

What Millions
of People Do
Online and Why
it Matters

click

Unexpected Insights for Business and Life

BILL TANCER

"BILL TANCER IS THE KING OF MEASURING ONLINE RESEARCH,
AND ONLINE RESEARCH IS THE MAIN STREET OF THE NEW WORLD,
WHICH MAKES BILL TANCER KING OF THE WORLD, OR SOMETHING LIKE THAT."

--STEPHEN J. DUBNER, co-author, *FREAKONOMICS*

click

**What Millions
of People
Are Doing Online
and Why It Matters**

BILL TANCER



For Lori

CONTENTS

Acknowledgments vii

Introduction 1

PART I: UNDERSTANDING OURSELVES 11

Chapter 1: PPC—Porn, Pills, and Casinos 13

Our fascination with skin, quick fixes, and easy money. Actions speak louder than words, as aggregate interaction with the underbelly of the Internet reveals patterns that aren't easy to ascertain from traditional sources.

Chapter 2: Getting to What We Really Think 33

Do platform stand, voting record, and integrity determine whom we vote for? Search-term data reveals a lot about a candidate's brand and the half-life of negative information. See how these insights may innovate the way we measure brand in the business world.

Chapter 3: Prom in January 49

Why do searches for "prom dresses" peak the first week of January? Internet searches reveal that our gut instincts regarding seasonality are often wrong, resulting in market inefficiencies. We have much to learn from consumers; we just have to observe collective search patterns.

Chapter 4: Failed Resolutions and the False Hope Syndrome 69

Our commitment to New Year change is surprisingly short-lived, with "diet" searches lasting less than a week. Outrageous claims as to how we can improve ourselves with no effort lead ultimately to the downward spiral of false hope.

Chapter 5: Celebrity Worship Syndrome 87

The Internet and specifically celebrity blogs have brought unprecedented public access to the lives of celebrities, fueling our obsession with the famous.

*Chapter 6: What Are You Afraid Of?
and Other Telling Questions* 101

We search on information about our fears—over a thousand unique ones—from fear of public speaking to fear of elbows and ceiling fans. Search engines have become a new nonjudgmental place for us to ask questions we are increasingly less likely to ask each other.

Chapter 7: Web Who.0 119

Is the 80/20 rule passé? With all the hype around Web 2.0, surprisingly few Internet users actively create consumer-generated media, giving rise to the new 1-9-90 rule.

PART II: WHAT'S POSSIBLE WITH WHAT WE KNOW 139

*Chapter 8: Data Rocks and the
Television-Internet Connection* 141

Laptops have become increasingly popular in the living room. Our Internet behavior reveals how we react online to what we see on the tube. As the television-Internet gap closes, we learn what motivates us to interact.

*Chapter 9: Women Wrestlers and
Arbitrating Financial Markets* 155

Does volume of search terms translate to popularity? Is it possible to predict reality television show votes from how Internet users search? Near real-time Internet data provides a time advantage over traditional leading economic indicators.

Chapter 10: Finding the Early Adopters 171

New technology spreads through society in predictable segments starting with Innovators and Early Adopters. Internet behavior viewed in the correct light can help illuminate who the Early Adopters are and what they're doing today.

*Chapter 11: Super-Connectors and
Predicting the Next Rock Star* 185

Visitors to official band websites traditionally come from either social networks or search engines; graphing these two sources of traffic allows us to visualize Malcolm Gladwell's "tipping point."

Epilogue: Who We Are and Why It Matters 199

Notes 205

Glossary 211

Index 213

About the Author

Credits

Cover

Copyright

ACKNOWLEDGMENTS

If you looked around our house, you'd quickly discover my obsession with books. Much to my wife's consternation (although she has her own collection), there are books everywhere— on the bookshelves where they belong, on the nightstand, in the kitchen, stacked in the corners on every available surface. I love books, but the process of writing *Click* has made me realize that in my voracious consumption of printed text, I've really taken books for granted. I had no idea just how many talented individuals are required to take something from concept to finished product. It takes a village to write a book.

Click is based on my work at Hitwise. It's hard for me to think of Hitwise as a company; I consider my colleagues to be part of my extended family, spread over four continents. I was very fortunate to have Tessa Court as my internal champion at Hitwise, but equally fortunate to work with Andrew Walsh and Chris Maher in making this book a reality. I owe deep gratitude to Hitwise analysts, both past and present: Heather Hopkins, who inspired the “tipping point” chapter; LeeAnn Prescott, who first discovered the prom dress phenomenon; as well as Robin Goad, Heather Dougherty, Sandra Hanchard,

and Eva Stringleman, who served as never-ending sources of inspiration; and a special thanks to our head of media relations, Matt Tatham, for helping to get the word out.

This project would never have left the concept stage had it not been for my good fortune in landing one of the most talented literary agents on the planet, Melissa Flashman with Trident Media Group. Melissa's excitement about *Click* was both electric and contagious. As a first-time author, I hit the jackpot in landing Mel as my agent.

I knew we had found the right home for *Click* at Hyperion. Many thanks to Bob Miller and Ellen Archer for enthusiastically embracing the book's concept and championing what I had to say. Thanks to Gretchen Young for her thoughtful and considerate work in guiding me in the right direction, to Beth Gebhard and Alex Ramstrum for lining up amazing publicity, and to Jane Comins and Sarah Rucker and her team. Rick Willett, who provided copyedits for *Click*, deserves special recognition for the tedium of correcting my grammar and spelling and turning the manuscript into something readable. Thanks also to Elizabeth Sabo, who served as my lifeline to the Hyperion team.

Of course there are the friends who stood by me throughout the book-writing process and those whom I met along the way. Thanks to my office neighbor and lunch buddy Luke McGuiness, who was always a willing sounding board. I credit Rafael Zorrilla with coming up with the chapter title for the book's first chapter. PPC means something in online marketing-speak, but Rafael has his own definition, which I borrowed with his permission. Thanks also to Daisy Whitney for her inspiration and comic relief in the book-writing process, and thanks to the Roosters (Bill, Jim, John, Bruce, Norm, Tim, and Gregg), my Saturday bike group, for letting me ride in the back of the group so that I could think about what I was going to write next.

I can't forget the wonderful support of my family. To my

parents, Marty and Sheila Tancer, who always supported and encouraged me in all of my endeavors, and to my mother-in-law, Paulette Minden, whose optimism for *Click* was a constant source of inspiration.

Finally, as I finished *Click*, it occurred to me that there really should be a support group for spouses/significant others of writers. My amazing wife, Lori, to whom I've dedicated this book, didn't flinch when I sat her down and announced that I was going to write a book (to appreciate this you'd have to consider all the previous half-baked ideas I had announced in the past). Throughout the writing process, Lori has been there for me, has put up with the moodiness that can go along with writing, not to mention the endless conversations and questions about prom dresses, and when needed has served as drill sergeant when I was just too tired to continue. Thanks for always being there, and for giving me the inspiration and support that helped make *Click* a reality.

INTRODUCTION

On my way in to work in San Francisco from our house in San Mateo, something about the radio interview I was listening to just didn't sound right. The reporter was talking to a British psychologist about the third week in January, a week that he claimed was the most depressing week of the year. The interviewee detailed how he had reached that conclusion from a mathematical formula he'd developed that factored in failed New Year's resolutions, credit card debt from the holidays, and weather patterns. All factors taken together pointed to the third week in January as the most depressing of the year. It sounded like a convincing story; I just had a hard time believing it.

While I'm all in favor of mathematical formulas, was it possible to build one that would detect a society's swing toward depression? You probably couldn't even detect such a thing by surveying a large sample of the population—can we count on people to answer questions about depression honestly? No, I was having trouble buying this story, because I had just charted “depression” a few weeks earlier, and I knew that the most depressing week of the year occurred right around Thanksgiving. How did I know that?

Pulling from a sample of 10 million Internet users in the United States, I was interested in seeing when people search for the term “depression.” My thought process was as follows: If as a society we become more depressed during certain parts of the year, a certain percentage of us, in order to seek help, will turn to our nearest confidant, a place we can go to and ask questions without fear of being judged. We will search on Google, Yahoo! Search, or MSN Search.

My first chart of the term, over a two-year period of time, didn't quite add up. It showed a pattern that seemed to reflect the school calendar. This was probably due to the fact that school-children were using search engines for research papers about the Great Depression. Not to be deterred, I created a new chart, this time combining the volume of searches for common antidepressants, drugs like Zoloft, Cymbalta, Lexapro. The pattern was pretty clear, once you ruled out the Great Depression and factored in pharmaceuticals: A significant number of us turned to our search engines the week of Thanksgiving in hopes of understanding our holiday blues.

This morning's drive was pretty typical for me. I hear things on the news, read items in the newspaper, or catch snippets of discussions among coworkers, and I have a burning urge to run to my computer and see if I can shed light on the topic du jour based on our Internet searches or visits to particular websites. Sometimes random associations just pop into my head and I turn to search-term data to see if there is any relationship. These ventures into data serendipity have led to some very interesting questions, from why prom dresses spike in January to the connection between women wrestlers and the arbitrage of financial indicators. What day of the week is the most popular for visits to pornographic websites? What scares us the most? What questions are on our minds? How are new Web 2.0 sites creating society's new winners and losers?

The answers to these questions and many more like them

can be found in a massive database of Internet behavior. I spend my days (and many of my nights) deep in the recesses of this data. When I can't sleep, I find myself sifting through traffic charts, tables of searches, demographics, and psychographics, turning the different pieces of data around in my head, examining them from different angles.

What struck me at first as just a lot of data is actually something much greater. If you spend enough time considering all the different elements of our Internet use, you begin to compile an ever clearer picture of who we are and what exactly is on our minds. The "who" that I'm referring to is a sample of more than 10 million Internet users in the United States, and the database is an anonymized and aggregate collection of what this massive group of Internet users do online every day. The "what" is a company called Hitwise. I work as the general manager of global research for the company.

Sometimes I have a hard time remembering exactly what I did before working for Hitwise. But I do remember how one phone call, a sales call from Chris Maher, who at the time was the general manager of the U.S. market for the company, and was later to become the company's president, changed the course of my career. This one call would drive me to switch to a job that provided me with such rich data, I would quickly find myself in the executive suites of some leading companies in the United States and abroad, and I would be invited to lecture at top universities, give keynotes at industry events, author a weekly online column for *TIME*, and eventually write this book. Of course I didn't know that my ringing phone would lead to such a major life change; at the time I was annoyed that my concentration had been broken by the telephone ringing on my desk.

I was busy working for the head of sales and business development at my current employer, LookSmart, a search engine that employed a large team of editors to inject human input into

search results. As the head of market research, I was troubled by one massive project that had landed in my lap.

As a search engine, LookSmart derived its revenue from advertisers who wanted their advertisements to show up when consumers entered a particular word in our search box. So if I were in the business of selling Waterford crystal, when a consumer searched for this particular term or maybe just the term “crystal,” I would want an ad for my online crystal store to show up within the listings, with the hope that a searcher might click on it, and eventually buy crystal from my online shop. To accomplish this, I would enter into an agreement with LookSmart to pay a small fee, maybe \$0.10, when anyone clicked on my Waterford crystal ad on the LookSmart search results page.

From a sales perspective, there was a big problem with this business model. The LookSmart sales force (or really, the sales force of any search engine at the time) had no clue how much inventory they had of searches for a particular phrase, such as “Waterford crystal” or “crystal.” This problem meant that our sales force was essentially selling blind. They could waste days, weeks, or months selling a search campaign to an advertiser, only to learn that nobody was actually searching for the terms that the advertiser was interested in.

As we sat in the café portion of the LookSmart building, an old converted warehouse on Second Street in San Francisco (big open-beamed space reminiscent of the Dot Com era), Robert Goldberg, who headed up sales and business development for LookSmart at the time, explained the problem this way: “I have a sales team that’s selling inventory from a warehouse, but there’s no reporting system to tell me what’s in the warehouse; there’s not even a door or window into this warehouse—we’re just forced to guess what’s inside.” Imagine running a sales organization with absolutely no visibility into what you are selling. The potential for inefficiency is enormous.

We did have one set of data, which Robert had delivered to

my desk on a CD-ROM (the file was huge): a sample of our biggest set of inventory at the time, one week's worth of searches on Microsoft's MSN Search engine. This one slice of search activity, for one week out of one year, contained a staggering 40 million search terms. Somehow from this massive file of terms, I had to find a way to estimate what our inventory of retail terms was—how many searches there were for different models of cars, what people typed into a search engine when they were trying to navigate to an online dating service, and on and on. An hour had gone by and I hadn't moved from my computer screen, scrolling through a sample of the top ten thousand terms, which I had downloaded into Excel.

The call from Hitwise came at what I thought was the worst possible time, but it turned out to be just the right time. I had to get back to Robert with my ideas on estimating our search inventory and didn't have time to listen to a sales pitch. I apologized to Chris and we scheduled a call for the next day. But before he hung up, Chris asked about some of the challenges I was facing at LookSmart. Out of my desperation for answers I decided to let my new acquaintance in on the problem, but only by giving him a small piece of the puzzle. "I'm trying to estimate the amount of interest in the wedding sector," I said. (If I was going to have to listen to a sales pitch, maybe I could get something out of it!)

The next day I received a forty-five-page PowerPoint deck long in advance of our call (not a good sign), and I purposefully blocked off only thirty minutes for the call in case I needed to escape.

They had me by the second slide. At that time, my external sources of data were based on what I thought were large samples, between forty and a hundred thousand Internet users in the United States. On the second slide of the deck, the data I was looking at was based on a sample of more than 10 million Internet users. This was unbelievable. So much so that I wasn't sure if I did believe it.

A thought immediately struck me; I was trying to solve my problem of estimating inventory by looking at one week's worth of data, when the chart in front of me clearly demonstrated just how seasonal Internet activity can be. Wedding site traffic spikes predictably during the summer months (wedding season), but looking closer at patterns in wedding-related search terms, like "wedding dresses," shows a much different pattern, with spikes in searches occurring during the first weeks of the year. Aggregate search behavior could show us when consumer demand peaks for various goods and services. I was hooked.

Convincing my boss was a different matter. When I explained how this new vendor's data could provide us with granular detail on all of the top U.S. sites, he didn't think it was real. In order to convince him of the data's accuracy, I devised a plan. Since we controlled the main results for MSN's search engine, we had a fair amount of control over the amount of search traffic that a site might receive from MSN. During two days in November, as a test, we placed a site called TripAdvisor (which was small at the time) into the top position in MSN results. Without revealing the dates, I asked Chris to chart out the traffic for TripAdvisor over several months. We agreed that if he could pinpoint the two days using Hitwise's data, he had a new customer. Five minutes later, the chart came through, and sure enough he had nailed the exact days.

After signing the contract, as I began using Hitwise data at LookSmart, I found myself facing a dilemma. I now had at my fingertips a looking glass into the aggregate Internet behavior of one of the largest samples in existence. At the time I could study more than 500,000 websites in 164 different industry categories, and have the data updated every day with new information. I suffered briefly from analysis paralysis as I tried to figure out where to start in this database. But that didn't last very long. I quickly realized that insight lay everywhere, from mea-

asuring the effectiveness of a competitor's marketing plan, to assessing interest in a yet-to-be-released movie by determining the amount of traffic visiting an official movie site.

I found myself losing track of time, spending hours finding interesting facets in the data, serendipitously following paths to counterintuitive conclusions. I didn't have enough time for my actual job. It was at this point that I approached Hitwise about the idea of working for them full-time as the head of research, an evangelist to show other users how to use this amazing trove of data, and to educate the market in general about the value of online competitive intelligence.

ABOUT THE DATA

The stories and examples in this book are based on data gathered from the Hitwise Competitive Intelligence Service. Unless otherwise specified, I relied on the U.S. sample. (Hitwise captures data in the United States, the UK, Hong Kong, Singapore, Australia, and New Zealand.) The U.S. sample contains the usage behavior of more than 10 million Internet users. The sample is gathered in two ways. The primary methodology, based on agreements with multiple ISPs throughout the country, anonymizes and aggregates usage data on more than 7.5 million users. The data is gathered from a number of ISPs to get a cross-section of Internet users regionally. The ISP sample is supplemented with multiple opt-in panels, or groups of Internet users who have agreed to be monitored and have supplied demographic information. That information allows us to report on the demographic and psychographic profiles of visitors to various sites and categories of sites. Both the ISP and opt-in data update every day, providing usage information for the previous day; search-term data is updated on a weekly basis; and demographic and psychographic data are based on a four-week rolling average of usage.

ABOUT PRIVACY

When I first show a friend, journalist, or business prospect the data that we have access to, one of the questions that commonly arises is: What about user privacy? Isn't the data that we're gathering a violation of end-user privacy? As a business, we know it's in our best interests to protect the privacy of end-users. We accomplish this through several mechanisms. First, our ISP and opt-in data partners anonymize and aggregate the data that is sent over to our system to analyze. In other words, we can't track any of our data back to individual Internet users (nor would we want to). We're more interested in tracking overall trends and gathering insights into how Internet users in aggregate are using the Web. We also scrub our search-term data to remove any personally identifiable information (phone numbers, credit card numbers, Social Security numbers, and so on).

ABOUT THE BOOK

This book is presented in two key sections. In Part I, you'll get to ride along as we explore fascinating points of data that have come up over the years. From visiting the dark side of the Internet and our habits revolving around porn, pills, and casinos, to understanding our obsession with celebrity, you'll see how our Internet behavior sheds light on why we do the things we do. We'll talk about the questions that we pose to search engines and what they reveal about us, such as queries about our fears, things that we would like to learn to do, or just general questions like "Why is the sky blue?" We will also address how some of the latest innovations in Internet communications have changed the way we interact with one another.

In the second half of the book we'll explore both actual and theoretical uses of this data, from the tactical applications that define the industry of online competitive intelligence, to the exciting potential to use the data to identify hot Internet trends, even to predict the next big rock star.

The most suitable place to start our journey through Internet behavior has to be the underbelly of the Net. In the world of Internet marketing, the acronym PPC stands for “pay per click,” the business model that addresses how much a business would be willing to pay for the users who clicked on their search advertisement. In my world, however, and in Chapter 1, PPC also has a different definition: It refers to porn, pills, and casinos. Internet data that surrounds our vices demonstrates just how compelling competitive intelligence data can be, from understanding the ebb and flow of traffic to websites, to discovering who visits those sites and what exactly is behind their decision to visit.

1

UNDERSTANDING OURSELVES

- [La Grande Illusion \(BFI Film Classics\) for free](#)
- [On the Spectrum of Possible Deaths.pdf](#)
- **click The Man-Kzin Wars**
- [download Botany : An Introduction to Plant Biology.pdf](#)

- <http://wind-in-herleshausen.de/?freebooks/Cartomancy-with-the-Lenormand-and-the-Tarot--Create-Meaning---Gain-Insight-from-the-Cards.pdf>
- <http://pittiger.com/lib/On-the-Spectrum-of-Possible-Deaths.pdf>
- <http://pittiger.com/lib/The-13th-Apostle.pdf>
- <http://anvilpr.com/library/Botany---An-Introduction-to-Plant-Biology.pdf>