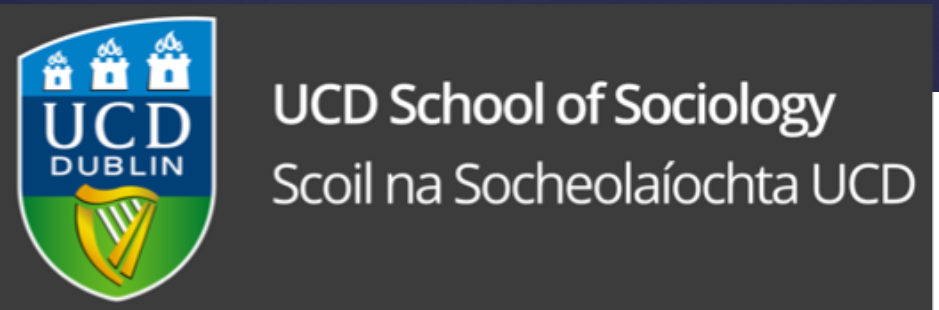
What matters to us most and why?

Studying popularity and attention dynamics via Wikipedia navigation data

Taha Yasseri and Patrick Gildersleve
School of Sociology, University College Dublin

19th August 2020

@TahaYasseri



Wiki Research

- Content

OPEN ACCESS Freely available online

PLOS ONE

A Practical Approach to Language Complexity: A Wikipedia Case Study

Taha Yasseri[1]*, András Kornai[2], János Kertész[1,3]

1 Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary, **2** Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, Hungary, **3** Center for Network Science, Central European University, Budapest, Hungary

- Editorial patterns

OPEN ACCESS Freely available online

PLoS one

Dynamics of Conflicts in Wikipedia

Taha Yasseri[1]*, Robert Sumi[1], András Rung[1], András Kornai[1,2], János Kertész[1]

1 Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary, **2** Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, Hungary

Viewership and Navigation

- Movies financial success
- Electoral popularity
- Disaster triggered collective attention
- Collective memory

- General navigation patterns and article typology

- Attention patterns in relation to news breakouts

OPEN ACCESS Freely available online

PLOS | ONE

Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data

Márton Mestyán[1], Taha Yasseri[1,2,3]*, János Kertész[1,3,4]

1 Institute of Physics, Budapest University of Technology and Economics, Budapest, Hungary, **2** Oxford Internet Institute, University of Oxford, Oxford, United Kingdom, **3** Department of Biomedical Engineering and Computational Science, Aalto University, Aalto, Finland, **4** Center for Network Science, Central European University, Budapest, Hungary

The Hollywood Reporter

MOVIES | TV | MUSIC | TECH | THE BUSINESS

NEWS | REVIEWS | TRAILERS | BOX OFFICE | HEAT VISION

How Wikipedia Can Help Track Movie Buzz

3:08 PM PST 11/8/2012 by Seth Abramovitch

Scientists have found a correlation between activity on the online encyclopedia and a movie's success.

Social media such as Facebook and Twitter have proved handy -- if scientifically iffy -- tools for tracking advance movie buzz. Now, a new study out of Hungary finds that Wikipedia could prove equally if not more valuable than those other sites in predicting

theguardian

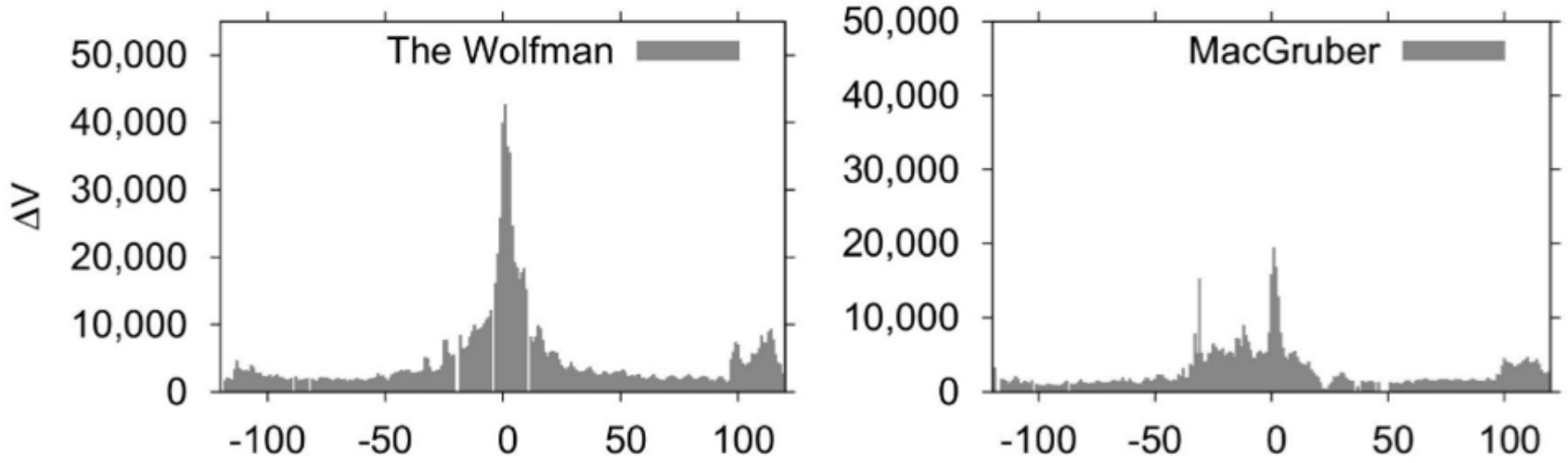
News | Sport | Comment | Culture | Business | Money | Life & style

News › Science › Mathematics

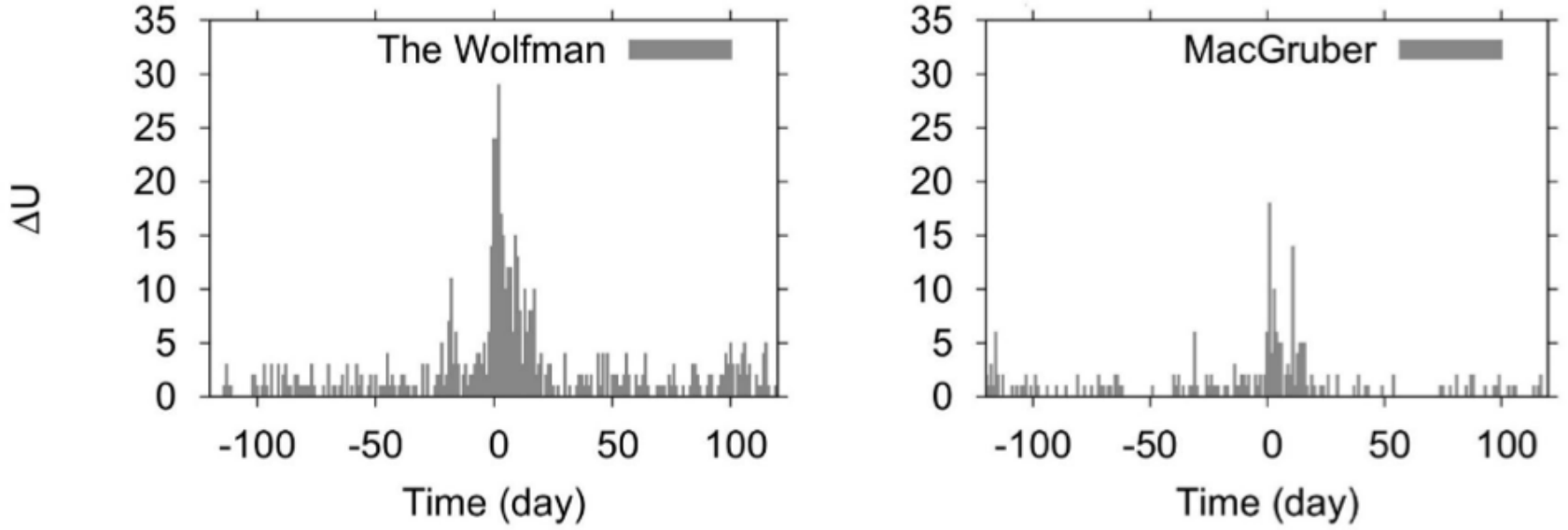
Wikipedia buzz predicts blockbuster movies' takings weeks before release

Mathematical model based on Wikipedia activity forecasts the earnings of the biggest movies with 90% accuracy

page views



unique users

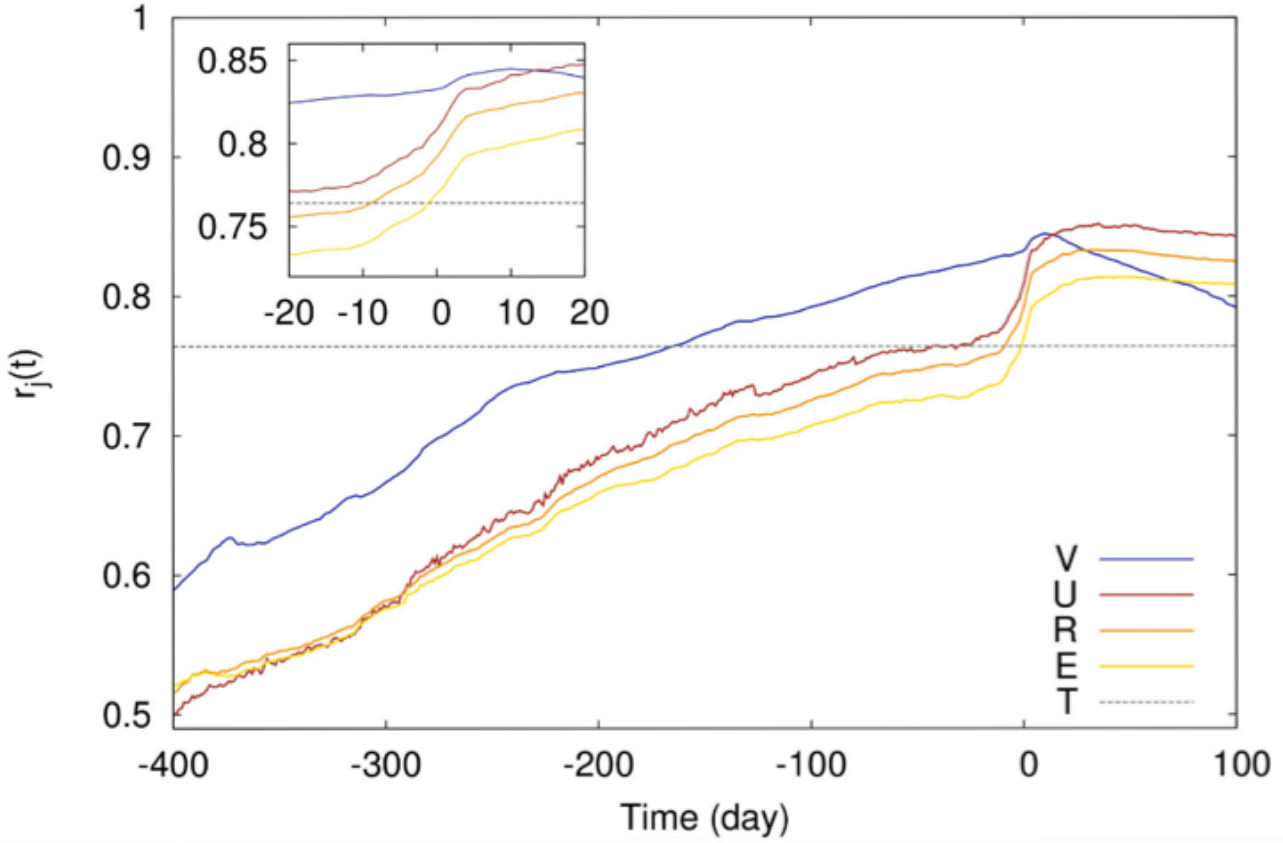


@TahaYasseri

312 movies, US Market. 2010

Inputs:

- Page views
- Edits
- Unique Editors
- Rigor
- Number of Theaters

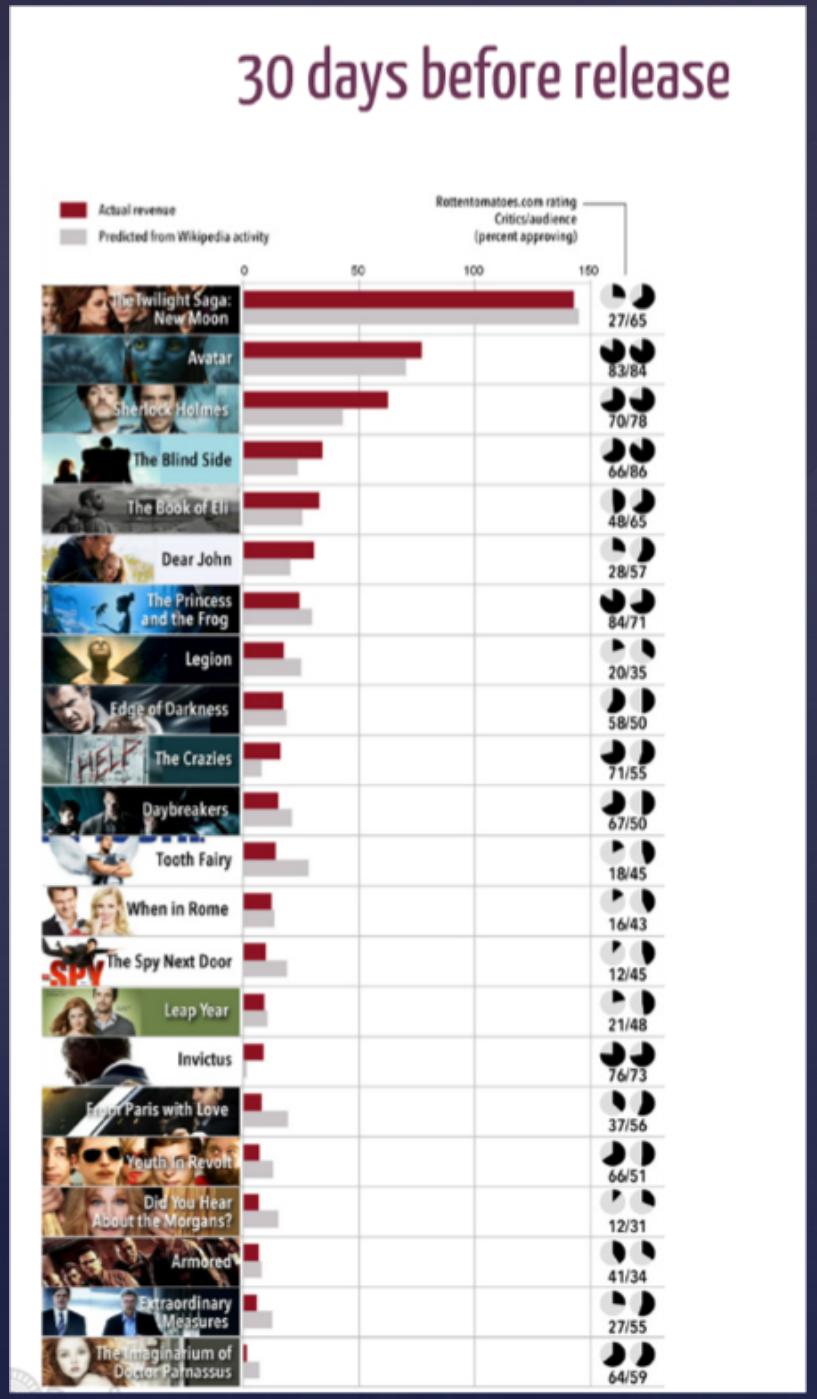


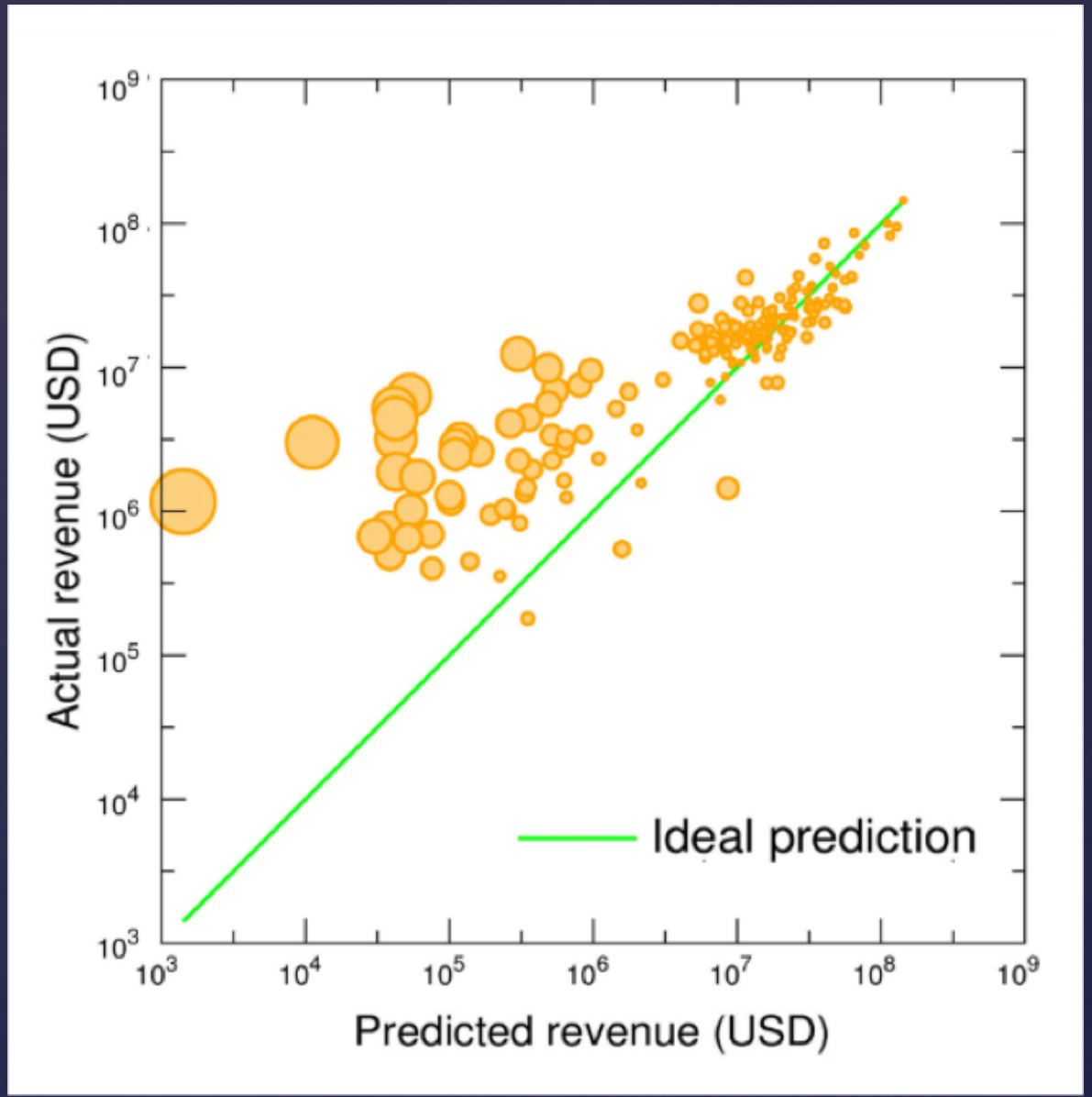
to Predict:

Linear Regression

- First weekend box office revenue

@TahaYasseri





@TahaYasseri

Yasseri and Bright *EPJ Data Science* (2016) 5:22
DOI 10.1140/epjds/s13688-016-0083-3

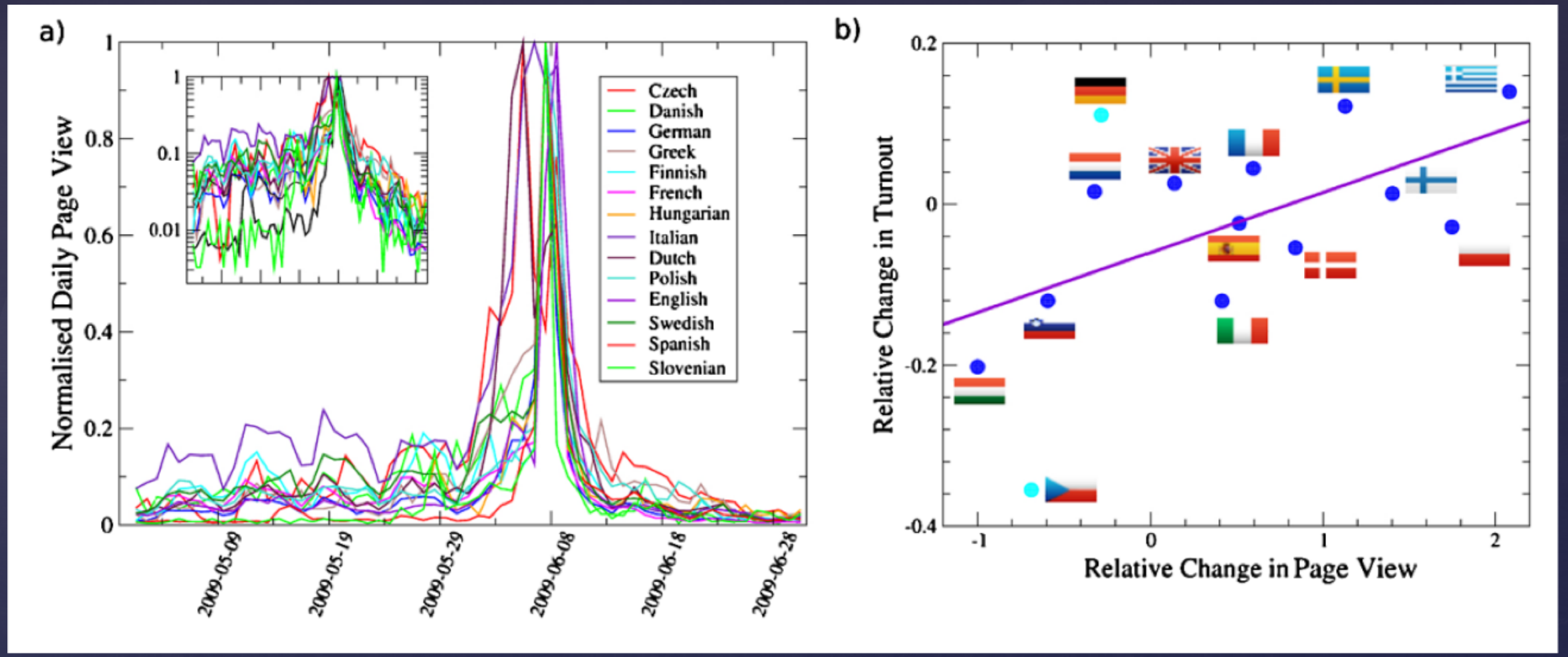
EPJ Data Science
a SpringerOpen Journal

REGULAR ARTICLE Open Access

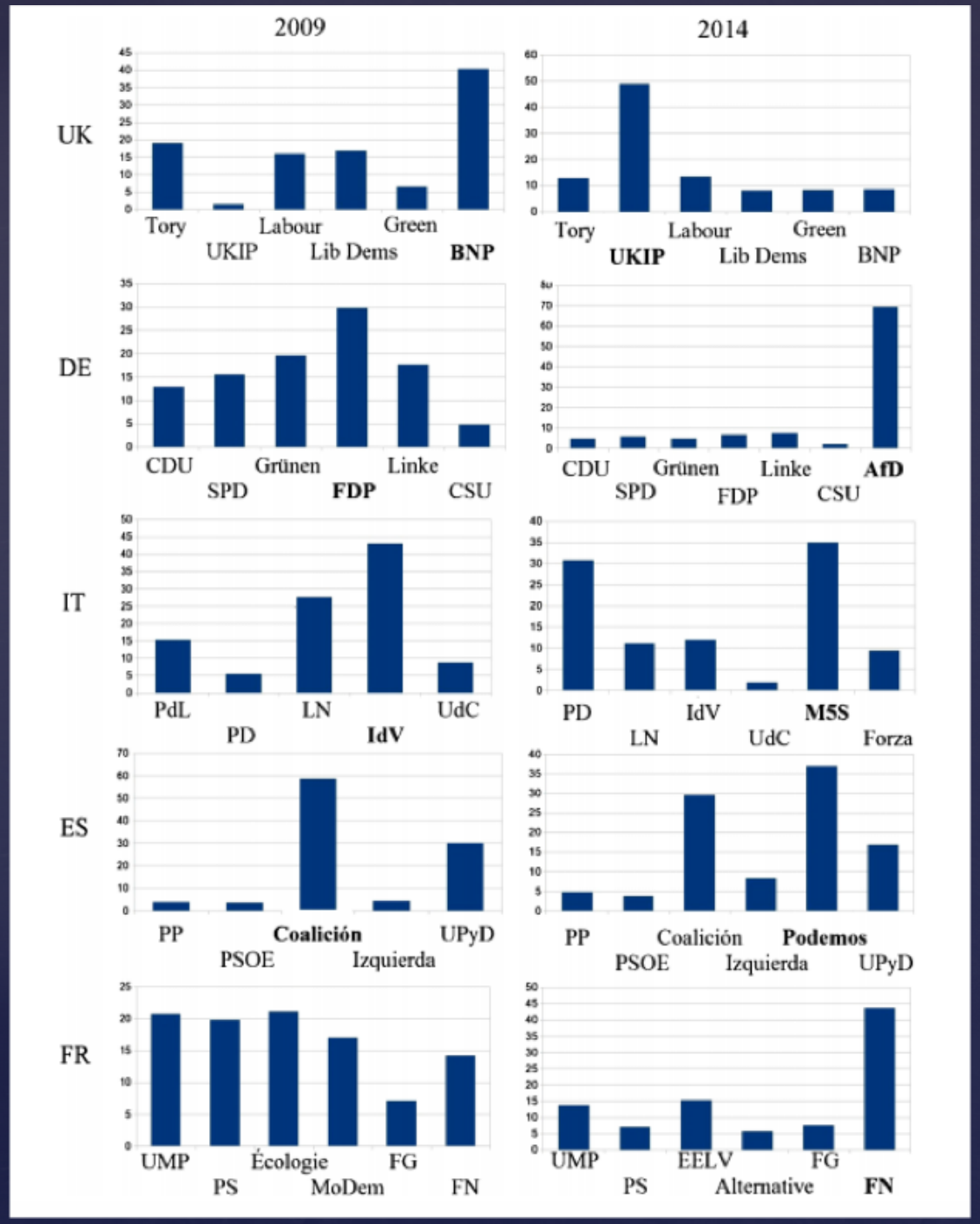
Wikipedia traffic data and electoral prediction: towards theoretically informed models

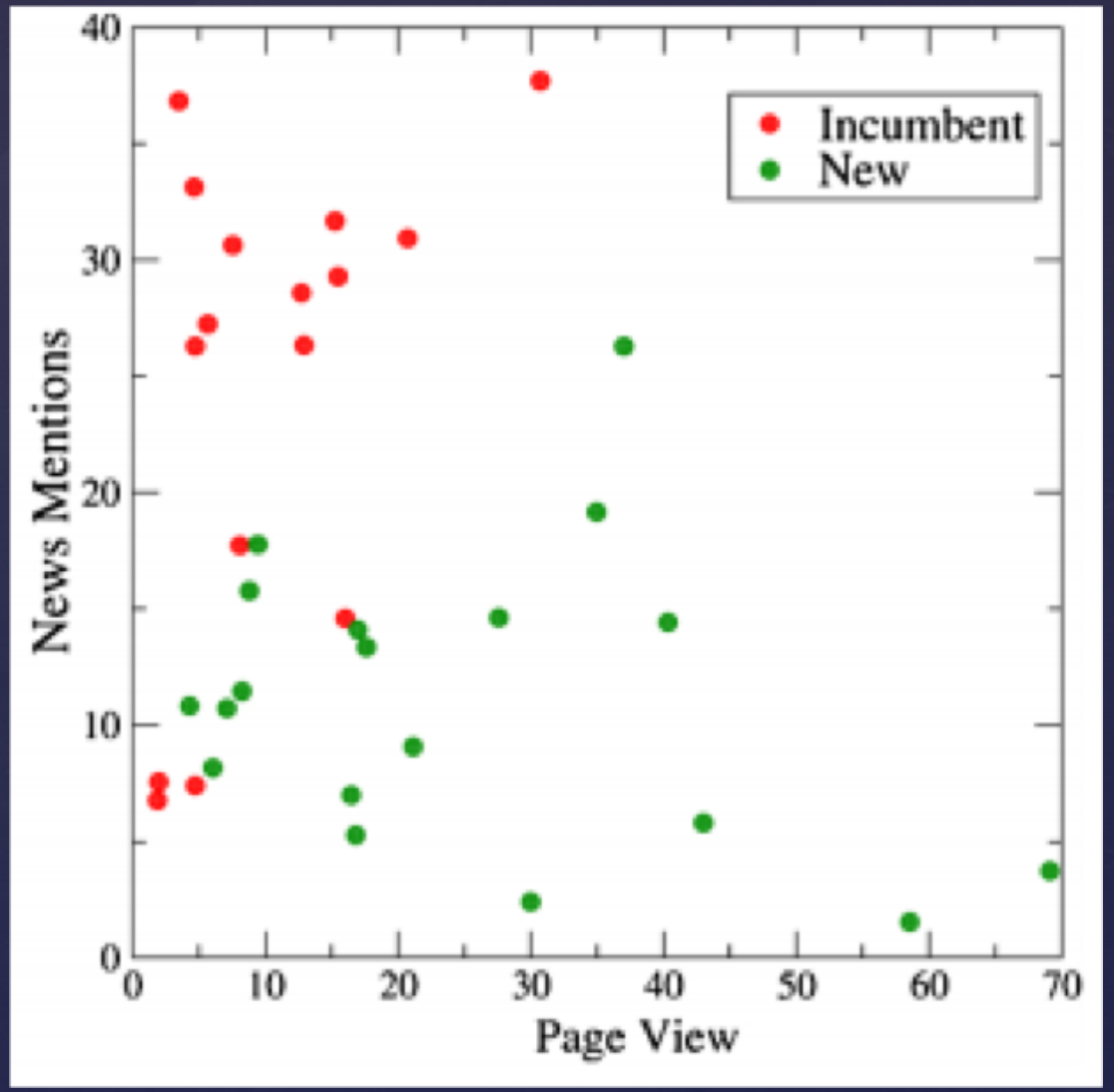
Taha Yasseri*† and Jonathan Bright†

@TahaYasseri



@TahaYasseri





@TahaYasseri

- Attention peaks very close to the election day
- Wikipedia viewership is very much driven by curiosity (rather than absolute popularity)
- The Wikipedia viewership is correlated with the relative **change** in vote share

Voters are cognitive misers who seek information only when considering changing their vote.

"Dynamics and biases of online attention: the case of aircraft crashes."

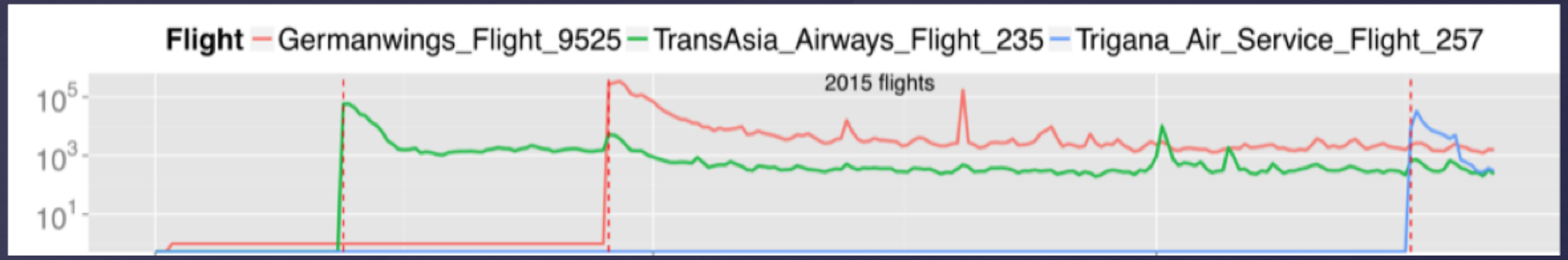
García-Gavilanes R, Tsvetkova M, & Yasseri T.

Royal Society Open Science **3**.10 (2016): 160460.

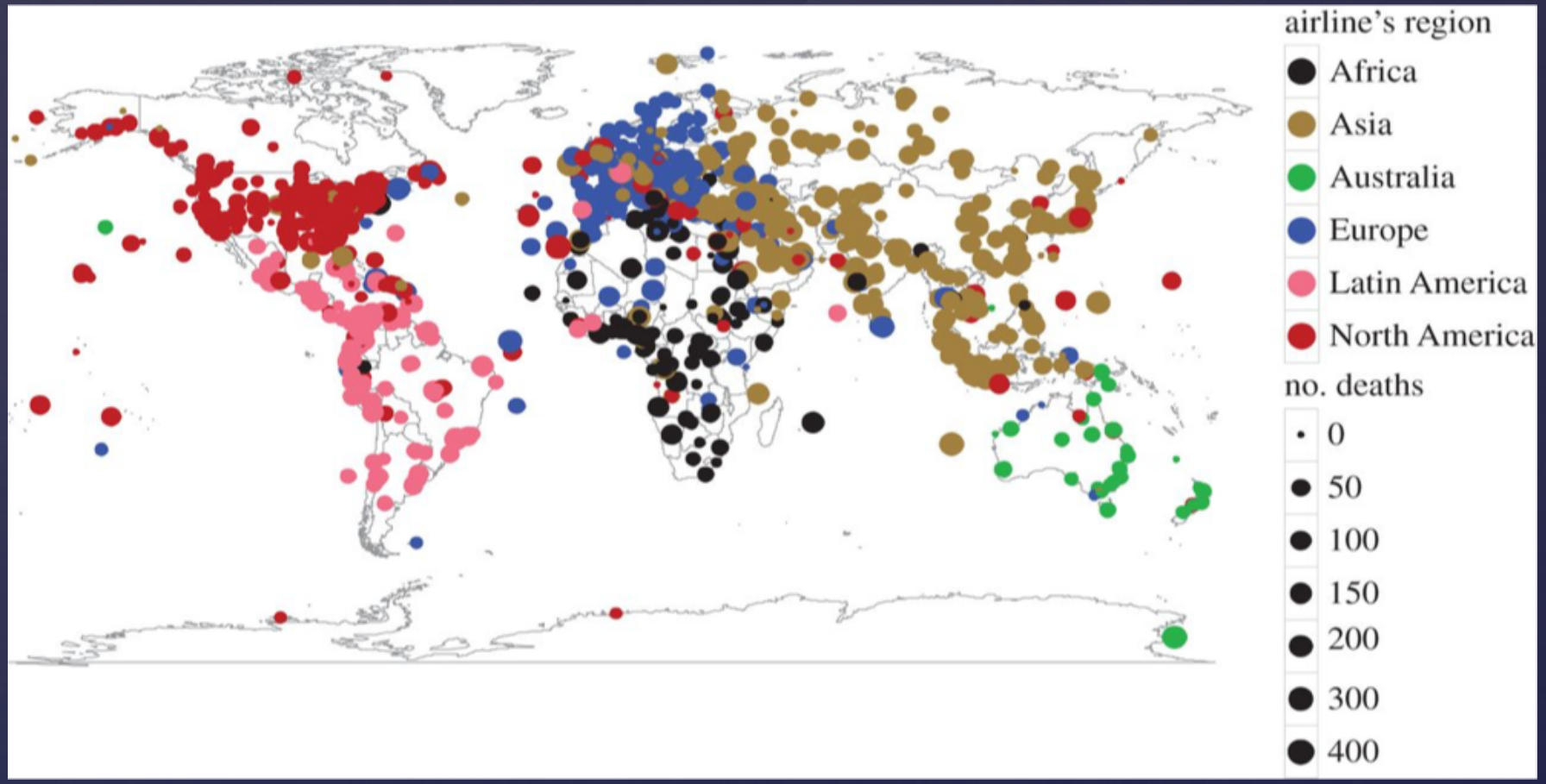
Germanwings Flight 9525, 24 March 2015, Prads-Haute-Bléone, France
Emmanuel Foudrot, Reuters

@TahaYasseri

Wikipedia page views of articles on aviation incidents



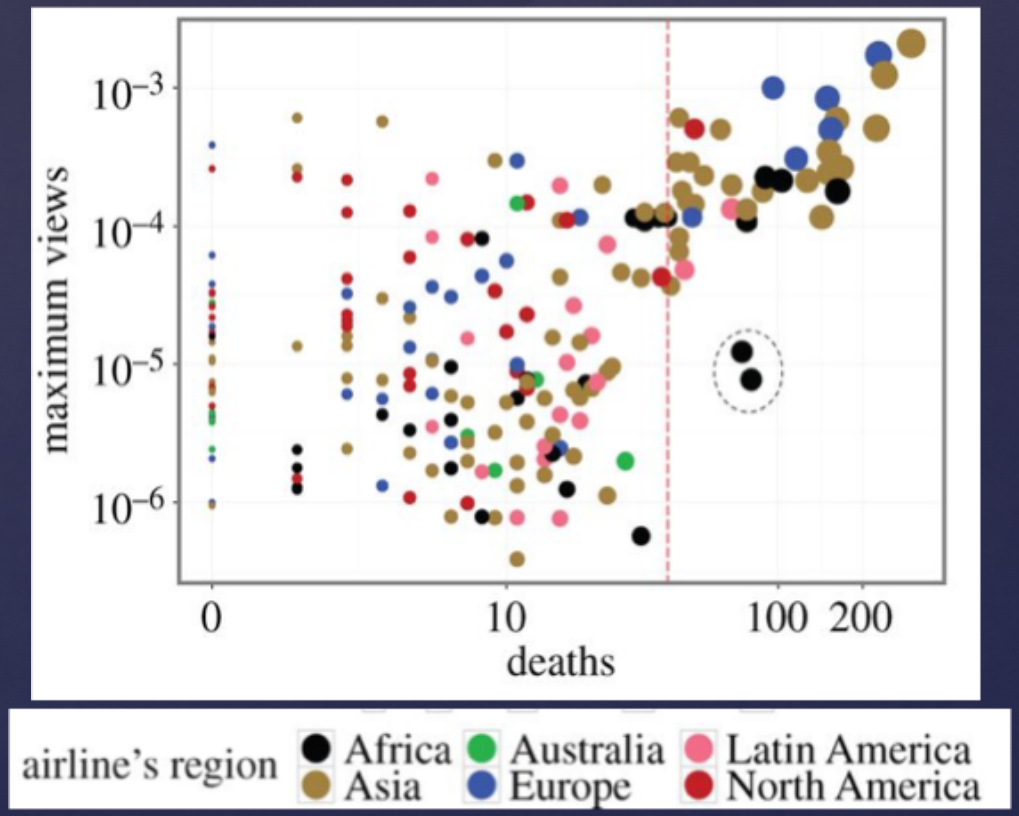
Wikipedia articles on aviation incidents



1496 geolocated incidents and accidents since 1897

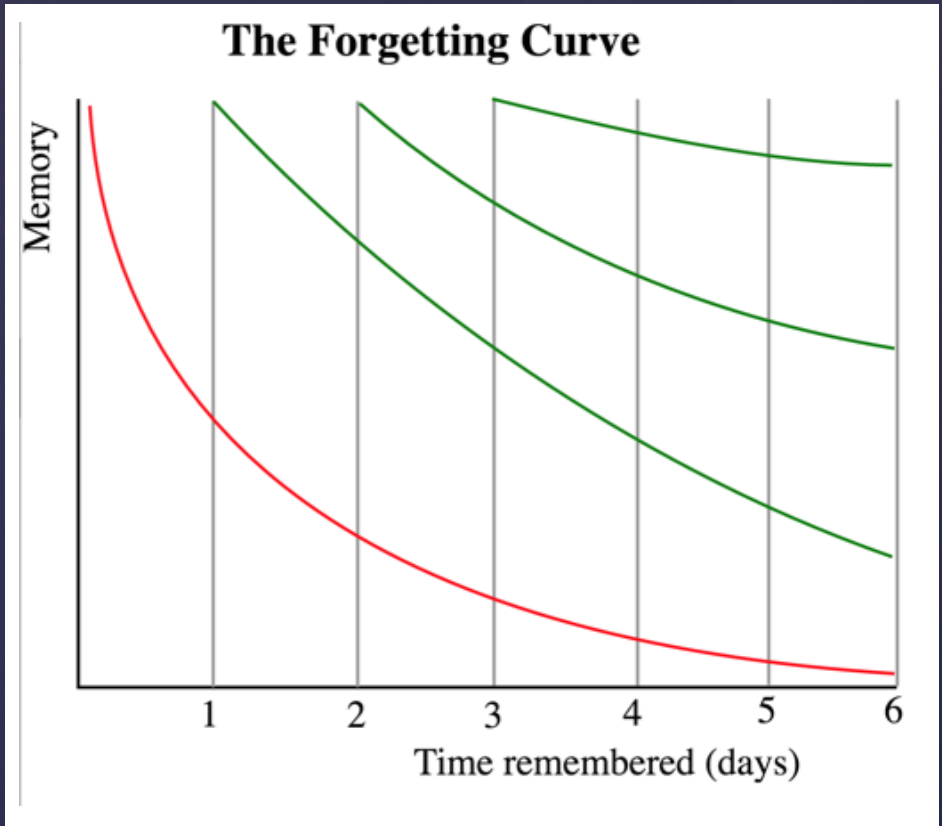
@TahaYasseri

Maximum attention vs Number of deaths



On average
a European death receives **26** times more
attention than an African death

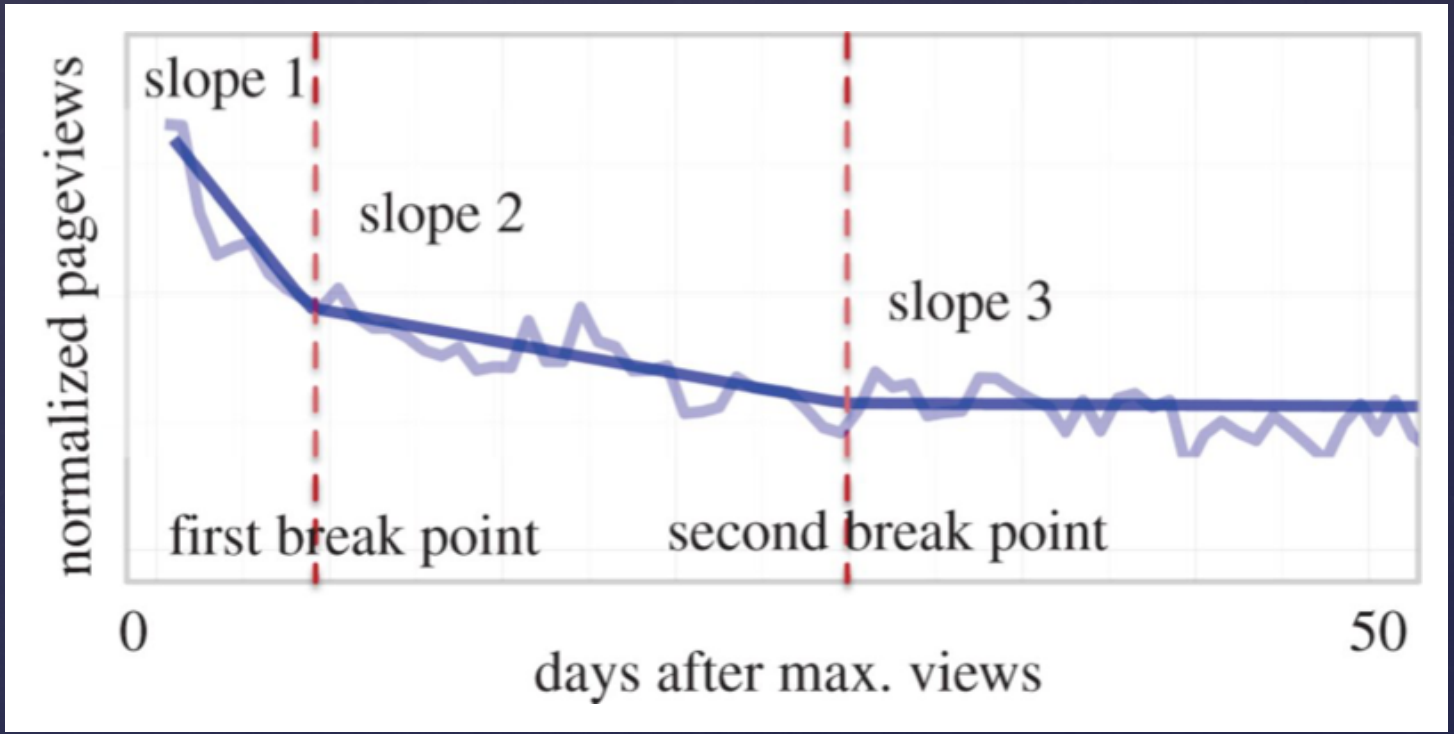
Dynamics of Attention



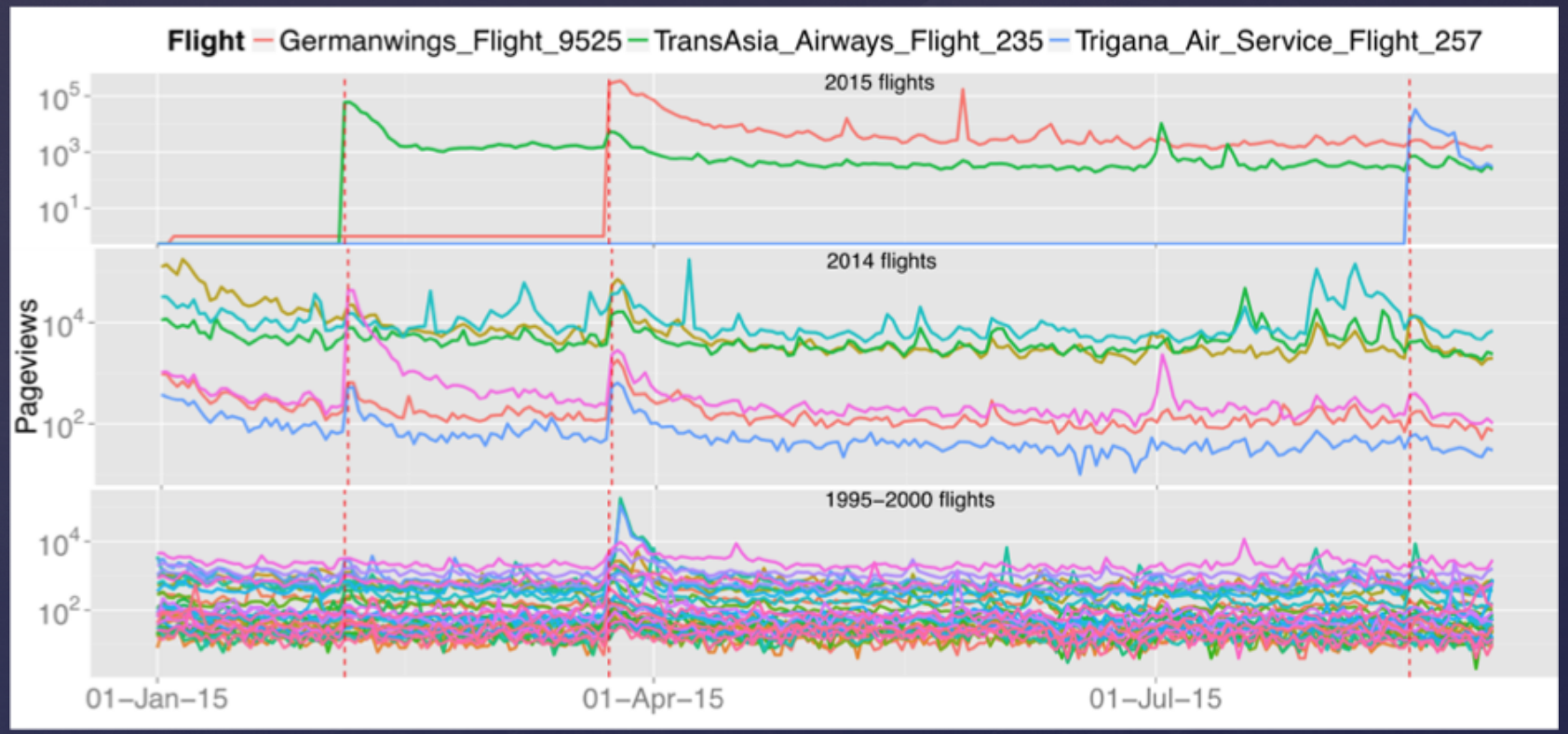
1850-1909

$$R = e^{-\frac{t}{S}}$$

Dynamics of Attention



How do current events remind us of the past events?

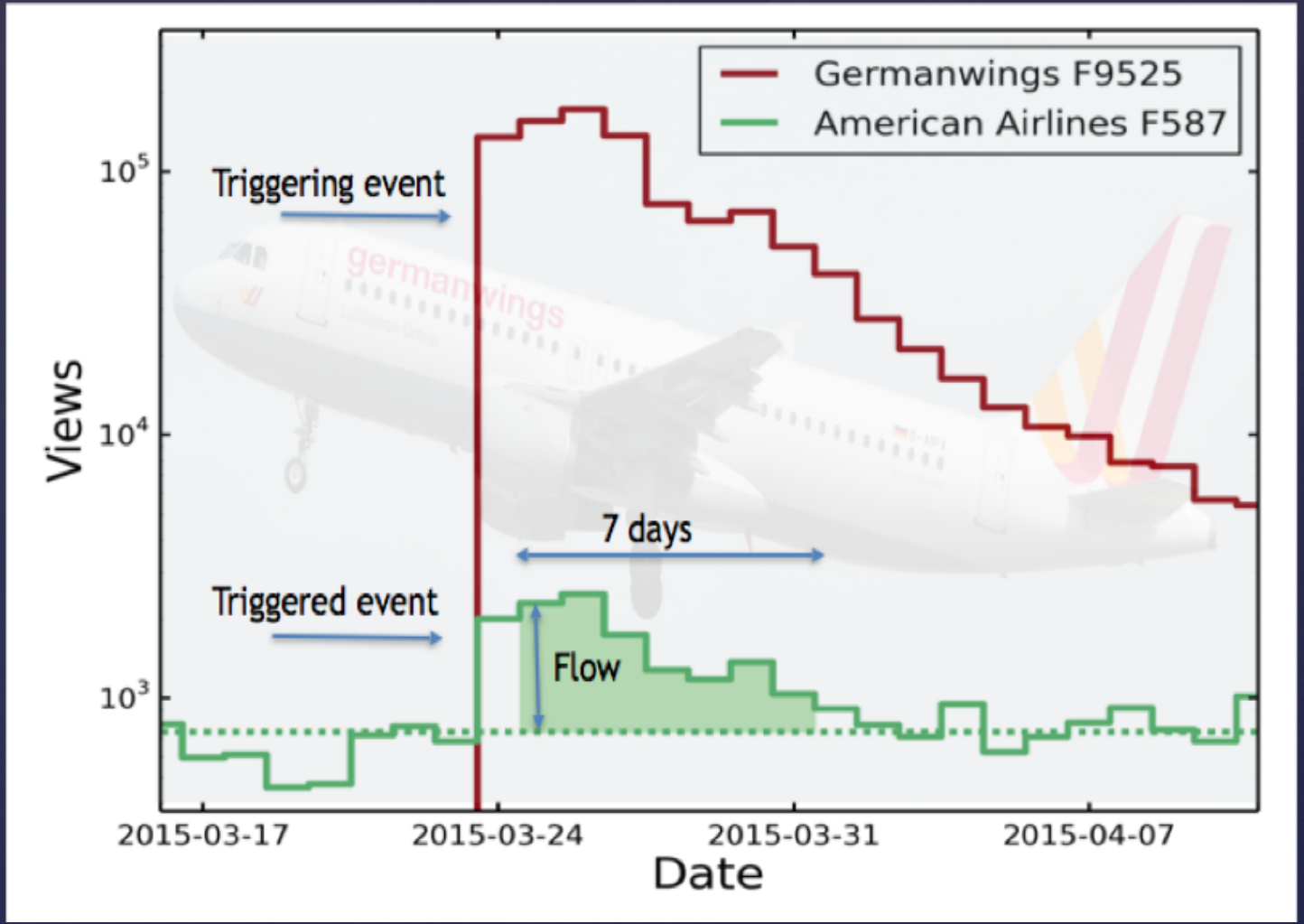


"The memory remains: Understanding collective memory in the digital age."

García-Gavilanes R, Mollgaard A, Tsvetkova M, & Yasseri T.
Science Advances 3.4 (2017): 1602368.

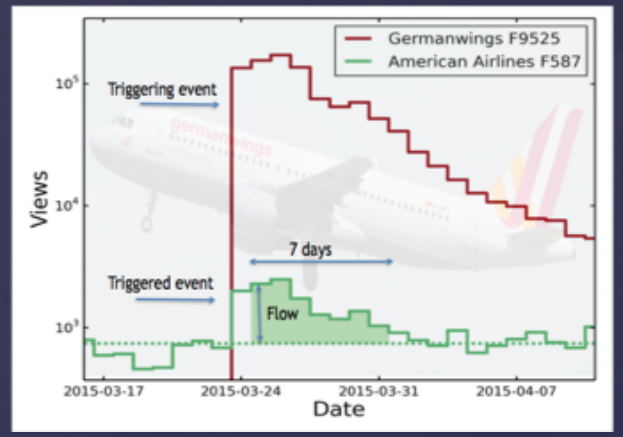
@TahaYasseri

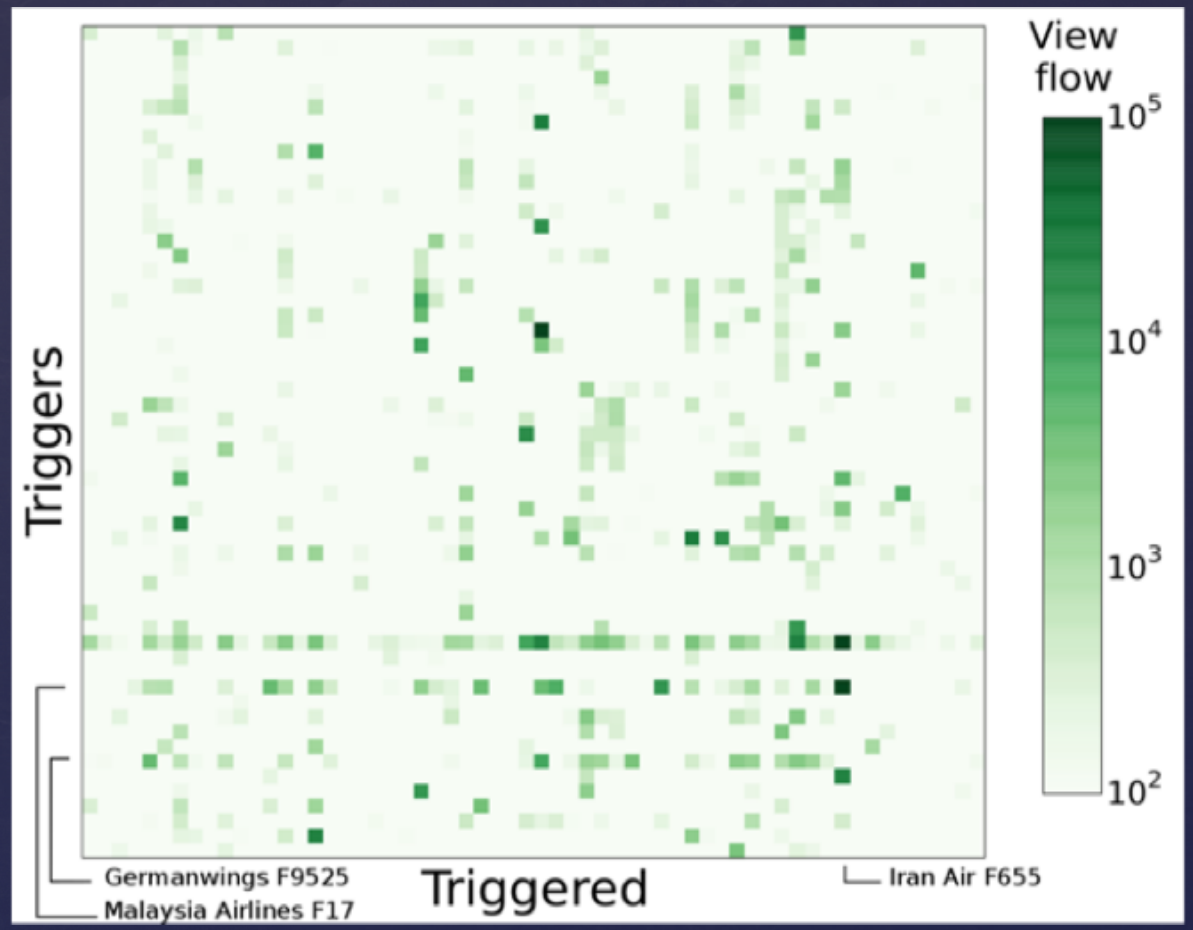
Flow of attention from a current event to a past event



@TahaYasseri

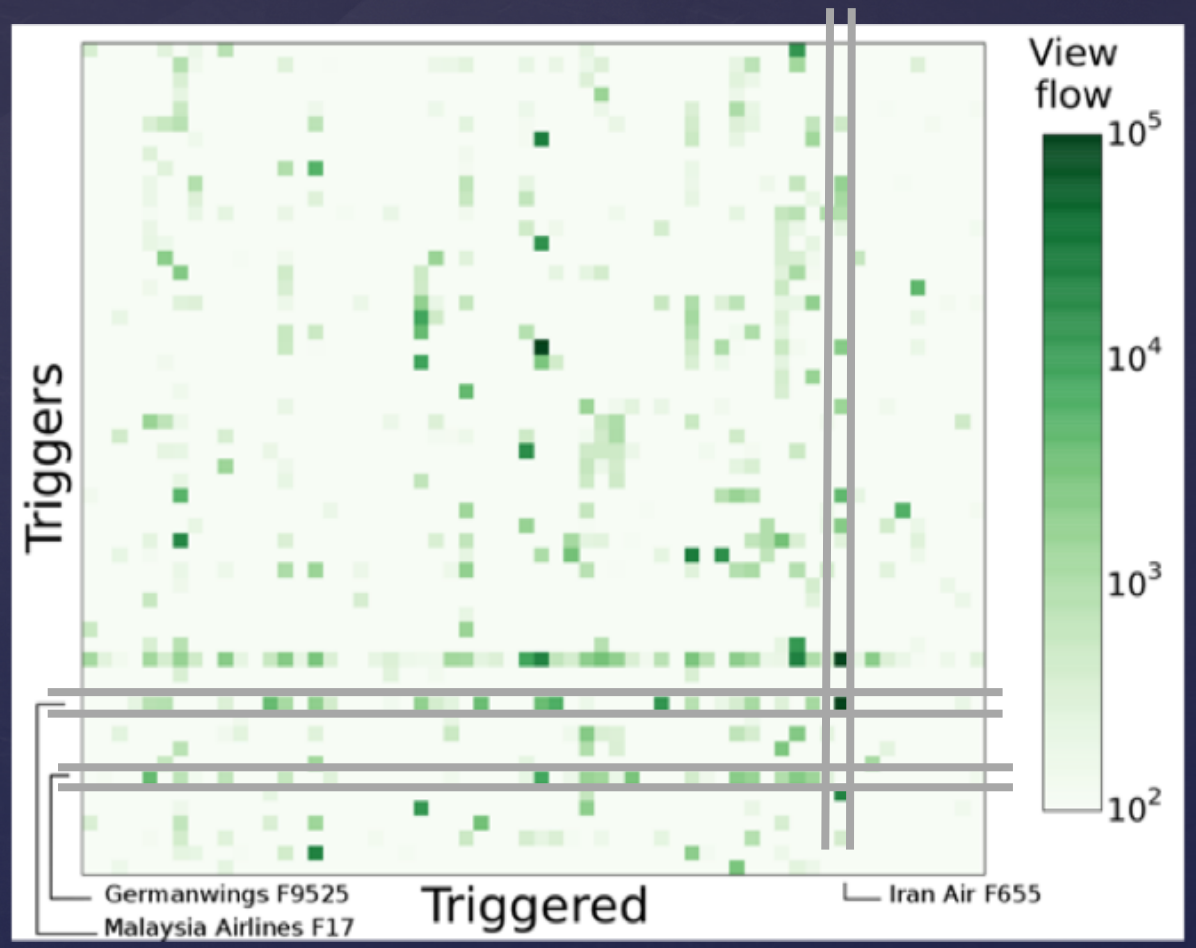
46,732 pairs of events





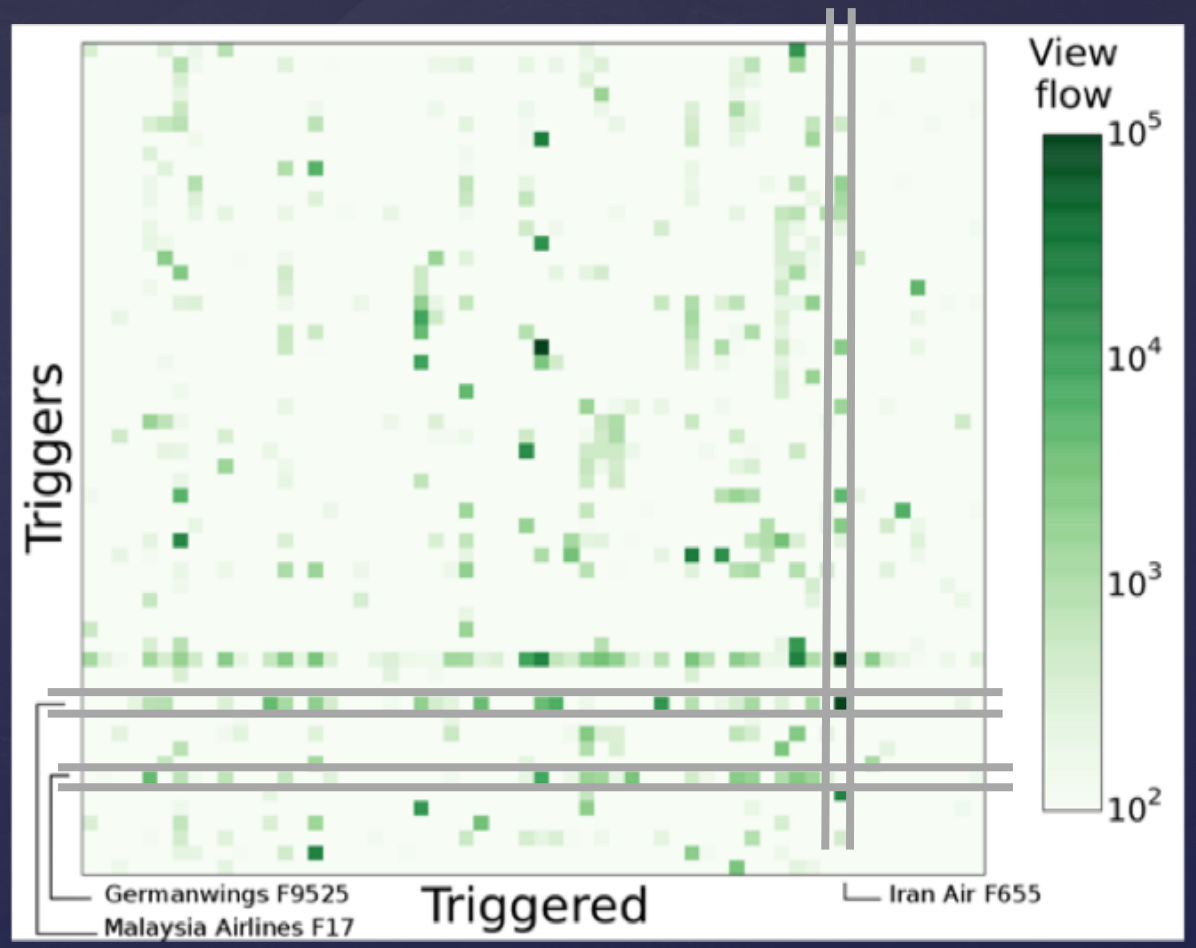
@TahaYasseri

46,732 pairs of events

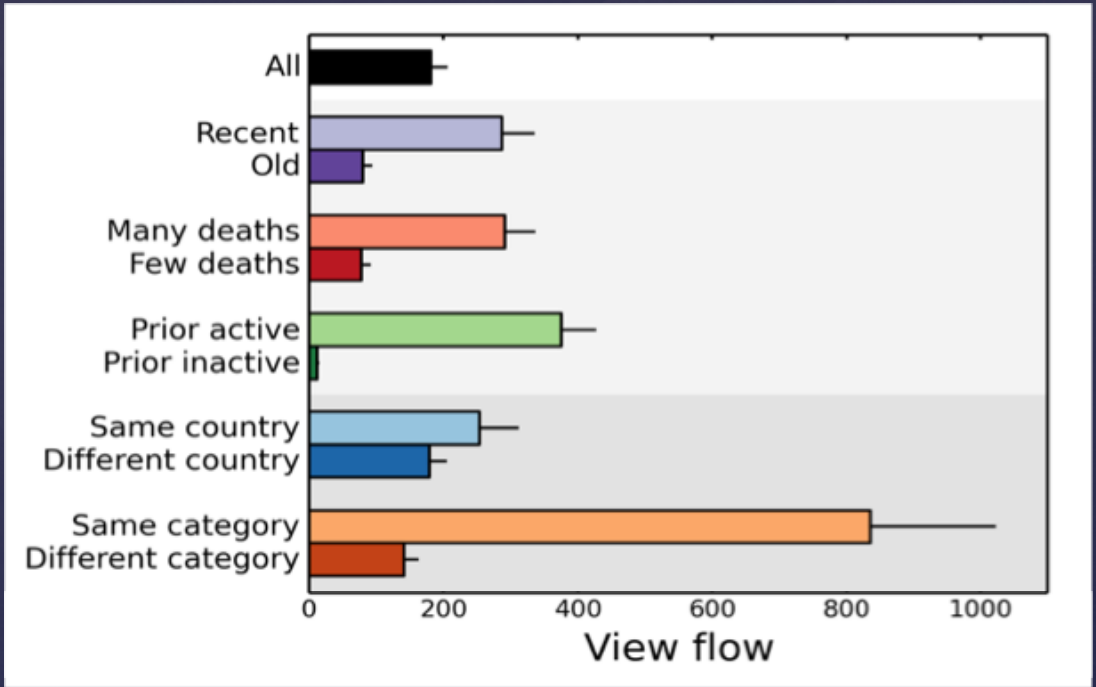


46,732 pairs of events

The sum of all the triggered attention to the past events is on average 140% of the attention to the current event.



Dependence of flow on different features



TransAsia Airways Flight 222

From Wikipedia, the free encyclopedia

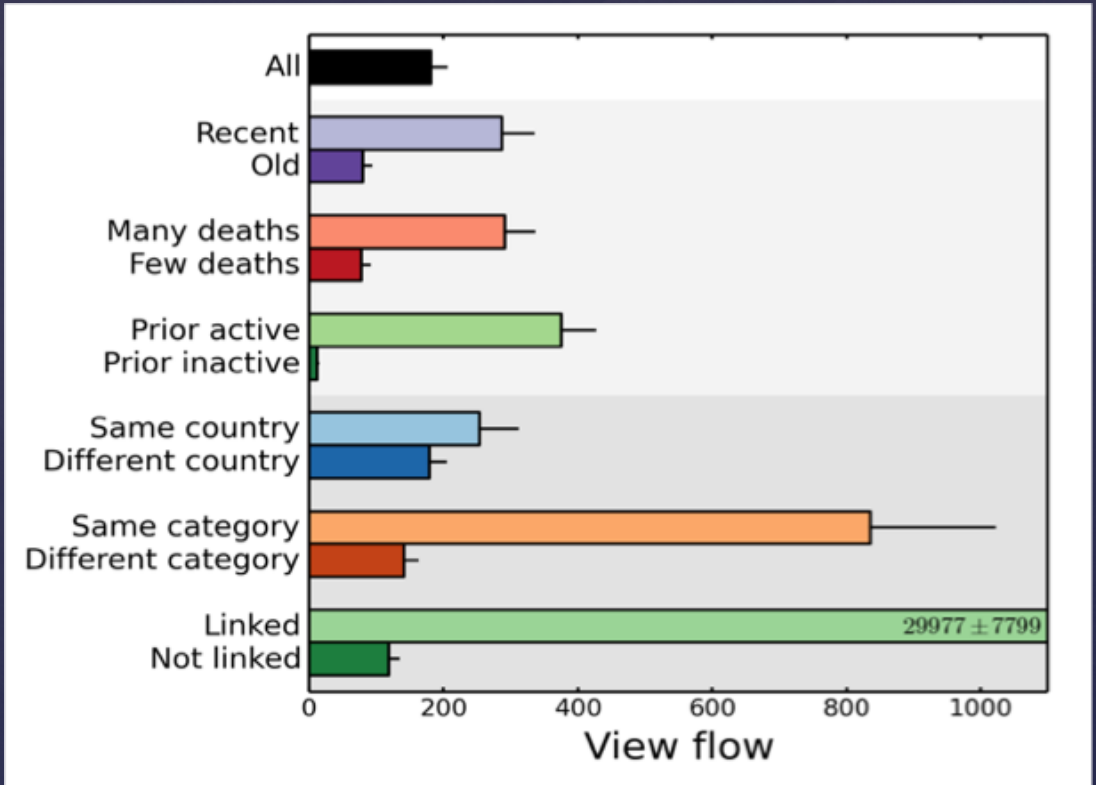
TransAsia Airways Flight 222 (**GE222/TNA222**) was a scheduled domestic passeng
International Airport in Kaohsiung to Magong Airport in Magong, Penghu Island. The a
weather at Magong Airport, Penghu Island, Taiwan, on 23 July 2014. The ATR 72-500

See also [edit]

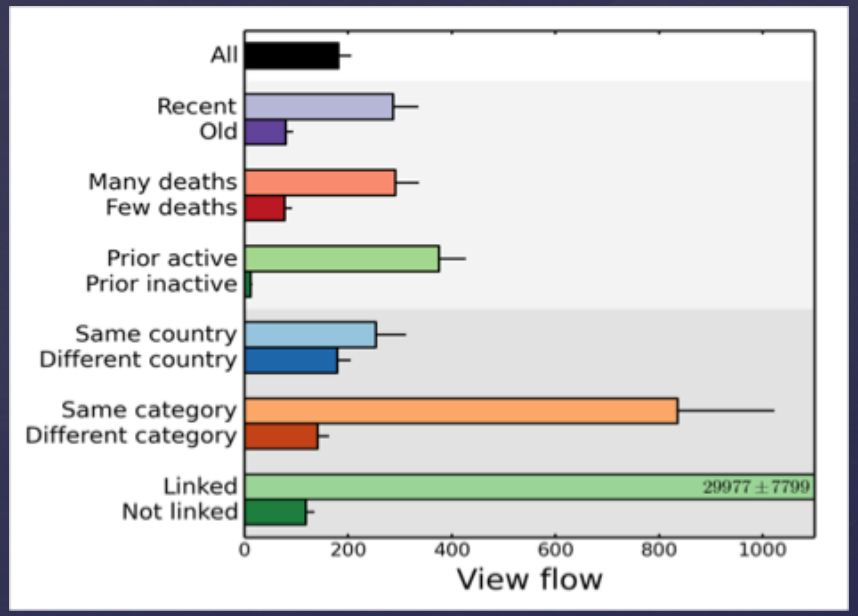
- Aviation Safety Council
- Flydubai Flight 981
- TransAsia Airways Flight 235

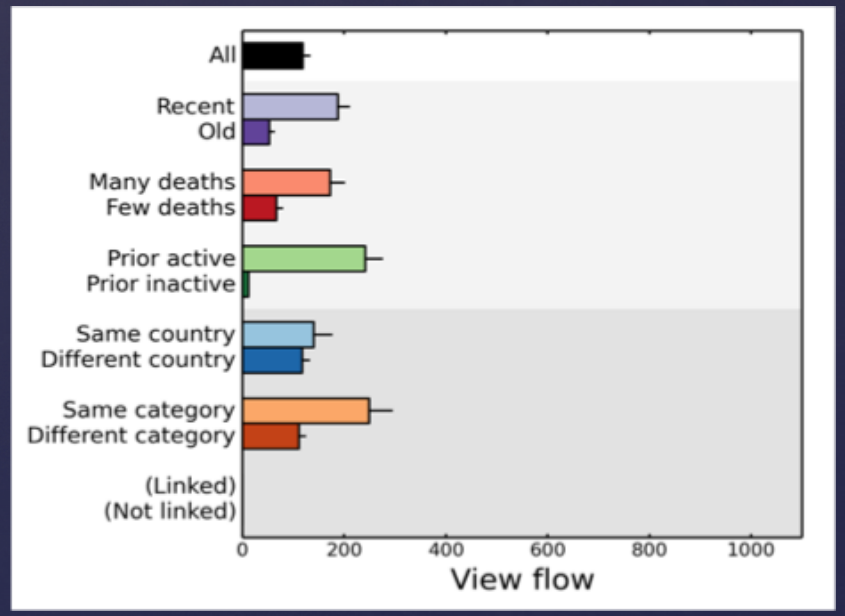
role of **hyperlinks**?

Dependence of flow on different features



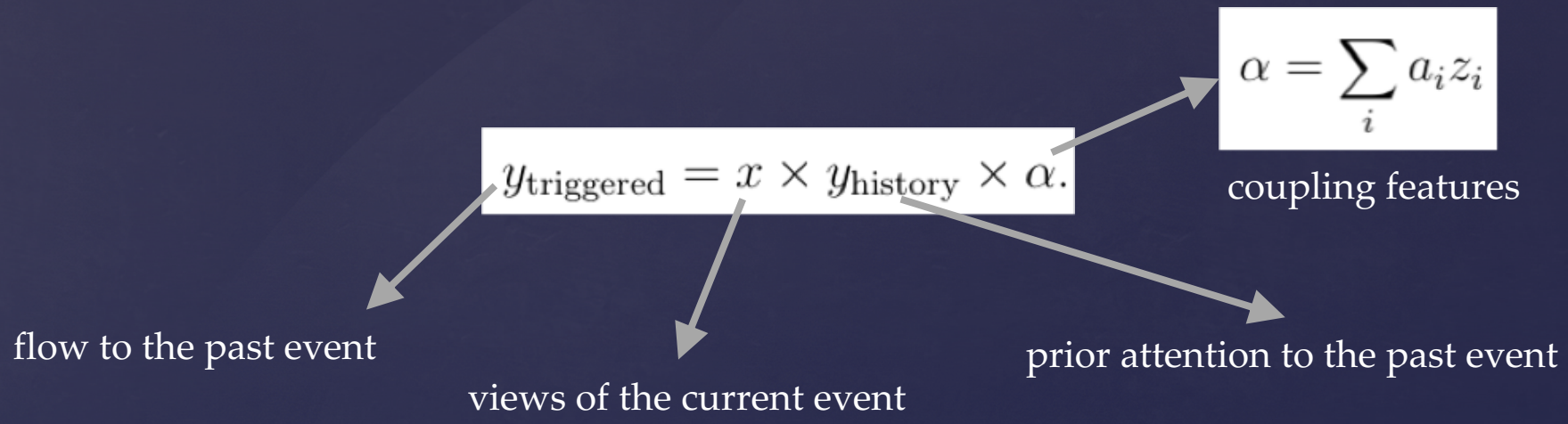
Dependence of flow on different features





pairs **without** a hyperlink

Modelling the flow



R=0.59

Collective Attention and Memory

- The level of attention is complex!
- The dynamics of attention is complex!
- We have a short attention span
- The attention span and decay rate are independent of the impact

However

- We have a long term memory
- It is possible to predict memory patterns