesemonger.com	0.90	109	4.25
www.cheeseplatesf.com	0.90	114	4.25
cheesetique.com	0.93	83	4.12
www.lafromagerie.co.uk	0.93	62	3.87
adrian1974fulga.wordpress.com	0.90	71	3.83
epicurefoodscorp.com	0.78	131	3.81
www.gourmetfoodstore.com	0.83	91	3.77
www.sheridanscheesemongers.com	0.85	71	3.65
www.sfgate.com	0.89	54	3.56
cheeseandchampagne.com	0.90	48	3.50
www.sciencedirect.com	0.69	161	3.50
about-france.com	0.98	34	3.48
www.cookipedia.co.uk	0.70	132	3.44
www.cheese.com	0.86	50	3.39
www.pennmac.com	0.76	87	3.38

III. Website Recommendation by Vertical




Sula

Site	Accuracy	# Triples	Score
ru.wikipedia.org	0.85	3,597	6.98
ja.wikipedia.org	0.73	3,032	5.88
en.wikipedia.org	0.53	36,489	5.57
www.opensourcesoftwaredirectory.com	0.91	323	5.26
wpedia.goo.ne.jp	0.63	2,846	5.03
www.ranker.com	0.81	308	4.66
packages.gentoo.org	0.65	942	4.45
uk.wikipedia.org	0.73	420	4.43
freecode.com	0.57	2,085	4.36
www.file.net	0.65	756	4.34
gpo.zugaina.org	0.65	711	4.28
gentoobrowse.randomdan.homeip.net	0.80	208	4.27
whatis.techtarget.com	0.56	1,417	4.08
www.starringthecomputer.com	0.87	101	4.02
www.linuxlinks.com	0.67	411	4.02
companies.findthebest.com	0.81	135	3.95
www.fileinfo.com	0.53	1,770	3.95
file.downloadatoz.com	0.62	571	3.92
www.zwodnik.com	0.55	1,187	3.92
www.system-tray-cleaner.com	0.70	273	3.92
bitnami.com	0.70	266	3.91

.....
Now, How to Compute KBT?

Key Idea in KBT



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikimedia Shop

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- Print/export

Languages

- Acèh
- Адыгэбзэ
- Afrikaans
- Alemannisch
- አማርኛ
- Ænglisc
- Англис
- العربية
- Aragonés
- ᱫᱷᱟᱱᱵᱟᱫᱽ

Create account Log in

Article [Talk](#) [Read](#) [View source](#)

United States

From Wikipedia, the free encyclopedia
(Redirected from USA)

For other uses, see [US \(disambiguation\)](#), [USA \(disambiguation\)](#), and [United States \(disambiguation\)](#).

The **United States of America** (**USA**), commonly referred to as the **United States** (**US**), **America** or simply **the States**, is a federal republic^{[10][11]} consisting of 50 [states](#) and a federal district. The 48 contiguous states and the federal district of [Washington, D.C.](#), are in central [North America](#) between [Canada](#) and [Mexico](#). The state of [Alaska](#) is the northwestern part of North America and the state of [Hawaii](#) is an [archipelago](#) in the mid-Pacific. The country also has five populated and nine unpopulated [territories](#) in the Pacific and the [Caribbean](#). At 3.79 million square miles (9.83 million km²) in total and with around 317

United States of America



Flag

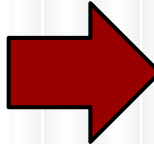


Great Seal

Motto:
"In God we trust" (official)^{[1][2][3]}
"E pluribus unum" (Latin) (traditional)
"Out of many, one"

Anthem: "The Star-Spangled Banner"
 0:00 MENU

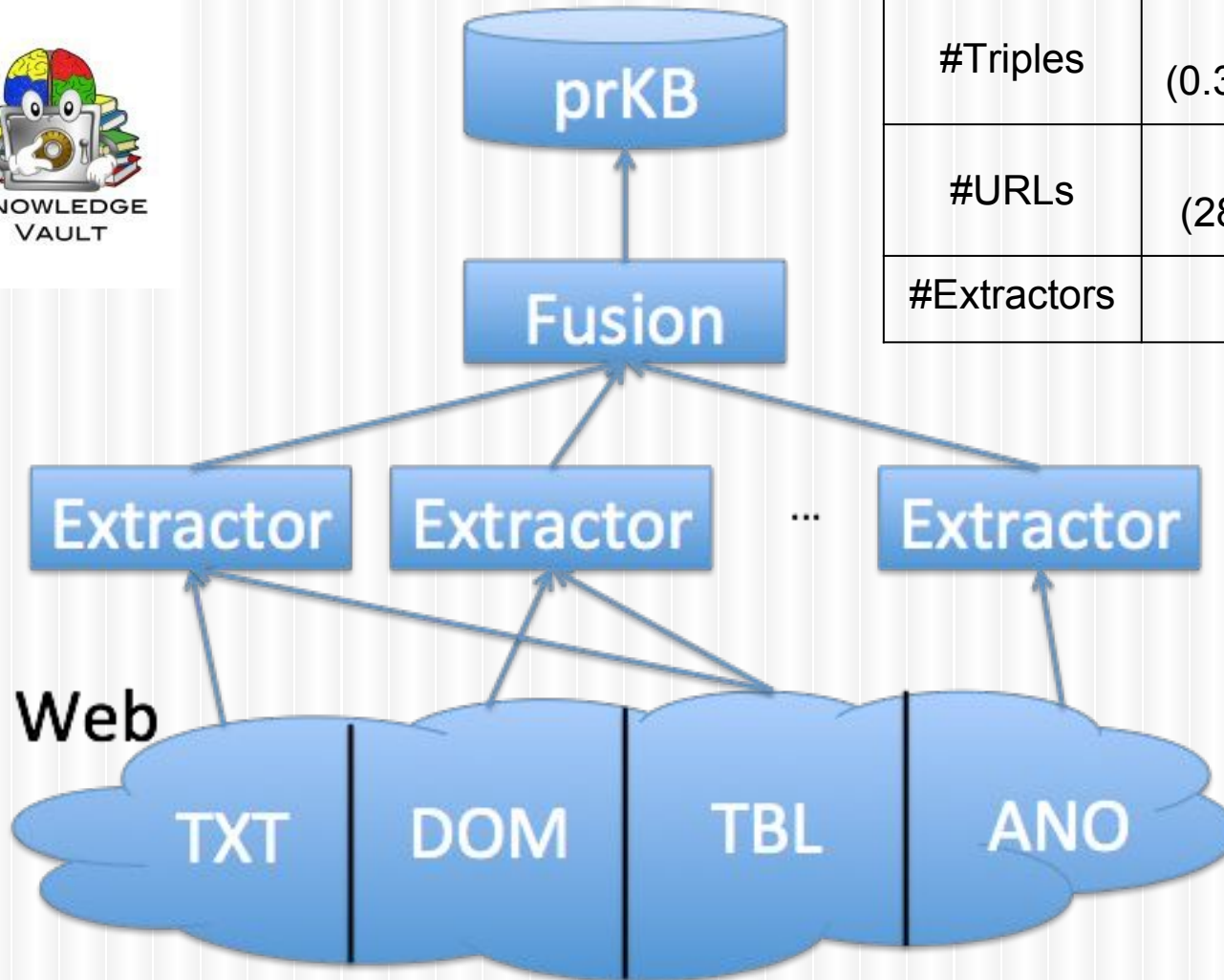
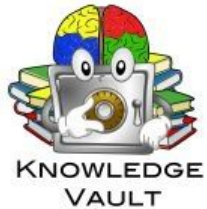




Fact 1	✓
Fact 2	✓
Fact 3	✗
Fact 4	✓
Fact 5	✗
Fact 6	✓
Fact 7	✓
Fact 8	✓
Fact 9	✓
Fact 10	✗
...	...
Accu	0.7


Knowledge Vault– Probabilistic Knowledge Fusion.....

[SIGKDD, 2014]
[VLDB, 2014]



#Triples	3.0B (0.3B w. $pr \geq 0.7$)
#URLs	2.5B (28M Websites)
#Extractors	16

KV Makes This Possible



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikimedia Shop

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- Print/export

Languages

- Acèh
- Адыгэбзэ
- Afrikaans
- Alemannisch
- አማርኛ
- Ænglisc
- Англис
- العربية
- Aragonés
- ᱫᱷᱟᱱᱵᱟᱫᱽ

Create account Log in

Article Talk Read View source ▼

United States

From Wikipedia, the free encyclopedia
(Redirected from USA)

For other uses, see [US \(disambiguation\)](#), [USA \(disambiguation\)](#), and [United States \(disambiguation\)](#).

The **United States of America** (**USA**), commonly referred to as the **United States** (**US**), **America** or simply **the States**, is a federal republic^{[10][11]} consisting of 50 [states](#) and a [federal district](#). The 48 [contiguous states](#) and the federal district of [Washington, D.C.](#), are in central [North America](#) between [Canada](#) and [Mexico](#). The state of [Alaska](#) is the northwestern part of North America and the state of [Hawaii](#) is an [archipelago](#) in the mid-[Pacific](#). The country also has five populated and nine unpopulated [territories](#) in the Pacific and the [Caribbean](#). At 3.79 million square miles (9.83 million km²) in total and with around 317

United States of America



Flag

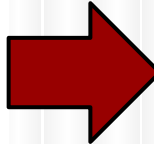


Great Seal

Motto:
"In God we trust" (official)^{[1][2][3]}
"E pluribus unum" (Latin) (traditional)
"Out of many, one"


Anthem: "The Star-Spangled Banner"



Fact 1	✓
Fact 2	✓
Fact 3	✗
Fact 4	✓
Fact 5	✗
Fact 6	✓
Fact 7	✓
Fact 8	✓
Fact 9	✓
Fact 10	✗
...	...
Accu	0.7

KV Makes This Possible



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikimedia Shop
- Interaction
 - Help
 - About Wikipedia
 - Community portal
 - Recent changes
 - Contact page
- Tools
- Print/export
- Languages
 - Acèh
 - Адыгэбзэ
 - Afrikaans
 - Alemannisch
 - አማርኛ
 - Ænglisc
 - Англис
 - العربية
 - Aragonés
 - ᱫᱷᱟᱱᱵᱟᱫ

Create account Log in

Article Talk Read View source ▼

United States

From Wikipedia, the free encyclopedia
(Redirected from USA)

For other uses, see [US \(disambiguation\)](#), [USA \(disambiguation\)](#), and [United States \(disambiguation\)](#).

The **United States of America** (**USA**), commonly referred to as the **United States** (**US**), **America** or simply **the States**, is a federal republic^{[10][11]} consisting of 50 [states](#) and a [federal district](#). The 48 [contiguous states](#) and the federal district of [Washington, D.C.](#), are in central [North America](#) between [Canada](#) and [Mexico](#). The state of [Alaska](#) is the northwestern part of North America and the state of [Hawaii](#) is an [archipelago](#) in the mid-[Pacific](#). The country also has five populated and nine unpopulated [territories](#) in the Pacific and the [Caribbean](#). At 3.79 million square miles (9.83 million km²) in total and with around 317

United States of America



Flag

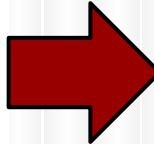


Great Seal

Motto:
 "In God we trust" (official)^{[1][2][3]}
 "E pluribus unum" (Latin) (traditional)
 "Out of many, one"

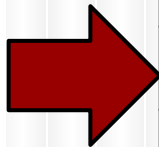
Anthem: "The Star-Spangled Banner"



Triple 1	1.0
Triple 2	0.9
Triple 3	0.3
Triple 4	0.8
Triple 5	0.4
Triple 6	0.8
Triple 7	0.9
Triple 8	1.0
Triple 9	0.7
Triple 10	0.2
...	...
Accu	0.7

Challenges



Triple 1	1.0
Triple 2	0.9
Triple 3	0.3
Triple 4	0.8
Triple 5	0.4
Triple 6	0.8
Triple 7	0.9
Triple 8	1.0
Triple 9	0.7
Triple 10	0.2
...	...
Accu	0.7

How to decide if a triple is indeed claimed by the source instead of an extraction error?

Extractions Can Be Wrong

- (Obama, nationality, Kenya)

2087 extractions:

- Example of a correct extraction

<http://beforeitsnews.com/obama-birthplace-controversy/2013/04/alabama-supreme-court-chief-justice-roy-moore-to-preside-over-obama-eligibility-case-2458624.html>

2006: Obama In Kenya: I Am So Proud To Come Back Home - [VIDEO HERE](#).

2007: Michelle Obama Declares Obama Is Kenyan And America Is Mean - [VIDEO HERE](#).

2008: Michelle Obama Declares Barack Obama's Home Country Is Kenya - [VIDEO HERE](#).

FLASHBACK: Obama Is The Original Birther! Obama In 1991 Stated In His Own Bio He Was Born In Kenya. [DETAILS HERE](#).

- Example of a wrong extraction

<http://www.monitor.co.ug/News/National/US+will+respect+winner+of+Kenya+election++Obama+says/-/688334/1685814/-/ksxagx/-/index.html>

US will respect winner of
Kenya election, Obama says

SHARE BOOKMARK PRINT RATING☆☆☆☆☆

Extractions Can Be Wrong

- (Obama, nationality, USA)

2481 extractions:

- Example of a correct extraction

<http://www.dogonews.com/2009/10/9/a-nobel-prize-for-our-awesome-president>

- Example of a wrong extraction

<http://blogs.telegraph.co.uk/news/timstanley/100169248/barack-obamas-life-story-contains-myth-not-truth-says-biographer-so-why-did-the-media-report-it-as-truth/>

Tim Stanley

Dr Tim Stanley is a historian of the United States. His new book about Hollywood politics is out in May. His personal website is www.timothystanley.co.uk and you can follow him on Twitter @timothy_stanley.

 Follow 15.6K followers

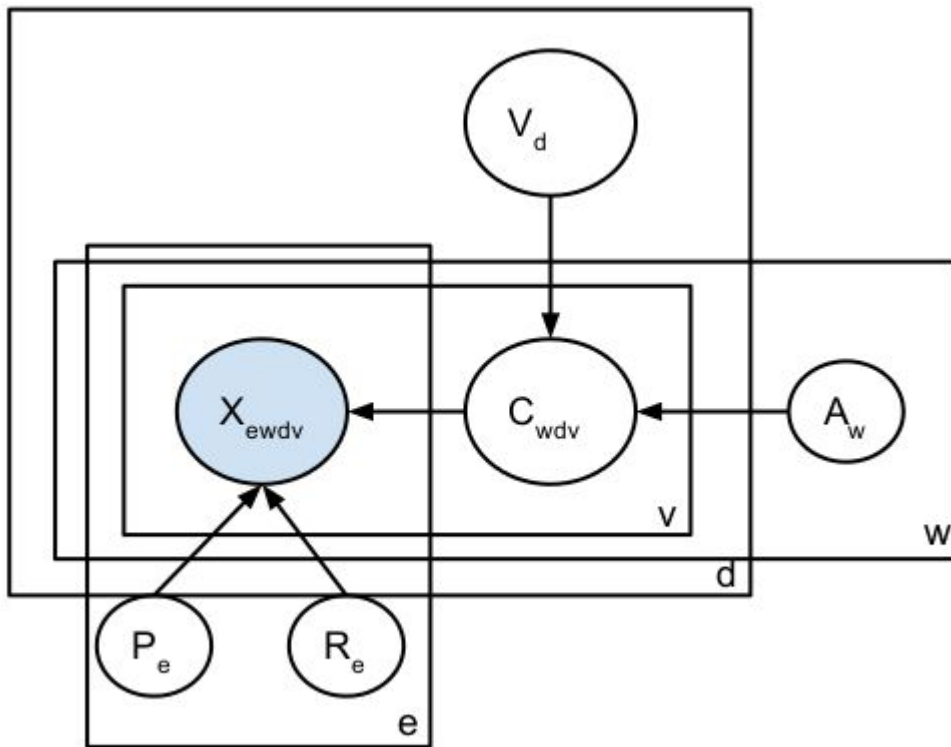


Barack Obama's life story contains 'myth, not truth', says biographer. So why did the media report it as truth?

KBT Strategies

1. Graphical model--predict at the same time
 - a. extraction correctness
 - b. triple correctness
 - c. source accuracy
 - d. extractor precision/recall
2. Un(Semi-)supervised learning (Bayesian)
 - a. leverage source/extractor agreements
 - b. trust a source/extractor w. high quality
3. Source/extractor hierarchy
 - a. Break down “large” sources
 - b. Group “small” sources

Graphical Model



Observations

- X_{ewdv} : whether extractor e extracts from source w the (d,v) item-value pair

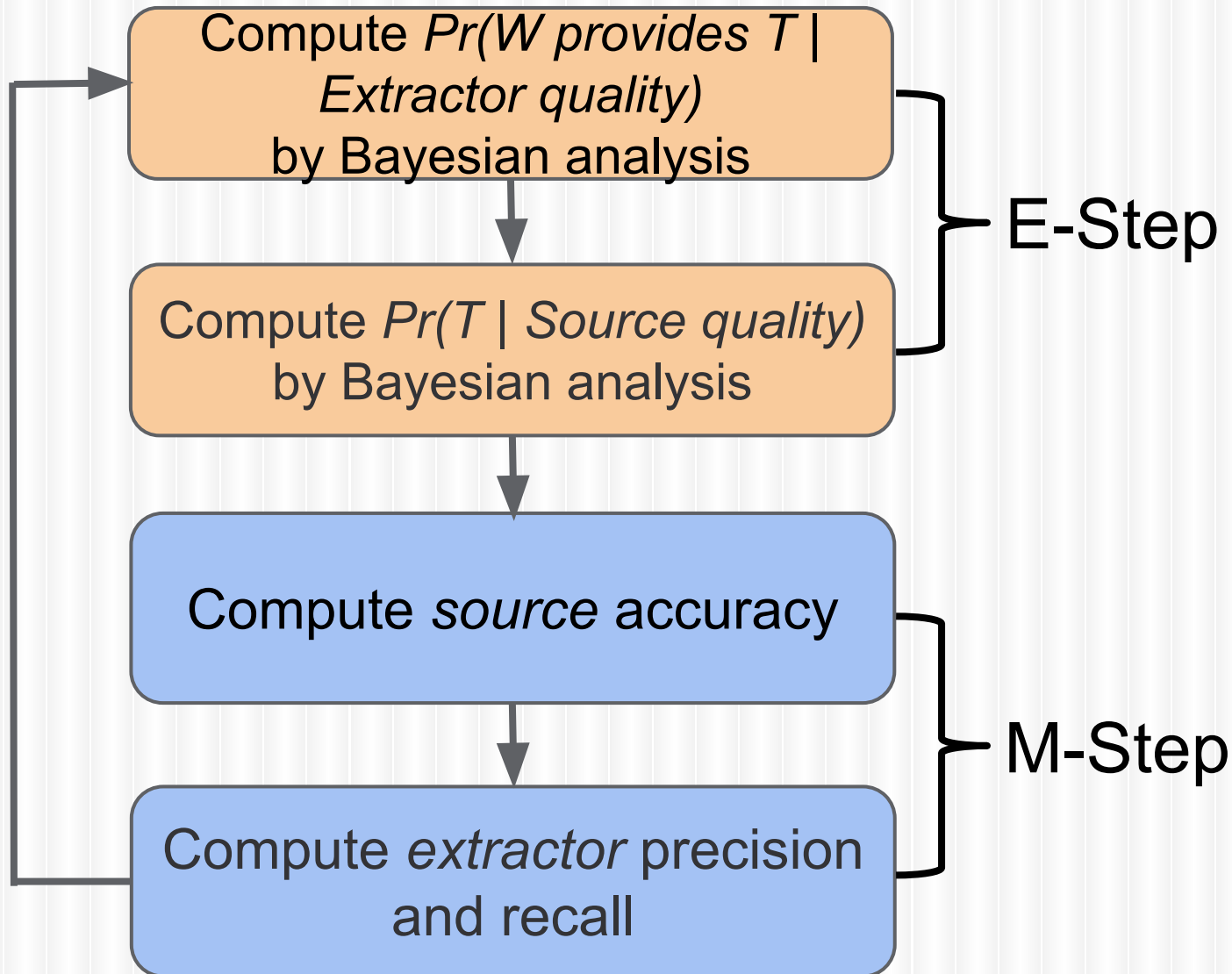
Latent variables

- C_{wdv} : whether source w indeed provides (d,v) pair
- V_d : the correct value(s) for d

Parameters

- A_w : **Trust** of source w
- P_e : Precision of extractor e
- R_e : Recall of extractor e

Algorithm



Web Source Trustworthiness



- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikimedia Shop

- Interaction
 - Help
 - About Wikipedia
 - Community portal
 - Recent changes
 - Contact page

- Tools
- Print/export

- Languages
 - Acèh
 - Адыгэбзэ
 - Afrikaans
 - Alemannisch
 - አማርኛ
 - Ænglisc
 - АӀсӀӀӀӀӀӀӀ
 - العربية
 - Aragonés
 - ᱯᱟᱨᱥᱤ

[Create account](#) [Log in](#)
[Article](#) [Talk](#) [Read](#) [View source](#)

United States

From Wikipedia, the free encyclopedia
 (Redirected from USA)

For other uses, see [US \(disambiguation\)](#), [USA \(disambiguation\)](#), and [United States \(disambiguation\)](#).

The **United States of America** (**USA**), commonly referred to as the **United States** (**US**), **America** or simply the **States**, is a federal republic^{[10][11]} consisting of 50 *states* and a federal district. The 48 contiguous states and the federal district of Washington, D.C., are in central North America between Canada and Mexico. The state of Alaska is the northwestern part of North America and the state of Hawaii is an archipelago in the mid-Pacific. The country also has five populated and nine unpopulated territories in the Pacific and the Caribbean. At 3.79 million square miles (9.83 million km²) in total and with around 317

United States of America




Flag Great Seal

Motto:
 "In God we trust" (official)^{[1][2][3]}
 "E pluribus unum" (Latin) (traditional)
 "Out of many, one"

Anthem: "The Star-Spangled Banner"
 0:00





	Triple Corr	Extraction Corr
Triple 1	1.0	1.0
Triple 2	0.9	1.0
Triple 3	0.3	1.0
Triple 4	0.8	1.0
Triple 5	0.4	0.9
Triple 6	0.8	0.9
Triple 7	0.9	0.8
Triple 8	1.0	0.2
Triple 9	0.7	0.1
Triple 10	0.2	0.1
...
Accu	0.73	

Predicting Extraction and Triple Correctness

- (Obama, nationality, Kenya)

2087 extractions:

- Example of a correct extraction ($Pr_extCorr=0.792$)

<http://beforeitsnews.com/obama-birthplace-controversy/2013/04/alabama-supreme-court-chief-justice-roy-moore-to-preside-over-obama-eligibility-case-2458624.html>

2006: Obama In Kenya: I Am So Proud To Come Back Home - [VIDEO HERE](#).

2007: Michelle Obama Declares Obama Is Kenyan And America Is Mean - [VIDEO HERE](#).

2008: Michelle Obama Declares Barack Obama's Home Country Is Kenya - [VIDEO HERE](#).

FLASHBACK: Obama Is The Original Birther! Obama In 1991 Stated In His Own Bio He Was Born In Kenya. [DETAILS HERE](#).

- Example of a wrong extraction ($Pr_extCorr=0.130$)

<http://www.monitor.co.ug/News/National/US+will+respect+winner+of+Kenya+election++Obama+says/-/688334/1685814/-/ksxagx/-/index.html>

US will respect winner of
Kenya election, Obama says

[SHARE](#) [BOOKMARK](#) [PRINT](#) [RATING](#) ★★★★★

- $Pr_tripleCorr=0$ (not enough support)

Predicting Extraction and Triple Correctness

- (Obama, nationality, USA)

2481 extractions:

- Example of a correct extraction ($Pr_extCorr=0.999$)

<http://www.dogonews.com/2009/10/9/a-nobel-prize-for-our-awesome-president>

- Example of a wrong extraction ($Pr_extCorr=0.261$)

<http://blogs.telegraph.co.uk/news/timstanley/100169248/barack-obamas-life-story-contains-myth-not-truth-says-biographer-so-why-did-the-media-report-it-as-truth/>

Tim Stanley

Dr Tim Stanley is a historian of the United States. His new book about Hollywood politics is out in May. His personal website is www.timothystanley.co.uk and you can follow him on Twitter @timothy_stanley.

 Follow 15.6K followers

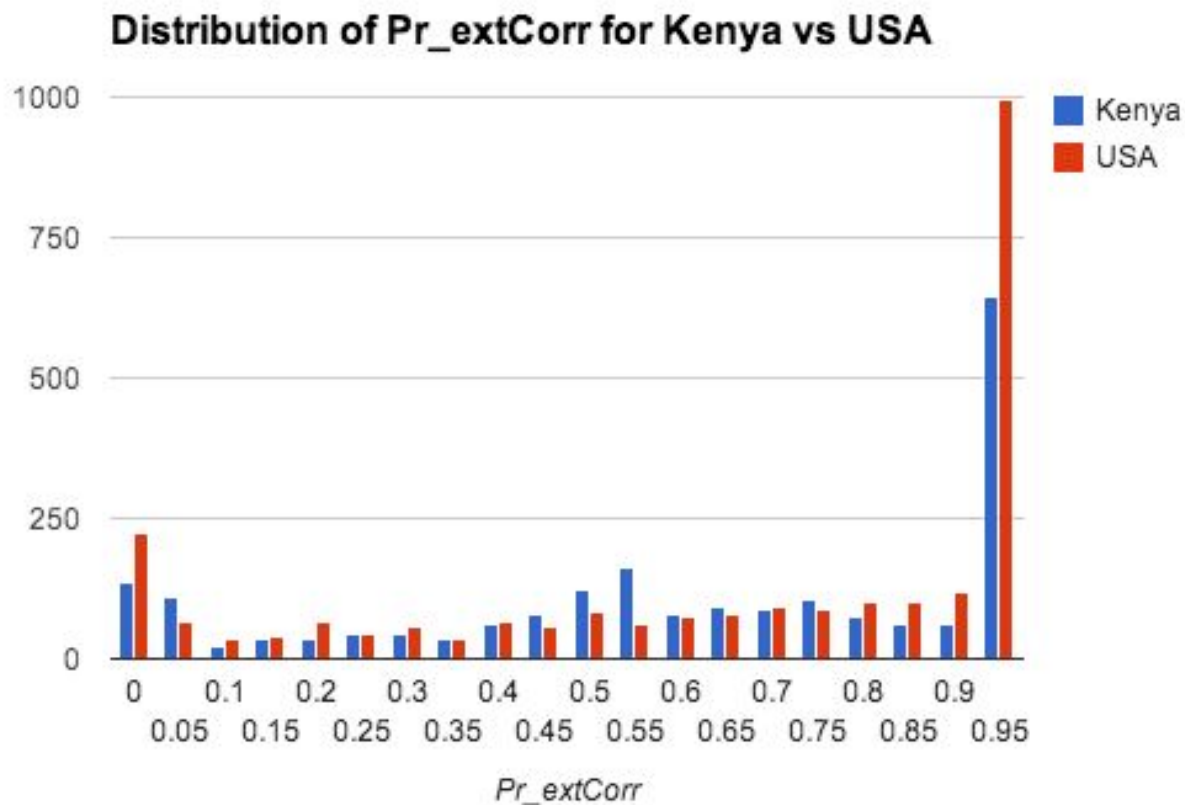


Barack Obama's life story contains 'myth, not truth', says biographer. So why did the media report it as truth?

- $Pr_tripleCorr=1$ (higher support)

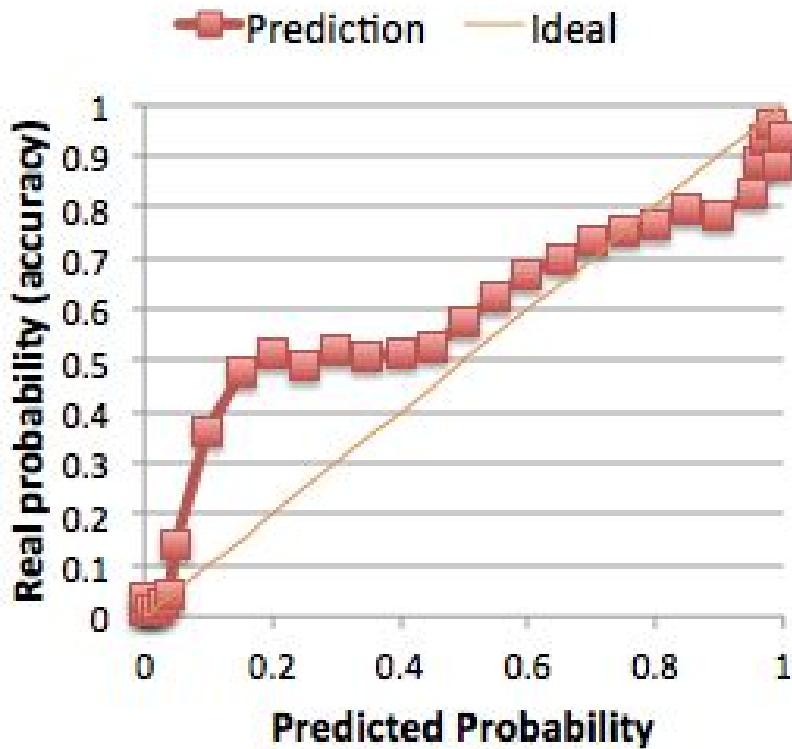
Predicting Extraction and Triple Correctness

Distribution of providers for Kenya and USA

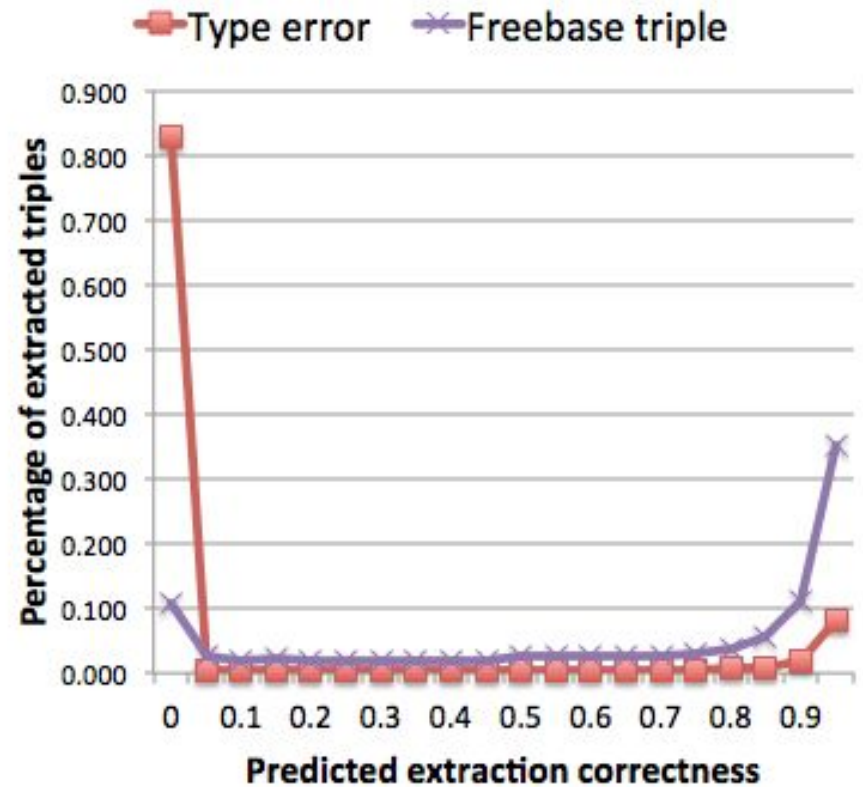


Predicting Extraction and Triple Correctness

Triple Correctness Prediction



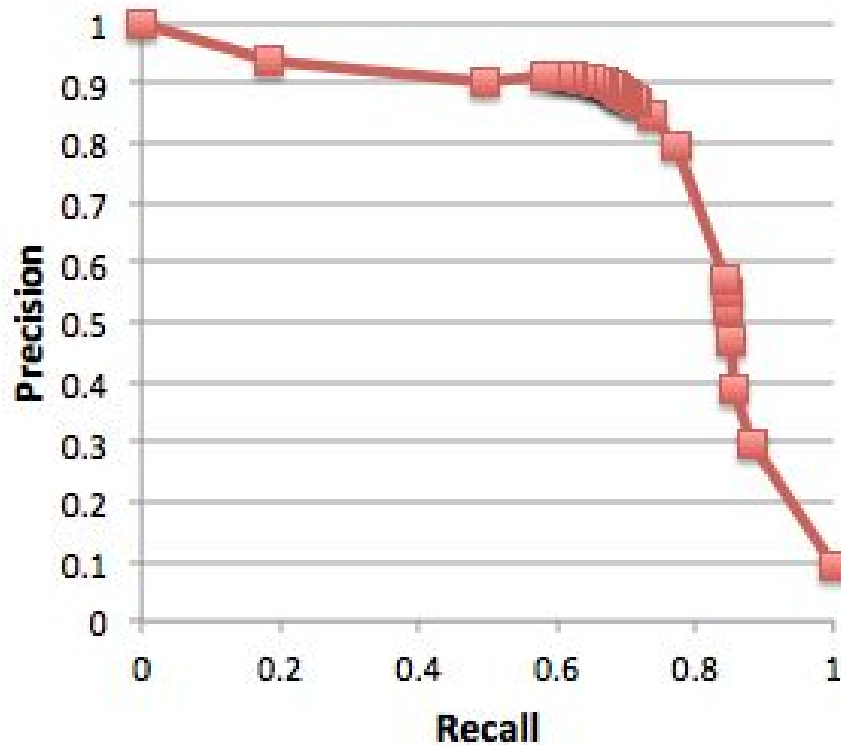
Extraction Correctness Prediction



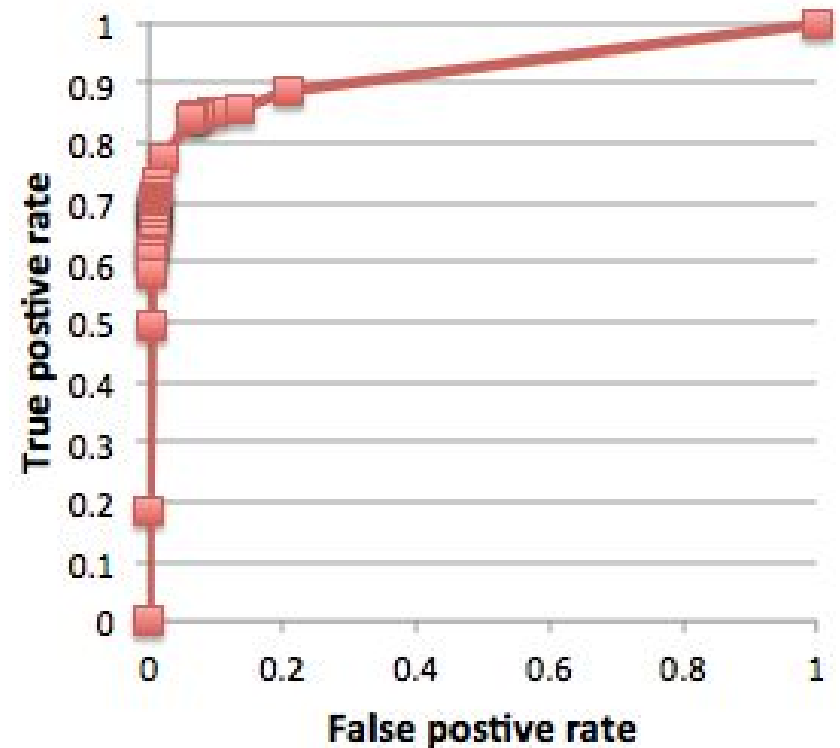
Predicting Triple Correctness

.....

PR-Curves



ROC-Curves



.....

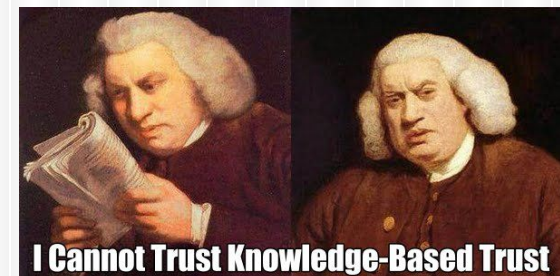
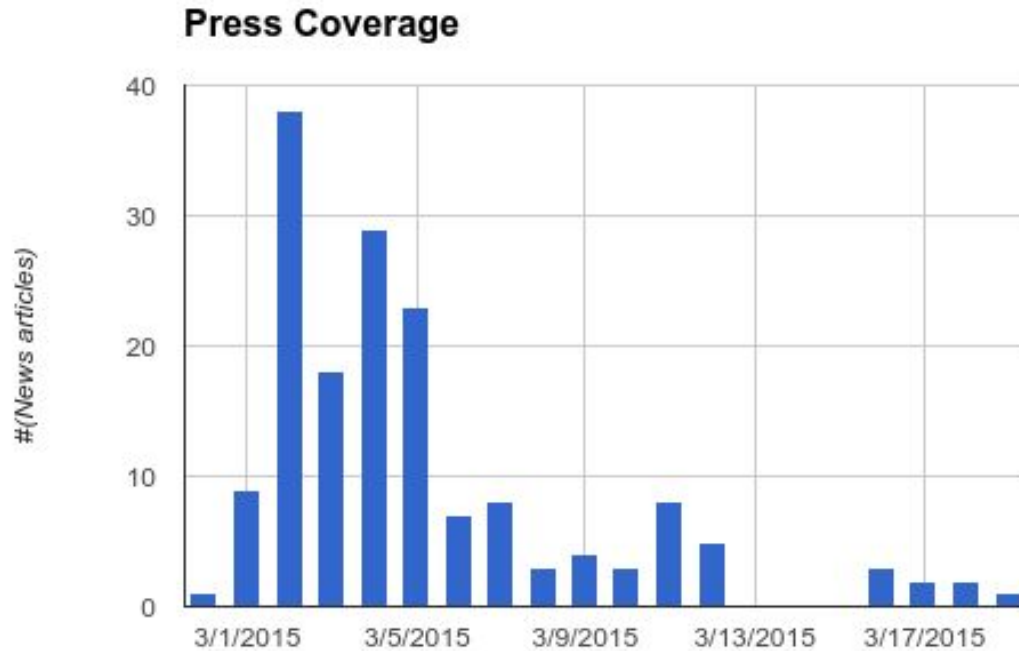
What is the Future of KBT?

Future Works

1. Extraction is still very sparse
 - a. 74% URLs each contributes fewer than 5 triples
 - b. We compute reliable KBT for <20% websites and <<5% webpages
2. Extraction is of low quality
 - a. Overall accuracy is as low as 11.5%
 - b. Low accuracy for some good sources because of undetected extraction errors

Call to arms --
Leave NO Valuable Data Behind

Press Coverage of the Paper



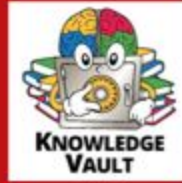
KBT Anecdote (Emails Dated 3/2015)

.....

... I read with interest your recent paper on KBT ...
Actually, that's false – I tried to read it, and did read all of
the parts that w... k characters. It
is quite an intere...

I'm writing becau... claiming that
YYY would be in... em that took
KBT into accoun... e 16 in your
paper ...

I'm writing with... you provide me
with the XXX's... ranking, and
how it compares to YYY's? ...



**KEEP
CALM
IT'S
ONLY
RESEARCH**

The Washington Post

Google has developed a technology to tell whether ‘facts’ on the Internet are true

It’s unclear exactly what Google plans to do with this new technology, if anything at all. Still, even the *possibility* of a search engine that evaluates truth is a pretty incredible breakthrough. And it definitely gives new meaning to the phrase [“let me Google that for you.”](#)

.....
THANK YOU!