Pulse
Pulse

The New Science of Harnessing Internet Buzz to Track Threats and Opportunities

DOUGLAS W. HUBBARD
My dedication for my third book is very similar to my earlier books. I am entirely motivated to write for and by my family: Janet, Evan, Madeleine, and Steven.

And once again, I would also like to dedicate this to my personal friends and colleagues in all branches of the U.S. Military for all their sacrifices.
Contents

Preface xi
Acknowledgments xiii

PART 1 INTRODUCTION TO THE PULSE: A NEW KIND OF INSTRUMENT

CHAPTER 1 A New Era for Measuring and Predicting Society 3
An Emerging Science: What Is This New Thing? 6
What The Pulse Isn't 8
The Major Areas of the Macro-Pulse 12
What Does All of This Mean for Us? 14
Notes 15

CHAPTER 2 The History of Seeing the Forest through the Trees 17
The King's Surveys 18
The Dawn of Stats, Theme-Maps, and Telegraphs 20
The Rise of the Machines 25
The Struggle to Become a Science 28
Notes 31

CHAPTER 3 Emergence of the Pulse and the New Research Discipline 33
Ubiquitous Use of the Internet 34
Digital Lives: Increasing Time and Activities per Person 39
Open Pastures: A New Field of Research 45
##Contents

Proving the Pulse: Addressing Misconceptions about the Data  
Notes  

CHAPTER 4  
Dynamics of the Pulse  
Incentives: Why the Pulse Exists  
The Systems behind Getting and Sharing Data  
Collaboration and Competition  
Power Law: Why a Few Sources Tell Us a Lot on the Internet  
Notes  

PART II  SOURCES OF THE PULSE  

CHAPTER 5  
What Our Surfing Says  
Tracking Flu Outbreaks: A Faster, Better, and Cheaper Method  
A Do-It-Yourself Pulse Tracking Example: Using Google Trends for Economic Predictions  
More Searchology: The Body of Research Grows  
Caveats and Miscellaneous Considerations for Searchologists  
Notes  

CHAPTER 6  “Friend” as a Verb  
How Do Social Networks Matter?  
The Dynamic Duo of Social Connections: Two Leaders in Practical Network Research  
Forecasting the Crowd  
Social Networks in the Pulse  
Notes  

CHAPTER 7  What We Say Online Matters  
An Introduction to Analyzing Buzz: Counting Tweets Predicts Movie Box Office Receipts  
Predicting the Broader Economy with Tweets  

## Contents

Predicting Markets with Anxiety ................................................................. 117  
Tools and Miscellanea for Tapping into the Global Mood ........................... 119  
Notes .............................................................................................................. 123

### CHAPTER 8
Three Potential Pulses: Where We Go, What We Buy, and How We Play .......... 125  
Our Flow and the Pulse: What the Movements of Millions of People Tell Us .... 126  
The Shopping Pulse ...................................................................................... 131  
Playtime and the Pulse .............................................................................. 136  
Notes .............................................................................................................. 140

### PART III  EFFECTS OF THE PULSE

### CHAPTER 9
Making the Pulse Practical ........................................................................... 143  
Rethinking Real-Time Decisions in the Pulse .............................................. 144  
A Brief Overview of the Economics of Timely Information .......................... 149  
Overcoming Cultural and Conceptual Obstacles: Maximizing the Value of the Pulse 154  
Implementing the Pulse ................................................................................ 156  
Notes .............................................................................................................. 159

### CHAPTER 10
The Future of the Pulse and Its New Challenges .......................................... 161  
Users and the Data They Share Will Increase ............................................ 162  
Services Will Be Better and Make More Data Public ................................... 165  
Research and Models ................................................................................ 169  
From Communicators to Tricorders—for Everyone ....................................... 171  
Evolving Opportunities and Challenges for Users of the Pulse ................. 173  
Notes .............................................................................................................. 177

**About the Author** ....................................................................................... 179

**Index** ....................................................................................................... 181
I first started thinking about a book on this topic when I finished writing the first edition of my first book—*How to Measure Anything: Finding the Value of Intangibles in Business* (John Wiley & Sons)—in late 2006. One chapter was titled “New Measurement Instruments for Management,” where I talked about the possibility of farming data from the Internet for practical business measurements. I wondered about such possibilities as using data about home sales published on sites like realtor.com or eBay to inform us—on a real-time basis—about that part of the economy.

Sometime in 2007 I learned about the tool Google Trends—a tool which shows how frequently certain search terms are used in Google. After the darkest days of the financial crisis in 2008, I became more interested in seeing if financial trends like this could have shown up in our search behavior on the Internet. At one point I typed “unemployment” into the Google Trends tool and compared it to the official unemployment figures from the Bureau of Labor Statistics. The patterns matched very closely, but what was really amazing was that the Google Trends data was available a month before the official data came out. I was hooked and started searching for more interesting ways that this free data source from Google could track useful economic trends—only much faster than the official economic reports.

Early in 2009, I started finding published research that had already been done on how Internet search data could be applied to forecasting economic trends and flu outbreaks. It seemed as if the frequency of articles like this published in respected, peer-reviewed economic journals was speeding up, and then I noticed that the popular science magazines *New Scientist* and *Science News* were beginning to report this sort of research about once a month or so. It looked like a critical mass was approaching.

Just within the last 24 months leading up to the publication of this book, the research really seemed to be accelerating, and more conferences were being held on the topic of using the Internet as a research tool. Finally, I told my publisher that I felt I had to write this book before someone else did. No one source discussed the recent research in all the different areas using data available on the Internet about our searches, relationships, and daily activities.
As you would expect, different researchers specialized in different aspects of this emerging field. I also found that many researchers were not even aware of work based on different methods of using the Internet. People who were experts at analyzing the text on Twitter to predict macro-trends were generally not aware of the research using location-based services on cell phones, while those experts, in turn, were not aware of research using tools like Google Trends or Yahoo!

This new field could be served, I thought, by identifying its multiple different components and how they are each used to track big trends. I divided up the kinds of tools along the lines of how we spend our time on the Internet: We search for information, we connect with others, we post our thoughts, we track and share our location, we buy things, and we play games. These are somewhat overlapping since some involve similar technologies, and they sometimes forecast or track similar things.

One of my biggest challenges in writing this book was the same as my other books: What do we call this new thing? One existing name for this field of research—computational social science—sometimes seems to be more specifically associated with social networks, which is only part of what I’m writing about in this book. I think that label is evolving and may eventually be the generally accepted term for all the areas of research I am writing about. In the meantime, I like the metaphor of a pulse. It is as if the combined system of the Internet and the people using it are a kind of organism. It has a rhythm, and if we can track that rhythm, we will see what was previously invisible to psychologists, economists, epidemiologists, law enforcement, linguists, anthropologists, and businesses. This mass of data from what is now about 2 billion Internet users will revolutionize these fields. This new measurement instrument will have as profound an impact for the social sciences as did the microscope for biology, the telescope for astronomy, or the radar and the satellite for meteorology.
Acknowledgments

I would like to thank the following people for helping make this a better book. Some are researchers who have taken the time to speak with me about their important research. Others are colleagues inside and out of this field who have given valuable input to this work on issues related to technology. Some are my analysts who have helped compile a lot of data. Thanks.

Gunther Eysenbach, University of Toronto
Bryan Routledge, Carnegie Mellon University
Brendan O’Connor, Carnegie Mellon University
Rammnath Balasubramanyan, Carnegie Mellon University
Noah Smith, Carnegie Mellon University
Mark Barros, Metric Junkie
Christina Carbone, European Commission Joint Research Center
Marshal Van Alstyne, Boston University
Nathan Eagle, txteagle Inc. and MIT Media Lab
Johan Bollen, Indiana University–Bloomington
Bing Pan, College of Charleston
Gregor Orelind, Alexa
Angela Lipscomb, SAS
Jeff Gilleland, SAS
John Bastone, SAS
Michael Bergines, eBay
Neel Sundaresan, eBay
Alan Mislove, Northeastern University

And my research assistants:
Michael North, Ashley Hubbard, and Joseph Hubbard