LETTERS

Edited by **Jennifer Sills**

Twitter: Big data opportunities

IN THEIR POLICY FORUM "The parable of Google Flu: Traps in big data analysis" (14 March, p. 1203), D. Lazer *et al.* remark upon recent failures of Google Flu Trends (GFT) and cast these as limitations of big data analysis in general. However, many of these limitations have been overcome by other big data systems. Specifically,

analyses that use Twitter for influenza surveillance account for the concerns of replicability, overfitting, construct validity, granularity, and temporal confounds that Lazer *et al.* have identified.

For example, the results of GFT cannot be replicated because they are based on proprietary data, whereas Twitter data are open. A community of researchers has replicated analyses based on these data [e.g., (*1*)]. Changes to the social media platform itself, such as reengineering the underlying algorithms, need not adversely impact replicability as long as these data remain open.

Previous articles have correctly remarked that GFT and keyword-based systems overestimate influenza prevalence by conflating signals of influenza awareness (such as media attention) with signals of actual infection (*2*, *3*). These signals are separable on Twitter. Our work (*4*) has recently shown that the rate of tweets indicative of actual influenza infection is strongly correlated with the U.S. Center for Disease Control's Influenza-Like Illness rates, even though these rates were not used for system development and despite focused media attention. Furthermore, our Twitter evaluations do not suffer from peak overestimation as does GFT. Finally, our analysis explicitly controls for seasonality and temporal autocorrelation, meaning that our results directly capture flu trends and not simply seasonal variations.

Finally, New York City's Department of Health and Hygiene successfully conducted a blind evaluation of our method using municipal data, also resulting in a strong correlation. Our system successfully demonstrates the ability to understand the prevalence of flu at local levels. Concerns that are specific to GFT should not be overly generalized to other big data analyses.

David Andre Broniatowski,[1]*
Michael J. Paul,[2] Mark Dredze[2]

[1]Department of Engineering Management and Systems Engineering, George Washington University, Washington, DC 20052, USA. [2]Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA.

*Corresponding author. E-mail: broniatowski@gwu.edu

REFERENCES

1. S. Tuarob, C. S. Tucker, M. Salathe, N. Ram, *J. Biomed. Inform.* **49**, 255 (2014).
2. P. Copeland *et al.*, *Proc. Int. Soc. Negl. Trop. Dis.* 3 (2013).
3. D. Butler, *Nature* **494**, 155 (2013).
4. D. Broniatowski, M. J. Paul, M. Dredze, *PLOS ONE* **8**, e83672 (2014).

Response

WE THANK BRONIATOWSKI, Paul, and Dredze for giving us the opportunity to reemphasize the potential of big data and make the more obvious point that not all big data projects have the problems currently plaguing Google Flu Trends (GFT), nor are these problems inherent to the field in general.

Our Policy Forum is meant to provide a constructive critique by highlighting possible pitfalls of big data analysis. These pitfalls are not the same for all big data sets, but are certainly not unique to GFT. We do agree that Twitter has substantial scientific potential and is distinctive in the public availability of its data. Indeed, one of us (A.V.) is using Twitter data for influenza surveillance in the context of the recent Center for Disease Control (CDC) "Predict the Flu Challenge" (*1*).

Twitter data provide an excellent representation of those who choose to express an opinion publicly, which can be of tremendous value for many research purposes. However, these data may be manipulated by both the service provider (such as Google) and the user (such as companies marketing a product), as we explain in our Policy Forum. In light of these trends, whether these data can be used to represent the entire United States population remains an open question.

Who uses Twitter and how they use it have changed markedly over the past several years. The algorithmic underpinning of Twitter (which identifies "what's trending") is subject to constant and invisible tinkering. The system is under constant attack, with armies of bots ready to produce content for the highest bidder (*2*, *3*). The norms of expression on Twitter are heterogeneous and still rapidly evolving—who feels the need to publicly express that they have flu symptoms on Twitter, and are these predispositions evenly distributed throughout the population (*4*)? Bodnar and Salathé's cautionary tale (*5*) on Twitter-based influenza surveillance clearly shows that seemingly irrelevant tweets (such as those about zombies) are moderately indicative of influenza prevalence, and that the choice of validation methods has a large effect on reported success.

It is possible that one day we will have reliable prediction of flu prevalence from social media. Certainly, this would require a careful evaluation and recalibration of methodologies, public and independent replication of results, and the explicit evaluation of error processes. Clearly, all big data projects do not have the same syndromes as GFT presently does, but by building strong collaborations and adhering to rigorous standards, we should be able to extract considerably more information from these highly informative new data sources.

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

[1]Lazer Laboratory, Northeastern University, Boston, MA 02115, USA. [2]Harvard Kennedy School, Harvard University, Cambridge, MA 02138, USA. [3]Institute for Quantitative Social Science, Harvard University, Cambridge, MA 02138, USA. [4]University of Houston, Houston, TX 77204, USA. [5]Laboratory for the Modeling of Biological and Sociotechnical Systems, Northeastern University, Boston, MA 02115, USA. [6]Institute for Scientific Interchange Foundation, Turin, Italy.

*Corresponding author. E-mail: d.lazer@neu.edu

REFERENCES

1. CDC, CDC Competition Encourages Use of Social Media to Predict Flu (www.cdc.gov/flu/news/predict-flu-challenge.htm).
2. J. Zhang, R. Zhang, Y. Zhang, G. Yan, "On the impact of social botnets for spam distribution and digital-influence manipulation" [2013 IEEE Conference on Communications and Network Security (CNS), 2013].
3. N. Bilton, "Friends, and influence, for sale online" (20 April

2014); http://bits.blogs.nytimes.com/2014/04/20/friends-and-influence-for-sale-online/.
4. Y. Liu, C. Kliman-Silver, A. Mislove, in Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (ICWSM'14) (Ann Arbor, MI, 2014).
5. T. J. Bodnar, M. Salathé, "Validating models for disease detection using Twitter," 1st International Workshop on Public Health in the Digital Age: Social Media, Crowdsourcing and Participatory Systems, WWW2013 (2013).

Leadership void for "designer baby" ethics

IN HIS PERSPECTIVE "Stirring the simmering 'designer baby' pot" (14 March, p. 1208), T. H. Murray discusses the ethical quandaries posed by looming genetic selection mechanisms designed to improve the health—and perhaps even the "quality"—of progeny. An underlying dilemma, in his view, is the lack of general agreement on an ethical framework to guide professional associations and public policy decision-makers in addressing the propriety of "designer" protocols. A critical step, he argues, is to agree on an appropriate forum for fleshing out various ethical claims and then selecting one among many sets of assumptions on which to base specific recommendations.

Murray suggests the President's Commission on Bioethics as a candidate for this forum, and he urges the Commission to intercede. However, the President's Commission is an advisory council, lacking the ability to coerce anyone to do anything, but rather advising the President of the United States on bioethical policy questions. If the committee issues a report the President does not like, it will gather dust. Even if President Obama did like the recommendations in a report, in the case of "designer babies," he would not have the constitutional power to impose his will in this area of human affairs. Many presidents have relied on such reports as tactics for galvanizing congressional action. But Congress, even when not in partisan gridlock, has been unable to enter the "reproductive liberty" environment with any degree of success.

Another candidate for this role could be the Food and Drug Administration (FDA), which has claimed [(*1*), p. 105] that human cloning falls within its purview. But is a reproductive strategy a food or a drug? More fundamentally, the key questions surrounding human cloning are matters of ethics and social accountability, and yet the FDA has no expertise in evaluating ethical, legal, and social issues (ELSI). In my view, even putative FDA jurisdiction over mitochondrial manipulation methodologies

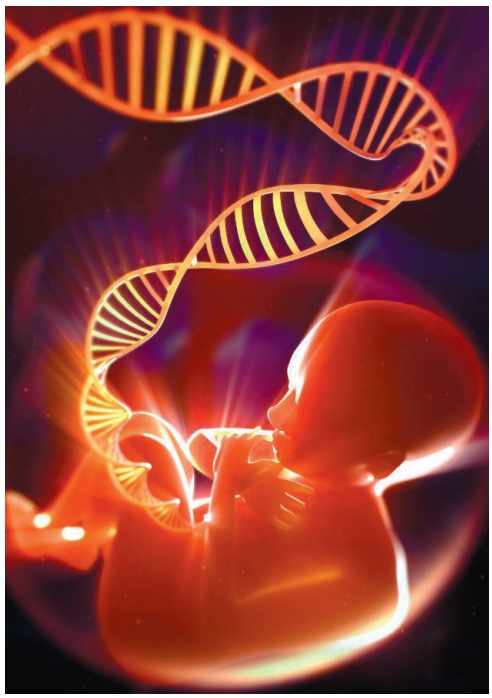

PHOTO: HENRIK JONSSON GRAPHIC DESIGN/ISTOCKPHOTO

has focused on how to do it without any pretense of considering ELSI strategies. Furthermore, the FDA does not conduct open meetings at which the citizenry can petition for a redress of concerns, which makes it poorly suited to address first-order "designer baby" issues.

Another, less likely, candidate is the U.S. Patent Office. With much bravado, the Patent Office announced that a cloned human being could not be patented [(*1*), p. 94], even though other synthetically produced life forms could be. In explanation, the Patent Office argued that no person—cloned or otherwise—satisfies the "utility" standard (i.e., a person cannot be considered a functional artifact serving some ulterior purpose). Moreover, to give somebody a patent—a 20-year monopoly—on a flesh-and-blood human makes that person a sort-of slave, and slavery is outlawed by the Thirteenth Amendment. It is not clear whether processes for converting unfertilized human eggs into preembryos without using male gametes are patentable material. Nor is it clear what ELSI authority or capacity the Patent Office possesses to promulgate policy in this area.

Given that the President's Commission, the FDA, and the Patent Office all leave much to be desired in terms of steering public policy on this issue, the leadership void described by Murray remains. It is an anxious business when the political system fails to engage significant policy agendas. Perhaps we have gone about it in the wrong way. We have assumed that "designer babies" are a matter for national action. Instead, state legislatures could serve as laboratory systems, thrashing through each and every "designer baby" ELSI question and passing appropriate legislation. The federal system could then serve as a marketplace for state opinion, where some sort of consensus will inevitably emerge.

Ira H. Carmen

Cortlandt Manor, NY, 10567, USA. E-mail: irahcarmen@gmail.com

REFERENCE

1. I. H. Carmen, *Politics in the Laboratory* (Univ. of Wisconsin Press, Madison, WI, 2004).

TECHNICAL COMMENT ABSTRACTS

Comment on "Quantifying long-term scientific impact"

Jian Wang, Yajun Mei, Diana Hicks

■ Wang *et al.* (Reports, 4 October 2013, p. 127) claimed high prediction power for their model of citation dynamics. We replicate their analysis but find discouraging results: 14.75% papers are estimated with unreasonably large μ (>5) and λ (>10) and correspondingly enormous prediction errors. The prediction power is even worse than simply using short-term citations to approximate long-term citations.

Full text at http://dx.doi.org/10.1126/science.1248770

Response to Comment on "Quantifying long-term scientific impact"

Dashun Wang, Chaoming Song, Hua-Wei Shen, Albert-László Barabási

■ Wang, Mei, and Hicks claim that they observed large mean prediction errors when using our model. We find that their claims are a simple consequence of overfitting, which can be avoided by standard regularization methods. Here, we show that our model provides an effective means to identify papers that may be subject to overfitting, and the model, with or without prior treatment, outperforms the proposed naïve approach.

Full text at http://dx.doi.org/10.1126/science.1248961