Gleaning provenance from ~~plagiarism detection~~ article similarity

Team BBCeX
@fantasticlife @ekponk @tristanf

What are we doing?

Measure the similarity of articles

Are they recycled churnalism or original reporting?

Use similarity to show clusters of articles and outlets

Use similarity to find the original source of an article

based on publication date

Our hypothesis

News articles are often plagiarised from other news outlets with little additional original material

People would trust news more if articles showed their provenance

Why?

We think similarity could be an indicator of provenance which is a signal of trust
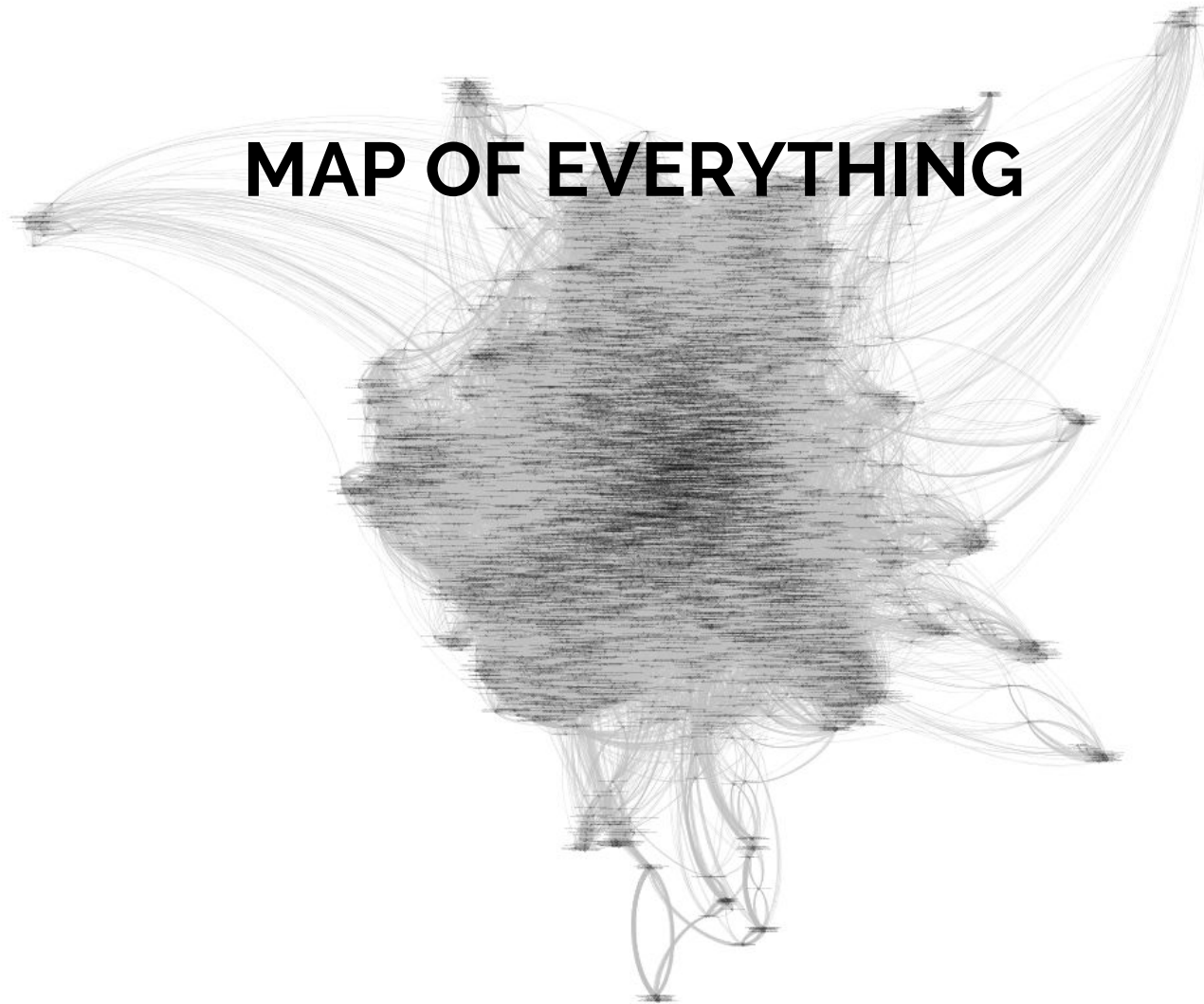
though not entirely sure if it's positive or negative!

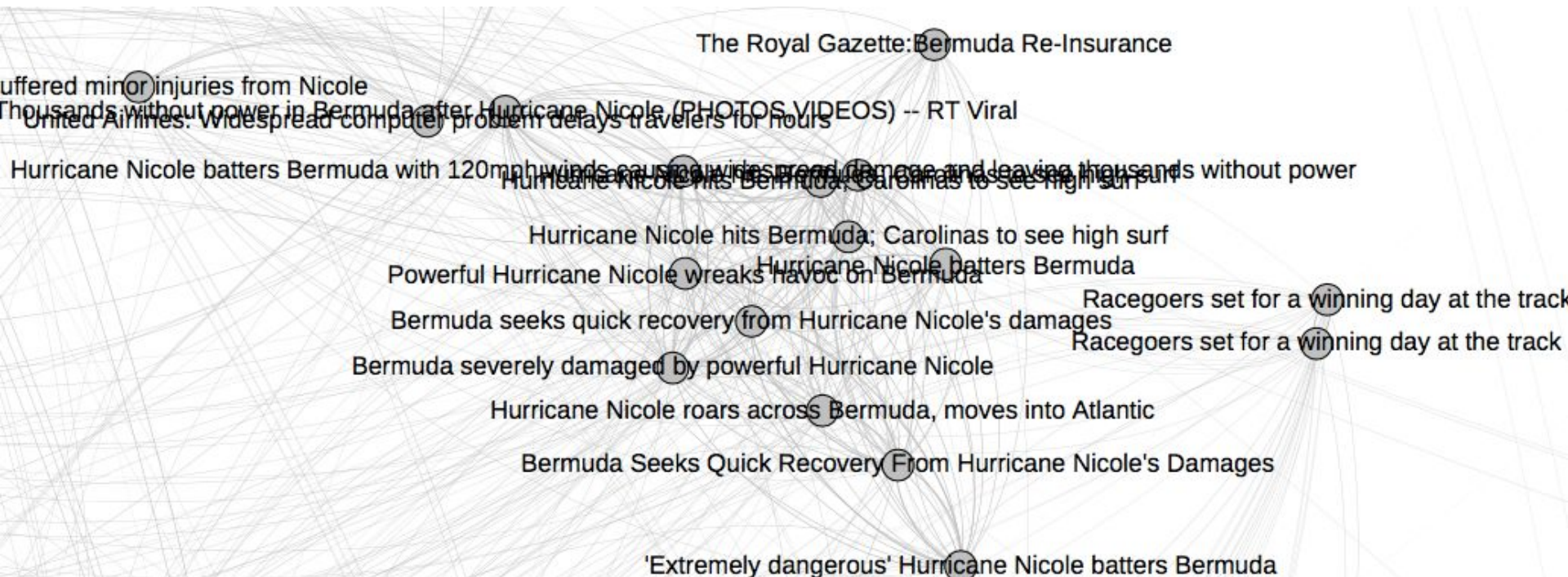*cf "Original reporting" and "Citations and References"*

Probably an indicator to be combined with other things

An investigation into feasibility and usefulness

MAP OF EVERYTHING

A CLUSTER

The Royal Gazette:Bermuda Re-Insurance

uffered minor injuries from Nicole
Thousands without power in Bermuda after Hurricane Nicole (PHOTOS VIDEOS) -- RT Viral
United Airlines: Widespread computer problem delays travelers for hours

Hurricane Nicole batters Bermuda with 120mph winds cause widespread damage and leaving thousands without power
Hurricane Nicole hits Bermuda; Carolinas to see high surf

Hurricane Nicole hits Bermuda; Carolinas to see high surf
Powerful Hurricane Nicole wreaks havoc on Bermuda
Hurricane Nicole batters Bermuda

Racegoers set for a winning day at the track
Bermuda seeks quick recovery from Hurricane Nicole's damages
Racegoers set for a winning day at the track

Bermuda severely damaged by powerful Hurricane Nicole

Hurricane Nicole roars across Bermuda, moves into Atlantic

Bermuda Seeks Quick Recovery From Hurricane Nicole's Damages

'Extremely dangerous' Hurricane Nicole batters Bermuda

US playing Syria terrorism card as part of perpetual struggle for global dominance – Assad

Published time: 14 Oct, 2016 03:02
Edited time: 14 Oct, 2016 04:57

Get short URL

© Mohamad Bayoush / Reuters

f 5.4K 🐦 reddit 6 StumbleUpon G+ t

The conflict in Syria is part of Washington's prime goal to impose global hegemony, President Bashar Assad told a Russian newspaper. Terrorists in Syria, he said, are being used by the US and its Gulf allies to assert control over the region, primarily to weaken Iran.

Time

Most similar

Russia TodayRussia Today

The Mirror

Irish Independent

Brisbane Times
Russia Today
RTE

STV
London Evening Standard
The Daily Telegraph
The Indian Express

Tampa Bay Times
U.S. News and World Report
Morocco World News
Montreal Gazette
U.S. News and World Report

South China Morning Post
Russia Today
The Sydney Morning Herald

The Punch

Least similar

There will be no mercy for the children of Aleppo (Photo: Getty)

Get Daily updates directly to your inbox

Enter your email   **+ Subscribe**

68
SHARES

1

Brutal Syrian dictator **Bashar al-Assad** has defied the global outcry over his horror bombing campaign by insisting his forces must 'clean' Aleppo of terrorists.

It came amid fears of another weekend of air-strikes against Aleppo by Russian and **Syrian** pilots could send the death toll soaring yet again.

Assad, speaking to a Russian publication, warned **Aleppo** is now a 'springboard' for a massive military push against so-called 'terrorist' organisations.

Hundreds have died in recent weeks as Russian and pro-Assad forces have forced a bloody stranglehold around as many as 250,000 starving and trapped civilians.
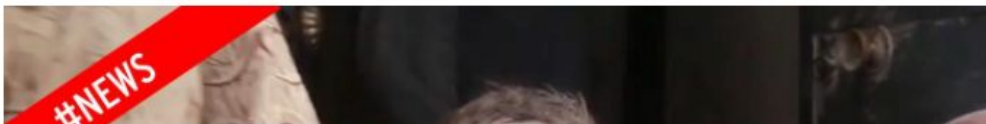
#NEWS

First lady Michelle Obama spoke forcefully for Hillary Clinton, and against Donald Trump, at a campaign rally in New Hampshire on Thursday. JIM COLE/AP

By MICHAEL CASEY, Associated Press

MANCHESTER, N.H. (AP) — Michelle Obama said Thursday that Republican presidential nominee Donald Trump's boast on a 2005 tape about his fame allowing him to "do anything to women" shook her to the core.

Speaking at a campaign rally in New Hampshire for Democratic nominee Hillary Clinton, Mrs. Obama called the comments "shocking and demeaning." She also dismissed Trump's claim that the remarks were simply "locker room talk," saying they were an affront to every woman, parent and citizen in United States.

"I can't stop thinking about this. It has shaken to me to my core in a way that I couldn't have predicted," the first lady said. "So while I would love nothing more than pretend that this isn't happening and come out here and do my normal campaign speech, it would be dishonest and disingenuous to just move onto the next thing like this was all a bad dream. This is not something we can ignore."

Mrs. Obama said the comments made by Trump are part of a long history of demeaning women.

"We have a candidate for president of the United States who over the course of his

Latest Videos

Fire Destroys Gatlinburg Attractions, Homes
WLEX - Lexington, KY

More videos:

Time

Most similar

Least similar

U.S. News and World Report
Philadelphia Daily News
Live Mint
The Hindu

CNN
Deccan Herald
Financial Express
The Indian Express
The Mirror
Malheur Enterprise Independent Newspaper U.S. News and World Report
Bangkok Post Irish Independent Economic Times
Tampa Bay Times The Denver Post The Indian Express Canberra Times Economic Times Economic Times

Michelle Obama tears into Trump

SHARE ARTICLE   f 000K   🐦   reddit 000K   ✉   💬   🖨 PRINT   A | A | A

Michelle Obama said Thursday that Republican presidential nominee Donald Trump's boast on a 2005 tape about his fame allowing him to "do anything to women" shook her to the core.

Speaking at a campaign rally in New Hampshire for Democratic nominee Hillary Clinton, Mrs. Obama called the comments "shocking and demeaning." She also dismissed Trump's claim that the remarks were simply "locker room talk," saying they were an affront to every woman, parent and citizen in United States.

SIMILAR SOURCES

But...

How do you distinguish fake news from original reporting?
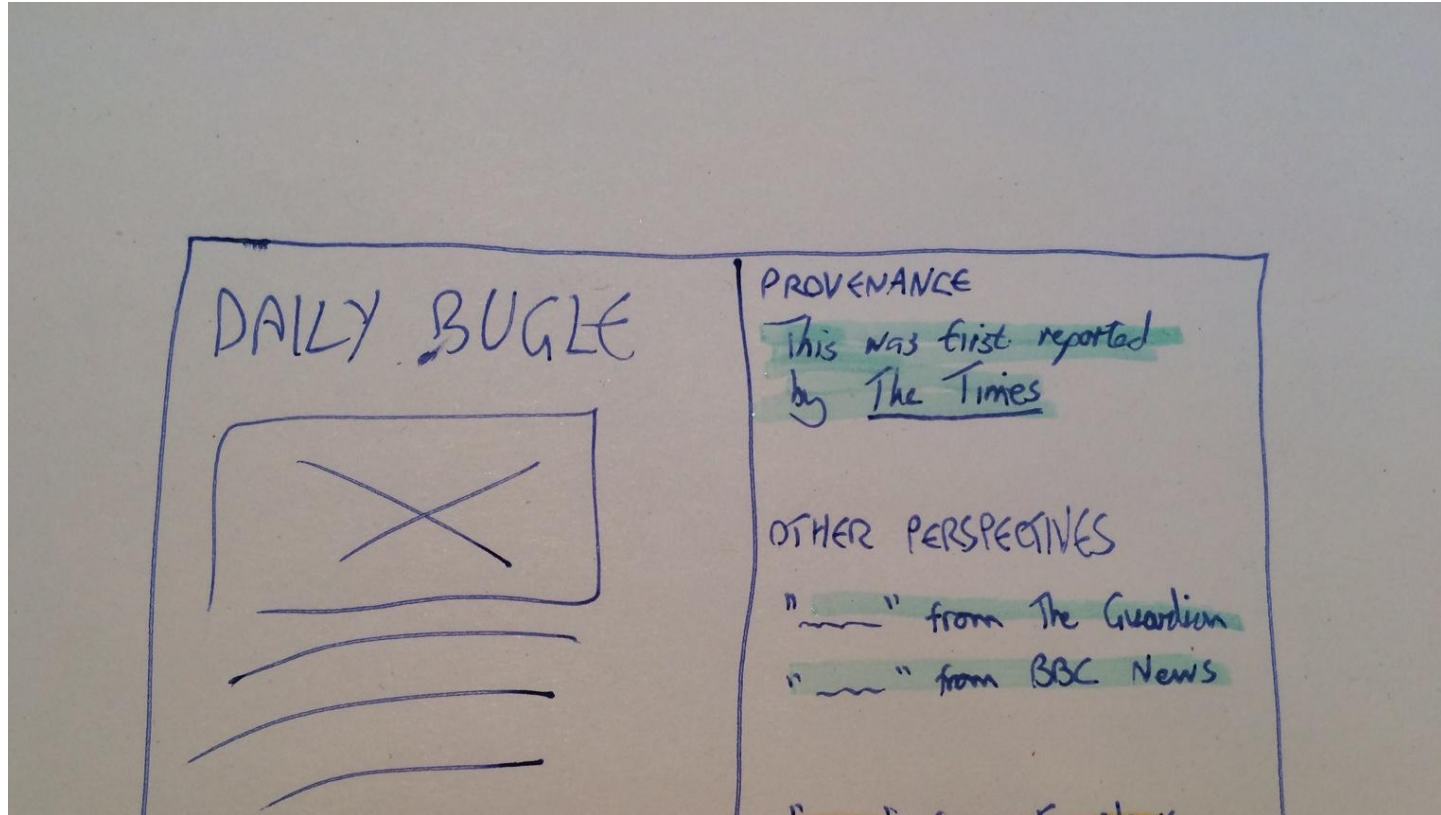They might both be the original or only source

Who would use it?

For consumers reading an article this could...

Show the source (or show if this is the source for others)

Show a churnalism rating (recycled vs original)

Link to other diverse perspectives (using clusters)

BROWSER EXTENSION?

Who would use it?

For publishers

   Show where similar or distinct to competitors

   Rewarding original journalism (getting credit)

Who would use it?

For aggregators and platforms

    Cluster sources

    Identifying the source

    Separate signal from noise

NEXT?

What would we do next?

Validate it

Test other similarity measures

Test with more articles

Add the wires

Update daily

Develop tools for users

Gleaning provenance from ~~plagiarism detection~~ article similarity

Team BBCeX
@fantasticlife @ekponk @tristanf