

# Radical Collaboration: The Science of CrowdSourcing and CrowdSourcing Science

Presentation to the Government-University-Industry  
Research Roundtable

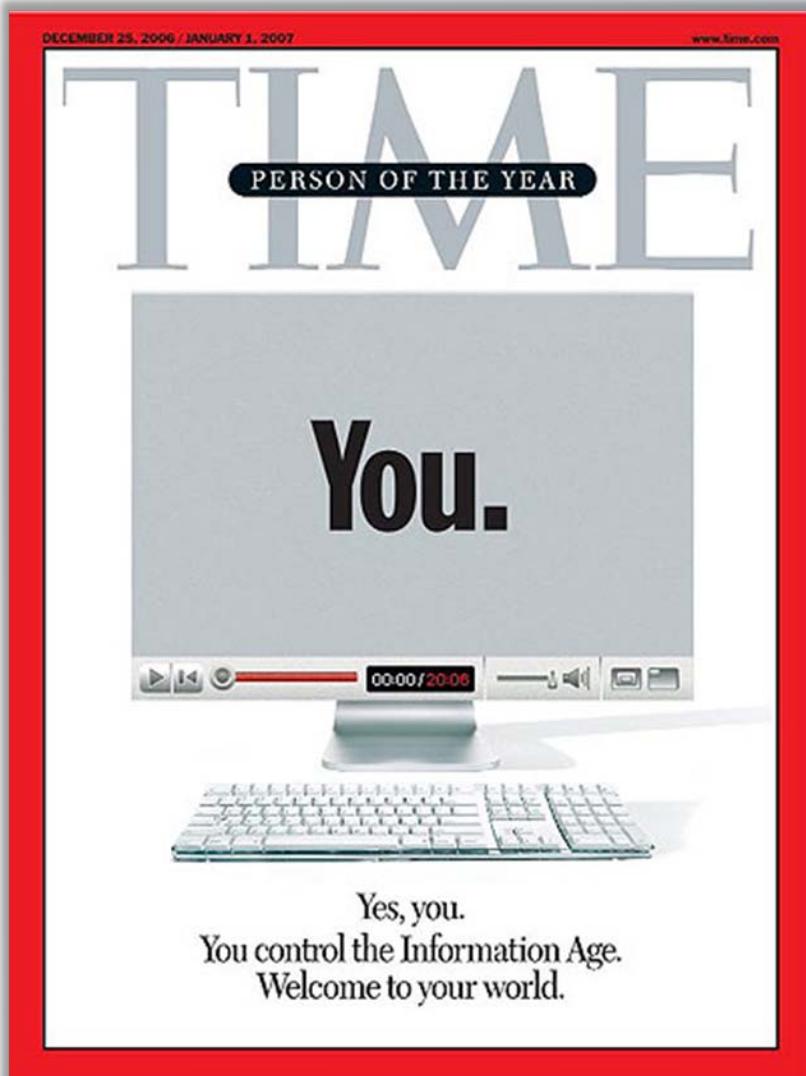
Richard P. Moser, Ph.D.  
Research Psychologist

Bradford W. Hesse, Ph.D.  
Chief, Health Communication and Informatics Research Branch

Behavioral Research Program

October 5, 2011

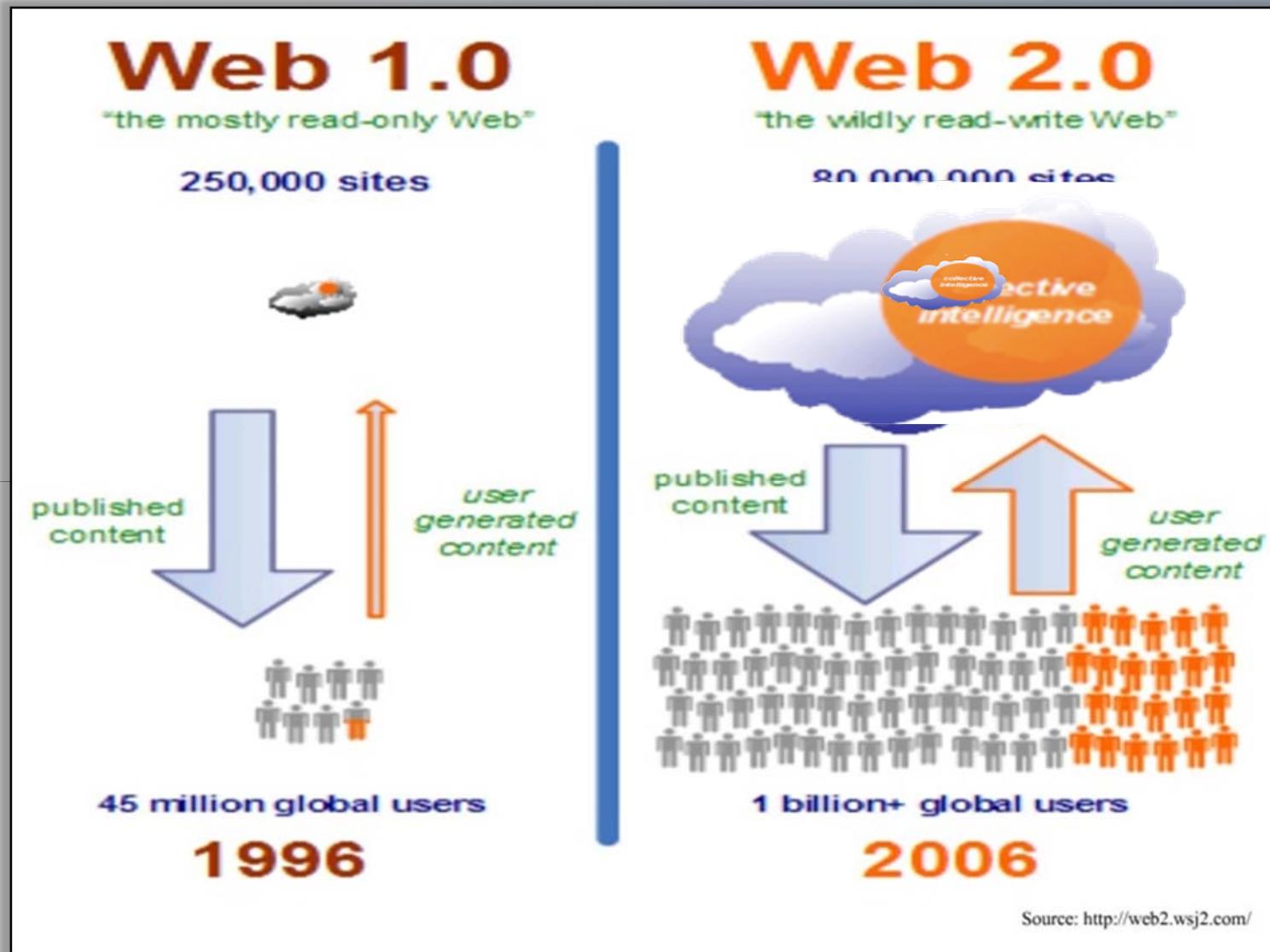
# Emergence of the Participative Web



- Characteristics of Web 2.0
- Platforms for Participation
- Harnessing “Collective Intelligence”
- Data as the new “Intel Inside”

*Time Magazine, 2006*

# Evolution of The Internet



Courtesy of Gunther Eysenbach, “Infodemiology & Infoveillance,” slideshare.net

# Science in a New Collaborative Environment

The image displays three separate screenshots illustrating different collaborative environments:

- Blog Interface:** A screenshot of a blog interface with a large "Blog" title and a cursor pointing at it. Below the title, there is a snippet of Latin text: "dolor sit amet, consectetur adipisci ut laoreet dolor veniam, qu".
- Wikipedia Homepage:** The main page of Wikipedia, showing links to various language versions of the encyclopedia. The English version is highlighted, showing 1,956,000+ articles. Other versions shown include Deutsch (625,000+ articles), Français (544,000+ articles), 日本語 (403,000+ articles), Nederlands (325,000+ articles), Español (257,000+ articles), Polski (415,000+ articles), Italiano (327,000+ articles), Português (290,000+ articles), and Svenska (245,000+ articles). The Wikipedia logo, a globe composed of text, is prominently displayed.
- Amazon.com Customer Review Average:** A screenshot of an Amazon product page showing the average customer review. The title reads "Amazon.com Customer Review Average:". Below it, "Customer Reviews" and "Monster Hunter Tri" are listed. The review summary shows "49 Reviews" with the following distribution:

Rating	Count
5 star	(32)
4 star	(12)
3 star	(4)
2 star	(0)
1 star	(1)

A "Average Customer Review" section shows a 4.7-star rating based on 49 reviews. Buttons for "Create your own review" and "Share your thoughts with other customers" are also present.

Copyrighted Material  
**A NEW YORK TIMES BUSINESS BESTSELLER**  
"As entertaining and thought-provoking as *The Tipping Point* by Malcolm Gladwell. . . . *The Wisdom of Crowds* ranges far and wide."  
—The Boston Globe

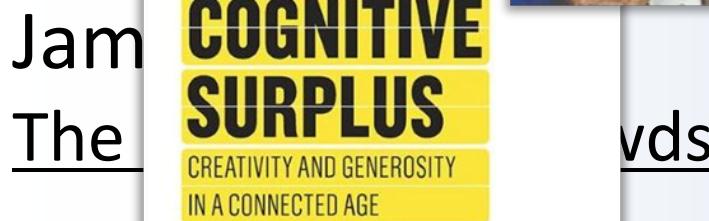
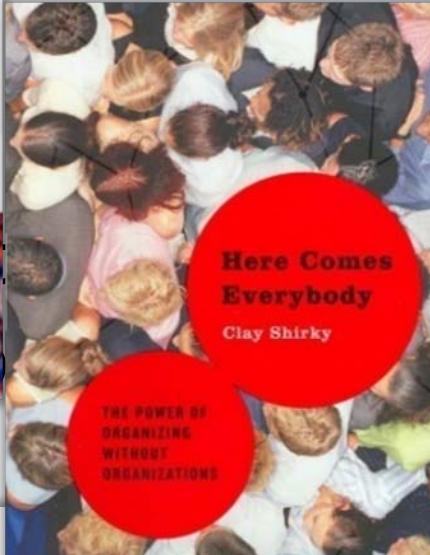
# THE WISDOM OF CROWDS

JAMES  
SUROWIECKI

WITH A NEW AFTERWORD BY THE AUTHOR



"...  
re  
the



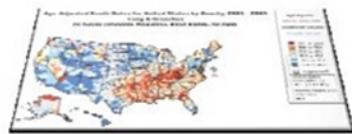
CLAY  
SHIRKY

# The Science of Crowdsourcing

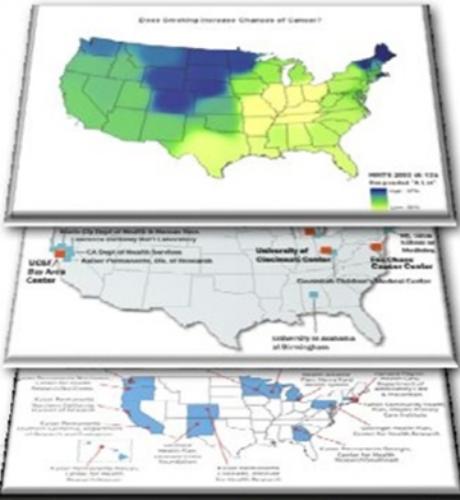
Conditions Necessary  
for  
tances, groups are  
• Independence of  
opinion  
are often smarter than  
m.”  
• Decentralization  
• Mutual goal  
• Diversity  
• Aggregation  
• Incentives (ROI)

# Crowdsourcing Science - Researchers

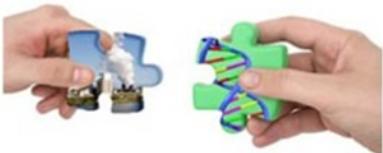
## GRID ENABLED MEASURES



Disease Registry Data (e.g., ePath)



PopSci Grid (caBIG®)



Information Exposure  
(e.g., Health Information National Trends Survey, HINTS)

Toxicant Exposures  
(e.g., Breast Cancer & Environment)

Health Care Environment  
(e.g., Cancer Research Network)



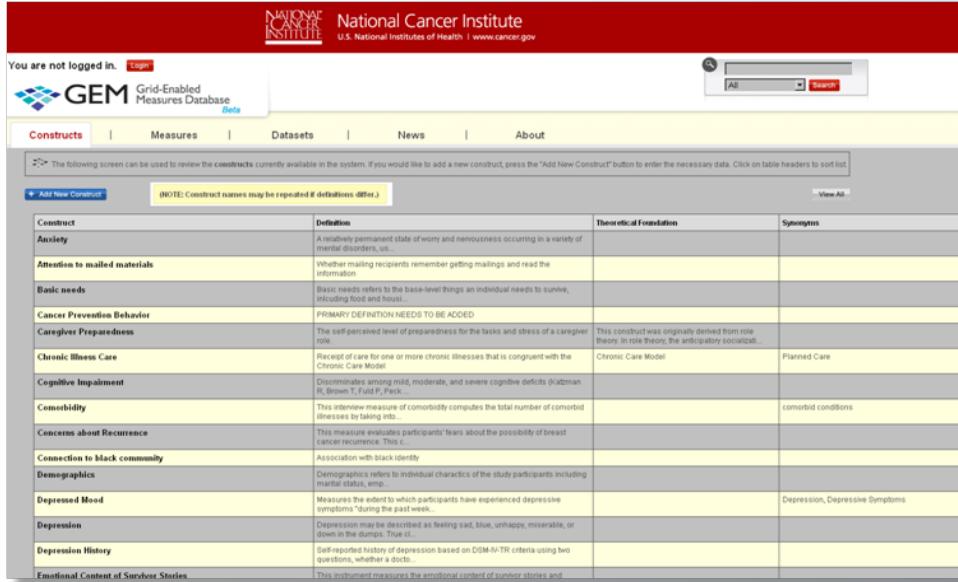
## Benefits

- Cost effective
- Coordinated
- Transparent

## Tensions

- No common vocabulary
- Individual vs. team incentives

# Grid Enabled Measures (GEM): Science 2.0



The screenshot shows the GEM Grid-Enabled Measures Database Beta website. The top navigation bar includes the National Cancer Institute logo and a 'You are not logged in.' link. Below the navigation is a search bar with a dropdown menu. The main content area is titled 'Constructs' and displays a table of constructs with columns for 'Construct', 'Definition', 'Theoretical Foundation', and 'Synonyms'. The table lists various constructs such as Anxiety, Attention to mailed materials, Basic needs, and Cancer Prevention Behavior, each with a detailed description and cross-references to other concepts.

Construct	Definition	Theoretical Foundation	Synonyms
Anxiety	A relatively permanent state of worry and nervousness occurring in a variety of mental disorders, us...		
Attention to mailed materials	Whether mailing recipients remember getting mailings and read the information		
Basic needs	Basic needs refers to the base-level things an individual needs to survive, including food and housing		
Cancer Prevention Behavior	PRIMARY DEFINITION NEEDS TO BE ADDED		
Caregiver Preparedness	The self-perceived level of preparedness for the tasks and stress of a caregiver role	This construct was originally derived from role theory. In role theory, the anticipatory processes...	
Chronic Illness Care	Receipt of care for one or more chronic illnesses that is congruent with the Chronic Care Model	Chronic Care Model	Planned Care
Cognitive Impairment	Discriminates among mild, moderate, and severe cognitive deficits (Fitzman, R, Brown T, Fuld P, Peck...		
Comorbidity	This interview measure of comorbidity computes the total number of comorbid illnesses by taking into...		comorbid conditions
Concerns about Recurrence	This measure evaluates participants' fears about the possibility of breast cancer recurrence. This is...		
Connection to black community	Association with black identity		
Demographics	Demographics refers to individual characteristics of study participants including marital status, ethnicity,...		
Depressed Mood	Measures the extent to which participants have experienced depressive symptoms during the past month		Depression, Depressive Symptoms
Depression	Depression may be described as feeling sad, blue, unhappy, miserable, or down in the dumps. True c...		
Depression History	Self-reported history of depression based on DSM-IV-TR criteria using two questions, whether a doctor...		
Emotional Content of Survivor Stories	This instrument measures the emotional content of survivor stories and...		

## Overall Goals:

To facilitate a virtual community of scientists using collaborative web technology to:

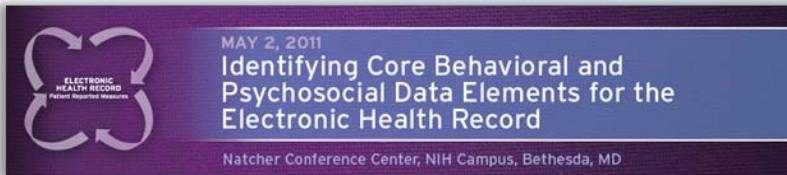
- vet and promote the use of shared measures– based on theoretically-meaningful constructs;
- share the resulting harmonized data.

<https://www.gem-beta.org>

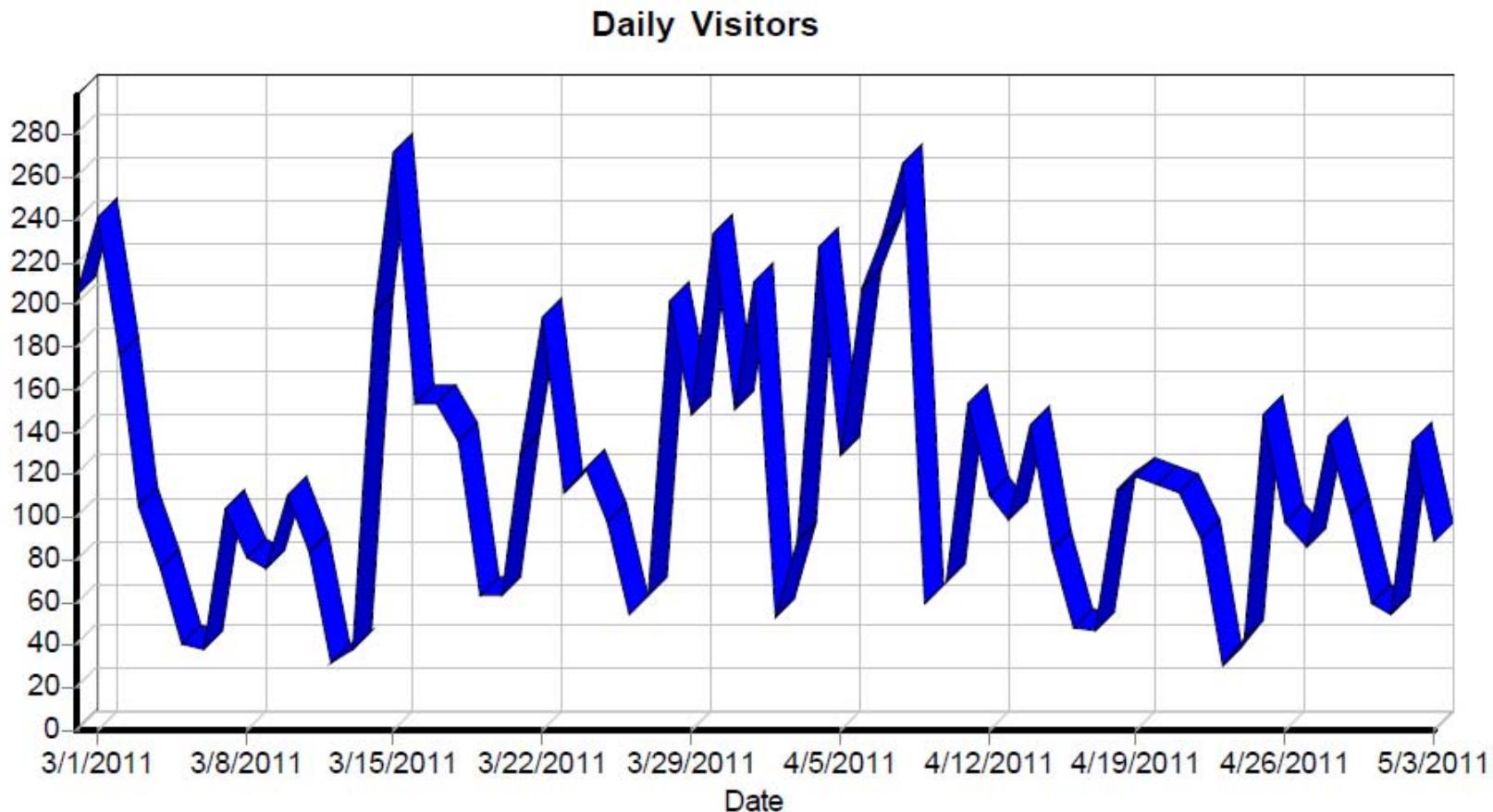
# Electronic Health Records (EHR) Campaign



- December, 2010 to May, 2011
- Goal: To develop consensus on common data elements for patient-reported health behaviors and psychosocial factors for EHRs
- Assessed 13 constructs
- Consensus meeting held at NIH May 2-3, 2011

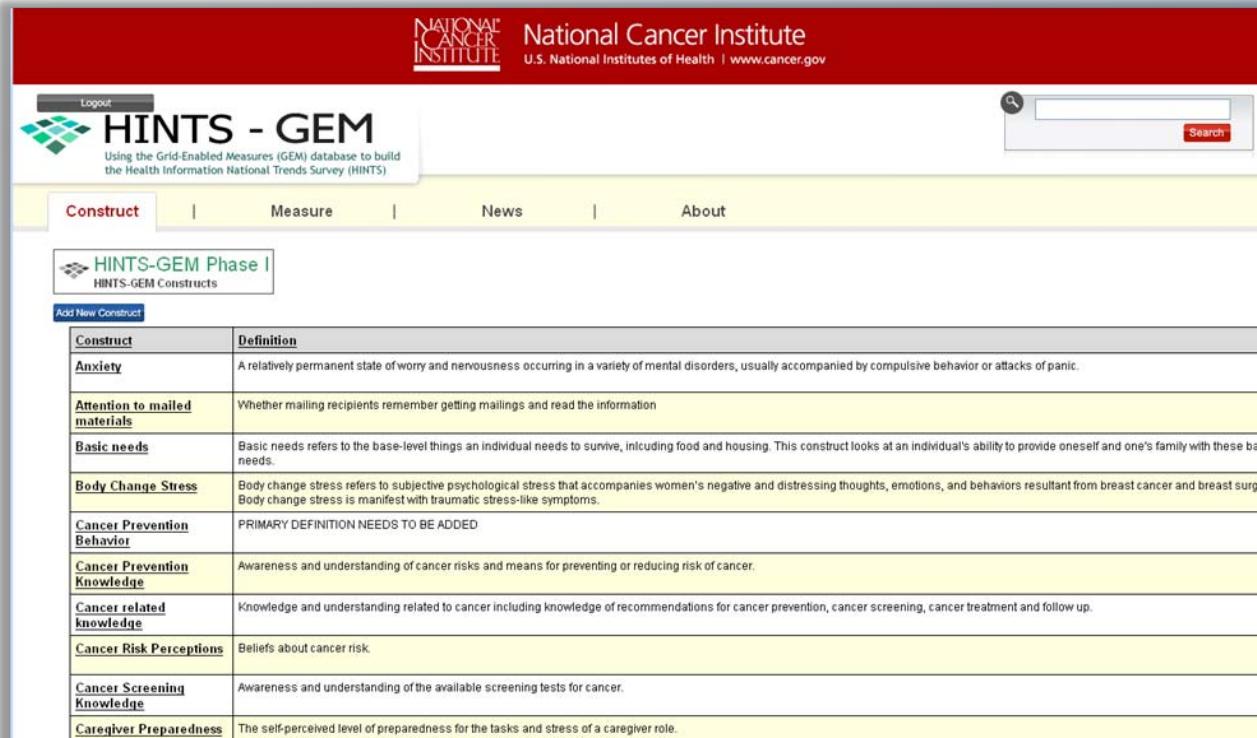


# Participation in EHR Campaign



- **Totals:**
  - 1,000+ registered
  - 296 comments (from 98 individuals)
  - 603 ratings (from 60 individuals)

# HINTS-GEM: Research Community Solicitation/Vetting of HINTS Survey Items



**HINTS - GEM**  
Using the Grid-Enabled Measures (GEM) database to build  
the Health Information National Trends Survey (HINTS)

Construct	Definition
<u>Anxiety</u>	A relatively permanent state of worry and nervousness occurring in a variety of mental disorders, usually accompanied by compulsive behavior or attacks of panic.
<u>Attention to mailed materials</u>	Whether mailing recipients remember getting mailings and read the information
<u>Basic needs</u>	Basic needs refers to the base-level things an individual needs to survive, including food and housing. This construct looks at an individual's ability to provide oneself and one's family with these basic needs.
<u>Body Change Stress</u>	Body change stress refers to subjective psychological stress that accompanies women's negative and distressing thoughts, emotions, and behaviors resultant from breast cancer and breast surgery. Body change stress is manifest with traumatic stress-like symptoms.
<u>Cancer Prevention Behavior</u>	PRIMARY DEFINITION NEEDS TO BE ADDED
<u>Cancer Prevention Knowledge</u>	Awareness and understanding of cancer risks and means for preventing or reducing risk of cancer.
<u>Cancer related knowledge</u>	Knowledge and understanding related to cancer including knowledge of recommendations for cancer prevention, cancer screening, cancer treatment and follow up.
<u>Cancer Risk Perceptions</u>	Beliefs about cancer risk.
<u>Cancer Screening Knowledge</u>	Awareness and understanding of the available screening tests for cancer.
<u>Caregiver Preparedness</u>	The self-perceived level of preparedness for the tasks and stress of a caregiver role.



## HINTS - GEM

Using the Grid-Enabled Measures (GEM) database to build  
the Health Information National Trends Survey (HINTS)

<https://secure.mmgct.com/hints-gem/>

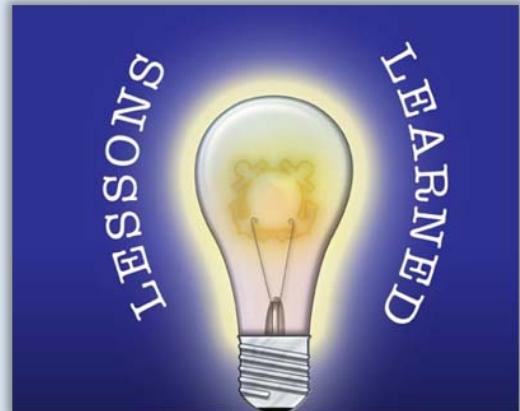
# HINTS GEM Results: Measures



- 1173 measures entered across 85 constructs
  - 526 measures from previous iterations
  - 647 new measures proposed
  - 60 alternative measures
- Final item pool submitted to OMB

Measure Disposition	Measures from Previous HINTS Iterations (n=526)	Measures Newly Proposed for HINTS 4 (n=647)
Recommended for inclusion	37.6%	41.7%
Recommended for exclusion	36.5%	6.3%
Under consideration	25.9%	51.9%

# GEM: Lessons Learned



- Small number of transgressions
- Most (public) work done by relatively few
- Infusion of new ideas
- Need to address barriers to participation



# Crowd-Sourcing Science: Mass Participation in Knowledge Generation from the Public

The image displays the Asthmapolis website, a red and white inhaler, and a map of Maryland pollution.

**Asthmapolis Website:**

- Header:** ASTHMAPOLIS
- Navigation:** About, How It Works, Blog, Press, Contact, Careers
- Section 1:** Tools to track, manage and research asthma. Includes a red and white inhaler image and a "Learn More" button.
- Section 2:** Why use Asthmapolis? Includes a video thumbnail of a person using an inhaler.
- Text:** Experts recommend that people with asthma track their symptoms, triggers and use of asthma.

**Map of Maryland Pollution:**

**Title:** Maryland Pollution

**Section:** Types of Pollutants

- Acid Mine Drainage
- Air Releases
- Arsenic
- Bio-Toxin Disposal Site
- Carbon Dioxide Emissions
- Chemical Agent Dump
- Effluents
- Hazardous Waste Handlers
- Landfill
- Nitrates
- Radon
- Sulfur Dioxide
- Saltwater Intrusion Zone

**Text:** map produced by Nick Arnold, 2009  
source: Environmental Protection Agency

# Crowdsourcing Science - Public

Army of women<sup>®</sup>  
LOVE/AVON Going Beyond a Cure  
a program of the Dr. Susan Love Research Foundation

Join the Army   Volunteer

Contact Us | Site Map

OUR GOAL OF ONE MILLION WOMEN WITH ONE VOLUNTEER

03590

SIGN UP

ONE MILLION WOMEN. RESEARCH GOAL. REVOLUTIONARY OPPORTUNITY.

Together we can prevent breast cancer.

ANNOUNCING THE LOVE/AVON ARMY OF WOMEN PUBLIC SERVICE ANNOUNCEMENT WINNER [READ MORE](#)

Click on thumbnails to watch video

About Us   Sign Up Today   Login   For Researchers   Army Town Hall   News & Events

Invite a Friend   Spread The Word

AOW Mission and Vision   AOW Current Projects   AOW Researchers   AOW Blog Updates   AOW Twitter Updates   AOW News and Events

331,000 women recruited since mid 2009;  
Goal is 1,000,000 women

## Benefits

- Reinvigorate public interest in science
- Overcome stumbling blocks to discovery (e.g., recruitment)
- Improve “n” / power

## Tensions

- Need to update sampling statistics
- Coordination needed between advocacy groups and science
- Quality Control

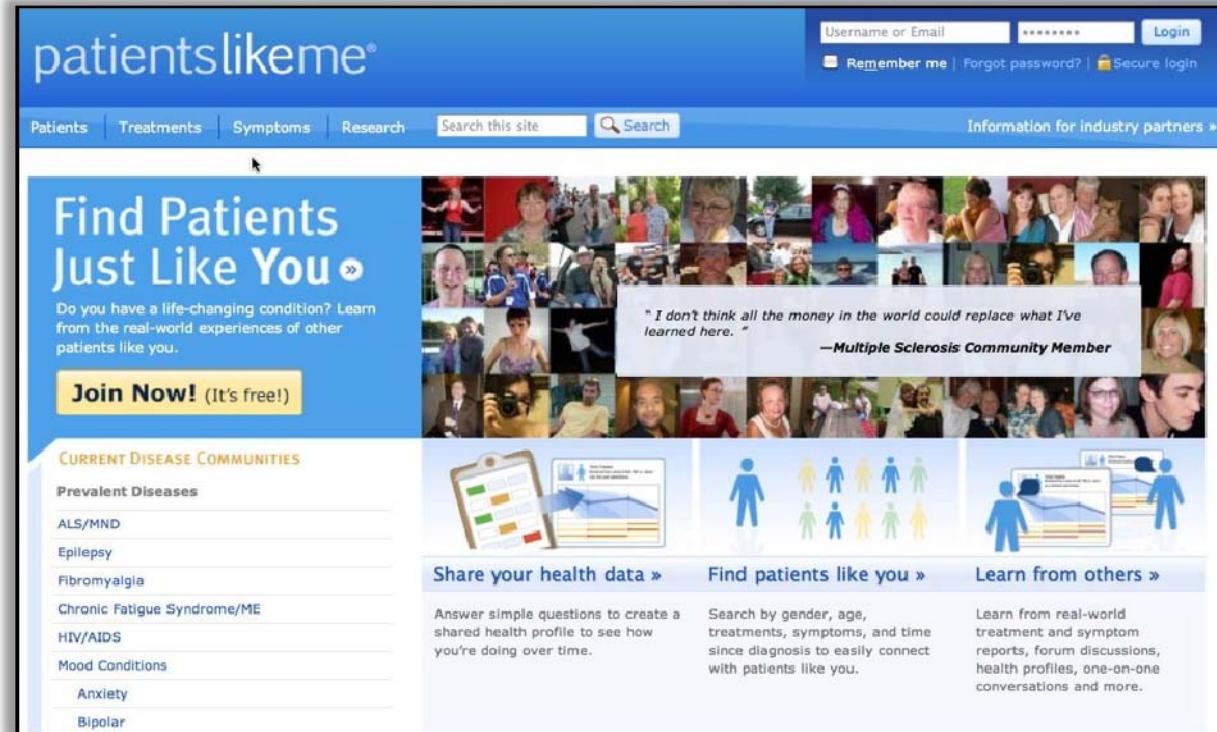
# “Citizen Science”: Patients Offering Personal Data in a Disease-Oriented Commons

## Benefits

- Improves likelihood of detecting side effects
- Permits social comparison
- May facilitate self-regulation processes

## Tensions

- Need to renegotiate boundaries between private sector, public sector, and citizens
- New rules needed for rigor & control



patientslikeme®

Username or Email  Login

Remember me | Forgot password? | Secure login

Information for industry partners »

Patients | Treatments | Symptoms | Research | Search this site | Search

## Find Patients Just Like You

Do you have a life-changing condition? Learn from the real-world experiences of other patients like you.

**Join Now!** (It's free!)

CURRENT DISEASE COMMUNITIES

- Prevalent Diseases
- ALS/MND
- Epilepsy
- Fibromyalgia
- Chronic Fatigue Syndrome/ME
- HIV/AIDS
- Mood Conditions
- Anxiety
- Bipolar

Share your health data »

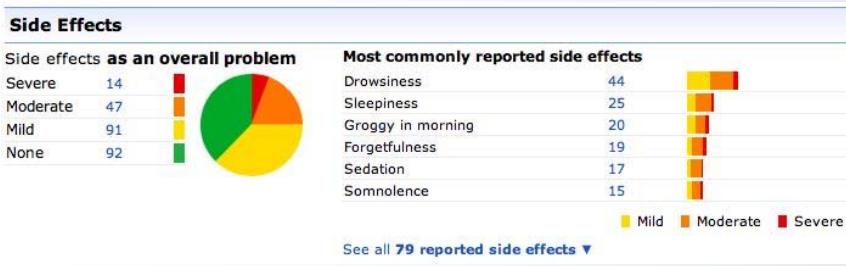
Find patients like you »

Learn from others »

Answer simple questions to create a shared health profile to see how you're doing over time.

Search by gender, age, treatments, symptoms, and time since diagnosis to easily connect with patients like you.

Learn from real-world treatment and symptom reports, forum discussions, health profiles, one-on-one conversations and more.

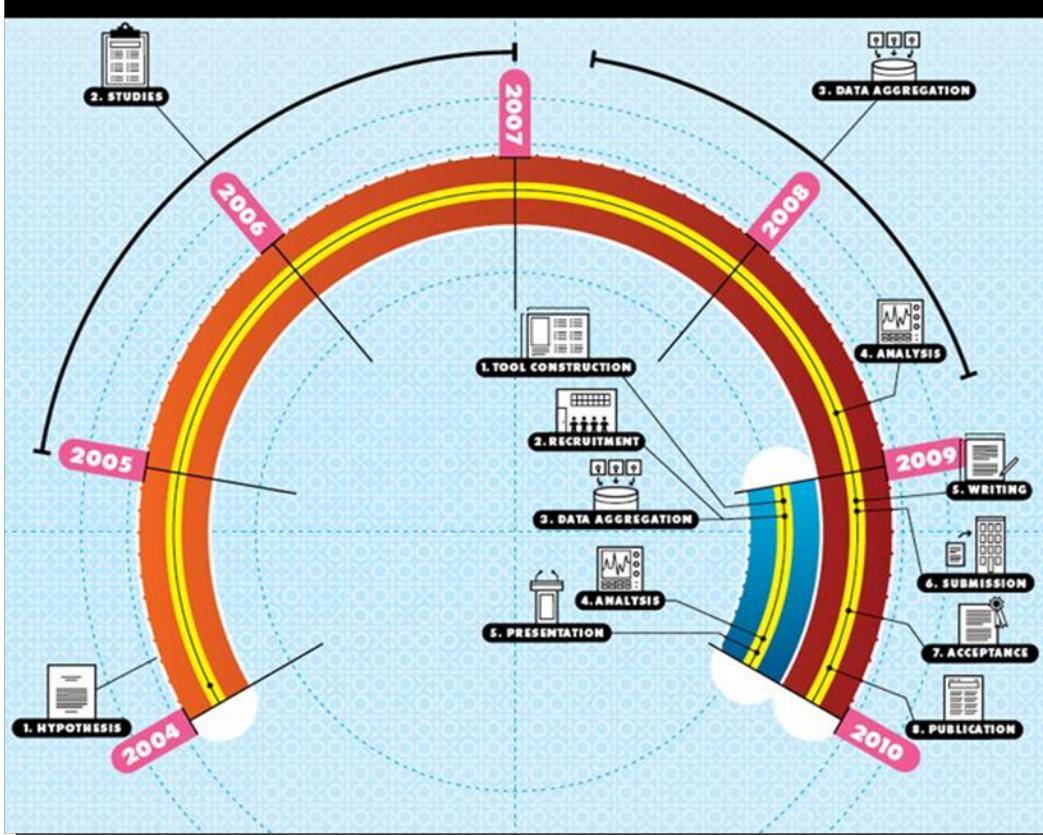


# Traditional vs. “New” Science

# WIRED

## High-Speed Science

Can a model fueled by data sets and computational power compete with the gold standard of research? Maybe: Here are two timelines—one from an esteemed traditional research project run by the NIH, the other from the 23andMe Parkinson’s Genetics Initiative. They reached almost the same conclusion about a possible association between Gaucher’s disease and Parkinson’s disease, but the 23andMe project took a fraction of the time.—*Rachel Swaby*



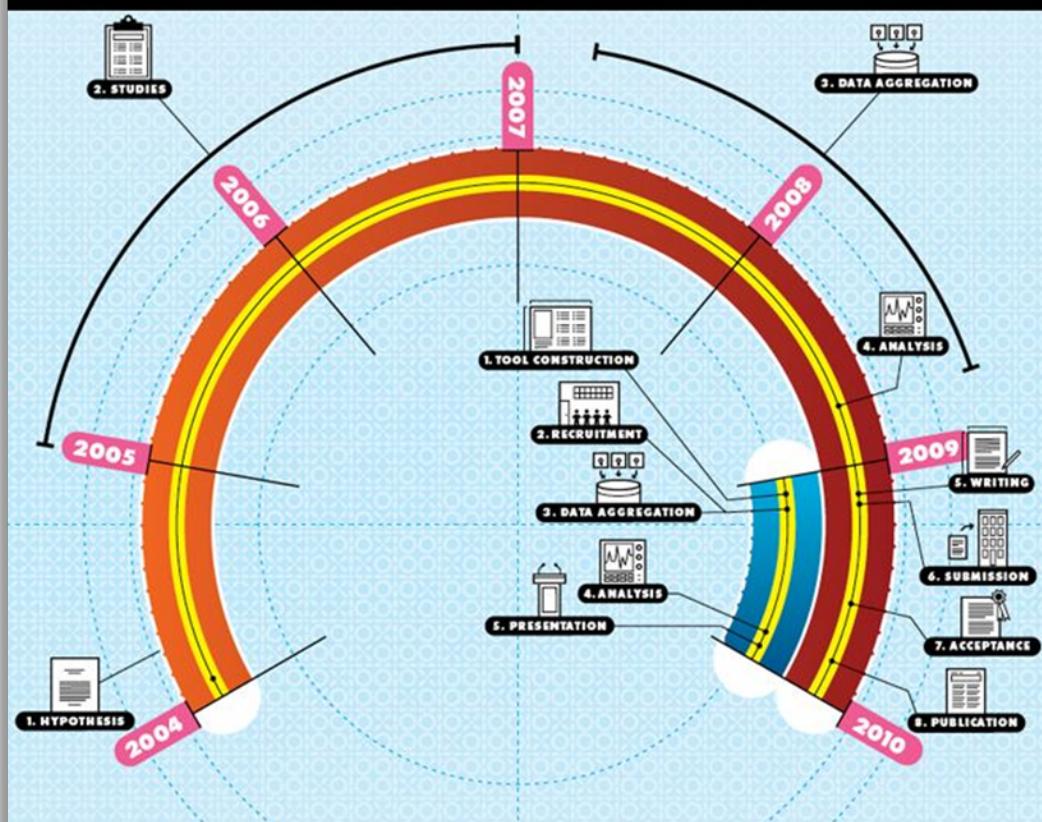
## Traditional Model

- Hypothesis
- Studies
- Data aggregation
- Analysis
- Writing
- Submission
- Acceptance (NEJM)
- Publication

**Time Elapsed: 6 years**

## High-Speed Science

Can a model fueled by data sets and computational power compete with the gold standard of research? Maybe: Here are two timelines—one from an esteemed traditional research project run by the NIH, the other from the 23andMe Parkinson's Genetics Initiative. They reached almost the same conclusion about a possible association between Gaucher's disease and Parkinson's disease, but the 23andMe project took a fraction of the time.—*Rachel Swaby*



## Genetics Initiative

- Tool construction
- Recruitment
- Analysis
- Presentation

Time Elapsed: 8 months



# What's Next?

# Is Science at the Table?

Science?

X



# Is Science at the Table? Proceed (with Caution)



- Emphasizing technology over science
- Finding the next 'killer app'
- No one size fits all
- Sampling issues/generalizability
- Need evidence-based applications

Nov, O., Arazy, O., and Anderson, D. (2011). Technology-Mediated Citizen Science Participation: A Motivational Model. *Proceedings of the AAAI International Conference on Weblogs and Social Media (ICWSM 2011)*. Barcelona, Spain, July 2011.

*Journal of Health Communication*, 16:10–31, 2011  
ISSN: 1081-0730 print/1087-0415 online  
DOI: 10.1080/10810730.2011.589882



## Realizing the Promise of Web 2.0: Engaging Community Intelligence

BRADFORD W. HESSE, MARY O'CONNELL,  
ERIK M. AUGUSTSON, WEN-YING SYLVIA CHOU,  
AND ABDUL R. SHAIKH

National Cancer Institute, National Institutes of Health, Bethesda,  
Maryland, USA

LILA J. FINNEY RUTTEN

Clinical Monitoring Research Program, SAIC-Frederick, Inc.,  
NCI-Frederick, Frederick, Maryland, USA

We pro  
contribu  
projects  
Compa  
in a w  
with ne  
volunte  
that p  
motivat  
intrinsic  
wherea  
Intrinsic  
were m  
which

*Discussions of Health 2.0, a term first coined in 2005, were guided by three main tenets: (a) health was to involve more participation, because an evolution in the*

# Crowdsourcing: What does the Future Hold?

Wednesday, June 24, 2009

## Big Brother crowd-sourcing

From NYT's [The Lede](#):

Update | 5:31 p.m. A reader, Jon, who said that he had been in touch with a "person in the Iranian government" told him that a specific Web site run by the Iranian government to identify protesters has been unable to get the site to load because it is under attack from supporters of the protesters. The reader asked to study photographs of protesters and then turn them in to the authorities.

Later: A reader points out that Gerd, running at least one page of photographs of faces of 25 protesters circled, and an unknown number of people.

