

**15% off upholstery sale.**  
 Crate&Barrel Save on over 2,000 fabrics and leathers. [Shop now >](#)



# Government Aims to Build a 'Data Eye in the Sky'

By JOHN MARKOFF  
Published: October 10, 2011

More than 60 years ago, in his "Foundation" series, the science fiction novelist Isaac Asimov invented a new science — psychohistory — that combined mathematics and [psychology](#) to predict the future.

[Enlarge This Image](#)



Brian Stauffer

Now social scientists are trying to mine the vast resources of the Internet — Web searches and Twitter messages, Facebook and blog posts, the digital location trails generated by billions of cellphones — to do the same thing.

The most optimistic researchers believe that these storehouses of "big data" will for the first time reveal sociological laws of human behavior — enabling them to predict political crises, revolutions and other forms of social and economic instability, just as physicists and chemists can predict natural phenomena.

"This is a significant step forward," said Thomas Malone, the director of the Center for Collective Intelligence at the Massachusetts Institute of Technology. "We have vastly more detailed and richer kinds of data available as well as predictive algorithms to use, and that makes possible a kind of prediction that would have never been possible before."

The government is showing interest in the idea. This summer a little-known intelligence agency began seeking ideas from academic social scientists and corporations for ways to automatically scan the Internet in 21 Latin American countries for "big data," according to a research proposal being circulated by the agency. The three-year experiment, to begin in April, is being financed by the [Intelligence Advanced Research Projects Activity](#), or Iarpa (pronounced eye-AR-puh), part of the office of the director of national intelligence.

The automated data collection system is to focus on patterns of communication, consumption and movement of populations. It will use publicly accessible data, including Web search queries, blog entries, Internet traffic flow, financial market indicators, traffic webcams and changes in Wikipedia entries.

It is intended to be an entirely automated system, a "data eye in the sky" without human intervention, according to the program proposal. The research would not be limited to

RECOMMEND

TWITTER

LINKEDIN

COMMENTS (6)

E-MAIL

PRINT

REPRINTS

SHARE



## Latest in My Network

What's This?

**Siva Vaidhyanathan** RECOMMENDED  
Occupy Wall Street and the Tea Party

**Siva Vaidhyanathan** RECOMMENDED  
Your Soaring Health Insurance Premiums

**Siva Vaidhyanathan** RECOMMENDED  
Derrick Bell, Pioneering Harvard Law Professor, Dies at 80

POPULAR NOW  
The Myth of Voter Fraud

**Crate&Barrel**  
**15% off.**  
 It's the **biggest upholstery event of the year.** Save on over 2,000 fabrics.  
 Now through October 16.  
[Shop now >](#)



## RSS Feed

[RSS](#) Get Science News From The New York Times >

## Readers' Comments

Share your thoughts.

[Post a Comment >](#)

[Read All Comments \(6\) >](#)

## TicketWatch: Theater Offers by E-Mail

Sign up for ticket offers from Broadway shows and other advertisers.  
[furd@mit.edu](mailto:furd@mit.edu) [Sign Up](#)  
[Change E-mail Address](#) | [Privacy Policy](#)

MOST E-MAILED

RECOMMENDED FOR YOU

130 articles in the past month

FRField3 All Recommendations 

1. GREEN COLUMN  
[Airlines Weigh the Advantages of Using More Biofuel](#)



2. [Recession Officially Over, U.S. Incomes Kept Falling](#)

3. EDITORIAL  
[The Myth of Voter Fraud](#)

4. THE TEXAS TRIBUNE  
[The Post-Census Blues Hit Voters \(and Mapmakers\)](#)



5. [For Romney, Social Issues Pose New Test](#)

political and economic events, but would also explore the ability to predict pandemics and other types of widespread contagion, something that has been pursued independently by civilian researchers and by companies like Google.

Some social scientists and advocates of privacy rights are deeply skeptical of the project, saying it evokes queasy memories of [Total Information Awareness](#), a post-9/11 Pentagon program that proposed hunting for potential attackers by identifying patterns in vast collections of public and private data: telephone calling records, e-mail, travel data, visa and passport information, and credit card transactions.

“I have Total Information Awareness flashbacks when things like this happen,” said David Price, an anthropologist at St. Martin’s University in Lacey, Wash., who has written about cooperation between social scientists and intelligence agencies. “On the one hand it’s understandable for a nation-state to want to track things like the outbreak of a pandemic, but I have to wonder about the total automation of this and what productive will come of it.”

Iarpa officials declined to discuss the research program, saying they are prohibited from giving interviews until contract awards are made later this year.

A similar project by their military sister organization, the Defense Advanced Research Projects Agency, or Darpa, aims to automatically identify insurgent social networks in Afghanistan.

In its most recent budget proposal, the defense agency argues that its analysis can expose terrorist cells and other stateless groups by tracking their meetings, rehearsals and sharing of material and money transfers.

So far there have been only scattered examples of the potential of mining social media. Last year HP Labs researchers used Twitter data to accurately predict box office revenues of Hollywood movies. In August, the National Science Foundation approved funds for research in using social media like Twitter and Facebook to assess earthquake damage in real time.

The accessibility and computerization of huge databases has already begun to spur the development of new statistical techniques and new software to manage data sets with trillions of entries or more.

“Big data allows one to move beyond inference and statistical significance and move toward meaningful and accurate analyses,” said Norman Nie, a political scientist who was a pioneering developer of statistical tools for social scientists and who recently formed a new company, Revolution Analytics, to develop software for the analysis of immense data sets.

Some scientists are skeptical. They cite the Pentagon’s ill-fated Project Camelot in the 1960s, which also explored the possibility that social science could predict political and economic events, but was canceled in the face of widespread criticism by scholars.

The project focused on Chile, with the goal of developing methods for anticipating “violent changes” and offering ways of averting possible rebellions. It led to an uproar among social scientists, who argued that the study would compromise their professional ethics.

In recent years, however, academic opposition to military financing of research has faded. Since 2008, a Pentagon project called the [Minerva Initiative](#) has paid for an array of studies, including research at Arizona State University into political opponents of radical Muslims and a University of Texas study on the effects of [climate change](#) on African political stability.

Social scientists who cooperate with the research agencies contend that, on balance, the new technologies will have a positive effect.



6. [Poker Inc. to Uncle Sam: Shut Up and Deal](#)

7. EDITORIAL  
[Future of the Gulf](#)

8. EDITORIAL  
[Protesters Against Wall Street](#)



9. [Lawmakers End Boycott of Parliament in Afghanistan](#)

10. NATIONAL BRIEFING | WEST  
[California: Governor Signs Ban on Shark Fins](#)

[Go to Your Recommendations »](#)  
[What's This?](#) | [Don't Show](#)



**What you get for ... \$1.25 million**

ALSO IN REAL ESTATE »

[What you get for ... \\$599,000](#)

[What you get for ... \\$280,000](#)

**nytimes.com**

REAL ESTATE

ADVERTISEMENTS



**SAVE 50%**  
ON HOME DELIVERY  
OF THE WEEKENDER

**The New York Times**

**SUBSCRIBE NOW ▶**

Advertise on NYTimes.com

Ads by Google

what's this?

**[Are You An SEO Firm?](#)**

Get More SEM Clients In Just

“The result will be much better understanding of what is going on in the world, and how well local governments are handling the situation,” said Sandy Pentland, a computer scientist at the M.I.T. Media Laboratory. “I find this all very hopeful rather than scary, because this is perhaps the first real opportunity for all of humanity to have transparency in government.”

But advocates of privacy rights worry that public data and the related techniques developed in the new Iarpa project will be adapted for clandestine “total information” operations.

“These techniques are double-edged,” said Marc Rotenberg, president of the Electronic Privacy Information Center, a privacy rights group based in Washington. “They can be used as easily against political opponents in the United States as they can against threats from foreign countries.”

And some computer scientists expressed skepticism about efforts to predict political instability with indicators like Web searches.

“I’m hard pressed to say that we are witnessing a revolution,” said Prabhakar Raghavan, the director of Yahoo Labs, who is an information retrieval specialist. He noted that much had been written about predicting [flu](#) epidemics by looking at Web searches for “flu,” but noted that the predictions did not improve significantly on what could already be found in data from the Centers for Disease Control and Prevention.

“You can look at search queries and divine that flu is about to break out,” he said, “but what our research has highlighted is that many of these new methods don’t add a huge lift.”

Other researchers are far more optimistic. “There is a huge amount of predictive power in this data,” said Albert-Laszlo Barabasi, a physicist at Notre Dame who specializes in network science. “If I have hourly information about your location, with about 93 percent accuracy I can predict where you are going to be an hour or a day later.”

Still, the ease of acquiring and manipulating huge data sets charting Internet behavior causes many researchers to warn that the data mining technologies may be quickly outrunning the ability of scientists to think through questions of privacy and ethics.

There is also the deeper question of whether it will be possible to discern behavioral laws that match the laws of physical sciences. For Isaac Asimov, the predictive powers of psychohistory worked only when it was possible to measure the human population of an entire galaxy.

GET MORE SEM CLIENTS. IN JUST  
3 Easy Steps. Join Google Engage!  
[www.google.com/ads/engage](http://www.google.com/ads/engage)

COMMENTS (6)

E-MAIL

PRINT

REPRINTS



**Get 50% Off The New York Times & Free All Digital Access.**

SPONSORED HEADLINES

[What's This?](#)

**Get Free E-mail Alerts on These Topics**

Lifescrypt.com

**Fibromyalgia Syndrome: Mystery Pain Explained**

Free Press

**Solar Car Design Powered by High Performance Computing**

CIO

**Supercomputer Nautilus 'Knew Where Bin Laden Was'**

Social Networking (Internet)

Sociology

Politics and Government

Science and Technology

Ads by Google

what's this?

### **Osteoporosis Exercises**

Increase Bone Density Levels With  
These 3 Simple Exercises. Start Now  
[StayingFit.com](http://StayingFit.com)

#### INSIDE NYTIMES.COM



BUSINESS »



Households Pay a Price for China's Growth

OPINION »

### Letters: The Hyperachieving Crowd

Readers respond to a Sunday Review article on our society's competitive nature.

BOOKS »



John Lennon's Primal Screams

OPINION »



Room for Debate: Have Americans Gotten Lazy?

MUSIC »



Michael Jackson, Celebrated and Sold

U.S. »



When the Uprooted Put Down Roots

[Home](#) | [World](#) | [U.S.](#) | [N.Y./Region](#) | [Business](#) | [Technology](#) | [Science](#) | [Health](#) | [Sports](#) | [Opinion](#) | [Arts](#) | [Style](#) | [Travel](#) | [Jobs](#) | [Real Estate](#) | [Autos](#) | [Site Map](#)

© 2011 The New York Times Company | [Privacy](#) | [Your Ad Choices](#) | [Terms of Service](#) | [Terms of Sale](#) | [Corrections](#) | [RSS](#) | [Help](#) | [Contact Us](#) | [Work for Us](#) | [Advertise](#)

◀ OPEN

**MORE IN SCIENCE** (2 OF 35 ARTICLES)  
**Silver Treasure, Worth Found in North Atlanta**  
[Read More »](#)