

# Complete Web Monitoring

*Alistair Croll and Sean Power*

**O'REILLY®**

Beijing • Cambridge • Farnham • Köln • Sebastopol • Taipei • Tokyo

# Table of Contents

Preface .....	xix
---------------	-----

---

## Part I. The Business Of Web Monitoring

<b>1. Why Watch Websites? .....</b>	<b>3</b>
A Fragmented View	4
Out with the Old, in with the New	4
A Note on Privacy: Tracking People	5
<b>2. What Business Are You In? .....</b>	<b>7</b>
Media Sites	8
Business Model	8
Transactional Sites	9
Business Model	10
Collaboration Sites	11
Business Model	11
Software-as-a-Service Applications	13
Business Model	13
<b>3. What Could We Watch? .....</b>	<b>15</b>
How Much Did Visitors Benefit My Business?	15
Conversion and Abandonment	16
Click-Throughs	16
Offline Activity	17
User-Generated Content	19
Subscriptions	20
Billing and Account Use	21
Where Is My Traffic Coming From?	21
Referring Websites	22
Inbound Links from Social Networks	22
Visitor Motivation	24

---

What's Working Best (and Worst)?	24
Site Effectiveness	24
Ad and Campaign Effectiveness	25
Findability and Search Effectiveness	25
Trouble Ticketing and Escalation	26
Content Popularity	26
Usability	27
User Productivity	27
Community Rankings and Rewards	28
How Good Is My Relationship with My Visitors?	28
Loyalty	28
Enrollment	29
Reach	29
How Healthy Is My Infrastructure?	29
Availability and Performance	30
Service Level Agreement Compliance	31
Content Delivery	32
Capacity and Flash Traffic: When Digg and Twitter Come to Visit	33
Impact of Performance on Outcomes	35
Traffic Spikes from Marketing Efforts	36
Seasonal Usage Patterns	37
How Am I Doing Against the Competition?	38
Site Popularity and Ranking	38
How People Are Finding My Competitors	40
Relative Site Performance	41
Competitor Activity	42
Where Are My Risks?	42
Trolling and Spamming	42
Copyright and Legal Liability	45
Fraud, Privacy, and Account Sharing	45
What Are People Saying About Me?	46
Site Reputation	47
Trends	47
Social Network Activity	48
How Are My Site and Content Being Used Elsewhere?	48
API Access and Usage	49
Mashups, Stolen Content, and Illegal Syndication	49
Integration with Legacy Systems	51
The Tools at Our Disposal	51
Collection Tools	51
Search Systems	52
Testing Services	52

<b>4. The Four Big Questions .....</b>	<b>53</b>
What Did They Do?	53
How Did They Do It?	54
Why Did They Do It?	55
Could They Do It?	55
Putting It All Together	56
Analyzing Data Properly	57
Always Compare	57
Segment Everything	57
Don't Settle for Averages	58
A Complete Web Monitoring Maturity Model	59
Level 1: Technical Details	60
Level 2: Minding Your Own House	60
Level 3: Engaging the Internet	60
Level 4: Building Relationships	61
Level 5: Web Business Strategy	62
The Watching Websites Maturity Model	62

## **Part II. Web Analytics, Usability, and the Voice of the Customer**

<b>5. What Did They Do?: Web Analytics .....</b>	<b>67</b>
Dealing with Popularity and Distance	68
The Core of Web Visibility	69
A Quick History of Analytics	69
From IT to Marketing	71
From Hits to Pages: Tracking Reach	73
From Pages to Visits: The Rise of the Cookie	74
From Visits to Outcomes: Tracking Goals	78
From Technology to Meaning: Tagging Content	80
An Integrated View	84
Places and Tasks	86
The Three Stages of Analytics	89
Finding the Site: The Long Funnel	90
Using the Site: Tracking Your Visitors	100
Leaving the Site: Parting Is Such Sweet Sorrow	106
Desirable Outcomes	108
Implementing Web Analytics	111
Define Your Site's Goals	112
Set Up Data Capture	114
Set Up Filters	129
Identify Segments You Want to Analyze	130
Tag Your Pages to Give Them Meaning	131

Campaign Integration	134
Go Live and Verify Everything	136
Sharing Analytics Data	139
Repeat Consistently	140
Start Experimenting	141
Choosing an Analytics Platform	144
Free Versus Paid	145
Real-Time Versus Trending	145
Hosted Versus In-House	146
Data Portability	146
The Up-Front Work	146
What You Get for Free	147
What You Get with a Bit of Work	147
What You Get with a Bit More Work	147
What You Get with a Lot of Work	147
Web Analytics Maturity Model	147
<b>6. How Did They Do It?: Monitoring Web Usability .....</b>	<b>149</b>
Web Design Is a Hypothesis	149
Four Kinds of Interaction	150
Seeing the Content: Scrolling Behavior	151
Scrolling As a Metric of Visibility	152
Proper Interactions: Click Heatmaps	154
Usability and Affordance	156
Analyzing Mouse Interactions	158
Data Input and Abandonment: Form Analysis	160
Individual Visits: Replay	163
Stalking Efficiently: What You Replay Depends on the Problem You're	
Solving	164
Retroactive Segmentation: Answering "What If?"	172
Implementing WIA	173
Knowing Your Design Assumptions	174
Deciding What to Capture	175
Instrumenting and Collecting Interactions	175
Issues and Concerns	181
What if the Page Changes?	181
Visitor Actions WIA Can't See	182
Dynamic Naming and Page Context	182
Browser Rendering Issues	183
Different Clicks Have Different Meanings	183
The Impact of WIA Capture on Performance	184
Playback, Page Neutering, and Plug-in Components	184
Privacy	185

Web Interaction Analytics Maturity Model	188
<b>7. Why Did They Do It?: Voice of the Customer</b> .....	<b>189</b>
The Travel Industry's Dilemma	190
They Aren't Doing What You Think They Are	191
What VOC Is	192
Insight and Clues	194
Subjective Scoring	195
Demographics	195
Surfographics	196
Collection of Visit Mechanics Unavailable Elsewhere	198
What VOC Isn't	199
It's Not a Substitute for Other Forms of Collection	199
It's Not Representative of Your User Base	199
It's Not an Alternative to a Community	200
It's Not a Substitute for Enrollment	200
Four Ways to Understand Users	200
Kicking Off a VOC Program	202
Planning the Study	202
The Kinds of Questions to Ask	204
Designing the Study's Navigation	207
Why Surveys Fail	208
Integrating VOC into Your Website	214
Trying the Study	215
Choosing Respondents	216
Deciding Who to Ask	222
Private Panels	223
Disqualifying Certain Visitor Types	224
Encouraging Participation	224
Getting Great Response Rates	224
Setting Expectations	229
Permission to Follow Up	229
Improving Your Results	230
Analyzing the Data	232
Integrating VOC Data with Other Analytics	239
Advantages, Concerns, and Caveats	241
Learning What to Try Next	241
Becoming Less About Understanding, More About Evaluating Effectiveness	241
You May Have to Ask Redundant Questions	242
Voice of the Customer Maturity Model	242

---

## Part III. Web Performance and End User Experience

<b>8. Could They Do It?: End User Experience Management .....</b>	<b>245</b>
What's User Experience? What's Not?	246
ITIL and Apdex: IT Best Practices	246
Why Care About Performance and Availability?	248
Things That Affect End User Experience	252
The Anatomy of a Web Session	254
Finding the Destination	256
Establishing a Connection	257
Securing the Connection	262
Retrieving an Object	262
Getting a Page	267
Getting a Series of Pages	272
Wrinkles: Why It's Not Always That Easy	272
DNS Latency	272
Multiple Possible Sources	274
Slow Networks	276
Fiddling with Things: The Load Balancer	284
Server Issues	286
Client Issues	292
Other Factors	297
Browser Add-ons Are the New Clients	297
Timing User Experience with Browser Events	299
Nonstandard Web Traffic	300
A Table of EUEM Problems	303
Measuring by Hand: Developer Tools	304
Network Problems: Sniffing the Wire	304
Application Problems: Looking at the Desktop	305
Internet Problems: Testing from Elsewhere	306
Places and Tasks in User Experience	311
Place Performance: Updating the Container	311
Task Performance: Moving to the Next Step	312
Conclusions	312
<b>9. Could They Do It?: Synthetic Monitoring .....</b>	<b>315</b>
Monitoring Inside the Network	315
Using the Load Balancer to Test	316
Monitoring from Outside the Network	317
A Cautionary Tale	317
What Can Go Wrong?	318
Why Use a Service?	319

Different Tests for Different Tiers	321
Testing DNS	321
Getting There from Here: Traceroute	321
Testing Network Connectivity: Ping	322
Asking for a Single Object: HTTP GETs	324
Beyond Status Codes: Examining the Response	324
Parsing Dynamic Content	325
Beyond a Simple GET: Compound Testing	328
Getting the Whole Picture: Page Testing	328
Monitoring a Process: Transaction Testing	330
Data Collection: Pages or Objects?	333
Is That a Real Browser or Just a Script?	334
Configuring Synthetic Tests	336
Test Count: How Much Is Too Much?	337
Test Interval: How Frequently Should You Test?	337
Client Variety: How Should I Mimic My Users?	338
Geographic Distribution: From Where Should You Test?	340
Putting It All Together	342
Setting Up the Tests	343
Setting Up Alerts	343
Aggregation and Visualization	346
Advantages, Concerns, and Caveats	348
No Concept of Load	348
Muddying the Analytics	349
Checking Up on Your Content Delivery Networks	349
Rich Internet Applications	349
Site Updates Kill Your Tests	350
Generating Excessive Traffic	350
Data Exportability	350
Competitive Benchmarking	351
Tests Don't Reflect Actual User Experience	351
Synthetic Monitoring Maturity Model	351
<b>10. Could They Do It?: Real User Monitoring .....</b>	<b>353</b>
RUM and Synthetic Testing Side by Side	355
How We Use RUM	356
Proving That You Met SLA Targets	356
Supporting Users and Resolving Disputes	357
"First-Cause" Analysis	357
Helping to Configure Synthetic Tests	358
As Content for QA in New Tests	358
Capturing End User Experience	358
How RUM Works	359

Server-Side Capture: Putting the Pieces Together	360
Client-Side Capture: Recording Milestones	360
What We Record About a Page	361
Deciding How to Collect RUM Data	367
Server Logging	368
Reverse Proxies	369
Inline (Sniffers and Passive Analysis)	370
Agent-Based Capture	373
JavaScript	374
JavaScript and Episodes	377
RUM Reporting: Individual and Aggregate Views	380
RUM Concerns and Trends	381
Cookie Encryption and Session Reassembly	381
Privacy	381
RIA Integration	382
Storage Issues	382
Exportability and Portability	383
Data Warehousing	385
Network Topologies and the Opacity of the Load Balancer	385
Real User Monitoring Maturity Model	385

---

## Part IV. Online Communities, Internal Communities, and Competitors

<b>11. What Did They Say?: Online Communities</b> .....	<b>389</b>
New Ways to Interact	389
Consumer Technology	389
Vocal Markets	390
Where Communities Come from	391
Digital Interactions	391
Making It Easy for Everyone	393
Online Communities on the Web	394
Deciding What Mattered	394
Email for Everyone, Everywhere	396
Instant Gratification	396
Everyone's a Publisher	398
Microblogging Tells the World What We're Thinking	399
<b>12. Why Care About Communities?</b> .....	<b>405</b>
The Mouth of the Long Funnel	406
A New Kind of PR	407
Broadcast Marketing Communications	407
Online Marketing Communications	410

Viral Marketing: Pump Up the Volume	411
Community Marketing: Improving the Signal	417
Support Communities: Help Those Who Help Themselves	420
What Makes a Good Support Community?	421
Risk Avoidance: Watching What the Internet Thinks	423
Business Agility: Iterative Improvements	424
A Climate of Faster Change	424
Getting Leads: Referral Communities	426
<b>13. The Anatomy of a Conversation .....</b>	<b>429</b>
The Participants: Who's Talking?	429
Internal Community Advocates	429
External Community Members	432
The Topics: What Are They Talking About?	438
The Places: Where Are They Talking?	448
Different Community Models	449
User Groups, Newsgroups, and Mailing Lists	451
Forums	455
Real-Time Communication Tools	456
Social Networks	459
Blogs	464
Wikis	465
Micromessaging	468
Social News Aggregators	474
Combined Platforms	476
Why Be Everywhere?	476
Monitoring Communities	477
<b>14. Tracking and Responding .....</b>	<b>479</b>
Searching a Community	481
Searching Groups and Mailing Lists	482
Searching Forums	485
Searching Real-Time Chat Systems	487
Searching Social Networks	487
Searching Blogs	490
Searching Wikis	490
Searching Micromessaging Tools	491
Searching Social News Aggregators	493
Cross-Platform Searching	495
Joining a Community	495
Joining Groups and Mailing Lists	497
Joining Forums	497
Joining Real-Time Chat Systems	498

Joining Social Networks	498
Joining Blogs	499
Joining Wikis	500
Joining Micromessaging Tools	501
Joining Social News Aggregators	501
Moderating a Community	501
Moderating Groups and Mailing Lists	502
Moderating Forums	502
Moderating Real-Time Chat Systems	503
Moderating Social Networks	504
Moderating Blogs	504
Moderating Wikis	505
Moderating Micromessaging Tools	505
Moderating Social News Aggregators	505
Running a Community	506
Running Groups and Mailing Lists	507
Running Forums	507
Running Real-Time Chat Systems	507
Running Social Networks	509
Running Blogs	510
Running Wikis	510
Running Micromessaging Tools	513
Running Social News Aggregators	513
Putting It All Together	514
Measuring Communities and Outcomes	515
Single Impression	515
Read Content	516
Used the Site	516
Returning	516
Enrolled	517
Engaged	518
Spreading	518
Converted	518
Reporting the Data	518
What's in a Community Report?	519
The Mechanics of Tracking the Long Funnel	520
Responding to the Community	527
Join the Conversation	528
Amplify the Conversation	530
Make the Conversation Personal	530
Community Listening Platforms	530
How Listening Tools Find the Conversations	532
How Tools Aggregate the Content	532

How Tools Manage the Response	537
Community Monitoring Maturity Model	538
<b>15. Internally Focused Communities</b> .....	<b>539</b>
Knowledge Management Strategies	540
Internal Community Platform Examples	541
Chat	541
Social Networks	542
Wikis	544
Micromessaging Tools	544
Social News Aggregators	545
The Internal Community Monitoring Maturity Model	546
<b>16. What Are They Plotting?: Watching Your Competitors</b> .....	<b>547</b>
Watching Competitors' Sites	549
Do I Have Competitors I Don't Know About?	549
Are They Getting More Traffic?	550
Do They Have a Better Reputation?	552
PageRank	553
SEO Ranking	553
Technorati	555
Are Their Sites Healthier Than Mine?	555
Is Their Marketing and Branding Working Better?	556
Are Their Sites Easier to Use or Better Designed?	558
Have They Made Changes I Can Use?	558
Preparing a Competitive Report	559
What's in a Weekly Competitive Report	559
Communicating Competitive Information	559
Competitive Monitoring Maturity Model	560

---

## Part V. Putting It All Together

<b>17. Putting It All Together</b> .....	<b>563</b>
Simplify, Simplify, Simplify	563
Drill Down and Drill Up	564
Visualization	564
Segmentation	565
Efficient Alerting	565
Getting It All in the Same Place	565
Unified Provider	566
Data Warehouse	566
The Single Pane of Glass: Mashups	570

Alerting Systems	577
Tying Together Offsite and Onsite Data	578
Visitor Self-Identification	578
Using Shared Keys	579
<b>18. What's Next?: The Future of Web Monitoring .....</b>	<b>581</b>
Accounting and Optimization	582
From Visits to Visitors	582
Personal Identity Is Credibility	585
From Pages to Places and Tasks	585
Mobility	586
Blurring Offline and Online Analytics	586
Standardization	586
Agencies Versus Individuals	587
Monetizing Analytics	588
Carriers	588
Search Engines	589
URL Shorteners	589
Social Networks	589
SaaS Providers	591
A Holistic View	591
The Move to a Long Funnel	592
A Complete Maturity Model	593
A Complete Perspective	596
The Unfinished Ending	597
<b>Appendix: KPIs for the Four Types of Site .....</b>	<b>599</b>
<b>Index .....</b>	<b>617</b>