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
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
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.

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

Over 1 billion Google searches a day

Predictive Analytics

Newspaper articles, Reports, etc...

~2 billion smartphones world wide

230 million tweets every day

Source: IBM
Over 1 billion Google searches a day

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

Weather information in real-time

Medicine and Public Health

Newspaper articles, Reports, etc...

Electronic Health Records (EHR)

~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)
TRADITIONAL DISEASE REPORTING

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

- Public
- Healthcare Workers
  Electronic Health Records
- Ministries of Health
- World Bodies (WHO, EMA)
- Laboratories
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 Flu Trends

Epidemiological information available 2-3 weeks ahead of traditional clinical tracking systems.
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 Ginsberg¹, Matthew H. Mohebbi¹, Rajan S. Patel¹, Lynnette Brammer², Mark S. Smolinski¹ & Larry Brilliant¹

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. Mohebbi¹ Correspondence and requests for materials should be addressed to J.G. or M.H.M. (Email: flutrends-support@google.com).
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

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
When Google got flu wrong.
nature.com/news/when-google-got-flu-wrong.
Snowden And The Challenge Of Intelligence: The Practical Case Against The NSA's Big Data

“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.”

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

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

\[
\text{logit}(P) = \beta_0 + \beta_1 \times \text{logit}(Q) + \epsilon
\]

where \( P \) is the percentage of ILI physician visits, \( Q \) is the ILI-related query fraction, \( \beta_0 \) is the intercept,
Assumptions in Google Flu Trends:

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

\begin{quote}
\textbf{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”.
\end{quote}
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!
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.

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

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

(with D. Wendong Zhang)
<table>
<thead>
<tr>
<th>Rank</th>
<th>Term</th>
<th>Rank</th>
<th>Term</th>
<th>Rank</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>influenza type a</td>
<td>35</td>
<td>is the flu contagious</td>
<td>68</td>
<td>fever in adults</td>
</tr>
<tr>
<td>2</td>
<td>bronchitis a</td>
<td>36</td>
<td>flu in children</td>
<td>69</td>
<td>decongestant</td>
</tr>
<tr>
<td>3</td>
<td>influenza a</td>
<td>37</td>
<td>fever flu</td>
<td>70</td>
<td>normal body</td>
</tr>
<tr>
<td>4</td>
<td>symptoms of pneumonia</td>
<td>38</td>
<td>take action tour</td>
<td>71</td>
<td>low body temperature</td>
</tr>
<tr>
<td>5</td>
<td>flu incubation</td>
<td>39</td>
<td>flu remedies</td>
<td>72</td>
<td>a fever</td>
</tr>
<tr>
<td>6</td>
<td>influenza incubation</td>
<td>40</td>
<td>flu report</td>
<td>73</td>
<td>influenza a symptoms</td>
</tr>
<tr>
<td>7</td>
<td>flu contagious</td>
<td>41</td>
<td>nasal congestion</td>
<td>74</td>
<td>dangerous fever</td>
</tr>
<tr>
<td>8</td>
<td>influenza contagious</td>
<td>42</td>
<td>fever reducer</td>
<td>75</td>
<td>is flu contagious</td>
</tr>
<tr>
<td>9</td>
<td>flu incubation period</td>
<td>43</td>
<td>sinus infections</td>
<td>76</td>
<td>lauderdale florida</td>
</tr>
<tr>
<td>10</td>
<td>tussinex</td>
<td>44</td>
<td>rhode island wrestling</td>
<td>77</td>
<td>hotel fort lauderdale</td>
</tr>
<tr>
<td>11</td>
<td>benzonatate</td>
<td>45</td>
<td>symptoms of influenza</td>
<td>78</td>
<td>webmail shaw ca</td>
</tr>
<tr>
<td>12</td>
<td>influenza symptoms</td>
<td>46</td>
<td>castaway bay</td>
<td>79</td>
<td>high fever</td>
</tr>
<tr>
<td>13</td>
<td>a influenza</td>
<td>47</td>
<td>coral by the sea</td>
<td>80</td>
<td>robitussin ac</td>
</tr>
<tr>
<td>14</td>
<td>sinus</td>
<td>48</td>
<td>cold or flu</td>
<td>81</td>
<td>bronchitis contagious</td>
</tr>
<tr>
<td>15</td>
<td>pneumonia</td>
<td>49</td>
<td>respiratory infection</td>
<td>82</td>
<td>indoor driving</td>
</tr>
<tr>
<td>16</td>
<td>flu fever</td>
<td>50</td>
<td>take action</td>
<td>83</td>
<td>tussinex pennkinetic</td>
</tr>
<tr>
<td>17</td>
<td>flu duration</td>
<td>51</td>
<td>respiratory flu</td>
<td>84</td>
<td>wrestling report</td>
</tr>
<tr>
<td>18</td>
<td>taste of chaos</td>
<td>52</td>
<td>soweto gospel</td>
<td>85</td>
<td>walking pneumonia</td>
</tr>
<tr>
<td>19</td>
<td>bronchitis symptoms</td>
<td>53</td>
<td>soweto gospel choir</td>
<td>86</td>
<td>days inn miami</td>
</tr>
<tr>
<td>20</td>
<td>symptoms of bronchitis</td>
<td>54</td>
<td>illinois wrestling</td>
<td>87</td>
<td>body temperature</td>
</tr>
<tr>
<td>21</td>
<td>how long does the flu last</td>
<td>55</td>
<td>how long is the flu contagious</td>
<td>88</td>
<td>phlegm</td>
</tr>
<tr>
<td>22</td>
<td>symptoms of the flu</td>
<td>56</td>
<td>cold symptoms</td>
<td>89</td>
<td>flu relief</td>
</tr>
<tr>
<td>23</td>
<td>taste of chaos tour</td>
<td>57</td>
<td>the taste of chaos</td>
<td>90</td>
<td>mt sunapee</td>
</tr>
<tr>
<td>24</td>
<td>influenza incubation period</td>
<td>58</td>
<td>is bronchitis</td>
<td>91</td>
<td>harlem globe</td>
</tr>
<tr>
<td>25</td>
<td>sinus infection</td>
<td>59</td>
<td>upper respiratory</td>
<td>92</td>
<td>levaquin</td>
</tr>
<tr>
<td>26</td>
<td>flu recovery</td>
<td>60</td>
<td>afrin</td>
<td>93</td>
<td>strep throat</td>
</tr>
<tr>
<td>27</td>
<td>chaos tour</td>
<td>61</td>
<td>painful cough</td>
<td>94</td>
<td>coughing</td>
</tr>
<tr>
<td>28</td>
<td>type a influenza</td>
<td>62</td>
<td>laprepsoccer</td>
<td>95</td>
<td>whistler snow</td>
</tr>
<tr>
<td>29</td>
<td>flu symptoms</td>
<td>63</td>
<td>upper respiratory infection</td>
<td>96</td>
<td>fever temperature</td>
</tr>
<tr>
<td>30</td>
<td>tussalon</td>
<td>64</td>
<td>amoxicillin</td>
<td>97</td>
<td>sales tax credit</td>
</tr>
<tr>
<td>31</td>
<td>type a flu</td>
<td>65</td>
<td>ski harness</td>
<td>98</td>
<td>glitches</td>
</tr>
<tr>
<td>32</td>
<td>treat the flu</td>
<td>66</td>
<td>robitussin dm</td>
<td>99</td>
<td>pennkinetic</td>
</tr>
<tr>
<td>33</td>
<td>treating the flu</td>
<td>67</td>
<td>treating flu</td>
<td>100</td>
<td>histinex</td>
</tr>
<tr>
<td>34</td>
<td>how to treat the flu</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
First week after being published online, it became the second most read paper in journal’s history! (After a paper published in 1998)
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

\[ \text{logit}[I(t)] = \sum_{i=1}^{n} a_i(t) \text{logit}[Q_i(t)] + e, \]
Santillana et al. American Journal of Preventive Medicine, 2014; 47 (3) pp 341-347
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

theguardian.com, Friday 4 July 2014 11.44 EDT
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

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.
Google Flu Trends is overstated, research finds. New study finds way to improve threefold — but says system must be revised.

**Featured Research**

**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.

**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.
Study: Methodology Changes Improve Google Flu Trend Accuracy

Monday, July 7, 2014

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).

**Related Topics:**
- Public Health
Big Data’s Potential in Public Health: Revisiting Google Flu Trends

July 7, 2014  Written by: Dan Gray  1 Reply

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).
Google Flu Trends gets a brand new engine

Posted: Friday, October 31, 2014

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
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).

Accurate estimation of influenza epidemics using Google search data via ARGO

Shihao Yang, Mauricio Santillana, and S. C. Kou

*Department of Statistics, Harvard University, Cambridge, MA 02138; †School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; and ‡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 epidemic, 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 1–3 wk 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,
<table>
<thead>
<tr>
<th></th>
<th>Whole period</th>
<th>Off-season flu</th>
<th>Regular flu seasons (week 40 to week 20 next year)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RMSE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGO</td>
<td>0.637</td>
<td>0.655</td>
<td>0.618</td>
</tr>
<tr>
<td>GFT (Oct 2014)</td>
<td>2.213</td>
<td>0.773</td>
<td>1.110</td>
</tr>
<tr>
<td>Santillana et al. (2014)</td>
<td>0.909</td>
<td>0.945</td>
<td>0.864</td>
</tr>
<tr>
<td>AR(3)</td>
<td>0.955</td>
<td>0.813</td>
<td>0.794</td>
</tr>
<tr>
<td>Naive</td>
<td>1.000 (0.354)</td>
<td>1.000 (0.600)</td>
<td>1.000 (0.339)</td>
</tr>
<tr>
<td><strong>MAE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGO</td>
<td>0.680</td>
<td>0.607</td>
<td>0.588</td>
</tr>
<tr>
<td>GFT (Oct 2014)</td>
<td>1.828</td>
<td>0.777</td>
<td>1.260</td>
</tr>
<tr>
<td>Santillana et al. (2014)</td>
<td>1.035</td>
<td>0.793</td>
<td>0.977</td>
</tr>
<tr>
<td>AR(3)</td>
<td>0.920</td>
<td>0.777</td>
<td>0.787</td>
</tr>
<tr>
<td>Naive</td>
<td>1.000 (0.206)</td>
<td>1.000 (0.425)</td>
<td>1.000 (0.259)</td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGO</td>
<td>0.984</td>
<td>0.984</td>
<td>0.988</td>
</tr>
<tr>
<td>GFT (Oct 2014)</td>
<td>0.874</td>
<td>0.989</td>
<td>0.968</td>
</tr>
<tr>
<td>Santillana et al. (2014)</td>
<td>0.970</td>
<td>0.959</td>
<td>0.982</td>
</tr>
<tr>
<td>AR(3)</td>
<td>0.963</td>
<td>0.968</td>
<td>0.971</td>
</tr>
<tr>
<td>Naive</td>
<td>0.960</td>
<td>0.951</td>
<td>0.954</td>
</tr>
<tr>
<td><strong>Corr. of increment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGO</td>
<td>0.744</td>
<td>0.796</td>
<td>0.793</td>
</tr>
<tr>
<td>GFT (Oct 2014)</td>
<td>0.706</td>
<td>0.863</td>
<td>0.702</td>
</tr>
<tr>
<td>Santillana et al. (2014)</td>
<td>0.671</td>
<td>0.782</td>
<td>0.688</td>
</tr>
<tr>
<td>AR(3)</td>
<td>0.386</td>
<td>0.585</td>
<td>0.569</td>
</tr>
<tr>
<td>Naive</td>
<td>0.438</td>
<td>0.602</td>
<td>0.570</td>
</tr>
</tbody>
</table>
The image contains a heatmap and a visual representation of time-series data. The heatmap displays a color gradient indicating negative and positive coefficients, with values ranging from < -0.1 to > 0.1. The time-axis is labeled with years from 2010 to 2014, and the data points are marked along this axis.
New flu tracker uses Google search data better than Google
Unlike defunct Flu Trends, the model is self-correcting and close to reality.
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.
Google, estadística y ‘big data’ para cazar brotes de gripe

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

Par Delphine Cuny Rédactrice

El modelo es capaz de producir estimaciones más precisas sobre brotes de gripe que cualquier otro método disponible, según los autores. / Sebastian Smit
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.
And on Aug 20th, 2015
Google discontinues Flu Trends indefinitely!

The Next Chapter for Flu Trends

Posted: Thursday, August 20, 2015

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.
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

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!
Beyond Google searches...

What are doctors searching for?

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?