Launching the next generation of digital disease surveillance tools

Mauricio Santillana

Faculty member, CHIP, Boston Children's Hospital Informatics Program
Associate, Harvard Institute for Computational and Applied Sciences
Instructor, Harvard Medical School

HARVARD School of Engineering and Applied Sciences
HARVARD MEDICAL SCHOOL
Boston Children's Hospital

Collaborators:

John S. Brownstein, Harvard Medical School and Healthmap co-founder
Sam Kou, Harvard Statistics Department
Michael Johansson, Centers for Disease Control, Puerto Rico.
Elsie Sunderland, Harvard School of Engineering and Applied Sciences
Caroline Buckee, Harvard School of Public Health
John W. Ayers, San Diego State University
Ben Althouse, Santa Fe Institute
Elaine Nsoesie, University of Washington
Sumiko Mekaru, Epidemico Inc.
Rumi Chunara, New York University
Clark Freifeld, Healthmap co-founder and Epidemico
Ruchit Nagar, Yale University

Students involved:

Shiaho Yang, PhD student in Statistics, Harvard Statistics
Clifton Dassuncao, PhD student in Public Health, HSPH MS' 2014
David W. Zhang, undergraduate applied math student SEAS. MS' 2013
Aditi Hota , undergraduate applied math student SEAS, BS'2014
Andre Nguyen, undergraduate applied math student SEAS, BS'2016
Tamara Louie, Master's student in Public Health, HSPH MS' 2015
Fred Lu, undergraduate applied math student SEAS, BS'2016

HARVARD School of Engineering and Applied Sciences

HARVARD MEDICAL SCHOOL

Boston Children's Hospital

Disclaimer: I am not a medical doctor, clinician, or epidemiologist.

Applied Mathematician and physicist with expertise in machine learning and processing of big data sets. Interested in improving the way public health and medical decisions are made.

…and I respect your privacy!

Big data

Trillions of sensors are monitoring, tracking, and communicating information from multiple locations in real-time

Newspaper articles, Reports, etc…

Predictive Analytics

30+ petabytes of user-generated data stored, accessed, and analyzed

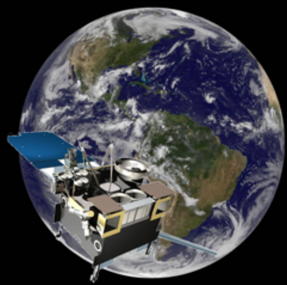
Over 1 **billion** Google searches a day

230 million tweets every day

~2 billion smartphones world wide

Source: IBM

Big data



Weather information in real-time

Newspaper articles, Reports, etc…

Medicine and Public Health

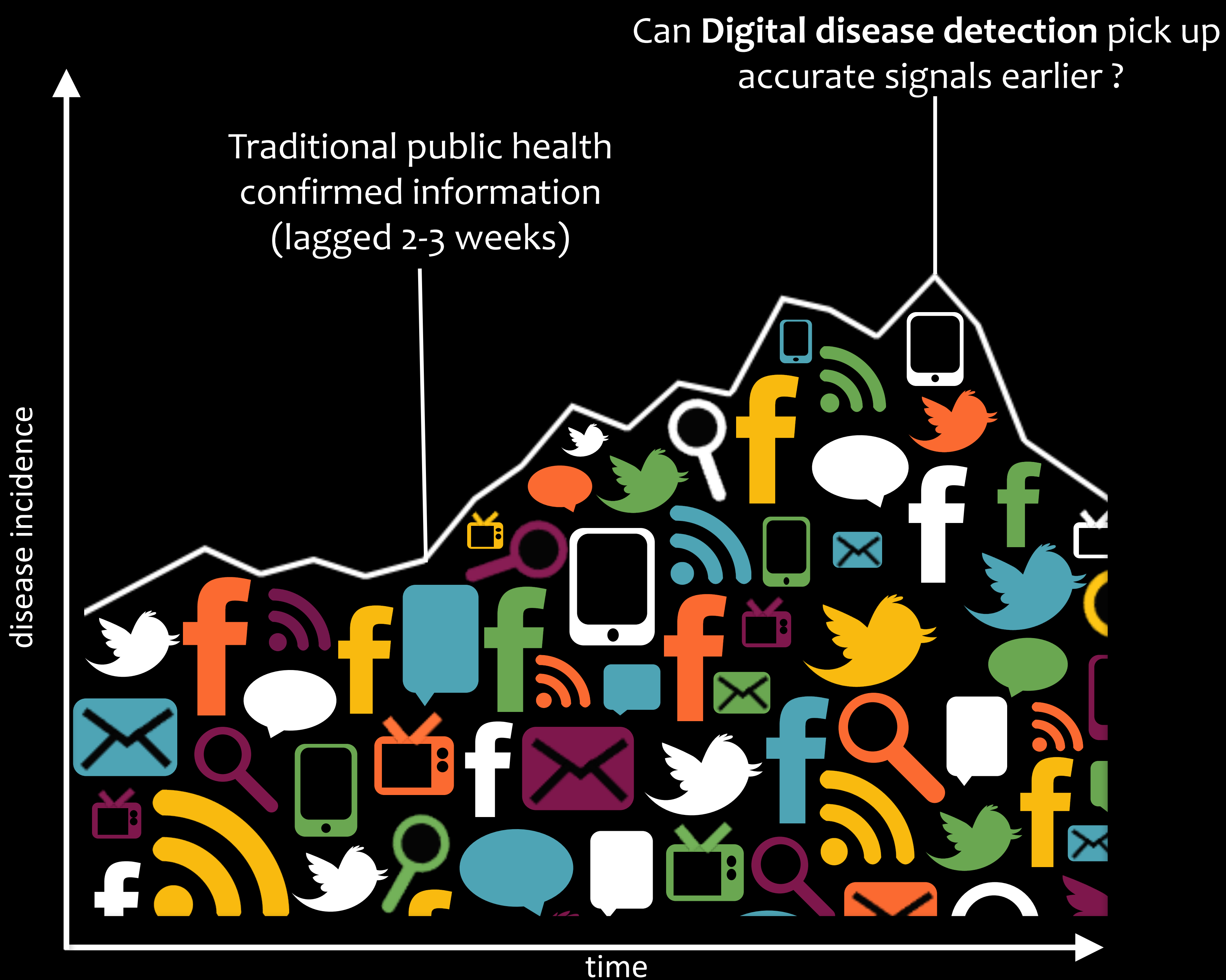
30+ petabytes of user-generated data stored, accessed, and analyzed

Electronic Health Records (**EHR**)

Over 1 **billion** Google searches a day

~2 billion smartphones world wide

230 million tweets every day

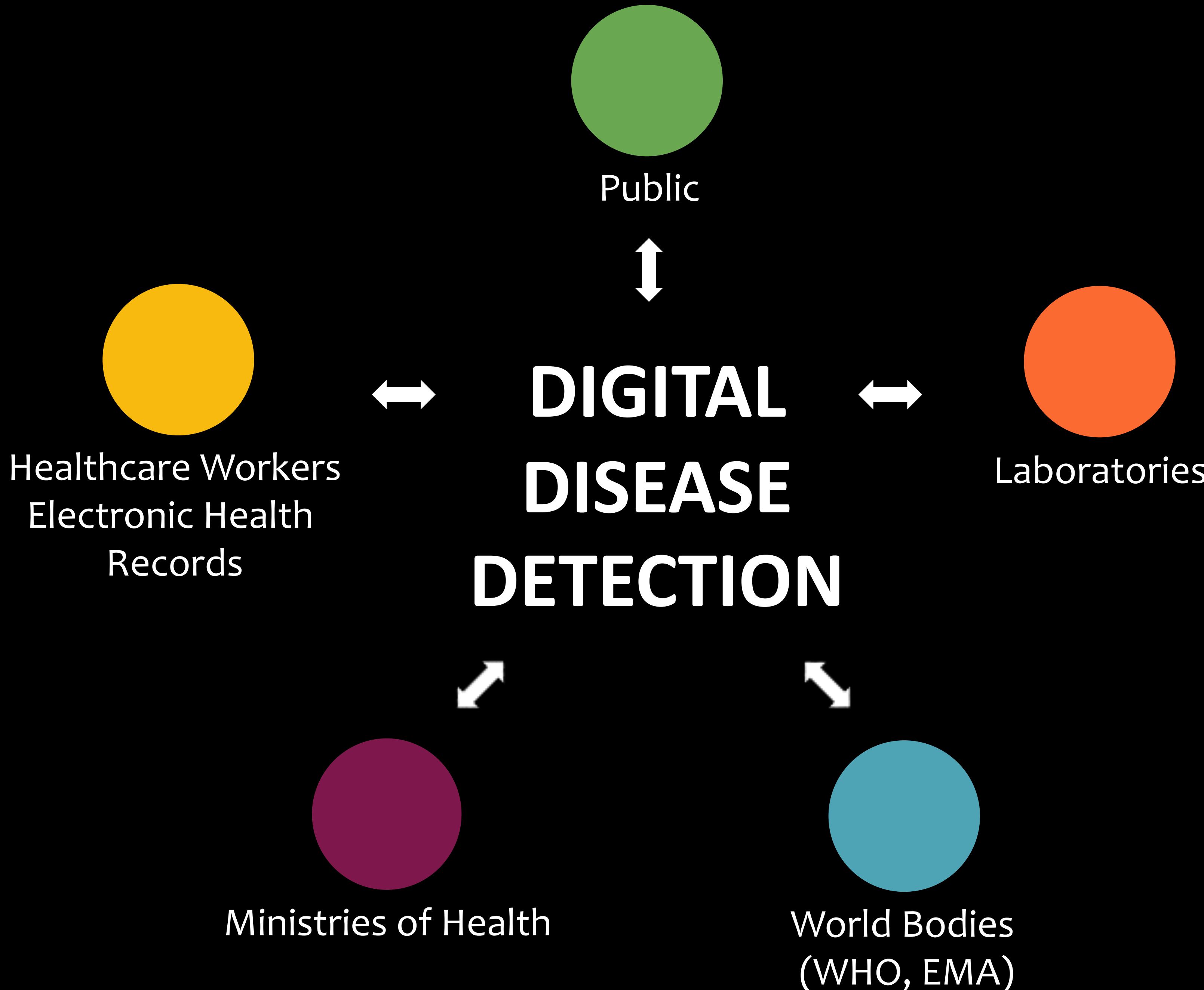


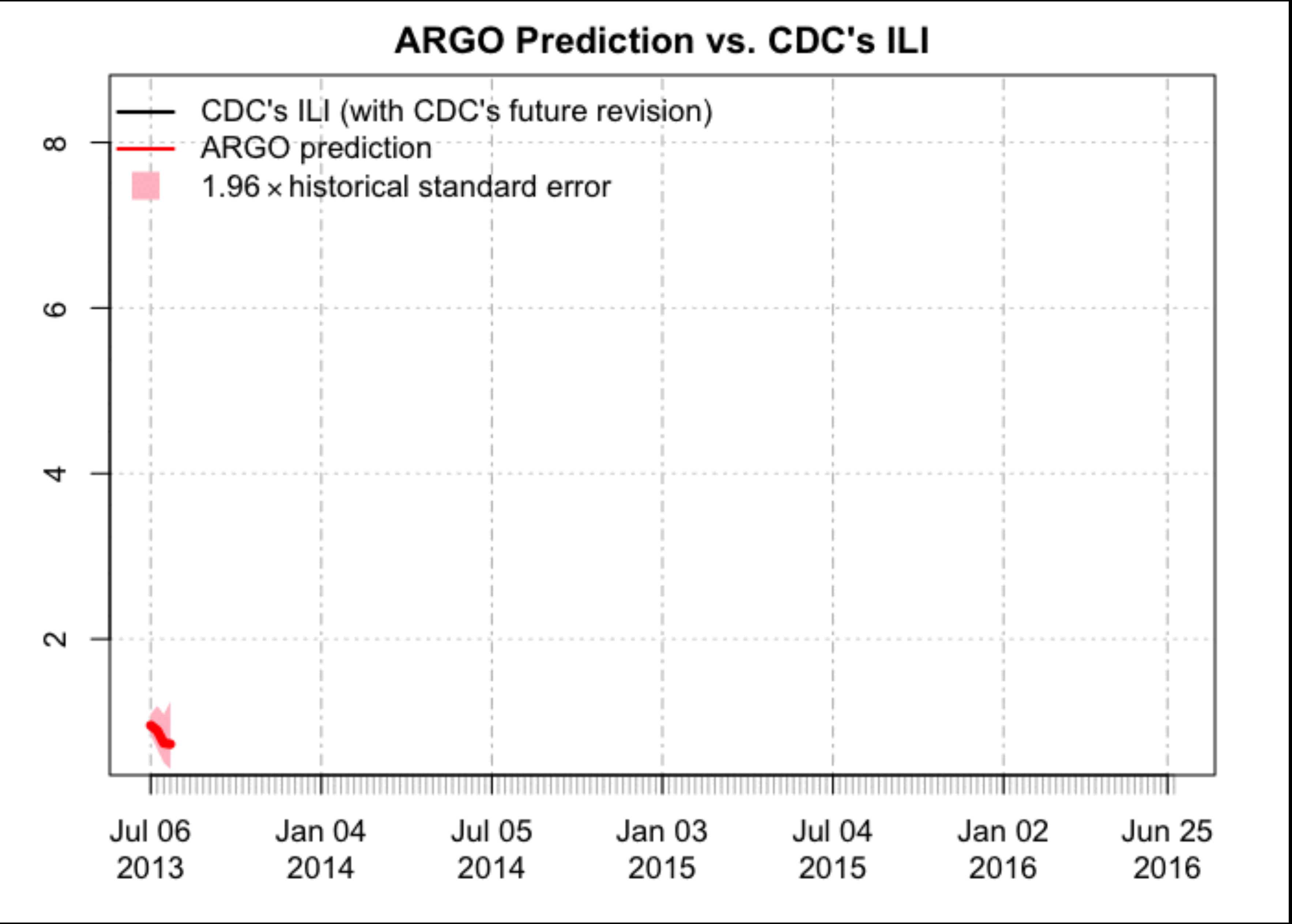
Can **Digital disease detection** pick up accurate signals earlier ?
Traditional public health confirmed information (lagged 2-3 weeks)
disease incidence
time

1. Public
2. Healthcare Workers, Electronic Health Records
3. Laboratories
4. Ministries of Health
5. World Bodies (WHO, EMA)

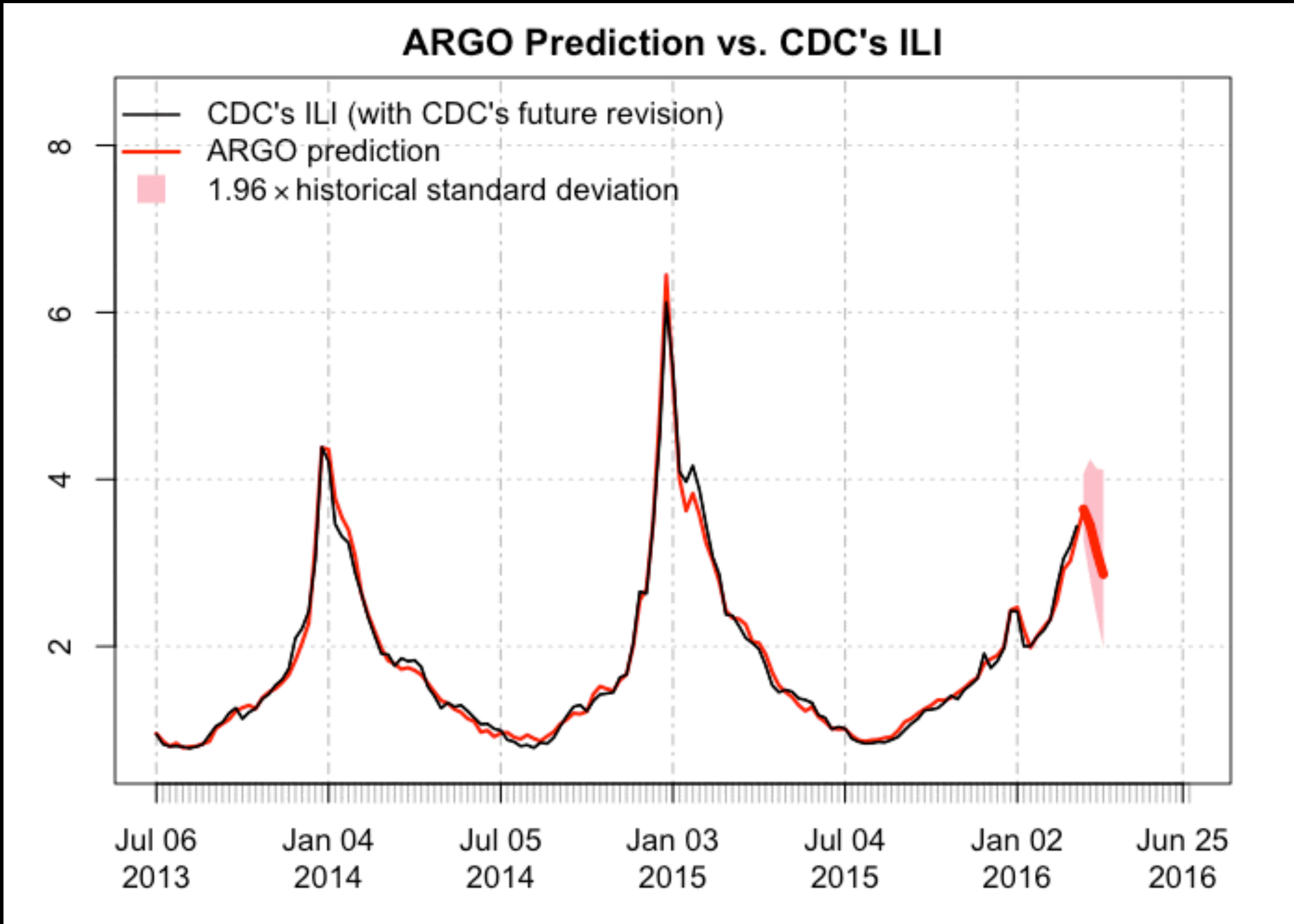
TRADITIONAL DISEASE REPORTING

1. Public
2. Healthcare Workers, Electronic Health Records
3. Laboratories
4. Ministries of Health
5. World Bodies (WHO, EMA)


Public
Healthcare Workers
Electronic Health
Records
DIGITAL
DISEASE
DETECTION
Laboratories
Ministries of Health
World Bodies
(WHO, EMA)

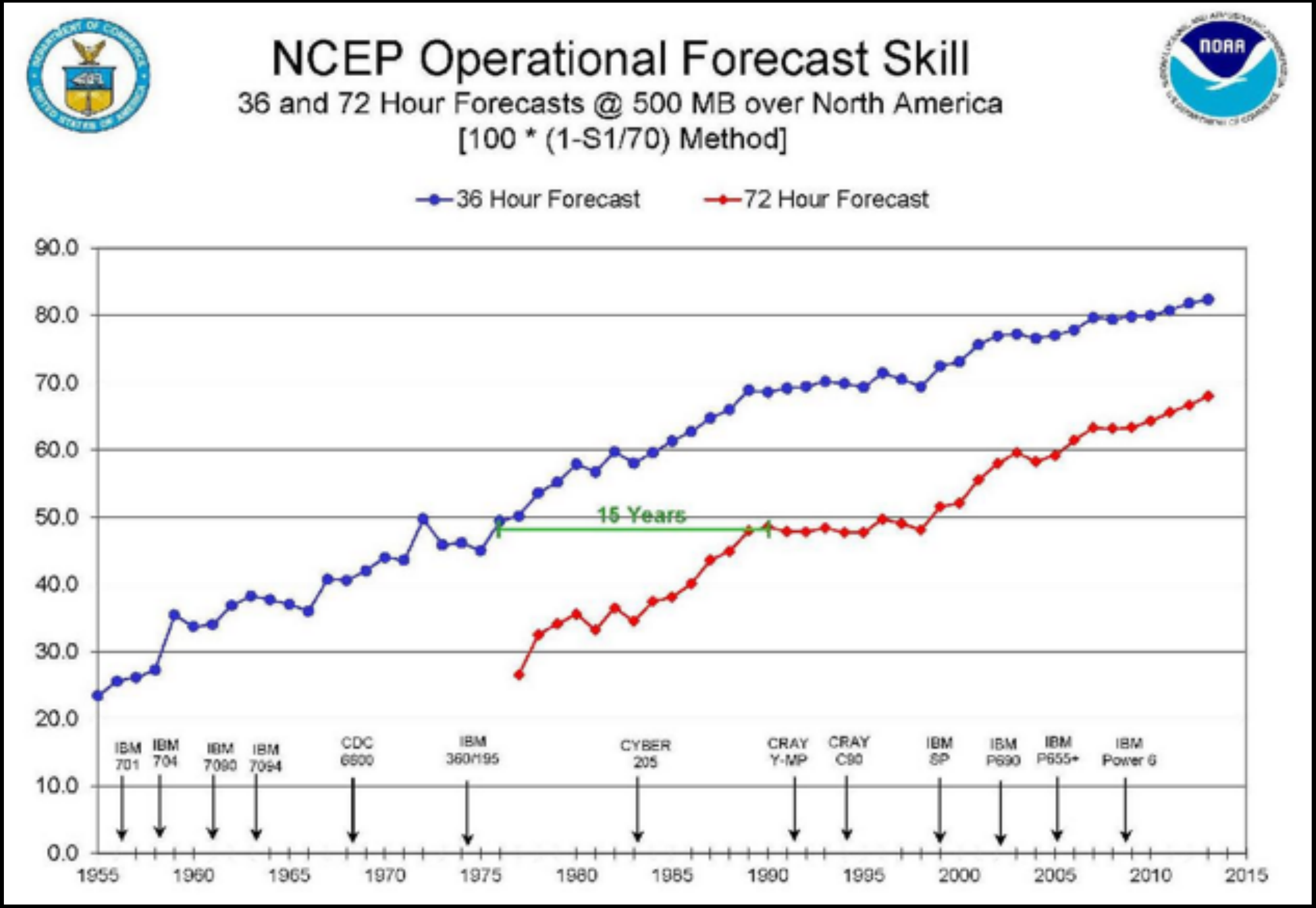


Video produced by Shihao Yang



ARGO Prediction vs. CDC's ILI
CDC's ILI (with CDC's future revision)
ARGO prediction
1.96 × historical standard deviation
8
6
4
2
Jul 06 2013
Jan 04 2014
Jul 05 2014
Jan 03 2015
Jul 04 2015
Jan 02 2016
Jun 25 2016

Real-time tracking vs predictions of disease incidence/risk
Similarities and differences with weather prediction

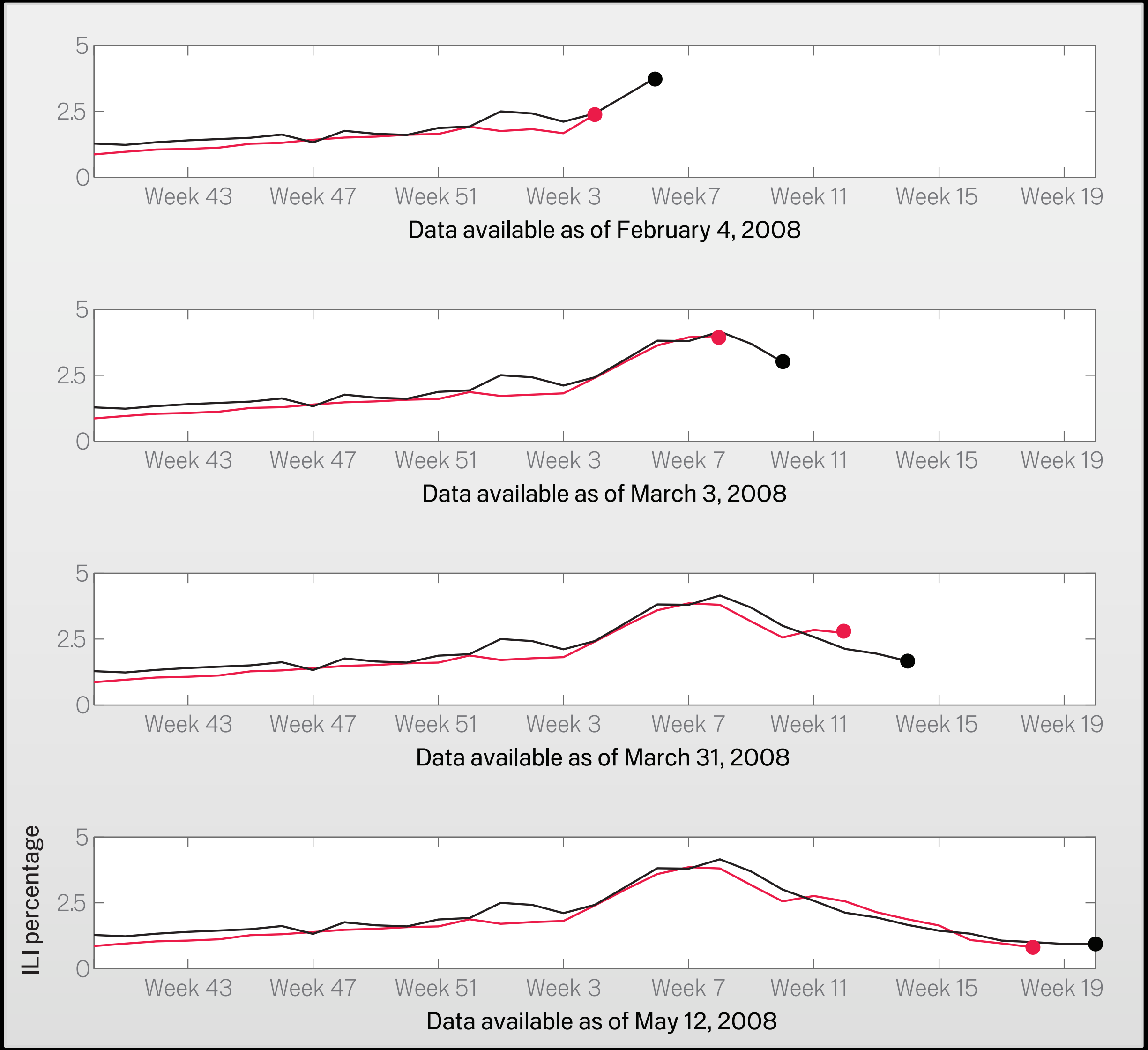


The promise of big data in public health

GOOGLE FLU TRENDS

google.org Flu Trends

Epidemiological information available 2-3 weeks ahead of traditional clinical tracking systems



nature International weekly journal of science

Letter

Nature **457**, 1012-1014 (19 February 2009) | doi:10.1038/nature07634; Received 14 August 2008; Accepted 13 November 2008; Published online 19 November 2008; Corrected 19 February 2009

Detecting influenza epidemics using search engine query data

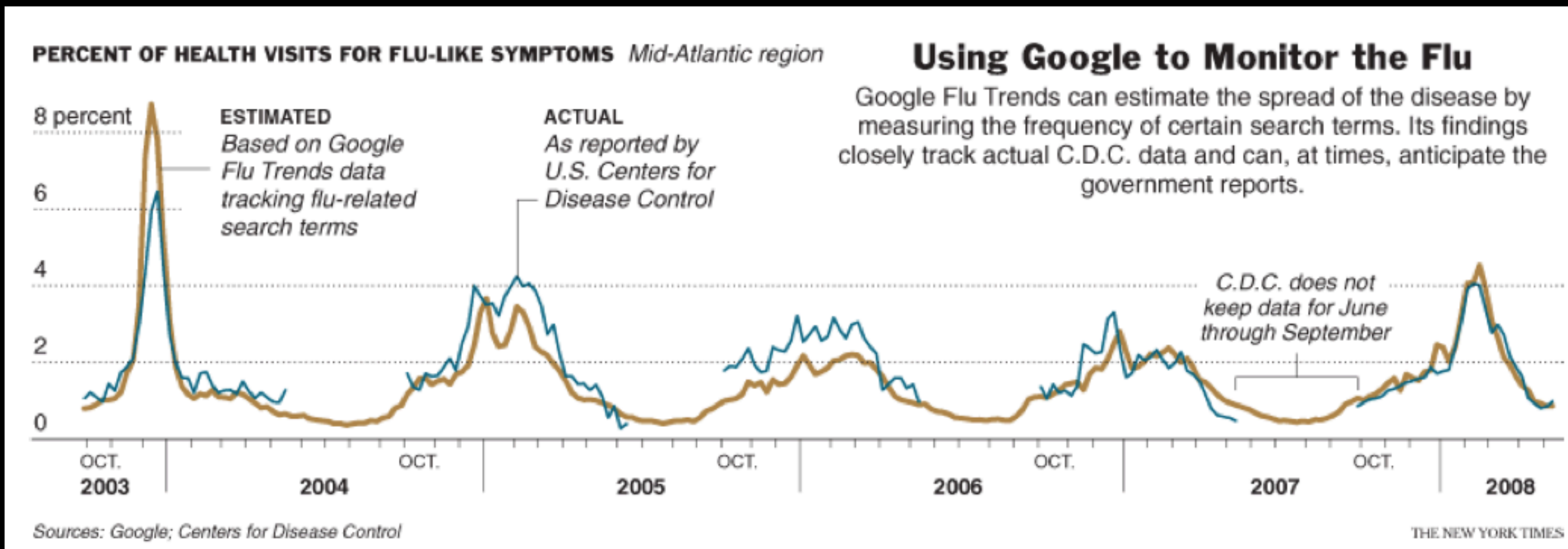
Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

1. Google Inc., 1600 Amphitheatre Parkway, Mountain View, California 94043, USA
2. Centers for Disease Control and Prevention, 1600 Clifton Road, NE, Atlanta, Georgia 30333, USA

Correspondence to: Matthew H. Mohebbi[1] Correspondence and requests for materials should be addressed to J.G. or M.H.M. (Email: flutrends-support@google.com).

The New York Times

Very promising retrospective comparison!



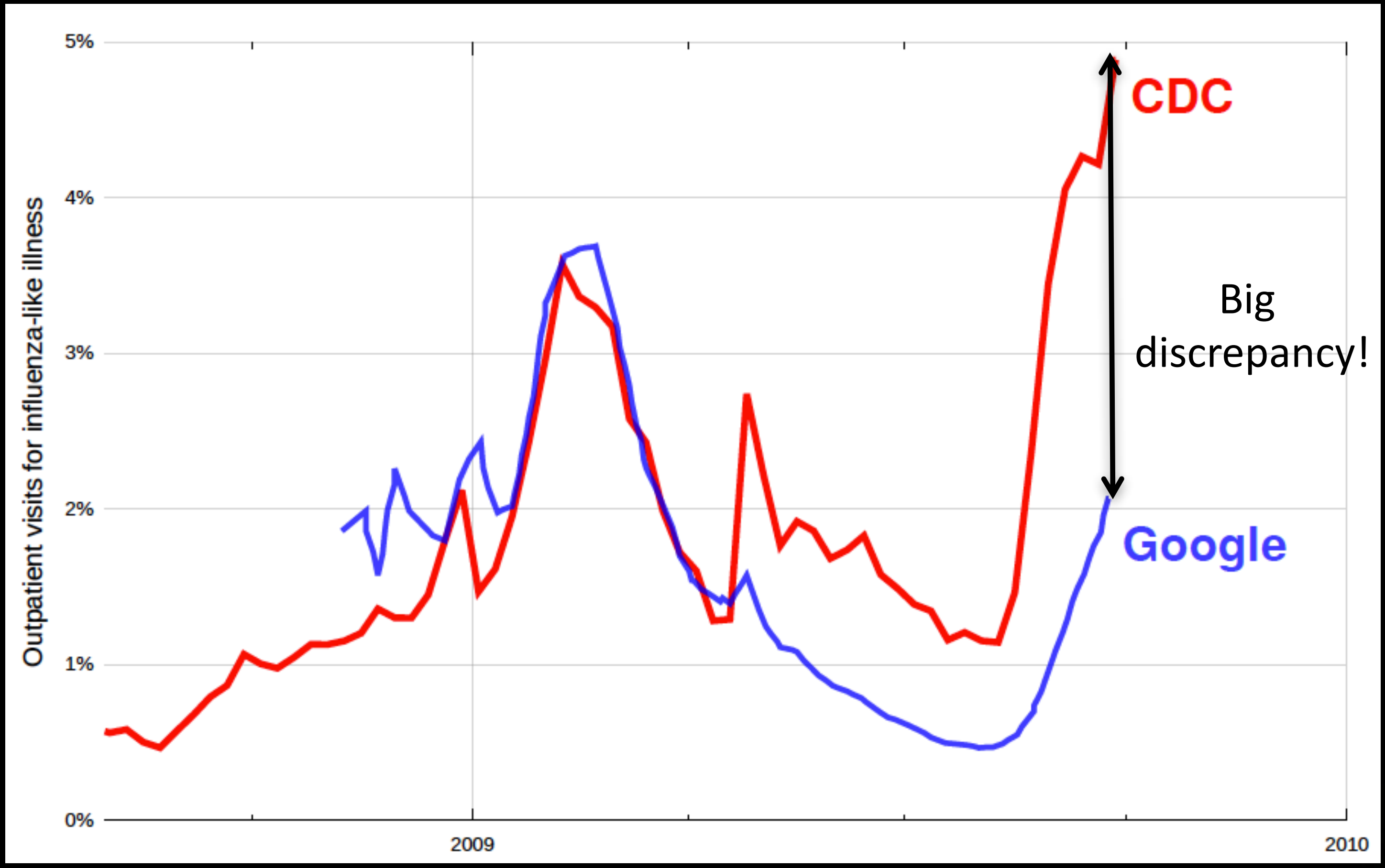
In April 2009, Dr. Brilliant said it epitomized the power of Google’s vaunted engineering prowess to make the world a better place, and he predicted that it would save untold numbers of lives.

Google Flu Trends

launched in November 2008

Real-time performance, first year…

Big errors seen during H1N1 pandemic (off-season)

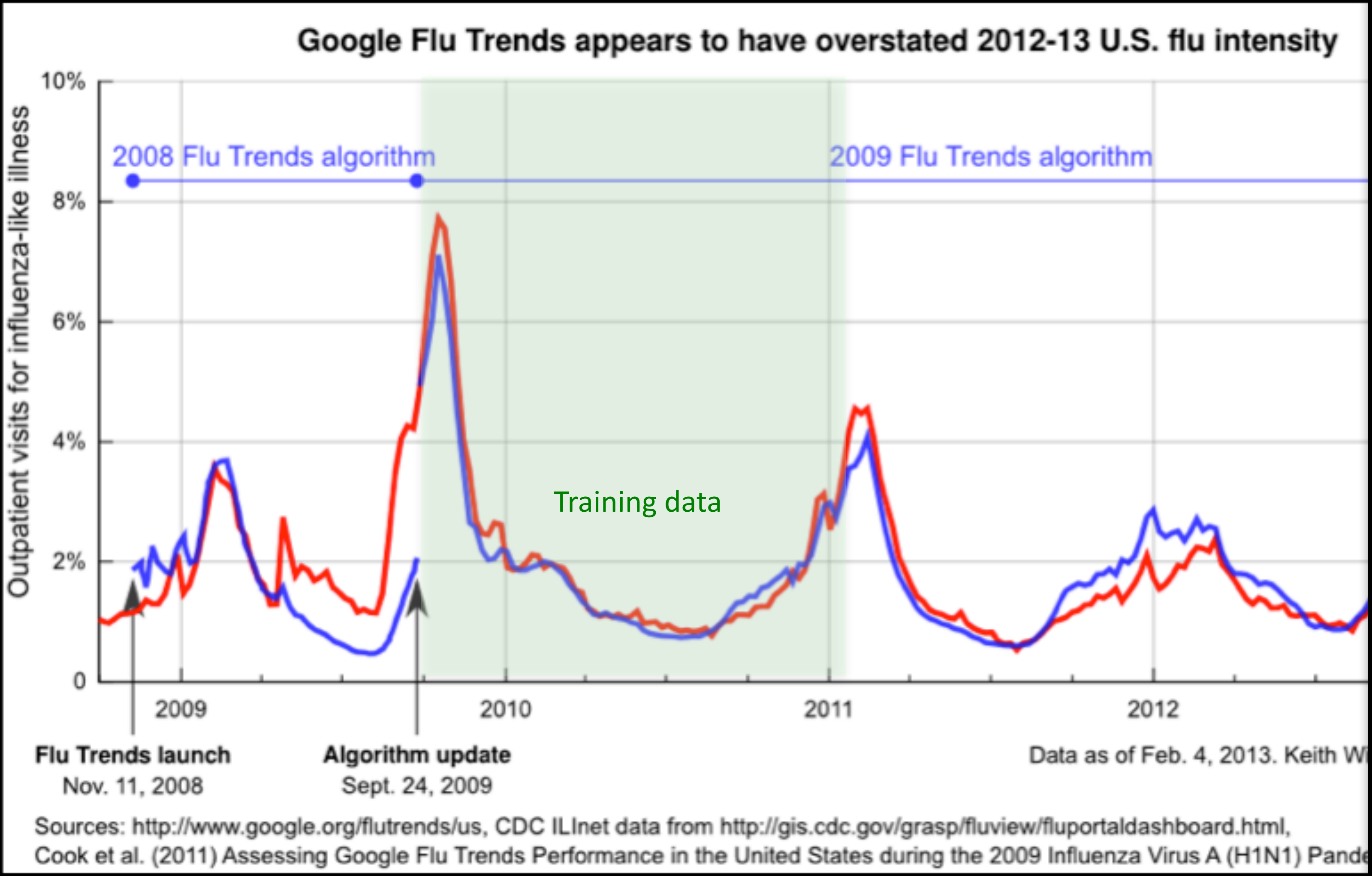


To some extent GFT was good at predicting seasons: fall-winter, not flu!

Plot obtained from:
http://blog.keithw.org/2013/02/q-how-accurate-is-google-flu-trends.html

What next?

need to remove (not useful) search terms

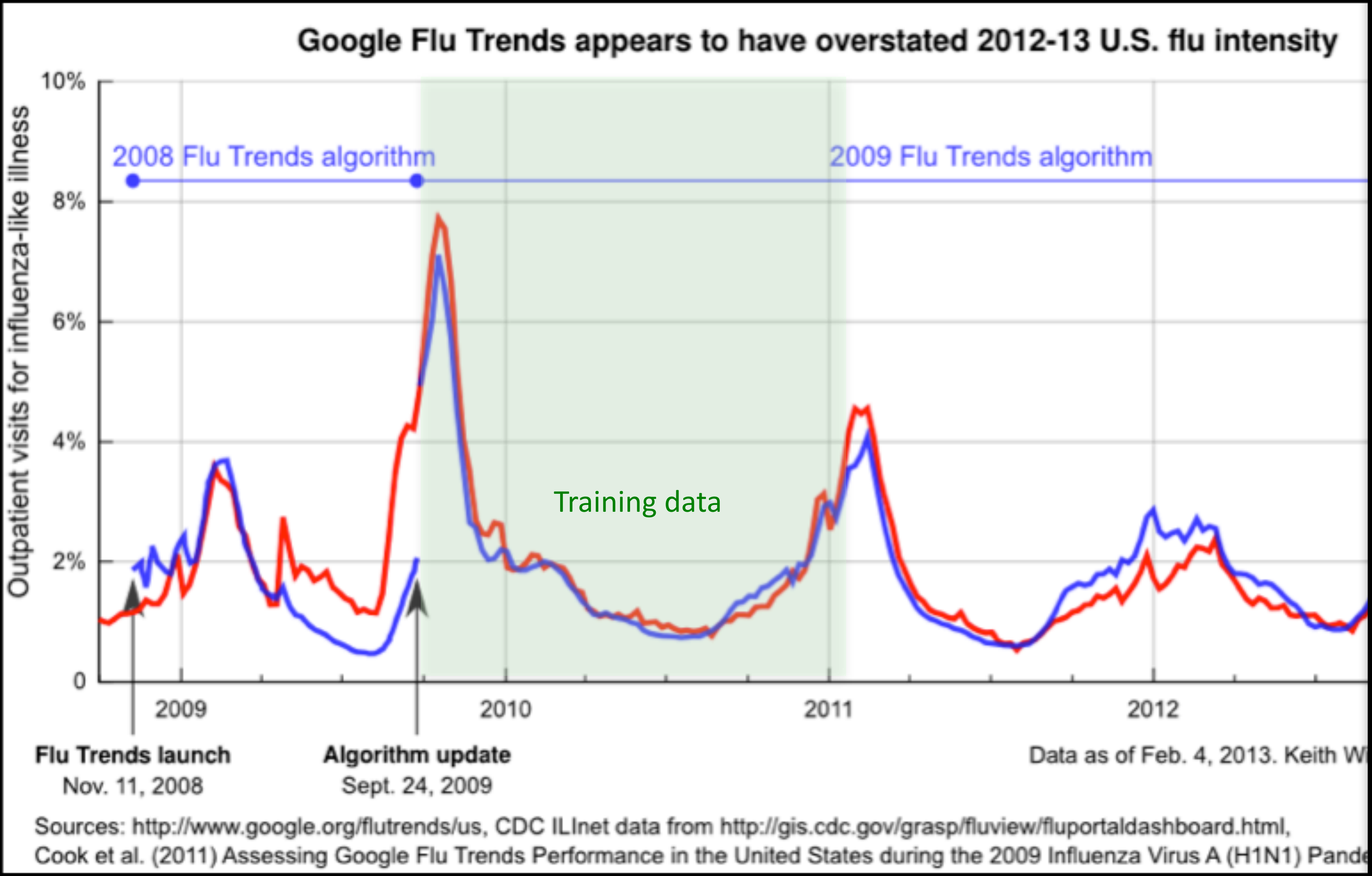


Fixes were reported in: Cook et al. (2011) Assessing Google flu trends performance in the U.S. during the 2009 influenza virus A (H1N1) pandemic. PLoS One

Plot obtained from: http://blog.keithw.org/2013/02/q-how-accurate-is-google-flu-trends.html

What next?

need to remove (not useful) search terms

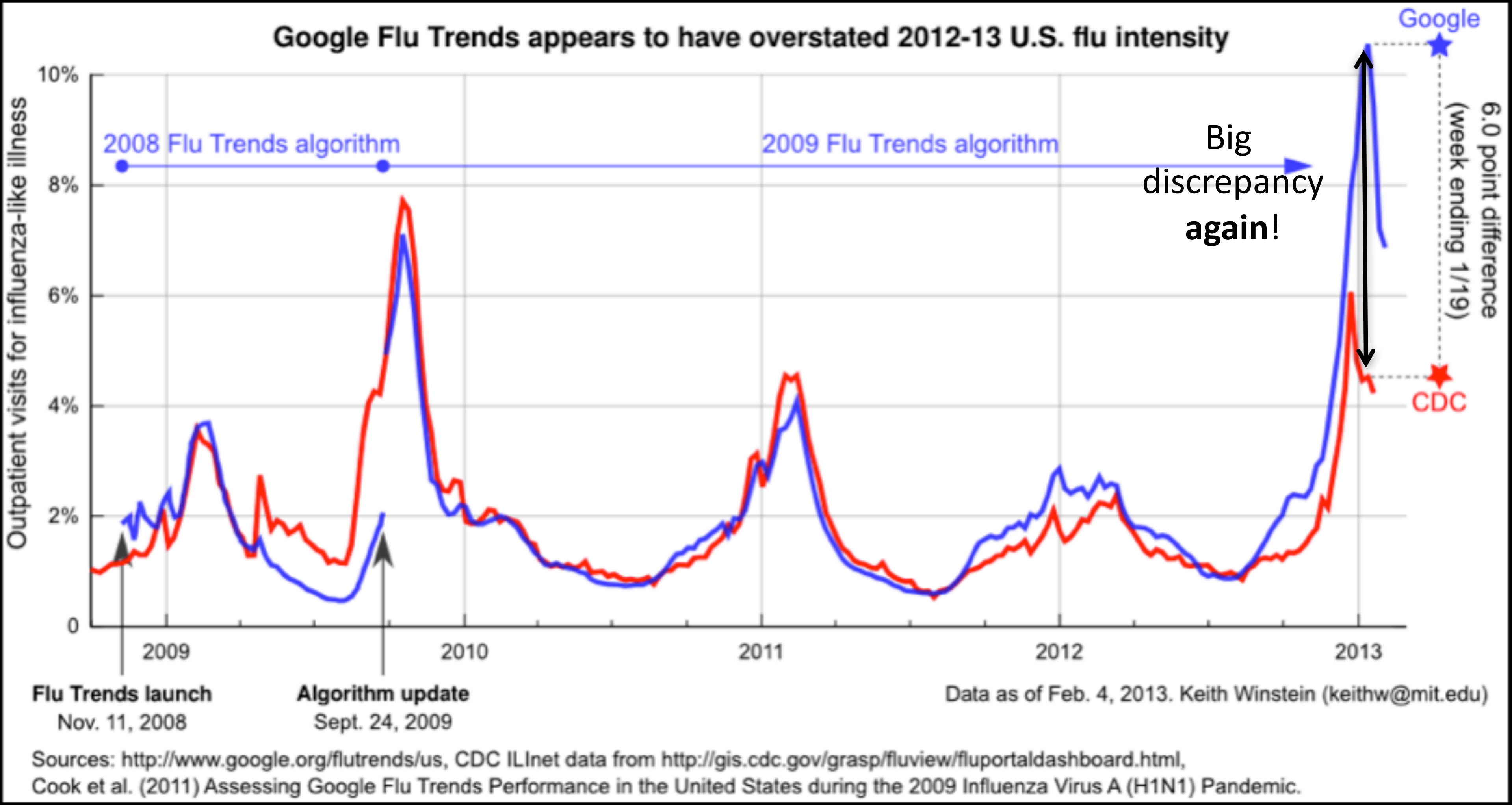


promising performance

Fixes were reported in: Cook et al. (2011) Assessing Google flu trends performance in the U.S. during the 2009 influenza virus A (H1N1) pandemic. PLoS One

Plot obtained from: http://blog.keithw.org/2013/02/q-how-accurate-is-google-flu-trends.html

What next? need to remove (not useful) terms. **Big discrepancies again!**



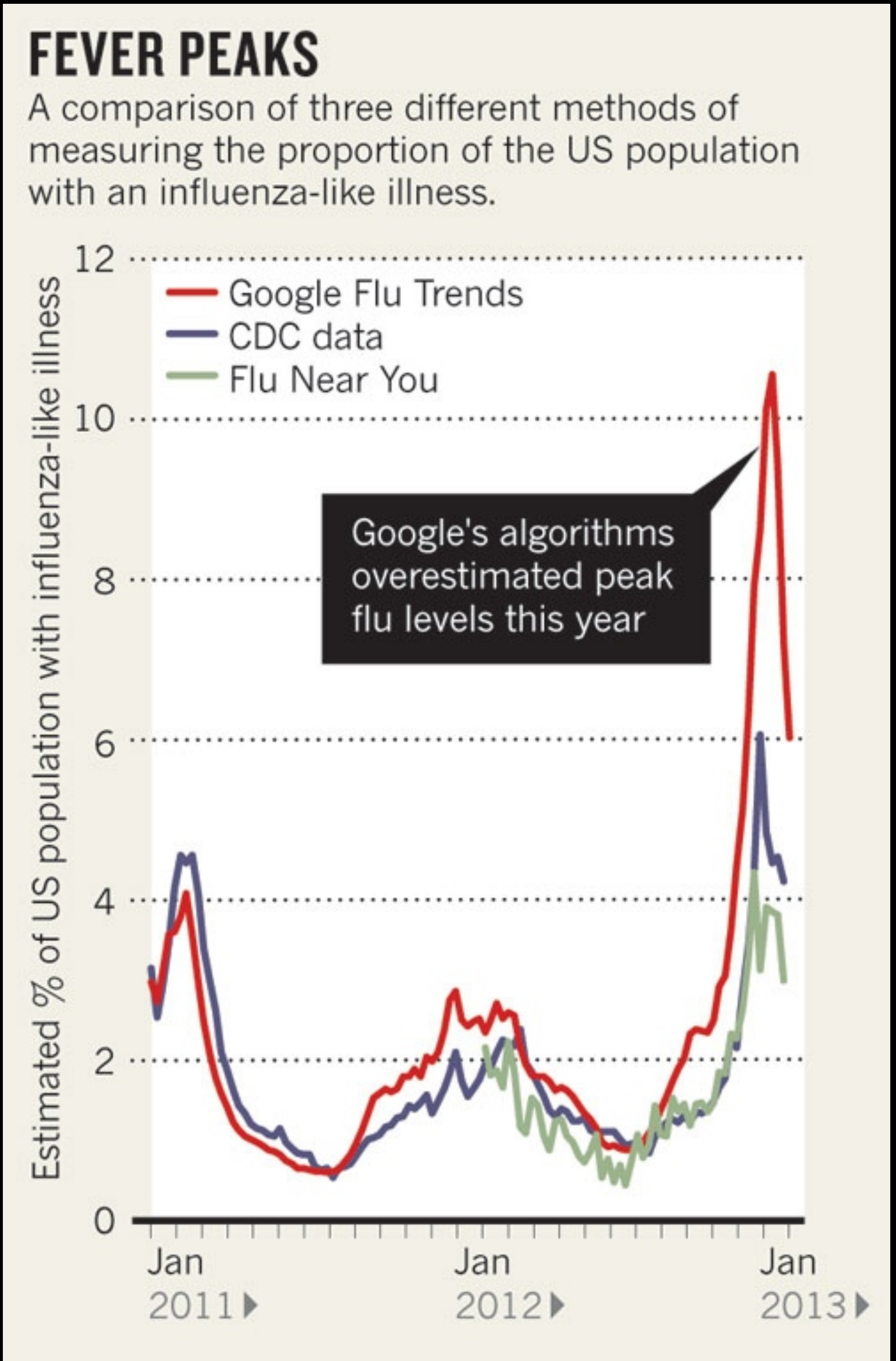
Fixes were reported in: Cook et al. (2011) Assessing Google flu trends performance in the U.S. during the 2009 influenza virus A (H1N1) pandemic. PLoS One

Plot obtained from: http://blog.keithw.org/2013/02/q-how-accurate-is-google-flu-trends.html

nature International weekly journal of science

When Google got flu wrong.

nature.com/news/when-google-got-flu-wrong.



Forbes

New Posts
+5 posts this hour

Most Popular
Forbes 400 Newcomers

Lists
The Forbes 400

63
Share

106
Tweet

61

Snowden And The Challenge Of Intelligence: The Practical Case Against The NSA's Big Data

12 comments, 7 called-out
+ Comment Now
+ Follow Comments

We should soon be able to keep track of most activities on the surface of the earth, day or night, in good weather or bad.

siliconANGLE
where computer science meets social science

{SILICON ANGLE}
CLOUD
MOBILE
SOCIAL
SERVICES
DEVOPS
RESEARCH

SiliconANGLE » Can Nate Silver's Data Culture Lead Us Out Of The NSA + Public Data Scare?

Can Nate Silver's Data Culture Lead Us Out of the NSA + Public Data Scare?

RYAN COX | SEPTEMBER 18TH

READ MORE

Lessons learned

Assumptions in Google Flu Trends:

1. Number of (influenza-like) ill people proportional to number of **total** searches of (Influenza-like illnesses) related terms

$$logit(P) = \beta_0 + \beta_1 \times logit(Q) + \varepsilon$$

where P is the percentage of ILI physician visits, Q is the ILI-related query fraction, β_0 is the intercept,

Assumptions in Google Flu Trends:

1. Number of (influenza-like) ill people proportional to number of **total** searches of (Influenza-like illnesses) related terms

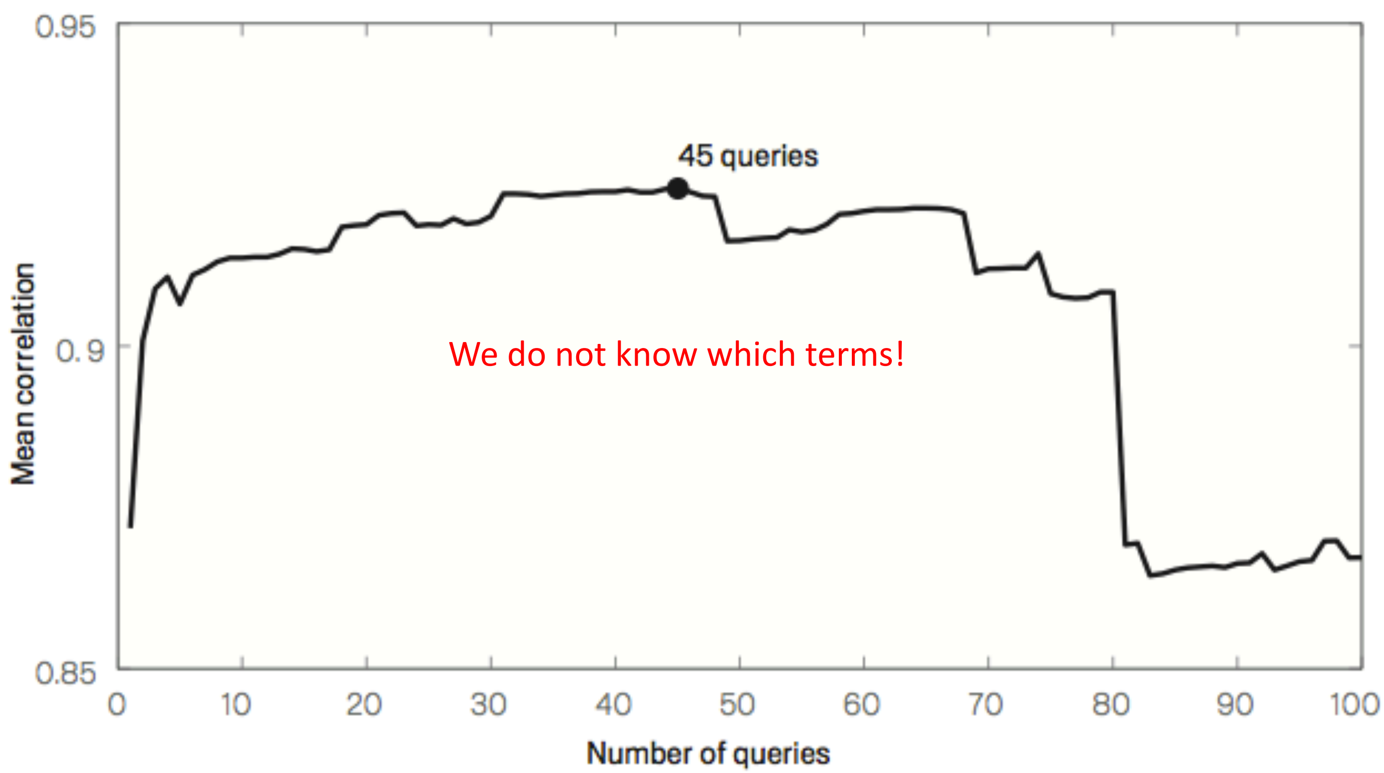


Figure 1: An evaluation of how many top-scoring queries to include in the ILI-related query fraction. Maximal performance at estimating out-of-sample points during cross-validation was obtained by summing the top 45 search queries. A steep drop in model performance occurs after adding query 81, which is “oscar nominations”.

Assumptions in Google Flu Trends:

2. Relationship between search volume and proportion of (influenza like) ill people is **static** (during a given year).

Assumptions in Google Flu Trends:

2. Relationship between search volume and proportion of (influenza like) ill people is **static** (during a given year).

Consequences: Model needed constant supervision by human experts

a. **Human experts** needed to **assess** relevance of individual search terms,

b. **Human Experts** needed to **recalculate** relationship between total number of searches and ill people, and

c. It is bound to **deliver poor predictions** at some point in the near future!

We proposed an alternative method and tested it using low quality input from Google Correlate in January 2013.

(with D. Wendong Zhang)

New model:

1. Each search term may contribute to prediction of ILI rate separately **(multi-variate approach)**

2. Relationship between search volume for each individual term and proportion of ill people is **dynamic** and should be found using supervised machine learning optimization techniques.

$$\boldsymbol{\beta}^{lasso} = \arg\min_{\beta} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{M} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{M} |\beta_j| \right\}$$

Every week the multiplicative coefficients (β's) would be automatically updated by expanding the training set (labeled data) as new information from the CDC became available.

Top correlated terms to CDC-reported data from 1/2004- 3/2009 (using Google Correlate)

1	influenza type a
2	bronchitis
3	influenza a
4	symptoms of pneumonia
5	flu incubation
6	influenza incubation
7	flu contagious
8	influenza contagious
9	flu incubation period
10	tussionex
11	benzonatate
12	influenza symptoms
13	a influenza
14	sinus
15	pneumonia
16	flu fever
17	flu duration
18	taste of chaos
19	bronchitis symptoms
20	symptoms of bronchitis
21	how long does the flu last
22	symptoms of the flu
23	taste of chaos tour
24	influenza incubation period
25	sinus infection
26	flu recovery
27	chaos tour
28	type a influenza
29	flu symptoms
30	tessalon
31	type a flu
32	treat the flu
33	treating the flu
34	how to treat the flu
35	is the flu contagious
36	flu in children
37	fever flu
38	take action tour
39	flu remedies
40	flu report
41	nasal congestion
42	fever reducer
43	sinus infections
44	rhode island wrestling
45	symptoms of influenza
46	castaway bay
47	coral by the sea
48	cold or flu
49	respiratory infection
50	take action
51	respiratory flu
52	soweto gospel
53	soweto gospel choir
54	illinois wrestling
55	how long is the flu contagious
56	cold symptoms
57	the taste of chaos
58	is bronchitis
59	upper respiratory
60	afrin
61	painful cough
62	laprepsoccer
63	upper respiratory infection
64	amoxicillin
65	ski harness
66	robitussin dm
67	treating flu
68	fever in adults
69	decongestant
70	normal body
71	low body temperature
72	a fever
73	influenza a symptoms
74	dangerous fever
75	is flu contagious
76	lauderdale florida
77	hotel fort lauderdale
78	webmail shaw ca
79	high fever
80	robitussin ac
81	bronchitis contagious
82	indoor driving
83	tussionex pennkinetic
84	wrestling report
85	walking pneumonia
86	days inn miami
87	body temperature
88	phlegm
89	flu relief
90	mt sunapee
91	harlem globe
92	levaquin
93	strep throat
94	coughing
95	whistler snow
96	fever temperature
97	sales tax credit
98	glitches
99	pennkinetic
100	histinex

AMERICAN JOURNAL OF
Preventive Medicine
A Journal of the American College of Preventive Medicine and Association for Prevention Teaching and Research

What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends?

Mauricio Santillana, PhD, MS, D. Wendong Zhang, MA, Benjamin M. Althouse, PhD, ScM, John W. Ayers, PhD, MA

 Am J Prev Med 2014;47(3):341–347 **341**

First week after being published online, it became the second most read paper in journal's history! (After a paper published in 1998)

AMERICAN JOURNAL OF
Preventive Medicine
A Journal of the American College of Preventive Medicine and Association for Prevention Teaching and Research

Articles in Press | **Most Read** | Most Cited

Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study

Vincent J Felitti, Robert F Anda, Dale Nordenberg, David F Williamson, Alison M Spitz, Valerie Edwards, Mary P Koss, James S Marks
Vol. 14, Issue 4
Published in issue: May, 1998

Abstract | Full-Text HTML | PDF

What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends?

Mauricio Santillana, D. Wendong Zhang, Benjamin M. Althouse, John W. Ayers
Vol. 47, Issue 3
Published online: July 1, 2014

Abstract | Full-Text HTML | PDF

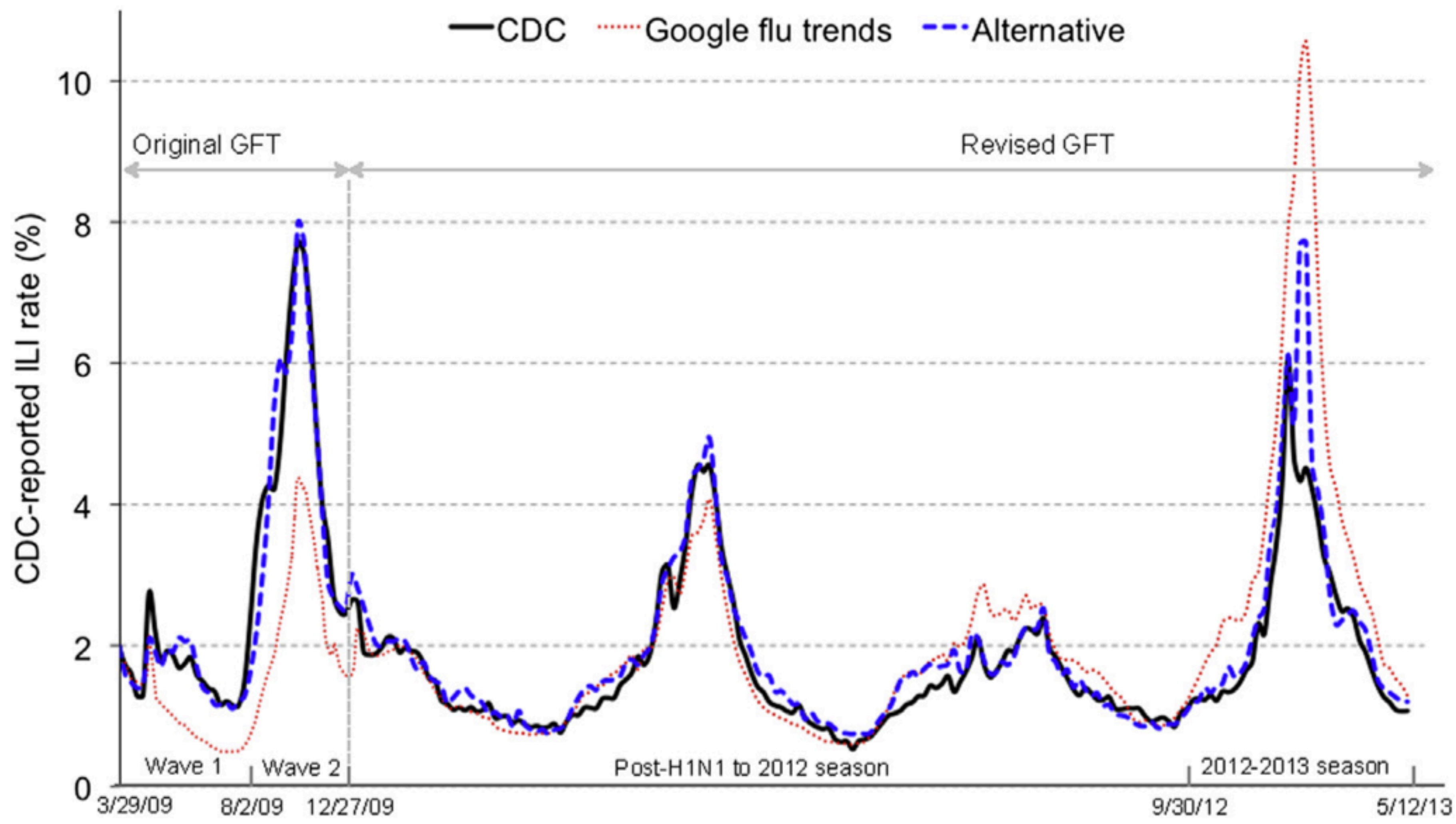
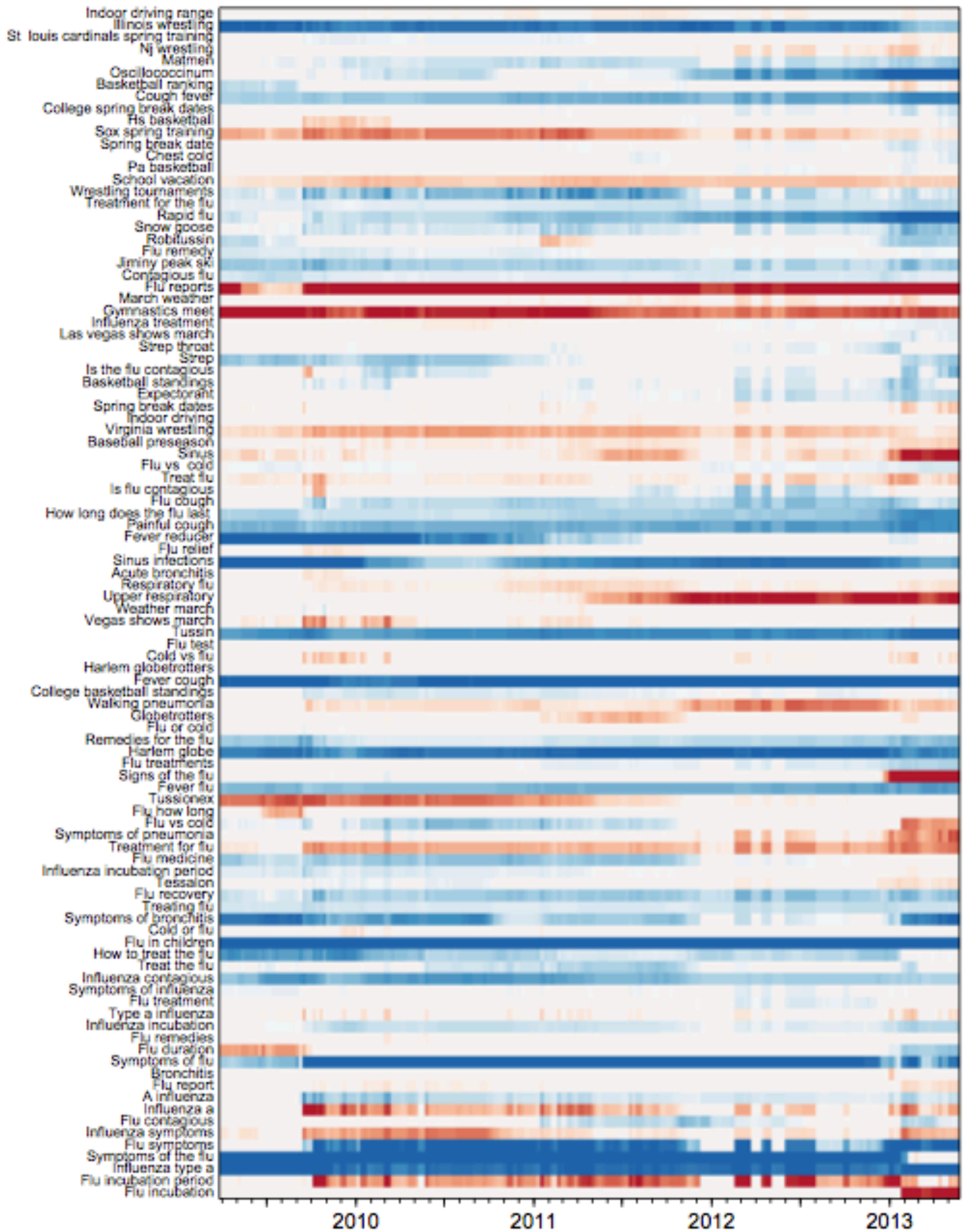
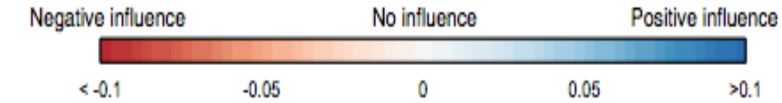


Figure 1. The alternative model outperforms Google Flu Trends

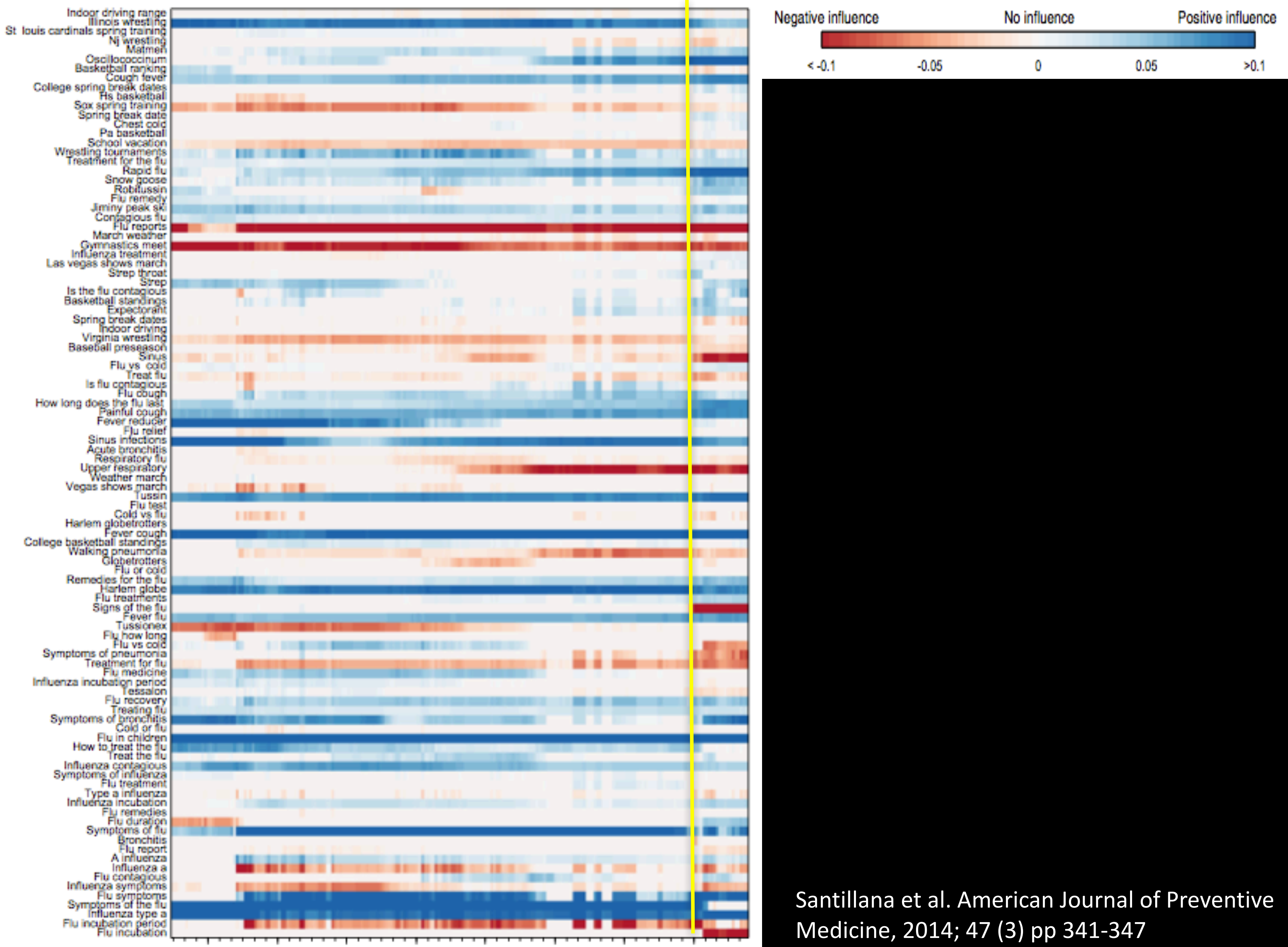
$$\mathrm{logit}[I(t)] = \sum_{i=1}^{n} a_i(t)\mathrm{logit}[Q_i(t))] + \mathrm{e},$$

Santillana et al. American Journal of Preventive Medicine, 2014; 47 (3) pp 341-347





Santillana et al. American Journal of Preventive Medicine, 2014; 47 (3) pp 341-347



Santillana et al. American Journal of Preventive Medicine, 2014; 47 (3) pp 341-347

theguardian

News | US | World | Sports | Comment | Culture | Business | Money

News > Technology > Google

Google Flu Trends promises are overstated, researchers say

New study finds way to improve Google Flu Trends accuracy threefold - but says systems must be more open

Charles Arthur

Follow @charlesarthur   Follow @guardiantech

theguardian.com, Friday 4 July 2014 11.44 EDT

theguardian

News | US | World | Sports | Comment | Culture | Business | Money

News > Technology > Google

Google Flu Trends promises are overstated, researchers say

New study finds way to improve Google Flu Trends accuracy threefold - but says systems must be more open

HealthData Management

NEW

POLICY & REGULATION | EHR | HEALTH INFO EXCHANGE | REVENUE CYCLE & PAYMENTS | CHRONIC CARE

Researchers Suggest Fixes to Google Flu Trends Analytics

A new study concludes that "revising the inner plumbing" of the Google Flu Trends disease surveillance system can improve the accuracy of forecasts for the severity of a flu season.

theguardian

News | US | World | Sports | Com

News › Technology › Google

Google Flu Trends p… overstated, research…

New study finds way to improve … threefold - but says systems mu…

ScienceDaily®

Your source for the latest research news

Featured Research *from universities, journals, and other organizatio…*

Finding real value in big data for public health

Date: July 2, 2014

Source: San Diego State University

Summary: Media reports of public health breakthroughs from big data have been largely oversold, according to a new study. But don't throw away that data just yet. The authors maintain that the promise of big data can be fulfilled by tweaking existing methodological and reporting standards. In the study, the research team demonstrate this by revising the inner plumbing of the Google Flu Trends (GFT) digital disease surveillance system, which was heavily criticized last year (see here and here) after producing erroneous forecasts.

Share This

- Email to a friend
- Facebook
- Twitter
- LinkedIn
- Google+
- Print this page

HealthData Management

POLICY & REGULATION | EHR | HEA… EXC…

Researchers Sug… Flu Trends Analy…

A new study concludes that "revising the in… surveillance system can improve the accur…

Related Topics

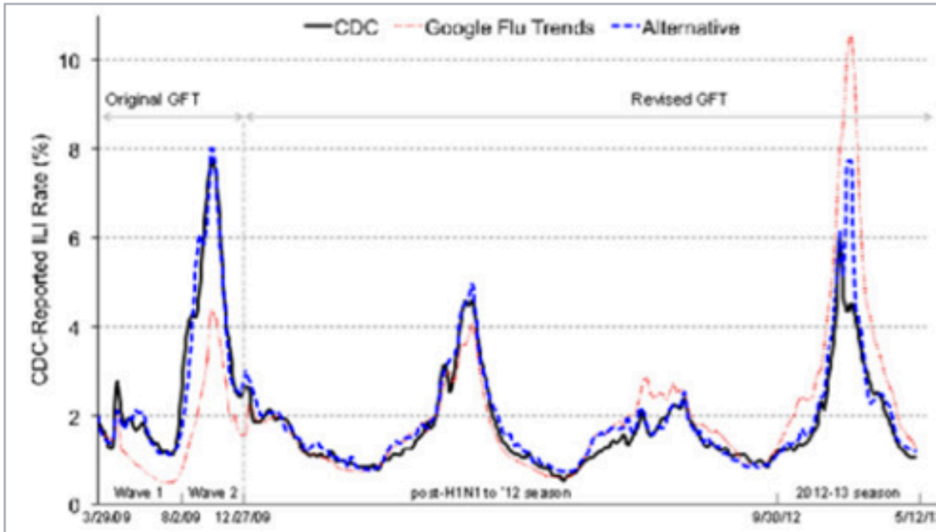
Health & Medicine
- Health Policy
- Public Health Education

Computers & Math
- Computers and Internet
- Computer Modeling

Science & Society
- Public Health
- Surveillance

Related Articles

- Public health
- Data mining



A graph depicting Google Flu Trends.

theguardian

News | US | World | Sports | Con

News › Technology › Google

Google Fl...

oversta...

New study ...

threefold - ...

Hea...

Ma...

POLICY & REGULATION

Resea...

Flu Tr...

A new study concludes that "revising the in...

surveillance system can improve the accur...

ScienceDaily®

Your source for the latest research news

Featured Research

from universities, journals, and other organizatio...

iHealthBeat

Reporting Technology's Impact on Health Care

HOME | INSIGHT | PERSPECTIVES | PICTURE OF HEALTH

NEWS ARCHIVE

SHARE EMAIL PRINT REPUBLISH

Study: Methodology Changes Improve Google Flu Trend Accuracy

Monday, July 7, 2014

RELATED TOPICS:

- Public Health

The accuracy of Google Flu Trends' disease surveillance system can be improved through simple changes in three different methodologies used by the system, according to a new study published in the *American Journal of Preventive Medicine*, *Health Data Management* reports (Goedert, *Health Data Management*, 7/7).

This

...mail to a friend

...acebook

...witter

...nkedIn

...oogle+

...int this page

A graph depicting Google Flu Trends.

Related Articles

- Public health
- Data mining

theguardian

ScienceDaily®

dataconomy

NEWS EVENTS OPINION START UPS INDUSTRY RESOURCES ABOUT JOBS

BI & ANALYTICS MACHINE LEARNING DATA SCIENCE DATABASE TECHNOLOGY DATA SECURITY UNDERSTANDING BIG DATA

Search

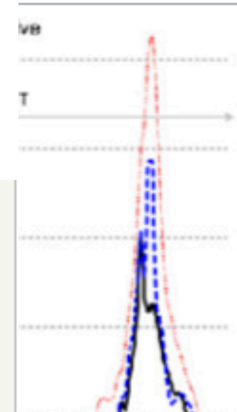
Big Data's Potential in Public Health: Revisiting Google Flu Trends

July 7, 2014 Written by: Dan Gray 1 Reply

nd other organizatio

This

nail to a friend
acebook
vitter
nkedIn
oogle+
int this page



Resea
Flu Tr

RELATED TOPICS:

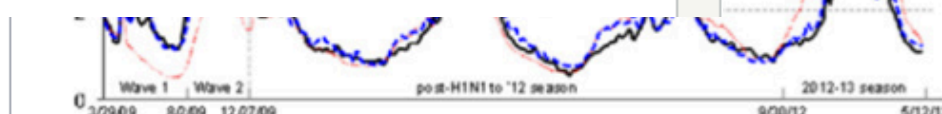
- Public Health

The accuracy of Google Flu Trends' disease surveillance system can be improved through simple changes in three different methodologies used by the system, according to a new study published in the *American Journal of Preventive Medicine*, *Health Data Management* reports (Goedert, *Health Data Management*, 7/7).

A new study concludes that "revising the in surveillance system can improve the accur

Related Articles

- Public health
- Data mining



A graph depicting Google Flu Trends.

Google Research Blog

The latest news from Research at Google

Google Flu Trends gets a brand new engine

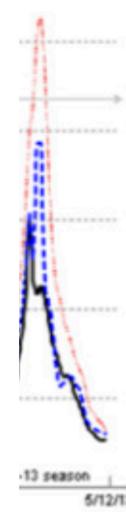
Posted: Friday, October 31, 2014

+1 222 | Tweet 161 | Like 104

Posted by Christian Stefansen, Senior Software Engineer

Each year the flu kills thousands of people and affects millions around the world. So it's important that public health officials and health professionals learn about outbreaks as quickly as possible. In 2008 we launched Google Flu Trends in the U.S., using aggregate web searches to indicate when and where influenza was striking in real time. These models nicely complement other survey systems—they're more fine-grained geographically, and they're typically more immediate, up to 1-2 weeks ahead of traditional methods such as the CDC's official reports. They can also be incredibly helpful for countries that don't have official flu tracking. Since launching, we've expanded Flu Trends to cover 29 countries, and launched Dengue Trends in 10 countries.

The original model performed surprisingly well despite its simplicity. It was retrained just once per year, and typically used only the 50 to 300 queries that produced the best estimates for prior seasons. We then left it to perform through the new season and evaluated it at the end. It didn't use the official CDC data for estimation during the season—only in the initial training.



Google Flu Trends heavily criticized in a paper published by Alex’s research team

POLICYFORUM

BIG DATA

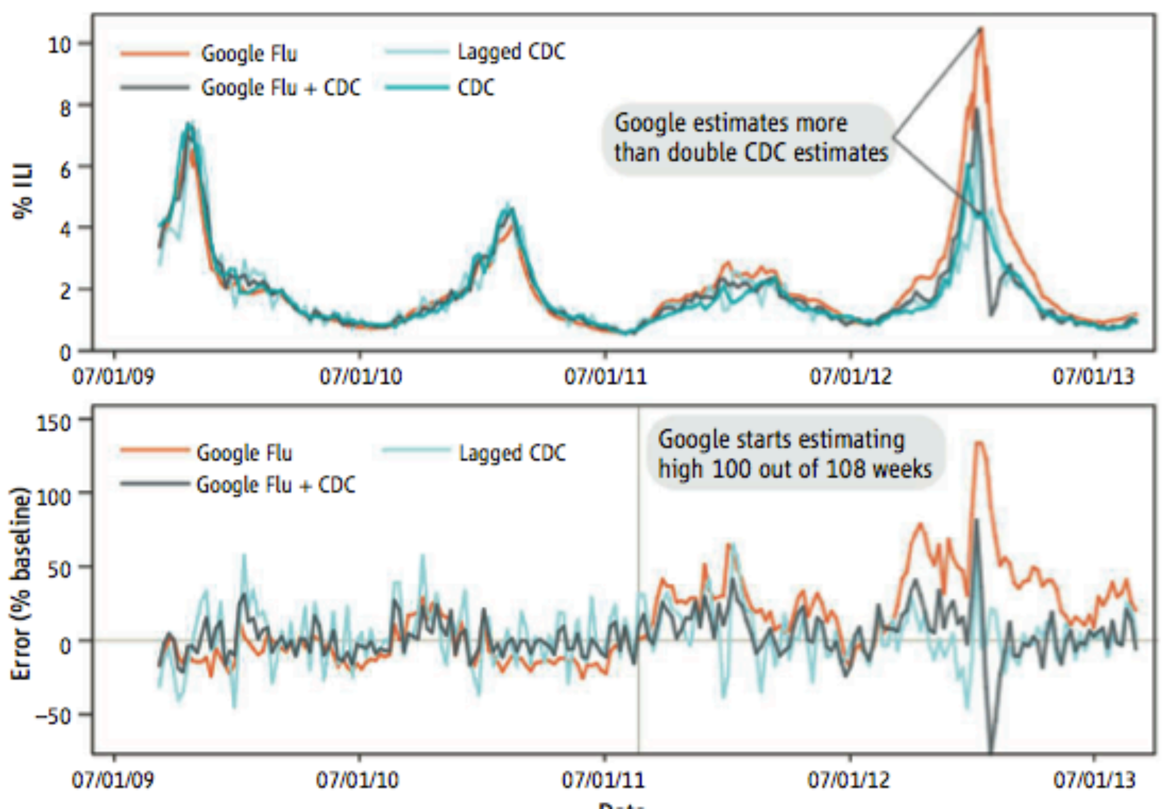
The Parable of Google Flu: Traps in Big Data Analysis

Large errors in flu prediction were largely avoidable, which offers lessons for the use of big data.

David Lazer,[1,2*] Ryan Kennedy,[1,3,4] Gary King,[3] Alessandro Vespignani[5,6,3]

www.sciencemag.org SCIENCE VOL 343 14 MARCH 2014

Published by AAAS



1. Lagged (CDC-based) models capable of outperforming GFT.

2. GFT + lagged CDC can outperform GFT (recalibrating importance of GFT)

3. Google search engine itself changed 86 times in June and July 2012 potentially leading to changes in Google search results (independent variable)

4. Feedbacks (recommended search terms depend on previous searches)

We recently established a new standard by Incorporating historical information (via autoregressive terms)

PNAS PNAS

Accurate estimation of influenza epidemics using Google search data via ARGO

Shihao Yang[a], Mauricio Santillana[b,c,1], and S. C. Kou[a,1]

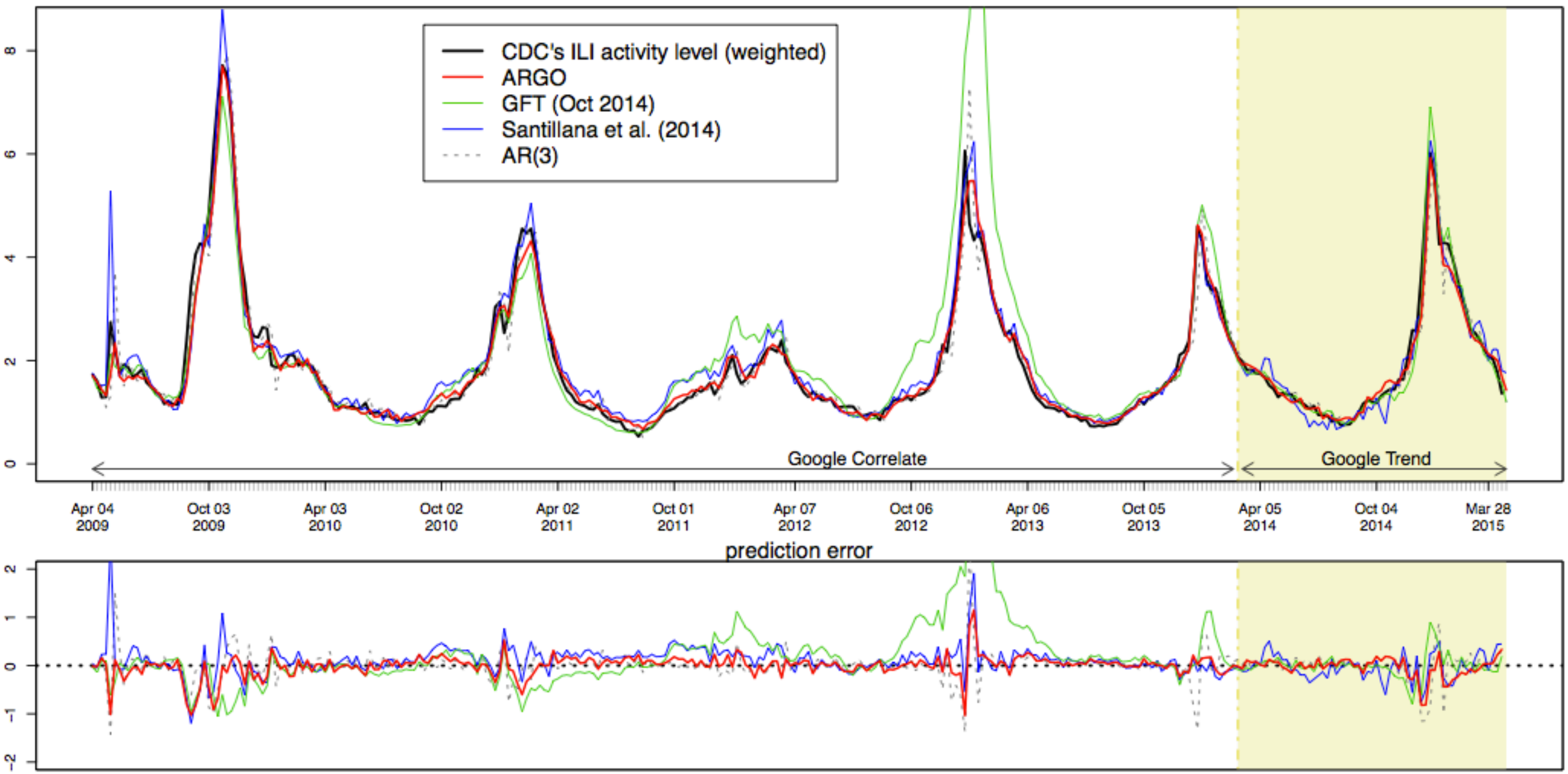
[a]Department of Statistics, Harvard University, Cambridge, MA 02138; [b]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; and [c]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA 02115

Edited by Wing Hung Wong, Stanford University, Stanford, CA, and approved September 30, 2015 (received for review August 6, 2015)

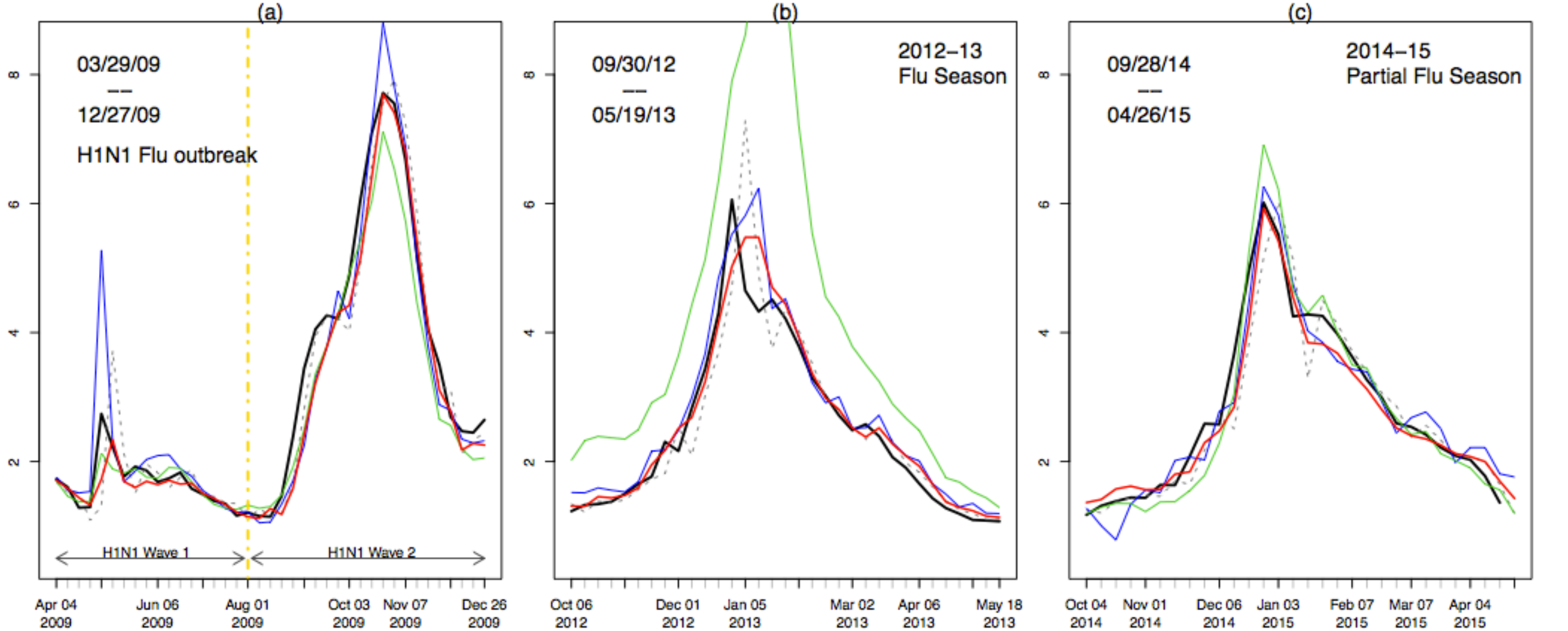
Accurate real-time tracking of influenza outbreaks helps public health officials make timely and meaningful decisions that could save lives. We propose an influenza tracking model, ARGO (AutoRegression with GOogle search data), that uses publicly available online search data. In addition to having a rigorous statistical foundation, ARGO outperforms all previously available Google-search–based tracking models, including the latest version of Google Flu Trends, even though it uses only low-quality search data as input from publicly available Google Trends and Google Correlate websites. ARGO not only incorporates the seasonality in influenza epidemics but also captures changes in people's online search behavior over time. ARGO is also flexible, self-correcting, robust, and scalable, making it a potentially powerful tool that can be used for real-time tracking of other social events at multiple temporal and spatial resolutions.

CDC's ILI reports have a delay of 1–3wk due to the time for processing and aggregating clinical information. This time lag is far from optimal for decision-making purposes. To alleviate this information gap, multiple methods combining climate, demographic, and epidemiological data with mathematical models have been proposed for real-time estimation of flu activity (18, 21–25). In recent years, methods that harness Internet-based information have also been proposed, such as Google (1), Yahoo (2), and Baidu (3) Internet searches, Twitter posts (4), Wikipedia article views (5), clinicians' queries (6), and crowdsourced self-reporting mobile apps such as Influenzanet (Europe) (26), Flutracking (Australia) (27), and Flu Near You (United States) (28). Among them, GFT has received the most attention and has inspired subsequent digital disease detection systems (3, 8,

PPLIED
HEMATICS

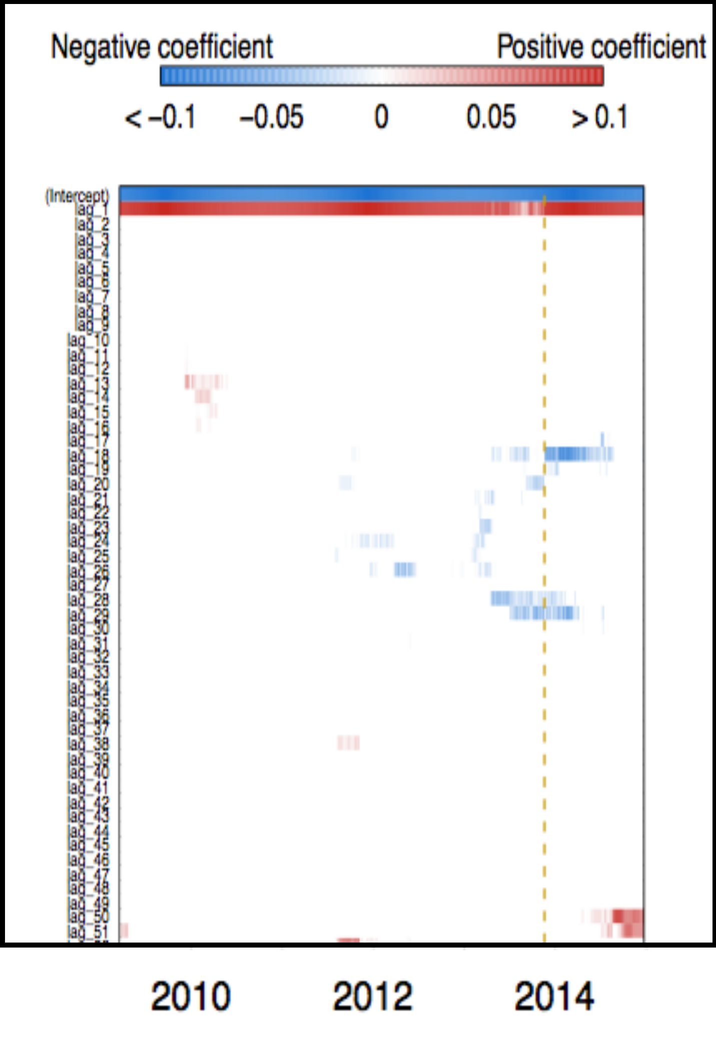


CDC's ILI activity level (weighted)
ARGO
GFT (Oct 2014)
Santillana et al. (2014)
AR(3)
Google Correlate
Google Trend
Apr 04 2009
Oct 03 2009
Apr 03 2010
Oct 02 2010
Apr 02 2011
Oct 01 2011
Apr 07 2012
Oct 06 2012
Apr 06 2013
Oct 05 2013
Apr 05 2014
Oct 04 2014
Mar 28 2015
prediction error



(a)
03/29/09
--
12/27/09
H1N1 Flu outbreak
H1N1 Wave 1
H1N1 Wave 2
Apr 04 2009
Jun 06 2009
Aug 01 2009
Oct 03 2009
Nov 07 2009
Dec 26 2009
(b)
09/30/12
--
05/19/13
2012–13
Flu Season
Oct 06 2012
Dec 01 2012
Jan 05 2013
Mar 02 2013
Apr 06 2013
May 18 2013
(c)
09/28/14
--
04/26/15
2014–15
Partial Flu Season
Oct 04 2014
Nov 01 2014
Dec 06 2014
Jan 03 2015
Feb 07 2015
Mar 07 2015
Apr 04 2015
2
4
6
8

	Whole period	Off-season flu	Regular flu seasons (week 40 to week 20 next year)				
		H1N1	2010-11	2011-12	2012-13	2013-14	2014-15 partial
RMSE							
ARGO	**0.637**	**0.655**	**0.618**	**0.830**	**0.679**	**0.308**	**0.593**
GFT (Oct 2014)	2.213	0.773	1.110	3.023	4.451	0.981	0.683
Santillana et al. (2014)	0.909	0.945	0.864	1.688	0.918	0.495	0.683
AR(3)	0.955	0.813	0.794	1.051	1.191	0.966	0.924
Naive	1.000 (0.354)	1.000 (0.600)	1.000 (0.339)	1.000 (0.163)	1.000 (0.499)	1.000 (0.350)	1.000 (0.500)
MAE							
ARGO	**0.680**	**0.607**	**0.588**	**0.760**	**0.653**	**0.406**	**0.673**
GFT (Oct 2014)	1.828	0.777	1.260	3.277	5.028	0.884	0.726
Santillana et al. (2014)	1.035	0.793	0.977	1.782	0.897	0.634	0.872
AR(3)	0.920	0.777	0.787	0.951	0.988	0.915	0.924
Naive	1.000 (0.206)	1.000 (0.425)	1.000 (0.259)	1.000 (0.135)	1.000 (0.325)	1.000 (0.213)	1.000 (0.332)
Correlation							
ARGO	**0.984**	0.984	**0.988**	**0.924**	**0.968**	**0.993**	0.981
GFT (Oct 2014)	0.874	**0.989**	0.968	0.833	0.926	0.969	**0.984**
Santillana et al. (2014)	0.970	0.959	0.982	0.898	0.960	0.982	0.967
AR(3)	0.963	0.968	0.971	0.877	0.903	0.928	0.939
Naive	0.960	0.951	0.954	0.887	0.924	0.923	0.929
Corr. of increment							
ARGO	**0.744**	0.796	**0.793**	0.309	**0.532**	**0.944**	0.851
GFT (Oct 2014)	0.706	**0.863**	0.702	0.484	0.502	0.849	**0.910**
Santillana et al. (2014)	0.671	0.782	0.688	**0.599**	0.375	0.882	0.738
AR(3)	0.386	0.585	0.569	0.077	0.011	0.414	0.498
Naive	0.438	0.602	0.570	0.095	0.134	0.415	0.518

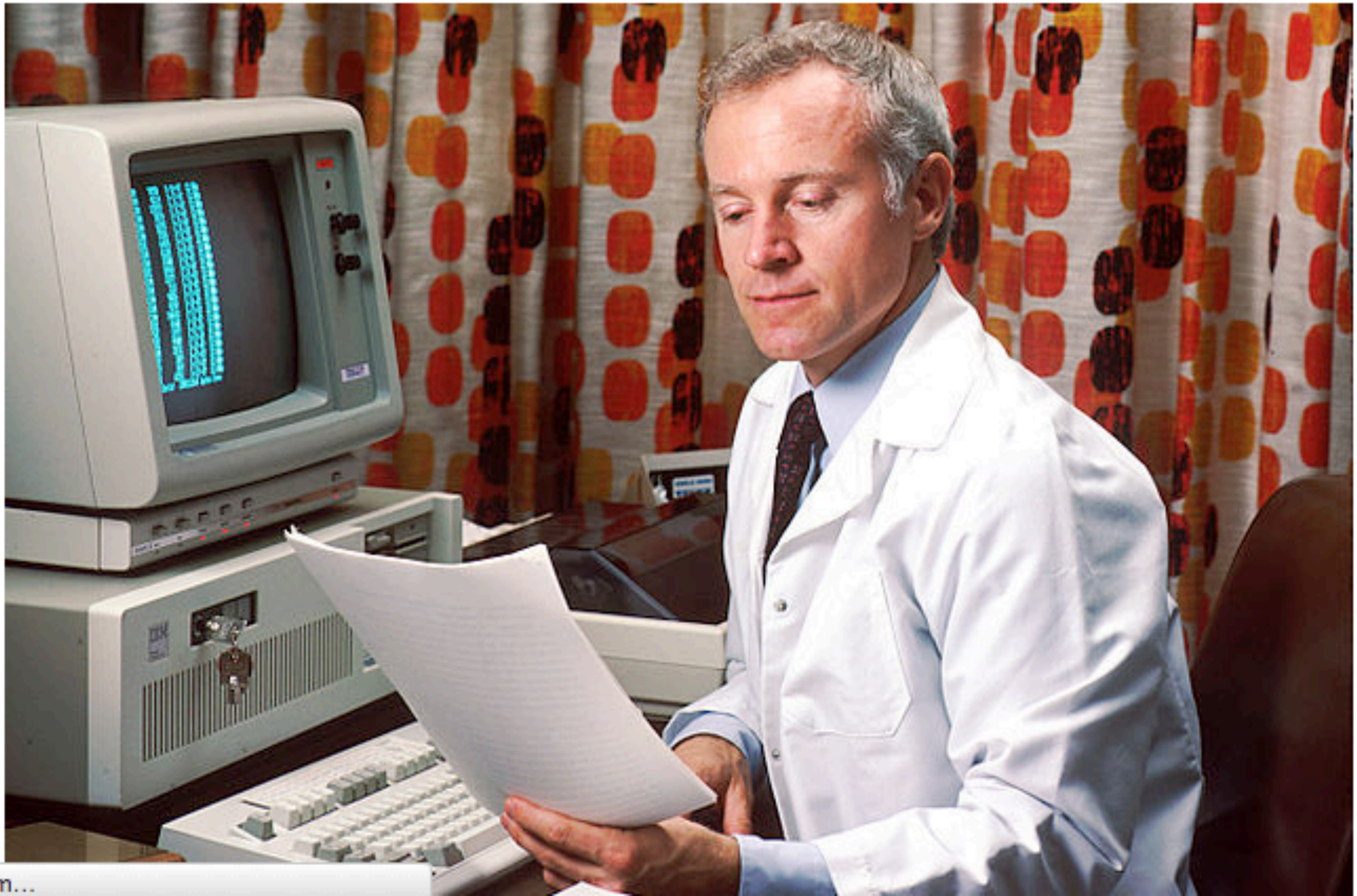

Negative coefficient
Positive coefficient
< −0.1
−0.05
0
0.05
> 0.1
(Intercept)
lag_1
lag_2
lag_3
lag_4
lag_5
lag_6
lag_7
lag_8
lag_9
lag_10
lag_11
lag_12
lag_13
lag_14
lag_15
lag_16
lag_17
lag_18
lag_19
lag_20
lag_21
lag_22
lag_23
lag_24
lag_25
lag_26
lag_27
lag_28
lag_29
lag_30
lag_31
lag_32
lag_33
lag_34
lag_35
lag_36
lag_37
lag_38
lag_39
lag_40
lag_41
lag_42
lag_43
lag_44
lag_45
lag_46
lag_47
lag_48
lag_49
lag_50
lag_51
2010
2012
2014


influenza.type.a
symptoms.of.flu
flu.duration
flu.contagious
flu.fever
treat.the.flu
how.to.treat.the.flu
signs.of.the.flu
how.long.is.the.flu
symptoms.of.the.flu
flu.recovery
cold.or.flu
flu.medicine
flu.or.cold
normal.body
is.flu.contagious
treat.flu
body.temperature
is.the.flu.contagious
reduce.fever
flu.treatment
flu.vs.cold
how.long.is.the.flu.contagious
fever.reducer
get.over.the.flu
treating.flu
flu.vs..cold
treatment.for.flu
human.temperature
dangerous.fever
the.flu
remedies.for.flu
influenza.a.and.b
contagious.flu
how.long.does.the.flu.last
fever.flu
oscillococcinum
flu.remedies
how.long.is.flu.contagious
flu.treatments
influenza.symptoms
cold.vs.flu
braun.thermoscan
fever.cough
signs.of.flu
how.long.does.flu.last
normal.body.temperature
get.rid.of.the.flu
type.a.influenza
i.have.the.flu
taking.temperature
flu.versus.cold
bronchitis
flu.germs
cold.vs..flu
flu.and.cold
thermoscan
flu.complications
high.fever
flu.children
the.flu.virus
how.to.treat.flu
pneumonia
flu.headache
flu.cough
ear.thermometer
how.to.get.rid.of.the.flu
flu.how.long
symptoms.of.bronchitis
cold.and.flu
over.the.counter.flu.medicine
treating.the.flu
flu.care
how.long.contagious
fight.the.flu
reduce.a.fever
cure.the.flu
medicine.for.flu
flu.length
cure.flu
exposed.to.flu
early.flu.symptoms
remedies.for.the.flu
flu.report
incubation.period.for.flu
break.a.fever
flu.contagious.period
influenza.incubation.period
cold.versus.flu
flu.in.children
what.to.do.if.you.have.the.flu
medicine.for.the.flu
flu.and.fever
incubation.period.for.the.flu
do.i.have.the.flu
cold.symptoms
how.to.reduce.a.fever
upper.respiratory
rapid.flu
fever.dangerous
2010
2012
2014

New flu tracker uses Google search data better than Google

Unlike defunct Flu Trends, the model is self-correcting and close to reality.

by **Beth Mole** - Nov 9, 2015 3:35pm EST



m...

New flu tracker uses Google search data better than Google

MOTS-CLÉS

La revanche du big data : Harvard plus forte que Google pour prédire la grippe

Des chercheurs de la prestigieuse université américaine ont conçu un modèle statistique deux fois plus efficace que la méthode Google. Le géant de l'Internet avait fermé cet été son projet, dont les prédictions avaient tourné au flop.

Par Delphine Cuny Rédactrice en chef adjointe. Publié le 11/11/2015 à 07h09

7 397 VISITES 7 RÉACTIONS • 1

m...

New

bett

Google, estadística y ‘big data’ para cazar brotes de gripe

G+1 20 Me gusta 154 Tweet

Un nuevo modelo que combina información epidemiológica y búsquedas de Google es capaz de predecir los brotes de gripe una o dos semanas antes que los métodos clínicos tradicionales. El modelo podrá servir para mejorar la toma de decisiones, como la distribución de personal y recursos hospitalarios en regiones que más lo necesiten.

Más información sobre: gripe brote epidemia Google estadística big data

SINC | Seguir a @agencia_sinc | 09 noviembre 2015 21:00

El modelo es capaz de producir estimaciones más precisas sobre brotes de gripe que cualquier otro método disponible, según los autores. / Sebastian Smit

Par Delphine Cuny Rédactric

m...

New
bett

Google, estadística y 'big data' para cazar brotes de gripe

G+1 20 Me gusta 154 Tweet

Un nuevo modelo que combina información epidemiológica y búsquedas de Google es capaz de predecir los brotes de gripe una o dos semanas antes que los métodos clínicos tradicionales. El modelo podrá servir para mejorar la toma de decisiones, como la distribución de personal y recursos hospitalarios en regiones que más lo necesiten.

Más información sobre: gripe brote epidemia Google estadística big data

HARVARD gazette

SCIENCE & HEALTH > HEALTH & MEDICINE

On top of the flu

Chance for advance warning in search-based tracking method

November 9, 2015 |

And on Aug 20^{th}, 2015

Google discontinues Flu Trends indefinitely!

Google Research Blog

The latest news from Research at Google

The Next Chapter for Flu Trends

Posted: Thursday, August 20, 2015

8+1 7

Instead of maintaining our own website going forward, we're now going to empower institutions who specialize in infectious disease research to use the data to build their own models. Starting this season, we'll provide Flu and Dengue signal data directly to partners including Columbia University's Mailman School of Public Health (to update their dashboard), Boston Children's Hospital/Harvard, and Centers for Disease Control and Prevention (CDC) Influenza Division. We will also continue to make historical Flu and Dengue estimate data available for anyone to see and analyze.

CIO

Most read:

NEWS

Google Flu Trends calls out sick, indefinitely

Google will pass along search queries related to the flu to health organizations so they can develop their own prediction models

By Fred O'Connor | Follow
IDG News Service | Aug 20, 2015 2:07 PM PT

MORE LIKE THIS

Google Begins Tracking Swine Flu in Mexico

Google's Panicky Flu Estimates Were Dead Wrong

VB

NEWS EVENTS RESEARCH SUBSCRIBE Sign up | Login

BIG DATA

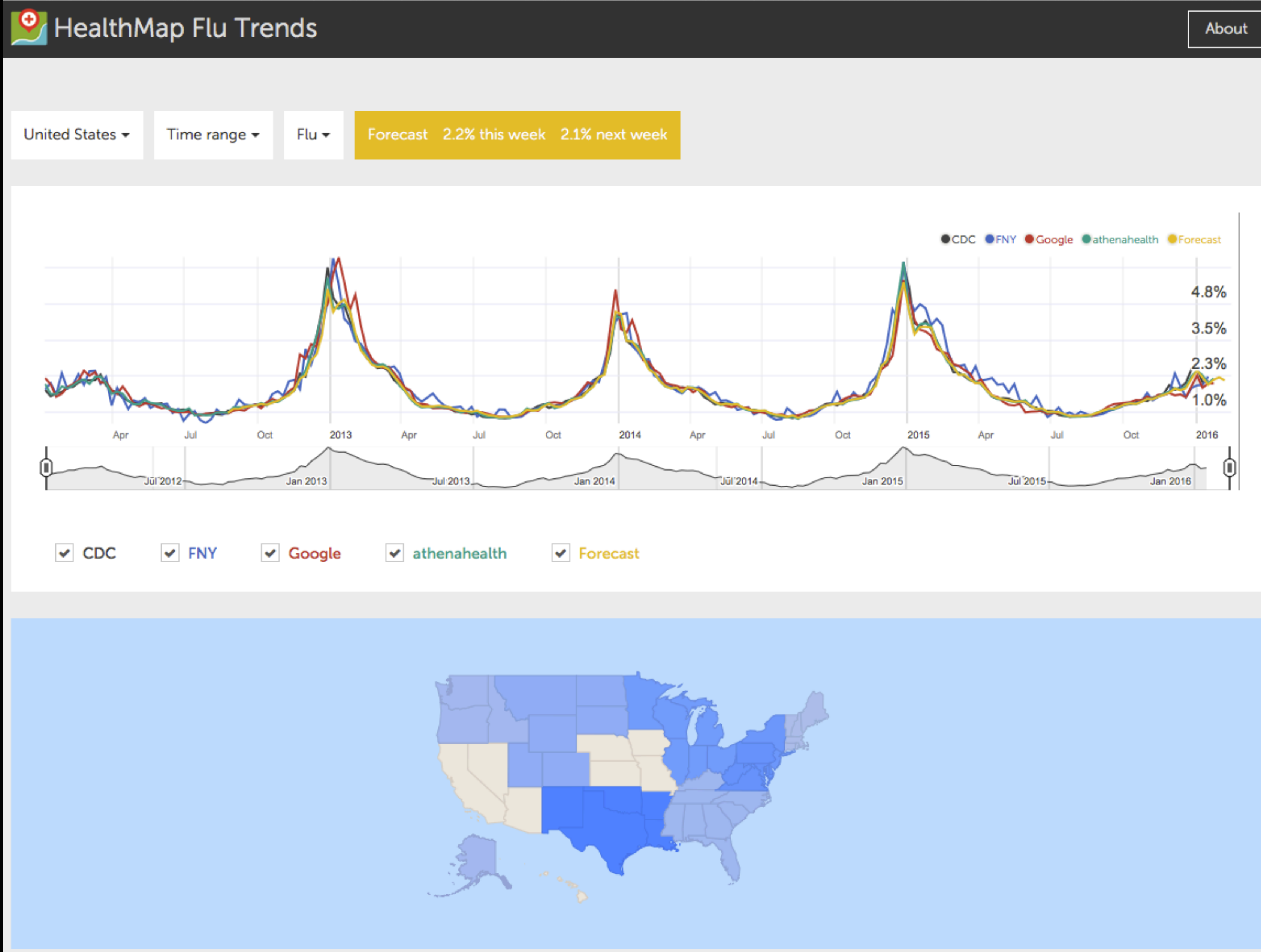
Google discontinues Flu Trends, starts offering data to researchers

JORDAN NOVET AUGUST 20, 2015 12:17 PM

TAGS: GOOGLE, GOOGLE FLU TRENDS

Our team at Boston Children's Hospital now has access to Google's search volumes, as one of the exclusive Google's partners.

We are creating a new improved disease forecasting platform

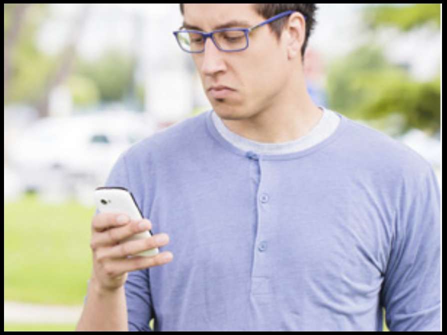


Thanks to Sue Aman, Rachel Chorney, Jeff Andre, Andre Nguyen, John Brownstein and Healthmap team!

HealthMap Boston Children's Hospital Until every child is well HARVARD MEDICAL SCHOOL

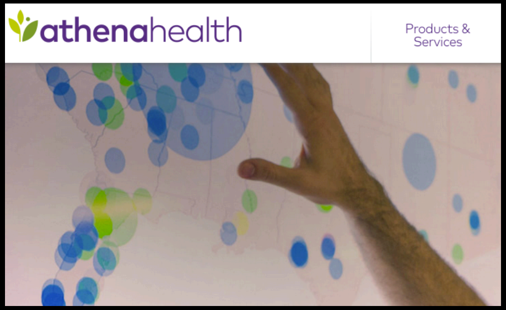
Beyond Google searches…

What are doctors searching for?



flu near you
do you have it in you?

What are people tweeting? What are they reporting on crowd-sourced disease surveillance apps?



Can we use Electronic Health Records (EHR) to track disease incidence? What lab tests or medications are doctors prescribing?