# Developing Web Analytics Best Practices for NIH Report

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15, 2011

*The Needs Assessment for Developing Best Practices for Trans-NIH Web Analytics was funded by the NIH Evaluation Set-Aside Program, (Reference number: 10-4031 OD-OCPL), administered by the Office of Program Evaluation and Performance, DPCPSI, Office of the Director*

# Contents

# Acknowledgments

Semphonic was honored to be selected to partner with the On-Line Information Branch of the Office of Communications and Public Liaison, Office of the Director, on the Developing Web Analytics Best Practices for NIH Project. We understand that this is the first time a project of this breadth and scope has been undertaken within a Federal Government agency.

I’d like to acknowledge the following people who have made this project possible and created a body of work that we hope will serve the NIH community for years to come:

- Ann Poritzky, Web Analytics Lead, On-Line Information Branch, whose vision, tenacity, and dedication to using Web analytics throughout NIH led to the creation and funding of this project. In addition, Ann’s leadership during the project, review of the project deliverables, and feedback about the new approaches we’re taking on digital analytics for a Federal Government agency have been invaluable and will no doubt enable the recommendations presented here to also be used by Federal agencies outside NIH.
- Dennis Rodrigues, Chief, On-Line information Branch, who supported Ann’s vision of this project and whose commitment to the use of Web analytics as a means to guide NIH Web strategy is an example of forward-looking leadership that will enable Federal agencies to thrive in the digital age.
- The Developing Web Analytics Best Practices Project Team members, who provided excellent feedback on the project’s presentations and research instruments: Cheryl Thompson, Web Manager, National Institute of Environmental Health Sciences; Craig Hicks, User Experience Designer, National Heart, Lung, and Blood Institute; Dennis Rodrigues, Chief, On-Line Information Branch, Office of the Director; Eric Djiva Kamal, Web Analytics System Administrator, National Library of Medicine; Fred Wood, Science Program Leader/Computer Scientist, National Library of Medicine; Ilene France, Digital Information Specialist, National Institute of Allergy and Infectious Diseases; James Petersen, Technical Writer/Editor, National Institute of Mental Health; Karina Boehm, Chief, Health Education Branch and Web Analytics Lead, National Institute of Dental and Craniofacial Research; Louise Marshall, Technical Information Specialist, National Institute of Mental Health; Mark Malamud, Web Analytics Program Manager and 508 Coordinator, National Heart, Lung, and Blood Institute; Sue Feldman, Web Analytics Program Manager, National Cancer Institute.
- My Semphonic colleagues who contributed their expertise and review of the Best Practices, as well as helping me develop new metrics and methods for conducting digital analytics for a Federal agency. I’d like to specifically recognize Gary Angel, Greg Dowling, Kevin Dupzyk, Jesse Gross, Chelsea Leonard, and Chris Meares.

Phil Kemelor, Semphonic

# Executive Summary

The Developing Web Analytics Best Practices for NIH Project is unique among U.S. Federal Government agencies in terms of its breadth and scale. No other Federal agency has conducted research to identify the Web and digital analytics information needs of the whole agency and, then, how to make Web analytics tools — often relegated to Information Technology (IT) departments — relevant and useful to the agency's communications and Web site development staff.

The NIH On-Line Communications Branch obtained evaluation set-aside funds for this initiative and selected Semphonic, a vendor-neutral Web analytics consulting firm, to conduct the research and develop recommendations and the reference materials that are the deliverables of this project. While the project recommendations require NIH Institutes and Centers (IC) to have Web analytics solutions, they do not require any particular vendor's product.

The main goal of the project was to develop specific recommendations to improve the quality, consistency, and comparability of Web site measurement among public-facing sites throughout NIH. Another important goal was to help NIH and the ICs use Web-measurement data for strategic guidance and as support for decisions about and investments in NIH's online initiatives.

To understand how Web site measurement at NIH works now and to improve it for the future, the project focused on answering these questions:

- How do NIH ICs measure Web sites overall now (for example, site traffic, usability, customer satisfaction)?
- What specific Web-measurement tools do ICs use (for example, Webtrends, Omniture, Google Analytics, American Customer Satisfaction Index (ACSI) surveys)?
- How can ICs do a better job of using Web and digital analytics?

Semphonic consultants worked closely with NIH staff. The project began by inviting the whole NIH Web and communications community to presentations about Web site and digital measurement at NIH.

The introductory presentations had about 300 attendees (in person and online). Between January and May 2011, Semphonic interviewed 76 NIH staff from the Office of the Director and 16 ICs, surveyed 131 NIH staff (online), and reviewed several ICs' Web-measurement-related reports. One hundred and twenty NIH employees attended presentations about the project's results.

The Semphonic Web Analytics Return on Investment (WAROI) Framework served as the basis for analyzing and assessing the NIH Web analytics efforts and for making recommendations. The framework's components are

- commitment of senior management,
- organizational resources and responsibilities,
- process for organization-wide and department-specific Web analytics governance,

- Web analytics metrics and deep-dive analysis,
- Web analytics solutions, methods, and opportunities, and
- Web analytics training, data analysis, and interpretation of data.

Definitions of and recommendations associated with each of the framework's components are listed below.

### Commitment of Senior Management

The use of Web analytics throughout organizations such as NIH succeeds when senior executives are committed to using analytics data to drive online strategies. At NIH, we found that many who work on IC Web sites would welcome a stronger commitment to Web analytics from Dr. Collins and from their own IC directors. NIH employees highlighted several challenges that prevent NIH and ICs from fully embracing Web and digital analytics as powerful resources for strategic planning or decision guidance. Many reported that a commitment to Web analytics from NIH senior managers was significantly absent.

#### *Recommendations*

- Senior management must develop a coordinated strategy for the Web and digital channels.
- Senior management must establish consistent and transparent funding for Web analytics and other digital analytics software solutions and programs.

### Organizational Resources and Responsibilities

Having people trained and available to create and interpret Web analytics data is generally an understood requirement. However, what is not often understood are the specific software, hardware, and human resources required to support Web analytics reporting and analysis. Analytics works well when enough of the right resources are dedicated to administration, project management, analysis, and technical support.

There are different levels of analytics services at NIH, but no central group coordinates Web or digital analytics strategies. Most ICs do not have dedicated staff to work with digital analytics data. Although analytics initiatives are much more powerful when they are standardized and coordinated, most ICs handle online measurement in an ad hoc manner.

We also found an overall sense of frustration among ICs that rely on Web analytics from central NIH IT organizations, such as OIT and especially CIT, and higher satisfaction with Web analytics hosting and services from other sources, such as contractors.

#### *Recommendations*

- Create a Digital Channel Program Office within the Office of the Director and an NIH Digital Channel Management Council.
- Expand the scope of Office of the Director (OD) On-Line Information Branch so that it guides NIH Web initiatives.

- Use a team approach for planning, implementing, and using Web analytics so that the tools and data are managed, maintained, and used correctly.
- Define user classes to clarify who has access to which data and in which formats and interfaces and to determine what training is necessary.
- Create Web analytics user groups within ICs.
- Link performance to Web metrics to increase accountability.

### Process for NIH-Wide and IC-Specific Web Analytics Governance

Successful Web analytics initiatives depend on a governance framework that includes systems for data collection, processing, and integration; metrics development and presentation; data analysis; and solution maintenance. Such initiatives are based on a collection of processes that must be created and followed by stakeholders, managers, and support teams.

At NIH, ICs face significant challenges in trying to make full use of analytics and to ensure they have enough employees to fulfill the required staff positions. Only a few ICs have processes for translating site goals into metrics that measure progress toward achieving those goals. ICs that do receive Web analytics reports tend not to share them widely. Although most do not have standards for collecting or preparing data to ensure consistency and accuracy, a couple of ICs do include measurement as an essential part of their Web site or content-development cycles.

#### *Recommendations*

- Formalize a project-planning process in which Web analytics metrics and reports are developed, produced, and tested in tandem with new Web applications, content, and outreach campaigns.
- Establish a standard approach for developing Web metrics.
- Develop data-collection frameworks and standards to ensure data integrity.
- Formalize protocols for basic Web analytics tasks such as presenting dashboards and data, responding to ad hoc data requests, and accessing the Web analytics solution.
- Support stakeholders' planning processes by developing strategies for distributing reports, interpreting them, and answering questions.
- Simplify and clarify the procurement process for obtaining analytics support.
- Develop a functional-specification template to map Web analytics requirements to Web analytics report building and data collection.
- Develop and maintain implementation standards for Web analytics data collection.
- Develop and maintain standard operating procedures for all who support the Web analytics initiative.

## Web Analytics Metrics and Deep-Dive Analysis

Effective measurement relies on providing the data and analysis required for driving strategic and tactical decision-making and presenting the data in ways that can be used by specific users, whether they are senior managers, marketers, designers, or content editors. At NIH, there are many different views about which measurements ICs should be using to assess Web content and communications and how information should be compared and shared.

Most ICs lack instructions about how to implement Web analytics, collect data in standard ways, interpret results, or prepare reports and dashboards. Nearly everyone we interviewed would like to get training in using Web analytics tools and analysis-related skills. They specifically asked for training in Web analytics, ranging from how to use the tools to analyzing and using data.

### *Recommendations*

- Use triangulation to manage multiple data sources.
- Implement the NIH Web Analytics Method. (See page 36 for description of NIH Web Analytics Method.)
- Increase the value of analysis by using segmentation.
- Institute success metrics, a goal-based Web analytics approach to evaluating NIH Web sites.
- Use the success metrics to better inform NIH and IC senior leadership and Web managers about the performance of Web sites.
- Define metrics for benchmarking within and across ICs' public-facing sites.
- Encourage a standardized approach for presenting success metrics through the use of dashboards.
- Use deep-dive analysis to complement dashboards and reports.
- Increase the use of A/B testing for content and Web site optimization.
- Begin to measure the investment in social media strategies.

## Web Analytics Solutions, Methods, and Opportunities

Web, social media, competitive intelligence, and online surveying are all driven by software, as are multivariate testing and search marketing (sometimes referred to as pay-per-click (PPC) search marketing). These solutions have tended to be "owned" by the stakeholder group that uses them most often. Because of this, NIH may be missing the opportunities for greater efficiencies of scale and overall return on investment (ROI) that the centralized purchase and distribution of software allows.

### *Recommendations*

- Implement a reliable Web analytics solution.
- Encourage the adoption of Google Analytics or Piwik (see **Appendix K**).

- Consider a two-tier Web analytics option. One tier—the basic tier—would be a free product such as Google Analytics. The second tier—the advanced tier—would be an enterprise-level tool, such as Adobe Omniture, Webtrends, IBM Coremetrics, or Comscore Digital Analytix.
- Use first-party cookies because, while not required for Web analytics, they add tremendous value.

## Web Analytics Training, Data Analysis, and Interpretation of Data

Web analytics and digital analytics data are challenging to use because the numbers are associated with activities that do not mean very much to most people. How do you interpret whether a certain number of page views is good or bad? How do you use the information to figure out what to do? This is why developing clear language around Web analytics is essential to their successful adoption and use throughout any organization. Digital analytics training and education should focus on how to interpret the data so that they are understandable and useful and how to produce analytics reports.

### *Recommendations*

- Initiate monthly meetings of Web analytics team members and the supporting Web content, design, and technical teams to discuss current metrics, reporting, and analysis and other issues regarding the analytics initiative.
- Develop a strategy to provide interpretive analysis.
- The IC Web analytics manager and the Web analytics lead in the On-Line Information Branch should make a quarterly “State of the Web” presentation to respective management teams.
- Enhance the online Web analytics knowledge center.
- Develop an analytics cooperative.
- Develop an analytics training and education curriculum.

## More Information

For more detailed information, please refer to the Findings and Recommendations section and the appendices in the Developing Web Analytics Best Practices for NIH Report:

Appendix A. Survey Analysis: Developing Web Analytics Best Practices for NIH
Appendix B. NIH Web Analytics Requirements Summary
Appendix C. Process for Collecting Accurate Analytics Data
Appendix D. Success Metrics Calculations
Appendix E. Dashboard Content and Sample
Appendix F. Functionalism: A New Approach to Web Analytics
Appendix G. How to Conduct A/B Testing
Appendix H. Strategies for Social Media Measurement
Appendix I. Evaluation of Web Analytics and Measurement Tools
Appendix J. Strategies for Google Analytics Implementation

Appendix K. Piwik Evaluation: A Free, Open-Source Web Analytics Tool
Appendix L. Glossary for NIH Web Analytics Best Practices

## Roadmap

To encourage NIH to adopt the recommendations in the Developing Web Analytics Best Practices for NIH Project Report, we advise implementing them in phases. We suggest a timeline for planning and implementing the various recommendations, most of which are separate initiatives. Staff will need to examine each one, break it into subtasks, and develop steps for putting them into practice.

The wording of the roadmap recommendations is slightly different from the wording in the report so that the recommendations are easily understood as activities and tasks to undertake.

The roadmap is broken down by the three primary groups for whom the best practices and recommendations were written:

- NIH ICs
- Office of the Director, Office of Communications and Public Liaison, On-Line Information Branch
- NIH senior leaders

### *Phase One – Planning: Zero to six months after recommendations are accepted*

#### NIH ICs

1. Determine who will serve in the following roles on the IC's Web analytics team: Web analytics manager, Web analyst, and developer–system administrator (to work on the technical aspects of the Web-measurement tools).

2. Set up regular monthly meetings for the Web analytics team members, employees supporting Web content, and design and technical teams to discuss current metrics, reporting, and analysis and other issues related to the analytics initiative.

3. Identify the employees who will be the "power users" of the analytics tools and the report viewers who will guide decisions about who has access to which data, in which formats or interfaces, and what their training requirements will be.

4. Create user groups for Web-measurement tools within ICs.

5. Establish a standard approach for developing business-oriented Web metrics derived from goals, objectives, and understanding Web site audiences.

6. Develop a process in which Web metrics dashboards are produced and tested on new Web applications, content, and outreach campaigns.

7. Develop data-collection frameworks and standards to ensure data integrity.

8. Formalize processes for completing basic Web analytics tasks, such as delivering reports, responding to ad hoc data requests, and granting access to the Web analytics tools.
9. Plan ways to manage multiple data sets from the various available tools and research methods through a disciplined “triangulation” approach.
10. Use the standard NIH Web Analytics Method for developing goal-oriented metrics and conducting analysis.
11. Develop a strategy to provide interpretive analyses of results from tool-based and other measurement approaches.
12. Any ICs considering a transition to a new Web analytics tool should determine whether to implement a free Tier 1 solution, such as Google Analytics or Piwik, or a fee-based Tier 2 solution.

**Office of the Director, Office of Communications and Public Liaison, On-Line Information Branch**

1. Develop a functional-specification template, based on the instructions in **Appendix F**, to map Web analytics requirements to Web analytics report development and data collection. The template can serve as a model that all ICs can use.
2. Create an online Web Analytics Knowledge Center for presentations, documents, and articles about Web analytics, as well as NIH Web analytics case histories. Consider establishing a wiki for sharing analytics issues and successes and/or an internal blog written by members of NIH Web Metrics Group.
3. Initiate an analytics cooperative to give interested NIH employees opportunities to learn how to use tools, create metrics, and interpret data.
4. Develop an analytics training and education curriculum.
5. Assist ICs in evaluating their current Web analytics solutions and in making decisions about continuing with current analytics solutions, moving to a Tier 1 solution (either Google Analytics or Piwik), or purchasing a Tier 2 solution.
6. Consider procuring an enterprise-level, fee-based Tier-2 Web analytics solution that enables NIH ICs to take advantage of economies of scale and lower costs for larger number of users.

**NIH Senior Leaders**

1. Support the development of a coordinated communications strategy for all public-facing NIH Web and online initiatives at NIH.
2. Work to establish consistent and transparent funding for Web and other digital analytics activities and support.

3. Create the Digital Channel Program Office in the Office of the Director, Office of Communications and Public Liaison, On-Line Services Branch.

4. Appoint a Trans-NIH Digital Channel Management Council to govern the Web communications channel throughout NIH.

5. Expand the scope of the Office of the Director, On-Line Information Branch to include guiding NIH digital channel operations.

### *Phase Two: Transition – 7 to 12 months after recommendations are accepted*

**NIH ICs**

1. Continue with Phase 1 tasks.

2. Start preparing, interpreting, and distributing dashboards featuring success metrics.

3. Increase the use of A/B testing for content and Web site optimization tactics.

4. Begin to measure the ICs' investment in social media strategies.

5. Initiate quarterly "State of the Web" presentations by Web analytics managers to their respective management teams.

**Office of the Director, Office of Communications and Public Liaison, On-Line Information Branch**

1. Continue with Phase 1 tasks.

2. Initiate quarterly "State of the Web" presentations by the Web Analytics lead in the On-Line Information Branch to the management or communications team.

**NIH Senior Leaders**

1. Continue with Phase 1 tasks.

2. Simplify and clarify the procurement process for analytics tools and support.

### *Phase 3: Adoption – 13 to 18 months after recommendations are accepted*

During Phase 3, NIH ICs; the Office of the Director, On-Line Information Branch; and NIH senior leaders should continue working the Phase 1 and 2 tasks listed above.

# Introduction

The Developing Web Analytics Best Practices for NIH Project is unique among U.S. Federal Government agencies in terms of its breadth and scale. No other Federal agency has conducted research to figure out the Web and digital analytics information needs of the whole agency and, then, how to make Web analytics tools — often relegated to IT departments — relevant and useful to the agency's communications and Web site development staff. We examined the analytics tools and measurement approaches in use at NIH, how the Institutes and Centers (IC) use the resulting data to make decisions, and how organizational processes and management decisions affect how widely and well NIH is using Web and digital analytics.

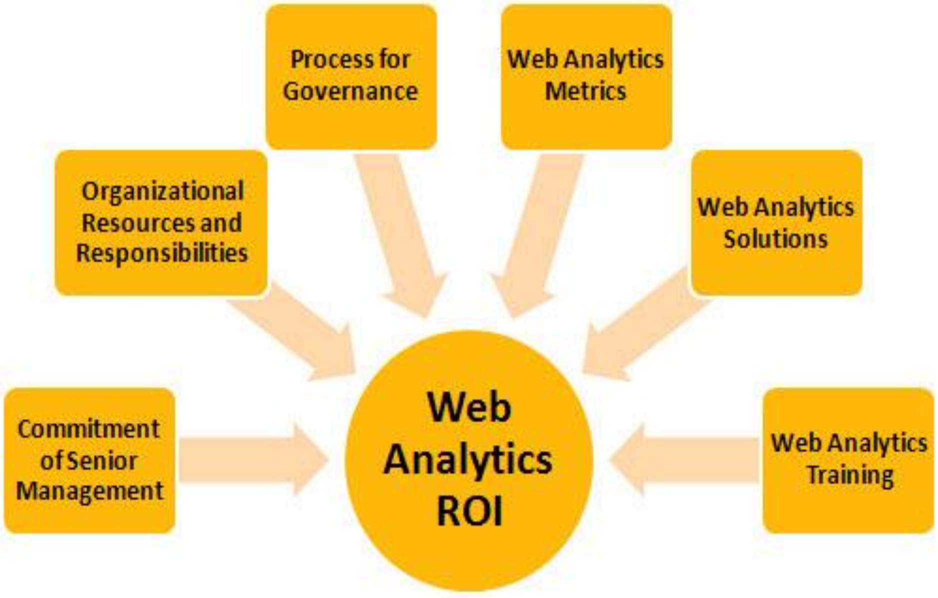
The main goal of this project was to develop specific recommendations to improve the quality, consistency, and comparability of Web site measurement among public-facing sites throughout NIH. Another important goal was to help NIH and the ICs use Web-measurement data for strategic guidance and support for decisions about and investments in NIH's online initiatives.

The NIH On-Line Communications Branch obtained set-aside funding for this initiative and selected Semphonic, a vendor-neutral Web analytics consulting firm to conduct the research and develop recommendations and the reference materials that are the deliverables of this project.

# Project Methodology

We based our approach to Developing Web Analytics Best Practices for NIH on an examination of six areas we identified as the foundation for a Web Analytics Return on Investment (WAROI) Framework (**Figure 1**) that can lead to a “best in class” Web analytics program.

***Figure 1.*** *Web Analytics Return on Investment Framework.*



- Commitment of Senior Management: The use of Web analytics throughout organizations such as NIH succeeds when senior executives are committed to using analytics data to drive online strategies.
- Organizational Resources and Responsibilities: Having people trained and available to create and interpret Web analytics data is generally an understood requirement. However, what is not often understood are the specific software, hardware, and human resources required to support Web analytics reporting and analysis. Analytics works well when enough of the right resources are dedicated to administration, project management, analysis, and technical support.
- Process for Governance: Successful Web analytics initiatives depend on a governance framework that includes systems for data collection, processing, and integration; metrics development and presentation; data analysis; and solution maintenance. Such initiatives are founded on a collection of processes that must be created and followed by stakeholders, managers, and support teams.
- Web Analytics Metrics and Deep-Dive Analysis: Effective measurement relies on providing the data and analysis required for driving strategic and tactical decision-making and presenting the data in ways that can be used by specific users, whether they are senior managers, marketers, designers, or content editors. We talk about Web analytics measurement in terms of key

performance indicators (KPI), digital analytics metrics, and deep-dive analysis, but these terms have multiple meanings. At NIH, there are many different views about which measurements we should be using to assess Web content and communications.

- Web Analytics Solutions, Methods, and Opportunities: Web, social media, competitive intelligence, and online surveying are all driven by software, as are multivariate testing, search marketing (sometimes referred to as pay-per-click (PPC) search marketing), and cross-selling platforms. These solutions have tended to be "owned" by the stakeholder group that is most often using them. Because of this, NIH may be missing the opportunities for greater efficiencies of scale and overall return on investment (ROI) that the centralized purchase and distribution of software allows.
- Web Analytics Training, Data Analysis, and Interpretation of Data: Web analytics and digital analytics data are challenging to use because the numbers are associated with activities that do not mean very much to most people. How do you interpret whether a certain number of page views is good or bad? How do you use the information to figure out what to do? This is why developing clear language around Web analytics is essential for their successful adoption and use throughout any organization. Digital analytics training and education should focus on how to interpret the data so that they are understandable and useful and how to produce analytics reports.

The Developing Web Analytics Best Practices for NIH Project used a comprehensive discovery process that included

1. **Stakeholder Interviews:** The Office of the Director's Office of Communications and Public Liaison invited NIH staff members to participate in the project, and Semphonic conducted 29 interviews (individual and group) with 76 staff members from the following Institutes, Centers, and Offices (IC):

   Center for Information Technology
   National Cancer Institute
   National Center for Complementary and Alternative Medicine
   National Center for Research Resources
   National Human Genome Research Institute
   National Institute on Drug Abuse
   National Institute of Environmental Health Sciences
   National Institute of Allergy and Infectious Diseases
   National Institute of Arthritis and Musculoskeletal and Skin Diseases
   National Institute on Deafness and Other Communication Disorders
   National Institute of Dental and Craniofacial Research
   National Institute of Mental Health
   National Institute on Minority Health and Health Disparities
   National Institute of Neurological Disorders and Stroke
   National Library of Medicine

NIH Clinical Center
Office of the Director

2. **Online Survey:** We created an online survey and made it available to staff involved in current Web initiatives throughout NIH and familiar with their IC's digital analytics activities. The survey received 131 responses (see **Appendix A**).

3. **Document Review:** We reviewed analytics reports from 12 ICs.

This report grew out of the discovery process and includes the following sections:

- Findings and Recommendations
- Appendix A. Survey Analysis: Developing Web Analytics Best Practices for NIH
- Appendix B. NIH Web Analytics Requirements Summary
- Appendix C. Process for Collecting Accurate Analytics Data
- Appendix D. Success Metrics Calculations
- Appendix E. Dashboard Content and Sample
- Appendix F. Functionalism: A New Approach to Web Analytics
- Appendix G. How to Conduct A/B Testing
- Appendix H. Strategies for Social Media Measurement
- Appendix I. Evaluation of Web Analytics and Measurement Tools
- Appendix J. Strategies for Google Analytics Implementation
- Appendix K. Piwik Evaluation: A Free, Open-Source Web Analytics Tool
- Appendix L. Glossary for NIH Web Analytics Best Practices

# Findings and Recommendations

This section is the central part of the strategic plan. It is organized around the six-point Web Analytics ROI Framework described in the Introduction.

Each finding is rated Good, Satisfactory, Evolving, or Needs Improvement, and recommendations follow the findings. Some of the findings here are not included in the Requirements Summary (**Appendix B**) because that document includes only findings from the stakeholder interviews and the NIH-wide survey. Here, we also include findings based on observations and references from Semphonic's work with other clients who face challenges similar to NIH's. **Appendices A** and **B** complement the recommendations.

Ratings:

- Good – This means that NIH or one or more ICs is exceeding goals and objectives in that framework area, and no significant improvement is required.
- Satisfactory – NIH or one or more ICs is meeting the goals and objectives related to that framework area. Some improvement is needed.
- Evolving – NIH or individual ICs are working toward meeting the goals and objectives related to the framework area and are improving.
- Needs Improvement – NIH or ICs are not doing enough work in this area, and the work that is being done doesn't appear to be leading to the needed improvements.

Please note that there is not a one-to-one mapping of findings to recommendations. One recommendation may address multiple findings, and multiple recommendations may address one finding.

## 1. Commitment of Senior Management

The use of Web analytics throughout organizations such as NIH succeeds when senior executives are committed to using analytics data to drive online strategies. We can estimate the level of commitment by looking at budgets, communication about analytics throughout the organization, and the direction that senior management gives staff about the use and value of analytics**.**

In general, we found that many who work on IC Web sites would welcome a stronger commitment to the Web from Dr. Collins as well as from their own IC directors. Addressing the following findings and requirements would help NIH use Web analytics to drive online strategy and tactics.

### *Findings*

- **Finding 1.1. The NIH mission focuses on research, but many ICs focus their Web sites on providing information to the public. – Needs Improvement**

  This dichotomy may affect the support for Web site initiatives and the analytics that could support those initiatives because some senior managers may not believe that the public information focus of the Web sites is critically linked to the NIH mission. Originally developed for researchers and grantees, most visible NIH Web sites are now more targeted to delivering health information to the public and health professionals and conducting outreach activities.

In some respects, this focus on public health is driven by the traffic-oriented approach to Web analytics. Constantly increasing page view and visit numbers reinforce the notion that the Web site is successful. However, on the basis of the conversations we had, it does not appear that the public health success of the sites is as meaningful to senior leadership as is the NIH mission focus on research. The disconnect between the NIH mission and Web site public health objectives has caused confusion about the goals and objectives of the sites within the ranks of those who manage them on a daily basis.

- **Finding 1.2. Many senior managers view the Web as a "feel good" effort that receives "love" but not enough financial support. – Needs Improvement**

  Stakeholders repeatedly indicated that their Web efforts were generally received well by senior management. However, they also indicated that budgets were generally flat year after year and that the Web initiatives were last in line for budget allocations.

- **Finding 1.3. When it comes to the relative effectiveness of Web sites, Web analytics and other online measurement practices do not often have as much influence on senior management as opinions held by friends, acquaintances, colleagues, or the executives in charge of the ICs. – Needs Improvement**

  Stakeholders cited multiple examples of senior management's preference to rely on opinions held by friends, acquaintances, or colleagues or articles they read rather than on Web analytics when assessing the effectiveness of a site or making changes to it.

  On the basis of our findings, we believe that in most cases, senior management is far more focused on promoting the NIH mission "offline," such as in newspapers and journals and at conferences, than online.

- **Finding 1.4. Some senior managers believe that the most important metric for Web sites is how much content can be published, rather than how visitors are using the content and accomplishing tasks that indicate successful use of the site. – Needs Improvement**

  The most prevalent metric being used for Web site development at NIH these days is how much content is published. A significant amount of time is spent on developing content or trying to get content published, but little is being done to determine, justify, or validate whether these efforts are necessary from a mission or return-on-investment (ROI) perspective.

- **Finding 1.5. Web analytics at NIH is based on a range of IC-specific efforts that may or may not have senior management support. – Needs Improvement**

  Web site managers have done what they can to gauge the success of their efforts. In many cases, analytics is a part-time endeavor that receives little consistent funding support. Funding models meant to provide centralized support from CIT and OIT haven't taken into account the need to train staff to use the Web analytics tools so their employees can provide valuable support. Other ICs depend on outside contractors for support, which seems to be a more successful model for obtaining value from Web analytics software. In addition, it's possible to apply for and secure set-aside funds for ad hoc analytics initiatives. In sum, the analytics efforts at NIH are a hodgepodge in scope and value. Although we could not quantify the cost of analytics initiatives in time and labor, our anecdotal findings suggest that efforts undertaken by individual ICs, also known as "siloed" efforts, result in many operational inefficiencies and lost opportunities for sharing resources and knowledge.

- **Finding 1.6. Budgets for the Web tend to be flat, and budget allocation is usually tied only to the previous year's allocations. – Needs Improvement**

  In most cases, the Web managers we spoke to indicated that budgets were flat or being cut, and it was not always clear how Web budgets were allocated. While we are currently in a tight budget environment, this situation has existed for years. It's apparently common practice to fund many other programs that relate to an IC's mission around research, and if there is money left over, Web programs receive the funding. However, there are examples of Web initiatives' receiving funding if IC directors understand the business imperative of the Web site or other online efforts. The funding of the Intramural Research Web site is such a case. This leads us to believe that NIH senior leadership is open to making more of a financial commitment to Web programs if a strong business case can be made for it. Improving the state of analytics could provide the "hard" data for that effort.

- **Finding 1.7. There needs to be a senior-level vision of how to use Web analytics data to create Web products that advance the NIH mission and serve visitors. – Needs Improvement**

  The potential for using Web data – analytics, surveys, market research – is vast and largely untapped. For the most part, NIH Web sites are like car companies, producing high quantities of content and creating inventory that is parked on large lots (or, in this case, Web sites). There is not enough being done to understand how the current content is used so that new information products of higher value to discrete audiences can be created. This is one of the reasons that nongovernment and commercial health sites garner more engagement, media recognition, and loyalty than NIH sites do.

  Currently, efforts to develop new Web sites or online products are undertaken by individual ICs, and there's no coordination among them. This is symptomatic of the lack of overall Web strategy, and it hurts NIH to use such a fragmented approach to the Web.

### *Recommendations*

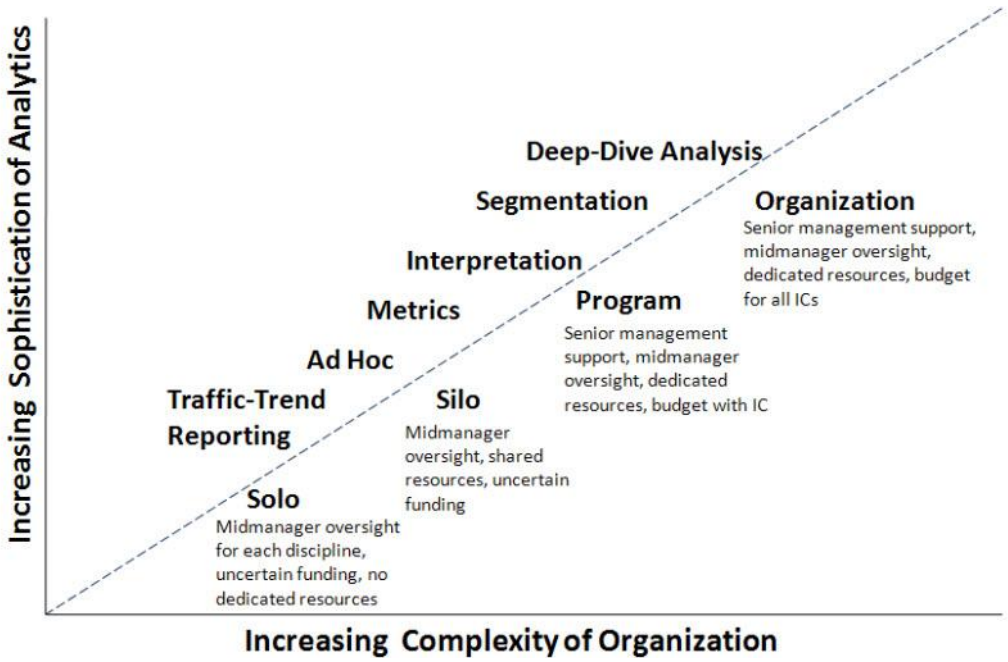
In the first annual Semphonic Profiles in Web Analytics Survey of the issues, challenges, strategies, and tactics Web analytics managers use, conducted in August 2010, we found that the single most important factor in successful Web analytics programs was *strong management commitment to analytics*. Our work with dozens of companies in the government, private, and nonprofit sectors has borne this out, as well as our work with organizations based in the United States and overseas. We see a direct correlation between management commitment and the ability of an organization to maximize the value of Web-visitor data.

In **Figure 2**, we show the relative value (y-axis) of how analytics data are used (see “Analytics” terms above the dashed line) when correlated with how analytics is conducted throughout the organization (see “Organization” terms, below that line). When there is less organization-wide structure to support the analytics, the relative value is less, such as when a single person is charged with all analytics functions (“Solo”) and can get only “Traffic-Trend Reporting.”

As you move up the relative-value scale, the value of analysis increases because there is more organization-wide recognition that it is a program worth supporting with sufficient resources. At NIH, most IC initiatives are within the Solo and Silo stages, with less than a handful approaching the Program level.

We found that NIH senior management is not strongly committed to Web analytics, which is consistent with the lack of commitment to consistent funding and support of the Web and digital channel initiatives we observed. Our recommendations address these observations.

***Figure 2.*** *Relative value of Web analytics data relies on management commitment to analytics.*


Deep-Dive Analysis
Segmentation
Interpretation
Metrics
Ad Hoc
Traffic-Trend Reporting
Organization
Senior management support, midmanager oversight, dedicated resources, budget for all ICs
Program
Senior management support, midmanager oversight, dedicated resources, budget with IC
Silo
Midmanager oversight, shared resources, uncertain funding
Solo
Midmanager oversight for each discipline, uncertain funding, no dedicated resources
Increasing Sophistication of Analytics
Increasing Complexity of Organization

- **Recommendation 1.a. Senior management must develop a coordinated strategy for the Web and digital channels.**

  One of the first steps NIH senior leadership should take toward having a successful Web channel is to provide strategic direction that defines a plan for accomplishing overall Web site goals. This may require a shift in perspective in how NIH positions itself not only in the offline world of conferences and research, but also in the online world of communication and applications. That, in turn, may require working with advisors and experts who understand the most effective ways to present health information online.

  This strategy should focus on

  - Providing leadership among government agencies and research organizations in how to use Web and other digital analytics to support a world-class Web channel.
  - Recognizing that the Web is a central channel through which NIH can accomplish its mission as a leading health and medical research institution.
  - Clarifying how the NIH mission can guide the Web site strategy.
  - Developing a strategy for an NIH brand that resonates throughout all IC digital initiatives.
  - Articulating clearly digital strategies for attracting, engaging, and keeping the channel relevant to NIH audiences, including researchers, health professionals, consumers, teachers, and students.
  - Developing a clear Web strategy that focuses on governance, branding, operations, and funding mechanisms.

  Not only will this Web analytics strategy help NIH ICs align in support of these goals, but it will ensure a consistent Web analytics approach across NIH.

- **Recommendation 1.b. Senior management must establish consistent and transparent funding for Web analytics and other digital analytics software solutions and programs.**

  It was difficult for us to find out the levels of budgeting and support for Web site and Web analytics initiatives. It did seem, though, that the funding is inadequate. We could not determine whether this was due to the actual funding or to how funds were spent. Because of today's tight budget environment, we recommend

  - Providing consistent funding for Web analytics initiatives, which could help senior-level and Web site management understand the overall return on investment (ROI).
  - Increasing transparency about Web budgeting and expenditures, which will help determine cost models that enable the development of ROI metrics (see the metrics section, beginning on page 20).

## 2. Organizational Resources and Responsibilities

Having people trained and available to create and interpret Web analytics data is a well-understood requirement at NIH. However, the need to have the software, hardware, and human resources to support Web analytics reporting and analysis is often not appreciated. This issue plays out in all the ICs, no matter what size they are.

Most responses to survey questions about the resources used for analytics at NIH indicated two things: that analytics isn't currently a full-time job at NIH and that this resource shortfall is felt throughout NIH and is certainly detrimental. In fact, when asked about the challenges of doing analytics, nearly 70% of respondents selected "Not enough staff to analyze data" or "Not enough time to analyze data."

This is not a problem purely in terms of resource hours to spend analyzing data; the lack of full-time analytics staff also prevents the development of a requisite level of expertise and familiarity with analytics data. Survey respondents repeatedly indicated that lack of expertise, training, and reliability of data are all major issues. If senior management could see analytics as a responsibility that requires a full-time staff member who can use tools and makes improving their analytic capabilities a primary component of professional development, this would go a long way toward addressing these concerns.

From our perspective, organizational impediments between or among Information Technology (IT) and Communication teams and gaps in staff knowledge are the biggest organizational challenges NIH needs to overcome to build a world-class analytics environment.

### *Findings*

- **Finding 2.1. Most ICs don’t have dedicated staff or contractors to work with digital analytics data. – Needs Improvement**

  Lack of staff resources for analytics is more the rule than the exception among ICs. Some ICs have staff that dedicate about half their time to analytics and receive contractor support. Others have very few resources for analytics. This has created a “have and have not” environment, where some ICs have the potential to do great work with analytics and others appear to be stuck in quagmires.

  Other factors that contribute to uneven staff involvement with analytics may stem from where and how the IC Web sites are hosted (technically) and who is managing them (centralized or decentralized). We found that it’s common at NIH to have a very small central team (often the Communications Office) that’s trying to “herd the cats” of branches and divisions that own Web sites hosted on servers the team doesn’t manage or even, by organizational authority, influence.

- **Finding 2.2. ICs do not use collective power to purchase analytics hardware, software, or services. – Needs Improvement**

  The NIH organizational structure that encourages ICs to act as independent business units may be effective for managing resources for funding research, but it’s not at all effective for managing Web analytics resources. It leads to a shared-knowledge gap rather than vibrant group learning, and the amount of time and money wasted by working independently is very likely great.

  In certain interviews, we found that some ICs are doing analytics work that could help others, but it was never shared. This was not because an IC held onto it for competitive reasons, but because the idea of sharing was not part of the institutional consciousness.

These same silo issues could be seen with even greater financial impact around software and services procurement. Collective purchasing could reduce the cost and enable "poorer" ICs to benefit from the resources of ICs with more generous budgets.

- **Finding 2.3. There is an overall sense of frustration among ICs that rely on Web analytics services from central NIH IT organizations such as OIT and CIT (more so from CIT); ICs with Web analytics hosting and services from other sources have a relatively higher level of satisfaction. – Needs Improvement**

  CIT and OIT provide minimal Web analytics support to dozens of ICs, and this support is generally considered to be subpar. CIT and OIT offer Webtrends log-file-based analysis as part of a package of hosting services it sells to ICs. This support involves the hosting of Webtrends on the Web server and running the monthly reports available from Webtrends.

  This level of service has created challenges for ICs that want to have their analytics supported by the centralized IT groups, including these:

  - Many data-quality settings within Webtrends are not put in place. This means that the data generated from the tool are inaccurate and unreliable.
  - Interested users don't have administrative rights or access to the Webtrends reporting interface, so they can't request or customize reports to make them more useful.
  - CIT and OIT staff have not received training in Webtrends, even though the tool is very complex and they need to know how to use it to do a good job.

  The current funding framework is one of the root causes of this situation. CIT and OIT must charge for all services, because the services they provide aren't a budget line item. For CIT, this has led to charging noncompetitive prices. OIT and CIT perform similar services, but OIT provides them to a smaller universe of Web sites. While it appears that OIT is more engaged and more interested than CIT in supporting its stakeholders' Web analytics efforts, the same funding framework exists for both organizations, and it does not look like the current funding framework will change.

  Within this environment, ICs have sought Web-hosting and Web-analytics-hosting services from IT groups within their IC or from contractors. There appeared to be a much higher level of satisfaction with this arrangement than with relying on centralized NIH IT groups. In addition, with the growing use of Google Analytics in the government, we saw an encouraging trend toward ICs' planning to implement it on their own or using their contractors to host, analyze, and manage their sites.

- **Finding 2.4. There's no central group coordinating Web or digital analytics strategy. – Needs Improvement**

  Many of the enterprise-level organizations we work with have a central Web-governance council. This group can have a range of authority and influence that includes determining an organization's strategy for how to use digital channels, determining priorities for development, allocating funds, setting standards, and providing a forum for senior-level leadership to come to consensus on how the organization’s overall goals can best be reached through digital channels. NIH does not have such governance. In most of the conversations we had with NIH staff, there was no indication of an overall strategy or a concept of Web governance. The Communication Director meetings and various working groups, such as the NIH Web Metrics Group and the Web Authors Group (WAG), are primarily forums for sharing information. They are not policy- or standard-setting groups.

The Office of the Director’s Office of Communications and Public Liaison provides guidance on Internet-based communication and Web analytics, but it does not have the authority to play a role in negotiations with vendors for tool sets. The office is also not in a position to negotiate on behalf of ICs with OIT and CIT for service level agreements for Web analytics. This affects the use of Web analytics because there are no agreed-upon standards for how ICs should use Web analytics, or even that they should use analytics. There is no mechanism to provide added value for purchasing (see **Finding 2.2.**) or direction on how Web teams may uniformly tackle Web analytics, either technically or as a decision-support tool. All current efforts to coordinate Web site measurements don’t appear to have any authority, or “teeth.” This is, in large part, because decision-making gets done by the IC directors and senior managers who do not appear to see the Web as a critical business channel (see management commitment section, beginning on page 1).

- **Finding 2.5. OMB memos released in June 2010 provided guidance on lifting the ban on the use of persistent cookies, but many NIH employees are not clear about what that means. – Evolving**

  OMB memo M-10-22 allows the use of persistent cookies on Federal Government Web sites. Persistent cookies enable you to count unique visitors and improve the ability to measure repeat visitors, frequent visitors, and new visitors. Unique-visitor counting allows you to segment visitor traffic in Web analytics. This means you can do more accurate analysis and, in turn, have better analytics data.

  The guidelines enacted by OMB require changes to Web site privacy policies so that visitors know how they can opt out of tracking and how the data will be used. Complying with the new guidelines makes it possible for Federal agencies to use Google Analytics because it uses first-party cookies to count site visitors.

  NIH was among the first agencies to make the required privacy policy changes. However, these changes have not been enacted across all ICs. Our understanding is that each IC must permit the changes and get them approved separately. This strikes us as inefficient.

### *Recommendations*

- **Recommendation 2.a. Create a Digital Channel Program Office within the Office of the Director and an NIH Digital Channel Management Council.**

We recommend these two entities together because we see them as strongly related to each other's roles, effectiveness, and success.

The Digital Management Council should have a senior-level endorsement and mandate to

- determine digital-branding, user-experience, and information-architecture strategy and standards at NIH
- set digital-channel policies
- develop new products and services based on the use of analytics data
- have decision-making authority for centralized funding for Web initiatives
- establish formal processes for Web initiatives at NIH and within the ICs.

The Council should have a budget funded by line-item contribution from all ICs that will be used to fund new initiatives. Approval of new initiatives should be based on an evaluation process guided by the NIH digital strategy to be developed by the Council.

We believe the Council should have representation from the following offices or groups:

- Digital Channel Program Office
- Communications Directors
- Office of the Chief Information Officer
- Office of Planning and Communication
- Office of Intramural Research
- Office of Communications and Public Liaison
- Public Information Office.

The Digital Channel Program Office will have an executive-level endorsement and mandate to

- make recommendations to the Council on policies, standards, and initiatives
- implement Council-approved policies, standards, and initiatives
- develop new products and services based on the use of analytics data
- make decisions about centralized funding for Web initiatives
- negotiate authority with vendors for favorable procurement
- develop relationships with consultants and contractors who may provide services to ICs
- establish formal processes for Web initiatives at NIH and within the ICs
- coordinate tasks with CIT, OIT, and other ICs that may have resources to share
- manage NIH-wide initiatives and tool sets for
    - Web analytics
    - market research
    - surveys
    - user-experience testing
    - content management
    - internal search
    - social media management and measurement
    - email delivery
    - 508 compliance
    - Web site operations management

    - RSS and multimedia measurement.

The Office should have a budget funded by line-item contribution from all ICs that will be used to fund staff and tools. The Office may also receive funding from the Council to implement approved new initiatives.

The establishment and funding of the new Office will likely be most helpful to ICs that do not have the staff and budget to fully build out their digital initiatives. The Office should be able to

- allow resource-challenged ICs to tap into a greater pool of staff, consultants, and contractors
- negotiate favorable pricing and a wider array of tool sets and services
- create centralized management of tools sets
- manage, coordinate, and negotiate training in tool sets and best practices
- negotiate standard service-level agreements for Web operations and Web Analytics with contractors, consultants, and NIH service organizations, such as CIT and OIT.

The Council or Office should not replace the individual work that ICs do on their Web sites. However, having a central Office in place that can take on some Web operational tasks may allow ICs greater flexibility in planning how to best use budgets and resources.

- **Recommendation 2.b. Expand the scope of Office of the Director (OD) On-Line Information Branch so that it guides NIH Web initiatives.**

  While we would like to see a new Council and Office governance and funding approach established as described above, we understand that this may be a long-term plan. In the short to medium term, we would like to see many of the governance tasks, the associated staff, and funding located in the On-Line Information Branch within the OD Office of Communications and Public Liaison.

- **Recommendation 2.c. Use a team approach for planning, implementing, and using Web analytics so that the tools and data are managed, maintained, and used correctly.**

  Lack of time and lack of support were cited as major issues that discourage the use of Web analytics within ICs.

  Most organizations committed to using Web analytics for driving decision-making understand that there must be a team approach to successfully implementing, using, and training others on how to use the data. We see pockets of this approach at NIH in ICs that are considered to be well funded.

  Although we would like to see the creation of a Digital Channel Program Office, we understand that ICs may need to continue working within the status quo. In that case, we suggest using the following team model to guide continued staffing with full-time equivalents (FTEs), consultants, and contractors or through outsourcing analytics entirely. The roles listed below may be filled by the same person; for example, a project manager may also be the Web analyst.

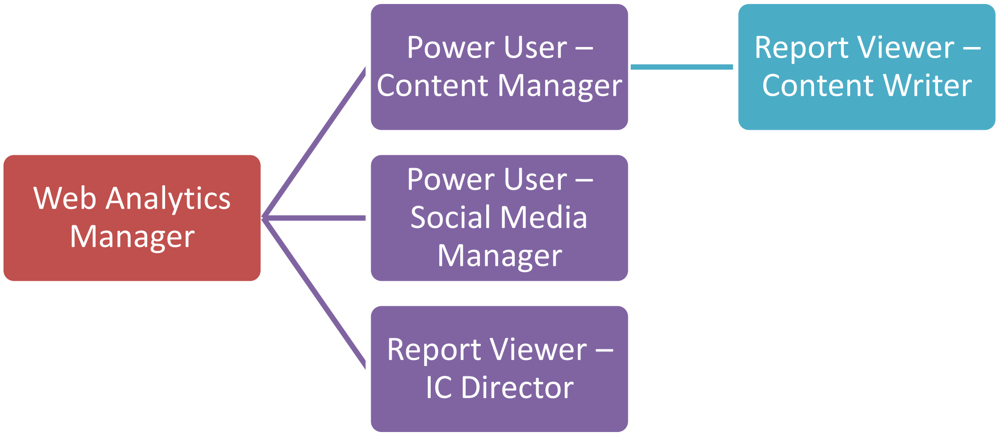
- Web analytics manager: This person will manage the overall Web analytics program; provide strategic direction for all Web analytics and audience and market research; and interface with senior management, supporting staff, and supporting operational teams, such as IT. Other duties will include
  - managing analytics team and technical support resources
  - developing research plans that include multiple Web site-measurement and testing methods for Web initiatives through the combined use of Web analytics, survey data, and focus-group, A/B, and user-experience testing
  - maintaining and managing vendor relationships
  - managing analytics and research education and training
  - promoting analytics and coordinating analytics initiatives among NIH ICs' sites at the management level.
- Web analyst: This person will be charged with the tactical aspects of conducting Web analytics tasks and activities. Duties will include
  - setting system permissions for people with access to reports
  - scheduling and publishing reports
  - providing diagnostic skills and the first level of support for system problems
  - developing metrics requirements with stakeholders
  - determining what data are required to develop metrics
  - analyzing reports, calculating metrics, and providing strategic and tactical recommendations based on metrics analysis
  - managing Web site-measurement projects that include initial implementation and new initiatives
  - conducting exploratory Web analytics projects
  - conducting training and educational assistance for analytics stakeholders
- Web analytics developer and system administrator: If the IC is using a JavaScript-based data-collection system hosted by the vendor, like Google Analytics, Webtrends OnDemand, or Adobe Omniture, this role should be held by a JavaScript developer. Duties should include placing JavaScript tags on CMS templates. If the IC is hosting, the role should be held by a systems administrator and/or database administrator, and duties should include
  - performing database maintenance and optimization tasks associated with software-based Web analytic applications, such as archiving, deleting expired data and backing up tape
  - monitoring data collection, data processing, and imports into the database
  - monitoring all software and hardware components of the site-measurement system
  - performing all required server maintenance tasks, such as installing security patches and updates.

- **Recommendation 2.d. Define user classes to clarify who has access to which data and in which formats and interfaces and to determine what training is necessary.**

  We recommend that there be an official definition of user classes for organizational considerations around tool permissions, training, and report distribution.

  A typical user-class structure is a hierarchy from Web analytics manager/Web analyst → power user → report viewer (see **Figure 3**). For example, the Web analytics manager may be in the Communications Office, the power user may be in a division, and the report user may be a subject-matter expert who writes for a division Web site.
  - The power user may access the Web analytics solution, have permissions control, and run reports and conduct analyses for people within the division or group.
  - Report users can view high-level reports or metrics calculated for them by either the Web analytics manager/Web analyst or a power user.

***Figure 3.*** *Example of a Web analytics user hierarchy.*


Web Analytics Manager
Power User – Content Manager
Report Viewer – Content Writer
Power User – Social Media Manager
Report Viewer – IC Director

- **Recommendation 2.e. Create Web analytics user groups within ICs.**

  There's an NIH-wide Web metrics group that provides a forum for information exchange and dialogue among people using Web analytics. We believe that this is a good model for IC-level Web analytics managers to use as they work to broaden the scope and use of Web analytics among their IC's divisions and offices. The group meets to learn about how to use data to optimize site design and content and to discuss issues that may be getting in the way of wider use of analytics.

- **Recommendation 2.f. Link performance to Web metrics to increase accountability.**

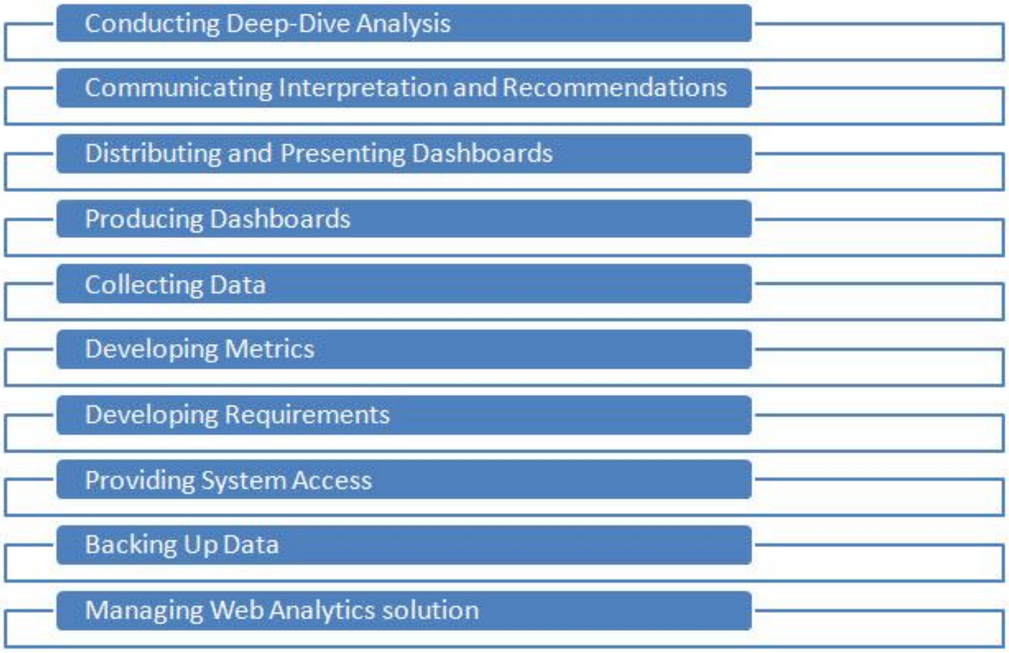
  We are beginning to see in organizations that have committed themselves to the importance of the online channel a corresponding change in employee evaluation. In this model, employee performance is linked to whether site content and outreach campaigns are meeting goals and targets that can be measured through Web analytics metrics. This approach should be studied further for NIH ICs after business-driven success metrics (described in the metrics section, beginning on page 20) are put into practice and considered reliable. Performance should not be linked to high-level metrics, such as increase in page views, visits, and visitors, since these are not linked to achievement of Web site goals and objectives.

### 3. Process for NIH-Wide and IC-Specific Web Analytics Governance

Successful analytics initiatives depend on a governance framework that includes systems for data collection, processing, and integration; metrics development and presentation; data analysis; and ongoing maintenance of the Web analytics solution. At NIH, there are significant challenges to overcome in managing these processes and ensuring that there are employees or contractors to fulfill the required roles.

Having the right roles, responsibilities, and resources assigned to Web analytics is crucial to success, but they must work in coordination with each other. Creating processes around Web analytics enables people to work well together, especially to ensure timely data collection, report building, distribution of dashboards, interpretation of data, and presentation of recommendations. If we were to look from a high level at the processes that contribute to a successful Web analytics initiative, we'd see that there are at least 11 processes—some working independently, some overlapping, and some requiring hand-offs or collaboration between the Web analytics program and another group. All these processes are dependent on each other, moving from bottom to top, as shown in **Figure 4**.

***Figure 4.*** *Processes that contribute to successful Web analytics initiatives.*


Conducting Deep-Dive Analysis
Communicating Interpretation and Recommendations
Distributing and Presenting Dashboards
Producing Dashboards
Collecting Data
Developing Metrics
Developing Requirements
Providing System Access
Backing Up Data
Managing Web Analytics solution

In general, there's not a great deal of consistent governance around Web analytics tasks and activities throughout NIH. As with other areas of Web analytics, much of this relates to available time and resources. Those ICs with more time and resources do tend to have more structure and governance around the performance of Web analytics tasks. Even in those ICs, though, we still see areas needing improvement, especially in how Web analytics fits into site development and campaign tracking.

### *Findings*

- **Finding 3.1. Few ICs have processes for translating Web site goals into metrics that can measure progress in achieving these goals. – Needs Improvement**

  Metrics-requirement development does not really exist throughout most of NIH. ICs receive what's available out-of-the-box from Webtrends and other analytics packages. We saw only rare instances of ICs' creating customized metrics based on business goals, and this occurred where the IC had some measure of control over its analytics tools. In most situations, ICs do not have control over their analytics tools and either get frustrated because it is difficult to go through the dozens of pages in out-of-the-box reports or to spend hours looking for data points in the reports. This situation is frustrating and discouraging to many.

  Because there has been little strategic direction around Web site development, most of the metrics Web managers use are for improving the Web site–user interface and redesigning information architecture. The redesign process tends to take a long time for the ICs we spoke with, and many sites have not had redesigns in several years. Some ICs have decided to take a phased approach to site redesign that enables them to make content and design upgrades section by section.

  The multipage documents produced by Webtrends and some of the contractors tend to provide high-level historical summaries. Although some stakeholders said that these reports were helpful in evaluating the types of content of interest to their users, most said they were hard to understand. The overall lack of confidence in the accuracy of Web analytics data also discouraged much activity around using Web analytics and developing new types of metrics. As a result, the development of new metrics at NIH appears to be at a point of stagnation.

- **Finding 3.2. Data-collection standards and preparation for data collection are sporadic. – Needs Improvement**

  There are no standards in place to enhance basic implementation of Google Analytics or to ensure consistent installations of Google Analytics data collection. Most of the Google Analytics implementations are handled by contractors, so there are no implementation standards for creating global NIH reporting or conventions around collection of data from internal search, forms, links, and campaigns. (For our discussion of data-processing issues associated with Webtrends, see the organization section, beginning on page 5.)

  If NIH chooses to implement Google Analytics globally, governance will need to be in place first for the assignment of data-collection variables so that there's a standard in place to create Web analytics metrics. We suspect that this issue could become important when there are more "hands in the pot" around Web analytics, especially among digital-agency contractors that may be launching sites and implementing the Web analytics code without any oversight or input from a central team.

- **Finding 3.3. Web analytics is not "baked into" the Web site or content-development lifecycle. – Needs Improvement**

  Even in more sophisticated digital organizations, Web analytics is "left out" of the Web-content-development process because analytics requirements were not included in the initial site or online application requirements. We believe that one of the reasons NIH Web teams are challenged to "triangulate" data sources such as Web analytics, voice-of-customer data, and user-experience test results is because there's no planning up front for how all the potential data sources may be used. This makes it difficult to put the pieces of the puzzle together after new sites, applications, and content are launched.

- **Finding 3.4. In most NIH ICs, one person or a small group of people uses Web analytics. – Needs Improvement**

  In the current environment, those who do receive analytics reports generally don't share them widely. They also don't create derivative analytics reports that could be shared. High-level traffic metrics are occasionally shared with IC directors and a variety of stakeholders. However, there are few examples where metrics have been distributed to a wider audience of content providers and communication teams. Some of this may be attributed to general lack of interest and trust in the data, and some to the difficulty in gleaning anything useful in the formats most commonly available. Similarly, the processes used for producing reports, providing analysis, answering questions, and following up with stakeholders to train them on how they might use analytics reports are generally ad hoc.

### *Recommendations*

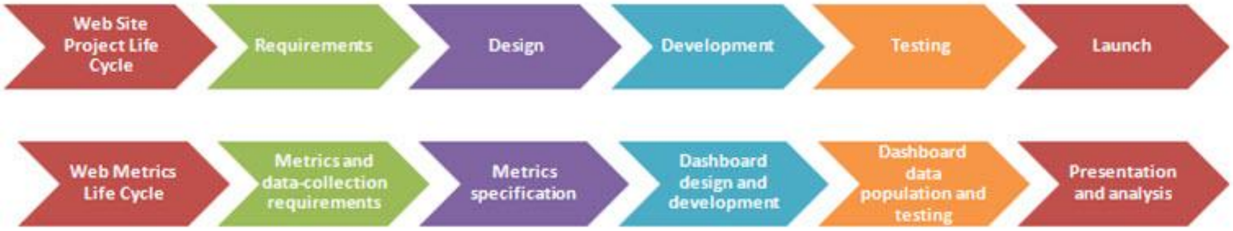
- **Recommendation 3.a. Formalize a project-planning process in which Web analytics metrics and reports are developed, produced, and tested in tandem with new Web applications, content, and outreach campaigns.**

  It helps to think about producing Web analytics within the context of existing processes.

  Formalizing the integration of analytics into Web content and development cycles should be a priority. This will ensure that analytics is considered early in the development process for site redesigns and application-development projects. For tasks that occur more frequently, such as introducing a new marketing campaign or new site content, we recommend documenting standard operating procedures that ensure that all elements of metrics development occur. We find that the breakdown in data collection caused by scheduling pressures or a lack of understanding about what's required to ensure data collection is common in all organizations.

We have found that the software-development-lifecycle process is a good model to use when developing Web analytics. It helps people who aren't familiar with analytics understand that it's similar in nature to other application-development and coding projects. In **Figure 5**, we map the process of developing analytics metrics and reporting to a basic Web-application or software-development lifecycle.

***Figure 5.*** *Comparison of the life cycles of Web site projects and metrics.*

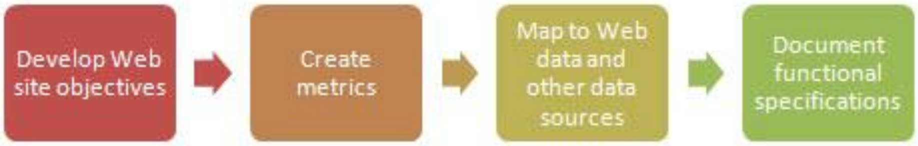

Web Site Project Life Cycle
Requirements
Design
Development
Testing
Launch
Web Metrics Life Cycle
Metrics and data-collection requirements
Metrics specification
Dashboard design and development
Dashboard data population and testing
Presentation and analysis

- **Recommendation 3.b. Establish a standard approach for developing Web metrics.**

  Within the larger project-development process for including Web metrics, subsidiary processes need to be considered for dashboard development and deep-dive analysis.

  **Figure 6** illustrates one standard way to address the development of metrics from Web site objectives to a functional specification that can be used by a Web developer to ensure data collection and by a report developer to create scheduled reports.

***Figure 6***. *A standard approach for developing Web metrics.*


Develop Web site objectives
Create metrics
Map to Web data and other data sources
Document functional specifications

  Deep-dive analysis requires a different process, beginning with a specific set of research questions that guide queries about Web analytics data sets.

- **Recommendation 3.c. Develop data-collection frameworks and standards to ensure data integrity.**

  Many teams, groups, and contractors are collecting and using Web analytics data at NIIH, but there are no standards for what data to collect, what data not to collect, what type of standard methods should be used for data collection, and so forth.

  If NIH ever wants to be able to compare IC traffic or report holistically on NIH Web sites in total, there must be agreement on standards for data collection that are used as guidelines by NIH staff, contractors, or agencies. We expect that the increased use of Google Analytics will lead to increased involvement from external contractors, making it even more critical to develop governance and data-collection standards. These standards should also include establishing conventions for campaigns, URLs, and page names throughout NIH. For example, a standard for campaign-tracking-code values should be established, and the parameter for the campaign-tracking-code variable (e.g. "src=", "s_src=", "s_subsrc=") should be identified and universally used. Content and campaign managers must be made aware of this parameter and guided as they incorporate these tracking codes into their incoming campaign URLs. Any vendors must also be made aware of the parameter.

To ensure precise and consistent reporting, a core list of parameters should be compiled and distributed to all content-management teams. For example, if a content manager wants to distinguish between a clinical research information page and a grant information page, there should be consistent naming conventions for URLs and pages.

For guidelines on how to collect accurate analytics data, please refer to **Appendix C.**

- **Recommendation 3.d. Formalize protocols for basic Web analytics tasks such as presenting dashboards and data, responding to ad hoc data requests, and accessing the Web analytics solution.**

  Some NIH Web analysts are busier than others with fielding requests for reports. We recommend that Web analysts create a process for logging in and evaluating all requests so that there's a recognized workflow for these tasks. It's important to make a formal announcement of the process, including documentation that explains to stakeholders how to obtain Web analytics reports, to avoid bottlenecks in producing reports and responding to requests.

  New requests should be evaluated within a consistent framework. For example,

  - Request
    - New requirements or revision?
    - Purpose: implementation, reports, analysis, data issues, new content, or application?
  - Evaluation
    - Rate by a priority system, or first-come, first-served
    - Determine task, resource, and level of effort (site size, data source, type of presentation, type of analysis, development of filters, customization, calculated metrics)
  - Response
    - Follow up with requestor, and name an expected delivery date.

- **Recommendation 3.e. Support stakeholders' planning processes by developing strategies for distributing reports, interpreting them, and answering questions. Fulfilling report requirements through the production of a dashboard or requested report does not end with the email distribution to the stakeholder.**

  We recommend that NIH Web analysts consider strategies to follow up with stakeholders after report delivery to review the data, provide analysis, and answer questions to ensure that the reports are not only getting used, but getting used correctly. This is an opportunity both to educate and advocate for the value of Web analytics. For example, new reports could be delivered with the instruction that the Web analytics program will contact the recipient in a week to set up a meeting to discuss the results. This analysis session could take place a few times or until the stakeholder seems to "get it." It could also be an opportunity to hear first-hand from stakeholders about new questions they have about the data and new requirements they need to fulfill.

- **Recommendation 3.f. Simplify and clarify the procurement process for obtaining analytics support.**

  ICs are often stymied by how to obtain resources for contracted analytics support. Some have stories of waiting months for requests to go through the procurement process. No one seemed to know the reasons for the delay, but it's clear that this is a bottleneck that requires attention. A dedicated central office (see the organization section, beginning on page 5) might alleviate this issue. However, establishing this office will require an understanding of procurement issues and processes.

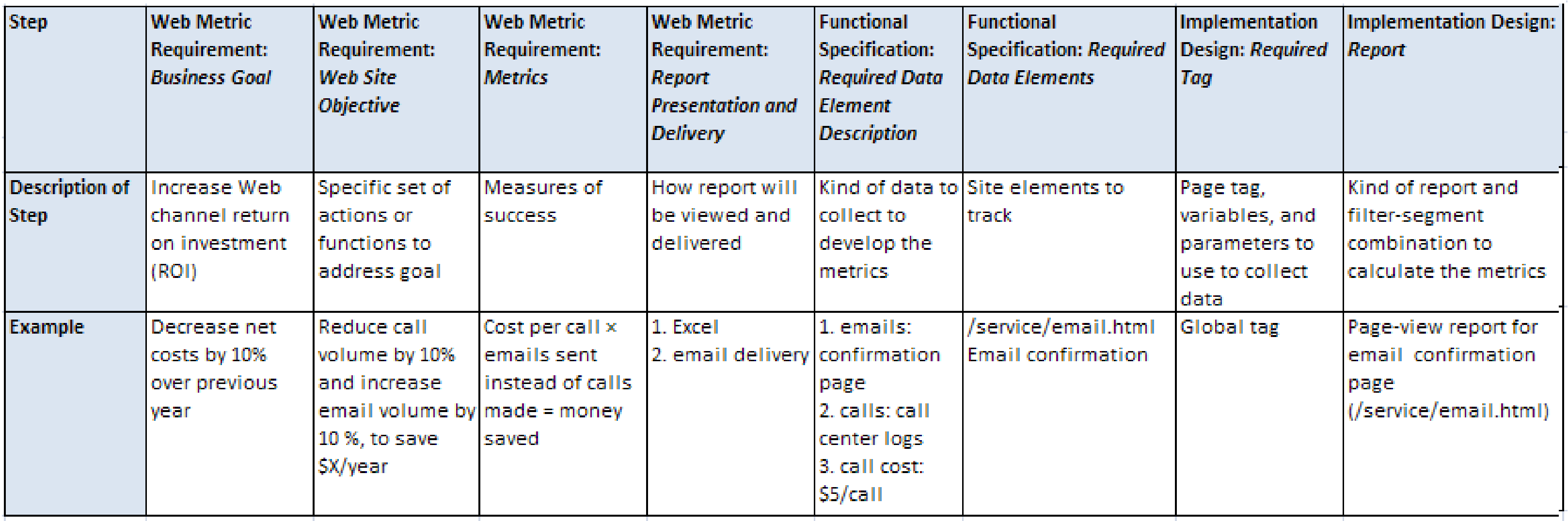
- **Recommendation 3.g. Develop a functional-specification template to map Web analytics requirements to Web analytics report building and data collection.**

  Functional specifications provide complete documentation, from reporting to data collection, and standards to use when developing reports and metrics. **Figure 7** is an example of documentation that brings together all the elements of Web analytics reporting. Functional specifications are a logical outgrowth of the Web metrics recommendation process.

- **Recommendation 3.h. Develop and maintain implementation standards for Web analytics data collection.**

  The purpose of these standards is to provide guidance to the internal team, agency, and third-party vendors that may be managing the Web analytics data collection. If standards are not implemented, there's a real potential for lack of coordination in assigning variables (for JavaScript tagging) and processing logs (for Webserver log analysis). The standards provide specific direction where possible, guidelines as appropriate, and explicit customization when it's demanded or desirable.

***Figure 7.*** *Web metrics functional specification example.*



| **Step** | **Web Metric Requirement: *Business Goal*** | **Web Metric Requirement: *Web Site Objective*** | **Web Metric Requirement: *Metrics*** | **Web Metric Requirement: *Report Presentation and Delivery*** | **Functional Specification: *Required Data Element Description*** | **Functional Specification: *Required Data Elements*** | **Implementation Design: *Required Tag*** | **Implementation Design: *Report*** |
|---|---|---|---|---|---|---|---|---|
| **Description of Step** | Increase Web channel return on investment (ROI) | Specific set of actions or functions to address goal | Measures of success | How report will be viewed and delivered | Kind of data to collect to develop the metrics | Site elements to track | Page tag, variables, and parameters to use to collect data | Kind of report and filter-segment combination to calculate the metrics |
| **Example** | Decrease net costs by 10% over previous year | Reduce call volume by 10% and increase email volume by 10 %, to save $X/year | Cost per call × emails sent instead of calls made = money saved | 1. Excel<br>2. email delivery | 1. emails: confirmation page<br>2. calls: call center logs<br>3. call cost: $5/call | /service/email.html<br>Email confirmation | Global tag | Page-view report for email confirmation page (/service/email.html) |

- **Recommendation 3.i. Develop and maintain standard operating procedures for all who support the Web analytics initiative.**

  A Web analytics standard operating procedures (SOPs) document should include descriptions and guidance for all those who support the Web analytics initiative. Its value is as a reference, especially for new staff, since it provides an index to all associated documents both online and offline, contact information, checklists, and so forth. SOPs require scheduled updating and provide a useful way to ensure that processes are documented. A Web analytics SOP for an IC could include the following types of information:
  - data-collection standards, such as query-string campaign conventions, page-title conventions, data filters, and methodology for capturing video and audio
  - administrative settings and permissions structure for user groups
  - vendor contract details
  - report set definitions
  - metrics calculations
  - report-distribution settings and schedule
  - contact lists for all stakeholders, including organizational, vendor, and consultant support
  - scheduled presentations to stakeholders and management with examples provided
  - processes for developing new metrics and reports
  - processes for stakeholders to follow to request new data collection and reporting.

## 4. Web Analytics Metrics and Deep-Dive Analysis

Effective Web analytics relies on having the right data and analysis for strategic and tactical decision-making and on presenting the data in ways that specific users can understand, whether they are senior managers, communications directors, Web team managers, marketers, designers, or content editors. We talk about Web analytics measurement in terms of key performance indicators (KPIs), metrics, and analysis, but these terms have multiple meanings. At NIH, there are many different views on which measurements are valuable for Web content and communications.

Compared with the .com and .org sectors, .gov Web analytics is still rudimentary, especially with its reliance on traffic-based metrics. On the other hand, .gov staff members use many analytics tools and methods that in some ways are more diverse than the ones their .com and .org counterparts use.

In the stakeholder survey we conducted for this best practices initiative, most respondents indicated that consumption of content was the primary goal of their Web sites. However, in no case during the discovery phase of the project was the consumption of content ever linked to an increase in the budget or additional resources for the Web site, nor was it linked to lowering Web site or IC operational costs or to a verifiable assurance that the IC mission goals were being met.

In fact, just the opposite is true. Although almost every NIH Web site could point to increases in content consumption, Web site budgets are generally flat year over year, Web site staffing is lean and often not full time, and there is little recognition among senior management that the Web site has any quantifiable business value. In fact, much of the increase in content publishing and consumption on NIH sites is due to greater attention from the public, not from the NIH mission's implied core professional audiences of scientists and researchers:

> "NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce the burdens of illness and disability." (From http://www.nih.gov/about/mission.htm)
>
> "The goals of the agency are
> - to foster fundamental creative discoveries, innovative research strategies, and their applications as a basis for ultimately protecting and improving health;
> - to develop, maintain, and renew scientific human and physical resources that will ensure the Nation's capability to prevent disease;
> - to expand the knowledge base in medical and associated sciences in order to enhance the Nation's economic well-being and ensure a continued high return on the public investment in research; and
> - to exemplify and promote the highest level of scientific integrity, public accountability, and social responsibility in the conduct of science. "
> (From http://www.nih.gov/about/mission.htm)

People who were interviewed for this project or who participated in the survey pointed to the value of providing information to the public based on the results of NIH research. This is an intrinsically important function. However, funding of NIH Web site initiatives is often inconsistent.

In this section of the report, we'll review findings related to the use of Web metrics at NIH and provide recommendations for how to create metrics that go beyond the traffic reporting the ICs are generally using today.

### *Findings*

NIH spends millions of dollars on its Web sites, but it often knows very little about how the sites perform and even less about the sites' return on investment.

- **Finding 4.1. Using Web metrics to drive business decisions. – Evolving**

  ICs leverage basic, out-of-the-box Web analytics reports for overview reporting, whether they are from Webtrends, Google Analytics, or another tool. The metrics most often used are high-level traffic-trend data that track visits, page views, visitors, and referrals to the Web sites. Stakeholders say this information is useful in guiding decisions about site content, design, and navigation, as well as for outreach. They also say that it's difficult to get to the "nuggets" of truly useful information. They point out, and rightly so, that trending information and technical data are not very intuitive from a "plain language" perspective. In addition, the amount of data from the "out of the box" reporting is not well targeted to specific areas of the sites, so Web analysts need to comb through a ton of data to get to the really relevant parts. These challenges make it hard for staff to develop and then present meaningful recommendations to senior management.

- **Finding 4.2. Standard set of Web metrics for NIH. – Needs Improvement**

  There's no standard set of NIH-wide metrics for ICs to collect. For the most part, each IC operates separately and knows little about the metrics others are using. Many ICs get canned reports from CIT, OIT, and their contractors that are "spit out" from Webtrends on a monthly or quarterly basis. These reports contain the dozens of metrics available in Webtrends. These are presented either as high-level historical trends that occur over specific time periods or as "top" numbers, such as Top 50 Page Views and Top Referrals.

  Google Analytics is gaining popularity among ICs that have either been frustrated with Webtrends reporting or lacked the budget to own a Web analytics tool. Google Analytics provides its data "on demand" as soon as a user logs in to the account. The Google Analytics metrics are very similar to Webtrends', except that Google's are easier to access. In both cases, determining the basic set of metrics is largely driven by what's available from the analytics tool rather than by what's most valuable for decision-making.

  Some ICs have been using Web site goals to determine what metrics to adopt. For example, NCI, OD-OSE, NLM, NIAID, and NHLBI have gone through a process of defining goal-based metrics to use in presentations to senior management, content editors, and communication staff.

  Having a consensus around which metrics to use and how to use them would encourage shared learning opportunities among ICs and help foster a climate of decision-making based on Web analytics data. No matter which metrics NIH adopts as standard, ICs will have to be able to generate them from different types of Web analytics software.

- **Finding 4.3. Benchmarking of Web analytics data among NIH ICs. – Needs Improvement**

  ICs don't know how to compare their sites' performance with that of other ICs. Currently, the kind of information people in the NIH Web analytics community share with each other is only anecdotal. The tricky component in any benchmarking exercise is how to create benchmarks that "compare apples to apples." This concept is universally accepted in principle. For example, ICs with smaller budgets do not want to be compared with ICs with larger budgets; ICs that have a more-focused mission or a smaller inherent target audience do not necessarily want to be compared with ICs with a broader mission or a larger baseline audience.

  There's more to be gained from sharing data and information than from not sharing them. Sharing information leads to greater learning, and a small dose of competitive spirit often leads to better performance all around. The challenge is in coming up with benchmarks that are actually meaningful.

  NIH has a history of using the ForeSee American Customer Satisfaction Index (ACSI), a voice-of-customer (online survey) that uses the ForeSee provides scores ranging from zero to 100. While it makes interesting conversation to compare one site with another this way, there can be too much focus on the score and not enough on the factors about the Web site that influenced the score.

- **Finding 4.4. Benchmarking of Web analytics data from Federal agencies, nonprofits, and healthcare and educational institutions with similar missions. – Needs Improvement**

  Most ICs don't know how to compare their sites' performance with that of other Federal Government agencies and organizations with similar missions in the .edu, .org, and .com sectors. There's increasing interest in institutions with an online presence that compete with NIH for credibility and authority in their particular specialties. Although NIH ICs generally rank high in online search results, it's understood that this metric reveals just part of the entire story. There are questions about where visitors come from before they get to NIH, where they go afterwards, how often they go to NIH sites compared with others, and which organizations are the main "competition." These questions can't be answered through Web analytics solutions, and they are really important to IC directors and senior NIH management.

- **Finding 4.5. Social media measurement metrics and reporting. – Needs Improvement**

  Social media is garnering a great deal of interest at NIH, but it's challenging to measure the return on investment (ROI) and success of the social media campaigns. Many ICs have Facebook pages, Twitter feeds, and YouTube channels. Some are deriving an understanding of their site's traffic from the metrics available through these channels. A few have started to make the investment in specialized social media measurement tools such as Radian 6. However, few methodologies are available for developing metrics that are related to business goals or for validating social media strategies.

  Because social media is a new communication channel, there's currently more experimentation than planning or purpose going on with it. There's certainly a "shiny new thing" aura around it. As long as there's a popular fascination with social media, it will continue to garner attention from senior management.

  We expect these social media initiatives to be funded as a matter of course, in much the same way that Web site initiatives are funded — in most cases, minimally. However, if ICs are to start making targeted investments in effective communication channels, they will need to be able to evaluate social media as it compares with fixed Web and offline channels. This will require the development of standard metrics for social media and ways to better evaluate the efficiency and return on investment of social media initiatives. There needs to be a better understanding of the people in the social media audience, such as what they are expecting to get from social media sites.

- **Finding 4.6. NIH ICs use a combination of Web analytics, user-experience testing, focus groups, survey data, market research sources, and industry benchmarking to understand the effectiveness of their Web sites. – Evolving**

There is no comprehensive Web site-measurement and testing strategy that brings together Web analytics, audience and market research, and user-experience testing. “Triangulation” is a popular concept among people who work with Web site content, planning, and design. It refers to the practice of taking multiple sources of Web site evaluation data, such as Web analytics reports, user-experience testing, and voice-of-customer surveys, and using them in a complementary manner to inform decisions about Web site design. In addition, many ICs also use Web-based clipping services to keep track of the pickup of press releases and news stories by the media. These data are generally viewed independently of each other.

However, many are not quite sure about the best way to use these data sets to reach conclusions. In addition, the data sets are often viewed separately, making it hard to understand the relationships among the data, or the data sets are so large that it’s hard to find the important parts. There are also questions about process, method, and timing around the different analytics to use during the course of a Web initiative.

In short, the methodology for coordinating the use of multiple analytics approaches and tools needs improvement. Achieving this will help guide strategy and tactics for developing content, user experience, and outreach.

- **Finding 4.7. Contextual use of customer-survey data. – Evolving**

  There is a love-hate relationship with customer-survey data at NIH. The concept of survey data generates passionate responses, most notably among people using the ForeSee Results ACSI survey. Some swear by it and consider it the baseline set of metrics that should guide Web site change; others believe it has no validity at all due to small sample size; others would like to use it but have no budget for it; and still others use the data along with Web analytics and user-experience testing.

  One reason that ForeSee is so influential is that the Federal Consulting Group has made favorable implementation and cost arrangements for it. Another is the blanket clearance ForeSee has received from the Office of Management and Budget (OMB). This clearance allows ICs to survey more than nine people without additional OMB approval. It also allows for all custom survey questions without specific approval for each one.

  Some Web managers at NIH are interested in using survey instruments such as SurveyMonkey and Zoomerang for researching post–Web site visits, but until recently, these were constrained by OMB’s lengthy approval processes. New policies encourage and allow for quick approval of surveys and other tools for collecting customer feedback.

  Finally, some Web managers who are using ForeSee survey data are interested in integrating the data with Web analytics data. This is possible at a cost. Conceptually, it would allow Web analysts to link site behavior to the survey respondents beyond their survey session. This would help Web managers ascertain which pages, site sections, and tasks survey respondents are visiting over time, which would provide insights into the survey results.

- **Finding 4.8. Measuring Web site content effectiveness. – Evolving**

  Figuring out how to evaluate the relative value of site content is a challenge for site managers and editors. One person we interviewed observed that the only important metric for some NIH Web site managers is the volume of content they publish. Some managers also have either ad hoc or scheduled processes for reviewing site content and pruning and updating it based on its relevance and Web analytics data. Many others don't engage in this process and aren't encouraged to do so by their managers. Still other staff may not have time for content evaluation because of their workload and availability of resources.

  When asked on the survey about what a visitor to the site should accomplish, which is referred to as a "success event," 78% of respondents chose "consuming content" — far and away more than any other option. In interviews, "consuming content" meant page views and PDF downloads. Neither of these is a strong metric by itself. They could be strengthened if the visits and visitors could be divided into categories based on visitor behavior, such as whether the visitors are new or returning.

- **Finding 4.9. Return-on-investment metrics. – Needs Improvement**

  The relationship between a Web site's budget and the alignment of the site's performance with Web site goals and the return on investment (from the budget and resources) is not generally understood, nor is understanding it a priority. Conversations with IC Web site teams generally indicated that their budgets are set arbitrarily and are not based on how the sites perform. This contributes to the sense that there's little need to justify whether a site is operating efficiently from the perspective of labor and capital cost expenditure, and it sets the bar for success fairly low. Furthermore, there's a pervasive theme around the connection among cost, budget, and Web site operations that can best be articulated as, "We're the government. Our site is not measured or evaluated in terms of dollars." We take issue with this sentiment. While it may be true that NIH sites do not offer a product that people pay for, return on investment can be measured in other ways, such as

  - the Web as a cost-effective alternative to offline communications methods
  - the Web as a means to measure the success of an IC's mission-focused initiatives that influence budget allocations
  - the testing and evaluation of Web content and applications to determine whether similar content and applications should continue to be funded
  - justification of "indirect" revenue in the form of set-aside or grant funding for Web initiatives.

- **Finding 4.10. Measurement of site navigation and usability. – Evolving**

  Although many ICs consider navigation and usability top priorities, they aren't giving Web analytics the prominent role it deserves in supporting those functions. Navigation and usability metrics often go hand-in-hand in helping Web managers understand just how effective a site is in getting users to the information they are looking for and to the success events the site is designed to achieve. Many ICs use ForeSee survey and user-experience testing instead. Web analytics, if used at all, is a rather blunt tool for this. It can generate internal-search keyword reports, but these are used as a guide to determine what content is of interest. Exit-page reports are produced in some cases, but people don't necessarily understand how these data fit into a larger navigation or information-architecture strategy. Many people say they need path reports to understand navigation. However, effective path reports generally need to be configured or customized to be of any real value. Most standard path reports available in Web analytics solutions don't yield data that show how visitors traverse a site. The reports generally start on the home page and then show that visitors take many paths from there. It's difficult to discern whether one path is more important than any other.

- **Finding 4.11. Internal-search analysis. – Evolving**

  Search analysis is one of the most common measurement methods used at NIH, but its value in guiding search engine improvement is inconsistent. Nearly 59% of respondents in the survey for this project said they conducted search analytics. This is the same percentage who said they measure customer satisfaction. In most cases, "search analytics" means only that the search terms people use are collected and tabulated. This information is conveyed in "most popular search term" reports. Some ICs also use the custom questions in customer-satisfaction surveys to understand the relative effectiveness of search engine performance for certain terms. However, ICs don't often track general activity, such as numbers of started searches, completed searches, and searches per session, which can yield insight into the overall effectiveness of the search engine. Most Web analytics tools available on the market today allow this kind of tracking, but the data collection has to be customized. This customization isn't often done at NIH.

  To truly understand site search, it's often necessary to go beyond the reports on general use of searched terms and completion of searches. For example, understanding the number of site exits per term or the number of pages per term would reveal the relative success of the term itself in providing a visitor with the required information. This level of reporting requires more work than the standard metrics available in most analytics packages call for.

- **Finding 4.12. Use of visitor segmentation to understand visitors' behavior on Web sites. – Needs Improvement**

  ICs have trouble understanding visitor behavior on their Web sites at an audience-specific level. NIH Web teams know who their audiences are, to some extent, as shown in the survey results in **Figure 8**. However, the challenge is in understanding how these audiences use the NIH Web sites. If more were known about this, Web managers could develop more-targeted content, applications, and outreach programs. It could also help focus public-awareness campaigns, which would mean more effective allocation of campaign budgets. Although some visitor personas have been developed in conjunction with the ForeSee survey, they haven't been used for analyzing visit behavior across the NIH Web sites. Personas are fictitious characters created from demographic, behavioral, click-stream, and other data about Web site users. By representing qualitative and other visitor data with human faces and personal stories, personas help Web teams gain an understanding of their visitors' needs, goals, and challenges.

***Figure 8.*** *Who are the primary audiences for this Web site? (Check all that apply.)*

| **Answer Options** | **Response Percent** | **Response Count** |
|---|---|---|
| Advocates | 55% | 47 |
| Consumers | 65% | 55 |
| Educators | 49% | 42 |
| Health care providers | 52% | 44 |
| Journalists, reporters, or media representatives | 45% | 38 |
| Legislators and policy makers | 41% | 35 |
| Patients, family, and friends | 67% | 57 |
| Postdocs and fellows | 49% | 42 |
| Public health officials | 39% | 33 |
| Researchers, scientists, and investigators | 78% | 66 |
| Students | 54% | 46 |
| Staff | 40% | 34 |
| Other (please specify) | 18% | 15 |
| Answered question | 100% | 85 |

*Source: Developing Web Analytics Best Practices for NIH Survey, 2011.*

Typically, Web analytics uses both visitor and visit-based segmentation. Using geographic origin is also common and is used by some ICs. Others use knowing the source of a visit, such as whether it originated from a .com, .org, or .edu address. While ICs do use these methods to understand and interpret Web analytics reports, we didn't see them using the data to inform decision-making. In general, at NIH Web site decisions are made with a very broad audience in mind.

Visitor segmentation is very much based on the concept of the "unique visitor." NIH hasn't been able to count unique visitors because that relies on the first-party cookie, a code that allows the counting of individual computers. (*Persistent* cookies are first-party cookies that count computers, which we call "visitors," over a period of time. If a site serves a first-party cookie to a visitor, that visitor is recognized at subsequent visits and counted as a specific repeat visitor. Note that cookies count only the computer or device that accesses the Internet, not individual people.) Deployment of first-party visitor cookies was constrained until June 2010, when the Office of Management and Budget altered the policy to allow their use. NIH.gov made the necessary revisions to the privacy policy, and other ICs have followed suit. With this new freedom, NIH sites may now track first-time and repeat visitors. These are both common visitor-segmentation metrics.

- **Finding 4.13. Measuring visitor acquisition. – Evolving**

  ICs are devoting resources to online communication and marketing campaigns and are beginning to consider how to measure success. Many ICs that have Web sites managed by Communications Departments understand the need to measure effectiveness of visitor acquisition through marketing campaigns, email-newsletter registrations, search engine optimization, offline promotions, public relations campaigns, and other vehicles. **Figure 9** shows the online methods NIH ICs use to encourage people to visit their sites; **Figure 10** shows the offline options they use.

***Figure 9.*** *What ONLINE methods does your Institute or Center use to drive or encourage visits to the site? (Check all that apply.)*

| Answer Options | Response Percent | Response Count |
|---|---|---|
| Email newsletters or RSS feeds | 78% | 68 |
| Links from partner or affiliate organizations | 43% | 37 |
| Onsite promotion | 30% | 26 |
| Press releases (online) | 51% | 44 |
| Search engine optimization (SEO) | 40% | 35 |
| SMS text or other alerts | 0% | 0 |
| Social media (Facebook, Twitter, etc.) | 59% | 51 |
| Don't know | 6% | 5 |
| None | 5% | 4 |
| Other (please specify) | 13% | 11 |
| Answered question | 100% | 87 |

*Source: Developing Web Analytics Best Practices for NIH Survey, 2011.*

***Figure 10.*** *What OFFLINE methods does your Institute or Center use to drive or encourage visits to the site? (Check all that apply.)*

| Answer Options | Response Percent | Response Count |
|---|---|---|
| Brochures (printed) | 69% | 59 |
| Direct mail (postal) | 12% | 10 |
| Exhibits and trade shows | 68% | 58 |
| News articles (print) | 40% | 34 |
| Outdoor advertising | 9% | 8 |
| Press releases | 48% | 41 |
| Radio and television programs | 20% | 17 |
| Don't know | 6% | 5 |
| None | 11% | 9 |
| Other (please specify) | 14% | 12 |
| Answered question | 100% | 85 |

*Source: Developing Web Analytics Best Practices for NIH Survey, 2011.*

ICs that subscribe to the GovDelivery email-newsletter service can use its metrics, such as sends, opens, and click-through rates. Those that don't use GovDelivery use metrics from their email-fulfillment vendor, if the vendor provides them. Communications departments also use third-party clipping services to track how press releases to online and offline media outlets are being distributed. This information can be correlated with referral data recorded by Web analytics' tools, which can then be used to help understand the referral power of these media outlets. We didn't see many examples where this correlation was being done at NIH. While some ICs go as far as using campaign codes (unique codes appended to URLs) and track the success of specific campaigns, others don't track campaigns at all. Tracking campaign effectiveness is one of the best ways to understand how to allocate the budget: keep doing what works, and trim back what doesn't. This is an area that would benefit from metrics based on visitors' taking actions or accomplishing tasks on sites. Such metrics are called "conversion-oriented" metrics.

- **Finding 4.14 Tracking Section 508 compliance. – Good**

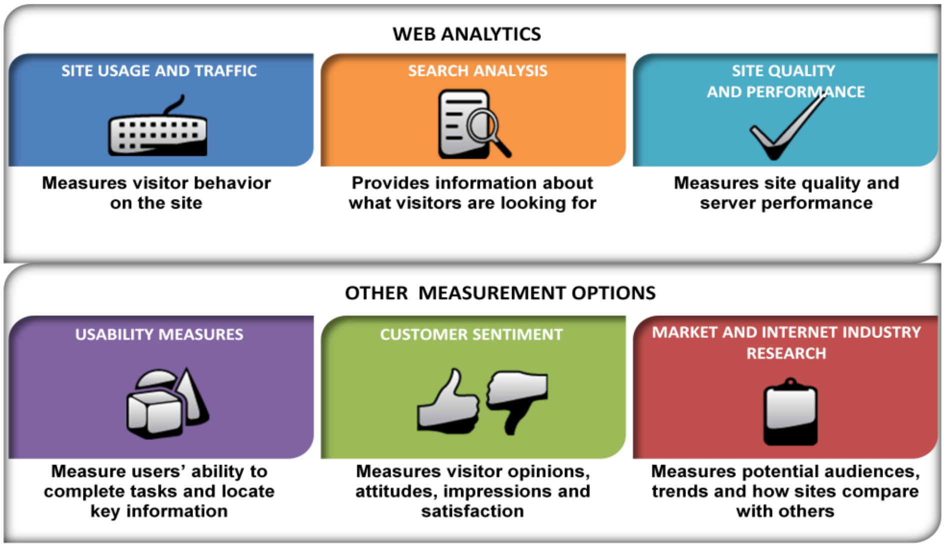
  Complying with Section 508 is a high priority because it's a legal requirement. In 1998, the U.S. Congress amended the 1973 Rehabilitation Act to require Federal agencies to make their electronic and information technology accessible to people with disabilities. Section 508 was enacted to eliminate barriers in information technology, make available new opportunities for people with disabilities, and encourage the development of technologies that will help achieve those goals. While this doesn't directly relate to Web analytics metrics, it's worth noting that many staff members who use analytics are also responsible for tracking their site's 508 compliance. In the scheme of things, 508 compliance is more important than Web analytics.

### *Recommendations*

NIH is not alone in viewing the core function of Web analytics as creating traffic-based reports that provide high-level views of visitor interactions with the Web. This perception undervalues the power of Web analytics.

We believe that NIH and the ICs need to take a comprehensive and cohesive approach to analytics. This approach must incorporate not only Web analytics, but all aspects of online measurement and testing. While the focus of this study is on Web analytics best practices, we must put Web analytics in context with these other methods and tools so that NIH ICs can leverage what's most relevant for evaluating online and digital initiatives. **Figure 11** summarizes the context for this section.

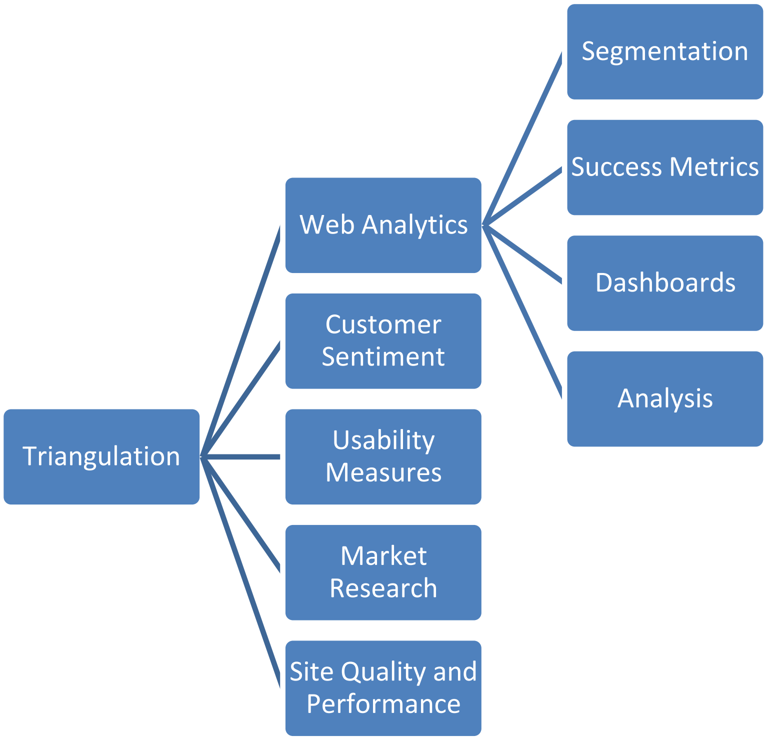
***Figure 11.*** *Tools used for Web site measurement at NIH.*


WEB ANALYTICS
SITE USAGE AND TRAFFIC
Measures visitor behavior on the site
SEARCH ANALYSIS
Provides information about what visitors are looking for
SITE QUALITY AND PERFORMANCE
Measures site quality and server performance
OTHER MEASUREMENT OPTIONS
USABILITY MEASURES
Measure users' ability to complete tasks and locate key information
CUSTOMER SENTIMENT
Measures visitor opinions, attitudes, impressions and satisfaction
MARKET AND INTERNET INDUSTRY RESEARCH
Measures potential audiences, trends and how sites compare with others

In our research, we found that NIH ICs struggle with both when and how to use these measurement methods. There's a lack of clear direction about where in the site and in the application-development lifecycle to strategically apply these methods and a lack of clarity about how the findings of each of these methods can inform each other.

We found that the methodology commonly used at NIH to describe the combined use of measurement is known as "triangulation," or using multiple data sets to both validate and complement results and conclusions. We consider triangulation to be an umbrella for all of the analytics methods, as illustrated in **Figure 12**.

***Figure 12.*** *Triangulation is a method for integrating several Web analytics data sets.*



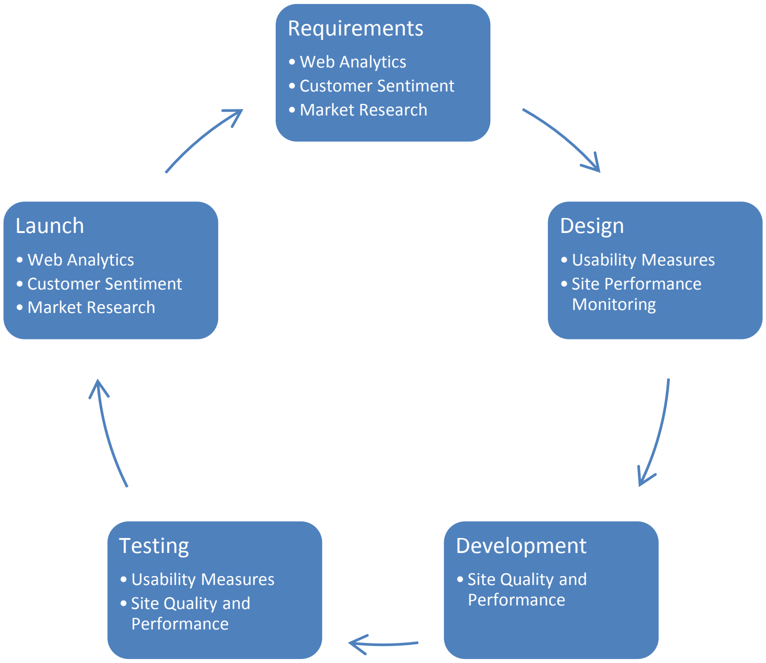
### Triangulation and the Web Site Development Process

Before we discuss how the different types of measurement work together, we need to describe how Web site development occurs. Basically, there are two types of Web development projects:

- content, with text pages with links, and
- applications, with interactive elements, such as forms, shopping carts, search engines, and data-selection wizards

The only real difference between these two is in size: some projects are small, such as creating a micro site, and some are large, such doing a major site redesign. However, the steps are the same for all projects. In **Figure 13**, we illustrate five basic steps in Web development projects and how to incorporate site measurement into them.

***Figure 13.*** *Web site development process and relevant site-measurement options.*



**Step 1: Requirements.** Requirements state what the Web site project will be and what elements it will include. Developing requirements may include preparing use-case scenarios to suggest how visitors will interact with the site.

How to triangulate site-measurement methods for requirements:

1. Use Web analytics for historical data to understand how visitor segments currently use content and applications to get ideas about the audience's interests.
2. Use focus groups, surveys, email, and call-center contacts to stimulate ideas for brainstorming about development.
3. Use market research to understand how your target audiences are using similar Web sites.

We recommend that by the time the design phase is complete, the development team has prepared research plans that consider all the elements to be used for measurement and testing in conjunction with the new initiative. The team should emphasize mapping how the qualitative and quantitative data would be used in tandem.

For example, if customer-satisfaction surveys are to be used to determine whether visitors are satisfied with the workflow of a new wizard application, a corresponding Web analytics funnel analysis should be conducted to validate, correlate, and expand upon the surveys' findings. A funnel analysis focuses on the effectiveness of the specific path you've laid out for visitors to take to complete a task. The analysis will show where on that path visitors exited.

Knowing what your visitors think, say, and care about is the single most important data asset you can have. Many of Semphonic's clients are becoming much more aggressive about collecting, consolidating, analyzing, and using customer-sentiment data. This is because everything customers tell us as they navigate through a site can help achieve our online business goals. From online surveys, comment cards, and call centers to focus groups and social media, organizations have access to enormous amounts of customer feedback and opinions. It's challenging to capture and use the data, and it requires new systems and analysis techniques that are different from traditional Web analytics.

We expect further challenges to materialize with the advent of social media tracking. However, the same practice of goal setting and clarifying how the data correlate with other data streams should help decrease the confusion about what source to believe and how much credence to give one source versus another.

**Step 2: Design.** Once the requirements are defined, the information architecture and page design can begin. Depicting initial designs graphically may help refine and focus the requirements.

How to triangulate site-measurement methods for design:

1. Use Web analytics for historical data to understand any site-navigation successes and challenges experienced by different segments of visitors.
2. Use Web site monitoring reports to avoid performance issues that can arise with large graphics files.

**Step 3: Development.** Once the design has been finalized, the development work can begin to include creating functional specifications and coding.

How to triangulate site-measurement methods for development:

1. Use Web site monitoring reports to avoid performance issues that can arise with densely coded pages.

**Step 4: Testing.** Conduct all quality-assurance tasks related to development.

**Step 5: Launch.** This is when the content or application is made available on the Web site and marketing programs are implemented. Once the site is launched, you can collect data on its effectiveness. This is also when you begin to evaluate how to improve the site or application and how to enhance strategies for driving more traffic to the site.

How to triangulate site-measurement methods for the launch:

1. Use Web analytics to compare new content or applications to your business goals or navigation ease and campaign effectiveness.
2. Use customer-satisfaction surveys to assess ease of use and value.
3. Use post-use surveys to assess the potential behavioral impacts of new content or applications.

4. Use market research to evaluate whether new content has changed the positioning of your site or applications relative to similar or competing sites or applications.

In the following set of recommendations, we first discuss how triangulation can be used throughout Web site development and measurement. Then, we make specific recommendations for how to approach Web analytics segmentation, metrics, dashboards, and analysis.

**Triangulation Recommendations**

- **Recommendation 4.a. Use triangulation to manage multiple data sources.**

  “Triangulation” means using multiple analytics data sets to both validate and inform one another so that you can make decisions about Web strategy and tactics. NIH ICs use Web analytics, surveys, user-experience testing, and focus groups to guide Web site redesign, as well as to provide input for developing new applications. Some ICs consider “use cases” as a way to plan and develop the way visitors will complete tasks on a site. Use cases are the common workflows that describe what visitors are supposed to do on the site. Combining planning and measuring processes at NIH would help guide ICs as they develop Web sites and analyze their effectiveness.

  NIH ICs should gain a great deal from correlating Web analytics, customer satisfaction, and post–Web use behavior through triangulation. Many of the issues organizations face with triangulation of data sources today stem from the fact that such efforts are often ad hoc. Through developing a research plan that’s tied to measuring specific activity, functionality, and effectiveness and that’s connected to specific goals and objectives, ICs should be able to focus the scope of their analysis and make it very useful.

  We also want to put the concept of triangulation into the context of the actual process of Web content and application development so that it’s clear when site measurement occurs within the development lifecycle and how the methods can inform each part of the lifecycle.

  Let’s start off with basic definitions of quantitative and qualitative tools and methods:
  - Quantitative Tools and Methods
    - Record performance data generated by the operation of a Web site.
    - Provide information about *WHAT* actions, events, and trends are happening.
  - Qualitative Tools and Methods
    - Record use data generated by users of a Web site.
    - Provide information about *WHY* events and trends are happening.

  Here’s a list of the quantitative and qualitative methods and tools used for analyses.
  - Quantitative Analysis Tools and Methods

    **1. Web Analytics:** Web analytics software is necessary for collecting information on
    - visitor activity
    - content and function usage
    - site promotion and marketing

- task completion
- site usability
- external search
- Internal search.

**2. Site Quality and Performance:** Web site monitoring is necessary for keeping track of Web server performance, such as server availability and response time and the site's compliance with Section 508 guidelines.

**3. Market Research:** Internet industry or market research and business intelligence companies recruit users and use Internet service provider (ISP) data to define their Internet and Web users' demographic and behavioral characteristics. This provides a clearer understanding of who is using your site, as well as your competitors' sites. Market research organizations include Experian Hitwise, ComScore, Alexa.com, and Google.

- Qualitative Analysis Methods and Tools

**1. Usability Measures — User-Experience and Usability Testing:** This kind of testing is used to evaluate Web site design for usability (for example, ease of finding information, reading a site, using site navigation, following workflow processes, filling out forms, and using search engines, multimedia, and graphics).

User testing is necessary for understanding

- user reactions to proposed site designs and navigations
- why users leave the site
- why users like or dislike the site
- why it's easy or difficult for users to find what they are looking for
- whether the site is intuitive enough for visitors to use without assistance
- whether site instructions are easy to understand or lead to frustration.

**2. Customer Sentiment—Email and Call-Center Contact:** Direct contact and feedback from site users can raise red flags about issues with site content, function, design, and navigation that need to be addressed. For every person who makes a comment, there are many others who have the same issues but do not comment.

**3. Customer-Sentiment Online Information Gathering—Surveys:** Customer-satisfaction surveys that deploy during a visit may be for understanding visitors' points of view as they relate directly to navigation or tasks they are trying to perform. Direct online surveys sent to visitors after use of the Web site provide insight into whether using the Web site elicited a change in behavior based on information received or on a specific task completed.

**4. Customer Sentiment—Focus Groups and Interviews:** Focus groups and interviews can provide information that you can use for improving and planning outreach campaigns, site content, and site functions.

Web analytics and the other measurement methods should all be used together through triangulation. We believe that the concept of triangulation makes more sense when it’s viewed within the context of the Web development lifecycle. Understanding the role and value of each type of measurement within the lifecycle helps determine how the information from each measurement method informs the lifecycle activity.

### NIH Web Analytics Method Recommendations

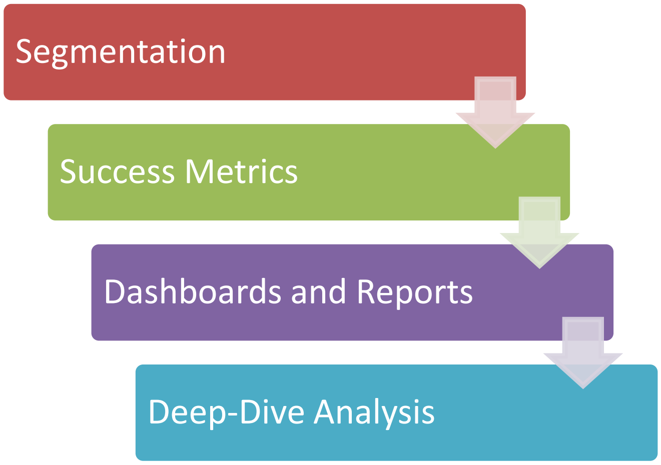
- **Recommendation 4.b.  Implement the NIH Web Analytics Method.**

  The focus of Web site development and analytics at NIH has largely been on public health–oriented traffic to the Web sites, even though the mission of NIH is geared more to the research and science communities. We believe that there needs to be more emphasis on understanding the behavior of visitors seeking health information, grants, and funding and scientific information as well as how the Web sites are optimized to serve these communities. Our recommendations in this section reflect this thinking.

  We also believe that there needs to be a focus on site performance in the context of return on investment. We’ll expand on this in the section on success metrics, see beginning on page 39.

  Finally, although we divided the metrics recommendations into specific sections, we believe that for NIH to be truly successful, all the recommendations must be implemented. Only in this way can NIH build a complete analytics framework. For example, if you don’t use segmentation in conjunction with success metrics, you won’t gain any real insight at the audience-centric level. Similarly, if you create dashboards without thinking through baselines and targets for the metrics, you’ll never be at a point where you can determine whether your efforts are succeeding. This is why we recommend a specific method for Web metrics development for NIH, which we’re calling the NIH Analytics Method (**Figure 14**). The rest of this section describes in detail each aspect of that method and how ICs can take advantage of it.

***Figure 14.*** *NIH Web Analytics Method.*



### Segmentation Recommendations

- **Recommendation 4.c. Increase the value of analysis by using segmentation.**

  While many NIH ICs are beginning to accurately track unique visitors, it's important to acknowledge that this metric is not very useful by itself. It tells you nothing about a visitor's characteristics. It's simply a tally. When you can define groups of unique visitors, you start learning how your site is being used and how well your outreach programs are working.

  Segmentation is a method for categorizing groups of visitors by interest, site activity, demographics, or other actual or arbitrary classifications. It enables you to

  - identify visitor trends for groups, rather than disparate individuals or a market of "everybody"
  - target markets to focus your site content and site marketing on
  - focus your spending on content and promotions that are aligned with the "right" target market for your site
  - develop content that matches visitors' level of experience with your organization
  - drive visitors to the content that can be most useful for them

  Segments can be developed based on visitor information, visitor behavior, or visit activity.

  **1. Visitor Information:** This information can come from data collected from visitors through registration or forms submitted on the site and can include the visitor's

  - industry or specialty
  - job title
  - state
  - country
  - city
  - organization (for example, jhu.edu, cdc.gov, gsk.com)

  **2. Visitor Behavior,** including

  - new visitors
  - repeat visitors
  - frequent visitors

  **3. Visitor Activity**, including

  - document views or file downloads based on the number of documents read or files downloaded
  - content section activity (for example, research funding, research findings, health information, clinical trials, careers)
  - external site referrals (for example, nongovernmental organization (NGO) partners, NIH ICs, government agencies, educational and health institutions)
  - search engine referrals (for example, Google, Yahoo)
  - external-search-term referrals: branded (NIH or IC named) and nonbranded (for example, disease, medical condition, news event)
  - internal-search terms used (terms used for basic and advanced searches)

- ForeSee survey responses, including respondents' satisfaction score and specific answers
- newsletter, RSS, and other ongoing communication registration (visitors sign up and receive scheduled communications from the IC)

These categories are not exclusive. For example, you can create segments that include visitor behavior and visit activity, such as repeat visitors who read at least two articles per visit. You then give the segments names that make them easily identifiable.

Once these segments have been set up in your Web analytics tool, you can run your metrics and reports against the segment. You can focus your research on how that segment performs. The results can tell you what type of content is most appealing to each segment.

At NIH, a few ICs are beginning to approach the use of segments through filtering so that they can create customized reports (for example, viewing the traffic from the .edu domain). In **Figure 15**, we list several segments that NIH ICs might want to use. Much of what shapes segment creation goes back to the basics of aligning Web analytics metrics to site goals and target audiences.

***Figure 15.*** *Possible audience segments and descriptions for NIH.*

| Segment Name | Visitor Information | Visitor Behavior | Visitor Activity |
|---|---|---|---|
| Grant-Funds Seeker | Visitors from .edu domains | New and Repeat Visitors over one month | Views three or more pages of Research and Funding content per visit |
| Grant-Funds Applicant | Visitors from .edu domains | New and Repeat Visitors | Clicks on link to grants.gov |
| Informed Health Care Professional | Specific hospital domain name, such as http://www.mountsinai.org/ | Repeat Visitors within a month | Views three or more pages of Health Care Professional Content per visit; visits three or more times per month |
| Engaged Consumer | All Visitors | Repeat Visitors within a month | Receives IC newsletter and clicks through to at least one article per newsletter |

Once the segments are created, they should be tested for three months to ensure that they're useful. After that, you should be able to gain some insights into how you could further engage with the segments you consider valuable. For example, with analytics information about specific visitor segments, an IC can create more-focused content. This will make it easier for those audiences to complete a specific task and for the IC to figure out a new outreach or interactive approach to increasing interest in the IC's programs and research.

The segments described in Figure 15 can be set up with most of the commonly used Web analytics tools, although each tool's approach is different. For example, in Google Analytics, it can be done through Advanced Segmentation; in Webtrends, it's through Custom Reports. You can then use segment reporting to find out how groups of visitors navigate the Web site, what content they consume, and how many times they visit before they submit an application or make contact. This information can help improve site navigation, content, and outreach to help you attract the most qualified research candidates.

**Success Metrics Recommendations**

- **Recommendation 4.d. Institute success metrics, a goal-based Web analytics approach to evaluating NIH Web sites.**

  NIH ICs could get a great deal more value from Web analytics if the metrics they used for dashboards and reports were drawn from the NIH mission, cost-related objectives, and content-measurement goals. With this in mind, we'd like to create metrics that can be used for the following purposes:

  - NIH-wide benchmarking: For comparing ICs over time.
  - IC-specific benchmarking: For comparing divisions within ICs over time.
  - Executive dashboards: For presenting to NIH senior-level management and IC directors
  - Content, outreach, and Web site effectiveness dashboards: For use by content managers and communication teams.

  The process of developing relevant metrics is based on a McKinsey and Company method for helping organizations define and measure their effectiveness against business objectives. In the 1980s, Jack F. Rockart of the Sloan School of Management clarified and widely promoted the approach. The result was a hierarchical performance measurement system of evaluation:

  - Defining organizational mission
    - *Identifying strategic goals that map to the mission*
      - *Determining concrete and measurable objectives*
        - *Defining critical success factors (CSFs) that are used to evaluate the objectives*
          - *Developing a key performance indicator (KPI) measurement and target that determine whether the CSFs are being met. (A KPI is a number derived from a count, percentage, or ratio. A target is a goal based on business objectives or historical trends.)*

For an NIH IC, this model could look something like this:

> Organizational mission – Inform the public about research activities and the impacts of this research.
> Strategic goal – Provide the most reliable and authoritative information
> Objective – Publish a weekly email newsletter
> CSF – Increase number of subscribers to the weekly newsletter
> KPI 1 – number of newsletter subscribers per month
> Target 1 – increase the number of newsletter subscribers by 10% each month
> KPI 2 – number of shares of newsletter by subscribers to nonsubscribers
> Target 2 – at least 100 shares per month

You can have several CSFs per objective, and a few KPIs and targets associated with each CSF.

As you can see, KPIs are only one part of a bigger picture and you can do a lot more with your current Web analytics numbers. It doesn't have to take a lot of work or sophisticated tools, just a different way of looking at the numbers.

The key difference between the traffic metrics commonly used today and the metrics described here is in the intention to use the new metrics to measure success against specific goals, objectives, and benchmarks. We call these success metrics, and they can be illustrated like this:

**Goal ⇨ Objective ⇨ Success Metrics**

Visitors accomplish objectives that are based on Web site goals. The objectives are the events measured by the success metrics. In the research we conducted for this project, we defined two types of objectives:

- **Task Completion** – Many NIH sites present tasks for visitors to complete, such as these:
  - email submission
  - registration
  - submitting grant applications
  - using site tools and calculators
  - submitting forms online

Completing these tasks helps validate the site's value and accomplish its mission. The tasks are relatively easy to identify on the Web site, and there is general consensus about their relative importance in meeting the NIH mission.

- **Content Consumption** – In the stakeholder survey we conducted for this best practices initiative, most respondents indicated that consumption of content was the primary goal of their Web sites. However, in no case during the discovery phase of this project was the consumption of content ever linked to increasing the Web budget, adding resources for the Web site, lowering Web site or IC operational costs, or producing a verifiable assurance that the IC mission goals were being met by the Web sites.

In fact, just the opposite is true. Despite the fact that content consumption of almost every NIH Web site has increased, Web site budgets are generally flat year over year, Web site staffing is lean and often not full-time, and there's little recognition among senior management of Web sites' quantifiable business value. In addition, much of the increase in content publishing and consumption on NIH sites is due to greater attention from the public, not from the NIH mission's main focus: professional scientists and researchers.

One difficulty in measuring content consumption accurately at NIH is that there's a great deal of content and, in most cases, no clear strategy for developing or promoting it. Content is often produced in response to individual initiatives rather than coordinated efforts.

- **Recommendation 4.e. Use the success metrics to better inform NIH and IC senior leadership and Web managers about the performance of Web sites.**

  Using success metrics will help NIH move away from meaningless traffic metrics and toward business-oriented measures that provide an opportunity to assess how effective a site is in achieving the NIH or IC mission and in maximizing return on investment. These metrics can also be used to establish comparative benchmarking. These approaches will give senior leadership more compelling cases for funding and for expanding Web content and outreach programs.

- **Recommendation 4.f. Define metrics for benchmarking within and across ICs' public-facing sites.**

  On the basis of our discussions with NIH staff, we recommend that ICs use this set of metrics for benchmarking. (Calculations and examples are in **Appendix D.**)

  - Task-Completion Effectiveness – Indicates visitor interest in and successful completion of tasks important to NIH or the IC.
    Definition: Measures completion of site tasks associated with the NIH or IC mission.
  - Mission-Critical-Content Effectiveness – Evaluates visitor interest in content that the IC or NIH considers important.
    Definition: Measures the number of times mission-critical content and/or video is viewed or files are downloaded.
  - Content-Relevance Barometer – Measures the number of site pages not viewed relative to the entire site's pages. Pages that are not viewed can then be revised, retired, or relocated.
    Definition: Measures the overall site content relevance based on viewing all site content pages.
  - Content-Distribution Score – Indicates whether visitors share content, which helps spread the IC brand and authority. We assume it is an indicator of the content's positive value and interest to the visitor.
    Definition: Measures whether visitors find content valuable and relevant enough to share.

- Navigation-to-Content Score – Indicates how much time visitors are spending trying to find content.
  Definition: Compares number of page views on navigation-oriented pages with content- or application-focused pages.

- Mission-Success Score – Evaluates engagement of site visitors based on the reading of mission-critical content and task completion.
  Definition: Measures per-visit engagement with site's mission-critical content and tasks.

- Return on Investment – Provides quantifiable cost-benefit analysis of budget and time expended on Web site operations.
  Definition: A way to express the value of Web site operations, content, applications, and labor relative to the budget for these initiatives.

### Dashboard Recommendations

This section describes how to apply specific metrics to the different audiences.

- **Recommendation 4.g. Encourage a standardized approach for presenting success metrics through the use of dashboards.**

  Many at NIH feel stymied by how to interpret Web analytics data, which keeps them from optimizing Web content, navigation, and marketing initiatives. Having a standardized approach would really help. We propose a set of recommendations and guidelines to data interpretation, based on common industry practice, in **Appendix E.** Dashboard Content and Samples.

  We recommend four primary dashboards for NIH ICs as well as for NIH.gov. These include
  - Executive Dashboard: For getting a global view of NIH and IC site activity; primarily for senior management.
  - Content Dashboard: For understanding visitor activity relative to the content of site sections and, then, for managing content and planning new content and application-development initiatives.
  - Web Site Optimization Dashboard: For understanding how site navigation and internal-search use perform for segmented and nonsegmented visitors, as well as across the whole site or within specific site sections.
  - Outreach Dashboard: For understanding the effectiveness of visitor-acquisition campaigns and search engine and referral strategy and tactics.

  These dashboards, including the metrics they use, are organized by dashboard type in **Appendix E**.

  The dashboards and reports we recommend in previous sections are only one part of the NIH Web Analytics Method. These reports will help you quickly gauge the effectiveness of your Web initiatives. However, their purpose is not to explain *why* traffic is increasing or decreasing or where on the site a bottleneck is causing visitors to exit or drop out of an ordering or registration process. To obtain that information, you need to do deeper analysis. The following section describes a variety deeper analytics approaches.

**Deep-Dive Analysis Recommendations**

- **Recommendation 4.h. Use deep-dive analysis to complement dashboards and reports.**

  Deep dive analysis complements regularly scheduled dashboards. Whereas dashboards provide an "at a glance" view of "what" is happening on the site, analyses such as the ones recommended, provide the "why" that explains trends, spikes and dips. This is the type of insight that is most critical in planning large scale changes to the site, development of new applications and determination of how to best attract the target audience segments to NIH web sites.

  Deep-dive analysis generally entails using a finite set of data, such as from a month or a quarter, developing different hypotheses about cause, and drilling down into specific site elements. To do this, you need a sophisticated analysis tool such as Adobe Omniture SiteCatalyst and DataWarehouse or Discover, Webtrends Insight, Unica NetInsight, or Comscore Digital Analytix. You can do a limited amount of deep-dive analysis within Google Analytics.

  We recommend the following six deep-dive analyses for use at NIH.

  - Visitor-Segment Analysis. Conduct behavioral profiling of known user types, such as visitors who register for ongoing communications or who submit grant proposals. This analysis will determine the following attributes of these user types: total number of visitors, content consumption, tasks performed, and how they come to the site. From this initial data set, you can determine whether there are patterns or anomalies that suggest strengths or challenges in engaging these audiences and then be in a position to develop recommendations for site or marketing optimization.

  - Campaign Analysis. Determine the effectiveness of campaigns and outreach channels in attracting visitors to the site. This approach reviews search, email, social media, and all other methods for attracting visitors through an analysis of both visit and visitor segments. Results from the analysis inform strategy and tactics for developing new campaigns and suggest potential A/B testing scenarios for landing pages and campaign vehicles.

  - Internal-Search Effectiveness. Analyze the internal-search tool to determine how it can be modified to improve the usability your site. The following activities are generally included as part of this analysis.

    - Researching how internal search fits into the general site structure, whether it is a primary or fallback routing mechanism, and what its primary use cases are.
    - Identifying common page types that users search from and other content viewed during search sessions.
    - Analyzing page exit rate by keywords and keyword groupings to determine the primary weaknesses of internal-search results.
    - Breaking down search use by registered user type and determining whether different users navigate the site differently.

  - Page-Value Analysis. Create a scoring or value scale for mission-critical content pages (by source, site area, content, ad type, etc.) to quantify and evaluate which pages visitors should be routed to. This could be coupled with creating a campaign report or analysis where you use the page value to determine the ROI of each campaign.

  - Content "Cross Sell." Identify the types of content that tend to drive people to read additional articles and the types that are typically "dead ends." Also, determine how well the site currently routes people from one article to another and recommend ways to increase engagement.

  - Functional Analysis. Functionalism is an analysis technique based on the proper function of each page on a site. This approach breaks a Web site into its constituent pieces and then assigns one or more specific functions to each piece. These functions can be things like navigation (e.g., route visitors to a specific place), motivation (e.g., convince a user to do something), or information (e.g., provide a visitor with some piece of information). On the basis of the functions of the page, it’s assigned a particular page type from a set of common templates. Once a page type is assigned, the success of a page is measured by critical success factors specific to the functions it was designed for. The power of functionalisms resides in this assignment of type-specific, easily measurable success factors. For more information about functionalism, see **Appendix F.**

**Other Recommendations**

- **Recommendation 4.i. Increase the use of A/B testing for content and Web site optimization.**

  A/B testing is a relatively simple process that enables evaluation of content and design to maximize site visitor satisfaction and use of the web site. We have provided a number of best practice recommendations for conducting and using A/B testing in **Appendix G.**

- **Recommendation 4.j. Begin to measure the investment in social media strategies.**

  NIH ICs are undertaking a number of social media initiatives, such as the launching of Facebook pages, and using YouTube and Twitter. Although much of this is viewed as an experiment, there will be an increasing impetus to determine the relative return on investment of the time to maintain these new digital channels. We have provided best practice recommendations for social media measurement in **Appendix H.**

## 5. Web Analytics Solutions, Methods, and Opportunities

Web, social media, competitive intelligence, and online surveying are all driven by software solutions, as are multivariate testing, pay-per-click (PPC) search marketing, and cross-selling platforms. Although we have purposely focused on nontool aspects of Web analytics at NIH, access to good tools is, of course, critical to conducting good analytics. We describe our findings and provide recommendations in this section. **Appendix I.** provides evaluations for a dozen tools that NIH ICs have expressed interest in using for both current and future analytics initiatives.

### *Findings*

- **Finding 5.1. Centralized technical support cannot support IC Web analytics requirements.– Needs Improvement**

  We found that the funding mechanisms and technical support constraints around Web analytics at NIH were so severe that they've had a crippling effect on Web teams' ability to use Web analytics to any reliable or useful degree. This has also contributed to the relatively immature use of Web analytics at NIH.

  We believe that the status quo of using the fee-based Webtrends log file solution is not tenable because of the current gaps of support, poor access, and lack of flexibility to do anything other than get "canned" automated reports of poor-quality data.

- **Finding 5.2. Lack of centralized funding mechanisms creates a "have" and "have-not" disparity between ICs that can afford their own tools and those that cannot.– Needs Improvement**

  Individual solutions have tended to be "owned" by the stakeholder group that's most often using them. Because of this, NIH may be missing opportunities for greater efficiencies of scale and overall return on investment (ROI) through centralized purchase and distribution of tools.

### *Recommendations*

- **Recommendation 5.a. Implement a reliable Web analytics solution.**

  Although metrics recommendations are tool neutral, a reliable Web analytics solution is still required.

  Because there's a range of Web analytics tools available to NIH, it's important that our recommendations be, in large part, tool agnostic. That way, the recommendations can be used with a range of Web analytics packages, such as Adobe SiteCatalyst, Webtrends, and Google Analytics, as well as "home-grown" tools such as Web-server log parsing programs. (It could be more time consuming or challenging to implement the recommendations using home-grown tools rather than other kinds, however.)

  With these metrics recommendations in mind, it should be understood that the relative ease of implementing the metrics described below will likely vary based on the tools used for
  - data-collection efficiency
  - data-parsing and -segmentation flexibility
  - report-presentation options

- **Recommendation 5.b. Consider a two-tier Web analytics option.**

  One tier—the basic tier—would be a free product such as Google Analytics. Google Analytics, while not without faults, provides NIH with an option that is easier to implement and maintain and lends itself well to standardization and distribution throughout all ICs for the production of top-level metrics (as described in the success metrics section, beginning on page 39. ICs that do not want to use a software-as-a-service option, where the analytics tool is hosted at the analytics vendor's site; we recommend the open-source tool Piwik. Piwik looks a lot like Google Analytics and collects data in the same way. The difference is that this can be hosted on NIH servers. Please refer to **Appendix K.**

The second tier—the advanced tier—would be an enterprise-level tool, such as Adobe Omniture, Webtrends, IBM Coremetrics, or Comscore Digital Analytix. These tools provide far more capable and scalable segmentation features, as well as the ability to do deeper analyses. In addition, these tools are set up to integrate with other data streams, such as ForeSee survey data. Both tools are described in greater detail in **Appendix I.**

We believe that enterprise analytics and measurement tools should be bought as part of a greater NIH site license in order to obtain favorable pricing for ICs that want the higher-power analytics.

- **Recommendation 5.c. Encourage the adoption of Google Analytics.**

  Over the past year, Google Analytics has emerged as a real alternative to the Webtrends log-file tools that have become fairly common at NIH. While we don't see Google Analytics as a panacea, its benefits as a free and accessible tool outweigh potential drawbacks, and we recommend that ICs consider Google Analytics implementation if they need to have a Web analytics solution in place quickly. For a discussion of our recommended approach and issues to consider, please refer to **Appendix J.**

- **Recommendation 5.d. Use first-party cookies because, while not required for Web analytics, they add tremendous value.**

  The benefits of the metrics recommendations will be most fully realized if the NIH Web servers hosting the IC sites serve first-party cookies. These will help ICs define unique visitors with some level of accuracy.

## 6. Web Analytics Training, Data Analysis, and Interpretation of Data

Web and digital analytics data are challenging to use. How do you interpret whether a certain number of page views is good or bad? How do you use the information to figure out what to do? This is why developing clear language around Web analytics is essential for the successful adoption and use of analytics throughout any organization. Training staff how to produce analytics reports and providing guidance on how to interpret the data so that they're understandable and useful are the chief areas of focus for digital analytics training and education.

### *Findings*

- **Finding 6.1. There is no standard documentation at NIH that allows Web analytics processes to be replicated. – Needs Improvement**

  Web analytics at NIH is currently performed on an IC-by-IC basis. We believe that ICs could be learning from each other about ways to improve their processes for implementing analytics practice. During our discovery phase at NIH, we found few examples of documented measurement processes. Nevertheless, as the NIH Web analytics program grows and becomes more comprehensive, it will become important for these processes to be formally documented:

  - **Web analytics requirements**: A description of organizational needs that can only be met through Web analytics.
  - **Cross-functional-team processes**: Collection of protocols for situations that require using the resources of several different teams or staff members.
  - **Functional specifications**: Descriptions of how business requirements map to specific analytics reports and how these reports are generated by the Web analytics tool.
  - **Standard operating procedures (SOPs):** Protocols for Web operations.

- **Finding 6.2. There is little documentation about Web analytics requirements for implementation and data-accuracy standards. – Needs Improvement**

  We discuss this is in more detail in **Appendix C.**

- **Finding 6.3. Web analytics reporting that goes to senior-level management focuses not on business objectives, but on site traffic that is not very relevant.– Evolving**

  In many cases, analytics reports generated from Webtrends and Google Analytics are being used primarily to report high-level traffic statistics. In the case of Webtrends, most of the reports are delivered as automatically produced documents 50+ pages long. To find any usable data, the Web analyst has to go through the report and pick out the data of interest. This is tedious and time consuming. Some analysts have gone a step further, creating more dashboard-oriented reports that compare time periods, incorporate metrics from social media channels, and focus on important content. These reports also take a long time to create with Webtrends. It’s no surprise that the difficulty in using the current tool set discourages people from developing reports.

  We’ve started to see some ICs create easier-to-produce and graphically attractive dashboards from Google Analytics.

  What tends to be missing from many of these reports, however, is an explanation of the results: why is a metric included? What does it mean? Is there an action to take? While glossaries are often included, these only serve to define an “out of the box” term, such as “page view,” “referral,” or “visit.” In most cases, interpretation of the data is left up to the report viewer. Terms, trends, and anomalies are rarely explained, nor are data-collection procedures or the issues affecting data quality and accuracy. Since little has been done about developing analytics requirements related to NIH business goals, interpretations are not generally provided in the context of business questions. Analysis is focused on numbers of page views and visits, and the stakeholder has to figure out what those numbers have to do with the business questions they may have about the Web site. This has contributed to the negative-to-ambivalent attitude about Web analytics and has perhaps exaggerated the perception that the data are inaccurate.

- **Finding 6.4. There is no NIH-sponsored training on how to use Web analytics tools and data. – Needs improvement**

  Very few people at NIH know how to use the tools, to create metrics, to develop reports, or to analyze and use data. Although many stakeholders we spoke to have taken advantage of numerous opportunities to learn about Web analytics from Web Manager University and industry events, they're really interested in learning more, especially around report and dashboard development, data interpretation, and data analysis.

- **Finding 6.5. The NIH Web Metrics Group sponsored by the On-Line Information Branch is a useful asset that could do more. – Evolving**

  Many of the people we interviewed were very positive about the meetings and information provided through the NIH Web Metrics Group, sponsored by the On-Line Information Branch, and would like to see this group do even more.

### *Recommendations*

- **Recommendation 6.a. Initiate monthly meetings of Web analytics team members and supporting Web content, design, and technical teams to discuss current metrics, reporting, and analysis and other issues regarding the analytics initiative.**

  Especially in the early stages of working with Web analytics data and reporting, it's important to have communication methods in place that contribute to organizational penetration and transparency of issues that may be affecting the development and production of reports. A monthly meeting fosters better communication among the Web analytics team and their collaborators by providing a good platform for discussing, assigning, and presenting analysis projects and determining the necessary levels of effort and resources.

- **Recommendation 6.b. Develop a strategy to provide interpretive analysis.**

  Interpreting Web analytics reports and metrics is a challenging task because stakeholders have different levels of understanding of how the Web analytics data relate to business goals. We suggest that NIH ICs develop a strategy that incorporates multiple methods for providing and communicating Web analysis, including these:

  - One-on-one sessions between the members of the Web Analytics Group and report viewers to discuss how to use reports and metrics to answer the business questions that came up during the requirements process.
  - Group sessions taught by the Web analytics manager that explain how to use the recommended report sets.
  - Annotated report sets with interpretive analysis by the Web analyst or power users.
  - Glossary of terms in reports and metrics.
  - Monthly interpretation of each IC's success metrics and recommendations written by the Web analytics manager and distributed throughout the IC (see metrics section, beginning on page 20.
  - Plain-language page titles provided in analytics reports instead of URLs.

We also suggest establishing monthly metrics meetings both at the IC level and within senior-management forums, such as communication director meetings. We hope that the recommended metrics for benchmarking will become a focal point of presentations and discussions and receive greater exposure as Web analytics becomes more relevant to senior leadership. These meetings could provide a forum for information exchange and education about analytics and help spread awareness throughout NIH about how to use the Web analytics

- **Recommendation 6.c. The IC Web analytics manager and the Web analytics lead in the On-Line Information Branch should make a quarterly "State of the Web" presentation to respective management teams.**

  This presentation should focus on key Web site initiatives and success metrics. It should be used to educate management about the importance of the Web initiative and the value of using analytics to measure progress and identify areas for improvement.

- **Recommendation 6.d. Enhance the online Web analytics knowledge center.**

  Many stakeholders said that a hindrance to applying analytics to their work is the lack of access to information about Web analytics—for interpreting data and understanding how analytics works and how to get more business value from Web analytics.

  Setting up an online, self-service Web analytics knowledge center, managed by the On-Line Information Branch or a cooperative of interested Web analysts, would address the lack of access to information. Stakeholders could educate themselves there. This resource would also be a repository for all presentations, documents, and articles about Web analytics, as well as NIH Web analytics case histories. In addition, a wiki could be established to encourage Web analysts to share their issues and success stories about analytics, and members of the Web Metrics Group could start an internal blog, for the NIH community.

- **Recommendation 6.e. Develop an analytics cooperative.**

  Create a user group or clearinghouse managed through the On-Line Branch whereby NIH staff interested in learning more about using analytics tools can be matched up with more experienced staff to learn how to use tools, create metrics, and interpret data. The newly trained staff could work on projects within their own ICs or be available to take on projects from other ICs that do not have resources to perform the relevant tasks.

- **Recommendation 6.f. Develop an analytics training and education curriculum.**

  We believe that NIH would benefit from a serious and aggressive Web analytics training program. The program could be aimed at Web analytics managers, tool administrators, potential power users (hands-on), and report viewers. We suggest the following two types of training: customized training for different user groups and standardized training on Web analytics methodologies.

- **Training by User Group Level**

This training needs to be customized to the different audiences that are using Web analytics. Some will only be using data; others will be using the tools to develop reports and analyses for others; and still others will be charged with maintaining and updating the data-collection and data-processing tasks. (See **Figure 16**.)

***Figure 16.*** *Types of user group Web analytics training.*

| Report Viewer | Power User | Implementation | Administrator | Deep-Dive Analysis |
|---|---|---|---|---|
| How to<br>▪ I interpret reports | How to<br>▪ use the tools<br>▪ analyze reports<br>▪ develop requirements<br>▪ build reports | How to<br>▪ customize codes<br>▪ assign data-collection variables<br>▪ do QA testing | How to<br>▪ set permissions<br>▪ manage tool settings<br>▪ troubleshoot data and report issues | How to<br>▪ choose a research method<br>▪ form queries |

- Report-Viewer Training: Focuses on how to interpret data in the IC dashboards described in the metrics section, beginning on page 20. This training can be customized to different stakeholder groups based on their specific dashboard metrics. IC Web analytics managers may conduct the training.

- Power-User Training: Focuses on how to use the tool for developing dashboards within the analytics tool. This training will impart skills required for creating reports, dashboards, calculated metrics, using calendar functions, sharing and distributing reports, using Excel reporting options and advanced tool functionality. This class may be taught by IC Web analytics managers who have gone through training. Basic training can be provided by software-tool vendors, partners, or consultants with experience.

- Implementation Training: Provided to developers and Web analytics managers so they understand how to configure and maintain the tools for data collection as well as how to conduct quality-assurance testing. Basic training can be provided by software-tool vendors, partners, or consultants with experience.

- Administrator Training: Focuses on setting up user groups and access permissions; creating report-distribution groups, report suites; and data profiles; and using data-processing configurations. Basic training can be provided by software-tool vendors, partners, or consultants with experience.

- Deep-Dive-Analysis Training: An add-on to Power-User Training. This is a one-to-one or one-to-a-few approach where students conduct an analysis project with guidance from the trainer. This class must be taught by an experienced analyst.

- **Training on Web Analytics Methodologies**

Web analytics training is not solely linked to the use of tools. Understanding the methodology and concepts behind the tools and reporting provides the necessary context and background for the practice of analytics. This training track features classes focused on theory that supports the use of tools and reports (**Figure 17**).

***Figure 17.*** *Types of training about Web analytics methodologies.*

| **Introduction to Web Analytics** | **How to Develop and Interpret Success Metrics** | **How to Conduct A/B Testing** | **How to Measure Social Media** | **Functional Analysis** | **How to Measure Effectiveness of Outreach Campaigns** |
|---|---|---|---|---|---|
| ▪ Survey course on terminology, data collection, tool sets, and business value | How to<br>▪ create<br>▪ set up dashboards<br>▪ analyze<br>▪ take action | ▪ Methods<br>▪ Tools<br>▪ Analysis | ▪ What to measure<br>▪ How to measure<br>▪ Tools to use | ▪ Methodology<br>▪ How to conduct analysis<br>▪ How to interpret | ▪ Search engine optimization<br>▪ Email marketing<br>▪ Offline campaigns |

- Introduction to Web Analytics: Survey course for the general-user population.

- How to Develop and Interpret Success Metrics: Based on success metrics methodology (see Appendix D), for Web analytics managers and analysts. The course describes the process for developing and using the success metrics process.

- How to Conduct A/B Testing: Based on the content in Appendix G. , the course describes business reasons for using A/B testing, demonstrates how to use testing in Google Optimizer and other available tool sets, and addresses available tools to use. Provides guidance on how to interpret and use the results for Web site and campaign optimization.

- Social Media Measurement: Based on the content in Appendix H., the course describes available tools and best practices for conducting social media measurement on the NIH Web site, as well as on Facebook, Twitter, YouTube, and the blogosphere.

- Functional Analysis: Based on the content in Appendix F., this course instructs students on how to classify their site by function and use this information to perform Functional Analysis as well as how to segment the site to enable efficient production of success metrics.

- Measuring Effectiveness of Outreach Campaigns: This course reviews how to plan Web analytics data collection for campaigns and to use the data for making decisions about campaign strategy and tactics.

# Developing Web Analytics Best Practices for NIH

## Appendices

# Appendix A.
# Survey Analysis: Developing Web Analytics Best Practices for NIH

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com

December 15 , 2011

# Appendix A. Survey Analysis: Developing Web Analytics Best Practices for NIH

## Introduction

Semphonic has been working with the NIH to explore the current state of Web measurement and to develop a strategy and framework for the use of Web analytics.

As part of this project, Semphonic worked with NIH staff to conduct a survey that would help establish a baseline understanding of the relationships among NIH stakeholders, staff, and Web analytics. The 24-question survey received 131 responses from individuals who identified themselves as having jobs ranging from Web team leaders to editors and more. This document analyzes the responses with a particular emphasis on questions such as (but not limited to):

- Are goals and audiences in alignment or in conflict?
- Are the top tasks to be completed on NIH sites in alignment with these goals and audiences?
- What are risks and rewards of having analytics staff members working part- or full-time?
- What is the impact of 508 compliance on building sites that are effective at meeting user needs?
- To what extent are stakeholders educated about Web analytics tools and data?

This document features an Executive Summary to highlight overall findings, followed by a question-by-question breakdown of key findings. When respondents chose from a predetermined set of responses, we summarize the results in the report graphically. Alternatively, for open-ended questions, respondents answered in their own words, and we include characteristic quotes to give some color to the analysis.

## Executive Summary

The survey results revealed the following key findings:

- Web analytics is most often a part-time (or less) job at NIH.
- There are opportunities for improvement in the levels of organizational knowledge and penetration of Web site measurement at NIH.
- Compliance with the Section 508 Amendment of the Rehabilitation Act of 1973 Disabilities Act is a necessity that commands a lot of attention and takes priority over other types of Web analytics.
- Web measurement stakeholders at NIH have a unified and consistent idea of who their audience is and what their goals are.
- Web content is of primary importance for most NIH sites, but understanding how visitors access and interact with content is a complex issue and often not addressed holistically.

- Institutes and Centers at NIH have a strong desire for a standardized set of best practices as well as a forum for exchanging Web-measurement information and expertise.

We discuss each of these findings below and include supporting evidence taken directly from the survey responses.

## Key Findings

### *Web analytics is most often a part-time (or less) job.*

Responses to survey questions about analytics resources at NIH indicated two things: analytics is not currently a full-time job at most NIH ICs, and many respondents reported that their IC lacked people with Web analytics skills and time to do analysis and reporting. Respondents felt both of these issues were detrimental. Question 21 asked about the challenges of doing analytics, and nearly 70% of respondents selected "Not enough staff to analyze data" or "Not enough time to analyze data."

This is not a problem purely in terms of resource hours to spend on data; the lack of full-time analytics staff also prevents the development of a requisite level of expertise and familiarity with analytics data. Survey respondents repeatedly indicated that lack of expertise, training, and reliability of data is a major issue. Seeing analytics as a responsibility that requires full-time staff who can learn to use tools and who make improving their analytic capabilities a primary component of professional development would go a long way toward addressing these concerns.

**Supporting Evidence:** In Question 13, respondents were asked how much of their time was spent on Web site measurement or evaluation; 79% of respondents selected "less than 25% of time," "rarely," or "never."

- In the same question, only 5% of respondents selected "75% of my time" or "100% of my time was spent on analytics."
- Question 21 asked respondents to indicate what challenges prevent them from using Web analytics more frequently; 68% of respondents selected "Not enough staff to analyze the data," and 69% of people selected "Not enough time to analyze the data."
- Question 23 asked an open-ended question about the challenges people have been facing. Many respondents mentioned the lack of staff time and resources. A typical response was "time to digest and analyze the Omniture reports, [and] search engine logs to guide site improvements."

### *There are opportunities for improvement in the levels of knowledge and widespread use of Web site measurement at NIH.*

Questions 15–19 asked respondents to indicate what tools they use for different types of analytics, including Web analytics, social media, and voice of customer. In all cases, at least 25% of respondents indicated that they didn't know. This indicates that respondents' understanding of different types of analytics could be improved. Furthermore, survey respondents suggested repeatedly that they did not have good data and/or did not trust the data they were seeing from analytics tools like Webtrends. In

Semphonic's experience, mistrust of data comes from a lack of understanding of how data collection works and/or poor presentation of data.

These findings are consistent with the overall survey findings about lack of analytics resources and education throughout NIH. The few full-time or even three-quarter-time staff is crippling . If there were more full-time staff, they could afford to develop and deliver reports to staff members at various levels and in various roles in an Institute or Center. They could also educate organizational stakeholders as part of their jobs.

**Supporting Evidence**

- The percentage of "Don't know" answers about the use of different tool types was striking.
  - Question 15 asked about the use of Web analytics, and 27% selected "Don't know." Question 16 asked about the use of search analytics, and 41% selected "Don't know."
  - Question 17 asked about the use of social media analytics, and 29% selected "Don't know."
  - Question 18 asked about the use of customer-sentiment measurement, and 28% selected "Don't know."
  - Question 19 asked about the use of Internet industry/market research, and 44% selected "Don't know."
- Question 20 asked how often Web measurement is used to guide decisions; 39% of respondents selected "Some of the time," while 25% said they didn't know.

In the open-ended question about challenges that respondents related to Web site measurement strategies or tools (Question 23), one response was "Getting upper management to focus on the data and make major decisions based on it."

### *Compliance with 508 is a necessity that commands a large amount of attention and takes priority over other types of Web analytics.*

Just as the lack of full-time analytics staff prevents NIH Web measurement practices from maturing as they should, the need to satisfy 508 compliance guidelines prevents most NIH sites from making purely data-driven decisions about site design and user experience.

When asked in Question 14 about tools used to guide Web site strategy and tactics, "Web Analytics" and "508 compliance monitoring tool" were the two most popular responses. Both were selected by roughly 75% of respondents. This need to serve two masters is interesting; while this survey does not provide data that indicate how the 508 tool is used, most compliance monitoring does not offer any real insights into how to best serve audiences or achieve site goals. Exploring how the 508 compliance tool is currently used, as well as how it can be used in conjunction with Web analytics, probably deserves further exploration.

**Supporting Evidence**

- When asked what measurement methods guide Web tactics and strategy, 77% of respondents answered that the 508 compliance-monitoring tool played a role. Additionally, 76% selected Web analytics.
- Question 22 asked how measurement has helped with making decisions. Many respondents framed their answers in terms of how analytics helped with redesigns, for example:
  - "We continually make improvements to the design and navigation based on survey and other data.
  - "Helped in redesign to guide the changes. It also helps us understand what our visitors are looking for and interested in."
  - "The Web site measurements have been very effective for content and layout changes we have made in our pages."

Any redesign, change in content, method of navigation, or graphical update must comply with 508 standards.

### *Web-measurement stakeholders at NIH have a unified and consistent idea of who their audiences are and what their goals are.*

Question 8 asked respondents to indicate the key audiences for their sites. The responses to this question were highly stratified, with each of the several possible responses selected by at least 40% of respondents. In other words, respondents indicated that on the whole, they were concerned about multiple audiences, such as researchers *and* patients; consumers *and* advocates. The audiences include just about everyone – not in equal amounts, necessarily, but in significant enough numbers that all of them should be carefully considered in defining site goals.

Fortunately, the goals mentioned by respondents seemed to be in line with these recognized audiences. A common theme throughout the survey responses was the primacy of preparing and disseminating content. In the responses to Question 7, about site mission and goals, one respondent said their goals were to "Showcase scientific research being conducted at NIH; serve as an interactive communication where ideas are exchanged, opinions voiced, and issues examined. Create a forum that promotes cross-fertilization of research insights and collaboration across institutes and centers."

When respondents were asked in Question 9 about the top tasks for users on their sites, the most frequent response was content consumption. The second most popular response was "To apply for research, grants, or training," and the third most popular was "Order publications or other NIH products."

In addition to establishing the primacy of content on NIH sites, these responses underscored the importance of education and furthering the NIH mission.

The emphases on content and education are both consistent with respondents' statements about the audiences they intend to serve and the goals of their sites.

**Supporting Evidence**

- Question 7 (open-ended) asked survey respondents to describe their Web site's main mission or business goals. The responses focused on the dissemination of content to the general public and to health industry professionals — to promote overall public health, per the NIH mission.
- Question 8 (multiple-choice) asked respondents to identify their sites' audiences from these . Choices were advocates, consumers, educators, patients, health care providers, researchers, and public health officials. These are all consistent with the broad mission addressed in Question 7. When asked in Question 9 about top site-based tasks, 78% of respondents chose "Consume Content."
- All other popular answers to Question 9 included consuming site content and education, either directly through NIH content and publications or indirectly through supporting grants and other research work.

***Web content is of primary importance for most NIH sites, but figuring out how visitors access and interact with content is complex and often not addressed holistically.***

Questions 15–19 surveyed respondents about the use of tools for different types of analytics (Web analytics, search analytics, social media, customer sentiment, and Internet industry/market research). In all categories, while Web analytics tools were a fairly popular response, niche and platform-provided tools were extremely popular. For example, in the social media question, direct, platform-provided Facebook and bit.ly statistics were often mentioned.

This suggests two things: that NIH ICs often select tools that deliver "direct" statistics from service providers, like Facebook and Twitter. This means that that the ICs are overloaded with data sources that are difficult to combine. The direct tools provide statistics, but, since they are offered for free by service providers, they have no incentive to add value beyond the direct delivery of data.**[** Given that NIH has limited resources for working with Web data, using these tools probably prevents NIH from getting any metrics other than low-level, baseline statistics. While it is useful to have data on search, social media, and voice of customer, it is even more useful to be able to see the impact of these data points on each other. (The survey did not reveal anything about efforts at NIH to integrate these various data points, and this certainly warrants further analysis.)

**Supporting Evidence**

- Respondents indicated that NIH ICs often use the simple tools provided by platforms themselves or direct product feedback, rather than tools that add analytical value and can really enrich the ways data are understood:
  - For search analytics (Question 16), 31% said they use "Search engine reports."
  - For social media (Question 17), 30% use Facebook statistics, 21% use bit.ly statistics, 13% use YouTube statistics, and 8% use blog statistics. Compare this with the 3% using Radian 6, a leading, fee-based social media measurement tool.
  - For customer sentiment (Question 18), 37% said they use comments collected online and 22% use call center feedback.

- In Question 21, 19% of respondents indicated, that they do not know how to use available analytics tools.
- In Question 23, one respondent indicated that they have "too much data." In Question 24, another respondent said, "Currently we have to fend for ourselves in a sea of expensive products that we don't understand enough to get the most from."

### *Institutes and Centers at NIH have a strong desire for a standardized set of best practices, as well as a forum for changing Web-measurement information and expertise.*

The last survey question asked one thing that would help respondents make more effective use of Web measurement. Multiple respondents asked for best practices, with additional mentions of a desire for a centralized set of NIH tools that could be used for Web measurement. One respondent stated, "If there was some sort of best practices checklist I could use to ensure we are covering all of our bases in terms of info we should use to make decisions, that would be helpful, or NIH staff experts available for in person consultation to review current practices against best practices, and discuss possible options for addressing any discrepancies."

Similarly, with regard to tools, one respondent said, "It would be nice to have some centrally provided measurement tools. Currently we have to fend for ourselves in a sea of expensive products that we don't understand enough to get the most from."

**Supporting Evidence**

- In addition to the quotes cited in the paragraph above, other responses included:
    - “I would like to be able to compare our data to that from other, comparable sites.”
    - “Implement a Web measurement tool (NIH-wide) that is (1) easy to use and (2) does not have all of the extraneous spider data.”
    - “An NIH Service Center model would be tremendously helpful for all things related to analytics and third-party software, including social media. Since we are a very small office, it would be nice if we could contact someone at NIH who would actually process all the documents necessary for us to implement any new tool/software. For example, we would like to use Google Analytics, Facebook, etc. It would be wonderful if we could make a request and have an outside person at NIH take care of drafting our privacy policy, getting approval for third party software, ensuring we are in compliance with all requirements, and guiding us through implementation. We just don't have the expertise or time to implement many of the things that would be beneficial to us. The WAG and Web Metrics group are helpful, but we still lack the time and technical expertise in-house to fully understand the vast amount of information coming out of these groups on a daily basis. A service center would be a real plus.”
    - “A Google Analytics dashboard that would allow comparisons across ICs.”

## Conclusions

Web-measurement stakeholders and staff at NIH have a good idea of their audiences, goals, and top tasks on their Web sites but are under-prepared, -educated, -staffed, and -resourced for moving from general awareness to a state-of-the-art analytics organization.

The two key areas of improvement are resources and education. From a resource standpoint, the survey revealed that Web measurement is not a responsibility that commands much staff or many labor hours. This prevents the development and delivery of reports, and, perhaps more important, it prevents the development of the expertise necessary to collect understandable, trusted data.

In terms of education, the survey respondents repeatedly indicated that they would welcome training opportunities. The number of tools and different types of data involved in measuring all aspects of a Web site can be overwhelming—especially in situations like those at NIH, where every decision driven by measurement must also comply with 508 requirements. Along with training, respondents would like a set of best practices to support their attempts to use measurement data and tools consistently across NIH.

While the survey indicated that there are many deficiencies in the Web analytics practice at NIH, there are also real opportunities to improve the situation quickly through education, organization-wide standardization, and, as the capabilities of staff improve, devoting more staff hours to working with data.

## Question Analysis

In the following section, there are analyses of individual questions.

### *Questions 1–5: Respondent Information*

**The first four questions of the survey asked respondents for their names, job titles, and email addresses; the Web sites they work on; and the site URLs. Question 2 provided some context around the roles of the individuals who completed the survey.**

Forty-nine of the 131 respondents answered this question, and there was a great variety in the responses. Only 14 job titles included the word "Web," and only two included "Web analytics." This indicates that Web analytics at NIH is still very much a task that falls on staff members with other Web and non-Web related responsibilities.  At the same time, if the titles of the respondents are any indication, it seems that Web analytics is important to staff members at a variety of levels and on a variety of organizational teams. For example, 10 of the 49 respondents were managers and 5 were directors; 6 identified themselves as editors; and 4 were in public affairs. Other respondents indicated that their roles ranged from librarians, archivists, and technical writers to staff scientists.

### *Questions 6–9: Web Site Profile*

While most respondents work on the main IC sites, many work on sites that serve either more highly specific purposes or purposes that are applicable across multiple ICs. This means that NIH probably needs a large number of carefully considered models of success and that an overarching framework for measurement, including common success events and standard metrics, would be very useful.

**Question 7: What are the Web site's main business or mission goals?**

There were two key elements to the responses to this question: 1) nearly every response indicated that the baseline purpose of their site was the dissemination of information and 2) their site's content is tailored to the needs of several distinct audiences. Two of the more eloquent, overarching responses were these:

> "To advance NIH's mission to seek fundamental knowledge about the nature and behavior of living systems and to apply that knowledge to enhance health, lengthen life, and reduce the burdens of illness and disability."

and:

> "Showcase scientific research being conducted at NIH; serve as an interactive communication where ideas are exchanged, opinions voiced, and issues examined. Create a forum that promotes cross-fertilization of research insights and collaboration across institutes and centers".

One respondent stated that their site's purpose was to "serve and inform grantees, the public, the media, [and] also Congress and advocacy groups." Another respondent identified their audience as including the two simple categories of "public" and "researchers." In fact, there appear to be several different types of visitor breakdowns:

- NIH employees and members of the public at large
- researches, health professionals, and the public at large
- ethnic groups (for example, the Latino community served by cancer.gov en español)
- educational vs. purely informational

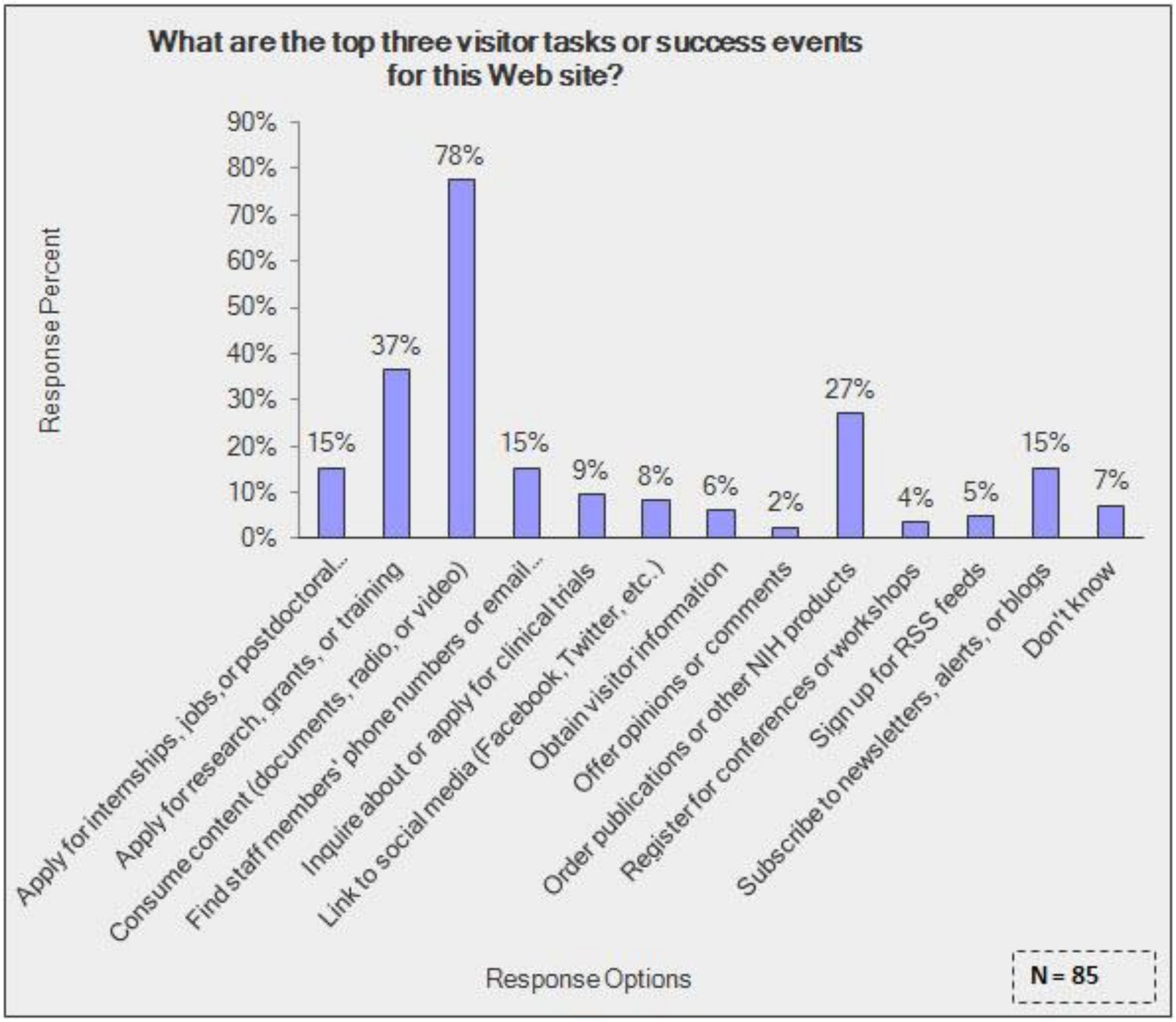
**Question 8: Who are the primary audiences for this Web site?**

This question was answered by 85 respondents, who could check all options that applied to their sites. While "Researchers, scientists, and investigators" was the most popular option, the least popular one ("Public Health Officials") was still relevant to nearly 40% of respondents. The underlying message here reinforces the point that NIH sites are responsible for informing several different audience groups, each with distinct needs.

**Question 9: What are the top three visitor tasks or success events for this Web site?**

Respondents to this question were asked to select up to the three options. Far and away the most popular response was the "Consume content." Many of the next most frequently selected responses suggest that, furthermore, NIH Web sites have a mandate not just to enable consumption of content, but to keep audiences informed moving forward ("Subscribe to newsletters, alerts, or blogs") and to enable them to use the content in important ways that fulfill the NIH mission referenced in the discussion of Question 7 ("Apply for internships, jobs, or postdoctoral programs" and "Apply for research, grants, or training").

***Figure A.1.*** *Question 9 from the Developing Web Analytics Best Practices for NIH Survey*



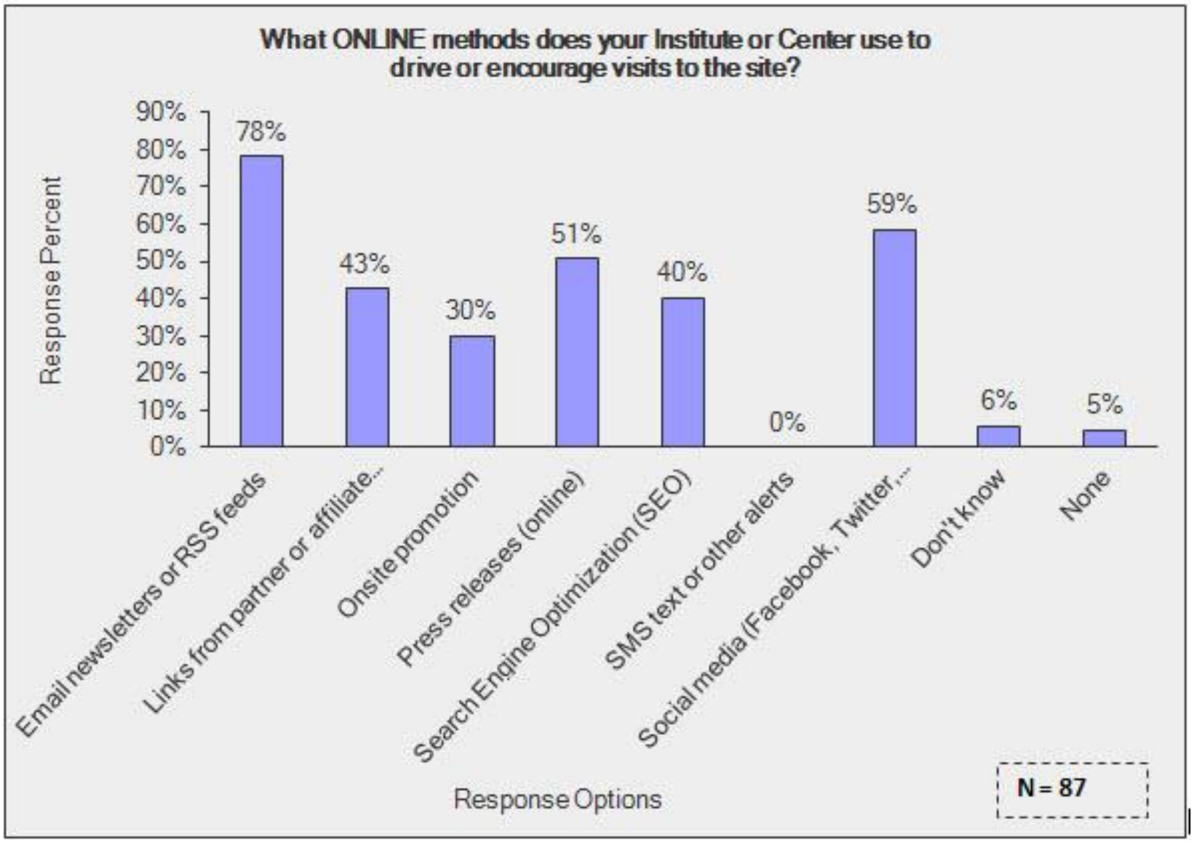
### *Questions 10–11: Generating Traffic*

**Question 10: What ONLINE methods does your Institute or Center use to drive or encourage visits to the site?**

Like many content-based Web sites—public or private—NIH combines newsletters and RSS feeds to drive traffic to the site. Perhaps less characteristic of standard Internet protocol is NIH's use of press releases as a popular tool. Also uncharacteristic is the low number of sites that use search engine optimization (SEO), a standard practice considered a "must have" for sites that are serious about generating traffic through non-paid search.

It is encouraging that the NIH is well aware of the growing importance of social media in disseminating content and purpose to an active and engaged online audience.

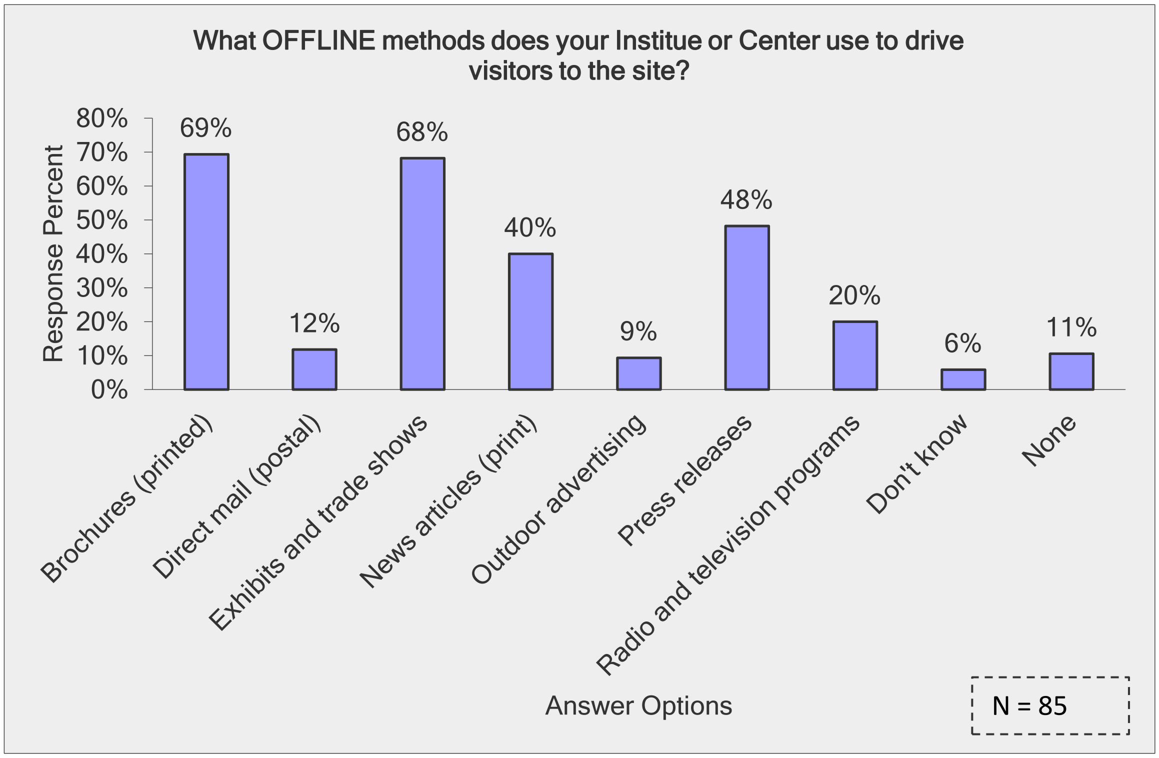
***Figure A.2.*** *Question 10 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 11: What OFFLINE methods does your Institute or Center use to drive visitors to the site?**

NIH Web sites use a variety of offline methods to encourage users to visit their sites. Many of these are print materials, but "Radio and television programs" did receive a fair number of responses. Note that "press releases" are important here, as they were in the question about online methods. The fact that there is overlap between the online and offline options for generating traffic begs the question of whether NIH studies and/or understands the interplay between its online and offline methods of generating user traffic. This topic probably deserves further exploration.

*Figure A.3. Question 11 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



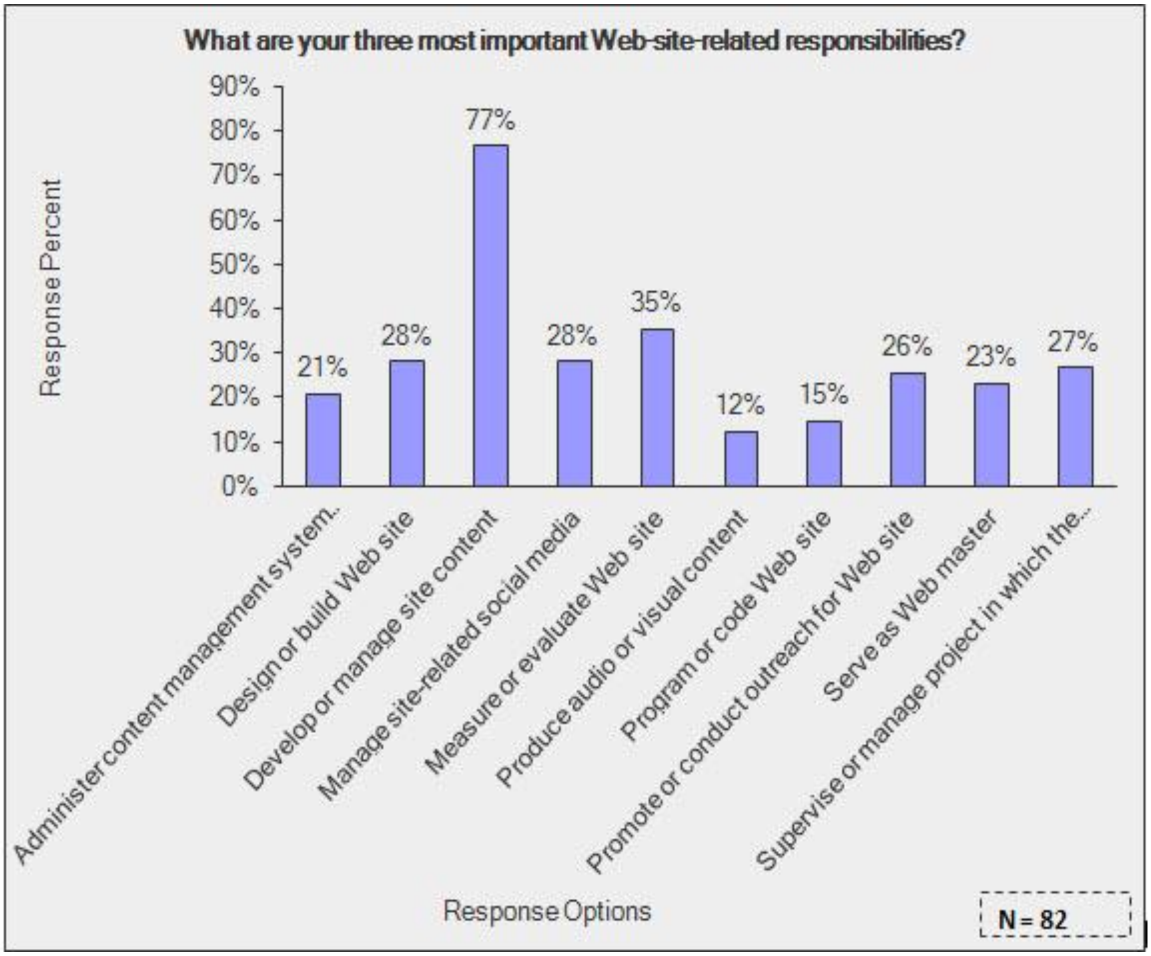
### *Questions 12–14: Web Analytics Responsibilities*

**Question 12: What are your three most important Web-site-related responsibilities?**

Responses to this question further reinforce the idea that most NIH Web analytics stakeholders and staff are not in roles that focus primarily on measurement. However, "Measure or evaluate Web site" was the second most popular response, indicating that the level of penetration of analytics at NIH is not insignificant.

It was not surprising that many of the most popular responses were similar to responses to previous questions, since they emphasized the importance of content and outreach, particularly via newer methods such as social media.

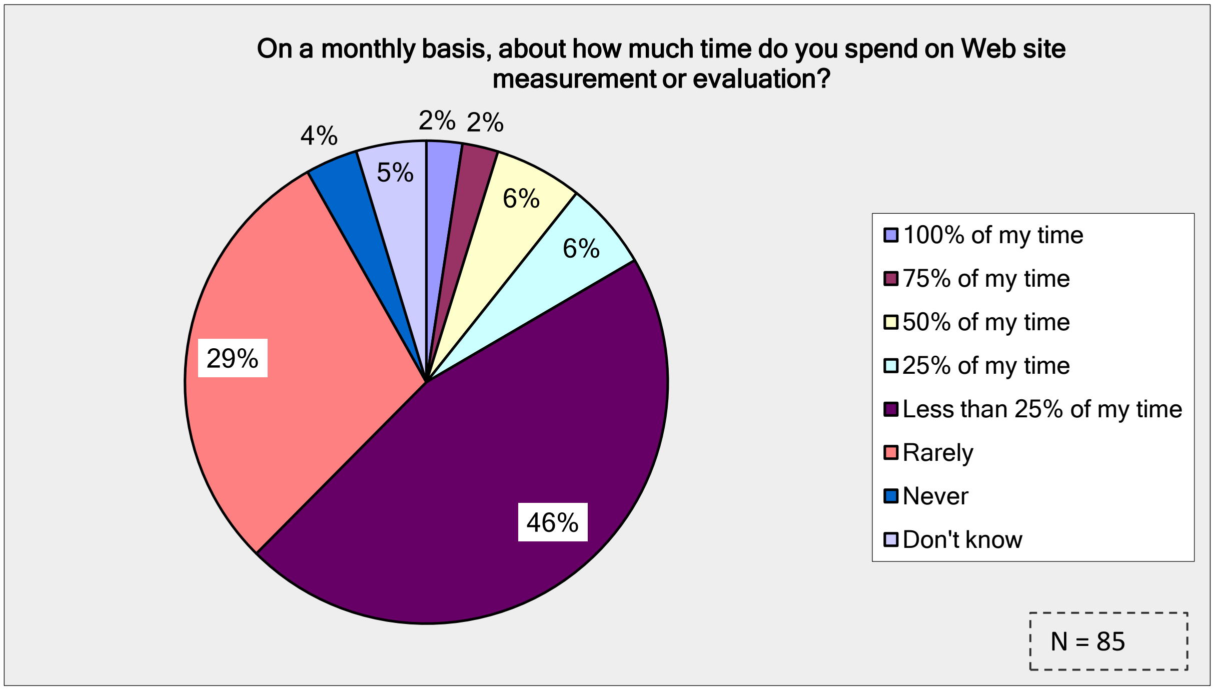
***Figure A.4.*** *Question 10 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 13: On a monthly basis, about how much time do you spend on Web site measurement or evaluation**

Web analytics work does not command full-time staff at NIH. Approximately 75% of respondents indicated that they spend "less than 25% of [their] time" on measurement or "rarely" spend time on it. Staff members who spend all their time on analytics represent less than 3% of total responses

***Figure A.5.*** *Question 13 from the Developing Web Analytics Best Practices for NIH Survey, 2011*

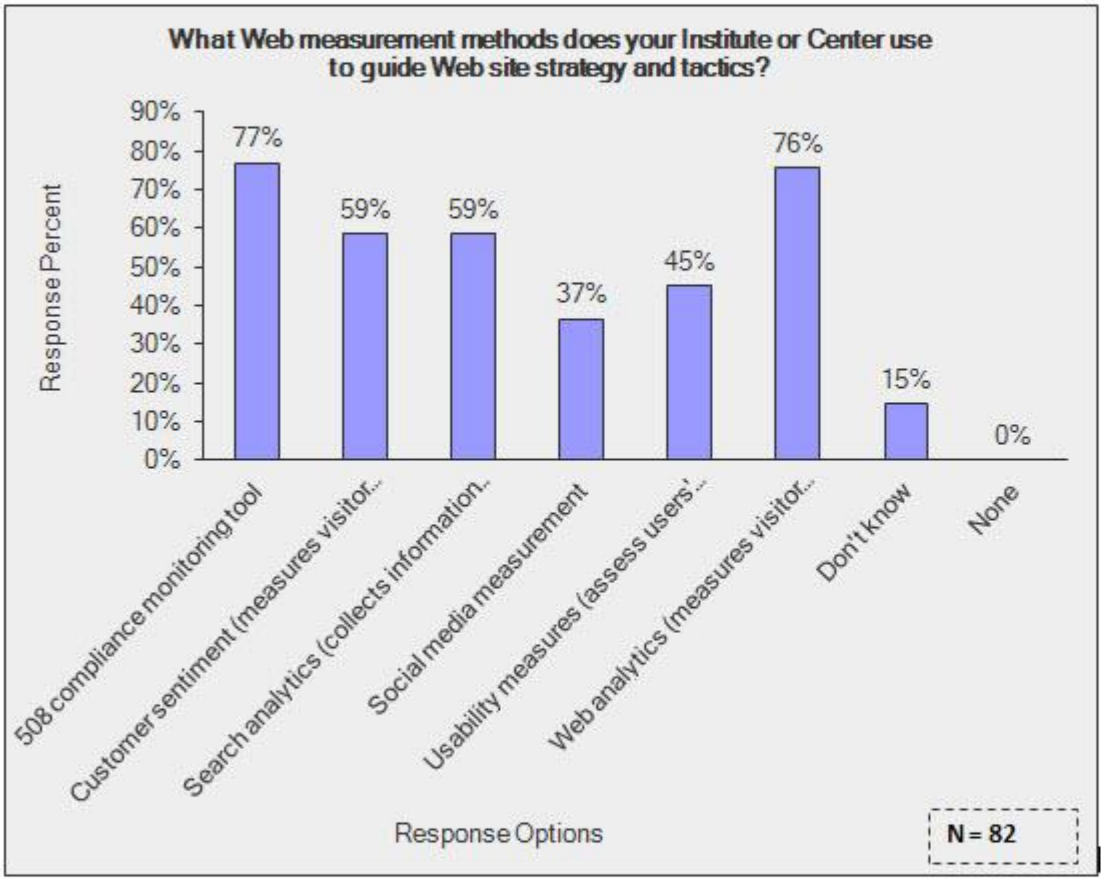


.

**Question 14: What Web-measurement methods does your Institute or Center use to guide Web site strategy and tactics?**

Respondents to Question 14 about measurement methods were allowed to select as many responses as applied to their Web sites. The result, as born out in the data, is an indication that most of the different methods suggested have a reasonable level of penetration at NIH. The two most popular methods were "Web analytics" and "508 compliance monitoring tools." This is implies an interesting duality in the roles that public sites have to fill. While they have the Web analytics data that most nonpublic sites would use as the basis of their site-optimization decisions, public sites have to fulfill the additional requirement of compliance. It would be interesting to find out whether—and if so, to what extent—do 508 compliance requirements constrain the ability of public sites to implement certain types of strategies or initiatives.

***Figure A.6.*** *Question 14 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



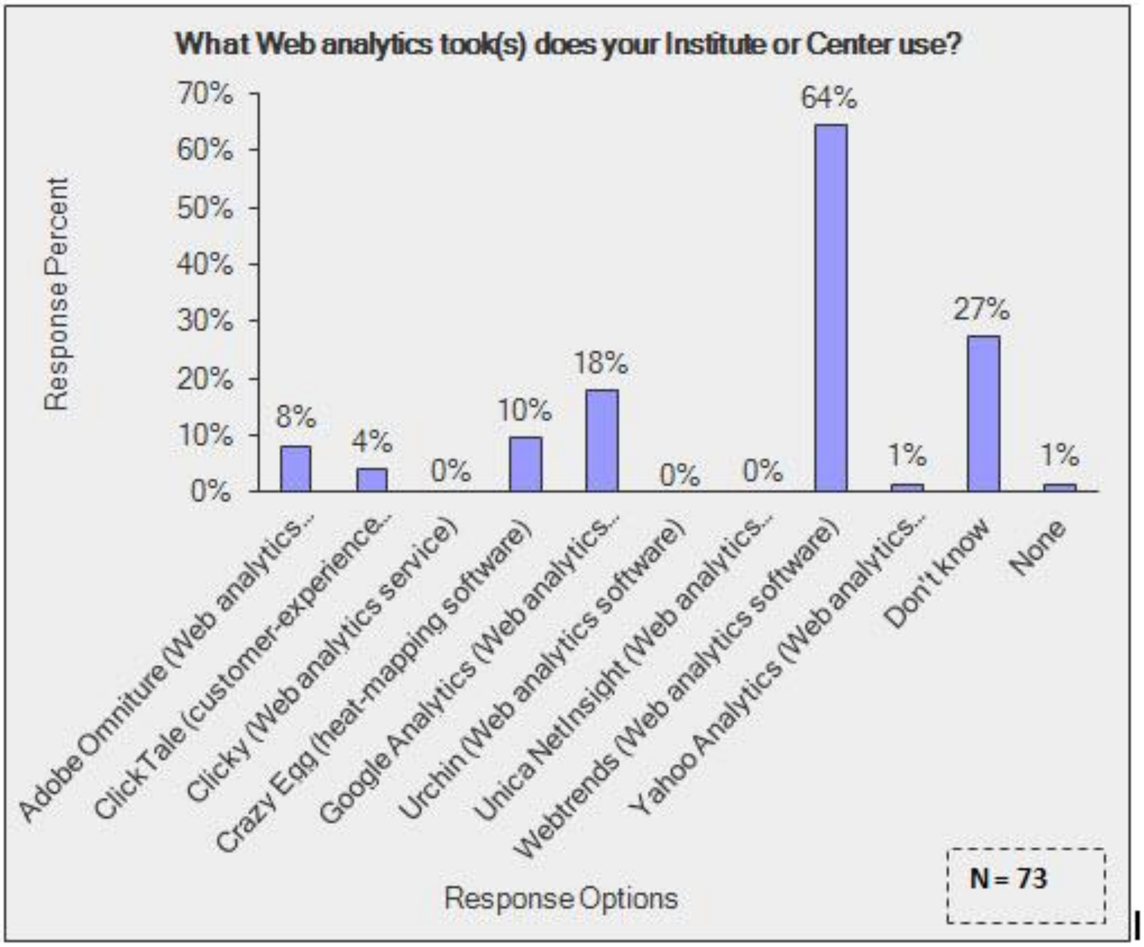
## *Questions 15–19: Tool Usage*

**Question 15: What Web analytics tool(s) does your Institute or Center use?**

Among the respondents to this question, representing a little more than half of all survey respondents, Webtrends was by far the most popular choice. Nearly 30% of respondents indicated that they did not know what tools were used by their Web site. This result, coupled with the fact that a significant number of individuals did not respond suggests that a large number of stakeholders and staff were not aware of the basic elements of their site's analytics program.

A second interesting finding from Question 15 is the level of Google Analytics use at NIH. Google Analytics is a free tool that represents an excellent stand-alone solution and, alternatively, can function as a checking or auditing system against a more expensive paid solution. In either case, as a solution with costs measured in staff and level of effort, rather than server calls and contract options, should be much more widely used. Since Google Analytics has only been available to the Federal Government for a year, we think that nearly 20% usage is noteworthy. We expect this percentage to increase significantly over the next 12 months.

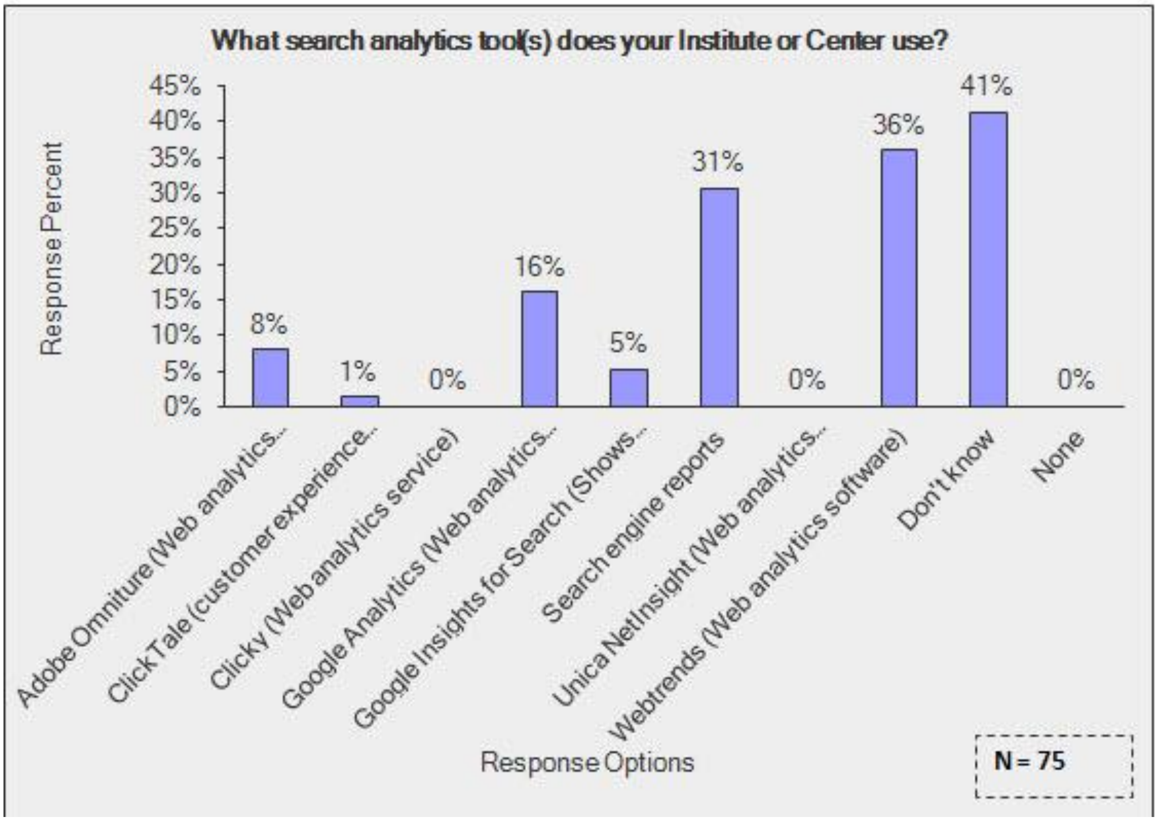
***Figure A.7.*** *Question 15 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 16: What search analytics tool(s) does your Institute or Center use?**
Responses to this question indicate both that there is a low level of awareness of search analytics and that when search analytics are used, the primary tool is still very often a Web analytics tool.

The level of reliance on search engine reports is also high, which begs the question of how the data are combined and used with other site data.

***Figure A.8.*** *Question 16 from the Developing Web Analytics Best Practices for NIH Survey, 2011*

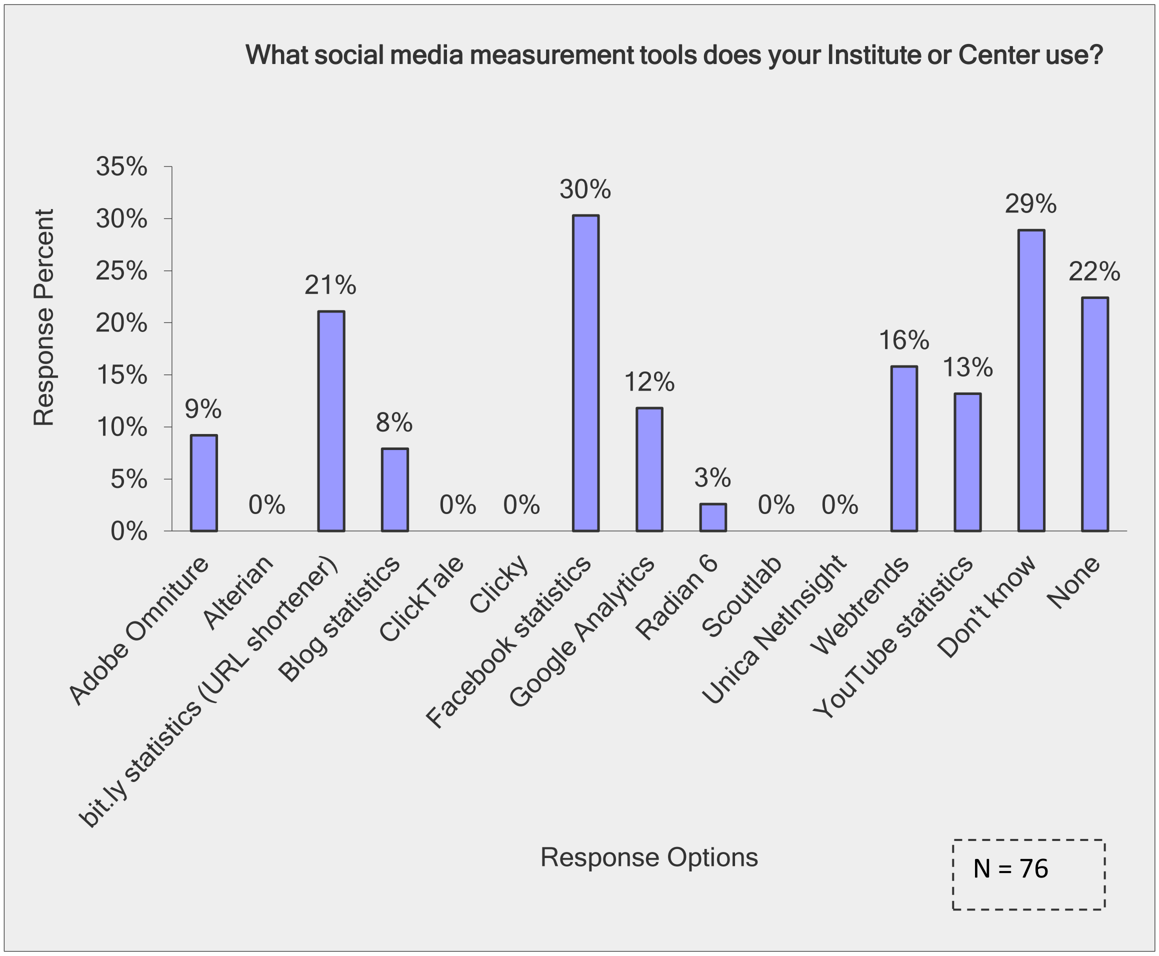


**Question 17: What social media measurement tool(s) does your Institute or Center use?**

Responses to this question suggest that for social media, the use of platform-specific and proprietary tools is more common than that of full-service Web analytics tools. Facebook statistics and bit.ly statistics were both more popular responses than the most popular Web analytics tool response. YouTube statistics were nearly as popular as Webtrends, which was the second most popular Web analytics tool, and more popular than both Omniture and Google Analytics—probably the two biggest players in the Web analytics space.

A high percentage o the responses were "Don't know" and "None." This *could* indicate that NIH is a little behind in terms of recognizing the importance of using and measuring social media options.

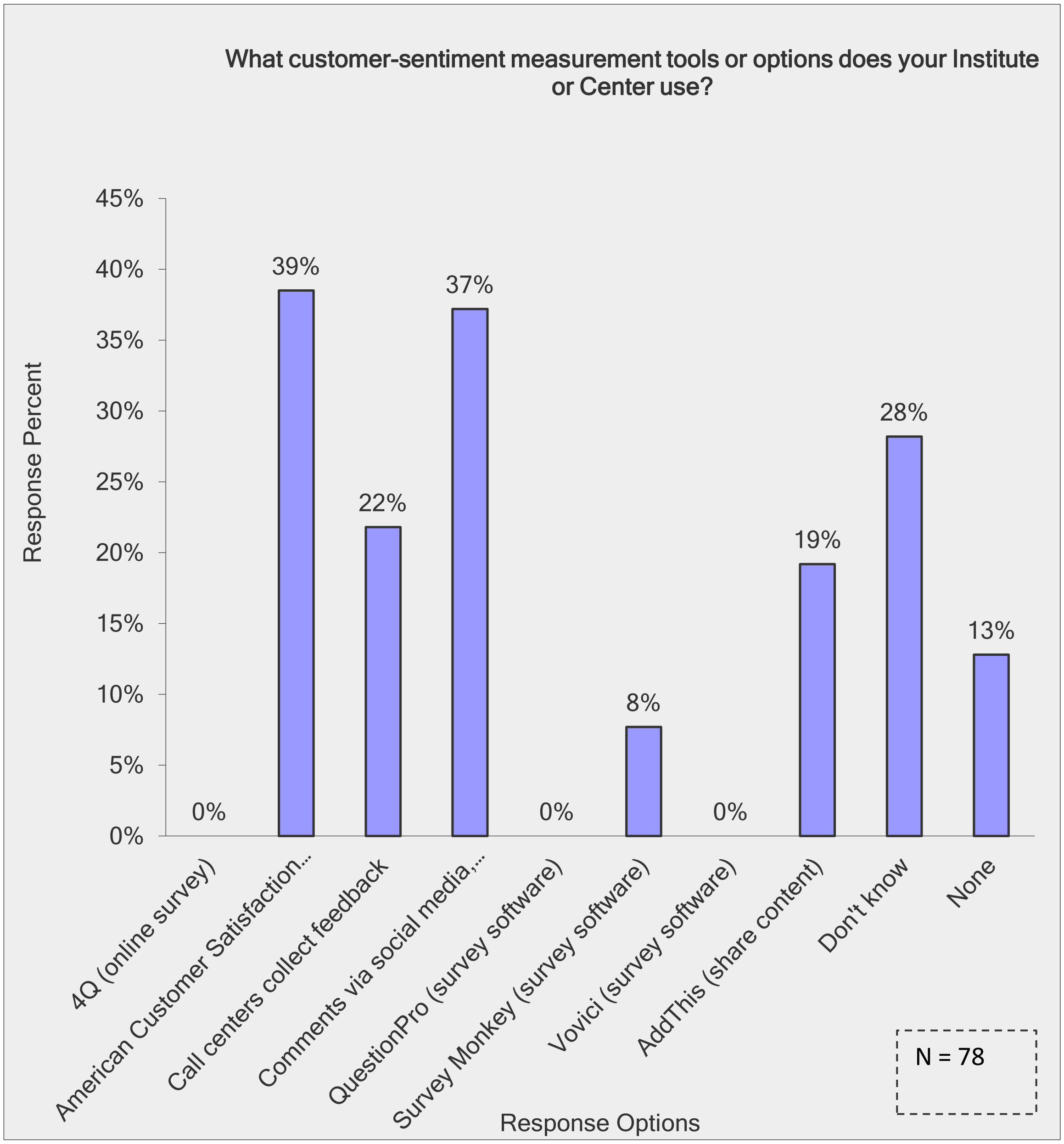
***Figure A.9.*** *Question 17 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 18: What customer-sentiment tools or options does your Institute or Center use?**
Teams at NIH are primarily using analysis of direct comments and the ForeSee Customer Satisfaction surveys to gauge customer sentiment, rather than direct surveys. Using direct surveys, which ask users questions about specific aspects of their experiences as users, provides more relevant data than ForeSee Customer Satisfaction surveys and less biased data than other observational methods. SurveyMonkey was selected a few times—by approximately 8% of the 78 respondents who answered this question. The majority of the other tools selected represent particular user actions that can be directly observed, like AddThis, commenting, and call-center feedback.

The customer satisfaction tool ACSI was the most popular response. For many, it serves as a benchmarking tool.

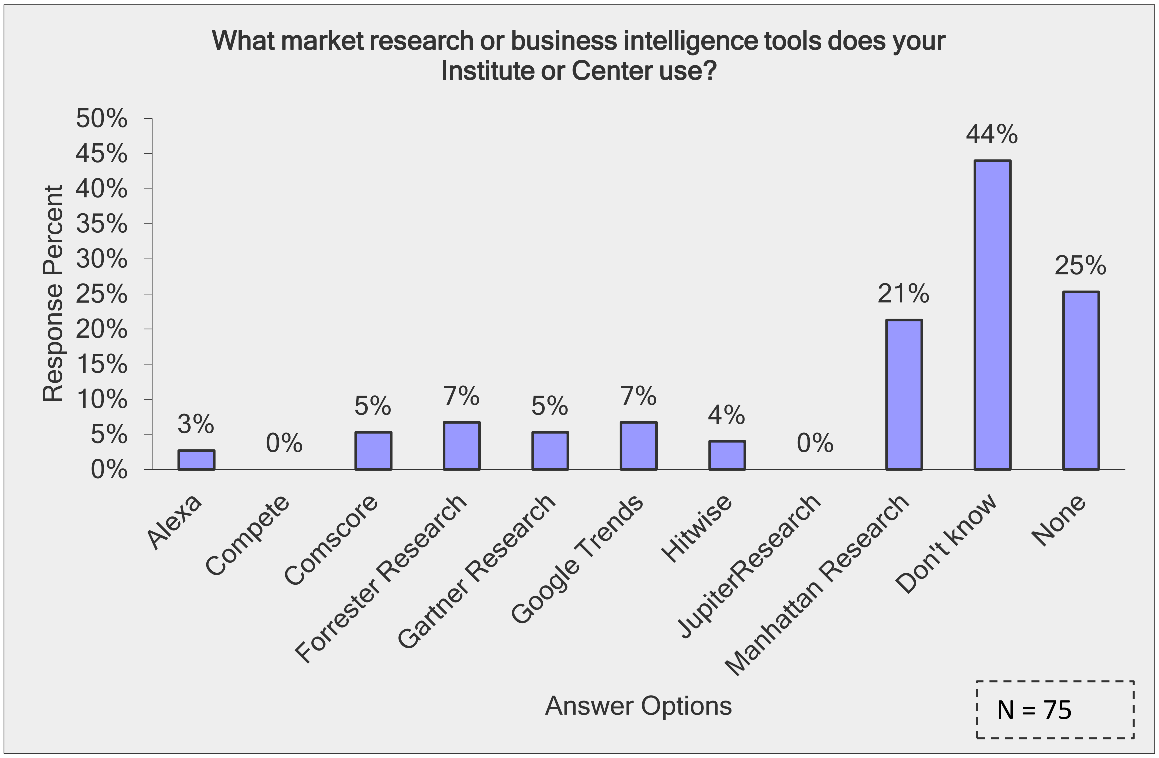
***Figure A.10.*** *Question 18 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 19: What market research or business intelligence tools does your Institute or Center use?**

The level of use of Internet industry or market research or business intelligence tools at NIH is very low. Only one tool, Manhattan Research, was chosen by more than 7% of respondents. Furthermore, "Don't know" and "None" were both chosen by more than 25% of respondents.

***Figure A.11.*** *Question 19 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



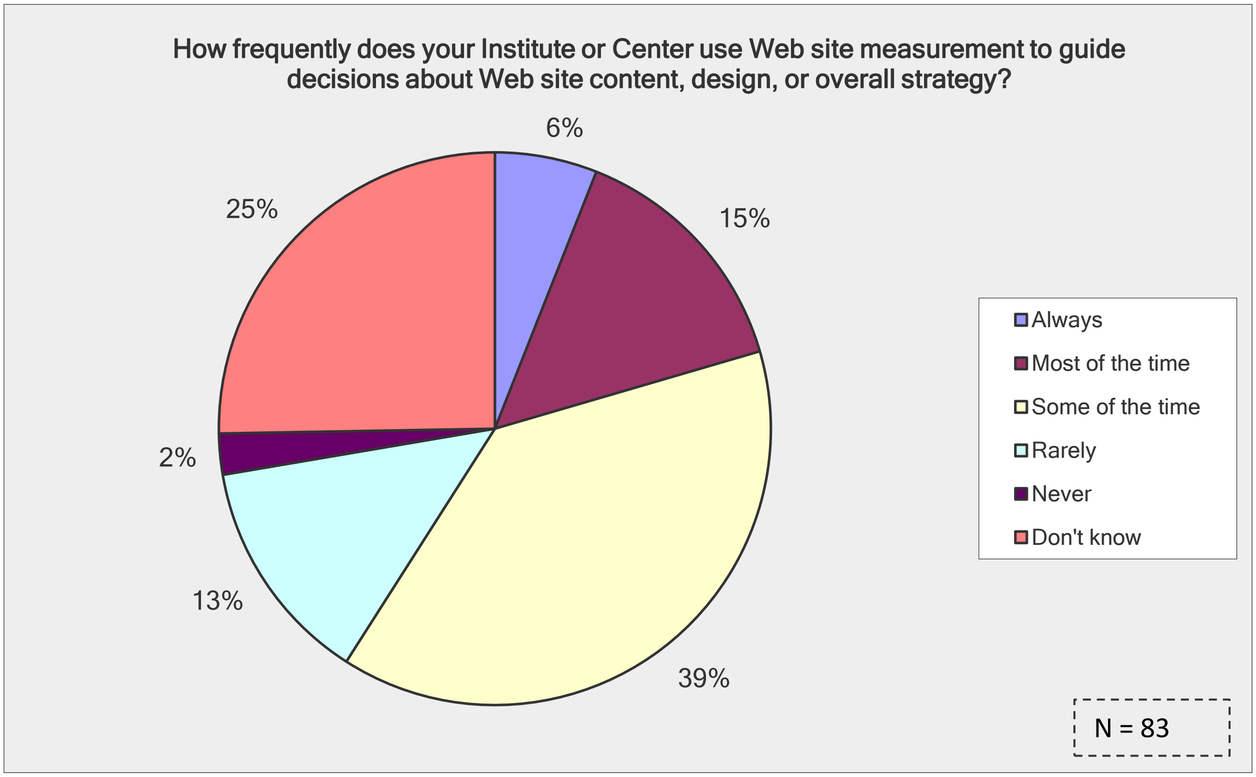
### *Questions 20–24: Overall Evaluative Thoughts on Web Analytics*

**Question 20: How frequently does your Institute or Center use Web site measurement to guide decisions about Web site content, design, or overall strategy?**

The responses to this question fall in line with findings from previous questions that address a similar topic: Web analytics at the NIH are not seen, or at least are not used, as vital and necessary tools. "Most of the time" and "Always" responses accounted for only 20% of total responses, while "Some of the time," the most popular response, was selected by roughly 40% of respondents.

Only a very small percentage of the responses were "Never." Most indicated that Web measurement is used to some extent.

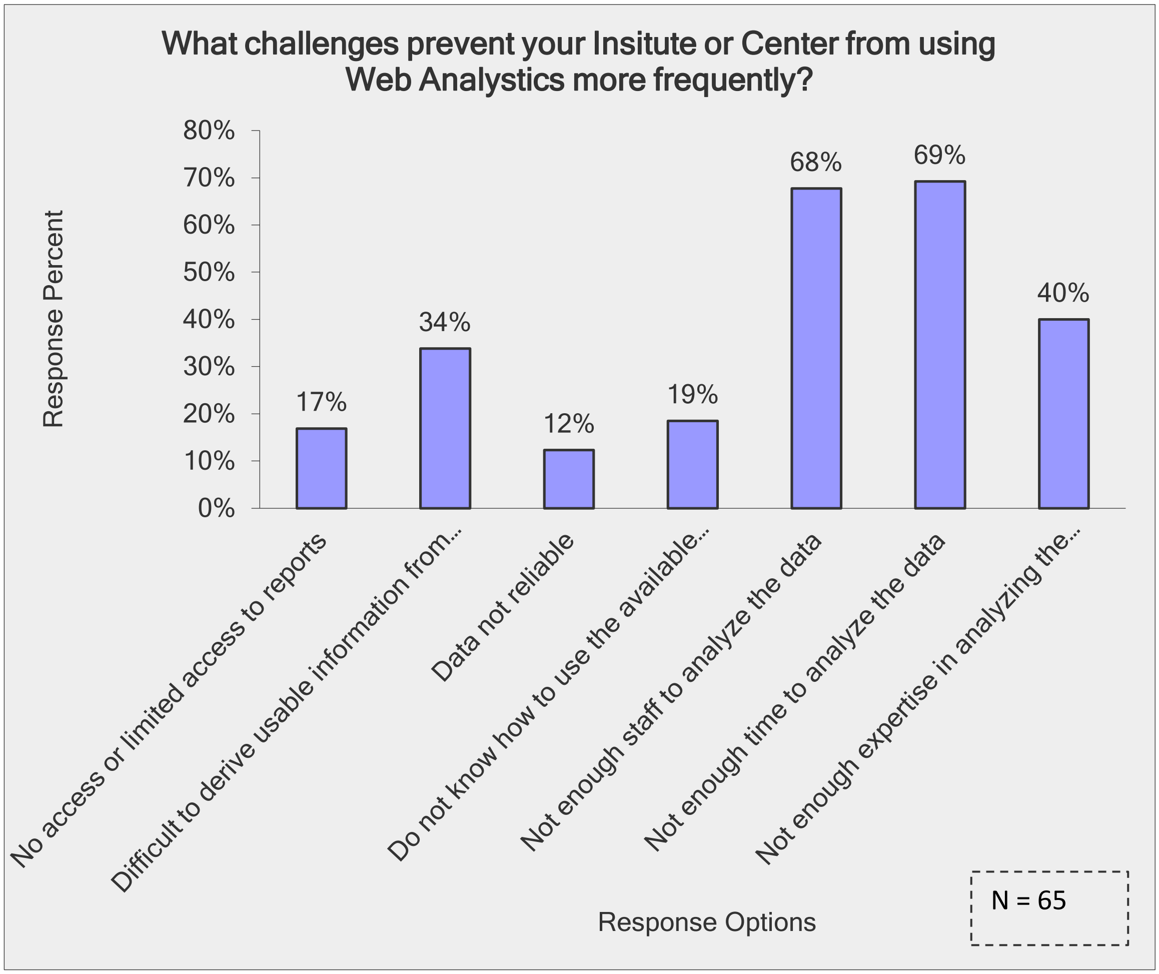
***Figure A.12.*** *Question 20 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 21: What challenges prevent your Institute or Center from using Web analytics more frequently?**

The responses to this question, which allowed respondents to select as many options as applied, give the overwhelming sense that the challenges of Web site measurement at NIH are the result of a lack of the necessary time and expertise for working with data. The two most popular responses emphasized a lack of staff and time, and the next most popular responses indicate a lack of expertise and general difficulty in working with data. These should be seen as direct results of the lack of staff and time.

***Figure A.13.*** *Question 21 from the Developing Web Analytics Best Practices for NIH Survey, 2011*



**Question 22: How has Web site measurement helped your Institute or Center with making decisions about adding content?**

Most responses to this open-ended question indicated that Web site measurement had in fact been useful. Such responses seemed to take one of two approaches.

First, NIH teams are using Web site measurement primarily as a tool for examining usability and guiding redesigns. One respondent stated, "We've used a combination of Webtrends, ACSI, and general feedback to determine who our audience really is and how they are navigating the site. For example, we know a high percentage of visitors use our internal search so we spent time perfecting our Google results and custom interface." Another respondent stated, more succinctly, "We rely almost exclusively on our data in making decisions about our navigation and architecture." As in the former response, several respondents indicated that the data used to make usability decisions are gathered from multiple sources.

Alternatively, many teams use Web site measurement as a tool for measuring content-related behaviors and interactions. Examples of this include

- "Unfound terms from user search has occasionally led to story ideas for newsletter."
- "We look at the number of visits each month and top downloads to see if our promotional efforts to drive traffic to the site have caused an increase in traffic."
- "The data have been very helpful to move discussions from ‘your opinion vs. my opinion’ to what works best for the people we're trying to serve."
- "Helped identify quick links to frequently requested reports, removal of infrequently used reports."

Unfortunately, several stakeholders and staff gave responses indicating that Web site measurement has not yet reached its full potential. Answers such as "Still working on this" and "It hasn't—no time or experience" and the fact that 83 respondents chose to skip this question indicated frustration with an utter lack of help with Web site measurement.

**Question 23: What was your biggest challenge related to using Web site measurement strategies or tools? How did you overcome the challenge?**

Respondents to this question cited several different challenges, but most fell into one of three categories: staff and resource, tool, and organizational challenges.

Staff and resource challenges are exactly what they sound like—cases where time and expertise are not as readily available as would be optimal. One respondent summed things up succinctly, stating, "We have severely limited staff, resources, and it's an ongoing challenge that probably won't be overcome anytime soon." In fact, most responses to the question indicated that challenges of this nature are by-and-large not overcome but are, rather, “lived with.”

Tool challenges include deficiencies in how analytics tools actually perform and are used. These range from a simple lack of understanding of the tool and suspicion of the data captured with it to the use of inadequate tools or the lack of resources to keep tools up to date. Of course, if the latter is the case, there is not a lot that can be done from within the Web analytics team. Instead, a reliance on IT staff and pleas for resources are probably in order. If lack of knowledge is the case, the key seems to be training. As one respondent said, "These challenges are not fully overcome, but we've made progress through training, experience, talking to others about their experiences, and a constantly evolving process of promoting analytics and learning what works and what doesn't."

Organizational challenges are situations where the limiting factors to the effectiveness of the Web measurement practice come from lack of organizational buy-in. The general consensus from the survey respondents was that the only real solution was persistence in making one's case for the value of Web measurement or, as one respondent put it, "being annoying/persuasive to budget holders."

**Question 24: What is one thing that your Institute or Center use Web site measurement tools and methods more effectively?**

Answers to this question varied greatly but ultimately boiled down to two now-familiar needs: resources and education. On the resources side, respondents once again cited a need for better tools, more staff, and more hours to work with data. In terms of education, among the things cited most were staff with expertise and training to develop such expertise.

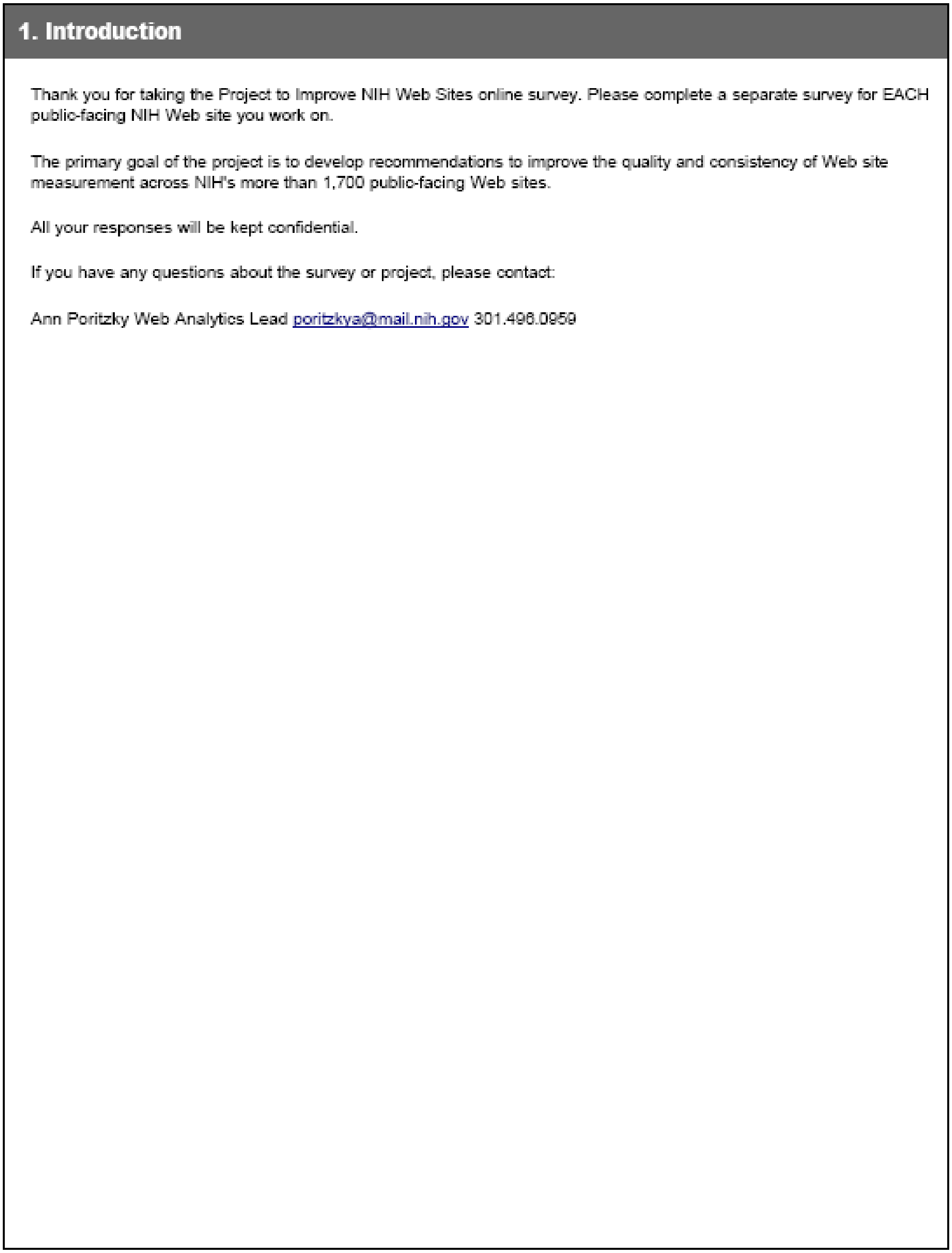
One new element among the answers to this question was a call for best practices and more collaborative efforts to share and benefit from knowledge across NIH. Some respondents looked for best practices to use as guidelines for work on their own sites. Others called for more open sharing of data, including "a consolidated site license for all of NIH" and "a Google Analytics dashboard that would allow comparisons across IC's."

### *Summary*

Survey responses both provided new information and corroborated the findings from the project's stakeholder interviews. In each case, respondents strongly articulated their interest in doing more with analytics, as well as the challenges they face in the current environment. Both sets of data are invaluable for guiding and developing the recommendations for the Developing Web Analytics Best Practices for NIH Project.

**The Developing Web Analytics Best Practices for NIH Survey, 2011** begins on the next page.

# The Developing Web Analytics Best Practices for NIH Survey, 2011



## 1. Introduction

Thank you for taking the Project to Improve NIH Web Sites online survey. Please complete a separate survey for EACH public-facing NIH Web site you work on.

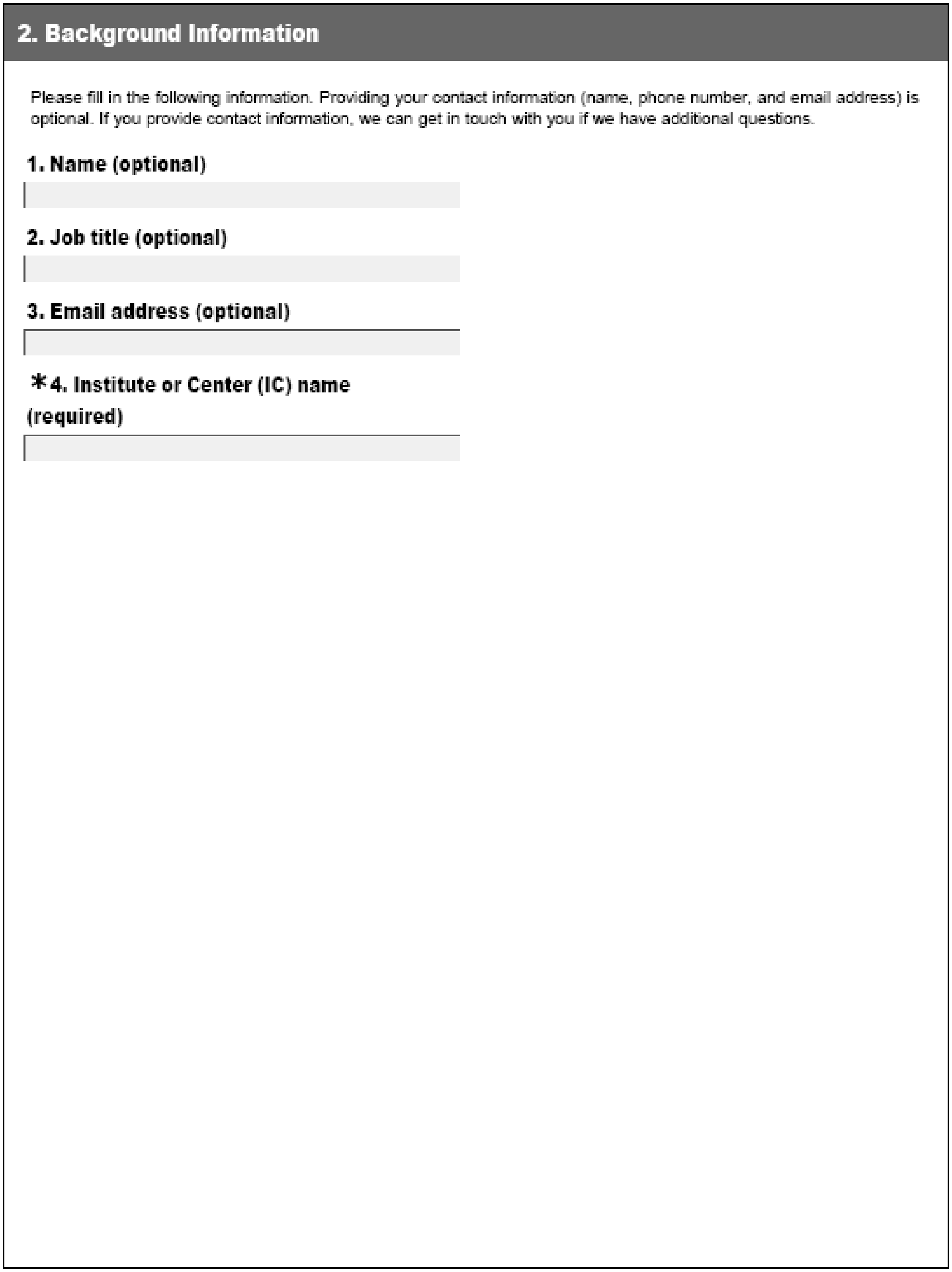
The primary goal of the project is to develop recommendations to improve the quality and consistency of Web site measurement across NIH's more than 1,700 public-facing Web sites.

All your responses will be kept confidential.

If you have any questions about the survey or project, please contact:

Ann Poritzky Web Analytics Lead poritzkya@mail.nih.gov 301.496.0959

# The Developing Web Analytics Best Practices for NIH Survey, 2011
(continued)



## 2. Background Information

Please fill in the following information. Providing your contact information (name, phone number, and email address) is optional. If you provide contact information, we can get in touch with you if we have additional questions.

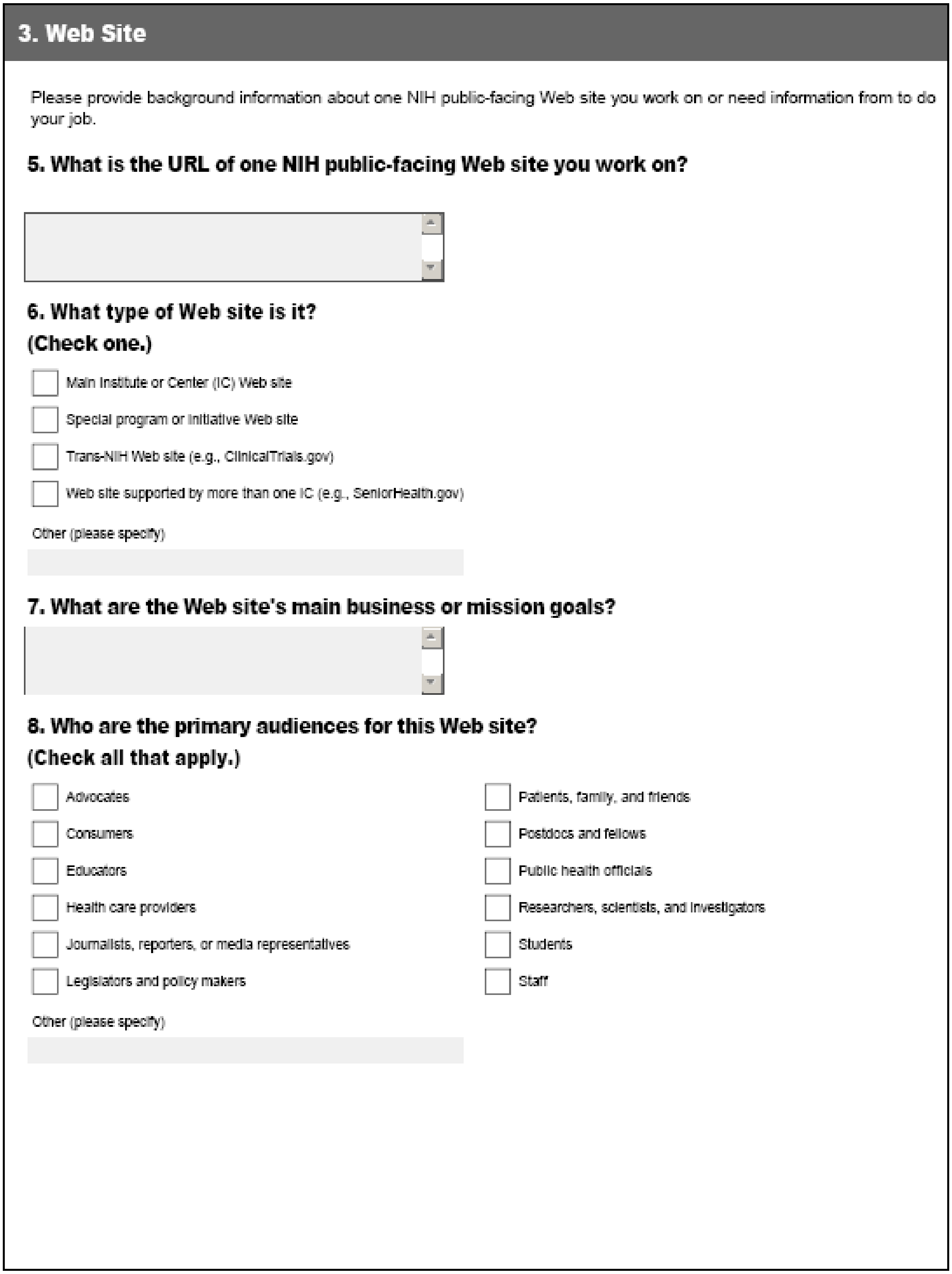
**1. Name (optional)**

**2. Job title (optional)**

**3. Email address (optional)**

***4. Institute or Center (IC) name (required)**

## The Developing Web Analytics Best Practices for NIH Survey, 2011
## (continued)



### 3. Web Site

Please provide background information about one NIH public-facing Web site you work on or need information from to do your job.

**5. What is the URL of one NIH public-facing Web site you work on?**

**6. What type of Web site is it?**
**(Check one.)**

☐ Main Institute or Center (IC) Web site
☐ Special program or initiative Web site
☐ Trans-NIH Web site (e.g., ClinicalTrials.gov)
☐ Web site supported by more than one IC (e.g., SeniorHealth.gov)

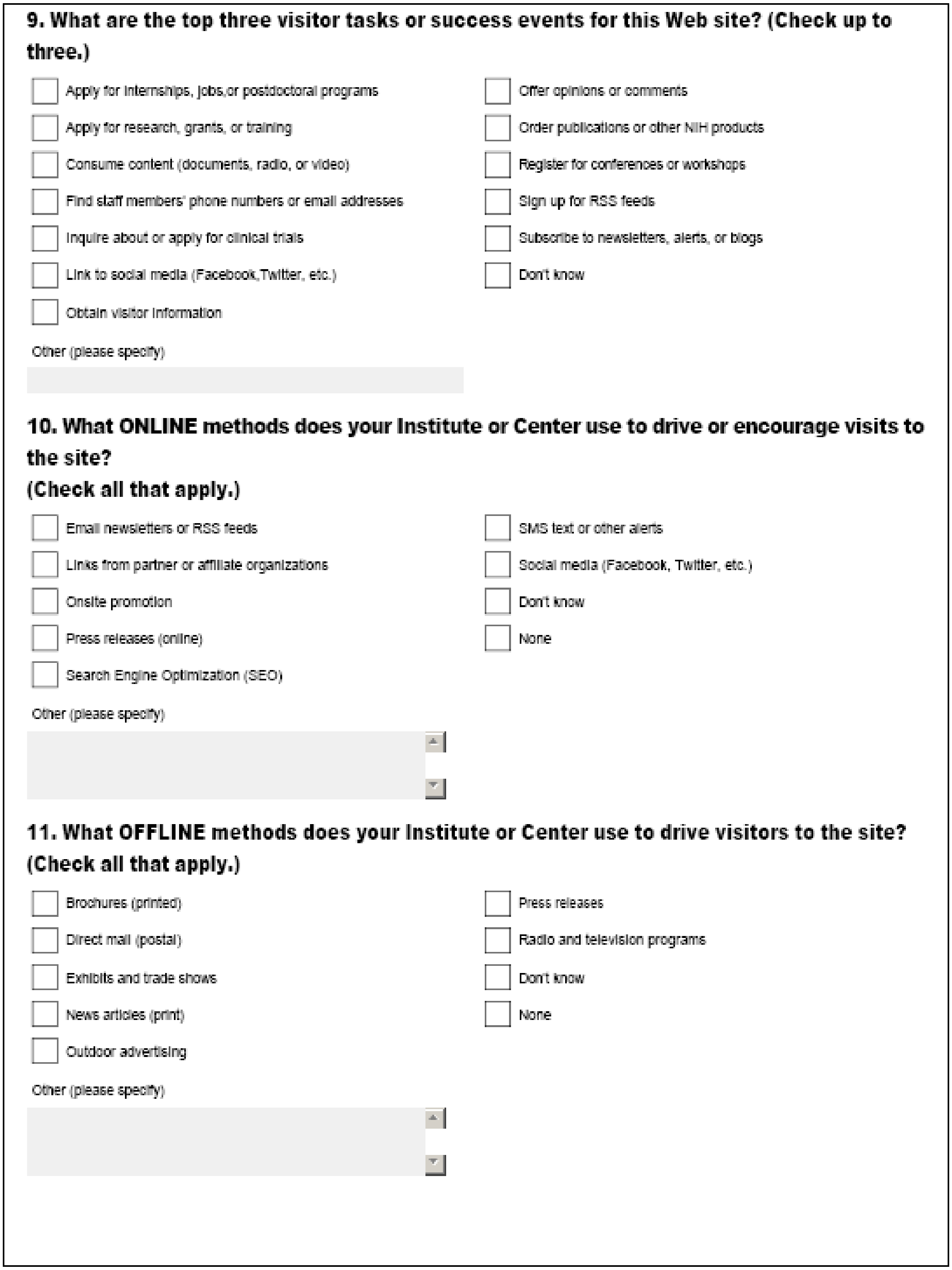
Other (please specify)

**7. What are the Web site's main business or mission goals?**

**8. Who are the primary audiences for this Web site?**
**(Check all that apply.)**

☐ Advocates
☐ Consumers
☐ Educators
☐ Health care providers
☐ Journalists, reporters, or media representatives
☐ Legislators and policy makers
☐ Patients, family, and friends
☐ Postdocs and fellows
☐ Public health officials
☐ Researchers, scientists, and investigators
☐ Students
☐ Staff

Other (please specify)



**9. What are the top three visitor tasks or success events for this Web site? (Check up to three.)**

☐ Apply for internships, jobs,or postdoctoral programs
☐ Apply for research, grants, or training
☐ Consume content (documents, radio, or video)
☐ Find staff members' phone numbers or email addresses
☐ Inquire about or apply for clinical trials
☐ Link to social media (Facebook,Twitter, etc.)
☐ Obtain visitor information
☐ Offer opinions or comments
☐ Order publications or other NIH products
☐ Register for conferences or workshops
☐ Sign up for RSS feeds
☐ Subscribe to newsletters, alerts, or blogs
☐ Don't know

Other (please specify)

**10. What ONLINE methods does your Institute or Center use to drive or encourage visits to the site?**
**(Check all that apply.)**

☐ Email newsletters or RSS feeds
☐ Links from partner or affiliate organizations
☐ Onsite promotion
☐ Press releases (online)
☐ Search Engine Optimization (SEO)
☐ SMS text or other alerts
☐ Social media (Facebook, Twitter, etc.)
☐ Don't know
☐ None

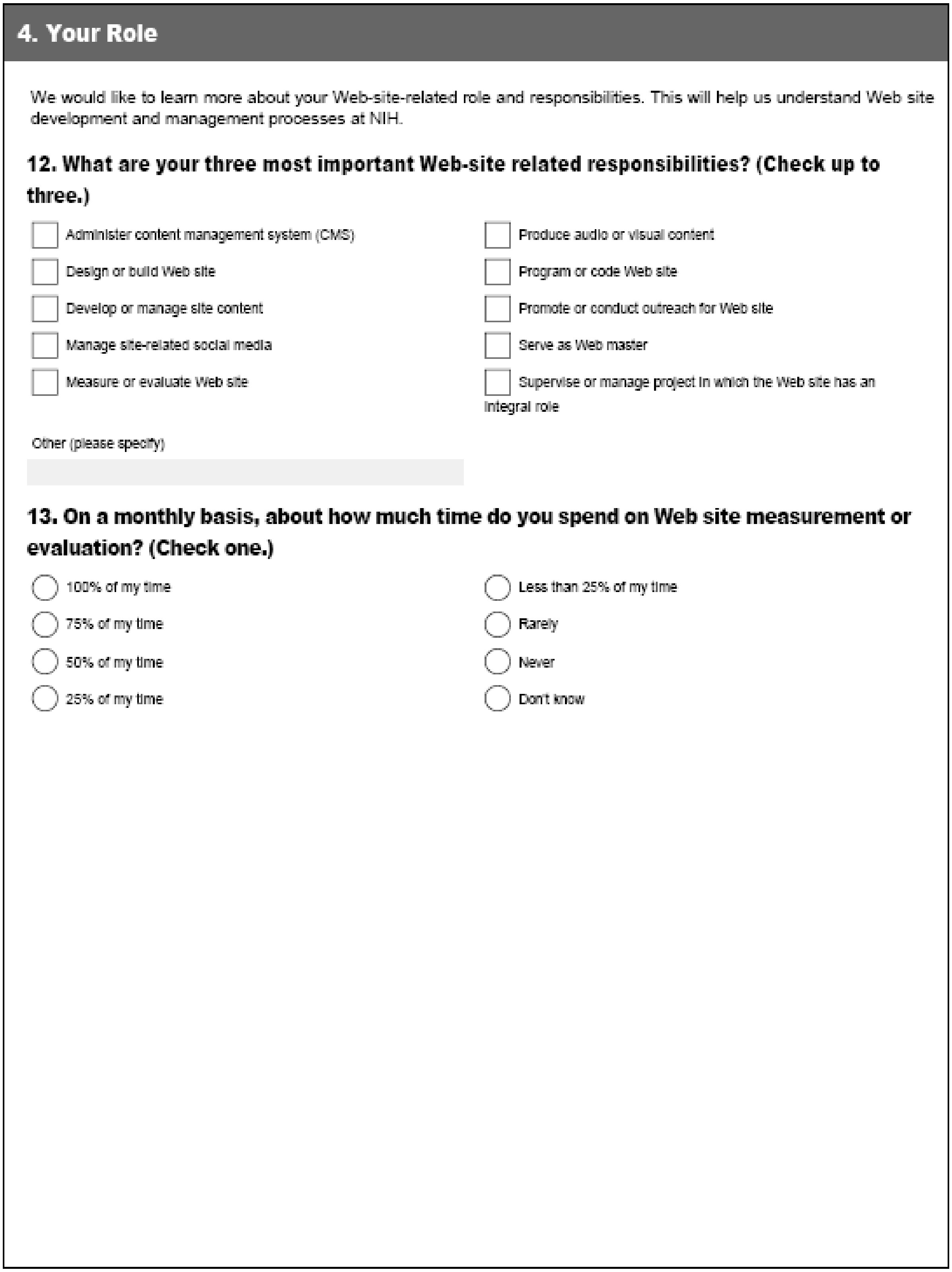
Other (please specify)

**11. What OFFLINE methods does your Institute or Center use to drive visitors to the site?**
**(Check all that apply.)**

☐ Brochures (printed)
☐ Direct mail (postal)
☐ Exhibits and trade shows
☐ News articles (print)
☐ Outdoor advertising
☐ Press releases
☐ Radio and television programs
☐ Don't know
☐ None

Other (please specify)

# The Developing Web Analytics Best Practices for NIH Survey, 2011
(continued)



## 4. Your Role

We would like to learn more about your Web-site-related role and responsibilities. This will help us understand Web site development and management processes at NIH.

**12. What are your three most important Web-site related responsibilities? (Check up to three.)**

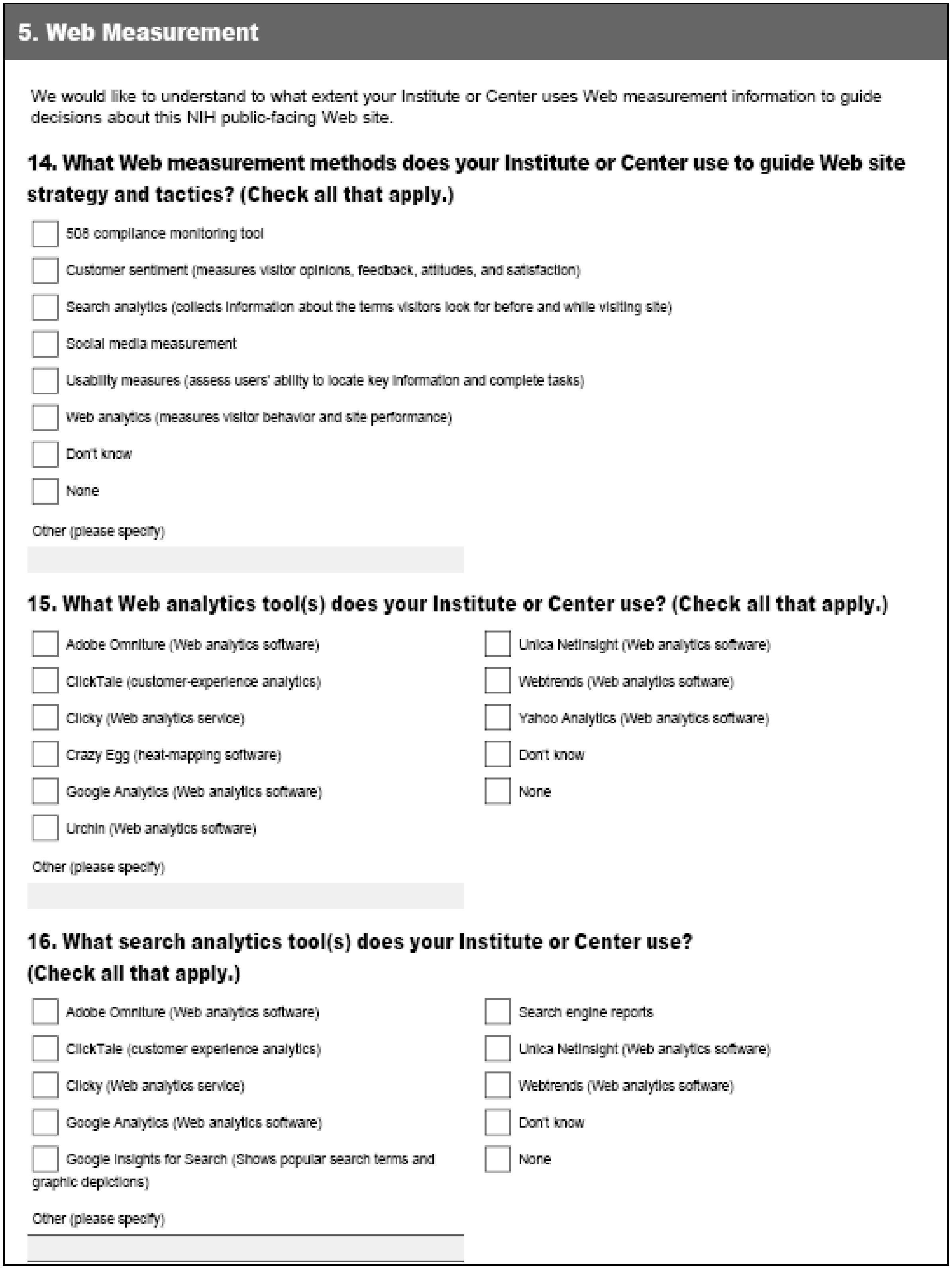
☐ Administer content management system (CMS)
☐ Design or build Web site
☐ Develop or manage site content
☐ Manage site-related social media
☐ Measure or evaluate Web site
☐ Produce audio or visual content
☐ Program or code Web site
☐ Promote or conduct outreach for Web site
☐ Serve as Web master
☐ Supervise or manage project in which the Web site has an integral role

Other (please specify)

**13. On a monthly basis, about how much time do you spend on Web site measurement or evaluation? (Check one.)**

○ 100% of my time
○ 75% of my time
○ 50% of my time
○ 25% of my time
○ Less than 25% of my time
○ Rarely
○ Never
○ Don't know

## The Developing Web Analytics Best Practices for NIH Survey, 2011
## (continued)



### 5. Web Measurement

We would like to understand to what extent your Institute or Center uses Web measurement information to guide decisions about this NIH public-facing Web site.

**14. What Web measurement methods does your Institute or Center use to guide Web site strategy and tactics? (Check all that apply.)**

☐ 508 compliance monitoring tool
☐ Customer sentiment (measures visitor opinions, feedback, attitudes, and satisfaction)
☐ Search analytics (collects information about the terms visitors look for before and while visiting site)
☐ Social media measurement
☐ Usability measures (assess users' ability to locate key information and complete tasks)
☐ Web analytics (measures visitor behavior and site performance)
☐ Don't know
☐ None

Other (please specify)

**15. What Web analytics tool(s) does your Institute or Center use? (Check all that apply.)**

☐ Adobe Omniture (Web analytics software)
☐ ClickTale (customer-experience analytics)
☐ Clicky (Web analytics service)
☐ Crazy Egg (heat-mapping software)
☐ Google Analytics (Web analytics software)
☐ Urchin (Web analytics software)
☐ Unica NetInsight (Web analytics software)
☐ Webtrends (Web analytics software)
☐ Yahoo Analytics (Web analytics software)
☐ Don't know
☐ None

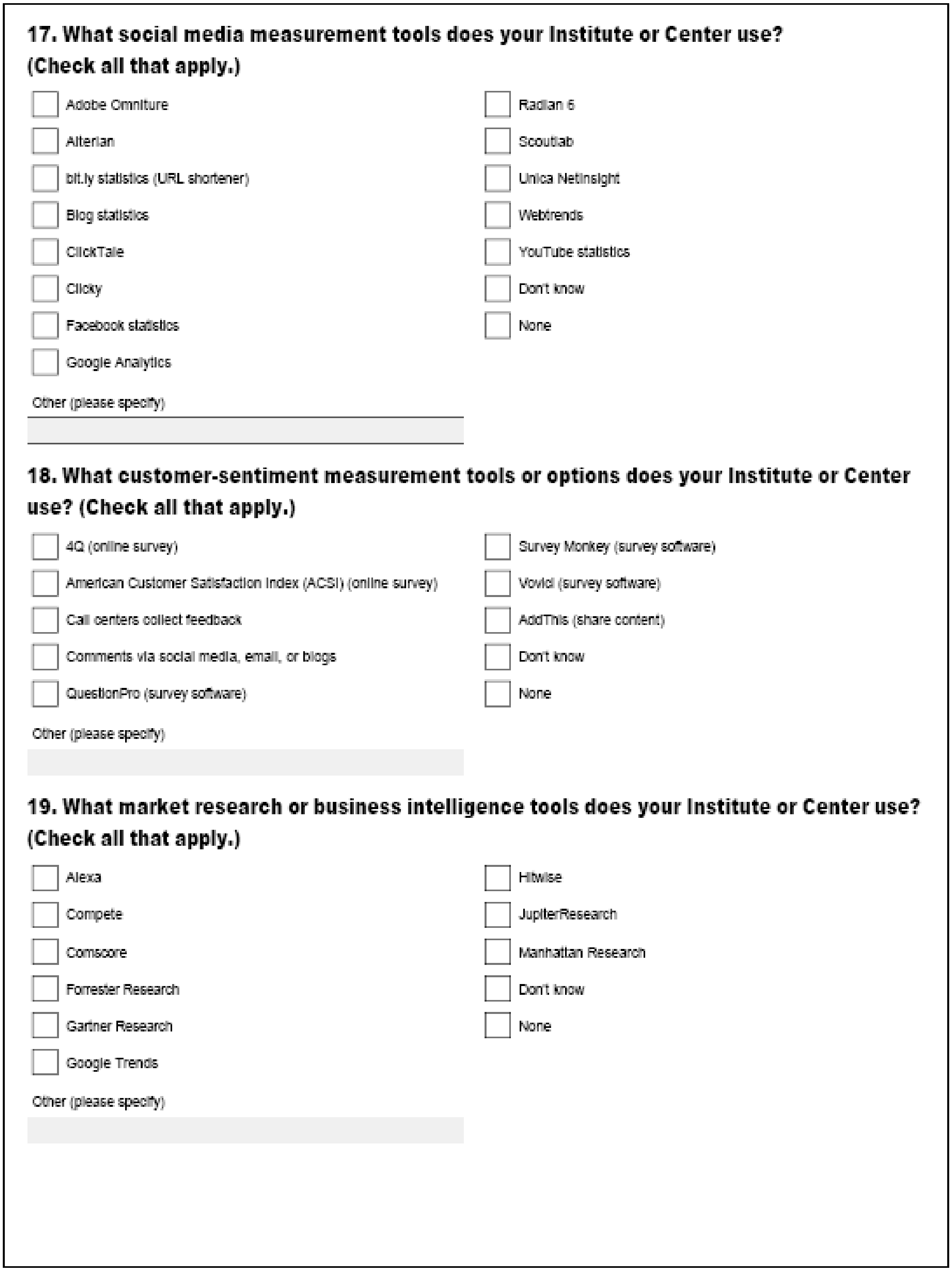
Other (please specify)

**16. What search analytics tool(s) does your Institute or Center use? (Check all that apply.)**

☐ Adobe Omniture (Web analytics software)
☐ ClickTale (customer experience analytics)
☐ Clicky (Web analytics service)
☐ Google Analytics (Web analytics software)
☐ Google Insights for Search (Shows popular search terms and graphic depictions)
☐ Search engine reports
☐ Unica NetInsight (Web analytics software)
☐ Webtrends (Web analytics software)
☐ Don't know
☐ None

Other (please specify)

# The Developing Web Analytics Best Practices for NIH Survey, 2011
(continued)



**17. What social media measurement tools does your Institute or Center use? (Check all that apply.)**

☐ Adobe Omniture
☐ Alterian
☐ bit.ly statistics (URL shortener)
☐ Blog statistics
☐ ClickTale
☐ Clicky
☐ Facebook statistics
☐ Google Analytics
☐ Radian 6
☐ Scoutlab
☐ Unica NetInsight
☐ Webtrends
☐ YouTube statistics
☐ Don't know
☐ None

Other (please specify)

**18. What customer-sentiment measurement tools or options does your Institute or Center use? (Check all that apply.)**

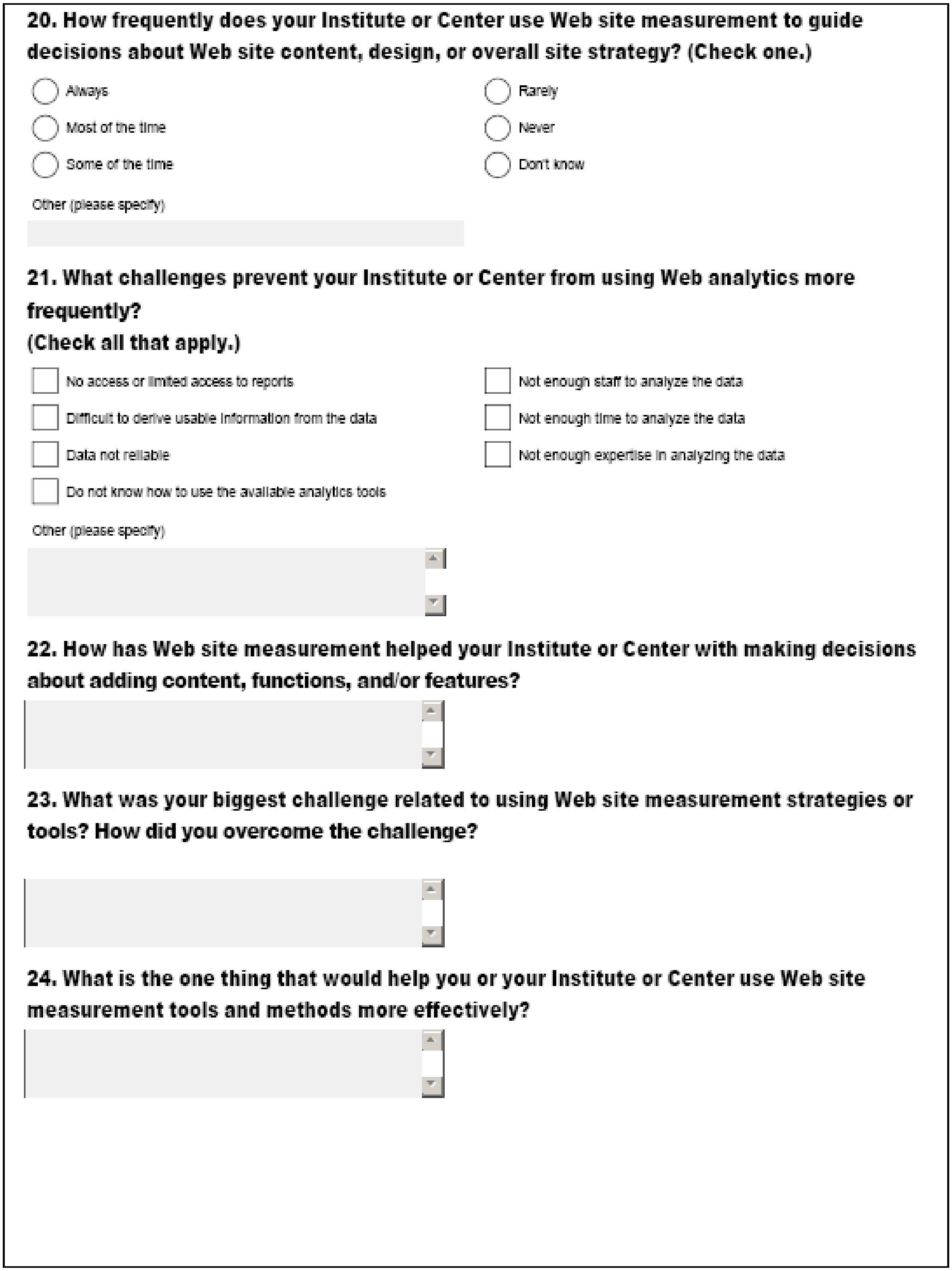
☐ 4Q (online survey)
☐ American Customer Satisfaction Index (ACSI) (online survey)
☐ Call centers collect feedback
☐ Comments via social media, email, or blogs
☐ QuestionPro (survey software)
☐ Survey Monkey (survey software)
☐ Vovici (survey software)
☐ AddThis (share content)
☐ Don't know
☐ None

Other (please specify)

**19. What market research or business intelligence tools does your Institute or Center use? (Check all that apply.)**

☐ Alexa
☐ Compete
☐ Comscore
☐ Forrester Research
☐ Gartner Research
☐ Google Trends
☐ Hitwise
☐ JupiterResearch
☐ Manhattan Research
☐ Don't know
☐ None

Other (please specify)



**20. How frequently does your Institute or Center use Web site measurement to guide decisions about Web site content, design, or overall site strategy? (Check one.)**

○ Always
○ Most of the time
○ Some of the time
○ Rarely
○ Never
○ Don't know

Other (please specify)

**21. What challenges prevent your Institute or Center from using Web analytics more frequently?**
**(Check all that apply.)**

☐ No access or limited access to reports
☐ Difficult to derive usable information from the data
☐ Data not reliable
☐ Do not know how to use the available analytics tools
☐ Not enough staff to analyze the data
☐ Not enough time to analyze the data
☐ Not enough expertise in analyzing the data

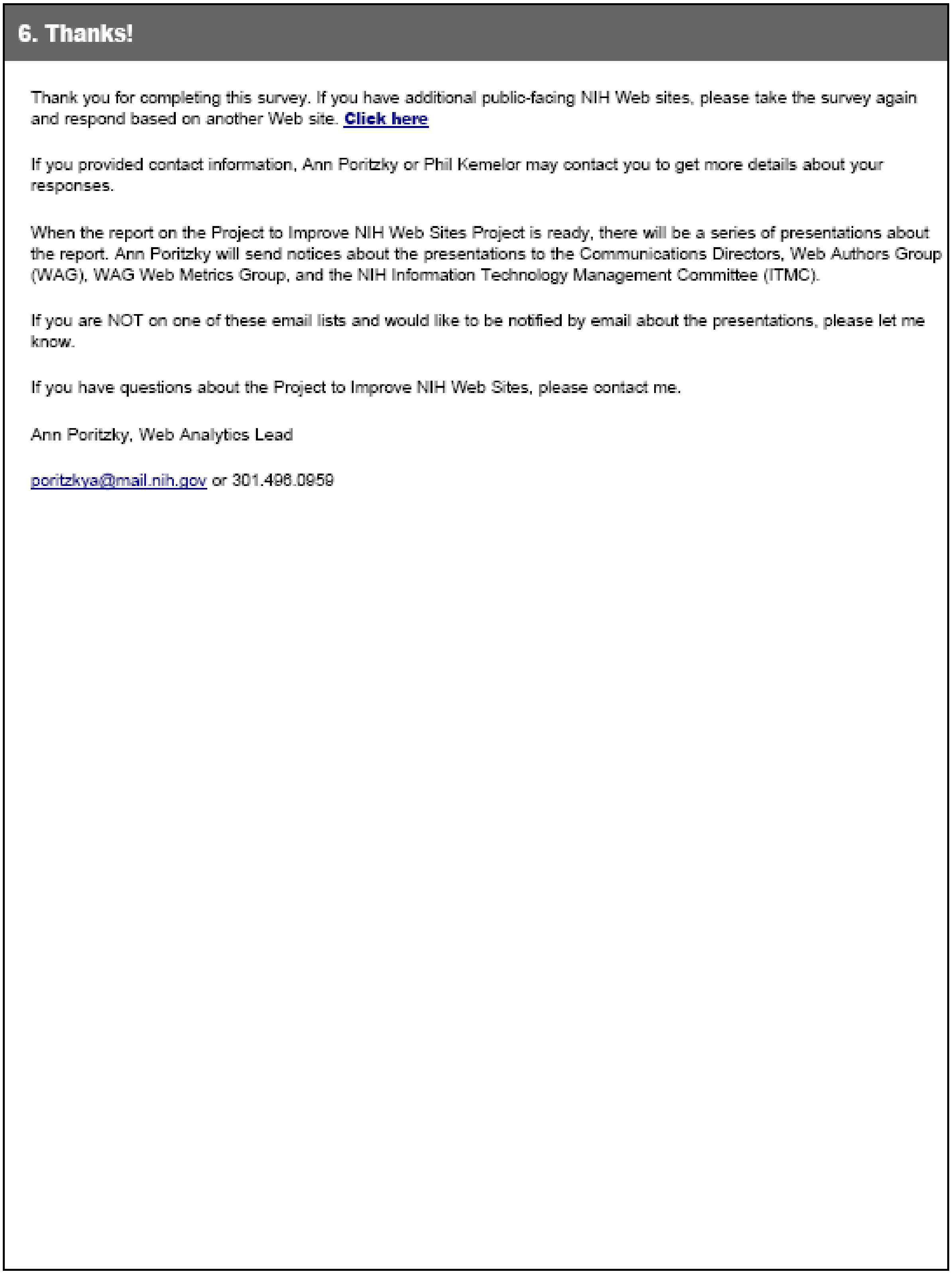
Other (please specify)

**22. How has Web site measurement helped your Institute or Center with making decisions about adding content, functions, and/or features?**

**23. What was your biggest challenge related to using Web site measurement strategies or tools? How did you overcome the challenge?**

**24. What is the one thing that would help you or your Institute or Center use Web site measurement tools and methods more effectively?**

# The Developing Web Analytics Best Practices for NIH Survey, 2011
(continued)



## 6. Thanks!

Thank you for completing this survey. If you have additional public-facing NIH Web sites, please take the survey again and respond based on another Web site. **Click here**

If you provided contact information, Ann Poritzky or Phil Kemelor may contact you to get more details about your responses.

When the report on the Project to Improve NIH Web Sites Project is ready, there will be a series of presentations about the report. Ann Poritzky will send notices about the presentations to the Communications Directors, Web Authors Group (WAG), WAG Web Metrics Group, and the NIH Information Technology Management Committee (ITMC).

If you are NOT on one of these email lists and would like to be notified by email about the presentations, please let me know.

If you have questions about the Project to Improve NIH Web Sites, please contact me.

Ann Poritzky, Web Analytics Lead

poritzkya@mail.nih.gov or 301.496.0959

# Appendix B.
# NIH Web Analytics Requirements Summary

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com

December 15 , 2011

# Appendix B. NIH Web Analytics Requirements Summary

## Introduction

The Developing Web Analytics Best Practices for NIH Project began with a comprehensive discovery process. We collected information from stakeholders and staff about Web analytics requirements and grouped them into several categories, including management commitment, organizational resources, governance process, Web metrics and analysis, training and analysis, and solutions that will enable the use of digital analytics both within the NIH ICs and across NIH.

This document summarizes the findings from the discovery process and the resulting stakeholder and staff requirements. The findings came from two primary sources:

1. **Stakeholder Interviews:** Semphonic conducted 29 interviews and meetings with 76 NIH staff invited to participate by the Office of Director's Office of Communications Public Liaison. Staff from the OD and 16 ICs participated in meetings. (See **Figure B.1** for a complete list of participating institutes, centers, and offices.)

2. **Online Survey:** We created an online survey and made it available to key staff involved in current Web initiatives throughout NIH and those familiar with their IC's digital analytics activities. The survey received 131 responses.

***Figure B.1.*** *NIH interviewees' Institutes, Centers, and Offices.*

| |
|---|
| Center for Information Technology |
| National Cancer Institute |
| National Center for Complementary and Alternative Medicine |
| National Center for Research Resources |
| National Human Genome Research Institute |
| National Institute on Drug Abuse |
| National Institute of Environmental Health Sciences |
| National Institute of Allergy and Infectious Diseases |
| National Institute of Arthritis and Musculoskeletal and Skin Diseases |
| National Institute on Deafness and Other Communication Disorders |
| National Institute of Dental and Craniofacial Research |
| National Institute of Mental Health |
| National Institute on Minority Health and Health Disparities |
| National Institute of Neurological Disorders and Stroke |
| National Library of Medicine |
| NIH Clinical Center |
| Office of the Director |

We organized this document by the categories of requirements we discovered. We will also use the categories in the final Needs Assessment and Best Practices for Trans-NIH Web Analytics report.

## Summary of Requirements

### *Commitment of Senior Management*

The use of Web analytics throughout organizations such as NIH succeeds when senior executives are committed to using analytics data to drive online strategies. We can estimate the level of commitment by looking at budgets, communication about analytics throughout the organization, and the direction that senior management gives staff about the use and value of analytics**.**

In general, we found that many who work on IC Web sites would welcome a stronger commitment to the Web from Dr. Collins as well as from their own IC directors. Addressing the following findings and requirements would help NIH use Web analytics to drive online strategy and tactics.

1. **Finding:** Many senior managers view the Web as a “feel good” effort that receives “love” but not enough financial support.

   **Requirement:** Help senior managers understand the need for their support so staff can develop goals, objectives, and tactics that map to an NIH Web strategy.

   **Requirement:** Offer senior managers opportunities to provide leadership among government agencies and research organizations in the use of Web analytics and support for state-of-art analytics program at NIH.

2. **Finding:** The NIH mission focuses on research, but most of the ICs focus their Web sites on providing information to the public. This dichotomy may affect the support for Web site initiatives and the analytics that could support those initiatives because some senior managers may not consider the public information focus of the Web sites as critically linked to the NIH mission. Originally developed for researchers and grantees, most visible NIH Web sites are now more targeted to delivering public health public health information, information for health professionals and conducting outreach activities.

   **Requirement:** Help NIH managers understand how the NIH mission guides Web site strategy and can help achieve agency-wide goals.

   **Requirement:** Work with senior managers to develop a clear Web strategy supported by senior management that focuses on governance, branding, operations, and funding mechanisms.

   **Requirement:** Help all levels of management understand the importance of having a Web strategy and recognize that the Web plays an essential role in accomplishing the NIH mission as a leading health and medical research institution.

3. **Finding:** When it comes to the relative effectiveness of Web sites, Web analytics and other online measurement practices many do not have as much influence on senior management as opinions held by friends, acquaintances, colleagues, or the executives in charge of the ICs.

   **Requirement:** Help senior-level managers to understand the importance of demonstrating their support for using online measurement, such as Web analytics, in decision-making rather than using only their intuition and experience.

4. **Finding:** Some senior managers believe that the most important “metric” for Web sites is how much content can be published, rather than how visitors are using the content and accomplishing tasks that indicate successful use of the site.

   **Requirement:** Make a compelling business case to senior managers that Web analytics and other digital analytics software solutions and practices are worth supporting financially

   **Requirement:** Encourage senior managers to establish consistent funding for Web analytics and other digital analytic s software solutions and program management support**.**

5. **Finding:** At NIH Web analytics is based on a range of IC-specific efforts that may or may not have senior.

   **Requirement:** Web analytics direction and initiatives must come from the top of the organization rather than the middle, as is currently the case.

6. **Finding:** Budgets for the Web tend to be flat, and budget allocation is usually tied only to the previous year’s allocations.

   **Requirement:** Encourage senior managers to establish consistent funding to support an NIH-wide analytics approach to training and services, including a planning process that outlines goals and objectives for the Web and how to measure them.

7. **Finding:** There is no apparent NIH-wide vision of how to use measurement data to create Web products that advance the NIH mission and serve visitors.

   **Requirement:** Create a senior management council or entity dedicated to researching how to use Web data to create Web products that advance the NIH mission and serve the visitors.

   **Requirement:** Educate senior management about the difference between using Web traffic statistics as evidence of successful outreach to public audiences and using Web analytics data to aid decision-making about achieving NIH mission goals.

8. **Finding:** The Center for Information Technology (CIT) currently provides minimal support to dozens of ICs, and this support is generally considered subpar (see the next section, “Organizational …”). CIT must charge for all services because the services are not a budget line item. This creates an environment where CIT charges noncompetitive prices, making it more viable for ICs to seek outside support from contractors. This system does not work. The Office

of Information Technology (OIT) and CIT perform similar services, but OIT provides them to a smaller number of Web sites. While it appears that OIT is more engaged and more interested than CIT in supporting its stakeholders' Web analytics efforts, the same funding framework exists for both organizations.

**Requirement:** Encourage senior managers to define the scopes of CIT and OIT services and specify how to fund them.

9. **Finding:** Required approvals from OMB have made it very difficult to collect feedback or survey Web site visitors. In 2011, OMB issued new rules that allow for a "fast track" approval process.

   **Requirement:** Once the new fast track processes are in place, ICs should take full advantage of them to survey and collect other types of feedback from their Web site visitors.

**Relevant Comments on Commitment of Senior Management**

- "Our IC director feels that the Web site is the most important communication tool. [The director] supports it and promotes it, but doesn't have any idea what goes into the site or how to direct its development."
- "We're told by management to have a Web site, but there is no direction; it's just checking a box. People started throwing content up on the site before there was a communications strategy."
- "There is a constant dialogue about whether the information on the site is current and appropriate. Web metrics aren't a top priority; the top metric that we're measured against is just getting the information 'out there.'"

### *Organizational Resources and Responsibilities*

Having people trained and available to create and interpret Web analytics data is generally an understood requirement. However, often misunderstood are the software, hardware, and human resources required to support Web analytics reporting and analysis. These issues play out in ICs of all sizes.

1. **Finding:** Most ICs do not have dedicated staff or contractors to work with digital analytics data.

   **Requirement:** ICs need dedicated, experienced staff analyzing data, verifying data accuracy, and reporting to all staff on a regular basis. Ideally, there should be someone in this role at the NIH and IC levels.

2. **Finding:** Many ICs do not have the budget to consistently support high-quality analytics. Many efforts are sporadic and ad hoc because they are only done when there are leftover budgets or occasional set-aside funds.

   **Requirement:** Create a centrally funded analytics tool set that would enable smaller ICs to benefit from the use of Web analytics. (Accenture Maxamine, the HHS-funded 508-compliance tool, could serve as a model for a centrally funded digital analytics tool set. Accenture Maxamine is available to everyone at HHS at no additional cost.)

3. **Finding:** Stakeholders and staff understand that it is possible to reduce costs by eliminating redundant analytics tool sets and moving toward more enterprise-level license models.

   **Requirement:** Work to secure and ensure the use of tool sets and shared expertise to get the best value for the analytics-related budget investments throughout NIH.

4. **Finding:** There is an overall sense of frustration among ICs that rely on Web analytics from central NIH IT organizations such as OIT and CIT. (See previous section, "Commitment ….")

   **Requirement:** These IT organizations should provide service-level agreements that guarantee reliable and skilled Webtrends and other analytics support.

5. **Finding:** The nature of NIH — and much of the government — can foster isolation and siloing of information, which hinders sharing wisdom and the development of best analytics-related practices.

   **Requirement:** Increase opportunities for sharing best practices and lessons learned among ICs and other government agencies.

6. **Finding:** The OMB memos released in June 2010 allowed and provided guidance on using persistent cookies, but stakeholders and staff are not clear about what that means for them yet.

   **Requirement:** Clarify how the new OMB guidance (OMB-10-22) policies should be implemented to take advantage of the additional flexibility. Specifically, explain how to change privacy policies and help staff and stakeholders access to Web analytics data, archiving, and retention policies.

**Relevant Comments on Organizational Resources and Responsibilities**

- "An NIH Service Center model would be tremendously helpful for all things related to analytics and third-party software, including social media. Since we are a very small office, it would be nice if we could contact someone at NIH who would actually process all the documents necessary for us to implement any new tool/software. For example, we would like to use Google Analytics, Facebook, etc. It would be wonderful if we could make a request and have an outside person at NIH take care of drafting our privacy policy, getting approval for third-party software, ensuring we are in compliance with all requirements, and guiding us through implementation.
- "We just don't have the expertise or time to implement many of the things that would be beneficial to us. The WAG and Web Metrics Workgroup are helpful, but we still lack the time and technical expertise in-house to fully understand the vast amount of information coming out of these groups on a daily basis. A service center would be a real plus."

- “It would be nice to have some centrally provided measurement tools. Currently, we have to fend for ourselves in a sea of expensive products that we don't understand to get the most value.”
- “We need to add interactivity to the site, get into social media, as well as deal with the care and feeding of existing site content. This competes for time that would otherwise be spent on analytics. We haven't overcome this challenge yet.”

### *Process for NIH-Wide and IC-Specific Web Analytics Governance*

Successful analytics initiatives are dependent on a governance framework that includes systems for data collection, processing, and integration; metrics development and presentation; data analysis; and solution maintenance.

1. **Finding:** In most NIH ICs, one person or a small group of people uses Web analytics.

   **Requirement:** Create processes to enable interested staff to have access to the analytics tools and any training they need to use them.

2. **Finding:** The current NIH and the IC procurement processes make it difficult to purchase software and services for analytics except evaluation set-aside or ad hoc funding mechanisms.

   **Requirement:** Encourage the development of shared NIH procurement for analytics tools and support.

3. **Finding:** There is great interest in what ICs are doing about Web analytics and other digital analytics practices. Some believe that there is an atmosphere that encourages NIH staff to "silo,” rather than share information about Web sites.

   **Requirement:** Evaluate and promote best practices from individual ICs throughout NIH by using an organized and structured dissemination process.

4. **Finding:** Using Web analytics is often underemphasized because for most ICs, analytics tasks are not assigned to anyone and no one is held responsible for them.

   **Requirement:** Clearly define the Web analytics roles and responsibilities for staff throughout NIH and within every IC. These analytics roles and responsibilities can be part of job descriptions and performance reviews.

5. **Finding:** Several stakeholders and Staff reported that it is not clear to staff how to use Web analytics to improve content, the experience of site visitors, or marketing.

   **Requirement:** Improve the process for integrating what is learned through Web measurement into content development and site-maintenance activities.

**Requirement:** Use analytics to move discussions from "your opinion vs. my opinion" to what data show works best for the site visitors.

6. **Finding:** The process for receiving assistance from CIT is unclear, as is system CIT's accountability.

   **Requirement:** Create formal service-level agreements with CIT.

   **Requirement:** Strive to improve communication between ICs and CIT so that the necessary level of analytics reporting and tool customization is available for making decisions about Web sites supported by CIT.

7. **Finding:** Historically, the distribution of digital analytics data has been closely held by a few people in ICs. This is probably the result of the lack of access to analytics tools and a lack of recognition of the value of analytics data.

   **Requirement:** Develop self-service access to analytics for content editors and outreach managers within the ICs, as well as those who serve as Web analysts and Web analytics managers.

   **Requirement:** Establish a governance model that outlines best practices and processes for supporting the consistent development and use of Web analytics across all of NIH.

### *Web Analytics Metrics and Deep-Dive Analysis*

Effective measurement relies on providing data and analysis required for driving strategic and tactical decision-making, and presenting the data in ways that can be used by specific users. These users may be senior managers, marketers, designers, or content editors. We talk about Web analytics measurement in terms of metrics and analysis, but these terms have multiple meanings. At NIH, there are many different views on what constitutes valuable measurements for assessing Web content and communications.

Compared with the .com and .org sectors, government Web analytics are still quite rudimentary, especially with continuing reliance on traffic-based metrics. On the other hand, some .gov Web teams use many analytics tools and methods that are more flexible their .com and .org counterparts are.

1. **Finding:** The level of mistrust about the Web analytics data coming from Webtrends is high. People want reliable data they can derive usable information from.

   **Requirement:** Select, correctly implement, and maintain a scalable Web analytics tool that allows ICs to derive trustworthy information.

   **Requirement:** Provide trained and available staff or contractors who will implement Web analytics solutions consistently and ensure that data collection and processing yield accurate data for Web analytics reporting.

2. **Finding:** The overall lack of up-to-date tools, time, resources, and knowledge about how to implement the tools or analyze the data completely cause the metrics that are considered common in nongovernment organizations to be used inconsistently at NIH.

   **Requirement:** Establish a basic set of metrics that can be used across NIH and the available tools can provide data for.

   These may include metrics and tools for
   - analyzing search terms
   - obtaining data about use of PDFs and downloads
   - measuring effectiveness of navigation and information architecture
   - analyzing campaigns (outreach and marketing initiatives)
   - determining how well content is optimized for search engines
   - keeping track of referrals from other sites
   - geolocating site visitors
   - monitoring browser and operating system usage metrics
   - segmenting audiences by usage, demographics, internal and external visitors
   - defining site success events and ways to measure them
   - streaming audio starts, completes, stops
   - filtering out bots, spiders, images, and non-human traffic
   - monitoring and measuring  of multimedia and social media use

3. **Finding:** NIH ICs use a combination of Web analytics, user-experience testing, focus groups, survey data, market research sources, and industry benchmarking to understand the effectiveness of their Web sites. However, many are not quite sure of the best way to use these data sets to reach conclusions. In addition, the data from each tool are often viewed separately This makes it difficult to understand the relationship between results.

   **Requirement:** Improve the methodology for coordinating the use of multiple analytics methods and tools, which will help guide strategy and tactics for developing content, improving user experiences, and expanding outreach.

   **Requirement:** Work to integrate direct user feedback with Web analytics data and relate them to visitor behaviors.

   **Requirement:** If possible, integrate survey results, such as Foresee Results, with analytics data to see how survey respondents interacted with the site.

4. **Finding:** ICs primarily use historical Web analytics and other digital measurements reports to learn what happened on their Web sites. These data are not often used as business intelligence to guide the development of new content and applications. Too often, Web teams receive non-evidence-based directives from senior management to make site changes or improvements.

   **Requirement:** Use analytics to help create the business case for developing new Web applications, content, and functions.

**Requirement:** Develop a more complete understanding of how Web site objectives link to onsite tactics for Web site content development, design, and outreach so that ICs can use metrics to do the following:

- determine what visitors are looking for and are interested in
- identify top tasks and additional links within the site to better understand how to direct visitors through the site
- define thresholds for content usage to identify popular and little used content and archive infrequently used content
- conduct before-and-after testing on Web site redesigns
- create content for certain important target audiences
- increase the pool of high-quality grant applications from new investigators

5. **Finding:** Although most ICs are aware of the different audiences they serve in the offline world, it is challenging to transfer this awareness to analyzing the use of Web sites.

   **Requirement:** Select and analyze visitor segments to understand how to prepare content and outreach campaigns for selected audiences. These segments may be based on visitors' behavior on the Web site, visitors' geographic location, outreach campaign identifiers, or other site activities.

6. **Finding:** ICs do not know how to compare their sites' performance with that of other ICs, Federal government agencies, and organizations with similar missions in the .edu, .org, and .com sectors.

   **Requirement:** Develop plans to gather top-line Web site traffic statistics, such as page views, visits, email subscribers, downloads, and YouTube traffic and share it among ICs. Benchmark selected Web analytics data among NIH ICs, such as unique visitors and page views.

   **Requirement:** Compare NIH data against other government and comparable sites, such as those at National Science Foundation, Centers for Disease Control and Prevention (CDC), and the Mayo Clinic.

7. **Finding:** Social media is garnering a great deal of interest at NIH, but it is challenging to measure the return on investment (ROI) and success of the campaigns.

   **Requirements:** Develop guidance around standard metrics for social media and ways to evaluate the efficiency and ROI of social media initiatives. NIH ICs need a better understanding of their social media audiences are and what they are expecting to get from their social media Interactions

   Although basic traffic data (page views, visits, etc.) that do not provide insight into whether the Web site is achieving goals or providing a business benefit, these are the reports that most managers see. The common use of these reports is probably the reason they erroneously consider those data important.

**Requirement:** Help managers gain a better understanding of the types of Web metrics that can help them make strategic, evidence-based decisions.

**Requirement:** To learn about what is and isn't working on the site, gather feedback from program officers, call center and other offline staff who respond to calls from Web visitors.

**Relevant Comments on Web Analytics Metrics and Deep-Dive Analysis**

- "We want to have an engagement metric for how well the page is doing and then get buy-in from the content producer or sponsor that the content needs to change. Content editors and program managers are the people who are responsible for the content and are the ones that need to be convinced. We need to come up with measurements that will show them whether the content is effective and assist the mission.
- "Non-human visitors still not filtered out! We still don't have reliable numbers because we don't know how much traffic non-human visitors account for."
- "We get completely unreliable Webtrends reports. The numbers jump all over the place, and they also come months late."
- "Content is driven by the people who are working with our target audiences, but they get little feedback or evaluation and don't use analytics to inform decisions."
- "I'd like to see what is going on with analytics across NIH. This would help provide a more global strategy across the Web rather than Institute specific. It could encourage a pulling together of similar Institutes and encourage the development of unified content across all ICs.

### *Web Analytics Solutions, Methods, and Opportunities*

Web, social media, Internet Industry/Market research and online survey are all driven by software solutions, as are multivariate testing, search marketing, and cross-selling platforms. These solutions have tended to be "owned" by the NIH stakeholder group that uses them most. This individual use may cause NIH to miss opportunities for greater efficiencies of scale and overall return on investment (ROI) through centralized purchasing and distribution.

**Findings:** Many of the previous sections' findings, especially those in "Digital Analytics Metrics …," relate to the state of available Web analytics and digital analytics tools. The following requirements summarize and expand on these previous findings.

**Requirements:** Provide upgrades to Webtrends, as well as additional technical expertise to modify Webtrends to yield data that are more accurate. (For example, eliminate spiders, bots, and images,) and provide customized reports.

1. Find the right tools to measure social media.

2. Work on acquiring an enterprise license for Web analytics along with training to ensure uniform measurement of sites at NIH. The tools should be easy to use and allow for filtering of spider data.

3. Consider using a tool set for presenting analytics data, such as Cognos or Tableau, and provide guidance on when it is best to use each one.

4. Work on acquiring an array of tools for analytics, mobile, social, voice of customer (VOC), user experience, presentation, direct surveys, call center, testing, competitive research, and search engine optimization (SEO).

5. Find a lower-priced voice of customer (VOC) tool than ForeSee Results or obtain more favorable pricing for the American Customer Satisfaction Index (ACSI) survey.

### *Web Analytics Training, Data Analysis, and Interpretation of Data*

Web analytics and digital analytics data are challenging to use because the numbers associated with activities do not mean very much to most people. Managers and IC staff want to know how to interpret whether a number of page views is good or bad. They also want to know to use the information to figure out what to do.  This is why developing clear language around Web analytics is essential for the successful adoption and use of analytics throughout any organization. Training staff how to produce analytics reports and providing guidance on how to interpret the data so that they are understandable and useful are the chief areas of focus for digital analytics training and education.

1. **Finding:** There is no NIH-sponsored training on how to use Web analytics tools and data.

   **Requirements:** Develop and offer high-quality training on how to use Web analytics tools, as well as how to read and interpret the data.

2. **Finding:** There is some communication about how ICs are using Web analytics successfully, but there could be more.

   **Requirement:** Promote NIH case studies that demonstrate how Web analytics helped obtain high returns on investment (ROI) for online projects. .

   **Requirement:** Assemble and distribute a best-practices checklist on analyzing data.

   **Requirement:** Provide guidance on how to construct benchmarks, as well as how to convey data in compelling ways.

3. **Finding:** Web analytics reporting that goes to senior level management focuses not on business objectives, but on site traffic that is not very relevant.

   **Requirement:** Prepare Web analytics information so it is easy to explain to senior management. This will help ensure that Web initiatives continue to be funded. Also, create reports that allow managers to easily identify which communications channels are providing the best ROI, such as social, online, or print.

**Requirement:** Present an explanation or interpretation along with all Web analytics reports.

4. **Finding:** Often a few people hold the analytics data very closely. This delays distribution of relevant data to groups that could use the data to make decisions and cause data to become outdated.

   **Requirement:** Develop more dashboards so that nonmetrics staff can have easy access to analytics.

**Relevant Comments on Web Analytics Training and Analysis and Interpretation of Data**

"If there was some sort of best practices checklist I could use to ensure that we are covering all of our bases in terms of info we should use to make decisions, that would be helpful, or NIH staff experts available for in-person consultation to review current practices against best practices and discuss possible options for addressing any discrepancies."

## Next Steps

The diversity of the NIH community is well reflected in the findings and requirements described in this report. It is clear from the survey responses and conversations that no matter what level of analytics maturity ICs have, every IC can improve in at least one of the categories described above. These findings served as a basis for the Developing Web Analytics Best Practices for NIH report.

# Appendix C.
# Process for Collecting Accurate Analytics Data

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix C. Process for Collecting Accurate Analytics Data

## Introduction

Establishing a successful Web Analytics practice requires several things, and none is more important than reliable, high-quality data. Data quality can be affected at virtually every stage of the analytics process, and it starts at the point of data collection.

With this in mind, Semphonic has developed standards for data collection to provide some insight into critical aspects of this first and most important stage of Web measurement. This appendix is broken into three sections: the data-collection process, data-accuracy concepts in Web Analytics, and standard data points to collect.

## Data-Collection and Processing

Before discussing standards for data-collection accuracy, it helps to understand how data collection works and the factors that have an impact on data-collection methods. In this section, we review the different methods and what each can and cannot be used to collect

### *Data-Collection Methods*

There are two primary methods for collecting Web analytic data: JavaScript page tags and Web server logs. Another method known as “packet sniffing”—intercepting visitor clicks on the network itself—is not commonly used. Some vendors offer hybrid solutions that combine log-file and page-tag data collection.

Page tagging is the most popular method for data collection among Web vendors and is almost ubiquitous among Software as a Service (SaaS) solutions.

Web server log-file-based analytics solutions are offered by a smaller number of vendors and are more typically sold as traditional software hosted on sites rather than remotely by a vendor.

Let us contrast these two approaches more closely.

- **Web server Log Data Collection and Processing**

  A Web server records information about the requests it receives and the data it sends back to the requesting computer. This information is recorded in a log file. The log file is a text file with a continuous set of entries. The data can populate site-analysis reports, but typically, the analytics software will first import log files into a database. Most of the NIH ICs using Webtrends host the software and use log-file-based data collection. The process generally works as follows:

  1. Log files contain visit-request data received by the Web servers.

  2. The log files are batch-loaded and imported into a pre-processing server on a scheduled basis, typically ranging from hourly to daily.

3. The processed log files are imported and parsed into an analytics database.

4. Reports generated from the data are aggregated into summary tables or data cubes. Both methods provide a way for data to be available for reports that are often referred to as "out of the box."

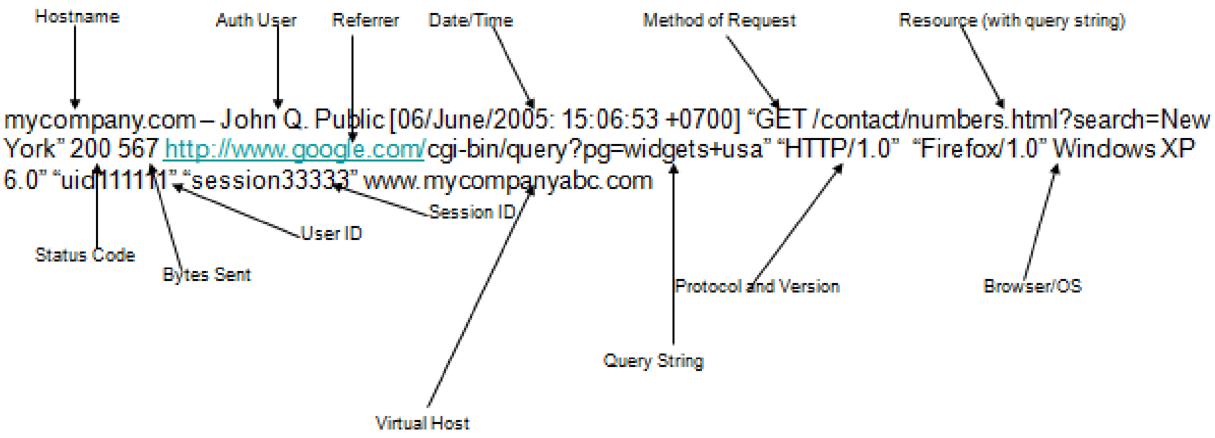
    a. Users may query summary table-based reports through available filters and/or predefined query parameters.

    b. Users may query nonaggregated data tables through available filters and/or predefined query parameters.

Server log-files can contain a vast amount of information about visits, sites, and server performance. Some of the information is useful for understanding site visits, such as what pages are being viewed and where visitors come from. Some of the information is useful for understanding how the Web server is functioning—that is, how the server is responding to requests.

Servers may use different field names for recording the same types of information and may be configured to capture different types of data. In other words, some log files may have more data than others do. **Figure C.1** shows the most fields that a Web server can populate with data.

***Figure C.1.*** *Data collected in a Web server log.*



```
mycompany.com – John Q. Public [06/June/2005: 15:06:53 +0700] "GET /contact/numbers.html?search=New
York" 200 567 http://www.google.com/cgi-bin/query?pg=widgets+usa" "HTTP/1.0"  "Firefox/1.0" Windows XP
6.0" "uid11111" "session33333" www.mycompanyabc.com
```

Most servers use similar terminology for defining log-file information. **Figure C.2** summarizes the data traditionally collected in a Web server log file.

***Figure C.2.*** *Log-file data terminology.*

| Information Collected in the Log File | It Tells You |
|---|---|
| Hostname and Internet protocol (IP) address | The Internet service provider (ISP) or organization where the visitor originates. The country where the ISP or organization is located. |
| Date and time | When the visitor viewed the pages or documents on the Web site. |
| Auth user (authenticated name of user) | The name of the visitor. You will *not* see this unless the visitor fills out a form that is not secure. This occurred more frequently in the early days of the Internet, but is very rare now, with today's focus on security and privacy. For security purposes, visitors' credentials are typically passed through the Web server to a database. |
| Resource | The URL of the page being requested. |
| Browser or operating system | The browser and operating system being used on the visitor's computer. |
| Method of request | Describes how the server retrieves Web pages. |
| Protocol and version | The standard format used to transfer information between the visitor's computer and the server. |
| Status code | How the server responded to the request: was it successful or was an error recorded? (Note that page-tag-based data-collection systems typically do not gather this information.) |
| Bytes sent | The size of the file sent by the server in response to the request. |
| Referrer | From where on the Internet the visitor linked to the page on your site. |
| User ID (cookie) | Unique users (computers) that visited the site. |
| Query string | For dynamic (that is, database-generated) content, this entails additional information passed in the URL. In some cases, query strings also hold search terms that are put into site search engines. |
| Session ID | A code for a unique visit that occurs on the site. |
| Virtual host | The name of your Web site (that is, the site from which data are being requested). |

- **JavaScript Page Tags**

  Page tags collect data directly from the site visitor's computer. This information is sent back to an analytics server where it is either

  - processed, compiled into log files, and imported into a database or
  - processed and imported into a database directly.

If your analytics software creates log files, they tend to be smaller than Web server log files because they contain less data.

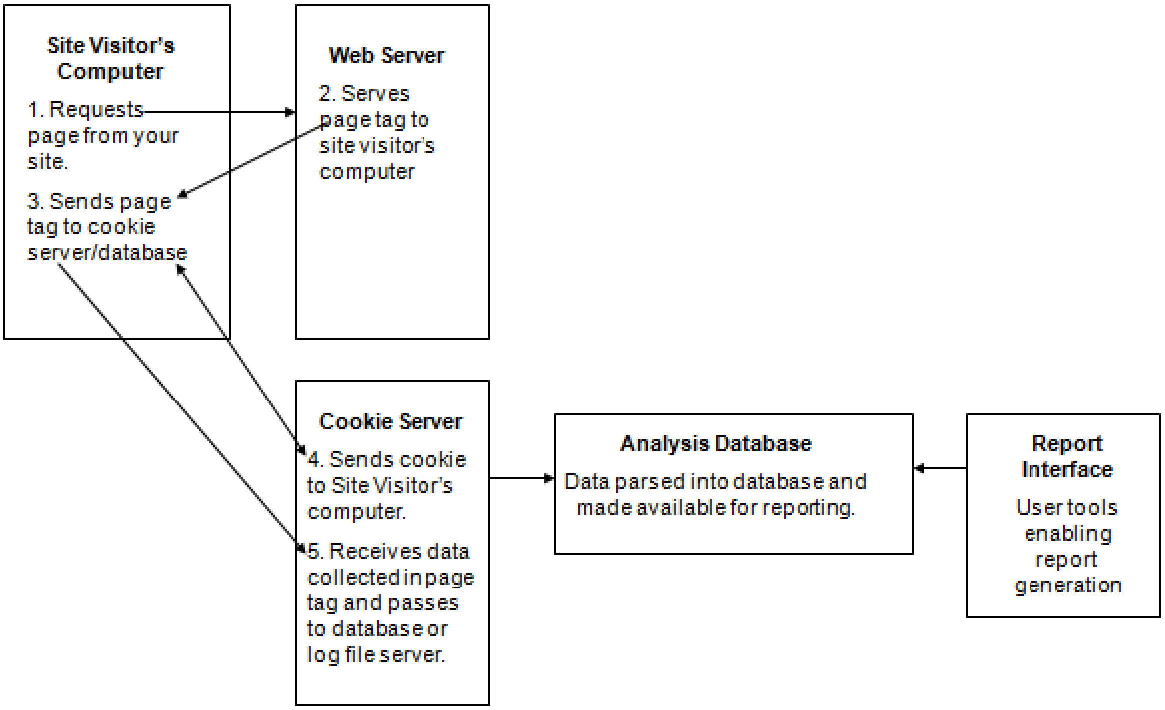
In general, JavaScript page-tag data collection and reporting occurs as follows (which also describes how first-party cookies work):

1. Visitor's computer requests page from Web server (www.yourcompany.com).
2. Web server delivers the page, which contains some embedded tags that contain JavaScript code.
3. The JavaScript page tag collects data from both the browser and the Web page while the page loads. Data may include the URL, type of browser, the IP address from where the visit originates, and potentially other data from the page itself.
4. Data collected by page tags are passed as a long query string to the Web analytics server via a 1x1-pixel invisible image.
5. The Web analytics server receives the data and sends a persistent first-party cookie to the visitor's computer. To ensure first-party-cookie status, a common practice is to have your organization create a subdomain for the analytics server in your domain name server (DNS), such as "analyticsserver.yourcompany.com." On subsequent visits, the visitor information contained in the page tag is always matched to the visitor's computer cookie. This enables the computer to "recognize" visitors' computers if they have visited the site before and not deleted the first-party cookie.
6. Analytics data are preprocessed into a log file that is parsed into an analytics database, or data are preprocessed and parsed directly into analytics database. This occurs in real time or near real time.
7. Reports generated from the data are aggregated into summary tables or data cubes.
    a. Users may query summary-table-based reports through available filters and/or predefined query parameters.
    b. User may query non-aggregated data tables through available filters and/or predefined query parameters

### *Tag-Based Web-Data Collection*

Much of the same information your Web server collects in a log file can also be collected in a page tag. (See **Figure C.3.**) There are notable differences between what a server log and page tags collect.

***Figure C.3.*** *Tag-Based Web-Data Collection*



Site Visitor's Computer
1. Requests page from your site.
3. Sends page tag to cookie server/database
Web Server
2. Serves page tag to site visitor's computer
Cookie Server
4. Sends cookie to Site Visitor's computer.
5. Receives data collected in page tag and passes to database or log file server.
Analysis Database
Data parsed into database and made available for reporting.
Report Interface
User tools enabling report generation

Page tags do not record

- Status, unless status-message pages are tagged. For example, a "Page Not Found – 404 Error" message (and any other server-error pages) must be page tagged in order to be collected.
- "Bytes sent" data. (Bytes indicate the size of the file sent by the server in response to the request.)
- Search engine robots and spider visits (most do not cause the JavaScript tag to collect data).
- Images, so you don't need to filter out images in a page-tag solution as you need to do in a log-file solution.
- Frames, unless you tag the frame. In log-file solutions, you need to filter out the frames to prevent them from getting counted.
- Information from visitors who have turned off JavaScript in their browsers (typically because of security or privacy concerns).
- Pages where visitors leave before the page tag is fired.

To put it simply: Web log files can collect everything that a Web server can record, so you need to filter out what you do not want. Page tags collect only what you want (and JavaScript can capture), so you need to create and place the tags properly to ensure you collect what you need. While log files are far more dependent on an organized and coherent Web site information architecture to create clean data, page tags allow you to create clean data that are not reliant on site structure. Instead, page tags can collect data based on whatever grouping or hierarchy makes sense or fits your analysis objectives.

Note that page tags can collect information from within a Web page and be coded variably for use on a template, site area, or even individual page.

Indeed, page tags allow for an extensive amount of customization to enable you to collect information on campaigns, search engines, Web 2.0 events, and other marketing and commerce activities. Log files do not provide the same level of data-collection flexibility. In addition, page tags typically produce much smaller log files than those generated by Web servers because they do not collect as much information.

See **Figure C.4** for an example of an Adobe Omniture SiteCatalyst page tag from cancer.gov. In this example, you see how the data are categorized into classifications such as s.channel, s.prop, evar, and events. These classifications, generally referred to as “variables,” convey to the Web analytics data processor where in the database the information goes, and then with what reports the data are associated.

***Figure C.4.*** *Example of an Adobe Omniture SiteCatalyst page tag from cancer.gov.*

```
<!-- ***** NCI Web Analytics ***** -->
<script language="JavaScript" type="text/javascript">
<!--
var s_account='nciglobal,ncienglish-all';
-->
</script>
<script language="JavaScript" type="text/javascript"
src="/JS/Omniture/NCIAnalyticsFunctions.js"></script>
<script language="JavaScript" type="text/javascript"
src="/JS/Omniture/s_code.js"></script>
<script language="JavaScript" type="text/javascript">
<!--
s.channel='NCI Home';
s.prop3='/';
s.prop6='Comprehensive Cancer Information';
s.prop8='english';
s.prop10=document.title;
s.prop25='01/01/1980';
s.eVar2='english';
s.events='event1';

var s_code=s.t();
if(s_code)
   document.write(s_code);
-->
</script>
<!-- ***** End NCI Web Analytics ***** -->
```

### *Data Sampling*

Some vendors offer tools that allow data sampling, so you can query and obtain reports against a subset of collected data. Tools that allow sampling may give you a way to control the sample size. For example, sampling up to 50% means that you keep 50% of the data collected and “throw out” the other 50%.

Why sample at all? If you are searching for trends within an enormous data set, sampling enables you to develop a meaningful report more quickly. Some vendors who offer data-mining tools have configured their reporting in this manner. Other vendors may offer sampling when hosting many sites with significant amounts of traffic. This is because querying against the raw data can take a long, long time (leading to wait times of as much as three to four days), so vendors offer sampling as a kind of shortcut for you and them. Of course, you can often pay more for your own private data warehouse (or you may host the software yourself), but even then, querying is time-consuming and you may find it expedient to sample before you run a query.

### *Visitor Identification*

User, or visitor, identification suggests that there is an opportunity to measure and analyze site usage on a “per-person” basis. Combined with registration, demographic and financial data, you can technically analyze down to the individual level; however, this is very labor intensive and expensive and is not allowed on .gov Web sites.

Identifying characteristics of visitors through the application of user-identification data or studying their site-visit patterns—helps define visitor segments. You may also focus qualitative research on members of these segments to obtain feedback on how to develop Web site design, content, features, and services.

The five methods that use analytics data to define site visitors are

- authentication
- serving a persistent cookie
- serving a session cookie
- capturing Internet protocol (IP) address only
- capturing a combination of IP address, browser, and operating system

#### Authentication

We typically authenticate site visitors by requiring them to have a unique ID and password for logging into a Web site. If visitors register, they may also provide a name, address, and other personal information, as required. Once this information is captured, you can accurately count visitors and, theoretically, map both online and offline data records to get a more complete picture of their behavior. To do this securely, you need to pass the user name and information into an offline database and synchronize the information with browser-activity information captured by the site-measurement tool. Because of the potential for accessing personally identifiable information (PII), authentication cannot be used at NIH.

**Persistent Cookies**

A persistent cookie is a unique string of text that analytics software places on the visitor's computer during the first visit and is then recorded again in subsequent visits. Any visit from that same computer is considered to be from a "unique visitor" or "unique user." If the visitor registers on the site and provides personal information, that information can be associated with the specific computer.

As the name implies, persistent cookies remain on the visitor's computer beyond that first visit to a site. Conversely, session cookies remain on a visitor's computer only for the length of that site visit.

Cookies in and of themselves do not breach a user's privacy (although they can be misused), nor do they pose security risks. They simply enable the tracking of a visitor's behavior on the site from which the cookie was issued. Visitors always remain anonymous unless they register on the site.

Cookies contain a unique identifier text string that typically serves as the primary key for tracking them across visitor sessions, as well as for linking Web visit data with other data sources, such as customer registration information. Setting cookies can be complex if you have a large site with multiple servers. In addition, cookies will not provide data that is 100% accurate. Here is why:

- Some visitors set their browsers not to accept session or persistent cookies.
- Some visitors use antispyware software that inhibits or rejects cookies.
- Some visitors delete cookies as part of their normal computer hygiene.
- The same individual may use a different computer at work and home. (This means that the same person gets two cookies from your site—one from each computer.)
- Several individuals may use home, library, and school computers. (This means that two or more people get one cookie from your site.)

Two types of cookies are generally used in capturing user-ID information: first-party cookies and third-party cookies. First-party cookies are served by the Web server hosting your site; third-party cookies are assigned by another Web server, such as that of an ad-serving network. In earlier days, SaaS vendors serving JavaScript tags used third-party cookies to enable identification of unique visitors. That is, they served the cookies from their own domains, much the way ad-serving systems still do.

With the advent of antispyware tools, the industry found that from 10 to 20 percent of users were deleting or rejecting third-party cookies. Vendors quickly responded by reconfiguring their solutions to use first-party cookies. Today, most vendors can use a first-party, rather than a third-party, cookie.

You will want to prepare your cookie strategies carefully and discuss your goals, objectives, and requirements with your analytics vendor so they will be addressed. You may need the vendor's assistance in implementing the correct cookie strategy for your site.

### Session Cookie

A session cookie is a unique string of text placed on a visitor's computer during every visit that expires after a short time or when visitors close their browsers. Many standard Web applications use session cookies because they enable the customer to enter items in a shopping cart, complete a multipage form, go to another page on the site, and then return to the form, without losing the information.

Again, once the visitor leaves, the cookie expires. If the person returns, the software issues a new session cookie.

Session cookies help you calculate the number of unique sessions. Although this does not provide any insight into visitor behavior over time, it can help you understand behavior on a per-session basis. Session cookies can help you link online and offline attributes to records, such as product IDs, hosts, referrers, and keyword search data for some levels of segmentation analysis.

### IP Address, Browser, and Operating System

This method combines the IP address with information about the user's browser, the browser version, and the operating system (OS). The additional browser, version, and OS information can help further segment users from the same IP address. This information is readily available from site-measurement reporting tools. It will be hard to identify unique users from organizations with standard browser and OS combinations, such as large companies and ISPs, because most of them may have the same IP address. For example, hundreds of users from ABC Company may use Internet Explorer 6.0 on a Windows XP 5.0 OS because that is the standard deployment throughout the organization. If these visitors access the Internet through a proxy server, they may all look like one person in the Web analytics data.

### Internet Protocol Address Alone

Data on every visitor coming from an Internet protocol (IP) address become readily available from site-measurement reporting tools. The concept of an individual user may not really be available if the visitor is browsing from a large enterprise, like America Online, or a large company or via an Internet service provider (ISP). In such cases, all users from that enterprise can be assigned the same IP address, so there is no way to distinguish unique users. Data from large organizations and ISPs can provide gross figures and, in some cases, narrow down the geographic locale of individual users.

Note that for the last two methods, in absence of any cookie data, visitor sessions will "time out" after 30 minutes. If there is no activity after that time, the session will end.

## Data-Accuracy Concepts in Web Analytics

When we speak about data accuracy, what comes to mind most often? It is probably whether your Web analytics software recorded the "right" number of page views, visits, or unique visitors. For this, you should understand the importance of *how* data are captured, because this affects whether the analytics tool is recording all the potential data or not. For example, incomplete page tagging or differences in Web server time-stamp configurations can account for irregularities in data accuracy. In addition, the prevalence of cookie deletion or rejection increases the potential for inaccuracies.

Most organizations simply accept inherent inaccuracies in Web analytics data. Even the most ardent practitioners will admit that data and reports are best used for trending guidance, rather than a “down-to-the-user” tally. While there may be aspects of how analytics tools calculate that you cannot control, there are some data-collection elements that you can control and plan.

In data capture and reporting, you may decide to capture several elements and then report on them together or separately. These elements could include

- internal traffic to your own pages
- search engine spiders and bots
- page redirects
- PDF, Excel, Word, and other binary files
- streaming audio/video
- link clicks
- HTTP error messages
- dynamic, database-generated content
- images
- frames
- Web 2.0 considerations
    - rich Internet applications (for example, Flash, Ajax)
    - podcasts and other content to be accessed offline
    - newsreaders: really simple syndication (RSS) and XML
    - social networks (for example, LinkedIn, Digg, Facebook, Twitter, YouTube)
    - user-generated content (for example, blogs, wikis)

It is essential to standardize *how* the data are being collected.

### *1. Determine a data-collection strategy, implement the strategy, and maintain the data-collection activities.*

It makes sense to develop a strategy for how to collect data and implement it so that reports are considered accurate and then to maintain ongoing data-collection activities. Unfortunately, this aspect is the single most likely one to be missed in accurate Web analytics reporting . The lack of a data collection strategy ultimately causes the most pain and is often the reason many decide that their Web analytics tools are not working.

There are many reasons why these activities do not get adequate attention. Most are related to organizational processes and the lack of clear definitions of roles and responsibilities. Another common issue is that organizations lack the needed technical support for their Web analytics solution’s data-collection processes.

Both page-tag and log-file solutions require carefully planned implementations and ongoing maintenance.

For JavaScript page-tag solutions, the need to make sure that every page on the site is covered by the tracking technology seems obvious, but what is often less obvious is the strategy for efficiently ensuring that the tags are deployed throughout the entire site. Whether server logs or JavaScript page tagging is being used, Semphonic recommends, where possible, putting the code in a global footer. This ensures that the code placement is comprehensive and easy to update.

Inserting the code into a global footer can get complicated when you start tracking applications because you may need to place tags within the code of the application. For example, you may need to place tags within the template code of a Web site's content-management system (CMS). Or the Web application must automatically generate the data to be collected by the tag. You may also need to insert data-collection code on the pages to collect data from forms, search engines, and links. In addition, to obtain all required data, you must pay special attention to tag syntax and placement throughout the site. Some vendors provide scripts to check for tag placement and syntax accuracy in conjunction with their implementation services and ongoing maintenance. In addition, there is a new niche industry offering services for checking for data-collection accuracy. Digital Diagnostics (formerly Maxamine) and ObservePoint are firms that offer those services.

To capture data accurately and completely with Web logs, you will need to design the site with a well-thought-out information architecture and use meta tags and query strings. For example, to ensure that the reports you produce from analyzing related content are coherent, the Web site hierarchy should be reinforced by the directory structure, such as "/sports/baseball/mets/04 22 07.htm," rather than a less distinctive structure, such as"/sports/article/1324325". To capture additional data, such as campaign IDs, search terms, and information from forms, you will need to rely on query strings that are appended to the URLs, such as this: "…/research/survey.htm?Question 1=Yes&Question 2=No&Question 3=Yes." In page-tag solutions, you can capture additional data through the page tags, as well as, in some cases, query strings.

Web analytics tools typically offer the ability to include and exclude data through filtering, which usually occurs during data processing or via filters that you apply once the data move into a database. You will need to set and maintain those filters manually.

People often underestimate the time and resources required for ongoing tag maintenance. You will need to put a process in place to make sure that tags get are updated and applied, as needed, on an ongoing basis.

### *2. Devise a method for removing bot traffic.*

Robotic traffic is a fact of online life. Search bots, in particular, are constantly indexing sites. They can inflate your site's data, particularly the numbers of visits and page views. Luckily, many JavaScript solutions automatically detect and remove bot traffic. Working with log files requires constantly updating a translation file for bots. This has become increasingly difficult over the years, as bots have proliferated exponentially.

### *3. Construct and apply a report-suite hierarchy or strategy.*

When tagging a Web site, it quickly becomes clear that certain collections of content under the "entire site" umbrella require their own comprehensive tracking and reporting. Determining how to treat this content to collect the wealth of data you need can become a major aspect of your Web analytics strategy.

In JavaScript page-tagging solutions, data collected from a site are organized into distinct groupings, often called "profiles" or "report suites." Within each report suite, data are processed up to the full capabilities of the tool. In other words, if the tool de-duplicates visits and visitors throughout the daily, weekly, and monthly periods, these de-duplicated metrics are valid across the entire report suite.

Most of the time, the best practice for a report-suite hierarchy is to have one (global) data bucket that collects *all* data from the site and a second, subsidiary *set* of report suites. Typically, rolling up the data from the subsidiary report suites to the global bucket can be done in two different ways. The first is for the report suite to be a simple roll-up of all of the subsidiary report suites (**Figure C.5.**). In this case, if you add the total traffic of all the subsidiary report suites together, you get the sum of all of the data. There is no de-duplication of the data. In other words, if you visit pages in Section A of the site and your visits are recorded in Report Suite A and then if you cross over and visit pages in Section B, where your activity is recorded in Report Suite B, you have recorded a "visit" and a "unique visitor" in each of the two sections. When these activities are added together into a global roll-up, the data become two visits and two unique visitors. For this reason, a global roll-up is only recommended if the individual report suites are clearly divided and you do not plan to add the traffic data from the separate suites together.

The second option is to have a global report suite. In this scenario, visits to different sections of the site are recorded as one visit by one unique visitor, and overlapping data are de-duplicated (**Figure C.6.**).

In the case of a global roll-up, the global view of data represents a simple aggregation:

***Figure C.5*** *Global roll-up.*

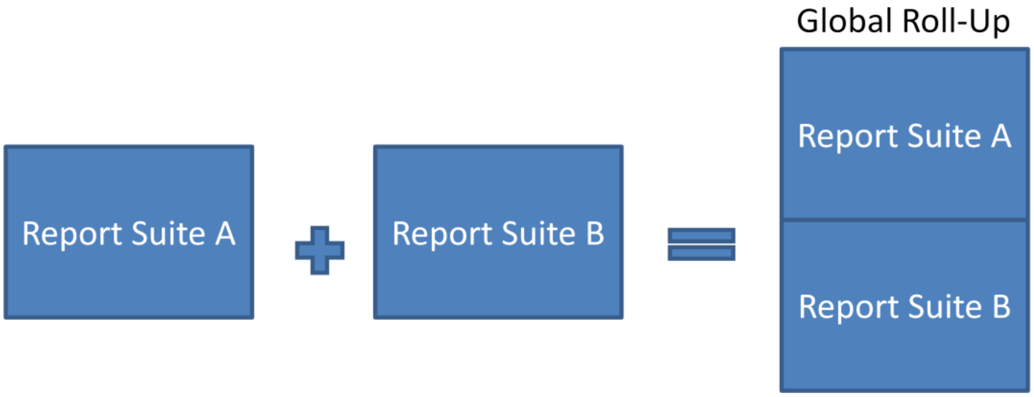
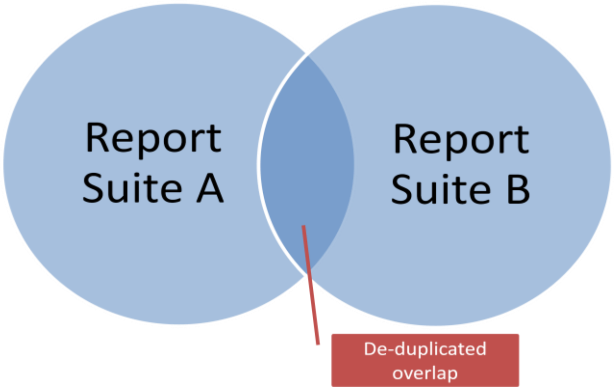


***Figure C.6.*** *Global report suite.*



### *4. Separate external, internal, and development traffic.*

Do you want to include NIH traffic and your IC traffic in your total traffic counts? Do you want to separate them into individual reports?

There are essentially three types of access to your site: external (from visitors outside your organization), internal from individuals in your organization, and internal from Web site developers. All this traffic needs to be reported differently. There is no analytic value to development traffic, though for training purposes, you need to capture the data it generates. Both external and internal traffic generate analytically interesting data, but the use cases are probably distinct so they should be treated differently.

Typically, internal and external traffic data can be separated in a couple of different ways. JavaScript tagging solutions often allow you to store data in different reporting suites or profiles. In both JavaScript solutions and log-file solutions, you can use IP addresses to parse out different types of traffic, assuming that individuals within your organization all fall within an organization-wide allotment of IP addresses.

### *5. If possible, use a first-party cookie.*

When using certain JavaScript tagging solutions, you will need to place a cookie on visitors' machines. There are several different types of cookies, and each has its own limitations. NIH ICs may use a first-party cookie as long as the site's privacy policy complies with OMB circular M-10-22. To see an example of such a policy, review the Privacy Policy statements on www.nih.gov and www.cancer.gov.

### *6. Use the Web-analytic-tool processing rules to "clean" data.*

To obtain "clean" data, particularly in log files, it is absolutely essential to use processing rules within the analytics tool. These rules enable you to exclude unwanted robots and spiders (non-human visitors), frames, images, and traffic from IP addresses, whether anchor tag views are counted as separate page views, and whether file downloads are counted separately or as page views.

## Data Points to Capture

In addition to standardizing how data are captured, certain aspects of user interactions with the Web site are extremely important to capture.

### *1. Devise a way to capture error pages and the URLs that land on them.*

Understanding how often, when, and where users are encountering 404 error pages can help you administer your site and improve the user experience. It is important that you can also determine what page the user was trying to access when the 404 page was served.

### *2. Standardize the character set and currency type used on the site.*

A few different data points are tracked almost universally and are influenced by several variables that can affect their capture. Character set and currency are two such variables. Both should be standardized across the site and data-collection mechanisms.

### *3. Associate a time stamp with all user actions and standardize the time zone being used for all data collection.*

From an analytics standpoint, it is important to associate a time stamp with each user action that takes place on the site. Since the site can be accessed from all over the world, it is also important to devise an approach that ensures that all time stamps are collected in the same way and are easy to interpret.

### *4. Capture standard page characteristics: page name, full URL, and page title.*

To some extent, all Web-measurement strategies revolve around the precise identification of unique pages and the ways visitors interact with them. Precise identification of pages requires a combination of three different data points: a "page name" (a "pretty name" for the page, which may be the full URL or page title), the full URL, and the page title. All three data points (or two, if the page name is the URL or page title) should be captured for every page load.

### *5. Identify sessionization and user IDs.*

Understanding behavior that occurs within the scope of both visits and visitors is crucial to gathering true insights from analytics data. To make this possible, both the session and the visitor involved with each action taken on the site need to be identified.

### *6. Capture the user’s IP address.*

Capturing the user's IP address helps identify unique users (see previous paragraph) and allows geolocation reports to be generated.

## Summary

This appendix reviewed the many different elements to consider when determining the accuracy of Web analytics data. Several dependencies, anomalies, and configurations affect how data are collected, processed, and, ultimately, counted. Please keep these in mind when you are purchasing and implementing an analytics solution, developing reports, and communicating results to your stakeholders.

# Appendix D.
# Success Metrics Calculations

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix D. Success Metrics Calculations

This appendix elaborates on the success metrics section in the Findings and Recommendations report by explaining how each metric is calculated and providing examples.

1. **Task-Completion Effectiveness** – Indicates visitor interest in and successful completion of tasks important to NIH or the IC.

   ***Definition:*** Measures completion of site tasks associated with NIH or IC mission.

   ***Calculation:*** Expressed as a percentage based on the total number of task-completion page views divided by total number of site page views. A higher percentage indicates more successful task completion.

   ***Calculation Notes:*** The completion page for each task will need to be identified, such as a "thank you" page or a submission confirmation.

   ***Example:*** During the month, there were 500,000 total page views. Of this total number, there were 10,000 task-completion page views. This equals a task-completion effectiveness of 2%.

2. **Mission-Critical-Content Effectiveness** – Evaluates visitor interest in content an IC or NIH considers important.

   ***Definition:*** Measures the number of times mission-critical content, including video, is viewed or files are downloaded.

   ***Calculation:*** Expressed as a percentage based on total views or downloads of the 25 content pages of the site that most relate to the NIH or IC mission, divided by the total number of all site content pages (not including portal, search, or navigation-based pages). A higher percentage indicates more successful content effectiveness.

   ***Calculation Notes:*** While it may be somewhat straightforward to define a list of top tasks, it is more challenging to choose what site content is most important. However, this is exactly what NIH ICs must do to fairly evaluate whether sites are meeting their missions through content consumption.

   Determining mission-critical content will encourage ICs to make internal decisions about relative value from a strategy-driven perspective. Making such decisions should create debate that draws attention to the important content rather than gross page-view numbers. This is because pages with the highest page views may not necessarily be the pages related to the IC's mission.

   For example, one site's content that focuses on new grants and research results may be considered important mission-related content, while content that focuses on jobs, training, and public health education may not.

***Example:*** During the month, there were 25,000 total page views and downloads of mission-critical content and 150,000 total page views of site-content pages, including downloads. This equals a mission-critical-content effectiveness of 17%.

3. **Content-Relevance Barometer** – Provides insight into the number of site pages that are not viewed compared page views for the entire site; suggests when to revise, retire, or relocate pages not viewed.

   ***Definition:*** Measures the overall site-content relevance and currency based on viewing of all site-content pages.

   ***Calculation:*** Percentage of all content pages receiving at least 20 page views divided by total number of site-content pages during a specific time. A higher percentage indicates higher content relevance.

   ***Calculation Notes:*** You have to be able to identify content pages. In this case, we mean pages that are primarily presenting information to be consumed, such as health articles, information about grants, and news articles as well as video, audio, and downloadable content. Content pages do not include

   - home pages
   - portal pages
   - search engine results pages
   - site maps
   - pages provided for navigational purposes

   Before we developed this metric, we recommended conducting a functional analysis (see **Appendix F**). This type of analysis sorts out and categorizes Web site pages according to their functions. This approach breaks a Web site into its constituent pieces and then assigns one or more specific functions to each piece.

   These functions can be navigation (e.g., route visitors to a specific place), engagement (grab the visitors' attention), re-direction (move the visitor to a new topic), and motivation (e.g., convince a user to do something) or information (e.g., provide some piece of information). Based on the page functions, each page is assigned a particular page type from a set of common templates that we have developed. Once assigned a page type, the success of the page is measured by the function the page was designed for. It is in the assignment of easily measurable Web page purposes that the power of functionalism resides. See **Appendix F** for a more detailed description of functionalism and how to implement the methodology.

   You also have to be able to identify pages that receive fewer than 20 or the selected number of page views. This requires that a filter be set in the Web analytics tool.

   ***Example:*** There are 30,000 content pages on a site; 1,800 have received fewer less than 20 page views during a month. This results in a content-relevance-barometer score of 94%.

4. **Content-Distribution Score** – Indicates whether visitors believe content is valuable and interesting enough to share; helps spread the IC brand and authority.

   ***Definition:*** Measures whether content holds value and relevance based on visitor opinion.

   ***Calculation:*** Content shares metric based on total number of AddThis (or similar share tracking) usage divided by page views. A higher percentage of content shares indicates that content was distributed more successfully.

   ***Calculation Note:*** If the site is not using AddThis, sharing events will need to be recorded by the Web analytics tool. This will require on-page tagging or use of query-string capture in the log-file-analysis tool.

   ***Example:*** During the month, there were 200,000 total page views and 12,000 AddThis shares. This equals a content-distribution score of 6%.

5. ***Navigation-to-Content Score*** – Indicates whether visitors are spending too much time trying to navigate to content.

   ***Definition:*** Compares number of page views on navigation-oriented pages with content- or application-focused pages.

   ***Calculation:*** Percentage of navigation-type pages viewed compared with all site-page views. This is derived from dividing navigation page views by all page views. A lower percentage indicates a more efficient visitor interaction with the site.

   ***Calculation Note:*** You need to identify navigation pages. This is similar to the exercise described for content relevance barometer above. In this case, we mean pages that are primarily presenting methods for finding content and applications on the site, including
   - home pages
   - portal pages
   - search engines
   - site maps
   - pages provided for navigational purposes

   Before we developed this metric, we recommended conducting a functional analysis (see **Appendix F**). Such analysis sorts out and categorizes Web site pages according to their functions. This approach breaks a Web site into its constituent pieces and then assigns one or more specific functions to each piece.

   These functions can be navigation (e.g., route visitors to a specific place), engagement (grab the visitors' attention), re-direction (move the visitor to a new topic), and motivation (e.g., convince a user to do something) or information (e.g., provide some piece of information). Based on the page functions, each page is assigned a particular page type from a set of common templates that we have developed over time. Once assigned a page type, the success of the page is measured by the function the page was designed for. It is in the assignment of easily measurable Web page purposes that the power of functionalism resides. See **Appendix F** for a more detailed description of functionalism and how to implement the methodology.

***Example:*** There are 40,000 page views of navigation pages out of 150,000 total page views for the month. This equals a navigation-to-content score of 26%.

6. **Mission Success Score** – Evaluates site-visitor engagement based on the amount of mission-critical content read and task completion.

   ***Definition:*** Measures per visit engagement with site's mission-critical content and tasks.

   ***Calculation:*** Page views per visit metric based on total number of page views divided by visits and downloads associated with mission-critical and task-completion metrics. A higher number of task completions and downloads per visit indicates greater mission success.

   ***Example:*** During the month, there were 200,000 total visits and 250,000 page views and downloads of content and task-completion pages and downloads. This equals a mission-success score of 1.25 pages per visit.

7. **Return on Investment** (ROI) – Provides quantifiable cost-benefit analysis of budget and time expended on Web site operations.

   ***Definition:*** Measures ROI in Web site operations, content, applications, and labor relative to the budget spent on these initiatives.

   ***Calculation of Task-Completion ROI:*** Number of task completions multiplied by a dollar value given to each task-completion type compared with the total spent on creating the task application or with the avoided-cost alternative to manual labor.

   ***Calculation Note:*** Total expenditures will need to be determined based on budget allocation, set-aside funds, hourly labor rates, etc. Dollar values for completion of tasks will need to be assigned based on the calculated cost of completing the task offline.

   While it may not be possible to get exact costs, such as a per hour rate for someone's time, you could determine estimates based on GS salary levels for staff members. If they are contractors, you can obtain this information from the contract documents. If neither of these information sources is available, you could default to a level-of-effort estimate based on the number of hours required to complete the production of the content or application associated with the task development. If the Web application or content development is part of a larger effort, such as a Web site section redesign, it may be that costs are all inclusive. Therefore, the total costs are aggregated and the metrics may be calculated on an aggregate basis rather than a per-task-completion basis.

***Example:*** There were 130 submissions of grant applications during the month from the NIH site. Each submission is valued at $200. This amount is based on the time saved by having the applicant enter the data directly into the database, rather than having staff transcribe the data from printed forms. This results in avoided cost or savings of $26,000.

***Another Example:*** Calculate and compare the cost and success rates associated with promoting clinical trials and grants in the offline channels and the online channels to determine which one was most effective at attracting qualified applicants.

# Appendix E.
# Dashboard Content and Sample

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix E. Dashboard Content and Sample

Standardized metrics presented in dashboards to specific audiences are a common method for staying up to date on Web site health in most organizations using Web analytics. We did not find this to be the case at NIH, where they are used in large part as a reference to help drive decisions around significant changes to a site, such as a redesign. While we understand that this may be related to the fact that NIH Web sites are dealing with a great deal of "evergreen" content, or content that does not require frequent updates, we also believe that the constraints facing content managers and Web site managers around resources and technology contribute to their lack of use. In addition, the challenges related to having data that are easy to use and interpret have been a deterrent. This has led to the following situation:

- dashboards have not been designed,
- people supporting Web analytics do not know how to create dashboards,
- metrics to display in dashboards have not been defined, and
- there is an overall lack of understanding about how to use Web analytics to improve Web sites.

We recommend that NIH use dashboards more for presenting data about IC Web sites. This could lead to more up-to-date content, more frequent refreshing of content, and better placement of content within Web sites. We believe that production of dashboards would alleviate three of the most significant barriers to the use of Web analytics, as described by NIH respondents to the project survey:

- not enough staff to analyze the data,
- not enough time to analyze the data, and
- not enough expertise to analyze the data.

A limited number of relevant and consistent metrics presented in a clean graphic format would begin to address these issues. Once established, the metrics' value in evaluating the Web sites could be more universally interpreted, analyzed, and understood.

We recommend four primary dashboards for NIH ICs and NIH.gov:

- Executive Dashboard: For senior management and others who want to see a global view of IC site activity.

- Content Dashboard: For understanding visitor activity with site content in site sections and for using the results for managing content, planning new content and initiatives for developing applications.

- Web Site Optimization Dashboard: For understanding how the use of site navigation and internal search performs for segmented and non-segmented visitors, across the whole site or in specific site sections.

- Outreach Dashboard: For understanding effectiveness of visitor-acquisition campaigns and search engine and referral strategy and tactics.

The metrics associated with the dashboards are presented below and include the following information that may be used to guide dashboard production:

- Title: Dashboard title
- Description: Name of the report or metrics to be used in the dashboard.
- Calculation: How the report or metrics will be created.
- Data Source: This refers to where the data will be coming from to populate the report. In most cases, the source is the Web site data, but may also include data from other sources, such as survey responses, market research, or social media tools.
- Baseline/Benchmarks: Setting targets is of critical importance for evaluating whether initiatives are making progress or require deeper evaluation and analysis. It is early in the process to begin setting arbitrary targets. We recommend that NIH ICs review trends over a three-month period and then start determining reasonable benchmarks and baselines themselves.
- Frequency: How often to produce reports.

This appendix also includes information on approaches for interpreting the meaning of success metrics and a sample dashboard.

## Benchmarking Dashboard

For comparison purposes between and within ICs, benchmarking dashboards should be used for IC-site-wide views and with defined visitor segments. (**Figure E.1.)**

***Figure E.1.*** *Suggestions for benchmarking dashboard contents*

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Content-Traffic-Comparison Summary | Number of visits, unique visitors, page views to specific ICs and content sections as percent of totals to all Web site content | ▪ Site Wide<br>▪ Site Section | Monthly trends/ Percentage increase goals | Monthly, Quarterly, Annually |
| Task-Completion Effectiveness | See Appendix D. Success Metrics Calculations | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Mission-Critical-Content Effectiveness | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Mission-Success Score | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Content-Relevance Barometer | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |

***Figure E.1 continued***

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Content-Distribution Score | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Return on Investment | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |

## Executive Dashboard

To be used by senior management and others who need a global view of NIH IC-site activity, the executive dashboard is based on defined visitor segments. Including reports based on individual site sections can make the dashboard even more useful and relevant. See sample executive dashboard, **Figure E.7.**

***Figure E.2.*** *Suggestions for executive dashboard contents.*

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Content-Traffic-Comparison Summary | Number of visits, unique visitors, page views to specific ICs and content sections as percent of totals to all Web site content | ▪ Site Wide<br>▪ Site Section | Monthly Trends/ Percentage Increase Goals | Monthly, Quarterly, Annually |
| Task-Completion Effectiveness | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Mission-Critical-Content Effectiveness | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Mission-Success Score | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Content-Relevance Barometer | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Content-Distribution Score | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Return on Investment | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Market Position | Page View, Visit, Visitor Comparison | Access to data from ComScore, Hitwise, Compete | Rank compared with other organizations | Monthly, Quarterly, Annually |

**Figure E.2 continued**

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Social Media Referrals | Referrals from social media sites | Data from analytics tool, social media tool, or market research service | Establish target percentage increases | Monthly, Quarterly, Annually |

## Content Dashboard

For understanding visitor activity with site content in site sections and for using the results for managing content and planning new content and application-development initiatives, content dashboards can be used for IC-site reviews and with defined visitor segments (**Figure E.3**).

***Figure E.3.*** *Suggestions for content dashboard contents.*

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Content-Traffic-Comparison Summary | Number of visits, unique visitors, page views to specific content sections and types of content (text vs. audio vs. video) as percent of totals to all Web Site content | ▪ Site Wide<br>▪ Site Section | Monthly Trends/ Percentage Increase Goals | Monthly |
| Content-Relevance Barometer | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Top Articles Read | Top 10: PDF views and Full-Text HTML Split per article and Totals for each article | ▪ Site Section | Per Segment | Monthly |
| Lowest-Ranked Articles Read | Bottom 10: PDF views and Full-Text HTML Split per article and Totals for each article (over 1 page view) | ▪ Site Section | All articles put on evaluation list for removal | Monthly |
| Zero-Page-View Articles | Bottom 25 PDF and HTML Split per article and Totals for each article (0 Page Views) | ▪ Site Section | All articles put on evaluation list for removal | Monthly |
| Content-Distribution Score | See Appendix D | ▪ Site Section | Establish target percentage increases per segment | Monthly, Quarterly, Annually |
| Mission-Critical-Content Effectiveness | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |
| Navigation/Content Index | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage decreases | Monthly, Quarterly, Annually |

**Figure E.3 continued**

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Top Video Views/Top Video Completes | Top 10 videos | ▪ Site Wide<br>▪ Site Section<br>▪ YouTube | NA | Monthly, Quarterly, Annually |
| Top Audio Downloads | Top 10 audio files | ▪ Site Wide<br>▪ Site Section<br>▪ YouTube | NA | Monthly, Quarterly, Annually |
| Top File Downloads | Top 10 file downloads | ▪ Site Wide<br>▪ Site Section | Establish target percentage increases | Monthly, Quarterly, Annually |

## Web Site Optimization Dashboard

This dashboard can be used to understand how site navigation and internal search usage performs for both segmented and nonsegmented visitors, as well as across the whole site or within specific site sections (see **Figure E.4**).

The metrics used here will help inform user experience testing prior to new user interface and information architecture initiatives. The metrics may also be used to provide information that will help refine the tuning of internal search engine performance.

This dashboard is to be used for IC-site review and with defined segments.

In delivering this dashboard to viewers such as content managers, User Interface (UI) designers, information architects, and search specialists, you will need to provide an analysis of the metrics or be able to view the data as a team in order to understand the story behind results.1

***Figure E.4.*** *Suggestions for Web site optimization dashboard contents.*

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency | How to Use |
|---|---|---|---|---|---|
| Search Usage | Number and percentage of searches started/submitted/ click through from results page | ▪ Site Wide<br>▪ Site Section (if site-section-specific search is available) | Establish target percentage increases of click through rates from result pages | Monthly, Quarterly, Annually | Overview on effectiveness of search page to provide relevant results from search queries |
| Advanced Search Usage | Number and percentage of searches started/submitted/ click through from results page | ▪ Site Wide<br>▪ Site Section (if site-section-specific search is available) | Establish target percentage increases of click through rates from result pages | Monthly, Quarterly, Annually | See actions listed for Search Usage. This information will also indicate whether advanced search is used as frequently as basic search. May need to do additional analysis of why advanced search is being used and by whom. |

**Figure E.4 continued**

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency | How to Use |
|---|---|---|---|---|---|
| Internal-Search Keyword Comparison | Top 10 keywords used in Search and Advanced Search | ▪ Site Wide<br>▪ Site Section (if site section specific search is available) | Repeated keywords analyzed to determine successful searches (additional content viewed vs. site exits) | Monthly, Quarterly, Annually | Indicates the keywords visitors are using. Also indicates whether search-related -trends were based on events or "evergreen" site content. May require additional analysis to understand the content selected from certain key words and how visitors interact with the site after reading search results. |
| Site Exits from Navigation Pages | Calculate the number and percentage of exits from navigation pages compared with all page views of that page | Navigation pages | Decrease in percentage across all navigation pages and specific navigation pages | Monthly, Quarterly, Annually | Indicate relative effectiveness of navigation pages in routing visitors to site content |
| Site Bounce Rate | Single page visits | ▪ Site Wide<br>▪ Site Section | Establish target percentage decreases | Monthly, Quarterly, Annually | Indicates that visitors landed on pages that were not of interest to them.<br>Requires additional analyses of visitor activities, such as where they came from (referrals) and where they went after single page visits, to understand the causes for bounce rates. |
| Navigation/ Content Index | See Appendix D | ▪ Site Wide<br>▪ Site Section | Establish target percentage decreases | Monthly, Quarterly, Annually | See Appendix D |

## Outreach Dashboard

This dashboard can help you understand the effectiveness of visitor-acquisition campaigns, search engine, and referral strategy and tactics (**Figure E.5**).

Segmentation: Visitor Type, New Visitors, Repeat Visitors, Organization, Geosegments

***Figure E.5*** *Suggestions for outreach dashboard contents.*

| Description | Calculation | Data Source | Baselines/ Benchmark | Report Frequency |
|---|---|---|---|---|
| Visitor Geolocation | Country breakout of new and returning site visitors | Web site | Month over month | Monthly |
| Acquisition-Source Comparison | Number of referrals and percentage comparison among the following sources as well as page views resulting from referral source from the following:<br>▪ Top 5 search engine referrals<br>▪ Search engine optimization (SEO): Source and keyword, brand and nonbrand – top 5<br>▪ Direct visits<br>▪ Referring sites: NIH, ICs, and external<br>▪ Display ad campaigns<br>▪ Emails: Direct emails and newsletters<br>▪ Twitter<br>▪ Facebook | Web site analysis, social media application programming interfaces (APIs), social media analysis tools | Establish target percentage increases | Monthly |
| Domain-Source Comparison | Based on visits, visitors, and page views to site content from the following sources:<br>▪ Top 5 .org<br>▪ Top 5 .com<br>▪ Top 5 .edu<br>▪ Top 5 .gov | Web site | Repeated sites analyzed as to cause; then goals set for percentage increases | Monthly |
| Referrals to Other NIH ICs | IC-site referrals to all other NIH properties based on number of total site exits to NIH properties compared with all visits | IC and NIH Web sites | Repeated sites analyzed as to cause; then goals set for percentage increases over month | Monthly |

## A Standardized Approach for Interpreting the Meaning of the Success Metrics

Many at NIH feel stymied by interpreting Web analytics data. While creating metrics that relate to business goals is a good first step, making decisions based on the metrics can present challenges. One of the challenges in using dashboards and reports is that there is an expectation that you will get a clear, black-and-white answer that shows you what to do. This does not necessarily happen. In fact, to really understand what happened on a Web site, you need to understand why something happened. This may require additional Web analytics analysis that goes beyond basic reporting and really digs into data by running specific, custom queries in the Web analytics tool. It may also require qualitative analysis based on surveys and user-experience testing.

Figure E.6. reviews the metrics recommended above and provides guidance on actions you can take based on your results.

***Figure E.6.*** *Recommended metrics, analyses and actions.*

| **Metrics** | **Analysis** | **Action** |
|---|---|---|
| Content-Traffic-Comparison Summary | ▪ Validates trends and targets against current outreach initiatives. | ▪ Investigate changes relative to trends and targeted levels, such as changes in<br>–outreach efforts.<br>–site hosting or access.<br>–site information architecture and/or navigation. |
| Task-Completion Effectiveness | ▪ Validates task-completion targets against outreach initiatives | ▪ Conduct funnel or fall-out analysis to see which part of the workflow visitors are leaving the process. This will help you to understand whether there are differences in behavior among visitor segments.<br>▪ Conduct campaign analysis to determine which campaigns are most and least effective at bringing in visitors who complete tasks.<br>▪ Conduct follow-up surveys and satisfaction surveys to understand why visitors did not start or complete tasks.<br>▪ Compare results of surveys with fall-out and campaign analyses to validate task and workflow weaknesses.<br>▪ Use findings to repair task completion processes and validate changes through user experience testing. |
| Mission-Critical-Content Effectiveness | ▪ Evaluates whether visitor interest in content that an IC or NIH considers important is trending upward or meeting goals set by management. | ▪ See actions for Task Completion Effectiveness |
| Content-Relevance Barometer | ▪ Provides insight into number of site pages that are not viewed as compared with page views for the entire site | ▪ Determine whether the pages should be revised, retired, or relocated to areas that are more accessible on the Web site. |
| Content-Distribution Index | ▪ Shows list of content that is most often shared and what methods visitors used to share content. | ▪ Conduct visitor-segment-based analysis to see which visitor segment is sharing particular types of content. Use this information to guide development of additional content that is similar to popular topics. |

**Figure E.6 continued**

| Metrics | Analysis | Action |
|---|---|---|
| Navigation-to-Content Index | ▪ Indicates whether visitors are spending too much time trying to navigate to content. | ▪ Conduct analysis to determine whether specific pages are causing the bottlenecks.<br>▪ Conduct usability tests to determine positive and negative attributes of navigation pages. |
| Mission-Success Score | ▪ Indicates how visitors much use mission-critical content and complete tasks that are important to them | ▪ Conduct follow-up surveys and focus groups to determine from audience segments what tasks and content they want from the site. |
| Return on Investment | ▪ Determine whether the Web site is providing the targeted return on investment (ROI) from a monetary perspective.<br>▪ Determine individual campaign ROI based on the  campaign's ability to bring in high- value visitors. | ▪ Conduct analysis to determine how the visitor segments are engaging with the tasks being evaluated for ROI in order to determine a most-valuable-visitor model. Plan to develop additional tasks that are of interest to this group.<br>▪ Conduct focus group and follow-up surveys to understand tasks and content of high interest to most-valuable visitor segments.<br>▪ Conduct campaign analysis to determine strengths and weaknesses of outreach efforts to drive audiences to ROI-based tasks.<br>▪ Conduct funnel or "fall-out" analysis for onsite processes to see where various types of visitors are leaving the process. This will help you to understand whether there are differences in behavior among visitor segments.<br>▪ Determine whether there may be lower-cost resources to continue work or to begin new work on applications and content related to ROI.<br>▪ Conduct ROI analysis for campaign comparison to determine best returns on acquiring visitors considered to be of highest importance (based on site goals). |

**Figure E.6 continued**

| Metrics | Analysis | Action |
|---|---|---|
| Market Position | ▪ Review relative to competitive Web sites. | ▪ Conduct focus groups and surveys to determine what specific audiences see in other sites that they are not seeing in NIH sites.<br>▪ Determine whether content or applications of interest to these groups are difficult to access or find. To assess difficulties, use Web analytics analyses of internal search, navigation, workflow, and campaign effectiveness. |
| Social Media Referrals | ▪ Provide relative impact of social media outlets in referring visitors to NIH sites compared with other referral methods and agreed-upon targets. | ▪ Assess value and activity of what social media referrals do on the site through a deeper analysis of Web site data based on visitor segmentation. |
| Top Articles Read | ▪ View popular articles to determine whether visitor interest changes over time. | ▪ Conduct content analysis based on visitor segments to understand what content is of most interest to groups during a time of time. Then use the findings to plan additional content initiatives.<br>▪ Determine whether campaigns or initiatives affect the relative popularity and longevity of article popularity; plan campaigns and site navigation accordingly. |
| Lowest-Ranked Articles Read | ▪ View least-read articles | ▪ Conduct analysis to determine reasons for poor readership through additional Web analysis and surveys.<br>Determine whether content strategy should shift into other areas based on analysis results. |
| Zero-Page-View Articles | ▪ View least-read articles | ▪ See actions listed for Lowest-Ranked Articles Read. Determine whether articles should be removed or revised. |
| Top Video Views/Top Video Completes | ▪ See Top Articles Read | ▪ See Top Articles Read |
| Top Audio Downloads | ▪ See Top Articles Read | ▪ See Top Articles Read |
| Top File Downloads | ▪ See Top Articles Read | ▪ See Top Articles Read |

**Figure E.6 continued**

| Metrics | Analysis | Action |
|---|---|---|
| Search Usage | ▪ Overview on effectiveness of search page to provide relevant results from search queries. | ▪ Conduct workflow, funnel, and fall-out analyses to determine whether certain types of searches and searches conducted by certain visitor segments result in more or less completion than others.<br>▪ Conduct satisfaction and follow up surveys to determine usability and search relevance factors that may affect successful search completion. |
| Advanced-Search Usage | ▪ See analysis listed for Search Usage.<br>▪ Will also indicate whether advanced search as frequently as basic search. | ▪ May need to do additional analysis of who is using advanced search and why.<br>▪ See actions for search usage. |
| Internal-Search-Keyword Comparison | ▪ Indicates the keywords visitors are using and trends based on "evergreen" or event-influenced interest. | ▪ May require additional analysis to understand the content visitors select from searches with these key words and how visitors interact with the site after reading the search results. |
| Site Exits from Navigation Pages | ▪ Indicate relative effectiveness of navigation pages in routing visitors to site content. | ▪ Conduct analysis of the specific navigation pages to understand potential trouble spots for certain visitor segments. |
| Site Bounce Rate | ▪ Indicates that visitors landed on pages that were not of interest. | ▪ Requires additional analysis to understand cause for bounce rates, such as review of referrals and where visitors went after the single page visit. |
| Visitor Geolocation | ▪ Used to evaluate visitors' original locations and to determine whether geographically targeted outreach campaigns are meeting objectives. | ▪ Conduct visitor-segmentation analysis to understand relative strength and weakness of site content, tasks, and outreach with geographic-based segments. |
| Domain-Source Comparison | ▪ Similar to Visitor Geolocation but using domain source as qualifier | ▪ Conduct analysis of domains referring visitors to site to understand relative strength and weakness of referring sites among visitors. |

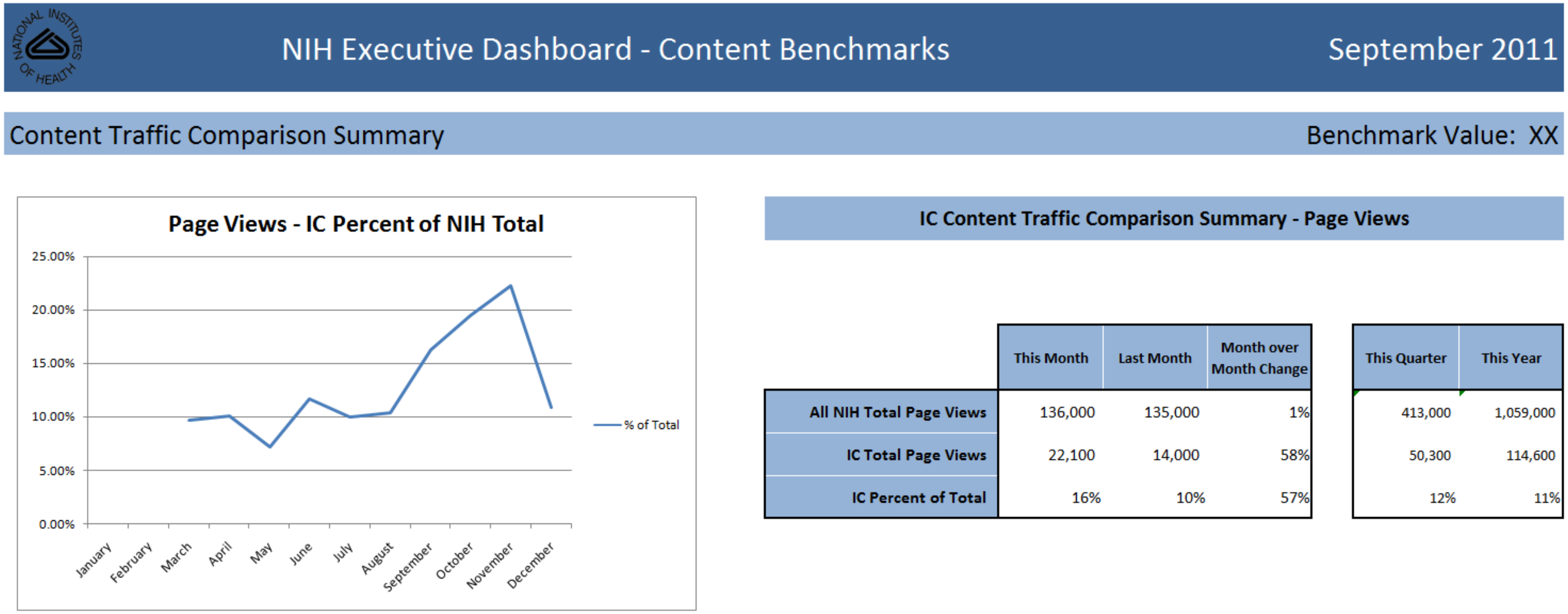
**Figure E.6 continued**

| Metrics | Analysis | Action |
|---|---|---|
| Acquisition-Source Comparison | ▪ Evaluates the performance of outreach sources compared with targets and each other. | ▪ Follow up findings with interviews and focus groups to better understand audience response.<br>▪ Conduct campaign analyses to compare acquisition and task-completion results based on response of different target audiences in order to plan future promotions.<br>▪ Review differences in acquisition rate as influenced by media, time period, content, and design of offer through campaign analysis.<br>▪ Review task-completion factors, such as page content, design, and navigational and transactional element placement on the page through user-experience testing. |
| Referrals to Other NIH ICs | ▪ Evaluate propensity of site visits to result in referral to other NIH sites. | ▪ At site strategy meetings , determine whether the high rate of referrals to other sites was intended.<br>▪ Determine through exit-link analysis whether there are opportunities to create better site experiences for visitors by combining content or tasks from ICs into one site.<br>▪ Conduct follow-up and customer-satisfaction surveys to determine whether use of site as referral to other sites is effective or desired. |

## Sample Executive Dashboard

See **Figure E.7** beginning on next page.

***Figure E.7*** *Sample Executive Dashboard*



NIH Executive Dashboard - Content Benchmarks

September 2011

Content Traffic Comparison Summary

Benchmark Value: XX

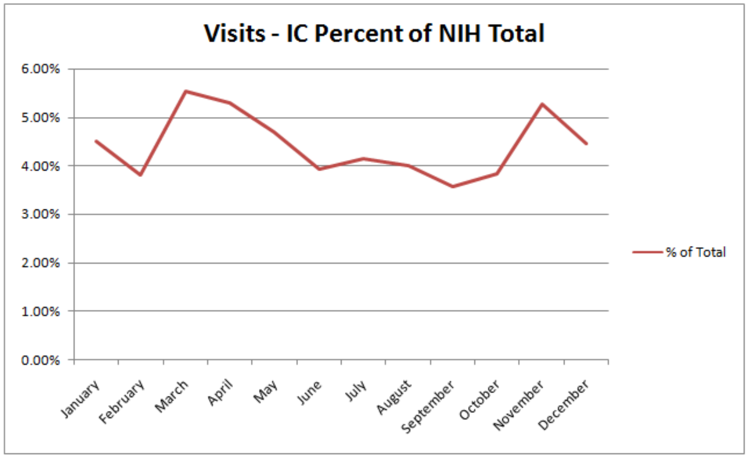
**Page Views - IC Percent of NIH Total**

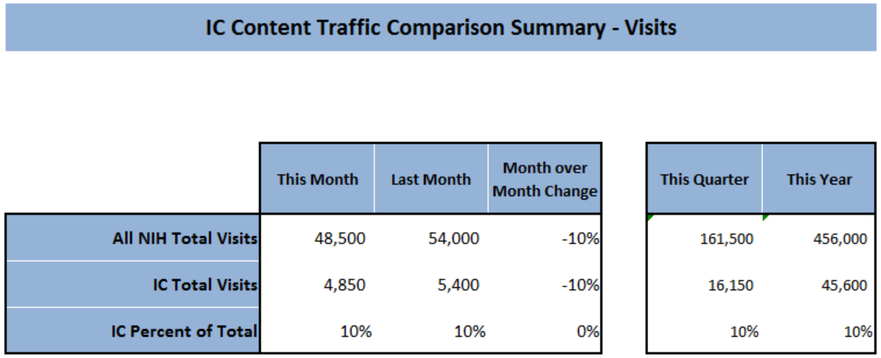
**IC Content Traffic Comparison Summary - Page Views**

| | This Month | Last Month | Month over Month Change |
|---|---|---|---|
| **All NIH Total Page Views** | 136,000 | 135,000 | 1% |
| **IC Total Page Views** | 22,100 | 14,000 | 58% |
| **IC Percent of Total** | 16% | 10% | 57% |

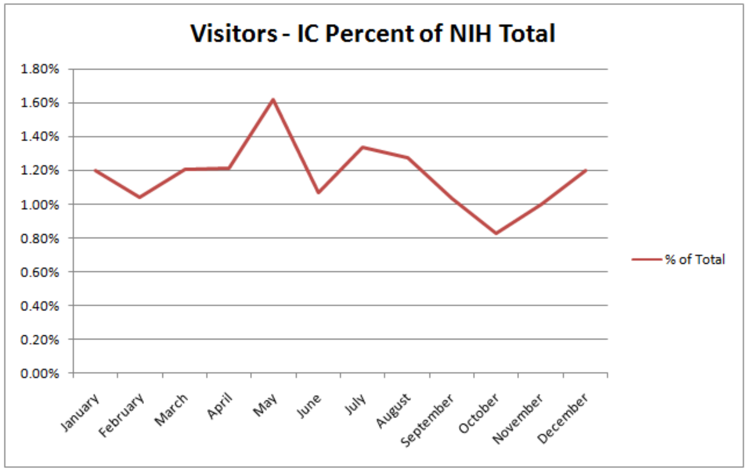
| This Quarter | This Year |
|---|---|
| 413,000 | 1,059,000 |
| 50,300 | 114,600 |
| 12% | 11% |

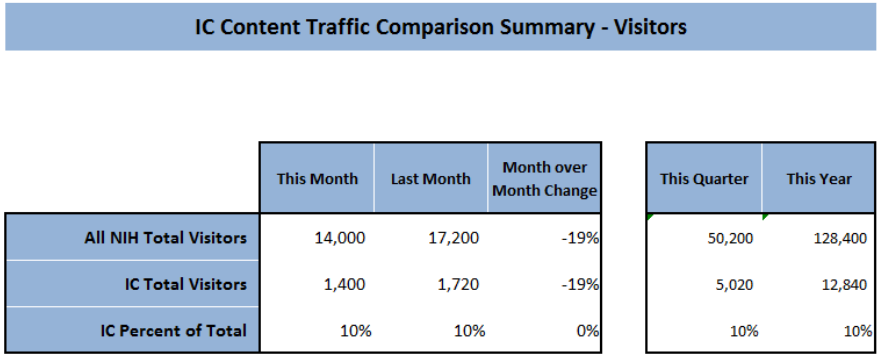
# Figure E.7 Sample Executive Dashboard (continued)




**IC Content Traffic Comparison Summary - Visits**

| | This Month | Last Month | Month over Month Change | This Quarter | This Year |
|---|---|---|---|---|---|
| All NIH Total Visits | 48,500 | 54,000 | -10% | 161,500 | 456,000 |
| IC Total Visits | 4,850 | 5,400 | -10% | 16,150 | 45,600 |
| IC Percent of Total | 10% | 10% | 0% | 10% | 10% |

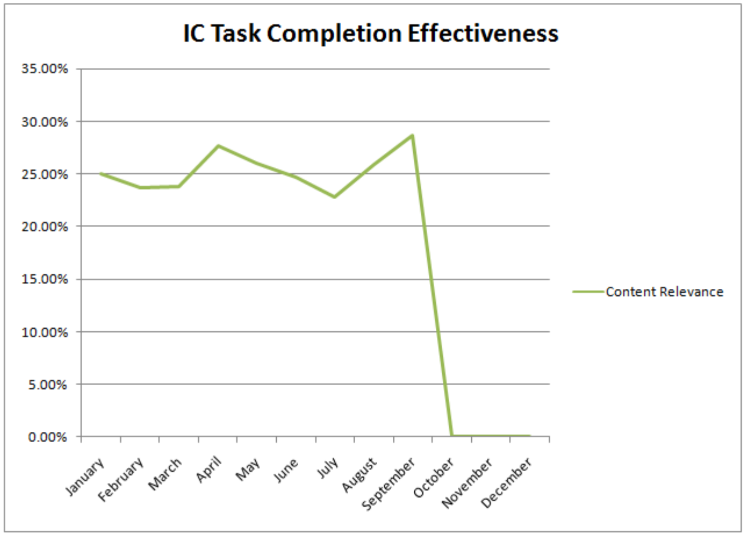



**IC Content Traffic Comparison Summary - Visitors**

| | This Month | Last Month | Month over Month Change | This Quarter | This Year |
|---|---|---|---|---|---|
| All NIH Total Visitors | 14,000 | 17,200 | -19% | 50,200 | 128,400 |
| IC Total Visitors | 1,400 | 1,720 | -19% | 5,020 | 12,840 |
| IC Percent of Total | 10% | 10% | 0% | 10% | 10% |

## Figure E.7 Sample Executive Dashboard (continued)

**Content Relevance** **Benchmark Value: XX**



**IC Content Relevance = Total Mission Relevant Pages/Total Content Pages**

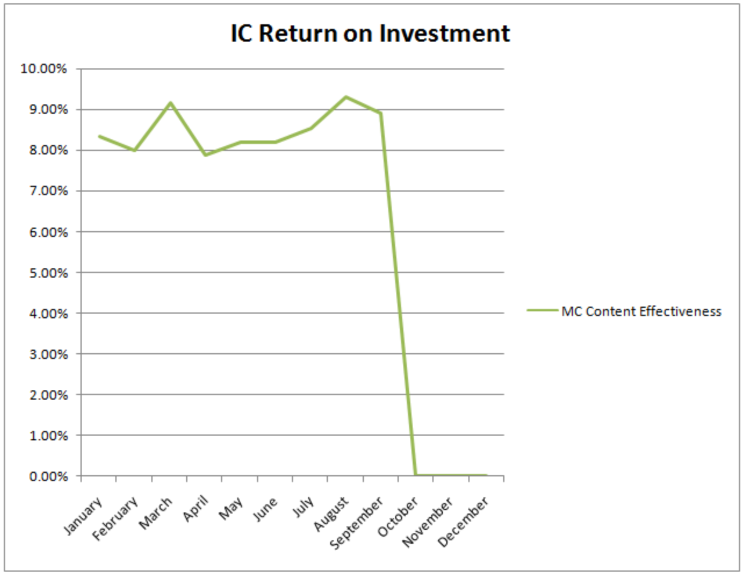
*This shows the percentage of relevant compared to total site pages. Relevant pages are those that received at least 20 views during the month.*

| | This Month | Last Month | Month over Month Change |
|---|---|---|---|
| **Total IC Content Pages** | 1,180 | 1,215 | -3% |
| **Total IC Mission Relevant Pages** | 338 | 314 | 8% |
| **Content Relevance Score** | 29% | 26% | 11% |

| This Quarter | This Year |
|---|---|
| 3,695 | 10,930 |
| 948 | 2,766 |
| 0% | 0% |

## Figure E.7 Sample Executive Dashboard (continued)

### Mission Critical Content Effectiveness Benchmark Value: XX



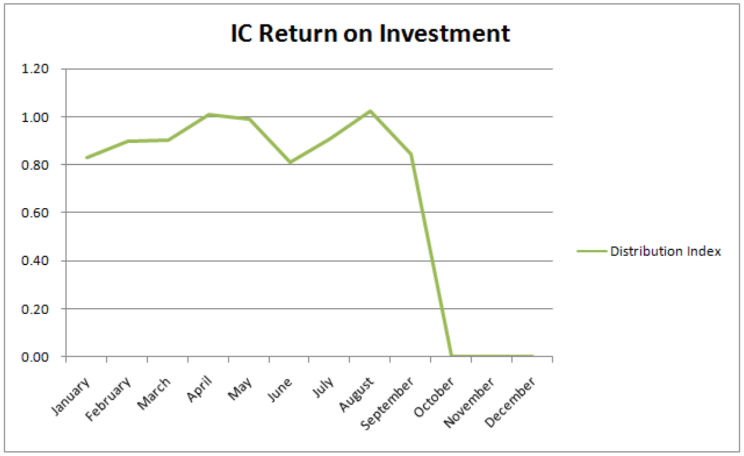
**IC Mission Critical Content Effectiveness = Total Mission Critical Pages/Total Content Pages**

| | This Month | Last Month | Month over Month Change | This Quarter | This Year |
|---|---|---|---|---|---|
| **Total IC Page Views** | 136,000 | 135,000 | 1% | 413,000 | 1,059,000 |
| **Total IC Mission Critical Page Views** | 14,000 | 11,300 | 24% | 48,000 | 111,450 |
| **Mission Critical Content Effectiveness** | 10% | 8% | 23% | 12% | 11% |

# Figure E.7 Sample Executive Dashboard (continued)

## Content Distribution Index

Benchmark Value: XX



### IC Return on Investment = Total Share Events/Total Content Pages

| | This Month | Last Month | Month over Month Change | This Quarter | This Year |
|---|---|---|---|---|---|
| **Total IC Content Pages** | 136,000 | 135,000 | 1% | 413,000 | 1,059,000 |
| **Total IC Share Events** | 47,000 | 51,000 | -8% | 238,000 | 615,500 |
| **Content Distribution Index** | 35% | 38% | -9% | 58% | 58% |

Appendix F.

# Functionalism:
# A New Approach to Web Analytics

July 2006

**Gary Angel, Joel Hadary, and Paul Legutko**

7430 Redwood Blvd.
Novato, CA  94945
(800) 763-2821

## Table of Contents

## *Introduction*

Despite dramatic improvements in the reporting and measurement capabilities of common web analytic tools, the actual practice of web measurement has not shown similar advances. The results of many analytic efforts are disappointing, and the usefulness of web measurement to an organization is still very dependent on having an outstanding individual practitioner. What is more, the practice of web measurement revolves around the usage of a disparate grab-bag of tools (pathing, page analysis, affinity, segmentation, events) with no particularly clear understanding of when, how or why a tool is the most appropriate choice.

These are all symptoms of an immature practice – one hindered by the lack of a standardized method that is highly likely to provide insight into any specific web site and analytic problem. This lack of strong methodology hinders every aspect of web measurement: it hides important insights, it makes training new analysts problematic, it makes sharing approaches and insights difficult and it causes tools to multiply features in ways that are often un-helpful and unnecessary.

Over the course of nearly a decade as a web analytics consulting company, we’ve developed a methodology – Functionalism – that solves these problems. This approach breaks up a web site into its constituent pieces and then assigns one or more specific functions to each piece. These functions can be things like navigation (e.g. route visitors to a specific place), motivation (e.g. convince a user to do something) or information (e.g. provide a visitor with some piece of information). Based on the functions of the page, it is assigned a particular page type from a set of common templates that we’ve distinguished over time in the measurement of different types of sites.

Once a page type is assigned, the success of a page is measured by Key Performance Indicators (KPI’s) specific to the functions it was designed for. Central to Functionalism is the concept that a measure of success for one page can be a measure of failure for a different page.  By categorizing and grouping pages by function, calculating and assigning KPI’s makes the failure or success of a page more straightforward and transparent. It is in this assignment of type-specific, easily measurable KPI’s that the power of Functionalism resides.

The advantages to Functional Measurement emerge in several different fashions. First, because each page on the web site can be measured directly against statistics related to its function, it’s much easier to determine the actual effect of design changes on the page’s success. When measuring a page change against larger site KPI’s (like total revenue), it is often virtually impossible to screen off exogenous effects to the point necessary to get any reasonable statistical measurement of change.

Perhaps even more important than the ability to usefully measure on a page-unit basis, is the conceptual framework that Functional Measurement places around the Design, Implement and Measure cycle. Designers, marketers and analysts can all immediately grasp the basic concepts in Functionalism – pages are built with a purpose in mind and measurement should be focused on that purpose. Because of this, measurement can be integrated in a seamless and REGULARIZED fashion into the design process. This simply isn't the way that web analytics is conducted now.

Finally, by providing a regularized method that can be consistently applied, Functionalism dramatically reduces the need for brilliant individual practitioners. This, after all, is the real purpose behind almost any methodology – and Functionalism provides the essential ingredients to achieve this: conceptual simplicity, near-universal applicability, a direct real-world path to implementation and consistent advantages in actual measurement capability over the majority of ad hoc measurement approaches.

## *Four Steps to Functionalist Web Analytics*

### Step 1: Classification

The first step in conducting a Function-based web analysis is a rigorous classification of the pages on the site. To assist with this, we've built up over the years a library of common types – and in most cases pages on a web site are easily mapped into one of these types.

Here are the most common web types we've identified:

**Engagers**: Pages whose primary purpose is to interest the visitor and get them to do something/anything on the site.

**Routers**: Pages whose function is to move visitors into specific places on the site. One of the differences between an Engager and a Router page is that the latter is built with the expectation that the visitor has come to the page looking for a particular type of information/service. Search should often be treated as a special class of Router page.

**Convincers**: Pages whose function is to "sell" the visitor on a product or service.

**Explainers**: Pages whose job is to help the visitor understand some aspect of a product or service.

**Informers:** Pages whose primary objective is to provide basic news and information about a product or industry. This differs from Explainers in that content is not specific to your own particular product or service.

**Billboards**: Pages providing various content to visitors, but whose primary business objective is to display third-party advertisements on a Cost-per-Impression basis.

**Sponsors:** Pages or series of pages which are themselves revenue-producing, whose content is mostly provided by a third-party who pays for inclusion within your website.

**Closers**: Pages that are supposed to get visitors to enter a conversion process.

**Converters**: Pages that are part of whatever is necessary to gather information/agreements and get a finished lead/sale/transaction.

**Re-Assurers**: Pages built to re-assure the visitor about some potentially problematic issue or concern (privacy policies are a common example).

**Tools:** Pages that are designed to collect or provide information as part of a non-sales process (like checking an account status or finding a location). There are many different kinds of tools and for a tool-rich site tool pages need to be sub-divided into more granular types.

**Completers**: Thank-you pages – designed to signal the completion of a process and – in some cases – drive to additional engagement.

For each of these pre-defined types, there is an appropriate set of measurements (derivable from most major web measurement solutions) that can be applied to measure both the comparative and absolute success of the page.

Every web site is different, however, and it's quite possible that a web site will contain pages whose function isn't well captured by one of the pre-defined types. That means the analyst will have to map appropriate measurements to the actual function of the page. While this isn't a trivial exercise, it's often much easier once the analyst sees how other functional types are treated. As with most problems, the most difficult part is understanding how to approach the problem. Once that's determined, the actual resolution is often quite a bit easier.

It's also important to keep in mind that a page need not be limited to a single function. Many pages, for example, are designed to be both "Convincers" and "Routers" – and a site's home page is frequently expected to be both an Engager and a Router. If, however, you think that a page is an Engager, Router, Convincer and Re-Assurer then it is probably the case that you're asking a single page to do too much. One of the beauties of Functionalism is that the process of classification can – itself – be an aid to good design and clear thinking about the web site.

The process of Classification is essentially manual. Most sites will work with the set of pages from the site based on their Content Management System or Top Pages type reports from their Web Measurement system. It helps to actually bring up each page on the site as you classify it. In most cases, the classification will be obvious. In other cases, multiple functions will suggest themselves. When this happens, all of the competing functions should generally be noted for measurement.

As with so many other problems of this sort, Excel is an excellent tool for keeping track of pages and classifications. Entering classifications into a tool makes it easy to re-sort pages based on classification groups. That's convenient, because much of the later tactical reporting needs to be at the page type level.

For very large sites, the exercise of classifying pages can be needlessly onerous. There is little value in classifying pages you aren't ever going to have time to seriously study – and whose volumes are too small to admit of even Functionalist performance measurement. There is no need to classify every page - concentrate on high-volume pages and pages that are clearly important to your site (pages that have high correlation to success goals for example).

Another technique you can use to save time is to classify groups of content by function. This technique is often quite appropriate and has the advantage that it allows you to treat a group pages (sometimes quite a large group) as a single entity with a single purpose. This would be even more useful if the common web measurement tools provided better support for building content groups.

While the best web measurement tools have some facilities for dealing with groups of content, these are frequently restricted to those marked off by directory structures in the site. And some tools have no virtually no content group analysis capabilities. The degree to which your tool supports these capabilities (and the extent to which your site structure matches your business function) will determine how often you might wish to treat groups of pages as a unit.

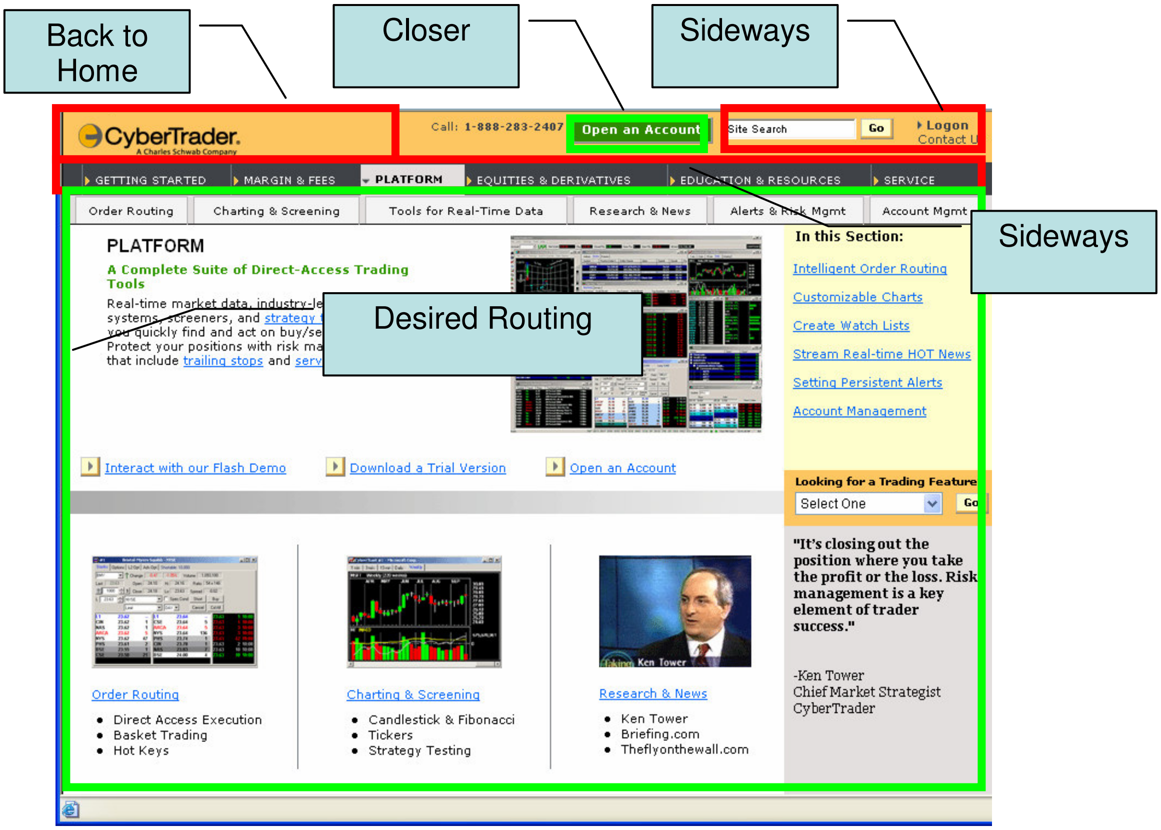
Finally, it's important to remember that there is no "right" answer to page classification. In many ways, the best answer is one of intention – what did you intend (or what would you like) the page to accomplish. While it's certainly within the realm of possibility that a page you intended to be a Router is actually a strong Convincer, this happens very infrequently. It's hard enough to build pages that get visitors to do what you'd like that the chances of getting a good result by accident are quite small!

**Step 2: Measurement Protocol**

For each Functional Type there are one or more defined KPIs that are suggested for getting the best measurement of that type of page. These suggested protocols for measurement will almost certainly require at least a little bit of site-specific tuning.

Let's take an example. One of the key measurements for a Router Page is the percentage of next pages driven from the Body of a Router Page (on first viewing and then on Subsequent views) compared to the next pages driven from Top or Left Navigation, Search, Back Button, side or bottom advertisements, and Site Exits. This is a simple thought – the job of a Router Page is to move people down into detailed content. If visitors use Search, go back to Home, or slide across the site using Top Navigation then the page hasn't done its job.

Given a page like the one shown below, the sections are fairly easy to describe.



PLATFORM

A Complete Suite of Direct-Access Trading Tools

Real-time market data, industry-le systems, screeners, and strategy you quickly find and act on buy/se Protect your positions with risk ma that include trailing stops and serv

- Direct Access Execution
- Basket Trading
- Hot Keys

- Candlestick & Fibonacci
- Tickers
- Strategy Testing

- Ken Tower
- Briefing.com
- Theflyonthewall.com

"It's closing out the position where you take the profit or the loss. Risk management is a key element of trader success."

-Ken Tower
Chief Market Strategist
CyberTrader

For measurement purposes, however, the analyst is going to have to classify next pages (or links) according to which of these basic navigation options each actually represents. The "Open an Account" button here is an interesting and very common case. When there is a single goal dominant on a site, it is common to include a "Closing" element – a drive to conversion – on every page. Usually, it's best to include this as a separate navigational category when evaluating Routers. These aren't bad outcomes, of course, but neither are they the directly intended outcome. If you saw that a Router page was driving lots of visitors directly into conversion, you might want to re-think the role of the page and include more “Convincer” elements. By measuring them separately, you get the best insight into the real performance of a page.

Chances are you’ve looked at the exit rate for your “Router” pages already. And you may already have looked at “Back-out” rates (most often a return to home). And if you are really good, you’ve probably also looked at “Search” rates. But the most common failure of a Router Page is actually sideways navigation – using the Top Navigation to go to another area of the site without any drill-down. Since top navigation isn’t obviously bad in the way that site exits are (or Search might be), most analysts never think to classify and aggregate links in this way.

The specific KPI’s for each page type are described in the page type templates included at the end of this paper. Most of these KPI’s are straightforward in their application and require almost no additional work. The classification of routing we described earlier is one of the exceptions.

Some of the other KPI’s that require additional work include:

Exit Propensity: Basic exit rates are often very un-informative when it comes to understanding page behavior. There are many reasons for this – almost all covered in the Functionalist paradigm. First, many high exit pages have specific jobs that make visitor abandonment natural. “Completer” pages (Thank you pages) are a classic example of this. “Engager” pages and “Converter” pages also tend to have exceptionally high exit rates. By comparing Exit rates within a paradigm, you get a much better sense of actual performance. In addition, we like to measure exits (within type) by depth of access. Web Sites tend to be an interest funnel – visitors are most likely to exit on the first page and are less likely to exit with each subsequent page viewed. The vast majority of web sites demonstrate this basic tendency. To counteract this effect in measurement, you can measure exit rate for pages grouped by average depth of access. We call this depth and type weighted Exit Rate the Exit Propensity and it’s a vastly better measure of exit performance than any other alternative.

Engagement Links: These present essentially similar issues to router links – you need to classify the links on the page.

Entry Rate and Entry Propensity: Like Exits, entry propensities can measure the degree of independence and integration of a particular page or set of pages, by gauging the intention and focus of the visitor. Particularly when compared to exit propensity and entry rate, entry propensity can measure how popular or effective a particular tool or explainer can be, contributing to long-term visitor loyalty to the site.

Dead-Ends: Also called back-outs, these are cases where a visitor re-traces a path back to a proceeding page. This is a common behavior on sites, and there are site designs (not usually optimal) that actually encourage this. One of the most common back-outs is to the home page. Where a measurement solution doesn’t provide the capability to measure back-outs generically (usually the case), the easiest solution is to focus on specific back-outs that are especially likely given where the page fits in the navigation scheme.

Re-Surface Routes: It is frequently useful to measure the visitor behavior of pages when they are being re-accessed. This can be a difficult issue for measurement Adaptation (making your measurement solution capture what you want) but it is also an issue for the Measurement Protocol in that you'll need to decided whether this is likely an important enough behavioral pattern to merit measurement. If it wasn't difficult to do in most tools, this might not be an issue. But since Re-surfacing behaviors are actually very hard to track with many tools, it's important to consider when it is actually worthwhile to make the effort.

Most analysts do not have the resources to study each page of a site individually, however, and the web analytics tools out there encourage analysis of many pages at once. Once the principal functions of pages have been identified, however, this categorization can be applied to the same KPI report across many pages. For example, one can download into Excel a list of pages with exits and visits directly from a web analytics tool, even by cutting-and-pasting. A quick calculation (exits/visits) might reveal the following chart:

| Page | Exit Propensity |
|---|---:|
| newhouse.com/mortgage_calculator | 95.93% |
| newhouse.com/find_a_house | 63.82% |
| newhouse.com/branch_locator | 63.70% |
| newhouse.com/todays_interest_rates | 62.04% |
| newhouse.com/homepage | 60.42% |
| newhouse.com/our_affiliates | 56.93% |
| newhouse.com/start_here | 56.72% |
| newhouse.com/real_estate_outlook_July | 56.07% |
| newhouse.com/about_our_services | 53.51% |
| newhouse.com/search | 51.86% |
| newhouse.com/Thanks_for_Registering | 50.95% |

A list like this – or, as more often the case, a list ten times this length – might be somewhat unintelligible. These are all high exit propensities, but does this mean that each page needs optimizing? If Functionalist Categories are applied to each page (an easy task in MS Excel), how to interpret this data becomes more clear:

| Page | Exit Propensity | Function | Evaluation |
|---|---:|---|---|
| newhouse.com/mortgage_calculator | 95.93% | Tool | OK |
| newhouse.com/find_a_house | 63.82% | Router | Poor |
| newhouse.com/branch_locator | 63.70% | Tool | OK |
| newhouse.com/todays_interest_rates | 62.04% | Informer | OK |
| newhouse.com/homepage | 60.42% | Engager | Poor |
| newhouse.com/from_our_affiliates | 56.93% | Sponsor | OK |
| newhouse.com/start_here | 56.72% | Closer | Poor |

| | | | |
|---|---|---|---|
| newhouse.com/real_estate_outlook_July | 56.07% | Informer | OK |
| newhouse.com/about_our_services | 53.51% | Explainer | Poor |
| newhouse.com/compute_closing_date | 51.86% | Tool | OK |
| newhouse.com/Thanks_for_Registering | 50.95% | Completer | OK |

Tools, Informers, Sponsors, and Completers should be expected to have a high Exit Probability, so seeing them high on this list should not be worrisome. But Routers, Engagers, and Closers should have a low exit propensity, so these are the pages to look at based on this table. This straightforward example shows a relatively quick and easy way to apply functionalism to web analytics while staying very close to the original data from your web analytics solution.

### Step 3: Adaptation

One of the most attractive features of Functionalism is that it is not built on top of any proprietary or esoteric measurement system. Most Functional analysis can be conducted from any of the common enterprise web analytics tools. There is no doubt that tools could better embody Functional concepts – and this would make analysis easier – but there is no reason to think that you will have to scrap your existing measurement solution.

In Step 3, you need to adapt your tool and the statistics it provides to the KPI's required for a Functional analysis. In most cases, this process is not especially difficult – but the vagaries of each tool will always necessitate some mapping.

In our routing example above, for instance, you'll need to decide how you are going to perform the analysis. Two obvious methods are using link based analysis or next page based analysis. Not every tool provides good link click reporting – so you may be forced into next page analysis. This usually isn't a problem and next page reporting has the advantage of being cleaner (in terms of naming conventions) and of capturing events (like back to home) that might otherwise be missed.

Analysis of first time vs. subsequent routing is quite a more difficult. In HBX, you might use an Event Sequence and in SiteCatalyst you might use PathFinder or a Data Warehouse request. In neither case, however, will the answer simply fall-out in a trivial fashion. And some web measurement tools won't support this break-out at all – so you'll be stuck lumping first routes in with subsequent routes.

The choices you make about Communication (reporting measurement) will also effect the amount and type of work you need to do for Adaptation. To measure individual pages, you don't typically need to worry about anything other than the stock problems of Adaptation – how to get the right numbers from your Measurement solution.

Moving Functionalist reporting to a higher level may not be so simple, however. Suppose you want to measure your overall site effectiveness at engagement by measuring the percent of prospects who reach either a Convincer or Converter page[1]. A large site may have hundreds of each of these page types spread across dozens of Content Areas. Attempting to create a unique visitor ratio presents a considerable challenge to any of the commonly available web measurement tools.

Similarly, not all tools make it easy to provide reporting on pages grouped by an outside classification of page type. The tools that work best for this are those that allow for reports to be built in Excel. This allows for a manual classification to be overlaid on top of the reporting so that there is no necessity to build the site classification scheme into the tag (though this will sometimes be useful).

---

[1] We're indebted for this example to the Public Site Team at Charles Schwab.

If your tool doesn’t support this kind of reporting, then you may need to limit how much of the Functionalist model is represented in your Management Reporting. Fortunately, as already mentioned, this in no way limits your ability to deploy Functionalist analysis in the tactical evaluation of key web site components.

### Step 4: Communication

There is not (and never will be) an absolute standard of performance for web sites or their functional components. It is impossible to say that a Router page should send 75% of its visitors into the Routing Area and that if it fails to do so the page needs work. Sites vary too much in the quality of their visitors, the demands they necessarily make on visitors and in the environment (brand, campaigns, marketplaces, etc.) that surrounds them to set any one performance bar.

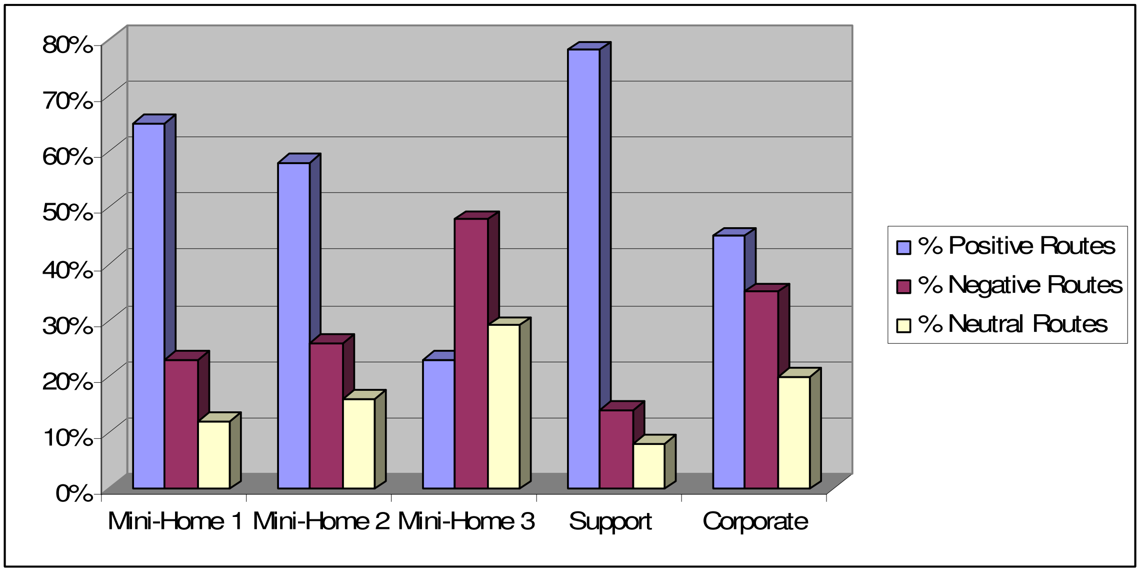
However, Functionalism can be (and is) tremendously effective in isolating comparative performance on the units of a single web site (or group of related web sites). Given mostly similar sourcing and environment, functional performance is very comparative. Since the primary purpose of Functional analysis is to direct the design process internal to a site, this is really all that is required. By comparing performance within a class of pages, it is simple to identify the best and worst-performing pages.

At the highest level, Functionalist Reporting should focus on two things:

- The comparative performance of content by type.
- The trended performance of types over time.

The first of these is designed to pinpoint the top targets for design effort. The second measure is designed to provide feedback on the effect of design changes and the degree to which incremental improvements are actually working.

Charts like these provide a simple, immediately understandable snap-shot into what is working and what isn't:



One of biggest advantages to Functionalism is the degree to which it makes fairly complicated measurement procedures immediately understandable to both designers and marketers. The best Functionalist reporting is designed to collapse most of the various measures into as a simple a representation of success at the core function as possible.

In addition to providing a powerful means of identifying acceptable standards of comparison for a site (no mean achievement), the Functionalist mindset can help a site define KPI's that make significantly more sense in terms of its real business.

The ultimate goal of a site is, of course, conversion. Many sites have a significant off-channel component that can be extremely difficult to measure. Customer Research (on or offline) can help plug this gap – but there are significant limitations in the ability of each to map real customer behavior and attitudes. Simple conversion proxies (like site visits or viewed x number of pages) are also valuable but often misleading. Replacing these definitions with Functionalist measures can make much clearer the extent to which visitors on the site are doing what you want them to.

## *Visitor Segmentation*

Current web analytics theory (if not always practice) says that visitor segmentation is at the heart of good analytics. Proper visitor segmentation can (and should) be integrated into your Functionalist approach.

However, one of the things we've realized in the course of many real-world engagements is how difficult it can be to build "intentionality" into an analysis. Persona-based models of behavior have many advantages – including how nicely they square with traditional marketing approaches. But they can be demonically difficult – often impossible – to implement/measure with the real-world web analytics tools that are available.However, where actions on a web site (or information contained about customers) does allow for meaningful visitor level segmentations to be built, these should definitely be applied to the Functionalist KPI's.

Some of the most common splits that are well worth attempting including Customers v. Prospects, Repeat Customers v. 1st Time Customers, Repeat Prospects v. Single Visit Prospects, Direct Converters v. Multi-Session Converters and Customer Support visitors. This is not meant as anything approaching an exhaustive list – and the range of possible segmentations is much greater than the number of web sites since each site will likely have at least a few segmentations that are uniquely of interest to that business.

The visitor segmentations that you use will be dependent on the exact nature of your business and your web site. In most cases, what you'll want to do is look for places on the web site that are strong clues to type or intention. By coding visitor populations based on those places, you should be able to carve out fairly usable visitor segments. You can then apply your Functionalist KPI's to each segment.

There are also many possible Functionalist-Based segments, where functionalist designations serve as the basis of a visitor segment. For example, a segment can be created to look only at visitors who see Tools, Informers, or Sponsors. Their subsequent multi-session behavior can indicate the relative value of pages that might otherwise be peripheral to your overall business goals. Similarly, Explainers and Closers are often visited in a different session from the session where the conversion takes place. Only a visitor segment, separate campaign, or creative conversion rules will capture this behavior and ultimately reveal the value of these pages.

In many cases, you'll decide that a page simply doesn't apply for a given segment. That's useful knowledge in that you can ignore its KPI's for that segment (but you'll probably want to look at why the segment is hitting the page). Equally likely is that you'll decide the KPI's for a single segment are all that you care about – allowing you to ignore the overall page performance and just measure the KPI's for one visitor type.

Just as there is no one right set of segmentations, there is not single set of right answers about related issues regarding segment exclusivity and segment duration. In some cases, especially when used for management reporting, it is very nice to have mutually exclusive segments that add up to a whole population. This is less commonly advantageous when doing analysis.

Segment duration (the amount of history you need) is even more variable. Every business has a different sales-cycle. Certainly one of the factors that every web marketer should be looking to understand is how long (time and sessions) it takes visitors to go from new prospect to converted customer. Even on operational or community sites, there is a similar path moving from new user to power user. Understanding the duration and intensity of this cycle for your business will help answer questions about how long your web analytics segmentations need to persist.

## ***Summary***

So much about the current state of web measurement is frustrating – to both practitioners and web marketers. Traditional analytics has focused on either simple measures of page performance (like exits or top paths) that are too misleading in most cases to be useful or on relationship to ultimate conversion, a much better technique, but one which is frequently impossible to do with any statistical meaningfulness.

There are huge sets of web analytic KPI’s but no process for understanding when and how to apply them. There is the inevitable demand for more data, better ties to customer behavior and the usage of alternate methods (like online surveys) that are all supposed to provide the insight web behavior doesn’t.

These demands aren’t unreasonable. They all make a certain amount of sense – but none of them will really provide a significant improvement in your web measurement unless and until web analysts and your marketing managers and your web designers all have a way to think about measurement that makes sense.

The Functionalist approach outlined here is the fruit of many years of actual, hands-on web analysis. It is designed to be applied in the real-world, with real-world tools already in the hands of most analysts. It is designed to provide a framework within which analysts can not only do their work, but can communicate their work out to the rest of an organization. It is built in a way that allows every player in the web marketing team to use it – both conceptually and practically – to do their job better.

It is, in our experience, a dramatically better way to do web analytics.

## *Functionalist Templates*

Page Type: **Engagers**

Functional Description: A page whose primary job is to grab the visitor's interest and get them to do something (almost anything!) on the web site. In general, the pre-supposition to an Engager page is that there is little known about the intent of the visitors landing here. Where more is known about what a visitor desires, there is usually a specific set of directions that the page is expected to drive toward and the page is generally better (or also) classified as a Router.

Sample Type: One of the most common types of Engager page is the Home Page on many web sites.

Functionalist KPI's: **% Engagement Links** (links to directed content – generally this excludes search), Exit Rate, Exit Propensity (a session depth-weighted measure of exit likelihood), Subsequent Page Consumption, Subsequent Success, % Return Visitors, Organic SE Entry %.

Notes: Many pages on publishing sites are essentially engagers. There is no particular desire for a visitor to move in a particular direction but there is always value to having a visitor consume more.

Measurement Issues: Engager pages are frequently Landing Pages. One of the questions in measuring effectiveness that often gets raised is how much engagement is due to a page and how much is due to general brand awareness among visitors. One helpful technique for isolating the actual engagement effectiveness in pages (and magnifying the effect of changes to make them more visible) is to segment a population group that is as neutral as possible. We have had good luck using population groups sourced from non-branded Search Terms.

Page Type: **Routers**

Functional Description: Pages whose primary purpose is to move visitors into particular sections of the site. The presumption is that there is fairly substantive information about what the visitor might be interested in and these alternatives are presented as navigational elements in the body of the page.

Sample Type: Pages accessed as the top page from a top-level navigation bar are often primarily router pages.

Functionalist KPI's: **% Body Routes**, % Routes by group (body, top, back), Exit Rate, Exit Propensity, % Re-surface (% of visitors who drill-down then come back up to the Router Page), %Re-surface Body Routes, %Re-surface Routes by group, %Re-surface Exits.

Notes: Routing page performance is often one of the most important elements in overall site performance. And unlike home pages, the linkage patterns from Router pages don't always get much study. Where analysis reveals particularly abysmal routing performance, this may indicate that a lot of detailed information isn't necessary - closing information and drives might be more appropriate. Where re-surface behavior is common, understanding the re-surface routing may suggest dynamic ad serving strategies on re-surface to effectively remap a router into a closer.

Measurement Issues: By far the trickiest aspect of router pages is measuring (and separating) re-surface from initial land behavior. This isn't always an issue, and before bothering with a more complex analysis, the analyst should check and see the percentage of visits that contain multiple pages views of a Router.

Page Type: **Convincers**

Functional Description: Pages whose function is to “sell” the visitor on a product or service.

Sample Type: Most product detail pages and pages accessed from Routers are “Convincer” pages.

Functionalist KPI’s: **Rate of Subsequent (multi-session) Drives to Closer/Converter**, Rate of Subsequent affiliated success completions, Rate of immediate drives to Closer/Converter, Rate of same-session drives to Closer/Converter, Rate of same-session drives to affiliated success, % affiliated Routes.

Notes: For many (perhaps most) sites, conversion is a multi-session process. This means that the performance of Convincers cannot reasonably be measured using single session statistics. This is not always the case, so one of the most important background measurement tasks you’ll need to perform is to understand the extent to which successful visitors are multi-session visitors. Where conversion cycles are commonly multi-session, the most common behavior pattern in actual conversion sessions is to bypass “Convincer” pages altogether – making them appear completely useless. A related issue is the decision to measure against drives to the Closer/Converter or to actual completions. Measuring to Closer/Converter drives (exclusively) can lead to a site that is very aggressive in driving traffic (perhaps not well-qualified) into a conversion process.

Measurement Issues: Many measurement tools are still quite limited in their ability to measure cross-session performance. The easiest way to get this measure for most sites/systems is to create visitor segments  based on visitors who visit the Closer/Converter/Completion pages. By comparing visits to “Convincer” pages in these segments to same page visits for the entire population you get a reasonable cross-session measure of “subsequent” drives.

Page Type: **Explainers**

Functional Description: Pages whose job is to help the visitor understand some aspect of a product or service.

Sample Type: Customer Support pages, FAQS, System Requirements pages, etc.

Functionalist KPI's: % Search Next Steps / % Search Next Steps all Site non-home, Exit Rate, Exit Propensity, affiliated Route %.

Notes: Explainer pages are difficult to measure because they often aren't tied to a specific outcome on the site. However, for sites with Search functionality, it is often illuminating to look at the percent of times visitors resort to broad search from an Explainer page. This is generally a measure of dissatisfaction with the information provided. It is also important to realize that most non-support Explainer pages occur within a larger framework of navigation – so they are also expected to continue routing within an area.

Measurement Issues: Isolating the set of Search Terms from an Explainer page is often a good way to understand potential sources of dissatisfaction. For Customer Support pages, it can also be helpful to think about the page as a "Completer" as well as an Explainer.

Page Type: **Informers**

Functional Description: Pages whose primary objective is to provide basic news and information about a product or industry

Sample Type: Industry News pages, Market Update pages.

Functionalist KPI's: Entry Propensity, Exit Propensity, Visitor Return Frequency, Page View Consumption of non-Informer pages, attrition, % of Organic SE entry, subsequent conversion.

Notes: Like tools, these sets of pages can stand alone as a source for site traffic. The independence, integration, intent and focus of these pages can tempt site designers to expend a lot of resources towards their optimization and attractiveness, but their ultimate objective must remain visitor retention and ultimate conversion.

Measurement Issues: It can take a long time to measure attrition and return frequency, and for particularly dynamic pages such as RSS feeds and hot topics, traffic is usually based more on what topic is being discussed rather than any design considerations. And, like tools, there will always be a percentage of visitor traffic whose only interest is the informer pages. These pages should be measured with some caution, therefore, as site designers and analysts might get very excited by high click-through and entry rates without realizing that the value of these pages lies in their more long-term results.

Page Type: **Billboards**

Functional Description: Pages providing various content to visitors, but whose primary business objective is to display third-party advertisements on a Cost-per-Impression basis.

Sample Type: Pages heavy with banner-ads; pages which host Google AdSense ads.

Functionalist KPI's: Page Views, Page View Consumption, Visit %, Visitor %, Page View Consumption in&out of Billboard pages; CTR to advertisements vs. lost traffic; attrition, visitor return frequency.

Notes: Basic traffic measurements such as page views are central to the performance measurement of these pages. Click-through-rates to these advertisements, however, must be carefully balanced with lost site traffic, particularly in the case of Google AdSense, which optimizes and pays on a cost-per-click basis. A fat check from Google every month is a tempting goal, but not at the expense of your e-commerce. Sites which optimize for ad-revenue can often see attrition rates start to soar and visitor loyalty tank.

Measurement Issues: It can be difficult to track Click-Through Rates to advertisements in most web analytics tools, since these are necessarily off-site links and are thus lumped together as site exits. The Best-Case scenario, where a visitor clicks on an ad but then back-buttons back to your site, is often invisible to web measurement without special tagging. And reconciling analytics data with data from your advertising providers is always a challenge.

Page Type: **Sponsors**

Functional Description: Pages or series of pages which are themselves revenue-producing, whose content is mostly provided by a third-party who pays for inclusion within your website.

Sample Type: Pages “brought to you by X”, pages “from our affiliates”.

Functionalist KPI’s: Page Views, Page View Consumption, Visit %, Visitor %, Page View Consumption in&out of Sponsor pages; CTR to sponsor pages; dead-end rates, visitor return frequency, SE Organic Entry %; Exit Propensity.

Notes: These pages are like separate sites-within-your-site. The principal page of these sponsors often acts as a separate mini-router and engager. While the primary objective of these pages is reached as soon as a visitor finds them, consumption of pages outside of sponsors is important to your overall site performance. Minimizing Exit Propensity while increasing CTR to sponsor pages is an important optimization process.

Measurement Issues: Ideally, a separate variable needs to be set up in order to roll-up sponsored pages. Since these pages and their designations change frequently according to sponsorship sales and resulting contracts, frequent tagging changes might be necessary, although these are, in practice, difficult to arrange. Since content on these pages is usually determined by the sponsors, there is also limited optimization possible to decrease exit rates.

Page Type: **Re-Assurers**

Functional Description: Pages built to re-assure the visitor about some potentially problematic issue or concern (privacy policies are a common example). The function of a re-assurer page is to get a visitor to re-join a process.

Sample Type: Privacy pages, secure shopping pages and why we need this information pages are all very common re-assurers.

Functionalist KPI's: Process Re-Join Rate, Exit Rate, Exit Propensity, Percent Process Viewers / Percent Re-Assurers.

Notes: The group of visitors who access Re-Assurers is almost always the least confident/qualified – so success rates for these pages are usually poor. One of the things to be aware of is the percentage of process viewers who check out a Re-Assurer. If this percentage is quite high, then it may make sense to move key information directly into the process. If, on the other hand, usage is very low and not successful then it may sometimes make sense to simple remove the pathways.

Measurement Issues: Re-Assurers are sometimes coded as rollovers or popups. It is often necessary to insure that these pages are tagged. In addition, popups often confuse pathing analysis in measurement solutions. Where a popup is present, make sure that you correctly analyze next steps from the original page.

Page Type: **Converters**

Functional Description: Pages that are part of whatever is necessary to gather information/agreements and get a finished lead/sale/transaction.

Sample Type: Shopping Cart pages are the most common example.

Functionalist KPI’s: Step-Drop off, Exit Rate, Affiliated Content Exit Rate, Re-Assurer Exit Rate, Avg. Step Time, Avg. Step time on Progression, Cross-Session Step-Drop Rate.

Notes: One interesting aspect of Converter drop-off is identifying how the drop-off takes place. Is it simply site exit, or are visitors going to re-assurer pages or back to convincer pages. These are often especially interesting when cross-tabulated with source – and can help identify places where you might need to bolster your information or pathways prior to driving conversion.

Measurement Issues: For the most part, Converters are fairly easy to measure. In some cases, intra-field abandonment can also be interesting – but this is less common than you might expect. Depending on the nature of the conversion process, it’s also important to find out the degree to which a conversion process is multi-session. Where multi-session behavior is common, in-session drop-off rates can be misleading. Because measurement tools frequently do a poor job of handling multi-session behavior in their funnels, we often use visit or visitor numbers for each page to create the basic step-drop off statistics.

Page Type: **Tools**

Functional Description: Pages that are designed to collect or provide information as part of a non-sales process (like checking an account status, executing a trade, paying a bill, or finding a location). There are many different kinds of tools and for a tool-rich site tool pages need to be sub-divided into more granular types.

Sample Type: Mortgage Calculator, Store Lookup, Mapping, etc.

Functionalist KPI's: Organic SE Entry %, 1st Step % (% of visitors whose first action is the tool – indicative of pre-determined desire to use the tool), Entry visitor qualification, Exit Rate, Exit Propensity, Affiliated Content Routing, Subsequent Conversion, Direct Conversion, frequency of use, and attrition.

Notes: Tools are created on a site for a whole range of reasons. In some cases, they are meant to be draws to a site – ways to get potential prospects in the door. Other times their function is to keep people on the site – visitors will want to do X within the context of their visit we'd better give it to them. In still other cases, they are services that generate revenue. And, of course, some tools have functions covered by other classes of page (wizard tools are essentially Routers for example). Naturally, the function of the tool will impact the appropriate ROI's for it.

Measurement Issues: Many tools are, by default at least, a black box to web measurement. You know a visitor went in and you where they came out. What happens in between is a mystery. In many cases, "black-box" tool analysis is perfectly adequate. But if the tool is extensive, you'll probably need to consider internal tagging. This usually isn't wildly difficult – but it also something that is much easier said than done – and much easier to do during implementation than after a tool is out the door.

Page Type: **Completers**

Functional Description: Thank-you pages – designed to signal the completion of a process and – in some cases – drive to additional engagement.

Sample Type: Order or Lead Confirmation pages.

Functionalist KPI's: Site Re-Engagement Rate, Exit Rate.

Notes: You're done. You got the order. Now there's nothing left but saying Thank You. Not so fast. Completers represent one of the most oft wasted spaces in the web channel. If you've just generated a lead or gotten a customer, there surely ought to be something else you'd like them to know: training, classes, options, support…something. Most Completers do a terrible job of re-engagement. And even a very good re-engager is going to have a very high exit rate. That's why it's so important to treat this as a specialized type of page.

Measurement Issues: This is usually a very straightforward next step analysis.

# Appendix G.

# How to Conduct A/B Testing

## Contents

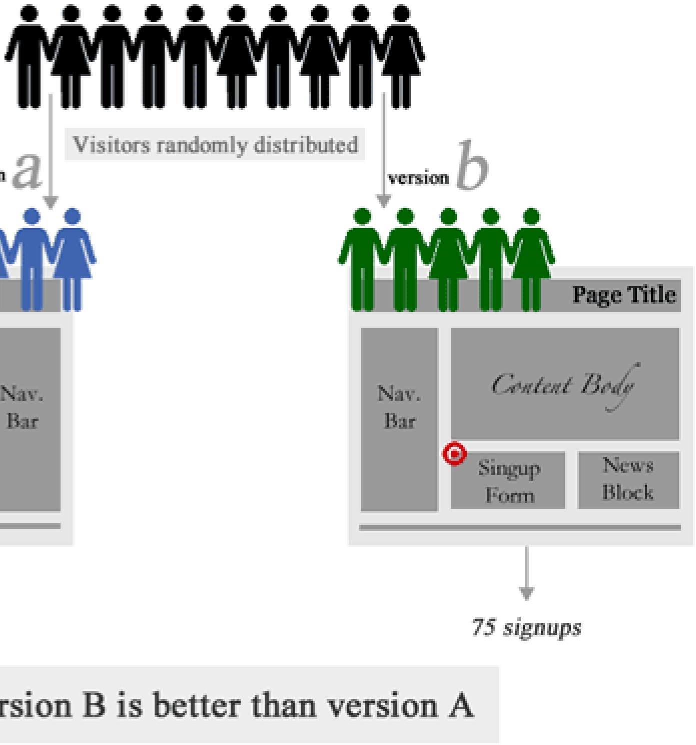
Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix G. How to Conduct A/B Testing

## A/B Testing Overview

What is A/B Testing? At its core, A/B testing is a process for comparing two versions of a Web page element, such as content, a graphic, or both, with a metric that defines success to determine which version (A or B) is better. A/B testing allows you to compare several alternative versions of the same Web page simultaneously and see which one produces the best outcome based on the measures you define as success. **Figure G.1** depicts an A/B test to determine whether the position of the email signup form on a Web page will increase the number of signups. In the example shown, the test provided data showing that Version B, with 75 signups, worked better than Version A.

***Figure G.1.*** *Example of A/B testing to determine best position for email signup form.*



## Why Do A/B Testing?

So why do A/B testing? We all make assumptions about visitor behavior that are clouded by our familiarity with our own sites. Testing your assumptions can verify assumptions with data and lead to great insights about your Web site's visitors. If done correctly A/B testing offers results that are completely based on data, not on personal opinions or "gut feelings." Such information will allow you to improve your site based upon direct visitor interactions and preferences.

## What to Test?

The choice of what to test will obviously depend on your Web site goals. For example, if your goal is to increase the number of email subscriptions, then you might choose to test the size and location of your “subscribe” button, the information collected on the email subscription form, or the look and feel of the subscription form, etc. The goal of A/B testing in the example above was to determine which factors would lead to more visitors signing up for your email subscription. The A/B test will offer insights into questions like:

- Are visitors having trouble finding the subscription button?
- Is the email subscription form too vague or too long?

These questions can be answered individually by testing the appropriate website elements. Furthermore, A/B testing does not have to include all of your users. Conducting tests with certain visitor segments can be a very productive approach. For example, if you notice that users referred to your site from Google (search) are not completing tasks as often as other visitors are, you can test two different types of task-oriented content just with the visitors from Google. Then you can check the A/B test completion rates to see if one of the revised versions increases the task completion rates for users referred by Google.

Even though every A/B test is unique, certain elements are frequently tested:

- Calls to action, including the content, font size, color and/or placement on the page
- Headlines or descriptions of content
- Forms, including the length and number of fields in forms
- Layouts and graphic designs of Web site
- Images on homepage, landing pages or product pages
- Amounts of text on page (short vs. long)

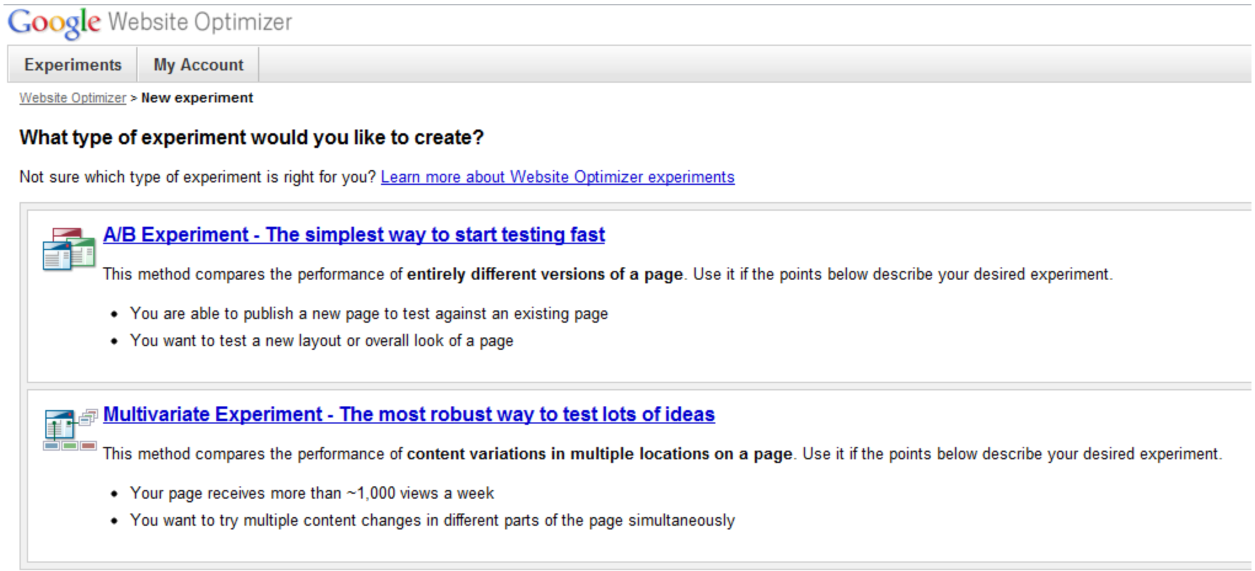
## Creating Your First A/B Test

Once you have decided to do an A/B test, the next step is to select the testing tool. There are a variety of tools available, both free and paid, such as Adobe’s Test and Target, Webtrend’s Optimize and Google’s Website Optimizer that allow you to do A/B testing on Web sites. One basic free tool that will require some coding in both HTML and JavaScript is Google Website Optimizer. Google Website Optimizer is a fairly straightforward tool and if you are using Google Analytics, it integrates nicely.

The following example shows how to set up and conduct a test with Google Website Optimizer. Other testing tools have similar processes.

Once you access the testing tool, the first step is to go to the site and select the type of test or experiment you want to create (**Figure G.2**).

**Figure G.2. Screen shot of Google Website Optimizer test-selection site.**



**What type of experiment would you like to create?**

Not sure which type of experiment is right for you? Learn more about Website Optimizer experiments

**A/B Experiment - The simplest way to start testing fast**

This method compares the performance of **entirely different versions of a page**. Use it if the points below describe your desired experiment.

- You are able to publish a new page to test against an existing page
- You want to test a new layout or overall look of a page

**Multivariate Experiment - The most robust way to test lots of ideas**

This method compares the performance of **content variations in multiple locations on a page**. Use it if the points below describe your desired experiment.

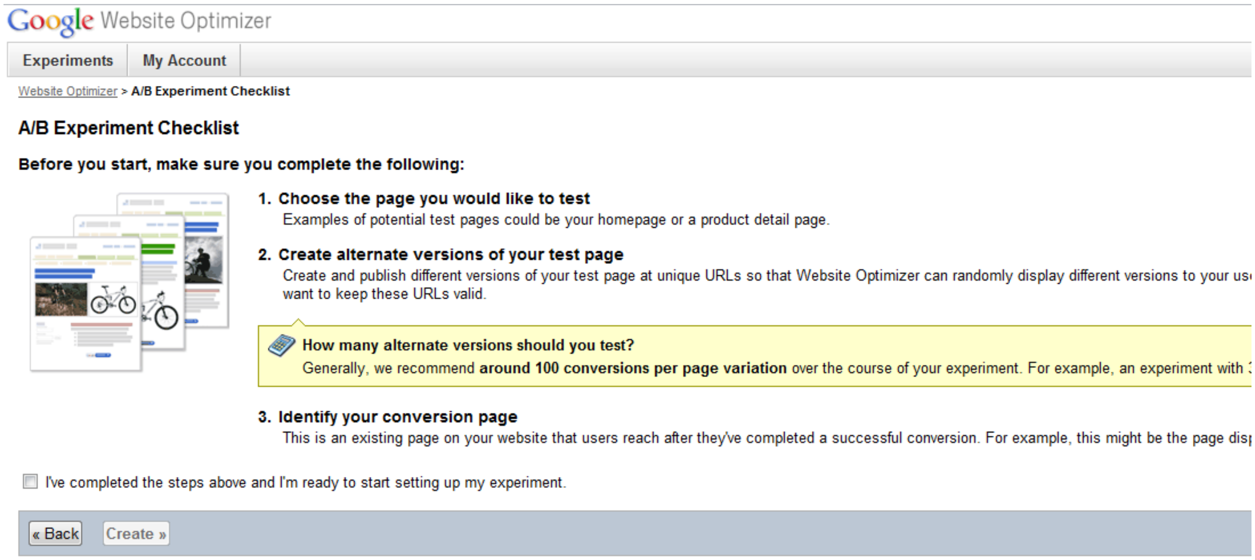
- Your page receives more than ~1,000 views a week
- You want to try multiple content changes in different parts of the page simultaneously

You can choose between conducting an A/B or Multivariate experiment. A/B testing will compare the performance of two different versions of a Web page element, such as content, graphics, font or the call to action. Multivariate testing compares the performance of content variations in multiple locations on a page. A/B testing is definitely the easier of the two testing options and the one we recommend trying first.

After choosing to do an A/B Experiment, the next step is to follow the A/B Experiment Checklist. You start by will choosing the page you will test. The second step is to create the two versions of the selected Web page. Each test page will need a unique URL so that Website Optimizer can randomly display the versions to your users. Your users may bookmark these URLs, so after the experiment finishes, you should keep both of the URLs working.

The third step is to identify your conversion page (the existing Web page that appears when visitors take the desired action during the test). For this example, the conversion page is the page the user will see upon completing the email subscription form (**Figure G.3**).

**Figure G.3. Screen shot of Google Website Optimizer after completing email subscription form.**



**A/B Experiment Checklist**

**Before you start, make sure you complete the following:**

1. **Choose the page you would like to test**
   Examples of potential test pages could be your homepage or a product detail page.
2. **Create alternate versions of your test page**
   Create and publish different versions of your test page at unique URLs so that Website Optimizer can randomly display different versions to your us
   want to keep these URLs valid.

**How many alternate versions should you test?**
Generally, we recommend **around 100 conversions per page variation** over the course of your experiment. For example, an experiment with

3. **Identify your conversion page**
   This is an existing page on your website that users reach after they've completed a successful conversion. For example, this might be the page dis

☐ I've completed the steps above and I'm ready to start setting up my experiment.

Once you have completed all three of the steps in the checklist, you check the box above that says, "I've completed the steps above and I am ready to start setting up my experiment."

Next, you will follow the instructions inside of Website Optimizer interface to name your experiment, identify the pages you want to test and designate your conversion page. After this step, you will have to install and validate the JavaScript tags that will be placed in the source code of your Web pages. This process is similar to placing JavaScript tags for Google Analytics. Once the tags are installed, you can check the status of the test implementation inside of Website Optimizer. Once the installation is complete and correct, you will then have time to review and edit the variations you have created.

When you are ready for the experiment to start, you will click on the button in Step 4 to begin the experiment. Although data collection will begin immediately, it may take an hour or so for results to appear in the Website Optimizer dashboard.

## Do's and Don'ts for A/B Testing

### *Do's*

- Do know the goals of your A/B test. Know exactly what you are testing before you begin. This will make the testing process and the data analysis much easier.
- Do make sure to run the test for the appropriate amount of time. Stopping the test too early can be detrimental because you might have gotten more meaningful results by waiting longer. Ending the experiment too late is not good either. This is because poorly performing variations can cost you conversions.
- Use a calculator to determine exactly how long a test should run. This is a good calculator: http://visualwebsiteoptimizer.com/ab-split-test-duration/

- Do make sure to show repeat visitors the same version. Your tool should have a mechanism for "remembering" what version a visitor has seen. Having a visitor see two different variations will cause misleading results. However, if visitors delete cookies, the testing software will not have any way to "remember" their previous visits.
- Do make your A/B test consistent across the whole Web site. If you are testing a sign-up button that appears on multiple pages, then each visitor should see the same version everywhere. Showing one version on page 1 and a different version on page 2 could also skew the results.
- Do many A/B tests. Each A/B test can only have one of three outcomes: Option A is better than Option B, B is better than A, or no statistical difference between A and B. The key to optimizing your site and increasing your conversion rates is to clearly define which version is better and continual testing. Such testing efforts can give a huge boost to achieving your digital goals.
- Do have a large enough sample size. If your sample size is too small, any data collected will be insignificant and unusable. Ideally, you want to have as many people visiting your site as possible. The larger the sample size, the more data you will collect and the more reliable your results will be. If you find you are not getting a large enough sample of visitors, you may need to run the test for a longer time.

### *Don'ts*

- Always test both versions simultaneously. If you test one version this week and the second version the next week, you are doing it wrong. Don't wait to test the variation until after you have tested the control version. Always split traffic between the two versions.
- Don't conclude your test too early. There is a concept called "statistical confidence" that determines if your results are valid and reliable. Ensuring statistical confidence will prevent you from reading too much into the results if you have only a few conversions for each version. Most A/B testing tools, including Google Website Optimizer, report statistical confidence.
- Don't let your gut feeling overrule test results. The winners in A/B tests are often surprising or unintuitive. On a green themed Web site, a stark red button could be the winner. Even if the red button is not easy on the eyes, don't reject it outright. Your test goal is to find the version that gives a conversion rate, not to judge the site aesthetics. Don't reject the results because of someone's arbitrary judgment about design.
- Don't just focus on big changes. Small changes can make an impact too. This is especially true if you run enough tests over time. While small, incremental changes, can add up to create a huge boost in goal completion.

## Summary

A/B testing is one of the primary tools in any data-driven environment. It will allow you to follow up your gut instincts or assumptions with hard data directly from your users. A/B testing can be used to test almost any aspect of your Web site. If done correctly it can lead to a well performing Web site, either from a revenue or usability standpoint.

# Appendix H.

# Strategies for Social Media Measurement

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com

December 15 , 2011

# Appendix H. Strategies for Social Media Measurement

## Overview

Measurement of social media is becoming increasingly sophisticated and comprehensive. However, many organizations have yet to develop a strategy for measuring, or they are struggling to get theirs properly implemented. This does not need to be the case, but, unfortunately, many organizations are not aware of the options now available to monitor, measure, and track social media.

To create a strategy for social media measurement, we first need to set goals. To set goals, we have to understand the types of conversations that occur in social media. Our users are listeners, responders, and speakers. For example, NIH may post a story on a Facebook profile page. NIH "fans," or users, can do one of three things with the NIH post:

- listen (consume the content),
- respond (comment on it), or
- speak (repost the content for distribution to their contacts).

By understanding the different types of interactions with social media, NIH can set clear goals for how to measure them.

### *Recommendations*

NIH will have to decide whether the goal of a given social media post is for the user to listen, respond, or speak. For example, one goal of posting on NIH's Facebook page could be for users to listen (so, they're directed from the Facebook fan page back to the NIH Web site to consume content). The appropriate way to measure this goal is to create two segments: one for directing traffic from NIH social media efforts and one for all referred social media traffic. Once the segment for direct traffic from NIH social media efforts has been created, NIH can run content-consumption success metrics against the referred social media traffic segment to see how it's performing relative to overall social media traffic.

Goals could include increasing subscriptions (responders), giving users an easier way to access information (listeners), or creating awareness around news and information about NIH (speakers). Even though NIH strives to allow easier access to information on its Web sites, users often want to consume content in the medium of their choice. By correctly measuring how users are interacting with its social media, an IC will be able to better understand whether it's meeting its goals by meeting users' needs.

To create segments of social-media-referring traffic, start by looking at your top referring sites and find out which social media sites already drive traffic to your site. Then, augment this list with other key social media sites. For example, poll your colleagues to determine which social media sites your IC is currently using or planning to engage in.

Once you have created the list of social media sites you want to track, you need to make sure you're adding campaign identifiers to your current social media initiatives. For example, if you have a Facebook page with X number of followers and you are posting links to your Web site on your profile page, you will want to add a campaign variable at the end of the link, like this http://www.examplesite.com/article.html?utm_source=facebook&utm_medium=wallpost&utm_campaign=NIHarticle. Google has a very valuable tool called Google URL builder, which allows you to build links and append campaign variables (in this example, campaign source is equal to Facebook, campaign medium is equal to wallpost, and campaign name is equal to NIHarticle) (http://www.google.com/support/googleanalytics/bin/answer.py?answer=55578). This gives you a way to attribute traffic that came from the IC's Facebook profile, which is an important step toward gauging the success of your social media initiatives.

### Social Media Metrics on Your Website

The key metrics to focus on when you are looking at social media and their effects on your Web site are similar to the metrics you use on other marketing campaigns, such as:

- page views/visit,
- time on site,
- content consumed, and
- bounce rate.

The most important metric of all, though, will be about whether IC goals are being met on the Web site. You can measure this for both segments of your social media traffic, from your social media channels and from social media in general.

#### *Recommendations*

Segment your traffic into those two social media segments: social media visits attributed to your efforts and all other general social media traffic. You can only do this if you set up the correct social media campaign variables discussed in the previous section. Analyze those segments against your current success metrics and goals as measured by your Web analytics tool. You can then also segment your two baseline segments by dividing them up according to social medium channel (Facebook, Twitter, YouTube, etc.). Once you have created these segments and subsegments, analyzing your social media traffic becomes simpler. For example, you will be able to follow the traffic patterns of each segment to see how they are performing over time. You can compare Twitter with Facebook traffic to see which one has a better goal-completion rate **Figure H.1**.

***Figure H.1.*** *Comparison of Facebook and Twitter goal-completion rates.*

| Traffic Type | Goal -Completion Rate |
|---|---|
| General Facebook | 35% |
| NIH-specific Facebook | 45% |
| General Twitter | 28% |
| NIH-specific Twitter | 36% |

In this case, Facebook does a better job than Twitter in both overall and NIH-specific campaigns. However, NIH-specific Twitter accounts do a better job of driving users to complete goals than the general Facebook traffic. This shows that Twitter is clearly useful for driving users to the NIH site.

A good way to think about measuring social media is through a three-step process:

1. set a goal,
2. set a strategy, and then
3. measure the outcomes.

Here is an example using email subscriptions:

- Goal: Increase email subscriptions.
- Strategy: Increase tweets and posts highlighting the advantages of becoming an email subscriber.
- Measurement: Analyze segmented social media traffic in order to calculate the goal-completion rate for email subscriptions. This will tell you how effective your push for email subscriptions was through each social medium.

Splitting out goal, strategy, and measurement means that you can look a little more closely at *why*, *how*, and *how effective* your social media efforts are.

### Social Media Metrics on the Social Media Site

Along with measuring the traffic referred to your site from social media sites, we also recommend that you measure how your organization is being described and discussed on social media sites throughout the Internet. Many of the available social media analytics tools from radian6, the high-end social-media-tracking software, to HootSuite allow you to bring Facebook Insights into their software solutions. Some social media tools also allow you to bring in Google Analytics data, including URL parameters (campaign coding) into their tools.

#### *Recommendations*

We recommend that you use these tools to track and measure the 10 key points of social media data across the Internet.

1. Support Social Media Workflow – Social media work is like a marathon, not a sprint. The only way to gain value from it is through repeated, regular involvement. To make this happen, a social media workflow needs to be developed and followed in accordance with the governance plan your organization has already instituted around social media.

2. Social Campaign Measurement – Social media can be used as a campaign channel to drive traffic back to your site. Using URL shorteners or campaign codes in your tweets, Facebook posts, and YouTube videos allows you to measure the impact of your social media efforts.

3. Nonsocial Campaign Measurement – Not all social media traffic can be attributed directly to your efforts. Sharing of content by users can also be a leading driver of traffic to your site. Being able to differentiate traffic from your social media efforts from traffic from all social media is critical to understanding the full impact of social media.

4. Crisis Alerts – Not all mentions of your brand in social media may be positive. It is important to track both trends and sentiment of the social media conversations taking place across the Web so you can be proactive when a wave of negative feedback is headed your way.

5. Brand Monitoring – "If you exist, then someone's already talking about you." Being able to monitor the conversations occurring in the social media space around your brand is extremely valuable. Monitoring the reputation of your brand online is an important aspect of any online marketing campaign today.

6. Social Channel Targeting – Specific messages require a specific audience, and certain conversations are conducted more effectively in certain channels. It is important that you target the different social channels according to the message you are trying to convey to that particular audience at that particular time.

7. Influencer Identification and Targeting – In the social media space, there is a sphere of influencers you may want to target based on demographics, geography, or expertise. It is important to be able to see who your influencers are and to join their conversation.

8. Content and Product Analysis – Social media are great for sharing and spreading your content to your loyal followers. Analyzing how your followers consume and spread your content can help you expand and improve your social media presence.

9. Competitive – Your competitors are also using social media to reach their audiences. Knowing how your reach and influence rank among your competitors' can be helpful as you make decisions about your site.

10. Consumer Research – Understanding your social media users better will allow you to attract and retain more followers. The more you understand your current followers, the more insight you will have into the type of follower you should attract to grow your audience.

### Summary

Any type of measurement strategy has to have a good foundation if it is going to work correctly. Setting goals is the most important step in creating your social media measurement strategy. These goals need to focus on the types of users you are trying to reach: listeners, responders, or speakers. Once the goals have been set, deciding how to meet those goals allows you to measure your social media efforts more effectively, on both your Web site and social media sites.

# Appendix I.
# Evaluation of Web Analytics and Measurement Tools

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

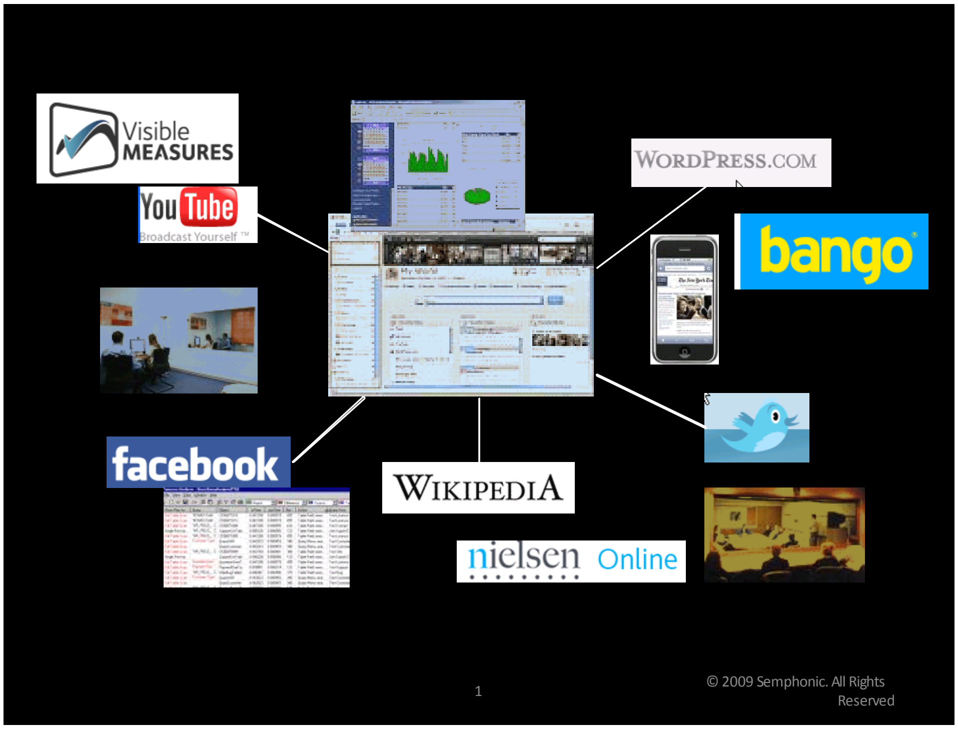
## Appendix I. Evaluation of Web Analytics and Measurement Tools

### The Challenge of Moving from Online to Digital

Until the last few years, NIH had been living in a relatively stable online world since the rise of the Web. Your Web site has been at the center of your online strategy. Return on investment, visitor satisfaction, and Web site effectiveness from Web analytics software, such as Omniture and Google Analytics, online surveying tools, such as Foresee and OpinionLab, and competitive intelligence services, such as Comscore and Hitwise, composed the measurement arsenal. Focus groups and user experience testing rounded out the entire measurement and evaluation tool box.

In the last few years, thanks to social media vehicles, such as blogs, Twitter, YouTube, and Facebook, and the increasing use of video content and mobile Web, the breadth of digital channels has expanded rapidly. (**See Figure I.1.**) So, while the concept of return on investment, visitor satisfaction, and Web site effectiveness might have been limited to your single Web site in the past, the explosion of new digital media has changed that. The result is that now, your Web site is the virtual "home" for your entire organization. All online and offline methods of contact and communication come together on your organization's Web site. With this convergence comes the increasing challenge of harnessing the visitor data available through these channels so that you can make decisions about how to run the growing complexity of your digital initiatives.

***Figure I.1.*** *Moving from Online to Digital*



When online analytics focused predominantly on the fixed Web, you had some basic choices to make about data collection (log file vs. page tag) and hosting model (licensed vs. subscription). Today, as NIH ICs plan and implement measurement of the complete digital-asset environment, they need to consider data-collection and hosting models, as well as data analysis and maintenance, of up to seven different data sources: online, social media, mobile, video, survey, market research/ Internet Industry Research intelligence, and offline data. This number of data sources could easily grow to include email marketing, call-center data, asynchronous online surveys, A/B testing, and search engine optimization (SEO). If these data sources are not managed by a central group, the chances of coordinating and managing data to obtain a holistic view of visitor experience and interaction are small.

Whether there is a centralized NIH analytics structure or analytics structure within individual ICs, there must be some coordination of tool sets in order to accommodate the metrics consistency and baseline reporting framework discussed in the metrics section of the NIH Best Practices for Web Analytics document (beginning on page 26).

The information provided in this appendix is meant to serve as a resource to support selection and decision-making about measurement and analysis tools. We selected these tools to provide options based on functionality and price so that ICs can determine the right “fit” based on budgets and requirements.

We suggest that the contents of this appendix be made into a wiki or similar reference to allow NIH staff and contractors to provide ongoing comments about their experiences with the tools. This will foster cooperative environment among those who use analytics and measurement tools at NIH.

## Web Analytics

### *Introduction*

Web analytics tools facilitate the collection, measurement, and reporting of Web data. The insights gained through using these tools help government users understand how visitors interact with their brand online and thereby drive promotion and outreach efforts and other key business decisions. Typically, Web analytics tools are the central hub of a digital-measurement strategy, with all other tools providing supplemental information.

**Tools Covered in This Section**

- Adobe Omniture Suite
- Google Analytics
- Urchin
- IBM/Unica NetInsight
- IBM/Coremetrics
- Webtrends
- ComScore/Digital Analytix
- Yahoo Web Analytics

### *Adobe Omniture Suite*

**SaaS Solution**

www.omniture.com

**Tool-Use Case**

- JavaScript-based suite of Web analytics tools that allow for highly customized tracking.
- Uses a combination of on-page coding with a central, cached JavaScript file that must be present on every page.
- Depending on contractual options, powerful analysis tools are available that allow for complex data manipulation and segmentation.
- Out of the box, provides basic data on traffic sources and visitor profile. With minor setup, provides more comprehensive data.
- Capable of handling large sites and multiple domains and subdomains.

**Tool Pros**

- **Highly customizable** – contract includes high number of customizable traffic variables, conversion variables, and success events.
- Capable of handling sites of any size.
- Additional tools within the suite allow for more complex analysis capabilities and are available on a contractual basis if deemed necessary.

**Tool Cons**

- Expensive.
- Requires significant technical and strategic up-front investment to use to its full capabilities.
- User Interface is complex and not entirely intuitive.

**Pricing**

The Omniture Suite is priced in two ways. First, the specific tools and capabilities are priced on a contract basis (e.g., which analysis tools, how many custom variables, etc.). Second, once the tool is implemented, there is a cost per server call, so there is a direct correlation between the volume of data collected by a site and the cost of the tool.

Contact Omniture for details.

**Recommendations**

- **Type of Site:** The Omniture Suite is intended for enterprise-level clients. It is most cost-effective when used by sites with high volumes of traffic, multiple domains and subdomains that need to be tracked, and complicated success models. Additionally, the suite features powerful analysis tools and thus should be leveraged on sites that will be benefit from regular analysis projects.
- **Organizational Support:** During the implementation phase, the Omniture Suite requires a significant amount of time for planning as well as dedicated, capable technical resources to code

for it. As a result, implementation is a full-time job on both the strategic and technical side for at least a few months.

Once an implementation has been completed, the data collected are only truly useful if there is a trained staff member who understands how to use the tool as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion . For example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *Google Analytics*

**SaaS Solution**

www.google.com/analytics

#### Tool-Use Case

- JavaScript-based suite of Web analytics tools.
- Tags must be present on every page.
- Provides a range of easy-to-use reports along with more advance and configurable options for segmentation, campaign reporting, and basic analysis.
- Provides significant out-of-the-box traffic and management reporting.
- Capable of handling most sites but samples data for reports on larger sites.

#### Tool Pros

- **Very attractive and easy-to -use interface.**
- Some nice turnkey "analysis" functions for highlighting key changes.
- Free and easy to install and configure.
- New additions provide customizable variables and excellent visit-level segmentation.

#### Tool Cons

- **Sampling and fairly tight limits on distinct values in a single variable.**
- Limited data integration, no ability to provide an event-level data-feed.
- **No visitor-level segmentation and limited variable customization.**
- **No service level agreement (SLA) and no direct vendor support.**

#### Pricing

Google Analytics is free in most cases.

#### Recommendations

**Type of Site:** Google Analytics is intended for a very broad range of sites. It is fully appropriate for most micro-sites and it can handle many medium to large sites in the right circumstances. It is particularly appropriate for organizations with many sites seeking a friendly, easy-to-deploy unified reporting solution. Because of its easy-to-use interface, it is an excellent tool when self-service reporting is one of the fundamental goals. Because of its lack of integration and customer-level data, it is a poor choice when a site needs to support visitor-level analysis or to integrate analytics with other systems.

Organizational Support: With the growth in sophistication of Google Analytics, more attention should be paid to implementations than has been common – particularly in multidomain situations. Most Google Analytics implementations are significantly easier than implementations in other enterprise-class tools. As with all Web analytics tools, however, a full-time Web analyst is highly recommended. Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *Urchin*

**On-Premise Solution**

www.google.com/urchin

#### Tool-Use Case

- On-premise solution that was the original model for Google Analytics.
- Uses log-files in most cases but data can be collected from page tags too.
- The feature set is similar to Google Analytics in most cases. We have noted aspects of Urchin that are different from Google Analytics in the following sections.

#### Tool Pros

- Ability to support log-file-based analytics such as 404 errors, robot reports, etc.
- Provides on-premise solution so there is no external passing of data.

#### Tool Cons

- Requires administrative setup and ongoing tuning as do most log processing systems
- Has lagged behind Google Analytics from user-interface and update perspectives.

#### Pricing

A license costs $10,000.

#### Recommendations

**Type of Site:** Urchin is not nearly as popular as its sibling, Google Analytics. Many organizations prefer a tag-based solution and, of course, Google Analytics is free. However, unlike Google Analytics Urchin can meet the additional requirement of not using page tags or passing data across the Internet.

**Organizational Support**: Urchin, like most log-processing systems, requires more attention to administrative setup than do tag-based solutions. Failure to maintain careful governance on a log-processing solution will nearly always result in poor data quality. Since the interface is now like Google Analytics, this tool requires less training for end users than solutions like Omniture and NetInsight.

### *IBM/Unica NetInsight*

**On-Premise**

www.unica.com

**Tool-Use Case**

- JavaScript-tagging AND log processing Web Analytics solution. Can even be used in hybrid mode with logs and tag data combined.
- Provides a single, powerful interface for most Web analytics tasks.
- Provides significant out-of-the-box reporting and advanced segmentation.
- Capable of handling large sites and multiple domains.

**Tool Pros**

- Single interface to support all users.
- Powerful and unlimited segmentation.
- Option to support tags and log processing provides high flexibility in deployment.
- Logical tagging setup provides easy implementation.
- Unlimited customization of variables and events.
- Ability to integrate visitor-level data.

**Tool Cons**

- Clunky interface that is more difficult to use and feels older than most competitors.
- Poor Excel integration.
- Limited administrative flexibility when deployed in SaaS configuration.
- Performance and latency can be an issue on larger data sets.

**Pricing**

Fairly expensive. Unica contracts are typically based on Cost per Million, or CPMM. CPMM costs are similar to Omniture’s costs, but the total cost is generally cheaper because there are many fewer add-ons.

**Recommendations**

**Type of Site:** NetInsight has been largely back-burnered by IBM, and it will be consolidated with the Coremetrics product. Because of this, we do not recommend NetInsight for organizations shopping for a solution.

**Organizational Support:** NetInsight’s tagging is logical and well thought out, making initial deployment easier than with some other systems. Administration of the on-premise solution is complex and demanding, definitely requiring at least a full-time employee and regular attention to both performance tuning and data-quality governance. Because the interface is complex, training end users on NetInsight can be more challenging than on most competitor solutions.

### *IBM/Coremetrics*

**Tag-Based SaaS Solution**

www.coremetrics.com

**Tool-Use Case**

- Provides a general-purpose reporting system as well as (additional cost) a more advanced segmentation and data-exploration system (Explore) for analysts.
- Out of the box, provides basic data on traffic sources and visitor profile. With minor setup, provides more comprehensive data than other options.
- Capable of handling large sites and multiple domains.
- Provides benchmarking data for some industries.

**Tool Pros**

- Excellent out-of-the-box attribution analysis.
- Robust suite of tools including several that tie directly to marketing optimization systems (Display and Testing).
- Structured approach to data model based around the customer.

**Tool Cons**

- Without Explore, it lacks sufficient segmentation and analysis capabilities. Explore limits nonsampled queries; this sometimes limits analysis.
- Explore provides fewer variable customizations than Omniture, Unica, or Comscore.
- Tagging is a bit more cumbersome and less flexible than with some other systems.
- Generation of automated and customized reports is frequently a "per-charge" item.

**Pricing**

Fairly expensive. Coremetrics contracts are typically Cost per million or CPMM-based. CPMM costs are similar to Omniture but the Total Cost is generally cheaper a little cheaper.

**Recommendations**

**Type of Site:** Coremetrics has been particularly strong in retail and business to business (B2B) settings. Where the company’s products have been popular, the system also benefits from the addition of useful competitive benchmarking. In general, Coremetrics reporting is strongest in ecommerce and campaign analysis. It is an appropriate solution for any ecommerce site —even the largest—and provides a considerable range of suite options for expansion.

**Organizational Support:** During the implementation phase, Coremetrics requires a significant amount of time for planning as well as dedicated, capable technical resources to code for it. As a result, implementation is a full-time job on both the strategic and technical side for at least a few months. Most companies choose to use Coremetrics Professional Services, and there are, in fact, fewer alternative support options than with Omniture or Google Analytics. Once an implementation has been completed, the data collected are only truly useful if there is a trained staff member who understands how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

## *Webtrends*

### On-Premise and Tagging Solution

www.webtends.com

### Tool-Use Case

- JavaScript-tagging AND log-processing Web Analytics solution.
- Provides a very attractive, user-friendly basic dashboard interface.
- Provides a powerful ad hoc reporting capability in a second interface.
- Capable of handling large sites and multiple domains.

### Tool Pros

- The dashboard interface is very well laid out and delivers highly intuitive views of the data
- An open system provides excellent integration opportunities and combines several public data sources out-of-the-box
- The full suite contains several powerful tools for warehousing, testing, social media, etc.

### Tool Cons

- The advanced reporting system is clunky and somewhat dated.
- On-premise solutions have lagged significantly behind the SaaS options
- Not all the pieces of the suite are closely integrated

### Pricing

Fairly expensive. Webtrends contracts are typically CPMM-based. CPMM costs are similar to Omniture's, but the total cost is generally cheaper because there are many fewer add-ons.

### Recommendations

**Type of Site:** Webtrends has invested heavily in social media reporting and open integration in the last few years. In addition, the Webtrends-management-reporting interface is the nicest on the market. This makes Webtrends an excellent choice for organizations concentrating on management reporting and cross-channel views of the data.

**Organizational Support:** Administration of the on-premise solution is complex and demanding, definitely requiring at least a full-time employee and regular attention to both performance tuning and data-quality governance.

Because the ad hoc interface is complex, training power users on the full reporting solution is a must.

Once an implementation has been completed, the data collected are only truly useful if there is trained staff who understand how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who  will use the tool in a more particular fashion. For example,  infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *Comscore Digital Analytix*

**Tag-Based SaaS Solution**
www.comscore.com

#### Tool-Use Case

- Provides a single, powerful interface for reporting and analysis.
- Out of the box, provides a rich set of standardized reports, unlimited segmentation, and the ability to customize reports as necessary.
- .
- Provides access to basic demographic data (age and gender) as well as collected behavioral data

#### Tool Pros

- Single interface to support all users.
- Aggregates all queries from detail data - this means that the reporting is significantly less limited than it is for tools that pre-aggregate the data.
- Near real-time data access.
- Access to demographic data in addition to the behavioral data.
- Can piggy-back on existing comScore tags.

#### Tool Cons

- Some user interface (UI) issues - especially for the U.S. market.
- Limited usage to date in the U.S. market, with uncertain adoption for the future.

#### Pricing

Fairly expensive. Comscore contracts are typically CPMM-based. CPMM costs are similar to Omniture's.

### Recommendations

**Type of Site:** Digital Analytix is a fairly new entrant into the Web analytics field. It is a powerful system, particularly in its ability to drive reporting and analysis from event-level, near-real-time data. It is particularly appropriate for media sites (where Comscore tagging is likely already in place) and for sites where near-real-time advanced reporting is a primary concern.

**Organizational Support:** Comscore's tagging is logical and well thought out, making initial deployment easier than on some other systems. However, the system requires significant customization (providing unlimited variables and segmentation).

Once an implementation has been completed, the data collected are only truly useful if there is trained staff who understand how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### Yahoo Web Analytics (YWA)

**Tag-Based SaaS Solution**
web.analytics.yahoo.com

*Tool-Use Case*

- JavaScript-based suite of Web analytics tools.
- Tags must be present on every page.
- Provides a range of easy-to-use reports along with more advanced and configurable options for segmentation, campaign reporting, and basic analysis.
- Provides significant out-of-the-box traffic and management reporting.
- Most appropriate for small-to-medium sized sites.

*Tool Pros*

- Free and easy to install and configure.
- Excellent out-of-the-box segmentation capabilities.
- Good integration with Yahoo Store.

*Tool Cons*

- Limited data integration, no ability to provide an event-level data-feed.
- No service level agreement (SLA) and no direct vendor support

*Pricing*

Yahoo Web Analytics is free in most cases.

*Recommendations*

**Type of Site:** YWA is similar in many respects to Google Analytics . Its interface is less intuitive and it lacks some of Google Analytics'niftier analysis features. On the other hand, it has always provided significantly better segmentation (a key analytics capability) than Google Analytics. It is fully appropriate for most micro-sites, and it can handle many medium sites in the right circumstances. It is particularly appropriate for organizations with close ties to the Yahoo Store. Like Google Analytics, it is an excellent tool when self-service reporting is one of the fundamental goals. Because of its lack of integration and customer-level data, it is a poor choice when a site needs to support visitor-level analysis or to integrate analytics with other systems. Unlike Google Analytics, it can be difficult to find YWA support and the range of third party expertise is much more limited since YWA has a much smaller market footprint than Google Analytics.

**Organizational Support:** Most YWA implementations are significantly easier than implementations in other enterprise-class tools.

As with all Web analytics tools, however, a **full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

## Mobile Analytics

### *Introduction*

Mobile Analytics tools collect, measure, analyze, and report data for Web content specifically accessed via mobile devices, such as smart phones. As consumers increasingly demonstrate a reliance on, and even preference for, mobile technology for accessing Internet-based content, it has become imperative to include these tools, in tandem with traditional Web data-collection methods, in order to develop a complete and accurate portrayal of Web-visitor behavior.

### *Overview*

- Marketplace is divided between Niche and Enterprise analytics service providers.
- Niche analytics firms focus on the mobile industry and are typically more adept at identifying devices and unique visitors.
- Enterprise analytics firms have depth of experience in analytics and can often provide "holistic" view of fixed/mobile/app activity.
- Niche vendors tend to focus only on mobile Web or mobile application measurement, not both.

**Tools Covered in This Section:**

- Adobe Omniture Suite
- AppClix
- Bango
- Google Analytics
- Localytics

- PercentMobile
- Webtrends

### *Adobe Omniture Suite*

**SaaS Solution**

www.omniture.com

**Tool-Use Case**

- Enterprise analytics vendor.
- Depending on contractual options, powerful analysis tools are available that allow for complex data manipulation and segmentation.
- Out of the box, provides basic data on traffic sources and visitor profile. With minor setup, provides more comprehensive data.
- Capable of handling large sites and multiple domains and subdomains.
- Offers both mobile-Web and mobile-application measurement.
- Mobile Web – data-collection methods:
  - standard JavaScript page coding per the fixed Web
  - serve- side image requests leveraging a mobile image beacon
  - libraries of tracking code for PHP and Java for client server-or-server to server data collection (non-JavaScript) applied directly to the Web page.
- Mobile application - data-collection methods:
  - system development kit (SDK) available for iPhone, Android, Blackberry, Windows Phone, and Symbian
  - open- source( SDK)– customization of tracking code
  - no app store monetization reporting.

**Tool Pros**

**Robust management of Visitor ID**— Customer managed > subscriber ID > cookie > user authentication (UA) and IP address.

- **Highly customizable** – Contract includes high number of customizable traffic variables, conversion variables, and success events.
- **Mobile application "plug-in"** allows for best practices data capture of custom variables with little coding effort.
- Capable of handling **sites of any size**.
- Additional tools within the suite allow for more **complex analysis capabilities** and are available on a contractual basis if deemed necessary.

**Tool Cons**

- **Moderate level of device details** and no device -level information combinations are available unless using Site Catalyst v15.
- **Campaign tracking** for mobile image beacons requires custom coding which is non-existent for mobile applications.
- **Expensive.**
- Requires **significant technical and strategic up-front investment** to use to its full capabilities
- **User Interface is complex** and not entirely intuitive.

**Pricing**

The Omniture Suite is priced in two ways. First, the specific tools and capabilities that are purchased are priced on a contract basis (e.g., which analysis tools, how many custom variables). Second, once the tool is implemented, there is a cost per server call, so there is a direct correlation between the volume of data collected by a site and the cost of the tool.

Contact Omniture for details.

**Recommendations**

**Type of Site:** The Omniture Suite is intended for enterprise-level clients. **It is most cost-effective when used by sites with high volumes of traffic, multiple domains and subdomains that need to be tracked, and complicated success models.** Additionally, the suite features powerful analysis tools and thus should be leveraged on sites that will be benefit from regular analysis projects.

Clients requiring a **consolidated view across all channels** (fixed Web , mobile Web, and mobile applications) would benefit from Omniture's comprehensive mobile measurement capabilities.

**Organizational Support:** During the implementation phase, the Omniture Suite requires a significant amount of time for planning as well as dedicated, capable technical resources for coding. As a result, **implementation is a full- time job on both the strategic and technical side for at least a few months.**

**The Omniture SDKs require dedicated developers** experienced in mobile application development frameworks (iOS, Android, Blackberry, Symbian, C#, etc.).

Once an implementation has been completed, the data collected is only truly useful if there is trained staff who understands how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *AppClix*

**Server-Based Solution**

www.appclix.com

### Tool-Use Case

- Niche analytics vendor – iOS (Apple's mobile operating system) ONLY.
- Out of the box, provides basic reporting data on traffic sources and visitor profile. With minor setup, provides ability for custom event tracking with parameters and values
- Mobile application - data-collection methods:
  - available for iOS
  - open source SDK – customization of tracking code
  - real time data collection
  - display orientation (iPad)
  - data Export API
  - full app store monetization reporting
  - campaign and referral tracking
  - off line data capture in batch mode.

### Tool Pros

- **iTunes integration** allows for revenue and sales-conversion tracking as well as reviews and rankings.
- **Not SaaS** – server-based data storage for data privacy/ownership.
- Capable of handling **sites of any size** by adding additional servers.

### Tool Cons

- **iOS (Apple mobile applications system) only**
- **No segmentation or advanced reporting interface.**

### Pricing

AppClix is priced three ways:

- single server per app license: 1 app ($299 per month) – unlimited ($599 per month);
- enterprise server license allowing for multiple server installations $1,999; and
- CloudClix - complete turnkey solution that resides on a server on the Amazon EC2 Cloud: 1 app ($149/mo) – unlimited ($299/mo).

### Recommendations

**Type of Site:** The AppClix solution is intended for iOS measurement only. **It is a standalone solution for tracking individual iOS applications.** Appropriate for sensitive data storage because it is a **product not a SaaS solution** and all data collected are owned and stored by the client

**Organizational Support:** Implementation is relatively straightforward and requires adding a small class to your existing iOS application. Migration from other mobile analytics platforms , including Flurry and Localytics, is simplified because interfaces are prebuilt to match these vendors.

Once an implementation has been completed, the available reporting is simple enough that a **full-time Web analyst is not required.** However, stakeholders who  will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training.

### *Bango*

**SaaS Solution**

www.bango.com

**Tool-Use Case**

- Niche analytics vendor.
- Out of the box, provides basic reporting data on traffic sources and visitor profiles. No custom event tracking or segmentation.
- Offers both mobile Web and mobile application measurement.
- Mobile Web data-collection methods:
    - server side image requests leveraging a mobile image beacon
    - libraries for PHP, JSP, and ASP.NET to generate image requests (non-JavaScript) ,based on applying tags directly on the Web page
    - link redirects used for outbound tracking.
- Mobile application – data-collection methods:
    - SDKs available for iPhone, Android, Blackberry, and Windows Phone
    - open source SDK – customization of tracking code
    - no app store monetization reporting.

**Tool Pros**

- **Strong carrier relations** due to existing mobile billing product allows for **enhanced data capture and augmentation.**
- **Custom visitor ID – "Bango ID"** formed from subscriber ID > Cookie > UA and IP.
- Data exports available via APIs.

**Tool Cons**

- No custom events – users must leverage query string (campaign) parameters (10).
- **No segmentation or advanced-reporting interface.**

**Pricing**

Bango is priced based on server calls ($49 per month 100K and $499 per month for 1 million). Enterprise solutions are also available that offer additional volume and additional data API access.

**Recommendations**

**Type of Site:** The Bango solution is intended for clients who do not want to use a JavaScript-based solution for mobile Web tracking and who would benefit from an enhanced Visitor ID methodology and carrier data augmentation.

**Organizational Support:** Implementation is relatively straightforward however developers will need coding experience in the language selected (PHP, JSP, ASP.NET) in order to establish the server side libraries and base tracking.

**The Bango SDKs require dedicated developers** experienced in mobile application development frameworks (iOS, Android, Blackberry, and C#)

Once an implementation has been completed, the available reporting is simple enough that a **full-time Web analyst is not required.** However, stakeholders who will use the tool in a more particular fashion – i.e., infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training.

### *Google Analytics*

**SaaS Solution**

www.google.com/analytics

**Tool-Use Case**

- JavaScript-based suite of Web analytics tools.
- Tags must be present on every page.
- Provides a range of easy to use reports along with more advanced and configurable options for segmentation, campaign reporting, and basic analysis.
- Provides significant out-of-the-box traffic and management reporting.
- Capable of handling most sites but samples data for reports on larger sites.
- Enterprise analytics vendor.
- Offers both mobile Web and mobile application measurement.
- Mobile Web –data-collection methods:
  - standard JavaScript page coding as per the fixed Web
  - libraries for PHP, JSP, ASP.net, and Perl for client/server data collection (non-JavaScript).
- Mobile application-data-collection methods:
  - SDKs Available for iOS and Android
  - open source SDK – customization of tracking code
  - no app store monetization reporting.

**Tool Pros**

- **No cost.**
- **Very attractive and easy-to-use interface.**
- Some nice turnkey “analysis” functions for highlighting key changes.
- Free and easy to install and configure.
- New additions provide customizable variables and excellent visit-level segmentation.

**Tool Cons**

- **Cannot run both versions of the tracking code (non-JavaScript and JavaScript) on the same page.**
- **Low level of device details** and limited device level sub relations.
- **Campaign tracking** for mobile image beacons requires custom coding and non-existent for mobile applications.
- **Mobile application tracking Requires significant technical and strategic up-front investment** to use to its full capabilities.
- **Data sampling that occurs without notice.**
- Limited data integration, no ability to provide an event-level data-feed.
- **No Visitor-Level segmentation and limited variable customization.**
- **No SLA and no direct vendor support.**

**Pricing**

Free.

**Recommendations**

**Type of Site:** The Google Analytics solution is intended for clients who want to maintain congruence with their existing Google Analytics implementations from the fixed Web and are looking to leverage existing investment in measurement strategy and coding. Additionally, clients who do not want to leverage a JavaScript-based solution for mobile Web tracking would benefit from a Google Analytics mobile implementation.

**Organizational Support:** Implementation is relatively straightforward however developers will need coding experience in the language selected (PHP, JSP,  and ASP.NET) in order to establish the server side libraries and base tracking.

**The Google SDKs require dedicated developers** experienced in mobile application development frameworks (iOS and Android).

Once an implementation has been completed, the data collected is only truly useful if there is trained staff who understands how to use the tool, as well as the nuances of data collection. **A full-time Web**

**analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *Localytics*

**SaaS Solution**

www.Localytics.com

**Tool-Use Case**

- Niche analytics vendor.
- Mobile application - data-collection methods:
  - available for iPhone, Android, Blackberry, Windows Phone
  - open source SDK – customization of tracking code
  - no app store monetization reporting
  - community Edition – Free
  - enterprise – Per app user fee Available for iOS (Apple's mobile platform).

**Tool Pros**

- **Custom event tracking.**
- Simple implementation **with real time data-collection.**
- Display orientation (iPad).
- **Data Export API available at Enterprise level with full session and user level details.**
- **Advanced segmentation capabilities (no need to define segments in advance).**

**Tool Cons**

- **Reporting interface somewhat limited.**

**Pricing**

Localytics is priced three ways:

- Community edition, free.
- Premium edition, $95 per month per application and includes real-time reports and email support.
- Enterprise edition - $895 per month for unlimited apps and includes unlimited event tracking and full data exports along with phone support and consulting services.

**Recommendations**

**Type of Site:** The Localytics solution is intended for **mobile application measurement** only and is **well suited to large enterprises** seeking **detailed segmentation and analysis capability** for their applications. The **data export capability at the user session level** allows for deep **integration with Web, customer relationship management (CRM)**, accounting, and advertising systems via data warehousing options.

**Organizational Support: I**mplementation is relatively straightforward and requires adding a small class to your existing mobile application.

**The Localytics SDKs require dedicated developers** experienced in mobile application development frameworks (iOS, Android, Blackberry, and C#)

Once an implementation has been completed, the data collected is only truly useful if there is trained staff who understands how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

### *PercentMobile*

**SaaS Solution**
www.percentmobile.com

**Tool-Use Case**

- Niche analytics vendor.
- Mobile Web - Data Collection Methods:
  - Standard JavaScript page tagging & Mobile image beacon
  - Libraries for PHP, ASP.NET, ColdFusion, Ruby, Drupal, and WordPress plugins
  - Available for iPhone, Android, Blackberry, Windows Phone
  - Open source SDK – customization of tracking code
  - No app store monetization reporting
  - Community Edition – Free
  - Enterprise – Per app user fee Available for iOS.

**Tool Pros**

- **Extensive device details with advance device segmentation capabilities and graphical views.**
- Simple implementation.

#### *Tool Cons*

- **No event tracking.**
- **Visitor ID limited to cookie based only (no Subscriber ID).**

#### *Pricing*

PercentMobile is priced two ways: 1) Starter Edition, free (100K server calls) and 2) Professional Editio, - $99 per month (extended calendar, export, saved reports).

#### *Recommendations*

**Type of Site:** The PercentMobile solution is intended for **mobile Web measurement** only and is **well suited to large enterprises** seeking **detailed device segmentation and analysis capability**.

**Organizational Support:** Implementation is relatively straightforward however developers will need coding experience in the language selected (PHP, JSP, ASP.NET) in order to establish the server side libraries and base tracking.

Once an implementation has been completed, the available reporting is simple enough that a **full-time Web analyst is not required.** However, stakeholders who will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training.

### *Webtrends*

**SaaS Solution**

www.Web trends.com

**Tool-Use Case**

- Enterprise analytics vendor.
- Depending on contractual options, powerful analysis tools are available that allow for complex data manipulation and segmentation.
- Out of the box, provides basic data on traffic sources and visitor profile. With minor setup, provides more comprehensive data.
- Capable of handling large sites and multiple domains and subdomains.
- Offers both mobile Web and mobile application measurement.
- Mobile Web – Data Collection Methods:
  - Standard JavaScript page coding as per the fixed Web
  - Server side image requests leveraging a mobile image beacon
  - Libraries for REST based data collection API (server to server & client side).
- Mobile Application - Data Collection Methods:
  - SDKs Available for iPhone, Android, Blackberry, and Windows Phone
  - Open source SDK – customization of tracking code
  - No app store monetization reporting.

**Tool Pros**

- **Robust management of Visitor ID -** Customer managed > Cookie > Webtrends Unique ID > User Authentication and IP address.
- **Privacy features** (visitor ID privacy options & opt-in/out support).
- **Highly customizable** – contract includes high number of customizable traffic variables, conversion variables, and success events.
- **Mobile Application 'Plug-in'** allows for best practices data capture of custom variables with little coding effort.

- Capable of handling **sites of any size**.
- Additional tools within the suite allow for more **complex analysis capabilities** and are available on a contractual basis if deemed necessary.

**Tool Cons**

- **Moderate level of device details.**
- Campaign tracking for mobile image beacons requires custom coding and non-existent for mobile applications.
- Requires **significant technical and strategic up-front investment** to use to its full capabilitie.

**Pricing**

The Webtrends solution is priced on a cost per server call, so that there is a direct correlation between the volume of data collected by a site and the cost of the tool.

Contact Webtrends for details.

**Recommendations**

**Type of Site:** The Webtrends solution is intended for enterprise-level clients. **It is most cost-effective when used by sites with high volumes of traffic, multiple domains and subdomains that need to be tracked and complicated success models.** Additionally, the solution features powerful analysis tools, and thus should be leveraged on sites that will be benefit from regular analysis projects.

**Organizational Support:** During the implementation phase, the Webtrends solution requires a significant amount of time for planning as well as dedicated, capable technical resources to code for it. As a result, **implementation is a full-time job on both the strategic and technical side for at least a few months.**

**The Webtrends SDKs require dedicated developers** experienced in mobile application development frameworks (iOS, Android, Blackberry, C#, etc.).

Once an implementation has been completed, the data collected is only truly useful if there is trained staff who understands how to use the tool, as well as the nuances of data collection. **A full-time Web analyst is highly recommended.** Additional stakeholders who  will use the tool in a more particular fashion – for example, infrequent users who will log in to run a small, specific set of reports – will likely need some minimal level of training, as well.

## Audio/Video Analytics

### *Introduction*

Interactive media continues to gain an ever-increasingly proportion of Web share, as marketers look for innovative methods of attracting and interacting with their intended audiences. Audio and Video Analytics tools allow for tracking, analyzing, and reporting on how users respond to interactive media content online, and provide marketers with a deep understanding of the effectiveness of their multimedia-based outreach, promotion and marketing campaigns.

**Tools Covered in This Section**

- Visible Measures
- PodTractor

### *Visible Measures*

**SaaS**

http://www.visiblemeasures.com

**Tool-Use Case**

- Hosted video measurement solutions to help organizations maximize return on investment on online video marketing campaigns.
- Using this solution, organizations s can identify the best sites to reach intended audiences, who are watching their video content and their relative performance on key metrics as compared to their competitors.
- Provides a comprehensive view of the video marketplace by tracking video performance data from hundreds of video-sharing sites.
- In addition to tracking general video data, captures in-stream audience behavior and provides audience engagement metrics for specific video content by integrating with site's video player. Moreover, this data is captured across paid (i.e., purchased advertising space), owned (i.e., hosted on branded sites), and earned (i.e., consumer-driven) media.

**Tool Pros**

- Not only provides rich, highly granular details of viewer activity on video marketing campaigns, but also synthesizes this data into easily digestible key metrics. For example, the Video Engagement Curve gives a quick visual representation of the total audience's cumulative interaction with a given video so users can readily identify which part of the video users find most compelling. The Initial Attention metric calculates the percentage of viewer drop-off at the beginning of a given video. Average Attention measures the rate at which viewers stop watching a video. Captivation identifies video "hot-spots" that inspire frequent rewinds and re-watches.
- Supports measurement of video on a wide variety of online video technologies, such as: Adobe Flash, HTML5 Video, Microsoft Silverlight, Apple Quicktime, and DivX.
- Brand advocacy metrics allow users to understand the online community's response to specific video campaign, as well as provide insight to overall sentiment and brand awareness.

**Tool Cons**

- Limited use cases—solution works best for complex video marketing campaigns with specific audiences to target.

- Customize data views limited. While the application does a good job at synthesizing data into easily digestible metrics, users cannot easily analyze data in a manner beyond these defined parameters.

**Pricing**

Contact Visible Measures directly for pricing details.

**Recommendation**

**Type of Site:** The Visible Metrics solution is most appropriate for enterprise-level organizations with large-scale online video marketing campaigns. Both advertisers and content publishers can surface significant value from the application. Additionally, organizations with strong concerns regarding industry competition will benefit from Visible Measures benchmarking and relative performance metrics.

**Organizational Support:** Visible Metrics provides extremely powerful analytics for individual video campaigns, as well as the online video landscape as a whole. While reporting and analysis is relatively straightforward, leveraging data into actionable marketing strategies requires more significant time and personnel investment.

### *comScore VideoMetrix*

**SaaS**

www.comscore.com/Products_Services/Product_Index/Video_Metrix

**Tool-Use Case**

- Third-party application that measures the video marketplace, by tracking video advertising, television programs, viral videos, and syndicated traffic.
- Using VideoMetrix provides a comprehensive view of the video marketplace by tracking video performance data from hundreds of video-sharing sites.
- In addition to tracking general video data, captures in-stream audience behavior and provides audience engagement metrics for specific video content by integrating with site's video player. Moreover, this data is captured across paid (i.e., purchased advertising space), owned (i.e., hosted on branded sites), and earned (i.e., consumer-driven) media.

**Tool Pros**

- Leverage Comscore's robust online data collection system, Unified Digital Measurement. Their methodology combines voluntary panel and census data, which reduces the impact of cookie deletion and rejection.
- Comscore's video data is widely accepted and relied upon for reporting video viewership trends.
- International coverage. Includes data from many regions outside the United States.

**Tool Cons**

- Only appropriate for large organizations.

**Pricing**

Contact comScore directly for a custom price quote.

**Recommendations**

**Type of Site:** VideoMetrix is only appropriate for enterprise-level organizations with large-scale online video marketing campaigns. Both advertisers and content publishers can surface significant value from the application.

**Organizational Support:** VideoMetrix provides extremely powerful analytics data for the video market landscape as a whole. Leveraging data into actionable marketing strategies requires more significant time and personnel investment, and therefore a fulltime analyst or marketer is recommended to manage the data.

### *PodTractor*

**SaaS**

http://podgarden.oneupweb.com/services/podtractor_enterprise.htm

**Tool-Use Case**

- Online, hosted analytics solution that measures, and thereby helps predict, podcast audience behavior.
- Enterprise-level application is intended for large volume podcast producers.

**Tool Pros**

- Provides a rich view of audience behavior through metrics on visitors, downloads, subscribers, loyalty, and trends over time.
- Implementing PodTractor does not require involvement from IT.
- Data is visualized with graphs and charts on a hosted dashboard, enabling users to quickly understand trends and overall audience behavior.

**Tool Cons**

Data are not exportable, which limits potential analysis opportunities.

**Pricing**

Contact PodTractor directly for pricing details.

**Recommendations**

**Type of Site:** PodTractor Enterprise is intended for organizations producing a large volume of podcasts, with specific marketing or branding interests in understanding podcast traffic and listener behavior.

**Organizational Support:** Using PodTractor requires relatively little organizational support. Implementation does not require involvement from IT. Additionally, data is automatically analyzed and visualized with graphs and charts, so the reporting phase is less time-intensive.

## Social Media

### *Introduction*

The landscape of the online space has changed dramatically over the past few years, with consumers now interacting with brands in places other than a single, company-owned Web site. Social networks allow users to discuss and share their thoughts, perceptions and experiences with a given brand. As a result, organizations can use these social networks to gain information about how consumers view their brands, as well as share information with consumers that may influence their perceptions. The tools discussed in this section facilitate these two-way brand-consumer interactions in various ways. Some, for example, provide analytics on social network user behavior, in ways that are similar to Web analytics data collection about an organization's Web site. In contrast, others facilitate the processes of generating and distributing of content onto various social media platforms.

#### Tools Covered in This Section

- Alterian
- bit.ly Statistics
- Google Analytics
- Facebook Insights
- Next Analytics
- Radian6
- YouTube Insight
- HootSuite

### *Alterian*

#### Social Media Monitoring

www.socialmedia.alterian.com

#### Tool-Use Case

- Create multiple profiles, with multiple search terms under each profile. These profiles allow users to capture, filter, and analyze social media conversations.
- Users can respond to social media conversations directly from the console.
- Provides data on key metrics such as: daily conversation volume for a specified brand or keyword, share of voice (i.e., sources with greatest impact on conversations), date comparisons, themes, demographics, and competitors.

#### Tool Pros

- Access to an extremely robust data warehouse with over 10 billion social media mentions, and corresponding data including date of publication and location of poster.

- Extremely flexible reporting, allowing users to view their data through several different lenses.
- Customizable sentiment analysis enables users to quickly understand the overall tone of conversations regarding their brand or campaign.

**Tool Cons**

- Advanced filtering and reporting options can overwhelm a novice user, so there is potentially a steeper learning curve than with other solutions.

**Pricing**

Packages start at $500 per month.

**Recommendation**

**Type of Site:** Alternian's platform offers a high-degree of social monitoring, which allows users to develop a rich understanding of how their brand is discussed and perceived in the social space. For this reason, the tool is most appropriate for large, highly visible organizations. It is only cost-effective for organizations that have a strong social media presence, or the resources to develop one.

**Organizational Support: T**o leverage the full extent of Alternian's social monitoring capabilities, a fulltime analyst or external consultant should dedicate at least half their time to managing campaigns.

### *bit.ly Statistics*

**URL shortener**

http://bitly.com/

**Tool-Use Case**

- Bit.ly is a URL shortened tool with built in tracking. The tool is most commonly used when shortening a link to be inserted into Twitter tweets or other microblogging platforms.
- Less commonly the tool is used to shorten and track links for Facebook or other social media platforms.

**Tool Pros**

- Built-in tracking for each shortened URL includes
  - Referring sites and clicks for your unique bit.ly URL
  - Referring sites and clicks for all bit.ly users that shortened the same target URL
  - Time-based run chart for clicks
  - Pie chart showing clicks by country
  - Twitter conversations w/ the target URL
  - A list of bit.ly users that also shortened the URL.
- API available to export and automate tracking.

**Tool Cons**

- Tracking is transparent and it is possible for anyone to view tracking for any bit.ly URL, regardless if they were the original creator/poster of the URL.

**Pricing**

Bit.ly is free, bit.ly Enterprise costs $995 per month.

**Recommendations**

**Type of Site:** Bit.ly can be used on a wide variety of sites, ranging from individuals tracking link activity from their personal Web content (e.g., Facebook page, Twitter account, or blog) to enterprise-level link tracking and benchmarking. Given its relatively high monthly cost, bit.ly enterprise should only be used by large organizations with large resources and budgets specifically for understanding and increasing their social media presence.

**Organizational Support:** Garnering useful marketing insights from bit.ly's link tracking, even at the enterprise level, requires less maintenance than other Web analytics tools, and therefore can be managed by an analyst with other responsibilities. Bit.ly also offers an open API, so data can be integrated into other analytical platforms, further decreasing the need for organizational support.

### *Google Analytics (Social Plug-in)*

**Social engagement tracking**

*www.google.com/analytics/*

**Tool-Use Case**

- Track social engagement and shared content on your Web site.

**Tool Pros**

- Allows analysts to see the level of site engagement for visitors who interacted with a social media on site compared to visitors that had no social engagement.
- Compares performance across all social media sources featured on your site.

**Tool Cons**

- Narrow scope of tracking since only social media interactions (for example, Facebook "Like" and "Send," and Del.cio.us bookmarks) that take place on your own Web site are tracked by the plug-in

**Pricing**

Free.

**Recommendations**

**Type of Site:** This tool is a plug-in for Google Analytics so it is most useful and therefore recommended for organizations already using Google Analytics. Since it only tracks social interactions stemming from a singular Web site, this tool alone cannot provide a complete depiction of how social media users are interacting with a brand. For this reason, it is recommended that larger organizations with highly-refined social media marketing ambitions should use a more sophisticated option. .

**Organizational Support:** Given the limited capabilities of this platform, little organizational support is required for its maintenance and use. A Google Analytics Web analyst can easily manage this data in additional to his/her regular responsibilities.

### *Facebook Insights*

**Facebook reporting tool**
http://facebook.com/insights

**Tool-Use Case**

- Allows Facebook page owners to access metrics around Facebook content to see a current snapshot or historical trends.

**Tool Pros**

- Provides tracking for
  - Traffic volume and trending
  - Fan volume and trending
  - Referring sources (both external and Facebook internal)
  - Content interactions including likes, comments, wall posts, video views, picture views and votes.
- API available to export/automate tracking

**Tool Cons**

- Data "silo" containing only Facebook tracking.
- Posts are tracked by total impressions and % Feedback, but to get a breakdown of the number of likes, comments, video views, picture views, etc. per post you must leverage the API.
- Difficult to export full text of comments (must be done through API).

**Pricing**

Free.

**Recommendation**

**Type of Site:** Facebook is one of the most popular social networks so understanding how Facebook users interact with an organization's fan pages or platform can help direct media marketing strategies. Moreover, Facebook Insights is automatically enabled for Facebook Page (for pages with 30 or more fans) and Facebook Platform, so it is recommended for any organization using these applications.

**Organizational Support:** Given the simplicity and limited capabilities of this tool, very little organizational support is required for its use and maintenance. Some developer resources may be necessary for organizations that would like to integrate the data with other platforms via Facebook's API.

### *Next Analytics*

**Excel integration tool**

http://excel.nextanalytics.com/

#### Tool-Use Case

- Report building applications for regular, ongoing, automated reporting.
- Exporting social media data to Excel (tool sometimes facilitates pulling more data and/or longer date ranges of data than would be available to export within the regular social media reporting interfaces).

#### Tool Pros

- Contains integrations for Facebook, Twitter, Google Analytics and LinkedIn within a single tool so that it is possible to automate tracking from a single social media source or multiple sources in a single file.
- The tool features a straightforward interface (compared to other Excel integration tools) and is fairly easy to use.
- Data refreshes are reliable, accurate and relatively fast.

#### Tool Cons

- The tool lacks a YouTube integration, however, the ability to automate data out of YouTube is on their product roadmap.
- All data comes out in daily granularity, so it is necessary to handle rolling up to weekly, monthly, quarterly, etc. levels using Excel techniques.
- Top video data within the main reporting console is limited to showing the percent of total views by video; viewing engagement levels by video involves looking into each video one by one.
- The tool lacks an automated "push" feature to refresh and distribute reports, instead a user must manually click refresh and handle distribution.

#### Pricing

Contact Next Analytics for pricing. Each additional license comes at a 10% discount.

#### Recommendations

**Type of Site:** Next Analytics' platform can be extremely useful for understanding how consumers are interacting with a particular brand across multiple social networking platforms. For this reason, it is recommended for organizations with enterprise-level dedicated cross-channel social media marketing campaigns.

**Organizational Support:** Organizations can leverage Next Analytic's platform to develop a deep understanding of their social media presence across several different channels. To ensure maximize return on investment, Next Analytics is most appropriate for enterprise-level businesses with outreach, marketing or promotional campaigns across multiple social media channels (e.g., Facebook, Twitter, and LinkedIn).

### *Radian6*

**Social media monitoring**

http://facebook.com/

#### Tool-Use Case

- A conversation management tool allowing the user to find brand or keyword mentions and directly respond to the conversations within a variety of social media tools including Facebook, Twitter and blogs.
- Analytics tracking tool for understanding brand or keyword mention volume, conversation channels and top influencers.

#### Tool Pros

- Recognized industry leader in social media monitoring.
- Easy to use interface with high quality data visualizations.
- Quick and simple Boolean-logic-based keyword-profile creation wizard allows for monitoring profiles to be set up after limited training and makes the tool appropriate for a broad audience.
- Quick data response time, typically returning keyword.
- Excel integration allows for insertion and updating of data from within Excel.

#### Tool Cons

- Limitations to the quantity of keywords and Boolean logic used to build tracking profiles
- Pricing based on total volume of posts within profiles, which makes broad tracking (including tracking competitors) cost prohibitive and/or substantially more than competitive tools
- Relatively limited analytic/reporting capabilities make this a tool that is not ideal for data analysts, other competitors in the space (i.e. My BuzzMetrics) serve as better "analyst workbenches."

#### Pricing

Price starts at $600 per month and is based on the volume of new, individual posts within a profile. A 50% discount is offered for registered and qualified charitable organizations.

### Recommendations

**Type of Site:** Radian6 is most appropriate for enterprise-level organizations with robust social media-based marketing campaigns.

**Organizational Support:** The level of organizational support required for using and maintaining a Radian6 account is highly dependent on the organization's size and tool use. A marketing professional or Web analyst is recommended to manage a Radian6 account. For a company similar in size to NIH, the minimum time required to maintain profiles and conduct basic reporting is estimated at eight hours per month. More advanced reporting or analytics would require closer to a 20-100 hour per month time investment, while using the tool for social outreach (for example, sending tweet or  posting on Facebook) is a minimum 20-hour-a-month investment.

## *YouTube Insight*

### Social media monitoring

http://www.youtube.com/my_videos_insight

### Tool-Use Case

- Allows YouTube channel owners to access metrics around YouTube videos to see a current snapshot or  historical trends.

### Tool Pros

- Provides quick, easy to read overview of channel as a whole.
- Data can easy be broken down by date range and/or location of users.
- Robust video level metrics including views, referring sources, demographics and location all at the individual video level.
- "Hot Spot" video tracking shows second by second "hot" and "cold" trend to identify video retention/drop off points compared to other YouTube videos of similar length.
- API available to export/automate tracking.

### Tool Cons

- Top video data within the main reporting console is limited to showing the percent of total views by video; viewing engagement levels by video involves looking into each video one by one
- Excel exports available only for summarizing total activity and for providing information on total number of videos; cannot get video specific data.

### Pricing

Free.

### Recommendations

**Type of Site:** YouTube Insight is extremely easy to implement, free, and requires very little maintenance. Therefore, it is recommended for that any organization using YouTube enable YouTube Insights.

**Organizational Support:** Very little organizational support is required to use YouTube Insights. The data available through this platform are most effective when used by managers to understand who views their YouTube campaigns to influence future video marketing endeavors.

## *HootSuite*

### Social media communications dashboard

http://www.hootsuite.com

### Tool-Use Case

- Facilitates online brand management, allowing users to track mentions and assess online presence with social analytics reports.
- Users can communicate with audiences across multiple social networks (e.g., Facebook, Twitter, WordPress, LinkedIn, etc.) from the dashboard in a single step.

### Tool Pros

- Streamlines online brand management by housing data for multiple channels on a single platform.
- Users can pre-schedule updates, automating communication on social networks.
- Extensive social analytics allow users to identify trends over time in their online presence and measure the effectiveness of specific marketing campaigns.

### Tool Cons

- Historically, users have encountered problems using HootSuite's URL shortener, ow.ly

### Pricing

A free account provides social analytics and can connect to five social profiles. For $5.99 per month, the Pro  greater social analytic capabilities, connection to unlimited social profiles, one free custom report, integration with Google Insights and Facebook Insights, RSS feeds to automatically update social networks with blog posts. Additional team members, enhanced support, and vanity shortened URLs are available for additional fees. The Enterprise plan, priced at $14.99/month, includes 10 free reports, 30 team members, prioritized support, personalized setup assistance, and enrollment for 10 users in an ongoing training program.

### Recommendations

**Type of Site:** Given the wide range of plans available, HootSuite is appropriate for personal use, small businesses, and enterprise-level organizations.

**Organizational Support:** HootSuite allows users to manage multiple channels from a single platform, thus decreasing the level of support typically necessary to manage an organization's online branding and marketing. Implementation and dashboard maintenance are relatively straightforward, so organizational resources will only be required for data analysis and reporting.

## Voice of Customer Surveys

### *Introduction*

Voice of Customer surveys are intended to help organizations understand the perspectives of their customers in order to design and implement better customer experiences. The tools discussed in this section facilitate all phases of the survey cycle: design, distribution, data collection, analysis and reporting.

#### Tools Covered in This Section

- Foresee Results  American Customer Satisfaction Index (ACSI) survey
- iPerceptions
- QuestionPro
- Vovici

### *Foresee American Customer Satisfaction Index (ACSI) survey*

**Full Service Online Survey Provider**

www.foresee.com

#### Tool-Use Case

- Create a full site-wide online survey instrument for research purposes.
- Provide ongoing site survey analytics from professional services.
- Create a site comparison benchmark of effectiveness based on the American Customer Satisfaction Index (ACSI) survey data.
- Deploy additional surveys optionally and at extra cost.

#### Tool Pros

- **Robust research methodology that provides detailed benchmarking and strong guidance on problem and optimization opportunities**
- Support around the configuration, design, targeting, and analysis of an online survey
- Robust survey tool that is very familiar to online users
- Mobile survey measurement tool available

#### Tool Cons

- **The methodology is restrictive,  forcing a long survey and limiting the amount of custom research that can be conducted.**

- The length of survey will drive up abandonment rates, a potential issue when integrating with behavioral data.
- Analytics are heavily dependent on vendor professional services.

### Pricing

Foresee Results ACSI survey  is a fairly expensive solution. It is typically purchased on a subscription basis with monthly analysis baked into the pricing.

### Recommendations

**Type of Site:** Foresee is appropriate to a very wide range of sites. Its rigorous methodology makes it an ideal choice for when comparison to industry or competitive set is a priority. It is also a solution that is geared toward a high-level of ongoing support - making it a good choice for cases where internal analysis resources are limited. It is not appropriate when the goal of a research program is to provide small, flexible survey instruments that are integrated into internal analysis programs.

**Organizational Support:** Foresee requires some internal support from the development phase and during the implementation process. The vendor will typically work closely with your technical team. Little ongoing technical support is required.

## *iPerceptions*

**Full Service Online Survey Provider**
www.iperceptions.com

### Tool-Use Case

- Create a full site-wide online survey instrument for research purposes.
- Provide ongoing site survey analytics from professional services.
- Deploy additional surveys optionally and at extra cost.
- Deploy short, limited surveys for self-analysis at no cost .

### Tool Pros

- **Flexible, attractive survey creation system with robust customization tools.**
- Models for both self-service and analyst supported.
- Excellent methodology for question progression and survey qualification.
- Attractive free offering for short, non-integrated surveys.
- Integrated customer satisfaction benchmarking (iPSI).

**Tool Cons**

- **There is a large product gap between free and the full survey solution.**
- **The customer satisfaction benchmarking is less accepted and more proprietary than the ForeSee Results' ACSI survey benchmarks.**

**Pricing**

iPerceptions is a fairly expensive solution. It is typically purchased on a subscription basis with monthly analysis baked into the pricing but it can be purchased as a pure technology offering with all analysis conducted by the client.

**Recommendations**

**Type of Site:** iPerceptions is appropriate to a very wide range of sites. Its performance and look make it an ideal choice for cases where a company is looking to deploy a sophisticated ongoing and customized site wide survey. For companies looking for a very easy low-end solution, 4Q is also appropriate though very limited.

**Organizational Support:** iPerceptions requires some internal support from development during implementation. The vendor will typically work closely with your technical team. Depending on your contract, you may need significant research support internally or you can choose to rely heavily on iPerceptions for analysis.

### *QuestionPro*

**Online Survey Provider**

www.questionpro.com

**Tool-Use Case**

- Create a full site-wide online survey instrument for research purposes.
- Create small, "point" surveys for specific research problems.

**Tool Pros**

- **Flexible, fairly robust survey creation tool**
- Solid online survey analytics reporting.
- Easy deployment model suited to a large range of different survey types.

**Tool Cons**

- **Not a turnkey solution with fully-baked in professional services**
- Less robust in basic implementation model - takes more work and internal knowledge to create and deploy a survey.
- No site comparison or benchmarking data.

### Pricing

QuestionPro is a relatively inexpensive solution. There are several different levels ranging from a free basic survey to a $100 a month subscription version.

### Recommendations

**Type of Site:** QuestionPro can work on a wide variety of sites and research programs.

**Organizational Support:** QuestionPro requires your research and technology teams to be self-supporting. Your teams will be responsible for the design, deployment and analysis of the surveys. It is a fairly easy technology solution to deploy and the construction and analysis of surveys is straightforward for anyone who is a professional in the discipline.

## *Vovici*

**Full Service Online Survey Provider**
www.vovici.com

### Tool-Use Case

- Create a full site-wide online survey instrument for research purposes.
- Provide ongoing site survey analytics from professional services.
- Manage customer Feedback Collection.

### Tool Pros

- **Flexible, attractive survey creation system with robust customization tools.**
- Models for both self-service and analyst supported with self-service predominating.
- Robust online reporting and analysis of survey results.
- Mobile survey measurement tool available.

### Tool Cons

- **There is no equivalent to ACSI benchmarking data, so comparisons with competitors or other Web sites aren't practical.**

### Pricing

Plans start at $1800 per year and can increase dramatically, into the six-figures. Contact Vovici directly for pricing information.

### Recommendations

**Type of Site:** Given the flexibility of their survey offerings, Vovici's platform is appropriate at an array of organizations with a wide range of use-cases. However, given the relatively high cost and resource demand, this application is best used for large surveying campaigns.

**Organizational Support:** Vovici requires some internal support from development during implementation. In most cases, you will need to support Vovici with your internal research team. It is a fairly easy technology solution to deploy and the construction and analysis of surveys is straightforward for anyone who is a professional in the discipline.

## User Experience

### *Introduction*

User Experience tools allows organizations to understand the entire visitor experience on their Web sites, from the user's perspective. By leveraging a user-experience tool, organizations can under what obstacles user encounters, what is causing abandonment, and identify areas where the site failed to do what the user needed. These insights help improve site usability, increase conversion rates, and mitigate customer dissatisfaction.

**Tools Covered in This Section**

- Tealeaf
- Clicktale

### *Tealeaf*

**Packet Sniffer Data Capture –** Captures all data flowing between a Web server and visitors. www.tealeaf.com

### *Tool-Use Case*

- Provides full text search of the underlying data - allowing an analyst or customer service rep to "replay" an actual visitor experience on the Web site.
- Provides rich reporting on the specific occurrences of events on the Website at a very detailed level.
- Can be used for both customer support and analysis/reporting.

### *Tool Pros*

- **Provides complete data capture, so every visit is available for analysis and replay. This makes it particularly useful for customer support**
- Wireline capture eliminates any need for tagging and any operational impact on the Website
- Rich text-based retrieval of sessions allows for single sessions to be easily identified and captured.

### *Tool Cons*

- **Expensive. Full data capture makes this a very high-end solution**
- Reporting is better suited to full-text retrieval than structured analysis. Finding behavioral trends or supporting broader reporting is possible but not ideal

### *Pricing*

Tealeaf is a fairly expensive solution. Costs are based on total traffic (and length of storage) but the system requires dedicated hardware installed on-premise.

### *Recommendations*

**Type of Site:** Tealeaf's capabilities are particularly well-suited to complex operational sites. Flat HTML sites will benefit little from Tealeaf's advanced replay capabilities. However, in situations where customers are navigating a range of options in an application-like environment, Tealeaf excels. The ideal profile is a fairly high-volume, transaction oriented Web site with a complex front-end.

**Organizational Support:** For customer support applications, Tealeaf needs to be integrated into your support operations. This will take significant training.

Administration and setup of the system is a task for a dedicated or part-time professional. Most organizations rely on Tealeaf during the implementation phase.

Analysis of Tealeaf data is best done in conjunction with a Web analytics system by dedicated analysts. Some basic training in the system is definitely advisable.

### *Clicktale*

**SaaS tagging solution**
www.clicktale.com

#### Tool-Use Case

- Using a tag, Clicktale captures a rich, client-side event-stream from a Web experience
- Provides structured search of the underlying data, allowing an analyst or customer service rep to "replay" an actual visitor experience on the Web site
- Provides rich reporting on the usage of pages including scrolling and mouse movements
- Primarily used for user interface (UI) analysis

#### Tool Pros

- **Lightweight and easy to implement**
- Client-side tracking provides detailed mouse-movement data that are not available in any other system
- Can be implemented on specific pages and can be sampled
- Provides very nice reporting of page usage including Mouse-Movement heat maps.

#### Tool Cons

- **Tag and software approach is often too heavy to capture all sessions, making it inappropriate for many customer support applications**
- Not all data are available for search
- Doesn't provide direct access to the data store

- Tagging methodology will only collect data after a full-page load - losing information that will be captured via a Wireline method.

**Pricing**

Clicktale is generally inexpensive to implement. It is priced on a per server call level. Since the pages and sample rates can be controlled, most organizations can set the level of investment they wish to make.

**Recommendations**

**Type of Site:** Clicktale's capabilities are appropriate for most sites. The tool can be fruitfully deployed on Home Pages, Landing Pages, Form Pages, and on entire sites. Most heavily trafficked sites will deploy the tool on either limited pages or sampled sessions both to control costs and limit performance impact (on both client-side and in the tool).

**Organizational Support:** Clicktale is a appropriate for analysts. It is relatively easy to learn but will probably require some basic introductory training.

## Presentation (like Tableau, Crystal Reports)

### *Introduction*

The positive effect of consumer behavior insights on brand awareness and the effectiveness of outreach and marketing efforts are greatly dependent upon how this information is communicated to relevant stakeholders and decision-makers. Even the most valuable insights are worthless if they are cannot be understood and transformed into action by the right parties. This, in turn, has caused organizations to focus not just on “what” is being presented, but “how” it is presented. The presentation tools covered in this section facilitate data analysis and reporting, specifically offering highly advanced data visualization capabilities.

**Tools Covered in This Section**

- Tableau
- Spotfire
- Crystal Reports

### *Tableau*

**Web and Distributed Reporting and Analysis System**

www.tableausoftware.com

**Tool-Use Case**

- Provide a rich, multi-dimensional customized interface to data
- Allow for desktop, distributed or Web access to corporate reports
- Combine multiple data sources and integrate into interactive dashboards

**Tool Pros**

- A powerful, visually attractive, and highly customized interface.

- A variety of delivery mechanisms that support most common use-cases
- Easy integration with most types of data
- Excellent performance against compact and mid-range data sources

### Tool Cons

- Limited scripting and advance analytic capabilities
- Full customization requires an experienced user and is not necessarily straightforward
- The interface can be confusing for low-experience users

### Pricing

There are different types of Tableau each with each pricing model. In general, this is a mid-range solution with pricing ranging from the low-teens to just under six figures depending on configuration and licensing model.

### Recommendations

**Type of Site:** Tableau is an advanced data analysis and exploration system. It is not for novices or for those using simple score cards or reports. It can support most advanced database systems and while it is not appropriate for big-data analysis, it can be used in any situation where sophisticated access to aggregated data is required.

**Organizational Support:** Customization of Tableau interfaces is a job best level to serious professional users. Tableau data processes and aggregations can be managed by power-users, but are often the domain of IT professionals. Users of Tableau range from serious full-time analysts to power-users to moderately sophisticated self-service consumers.

## *Spotfire*

**Web Reporting and Analysis System**

spotfire.tibco.com

### Tool-Use Case

- Provide a rich, multi-dimensional customized interface to data
- Allow for Web access to advanced dashboards and reports
- Combine multiple data sources and integrate into interactive dashboards

### Tool Pros

- **A powerful, visually attractive and highly customized interface**
- Ability to deliver interactive dashboards over the Web
- Easy integration with most types of data
- Excellent performance compared to t compact and mid-range data sources
- Ability to call and use more advanced statistical analysis capabilities inline

### Tool Cons

- Full customization requires an experienced user and is not necessarily straightforward
- The interface can be confusing for low-experience users

### Pricing

There is an individual version of Spotfire at $80/month. For common corporate use, Spotfire is a mid-range solution with pricing ranging from the low-teens to just under six figures depending on configuration and licensing model.

### Recommendations

**Type of Site:** Like Tableau, Spotfire is an advanced analysis and reporting system that can be custom-fitted on top of most types of data. The interface is not as polished as Tableau but the system provides readier access to powerful external and programmatic statistical analysis systems.

**Organizational Support:** Customization of Spotfire interfaces is a job best left l to serious professional users. Spotfire data processes and aggregations can be managed by power-users, but are often the domain of IT professionals. Users of Spotfire range from serious full-time analysts to power-users to moderately sophisticated self-service consumers.

## *Crystal Reports*

### Web Reporting and Analysis System

http://www.sap.com/solutions/sap-crystal-solutions/index.epx

### Tool-Use Case

- Provide a rich, multi-dimensional customized interface to data
- Allow for Web access to advanced dashboards and reports
- Combine multiple data sources and integrate into interactive dashboards

### Tool Pros

- A powerful report-oriented interface with excellent business charting and graphics
- Ability to deliver across a variety of mediums including interactive Web
- Easy integration with most types of data
- Excellent performance compared to  compact and mid-range data sources

### Tool Cons

- Stronger as a reporting engine than as an interactive data exploration tool

### Pricing

For common corporate use, Crystal Reports is a low to mid-range solution with pricing ranging from a few thousand dollars to just under six figures.

#### Recommendations

**Type of Site:** Like Tableau and Spotfire, Crystal is an advanced analysis and reporting system that can be custom-fitted on top of most types of data. The interface is very mature and the technology has been evolving for many years. Its strength is more as a customized reporting engine (across Web or automated delivery) than as a data-exploration platform. It provides a less natural interface to slice-and-dice data visually than either Tableau or Spotfire and less statistical capability. However, it provides more report-layout customization than competing solutions.

**Organizational Support:** Customization of Crystal Reports interfaces is a job best level to serious professional users. Crystal data processes and aggregations can be managed by power-users but are often the domain of IT professionals. Users of Crystal encompass the full spectrum of data consumers from very sophisticated to novice.

## Direct Surveys

### *Introduction*

Direct surveying is a widely accepted method for acquiring knowledge about a given population's sentiments, perspectives, and experiences. These types of insights are particularly useful for directing outreach and marketing efforts and other business decisions. The tools discussed in this section support all phases of the survey cycle. Specifically, these platforms support survey development, distribution, data collection, analysis, and reporting.

#### Tools Covered in This Section

- Qualtrics Survey Suite
- SurveyGizmo
- SurveyMonkey
- WorldApp Key Survey

### *Qualtrics Survey Suite*

#### SaaS Solution

www.qualtrics.com

#### Tool-Use Case

- Self-service, hosted enterprise survey solution intended for large-volume surveying such as market research, employee feedback, and academic research.
- Facilitates and supports the entire survey cycle with features for creating and distributing surveys, as well as analyzing and reporting results.
- Uniquely offers several multimedia, interactive question-types intended to engage survey-takers, making it especially appropriate for surveying campaigns with direct marketing or brand-awareness goals.

### Tool Pros

- Survey appearance highly customizable – Qualtrics features multiple-media question types (e.g., sliding scales, heat maps, image ranking, grading), customizable survey skins, and complete control over survey branding.
- Straightforward, "point and click" interface – even novice users can easily create and customize questions without knowledge of HTML, JavaScript, or CSS.
- Platform includes various features to maximize efficiency of data analysis and reporting. Raw data can be coded, scored, and compared by demographics. Flexible reporting tools allow for direct exporting of raw data into Excel or SPSS, and downloading reporting into MS PowerPoint, Word, Excel, or Adobe PDF.

### Tool Cons

- Expensive.
- Feature-intensive interface can potentially overwhelm users.
- Platform greatly prioritizes the "look and feel" of surveys, which may distract from quality of content and data accuracy. This may be especially detrimental to users without knowledge in surveying best practices.

### Pricing

Qualtrics is sold in annual license contracts at three levels, each allowing for increasing numbers of completed surveys. Every license includes an unlimited number of surveys, training, and support. The Branded option allows for up to 10,000 completed surveys and costs $10,000 per year. The Enterprise version is priced at $25,000 per year, allowing up to 40,000 completed responses and API database automation. Global Enterprise is recommended for those looking to standardize a single surveying platform across an entire organization. At this level, the solution will cost $55,000 for an annual license, with bandwidth for up to 250,000 completed responses. Qualtrics also offers discounted rates and more flexible parameters (e.g., more completed surveys at each level) for not-for-profit organizations.

### Recommendation

**Type of Site:** The Qualtrics Survey Suite is intended for enterprise-level clients conducting high-volume or multi-iterative surveying.  It is most cost-effective and priced accordingly, for large-scale surveying rather than a single campaign or project. Given its extensive question types, secure distribution options and branding control, Qualtrics is extremely powerful and flexible. There is a strong emphasis on survey appearance, with customizable survey "skins" and interactive questions. This makes Qualtrics a good option for organizations looking to engage survey-takers and reinforce their brand via their surveying endeavors.

**Organizational Support:** Survey creation, distribution and reporting with the Qualtrics' platform is extremely intuitive, and therefore surveys could be implemented by anyone in the organization, regardless of experience. The available features and tools even allow for considerable automation of the surveying process. However, creating survey content, interpreting data, and leveraging results effectively should be handled by those technically trained and experienced in survey design and quantitative analysis.

### *SurveyGizmo*

**SaaS Solution**

www.surveygizmo.com

#### Tool-Use Case

- Enterprise survey software that allows users to build, customize, distribute and analyze surveys from a hosted platform.
- Extensive feature set allows for customization of structure at both the question and survey level.
- Full-service consulting and technical support available for custom survey building, training, and custom reporting. These services are billed at an hourly rate.

#### Tool Pros

- Solution includes features specifically designed to support the needs of enterprise-level clients – integration with various other enterprise solutions (e.g., Salesforce, ExactTarget, iModerate, Cint), open access API, multiple team-based logins.
- High level of customer support, training, and consulting available. Tutorials and webinars are available free of cost for any user. Prioritized email and phone support is also available for the higher grade paid accounts.
- Advanced logic, branching, and piping capabilities – supports automatic restructuring of survey path based on various criterions. This allows surveys to be highly dynamic.

#### Tool Cons

- Complicated feature set – platform supports total survey customization and brand control, but leveraging these functionalities requires extensive knowledge of the interface.
- Steeper learning curve than other solutions for first-time users.

#### Pricing

SurveyGizmo is sold as a month-to-month subscription, with no contract required for the Pro and Enterprise plans. Users may upgrade or downgrade their account at any time. A free version is available, with limited functionality and features. The Personal plan, at $19 per month, allows unlimited surveys and responses, and up to 10,000 email invitations. The Professional package costs $49 per month, and supports up to 50,000 email invitations and advanced logics. For $159 per month, the Enterprise plan allows 500,000 email invitations, advanced reporting options, complete brand control, salesforce.com integration, and phone-based customer support. A Dedicated survey plan  is available for $7,888 per year, supports 40 unique users and 10 teams, 1,000,000 email invitations, a dedicated account representative, dedicated database, 10 hours of  initial support, and 2 hours a month of professional services.

### Recommendations

**Type of Site:** SurveyGizmo is most appropriate enterprise-level clients with high-volume campaigns that require complex structuring, branding control, and/or highly customized surveys. Additionally, SurveyGizmo is a good option for organizations willing to spend more for professional support.

**Organizational Support:** The level of internal organizational support necessary when using SurveyGizmo's platform depends on both the use-case and selected survey package. While the interface is relatively complicated, especially if using the complex structure options (e.g., advanced question logic, branching, etc.), users with the more expensive survey plans can leverage the high-quality customer support available.

## *SurveyMonkey*

### SaaS Solution

www.surveymonkey.com

### Tool-Use Case

- Do-it-yourself survey platform intended for small-scale, simple surveys. Surveys and response data are hosted on SurveyMonkey's site, and can be accessed online from any location.
- Users create original questions, selecting from available 15 question types, or use pre-populated survey templates.
- Surveys are assigned unique URLs that can be distributed on Web sites, social networks, and/or via email. With paid packages, users may customize the survey URL. Platform also includes tools for inviting users via email, Facebook, or timed pop-ups on a Web site.

### Tool Pros

- Low-cost--a free version is available, with limited features. Even with the most expensive plan, SurveyMonkey is cheaper than most other survey platforms.
- Simple user-interface and limited customization options make the user orientation process quick. First time users can easily implement a survey within a couple of hours.

### Tool Cons

- Limited question types and survey customization options.
- Brand control is limited. White-label surveys (without SurveyMonkey branding) are only available in the most expensive package.
- Reporting options are limited. Data cannot be compared by demographics, scored or recoded. Most data analysis must be conducted using external tools

### Pricing

The SurveyMonkey solution is available in four plans, each with ascending degrees of available features and pricing. The Basic plan is available at no cost, allowing for up to 10 questions and 100 responses per survey. Most features in this plan are unavailable, or reduced. The Select option costs $16.99 per month and allows unlimited questions and responses per survey. At this level, users may customize survey URL, increase security, export to Excel and PDF, and add skip logic to questions. The Gold package, priced at $24.99 per month, includes more advanced logic; customize end-of-survey redirects, free response text analysis, and SPSS integration. Lastly, the Platinum package, at $69.99 per month, allows complete brand control and phone support.

### Recommendations

**Type of Site:** SurveyMonkey's simple platform is most appropriate for small-scale, informal surveys. Small to mid-sized companies would benefit from using Survey Monkey for customer surveys, employee feedback, academic research, or course evaluations. Given the constraints on customizing survey appearance, Survey Monkey is not recommended for organizations with strict branding requirements.

**Organizational Support: Once survey content is developed, using SurveyMonkey's solution is simple and efficient, so anyone in the organization could oversee survey execution. It is important to note, however, that compared to other solutions, customer support and training is relatively limited.**

## *WorldApp Key Survey*

**SaaS Solution (also available On-Premise)**

www.keysurvey.com

### Tool-Use Case

- Fully hosted surveying solution, available with a basic set of features out-of-the-box and full customization via their professional services offerings.
- Supports the entire surveying cycle, from creation to reporting and analysis.
- Professional services, including: custom dashboards on the user interface, custom survey design, survey scanning, distribution, and training.

### Tool Pros

- High-touch customer support and professional services. Surveys are fully customizable, as WorldApp will develop unique survey designs for individual clients.
- Capable of integrating with various other enterprise solutions, deployment options with varying degrees of data security (e.g., on-premise, private label, or custom-configured servers) make Key Survey a good option for enterprise-level clients.
- Extensive features options allow for total flexibility and brand control. Users are not confined to template-based surveys structures.

**Tool Cons**

- Complicated Web site and extensive package options could cause initial implementation to be time-consuming and requiring input from many stakeholders.
- Most customizations require intervention from the professional services team, which drives cost significantly upwards.

**Pricing**

Key Survey is sold in annual licenses, in an ascending pricing structure, allowing for increasing numbers of administrative users. Because the solution emphasizes, and often requires, professional services pricing varies greatly.

Contact WorldApp directly for a price quote.

**Recommendation**

**Type of Site:** Key Survey is most appropriate for enterprise-level clients with high-volume, complicated surveying needs. The platform would be especially useful for organizations with complicated survey structuring, security and data management concerns, and/or looking to outsource some of the surveying process. Though it comes with an increased price tag, Key Survey offers a uniquely high level of customer support and consulting.

**Organizational Support:** Using this application will require a high-level of organizational support, as compared to other survey tools. This is especially true during the initial survey design phase. Given its exceptionally flexible capabilities, this platform requires a high degree of decision-making. Therefore, survey campaigns using Key Survey should be driven by employees with access to key decision-makers and at least moderate experience with survey design.

## Call Center Analytics

### *Introduction*

Call Center Analytic tools complement traditional Web analytics reports on of customer conversion by capturing offline conversations and other desired actions. . Using the various solutions discussed in this section, organizations can trace offline conversions back to the online referring sources. This allows organizations to understand the relative effectiveness of their various marketing campaigns, pay-per-click ads, and SEO efforts even when the actual sale occurs offline.

**Tools Covered in This Section**

- Mongoose Metrics
- Marchex
- Ifbyphone

### *Mongoose Metrics AccuTrack Session*

**SaaS Solution**

www.mongoosemetrics.com

### *Tool-Use Case*

- Fully hosted, enterprise level call tracking solution that allows users to track and analyze source and visitor data for all calls placed to a call center.
- A JavaScript snippet added to each Web page dynamically assigns a unique phone number to each new site visitor. This allows lead details, such as last URL visited, keyword searched, and Caller ID, to be recorded when the visitor places a call.
- Data can be accessed online within Mongoose Metrics' platform, downloaded to a CSV file or, using an API, integrated with a wide-variety of Web analytics, marketing automation, and customer relationship management solutions.

### *Tool Pros*

- Extremely compatible with various Web analytics platforms, such as Google Analytics, Omniture SiteCatalyst, Yahoo! Web Analytics, Lyris ClickTracks, Coremetrics, and Webtrends. Users can easily integrate data from their Mongoose Metrics campaigns into any of these solutions thereby making information more accessible and familiar to a wider audience within the organization.
- Reporting dashboard user-friendly and easy to decipher. Reports succinctly summarize data at the individual call-level. These data can be easily aggregated to develop a comprehensive view of which marketing campaigns, key words, or site pages are driving calls.
- Solution captures data for offline conversion events, which are typically missed by traditional Web analytics tools.

### *Tool Cons*

- Tool is most effective as a supplement to traditional Web analytics, and therefore less powerful as an isolated tool. This can be problematic for organizations without a strong Web analytics program already in place.
- While reporting dashboards offer rich, detailed views of individual calls and keywords, most analysis and data visualization (e.g., creating charts or graphs) must be done manually and outside of the platform.
- Implementation is relatively complicated. Since tool operates with JavaScript, developers will likely need to be involved in the initial implementation phase.

### *Pricing*

AccuTrack Session is sold in licenses based on call volume and site traffic. Contact Mongoose Metrics directly for a specific price quote.

### *Recommendation*

**Type of Site:** Mongoose Metrics' AccuTrack Session is most appropriate for enterprise level organizations with well-supported, established Web analytics programs. The tool is meant to supplement measurement of Web behavior to create a holistic view of how visitors convert to customers.

**Organizational Support:** Much like most Web analytics solutions, AccuTrack Session requires significant time and resources during the planning and initial implementation. Because data derived from this tool are easily integrated with Web analytics systems, however, reporting and analysis would not result in a significantly increased workload for the current Web analyst.

### *Marchex Call Analytics*

**SaaS Solution**

www.marchex.com

**Tool-Use Case**

- Fully hosted, enterprise level call tracking solution that allows organizations to understand which marketing campaigns, site pages, and keywords are driving calls to their call centers, and ultimately conversions.
- Organizations can track calls at the campaign-level by assigning a unique number to each marketing method (e.g., print ads, Web banners, emails, etc.).
- To understand which pay-per-click and organic search terms are driving calls, JavaScript placed on Web site landing pages dynamically assigns a phone number to each unique term.
- Tracked data are available in 50 different reports on Marchex's online platform. Users can also download data to Excel, send via email, or integrate with other solutions (e.g., Google Analytics or Omniture SiteCatalyst) via an API.

**Tool Pros**

- After implementation, tracking is automated and requires little maintenance.
- Largest inventory of tracking phone numbers (both toll-free and local numbers are available).
- Private labeling available, making this a good option for agencies looking to enhance the value of their marketing campaigns.

**Tool Cons**

- Reporting is static, so any data manipulation (e.g., visualization or analysis) must be done outside the solution.
- While Marchex provides data on a wide variety of metrics, these metrics cannot be readily combined to develop a comprehensive evaluation of each campaign's effectiveness.

**Pricing**

Marchex call analytics solution is priced based on the number of phone numbers and minutes used per month. Rates start at $9 per phone number, per month, and economize with the scale of use. Because these rates vary greatly, a Marchex sales representative must be contacted for a specific price quote.

**Recommendations**

**Type of Site:** The Marchex call analytics platform is appropriate for a wide-variety of organization types. Direct marketers in enterprise-level organizations and external agencies can use this platform to optimize pay-per-click search campaigns. Smaller organizations can benefit from tracking call sources at the higher marketing campaign-level.

**Organizational Support:** If tracking calls at the keyword-level, the implementation phase will require input from developers, as the dynamic tracking numbers operate using JavaScript code. After the initial setup, however, tracking call data becomes a predominately-automated process. Integrating Marchex into other web analytics platforms via an API further reduces the resource and time burden for users.

### *Ifbyphone*

**SaaS Solution**

www.ifbyphone.com

**Tool-Use Case**

- A package of voice applications for call tracking, call notifications, call forwarding, and call automation hosted on Ifbyphone's cloud-based platform.
- The SourceTrak suite of products track call data with varying degrees of detail. SourceTrack Basic assigns a unique phone number on any designated marketing venue (e.g., billboards, print ads, direct mailings) and tracks the number of calls originating from each number. SourceTrak Dynamic assigns unique phone numbers to specified online traffic sources (e.g., PPC search terms, organic search terms, domains, or ads), one of which is dynamically displayed on the Web site for each visitor based on the referring source. SourceTrak Dynamic Groups works like SourceTrak Dynamic, but for larger marketing campaigns, such as keywords within a PPC Ad group. SourceTrak Session tracks the most granular, keyword-level detail for an unlimited number of traffic sources.
- For dynamically changing phone numbers, a snippet of JavaScript code is placed on the organization's Web site and calls back to Ifbyphone's servers to determine the appropriate phone number to display.

**Tool Pros**

- Available as a Google Analytics application, allowing for a seamless integration into established Google Analytics. This decreases the additional workload for users
- Strong customer support.

- Relatively inexpensive compared to other call-tracking tools. Additionally, Ifbyphone's services are billed monthly on a per-minute basis, but organizations need not predetermine the amount of minutes they will use. This ensures that organizations only pay for minutes used.

**Tool Cons**

- The data analysis required to surface value from this solution must be conducted manually and externally from the solution. This issue, however, is easily remedied by integrating data into Google Analytics or other platforms via an API.
- Only available as a bundle with other voice marketing applications (e.g., call routing, call broadcasting, SMS services, call distribution, etc), so organizations may end up paying for services they do not need or use.
- Reporting interface is relatively unintuitive.

**Pricing**

Ifbyphone's call tracking services are sold in three levels of monthly subscriptions, bundled with other voice marketing applications. The Basic package, sold at $49.95 per month, includes call tracking and eight other services. The Advanced plan, at $59.95 per month, also includes Google Analytics integration, dynamic phone numbers, IVR, and voice broadcasting. At the highest level, the Complete plan is priced at $74.95 per month. This is the only package that includes a developer API toolkit. Every subscription includes 400 minutes of call time. Additional minutes used are billed at an economizing rate (i.e., the more minutes used, the cheaper each minute costs).

**Recommendations**

**Type of Site:** The Ifbyphone suite of products is most appropriate for small to mid-sized organization with basic call analytics needs. Additionally, as Ifbyphone's call tracking is bundled with other voice marketing services, their solution is most cost-effective for organizations interested in leveraging call center services.

**Organizational Support:** Because Ifbyphone's call tracking operating using JavaScript, a developer will likely be needed to assist the initial implementation. Additionally, identifying what should be tracked (e.g., keywords, PPC search terms, site pages, etc.) requires input from multiple decision-makers. If a Web analytics program is well-established within the organization, particularly using Google Analytics, leveraging Ifbyphone's call tracking services should not require a significantly increased resource investment. However, when used on its own, data must be extracted, analyzed, visualized, and reported manually.

## A/B and Multivariate Testing

### *Introduction*

A/B testing is the method used for comparing the effectiveness of content and content placement on two different pages so you can see what may be best in attracting and keeping site visitors. Multivariate testing takes this a step further in that you test multiple elements on one page, such as text, image type, placement and so forth. These tools provide you with the means to conduct the tests with your pages and measure the results.

**Tools Covered in This Section**

- Adobe Test and Target
- Sitespect
- Maxymiser
- Google Website Optimizer

### *Adobe Test and Target*

#### A/B and Multivariate Testing Solution

http://www.omniture.com/en/products/conversion/testandtarget

**Tool-Use Case**

- Create and test the effectiveness of alternative site (and email) graphics and/or content
- Determine best option for modifying graphic or content approaches for specific audiences or segments on your Web site
- Test multiple creative units simultaneously on a page

**Tool Pros**

- **Integration with Omniture Web analytics solutions for segmentation and variables**
- **Integration with Adobe Content Creation Systems**
- Full-service professional services consulting to assist in implementation and testing
- Rich targeting capabilities

**Tool Cons**

- Less sophisticated multivariate testing algorithms

**Pricing**

Test and Target is typically priced on a Cost per Million or CPMM basis - so cost is dependent on the volume of tests served. This can be problematic for high-traffic Websites. In general, it tends to be more expensive than its competitors' options per test-served. It often comes with a significant professional services component as well.

**Recommendations**

**Type of Site:** Test & Target is most attractive for sites that are running Omniture as their Web analytics solution or are looking for tight integration with Adobe creative tools. It is also appropriate in cases where the focus is on segmentation within smaller, high-value tests.

**Organizational Support:** Test and Target requires moderate expertise for the setup. However, like all testing systems, it requires significant expertise in testing methodologies and test design and interpretation to be run well.

### *Sitespect*

**A/B & Multivariate Testing Solution**

www. sitespect.com

**Tool-Use Case**

- Create and test the effectiveness of alternative site content and/or graphic. Determine the best options for modifying content or graphic approaches for specific audiences or segments on your Web site
- Test multiple creative units simultaneously on a page
- Supports both on-premise and hosted versions

**Tool Pros**

- Rich multivariate test design capabilities
- Turnkey integration of testing results into multiple Web analytics solutions
- Powerful ad hoc segmentation capabilities
- Intuitive interface for generating campaign creative options
- Minimally invasive approach to site-tagging

**Tool Cons**

- Integration from Web analytics to the testing tool (as opposed to the reverse)
- More complex page placements can involve more and more difficult work than tagging requires

**Pricing**

SiteSpect is typically priced on a CPMM basis - so cost is dependent on the volume of tests served. This can be problematic for high-traffic Websites. In general, it tends to be less expensive than its competitors are per test-served. It comes with a range of professional service offerings of which the basic Jumpstart package is the most common.

**Recommendations**

**Type of Site:** Sitespect is attractive for a wide range of sites seeking a straightforward means of implementing A/B and multivariate tests with a minimum of IT impact.

**Organizational Support:** Sitespect is relatively easy to deploy. However, like all testing systems, it requires significant expertise in testing methodologies and test design and interpretation to be run well.

### *MaxyMiser*

**A/B & Multivariate Testing Solution**

www. maxymiser.com

**Tool-Use Case**

- Create and test the effectiveness of alternative site content and/or graphic. Determine the best options for modifying content or graphic approaches for specific audiences or segments on your Web site
- Test multiple creative units simultaneously on a page
- Predictive modeling and behavioral targeting for personalization of pages for individual visitors' characteristics and preferences
- Automated process for creating and classifying segments for testing

**Tool Pros**

- **Tag Management style approach to site integration**
- **Separate module dedicated to automated Segmentation**
- Behavioral Targeting capabilities for both content and recommendations
- Intuitive interface for generating campaign creative's
- Minimally invasive approach to site-tagging

**Tool Cons**

- Integration with Web analytics solutions

**Pricing**

Maxymiser is typically priced on a CPMM basis - so cost is dependent on the volume of tests served. This can be problematic for high-traffic Websites. In general, it tends to be a mid-range solution. Professional Services are available to support both implementation and test management.

**Recommendations**

**Type of Site:** Maxymiser is attractive for a wide range of sites seeking a straightforward means of implementing A/B and multivariate tests. It provides a solution with a very robust segmentation approach - which is critical to more advanced testing.

**Organizational Support:** Sitespect is relatively easy to deploy. However, like all testing systems, it requires significant expertise in testing methodologies and test design and interpretation to be run well.

### *Google Website Optimizer*

**A/B & Multivariate Testing Solution**

***www.google.com/websiteoptimizer***

**Tool-Use Case**

- Create and test the effectiveness of alternative site content and/or graphic approaches
- Test multiple creative units simultaneously on a page

### Tool Pros

- **Simple, easy to use interface for the creation and execution of multivariate tests**
- **Tight integration with Google Analytics**

### Tool Cons

- Lack of Segmentation and Targeting Capabilities
- Limited sophistication in  multivariate testing methodology

### Pricing

Google Website Optimizer is free.

### Recommendation

**Type of Site:** Google's Website Optimizer is generally appropriate for smaller sites, sites running Google Analytics  as their primary Web analytics solution, and site's looking to "get their feet wet" in testing.

**Organizational Support:** Google Website Optimizer is quite easy to setup and run. However, like all testing systems, it requires significant expertise in testing methodologies and test design and interpretation to be run well. The ease of use of this system can make it deceptively easy and encourage sloppy or mistaken approaches to testing.

## Market or Internet Industry Research (Panel Research)

### *Introduction*

The platforms discussed in this section monitor and collect Web data and study online behavior, to provide organizations with third-party marketing data and insights. Organizations can leverage this data to develop a comprehensive view of their competitors, market share, and industry landscape. Moreover, these tools give organizations access to extremely robust data warehouses without the investment of their own resources.

### Tools Covered in This Section

- Compete
- Experian Hitwise
- Comscore

### *Compete*

**SaaS**

www.compete.com

### Tool-Use Case

- A hosted, online intelligence software system that provides site, search, and referral analytics.
- Site Profiles feature allows users to monitor and benchmark against another site on metrics that track visitor traffic, engagement, and demographics.

- Search Analytics analyzes search engine performance of a Web site or a specific keyword.
- Referral Analytics identify the channels that drive traffic to a given domain, as well as the downstream traffic from a given Web site.

**Tool Pros**

- Panel tracks the online behavior of over 2 million participants.
- Panel consists of a diverse, representative cross-section of United States Web users. This helps mitigate the impact of source bias on the analytics data.
- Results can be downloaded to a CSV for further manipulation.

**Tool Cons**

- Consumer panel is drawn exclusively from the United States, so Compete is not a good option for organizations with vested interest in other regions.
- Many of the key product features, such as referral analytics, are only available in the more expensive subscriptions.

**Pricing**

Compete's intelligence platform is sold in monthly subscriptions with varying levels of capabilities based on price. For $199/month, the Intro plan allows one user, site analytics and benchmarking, up to 50 search analytics reports per month, and data for the top 200 sites. The Standard plan includes licensing for one user, up to 100 search analytics reports, and data for the top 1,000 sites. For $499 per month, the advanced subscription includes one user, extended site analytics, advanced search analytics with up to 250 reports per month, referral analytics, and data for the top 15,000 sites. The Enterprise plan, for groups and companies, includes licensing for multiple users, dedicated phone support, expanded site analytics, unlimited search analytics reports, referral analytics, and data for the top 15,000 sites. Since the number of licenses at this level varies, customers should contact Compete directly for pricing.

**Recommendations**

**Type of Site:** Compete's competitive intelligence platform is specifically designed to make analytics data available to relatively smaller organizations. Using this solution, small to mid-sized organizations can leverage Web behavior information typically only accessible to large organizations with extensive budgets and resources. With the introduction of the Compete PRO Enterprise plan, this platform is also appropriate for enterprise-level clients with multiple users.

**Organizational Support:** Because the consumer panel is maintained by Compete, the implementation phase requires relatively little organizational support, as compared to many other Web analytics solutions. Generally, the tool's user interface is intuitive. However, effectively leveraging the data surfaced using this solution will require the involvement of a highly trained staff, experienced in data analysis.

### *Experian Hitwise*

**SaaS**

www.hitwise.com

**Tool-Use Case**

- A competitive intelligence system that tracks online behavior by collecting and aggregating data from ISP networks and opt-in panel partners.
- Can be used to analyze trends in online visitor behavior and measure Web site market share.
- Provides data on channels driving Web site traffic, downstream Web site traffic, site visitor demographics, search behavior, and the key players and rankings in over 160 industries.

**Tool Pros**

- Extensive filtering options on the various reports allow users to refine and customize data to reflect only relevant visitors, terms, or behaviors.
- Exceptionally large pool of data—Hitwise collects information from over 25 million people (10 million in the United States).
- Includes international user data, so this is a good option for organizations looking for data about users from other countries.

**Tool Cons**

- Extremely expensive.
- Given the money and time resources required, this solution is only accessible to large organizations.
- Extensive reporting options require a knowledgeable staff to be used appropriately and effectively.

**Pricing**

Waiting to hear back from sales representative, but quotes have ranged from $50,000—$60,000 per year.

**Recommendations**

**Type of Site:** Hitwise is an extremely powerful competitive intelligence platform. Using the available feature set, marketers can determine their relative position and performance within their industry identify the channels and search terms that drive users to their Web site, and understand the profile and behavior of visitors to their website. Moreover, similar information can be derived for their competitors. However, given the exceptionally high price tag, this solution is only appropriate for large, enterprise-level organizations with specific marketing and branding goals that can be achieved through analyzing the available data.

**Organizational Support:** In order to ensure a reasonable return on investment, any Hitwise implementation should be managed by a well-trained staff with experience in data analysis and a strong understanding of the organization's marketing goals. Given the robust, detailed information accessible, using this product will likely require a full-time analyst or marketing professional.

### *comScore*

**SaaS**

www.comscore.com

#### Tool-Use Case

- A wide range of marketing services based on market data collected by comScore via a robust, cross-sectional panel of Web users.
- Can be used to analyze trends in online visitor behavior and measure Web site market share.
- Collects Internet data by recruiting volunteers to install monitoring software on their computers, which tracks all their online behavior.

#### Tool Pros

- Widely considered the standard in the measurement of online behavior.
- Unified Digital Measurement system blends both panel and census-based measurement. Their proprietary method is advantageous because it is less affected by variables such as cookie deletion or blocking.
- Fully integrated suite of programs based on their collected data, so organizations can address all their digital measurement needs from a single platform.

#### Tool Cons

- Only appropriate for large organizations.
- Expensive.

#### Pricing

Given their highly diverse offerings, pricing can vary greatly. Contact comScore directly to discuss digital measurement needs.

#### Recommendations

**Type of Site:** comScore's digital measurement system is most appropriate for large organizations (i.e., Fortune 1000).

**Organizational Support:** comScore offers extremely powerful digital measurement tools, which can provide comprehensive insight and guidance for large-scale marketing campaigns. Therefore, a fulltime, highly-trained analyst is recommended to manage, analyze, report, and leverage the data.

## Search Engine Optimization (SEO)

### *Introduction*

Given that a great majority of visitors navigate the Web using search engines, the ranking and inclusion of an organization's Web content in organic (non-paid) search results is critical for driving site traffic, brand awareness, and ultimately customer acquisition. Search Engine Optimization (SEO) tools help organizations understand, manage, and improve their presence in search results.

#### Tools Covered in This Section

- Google Insights for Search
- Search Engine Reports
- SEOmoz

### *Google Insights for Search*

#### Service provided by Google (Beta)

http://www.google.com/insights/search

#### Tool-Use Case

- A Google service that calculates the number of searches that have been conducted on Google.com for a given search term, relative to all Google.com searches. Allows organizations to understand the impact of their marketing campaigns on search interest.
- User enters search term and sets date parameters. Application then provides a visual representation of search volume over time, regional interest, related search terms, and rising search terms
- Results are normalized to account for differences in total search volume.
- Data is reported on a scale from 0 to 100, where a score of 100 is the highest search volume achieved and all other scores are relative to that peak.

#### Tool Pros

- Free.
- No setup required. The service is available online for public consumption.
- Results can be downloaded to a CSV for further manipulation.

#### Tool Cons

- Limited analysis capabilities. Google does not provide raw search data, which limits users' ability to conduct most analyses.

#### Pricing

Free.

**Recommendations**

**Type of Site:** Google Insights for Search is a useful tool for any marketing campaign, regardless of size. Marketers can use the information extracted from this application to adjust pay-per-click and SEO campaigns. Additionally, Google Insights for Search is a helpful tool for understanding seasonality, brand recognition or association with keywords, and data that help you evaluate success of search campaigns in new markets.

**Organizational Support: G**oogle Insights for Search requires very little organization support. As an on-demand service, implementation or configuration are not necessary. Leveraging the results to optimize search campaigns, however, may require more experienced marketing professionals.

### *Search Engine Reports*

**Online Service**

http://searchenginereports.net/

**Tool-Use Case**

- A free online service that reports ranking of a designated domain based on one or more keywords.
- Operates using JQuery, AJAX, and the search engine's API on Amazon's EC2 Web Cloud.
- Users enter the domain URL and desired key words into the online application and click “Create Report.”
- Reporting rankings for both Google.com and Bing.com search engines.
- Also provides a detailed URL view of the ranking search result.

**Tool Pros**

- Free.
- No setup required. Users can quickly enter information and run reports from the Web site.

**Tool Cons**

- Limited analysis capabilities. Only reports ranking
- Cannot export reports to other applications.

**Pricing**

Free.

**Recommendations**

**Type of Site:** Search Engine Reports is a useful, “gut check” tool for a quick check of search engine rankings. It is a good option for small organizations with limited resources to allocate to search engine optimization. Reporting is relatively limited, however, so if possible, more robust SEO tools should be employed in addition.

**Organizational Support**: Search Engine Reports requires very little organization support. As an on-demand service, there is no need to setup or configure any software or application settings. However, because search engine optimization involves a high-degree of data analysis and experimentation, leveraging the results to optimize search campaigns may require more experienced marketing professionals.

### *SEOmoz*

**SaaS**

www.seomoz.org

#### Tool-Use Case

- A fully hosted online SEO software system that monitors search campaigns, provides recommended action steps and reports competitor rankings.
- Automatic weekly crawls of the designated Web site identify potential issues that may impact performance and report search engine rankings for specified keywords.
- Platform includes Open Site Explorer, which monitors competitors' ranking and link metrics.

#### Tool Pros

- Many reporting features can be configured to run automatically.
- Active community and social component. Allows users to stay abreast of SEO-related topics, continually improve knowledge and thereby make more informed decisions and ensure best practices for site content and keyword optimization.
- Provides organizations with actionable recommendations, including suggested keywords and quick site content adjustments that will optimize search results rankings.

#### Tool Cons

- Limited branding control over reports.
- Platform's dashboard is somewhat disorganized, so tools may be difficult to find.

#### Pricing

The Pro plan costs $99/month and allows for up to five campaigns, 300 keywords, and 10,000 pages crawled. The Pro Plus plan is priced at $199 per month, with 12 campaigns, 1,000 keywords, and 10,000 pages crawled. The Pro Elite plan runs $499 for 30 campaigns, 3,500 keywords, and 20,000 pages crawled. This plan also includes branded PDF reports. There is a 20% discount for plans purchased on an annual basis. The Pro plan can be tested with a 30-day free trial.

#### Recommendations

**Type of Site:** Given the robust array of analysis and link building tools and the varied pricing structure, SEOmoz is a good option for organizations of all sizes with a variety of use-cases, including in-house

SEOs, consultancies, and agencies. It is most cost-effective for organizations running at least five campaigns.

**Organizational Support:** Because the SEOmoz platform provides a high-degree of recommended action steps, training opportunities, and community support, there is less time and resource demand on analysts to make sense of the data. This means less-experienced employees could take the responsibility of maintaining.

## Internal Search

### *Introduction*

Internal search tools provide infrastructure for search on an organization's intranet and/or Web site. These tools are intended to increase the efficiency of accessing information and Web content, which in turn can drive customer satisfaction, increased conversion rates, and improved work productivity.

#### Tools Covered in This Section

- Google Search Appliance
- Google Site Search
- Adobe Search& Promote

### *Google Search Appliance (GSA)*

#### On-Premise

www.google.com/enterprise/search/gsa.html

#### Tool-Use Case

- On-premise solution for internal search within an organization's intranet or publicly on the Web.
- Indexes across the various silos that house information, such as intranets, portals, file shares, databases, content management systems, and business applications. Additionally, with the recent introduction of Cloud Connect, the GSA searches cloud based systems, including all Google apps and Twitter. In this manner, the GSA solution breaks down the barriers among these silos, so employees need only to search in one place to access all relevant information.
- For use on a Web site, the GSA allows site visitors to search site content in the same manner they would search on Google.com.

#### Tool Pros

- Highly scalable. Large organizations can seamlessly share indices across multiple appliances such that unified search results will be returned in all instances.
- Easily configures to comply with organizational security regulations. Search results only include information to which the user has access. The GSA supports single sign-on and authentication measures, as well as both early and late binding options.

- Highly customizable. Intelligent biasing features allow organizations to prioritize search results based on a variety of metrics, including: source, date, collection, metadata and node biasing. User-interface can be customized by integrating search results using an XML API.

**Tool Cons**

- Inappropriate for small or midsized organizations.

**Pricing**

Google Search Appliance is sold in licenses based on the number of documents searched. Each license includes all necessary hardware and software, hardware-replacement, product updates, and customer support. For up to 500,000 documents, the Google Search Appliance costs $30,000. A Google sales representative must be contacted to obtain pricing details for larger scale implementations.

**Recommendations**

**Type of Site:** The Google Search Appliance is intended for large enterprise organizations with high volumes of information stored in multiple systems and applications. This device is particularly appropriate for organizations with strong security concerns and the resources available to leverage the vast range of customizations available. For small to mid-sized organizations , the Google Mini provides similar indexing capabilities for up to 300,000 pages of content. The Google Mini costs $2,990 to $9,990 for a two-year license, depending on the number of documents indexed.

**Organizational Support:** Google Search Appliance is extremely intuitive to setup and configure, requiring little intervention from system administrators. A Web -based admin console can be used to make configuration changes. However, in order to employ GSA's extensive customization features, organizations will need to involve administrators with experience using API's.

### *Google Site Search*

**SaaS**

www.google.com/sitesearch

**Tool-Use Case**

- Provides site search functionality for Web sites using Google.com search technology.
- Customizable to match site appearance. Basic customizations available via a straightforward Web interface. Complete control over search results format can be achieved by using the XML API.
- Unlike Google.com, search results are displayed without ads.

**Tool Pros**

- Inexpensive.
- Extremely easy to implement on a Web site. Setup wizard allows users to easily input information and then paste generated code onto their site.

- Date and top results biasing allow administrators to influence rankings of search results based on the age of the content and area on the site.
- User-friendly and familiar. Given the ubiquity of Google.com search, site visitors are likely to easily navigate similarly formatted search results using Google Site Search.

**Tool Cons**

- Limited functionality. Google site search only provides indexing and search results.
- Does not provide analytics for visitor search behavior.

**Pricing**

Google Site Search is pricing plans are based on the number of search queries per year. For $100 annually, users can have up to 20,000 search queries. The next level allows 50,000 queries for $250 per year. For up to 150,000 queries an annual license costs $750. At $2,000 annually, the Google Site Search allows 500,000 queries. For organizations requiring more than 500,000 queries, contact Google sales directly.

**Recommendations**

**Type of Site:** The availability of varied pricing plans makes Google Site Search a good option for organizations s of all sizes. However, given the absence of reporting or analytical capabilities, this application is best for organizations that have other mechanisms for tracking visitor search behavior, such as tag-based Web analytics applications.

**Organizational Support:** Google Site Search requires relatively little organizational support, in both the implementation and maintenance phases. Configuring the application is simple using the Web -based setup wizard. The most time and resource intensive aspect of using Google Site Search will be ensuring the most up-to-date pages are indexed.

### *Adobe Search&Promote*

**SaaS**

www.omniture.com/SiteSearch

**Tool-Use Case**

- A hosted, scalable site search application intended to strategically optimize how visitors access information and products using internal site search in order to drive higher rates of conversion, engagement, and average order value.
- Part of the Adobe Online Marketing Suite, powered by Omniture.
- Offers a high level of control over visitors' search experiences. Marketers can analyze search behavior using built-in SiteCatalyst reporting, anticipate visitor intent and create formulas for generating search results with relevant and strategically-promoted products and content.

**Tool Pros**

- Search and navigation interfaces are dynamic, ensuring the highest quality, most efficient experience for each site visitor on every visit. Auto-complete and "did you mean" help guide search queries, while filtering and related content features help refine results.
- Admin console is both powerful and user-friendly. For example, marketers can easily build rules for specific search scenarios (e.g., a specific search term triggers the display of a given promotional banner) by replicating a visitor's experience on their Web site and adjusting content accordingly. Additionally, administrators can adjust the relative influence of natural relevance versus ranking rules (e.g., days old, inventory, or product revenue) using a sliding scale.
- Highly scalable—tool can handle millions of site pages and high search volumes.

**Tool Cons**

- Best used in conjunction with other applications in the Adobe Online Marketing Suite, thus making it less accessible to organizations not currently using these products.
- For maximum cost effectiveness, Search&Promote requires a significant time-investment for analyzing visitor search behavior and programming the application to surface the best results for each search instance.

**Pricing**

**Recommendations**

**Type of Site:** Adobe Search&Promote is most appropriate for enterprise-level organizations with high volumes of search traffic and strategic marketing endeavors. Organizations looking for a high-level of control, and personalization of, as well as extensive reporting analytics for, visitor search behavior will benefit from this application.

**Organizational Support:** Search&Promote’s admin interface is relatively intuitive and user-friendly, and therefore could be configured by anyone in the organization. However, to ensure maximum benefit from the tool, marketing professionals should be reviewing reporting data and making key decisions regarding relevance, ranking, rules, etc. This aspect of the application could require a significant time investment from these decision-making individuals.

## RSS

### *Introduction*

RSS analytics tools are primarily used to see how many RSS subscribers are associated with a particular feed. Most of the RSS analytics tools are actually RSS management tools and can be used to manage feeds in addition to measuring the feed usage.

### Tools Covered in This Section

- FeedBlitz
- FeedManager

## *FeedBlitz*

http://www.feedblitz.com/f/f.fbz?Rsshome

### Tool-Use Case

- Provides comprehensive RSS feed management services, including feed settings and customization, social media integrations, analytics, and the ability to diagnose and remediate issues with feed distribution.
- Reporting tools allows users to track subscriber counts and track links (including internal post links).

### Tool Pros

- Gives users a high level of control over their RSS feed features and settings, as compared to other RSS management systems.
- Automatically creates mobile versions of RSS feeds.
- Handles long feeds, as compared to other RSS management systems.
- Capable of reading feeds that are password protected with HTTP-authentication.

### Tool Cons

- Pricing based on email subscribers can drive cost up for popular, but individually managed, blogs.

### Pricing

Pricing is based on the number of email subscribers to the feed, regardless of the number of feed subscribers. Prices range from $1.49/month for 0-9 email subscribers to $600/month for up to 199,999 subscribers. Each additional 50,000 subscribers costs $100 per month.

### Recommendations

**Type of Site:** FeedBlitz is recommended for a wide range of uses, from personal blogs to enterprise-level organizations RSS feed management. Because the pricing structure is based solely on email subscriber count, users of any size have access to all of the tool's features.

**Organizational Support:** Little organizational support is required for managing RSS feeds using FeedBlitz, according to the company.  Feed manager only needs to allocate a few hours per month to set up, publish and maintain the feeds as well as review the analytics data.

## *FeedManager*

http://www.rapidfeeds.com/

### Tool-Use Case

- Using the FeedManager platform, customers can create, manage and publish RSS feeds.
- Platform's reporting capabilities allow users to track feeds over time, subscriber demographics and link clicks.

### Tool Pros

- Inexpensive.
- Automation features, such as prescheduled publishing of new content and Twitter updates.
- Allows for podcast integration in RSS feeds.

### Tool Cons

- Analytics are only available with the more expensive plans.

### Pricing

For $4.49 per month, the Basic plan allows for three RSS feeds, unlimited number of items, publishing scheduling, password protection, and auto-Tweeting. At $6.95 per month, the Pro plan, allows seven feeds, analytics, and branded feed URLs. The Enterprise plan costs $13.95 per month and includes unlimited feeds and prioritized customer support.

### Recommendations

**Type of Site:** FeedManager is appropriate for a wide range of organizations. FeedManager is currently used by several enterprise-level clients.

**Organizational Support:** Using FeedManager requires relatively little organizational support. Implementation and upkeep is extremely straightforward and will likely require only a few hours a month time commitment.

# Appendix J.
# Strategies for Google Analytics Implementation

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com

December 15 , 2011

# Appendix J. Strategies for Google Analytics Implementation

## Introduction

Google Analytics is gaining popularity among NIH ICs that have been either frustrated with their current Web analytics tools or lack the budget to purchase a fee-based solution.

The first section of this appendix describes the process for ICs to obtain Google Analytics. The second section presents information about how ICs can implement Google Analytics.

## Getting Started with Google Analytics at NIH

NIH ICs are allowed to use Google Analytics, but they must follow these steps to obtain and implement it.

**Step 1: Read policy guidance from OMB and HHS**: Links to both are below.

Office of Management and Budget (OMB) Guidance for Online Use of Web Measurement and Customization Technologies OMB M-10-22
http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-22.pdf

Michael W. Carleton, HHS Chief Information Officer and Senior Agency Official for Privacy Implementation of OMB M-10-22 and M-10-23
www.hhs.gov/ocio/policy/policydocs/implementation_of_omb_m-10-22_and_m-10-23.doc - 2010-12-21

**Step 2: Contact your IC's privacy coordinator or the NIH senior official for privacy about your plans to use Google Analytics.** Ensure that your plans comply with Federal and IC privacy policies and guidelines. Here is the link to NIH privacy coordinators:
http://oma.od.nih.gov/about/contact/browse.asp?fa_id=3

**Step 3: Contact your IC's information systems security officer (ISSO) to confirm that your plans comply with both NIH and Federal policies.**

**Step 4: Create a Google account with your NIH email address.** Then you can obtain your Google Analytics code and NIH analytics data using that account. Here is link to the Google account setup page: http://bit.ly/ij9fiu

**Step 5: Important! Get Google Analytics tool from Apps.gov**. All ICs must notify the HHS New Media Office when using Google Analytics. When your IC goes through the Apps.gov process to get Google Analytics, the HHS New Media Office will be notified automatically. This is the link to Google Analytics page on Apps.gov http://bit.ly/l0mxje.

**Step 7: Develop or modify your site's privacy policy or notice to reflect your use of Google Analytics.** Follow the policy guidance document OMB M-10-22 (listed under Step 1 above). You can use the NIH privacy notice as a guide, http://1.usa.gov/uzjRd2.

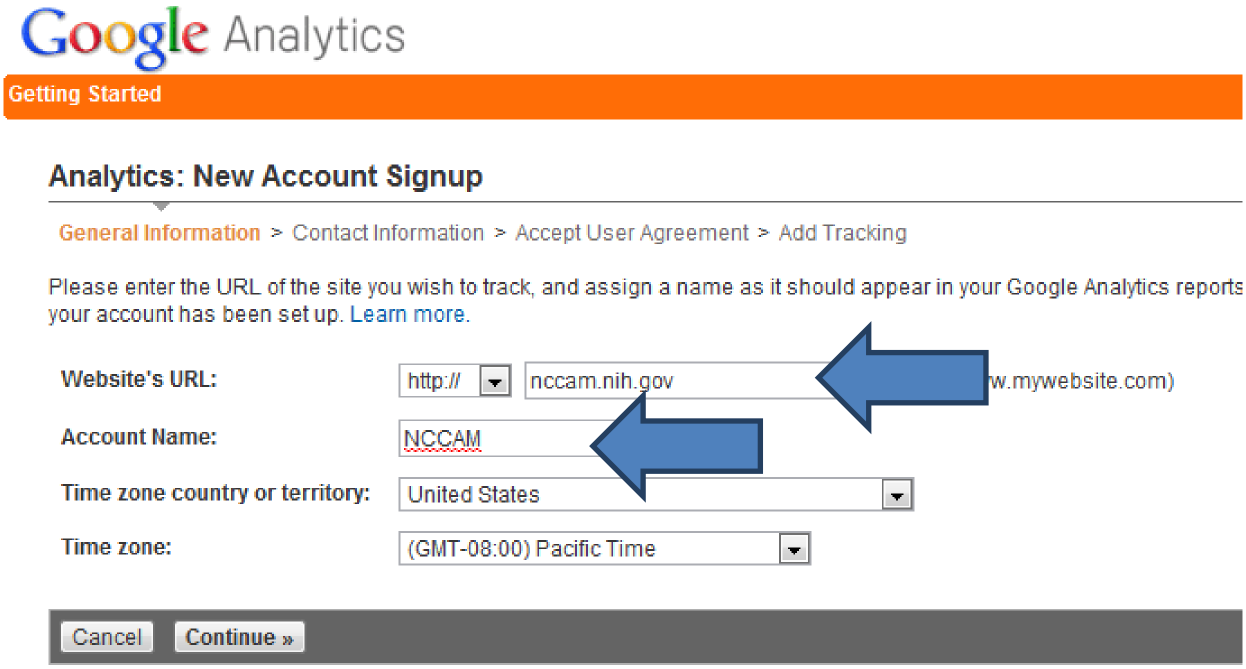
Once you complete the steps above, you are ready to start setting up Google Analytics. If you need additional help, please refer to the Google Analytics Help site, http://bit.ly/u6OWZn.

## Implementing Google Analytics for Institutes and Centers (ICs)

To implement Google Analytics, NIH ICs need to follow these steps to ensure high-quality data collection.

A. Create a new account for the specific IC Web site by visiting www.google.com/analytics and using the Google account you created in Step 4 above. See **Figure J.1**.

***Figure J.1.*** *Sample screen shot of Google Analytics new account signup page.*


Google Analytics
Getting Started

Analytics: New Account Signup

General Information > Contact Information > Accept User Agreement > Add Tracking

Please enter the URL of the site you wish to track, and assign a name as it should appear in your Google Analytics reports
your account has been set up. Learn more.

Website's URL: http:// nccam.nih.gov w.mywebsite.com)

Account Name: NCCAM

Time zone country or territory: United States

Time zone: (GMT-08:00) Pacific Time

For Web site URL, enter the URL of your top-level domain, as shown in **Figure J.1.** For account name, you will probably use your IC's name. Once everything is correct, click continue.

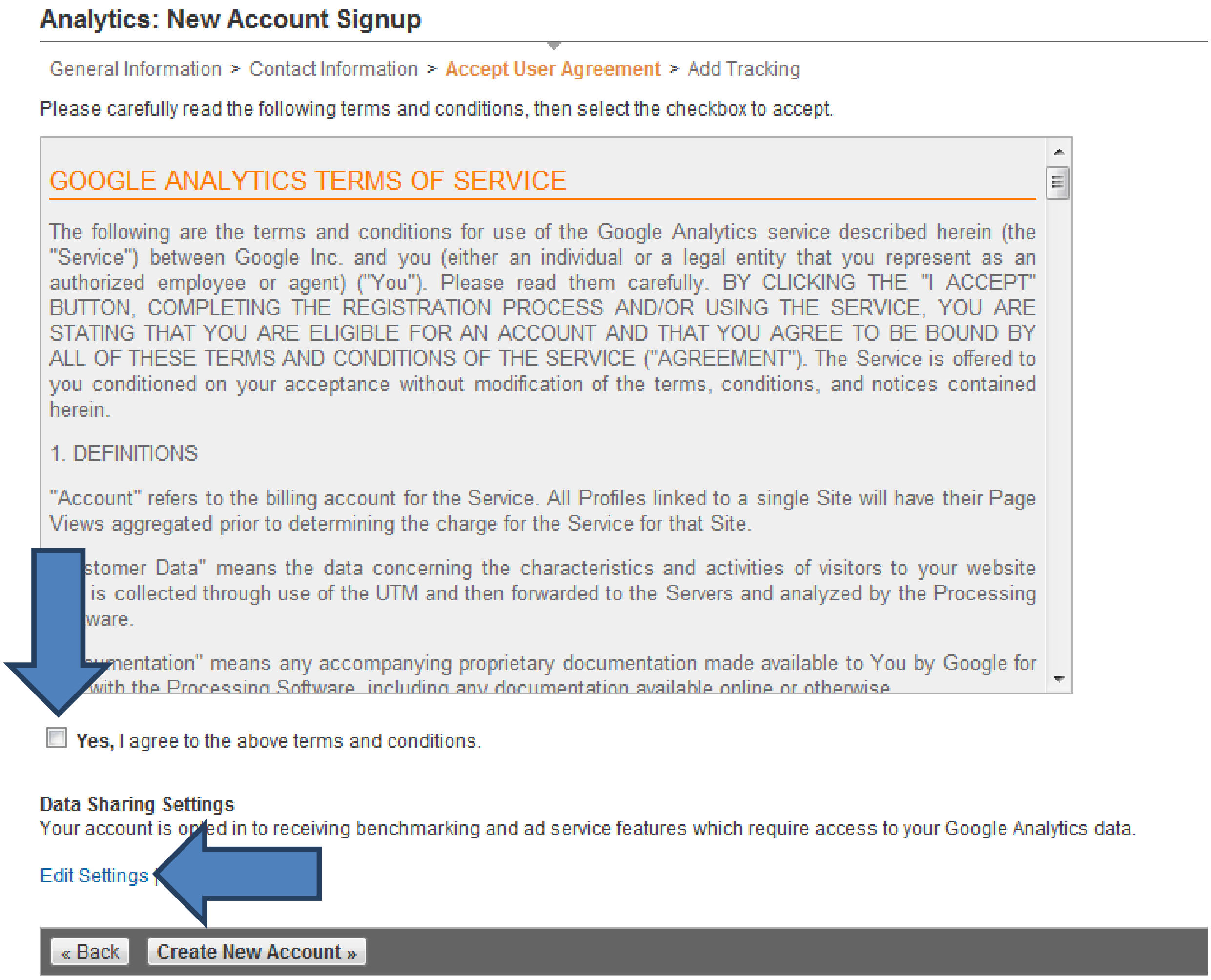
B. The next step will be to add the Web or communication director's contact information, including last name, first name, and country, and then click continue.

C. Next, accept the terms of service (TOS) for Google Analytics (see **Figure J.2**). HHS has already signed a "Federal -friendly" TOS agreement for Google Analytics that covers NIH. This means that your IC does not need to create a new one.

When you are installing Google Analytics, you will be asked to accept the TOS agreement. Once you accept it, the Google Analytics folks will automatically know that you are agreeing to the Federal TOS.

You will need to check the “Yes” box and then click on “Edit Settings” under “Data Sharing Settings.” You can either check the box that says, “Do not share my Google Analytics data" or “Share Google Analytics data anonymously with Google and others.” With the second option, Google will remove all identifiable information about your Web site data. Then, Google will combine your data with data from other anonymous sites in similar industries and create reports on aggregate data. Choosing this option will help Google improve its products and allow you to get benchmarking reports. Then click “Create New Account.”

**Figure J.2. Screenshot of Google Analytics terms of service page.**



Analytics: New Account Signup

General Information > Contact Information > Accept User Agreement > Add Tracking

Please carefully read the following terms and conditions, then select the checkbox to accept.

GOOGLE ANALYTICS TERMS OF SERVICE

The following are the terms and conditions for use of the Google Analytics service described herein (the "Service") between Google Inc. and you (either an individual or a legal entity that you represent as an authorized employee or agent) ("You"). Please read them carefully. BY CLICKING THE "I ACCEPT" BUTTON, COMPLETING THE REGISTRATION PROCESS AND/OR USING THE SERVICE, YOU ARE STATING THAT YOU ARE ELIGIBLE FOR AN ACCOUNT AND THAT YOU AGREE TO BE BOUND BY ALL OF THESE TERMS AND CONDITIONS OF THE SERVICE ("AGREEMENT"). The Service is offered to you conditioned on your acceptance without modification of the terms, conditions, and notices contained herein.

1. DEFINITIONS

"Account" refers to the billing account for the Service. All Profiles linked to a single Site will have their Page Views aggregated prior to determining the charge for the Service for that Site.

stomer Data" means the data concerning the characteristics and activities of visitors to your website is collected through use of the UTM and then forwarded to the Servers and analyzed by the Processing ware.

umentation" means any accompanying proprietary documentation made available to You by Google for with the Processing Software, including any documentation available online or otherwise.

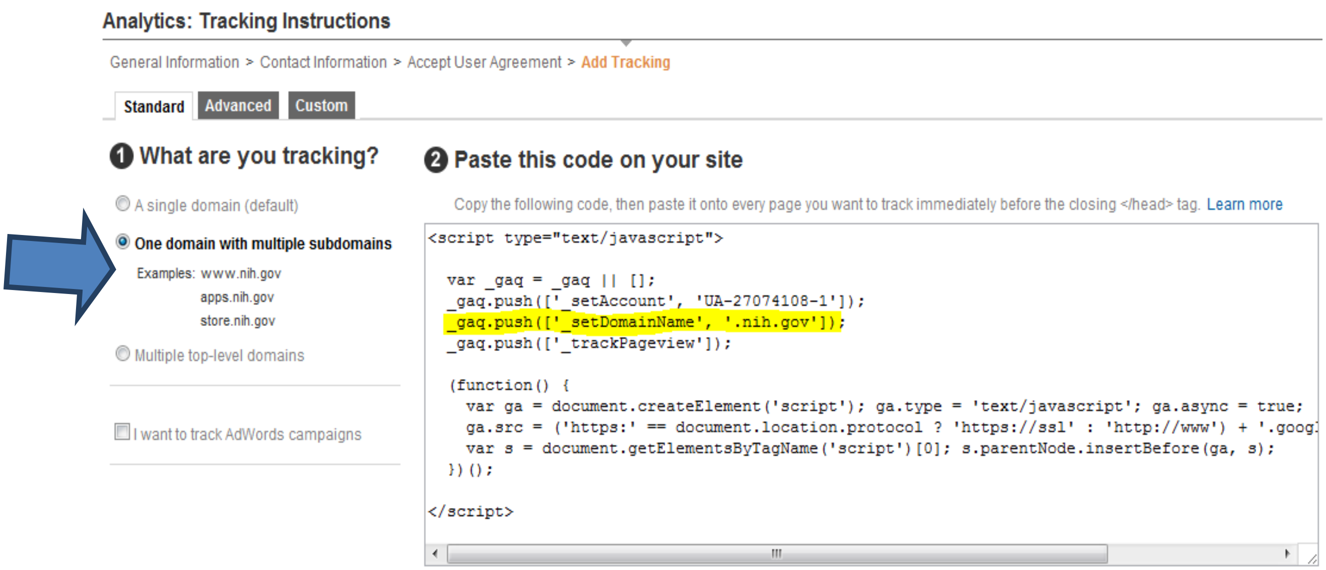
☐ **Yes,** I agree to the above terms and conditions.

**Data Sharing Settings**
Your account is opted in to receiving benchmarking and ad service features which require access to your Google Analytics data.

Edit Settings

D. In this step, you will receive the tracking code to be copied and placed onto the Web site. Select “One domain with multiple subdomains ,as seen in **Figure J.3.,** copy the code to a Word document or Notepad text file , and select “Save and Finish.”

***Figure J.3.** Screenshot of Google Analytics tracking instructions.*



Analytics: Tracking Instructions

General Information > Contact Information > Accept User Agreement > Add Tracking

1 What are you tracking?

○ A single domain (default)

● One domain with multiple subdomains

Examples: www.nih.gov
apps.nih.gov
store.nih.gov

○ Multiple top-level domains

☐ I want to track AdWords campaigns

2 Paste this code on your site

Copy the following code, then paste it onto every page you want to track immediately before the closing </head> tag. Learn more

```
<script type="text/javascript">

  var _gaq = _gaq || [];
  _gaq.push(['_setAccount', 'UA-27074108-1']);
  _gaq.push(['_setDomainName', '.nih.gov']);
  _gaq.push(['_trackPageview']);

  (function() {
    var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
    ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') + '.goog
    var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
  })();

</script>
```

Before pasting this code onto your site, you will need to confirm that the domain name contained between the single quotations is the correct domain for your Web site. If it is not, you will need to make a small change to the code, highlighted in **Figure J.3**. and shown below.

```
_gaq.push(['_setDomainName', '.nih.gov']);
```

For example, if your site's URL is nccam.nih.gov, then "nccam.nih.gov" must appear between the single quotations.

If the code says something else, you need to revise the Google code so the domain in the code follows this convention:

.(IC_site).nih.gov – for subdomains of NIH.gov
.(IC_site).gov – for standalone Web sites
Note that the dot "." needs to come before the IC-site domain name.

E. Next, copy and paste the corrected Google Analytics code on to all pages on your Web site including any subdomains, such as videolectures.nccam.nih.gov.

## Google Analytics Profiles

Google Analytics uses "profiles" to allow users to separate and view  Web site data. A Web site profile is a set of rules that define the reports that Google analytics users can see. Usually, the Web site profile corresponds with a domain, so there will be one profile for each domain. It is also possible to create profiles to track subdomains by setting up filters that enable Google Analytics to show data for a particular subdomain or section of the site. (Source: http://www.google.com/support/analytics/bin/answer.py?answer=55595)

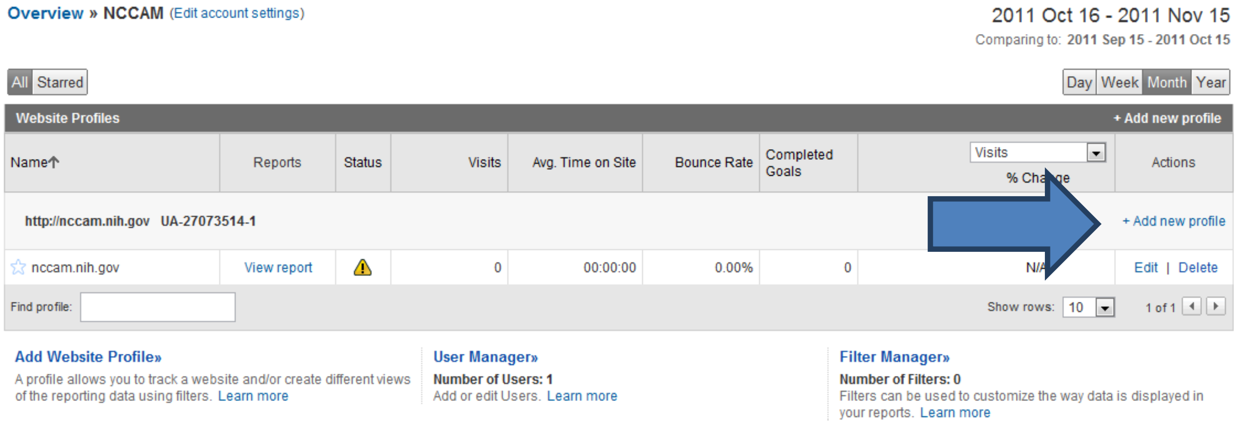
Profiles also allow the Google Analytics administrator to set permissions that allow certain users to access only selected profiles. For example, an administrator may want to prevent users working on one subdomain from looking at reports on another domain within an IC.

To create separate profiles, you will need to filter the data coming into Google Analytics. Because filtering the data occurs before they are presented in the reports, you should always keep one profile completely unfiltered as a backup in case of filtering errors. Using NCCAM as an example, the processes for filtering the data for certain subdomains and creating an unfiltered profile are shown below.

Follow these steps to create a new profile:

1. Create a profile for each of your subdomains inside Google Analytics. You can do this from the home page of your Google Analytics account by clicking “add new profile” (**Figure J.4**.).

***Figure J.4.*** *Screen shot of home page of a Google Analytics account.*



2011 Oct 16 - 2011 Nov 15
Comparing to: 2011 Sep 15 - 2011 Oct 15

| Name↑ | Reports | Status | Visits | Avg. Time on Site | Bounce Rate | Completed Goals | Visits<br>% Change | Actions |
|---|---|---|---|---|---|---|---|---|
| http://nccam.nih.gov UA-27073514-1 | | | | | | | | + Add new profile |
| ☆ nccam.nih.gov | View report | ⚠ | 0 | 00:00:00 | 0.00% | 0 | N/A | Edit \| Delete |

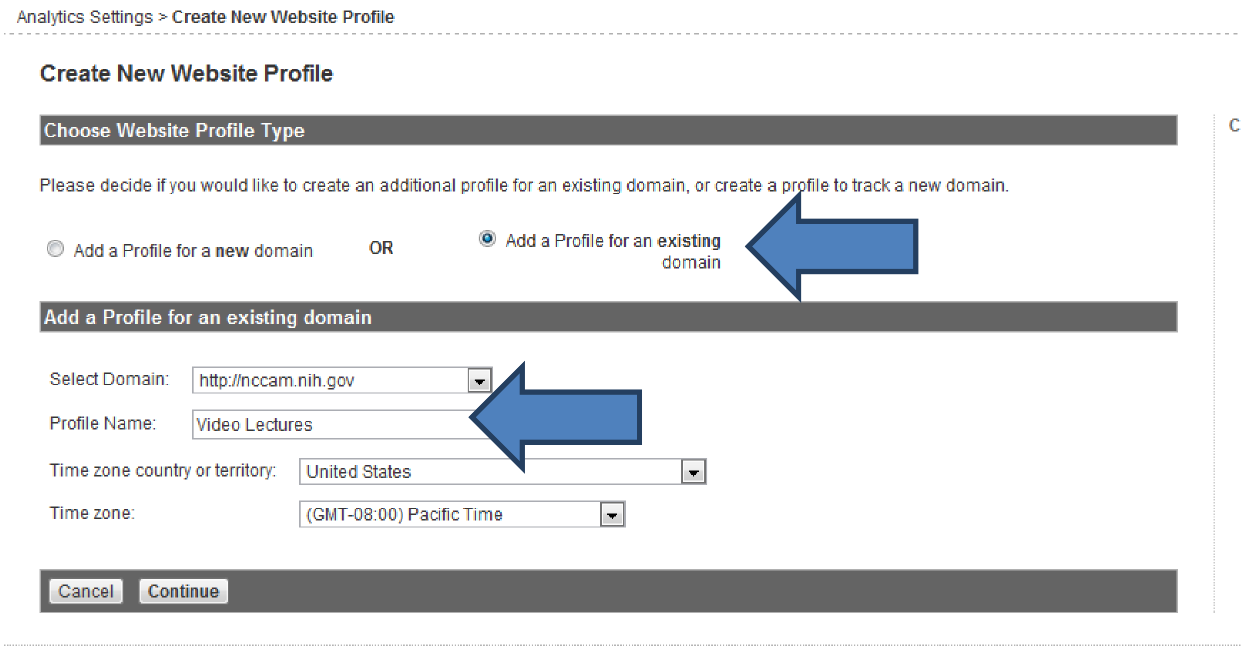
**Add Website Profile»**
A profile allows you to track a website and/or create different views of the reporting data using filters. Learn more

**User Manager»**
Number of Users: 1
Add or edit Users. Learn more

**Filter Manager»**
Number of Filters: 0
Filters can be used to customize the way data is displayed in your reports. Learn more

2. Confirm that “Add a Profile for an existing domain” is selected and then name your profile for the subdomain you are setting up. (See **Figure J.5**.)

***Figure J.5.*** *Screen shot of Google Analytics "Create New Web site Profile" page.*



**Create New Website Profile**

**Choose Website Profile Type**

Please decide if you would like to create an additional profile for an existing domain, or create a profile to track a new domain.

○ Add a Profile for a **new** domain OR ● Add a Profile for an **existing** domain

**Add a Profile for an existing domain**

Select Domain: http://nccam.nih.gov

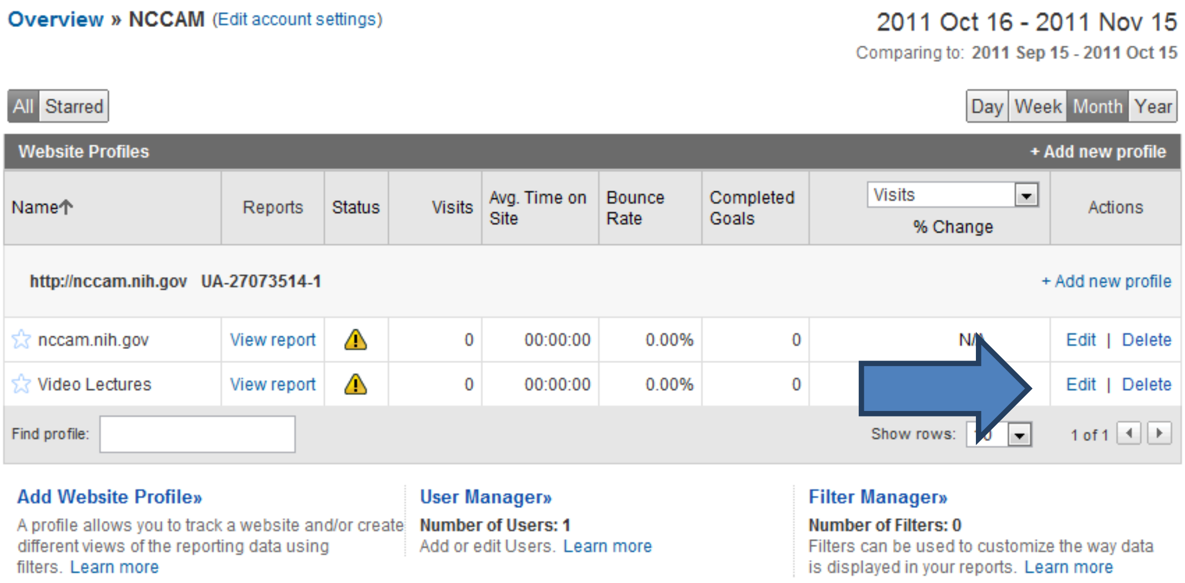
Profile Name: Video Lectures

Time zone country or territory: United States

Time zone: (GMT-08:00) Pacific Time

3. Once you have set up the new sub domain profile, create an advanced filter so traffic for just that subdomain will appear in the profile. To do this, click on "Edit" from the overview page of Google Analytics (**Figure J.6**).

***Figure J.6.*** *Screen shot of Google Analytics overview page.*



2011 Oct 16 - 2011 Nov 15

Comparing to: **2011 Sep 15 - 2011 Oct 15**

**http://nccam.nih.gov UA-27073514-1**

| Name | Reports | Status | Visits | Avg. Time on Site | Bounce Rate | Completed Goals | Visits % Change | Actions |
|---|---|---|---|---|---|---|---|---|
| nccam.nih.gov | View report | ⚠ | 0 | 00:00:00 | 0.00% | 0 | N/A | Edit \| Delete |
| Video Lectures | View report | ⚠ | 0 | 00:00:00 | 0.00% | 0 | | Edit \| Delete |

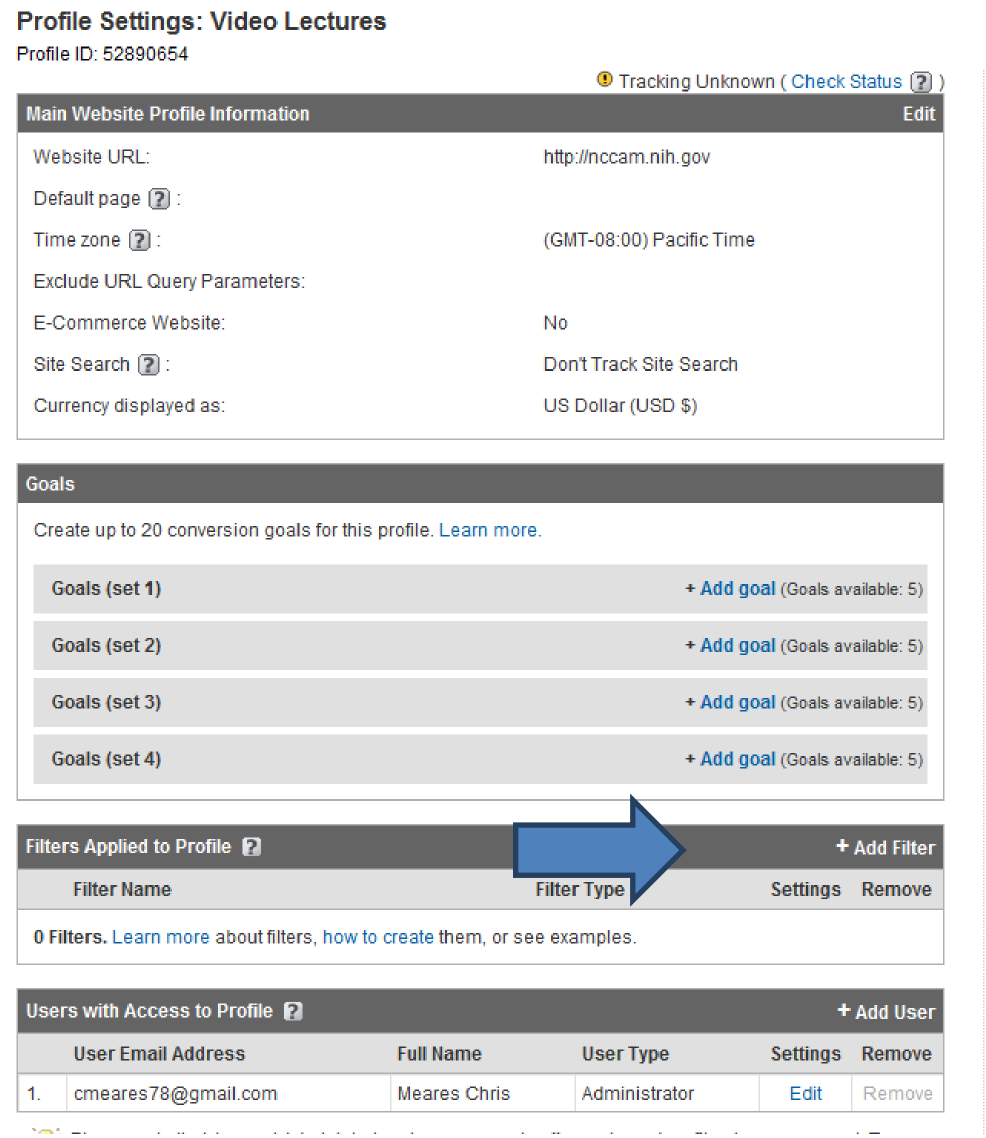
**Add Website Profile»**
A profile allows you to track a website and/or create different views of the reporting data using filters. Learn more

**User Manager»**
**Number of Users: 1**
Add or edit Users. Learn more

**Filter Manager»**
**Number of Filters: 0**
Filters can be used to customize the way data is displayed in your reports. Learn more

4. Create a filter to allow Google Analytics to populate the profile with data. Select "Add new Filter for Profile" and then name the filter (see **Figure J.7**).

***Figure J.7.*** *Screen shot of Google Analytics profile settings page.*



**Profile Settings: Video Lectures**

Profile ID: 52890654

Tracking Unknown ( Check Status ? )

**Main Website Profile Information** Edit

| | |
|---|---|
| Website URL: | http://nccam.nih.gov |
| Default page ? : | |
| Time zone ? : | (GMT-08:00) Pacific Time |
| Exclude URL Query Parameters: | |
| E-Commerce Website: | No |
| Site Search ? : | Don't Track Site Search |
| Currency displayed as: | US Dollar (USD $) |

**Goals**

Create up to 20 conversion goals for this profile. Learn more.

| | |
|---|---|
| Goals (set 1) | + Add goal (Goals available: 5) |
| Goals (set 2) | + Add goal (Goals available: 5) |
| Goals (set 3) | + Add goal (Goals available: 5) |
| Goals (set 4) | + Add goal (Goals available: 5) |

**Filters Applied to Profile** ? + Add Filter

| Filter Name | Filter Type | Settings | Remove |
|---|---|---|---|

**0 Filters.** Learn more about filters, how to create them, or see examples.

**Users with Access to Profile** ? + Add User

| | User Email Address | Full Name | User Type | Settings | Remove |
|---|---|---|---|---|---|
| 1. | cmeares78@gmail.com | Meares Chris | Administrator | Edit | Remove |

5. If you are filtering for a subdomain, such as videolectures.nccam.nih.gov, then create a custom filter by selecting "Custom Filter" under "Filter Name" (see **Figure J.8**).

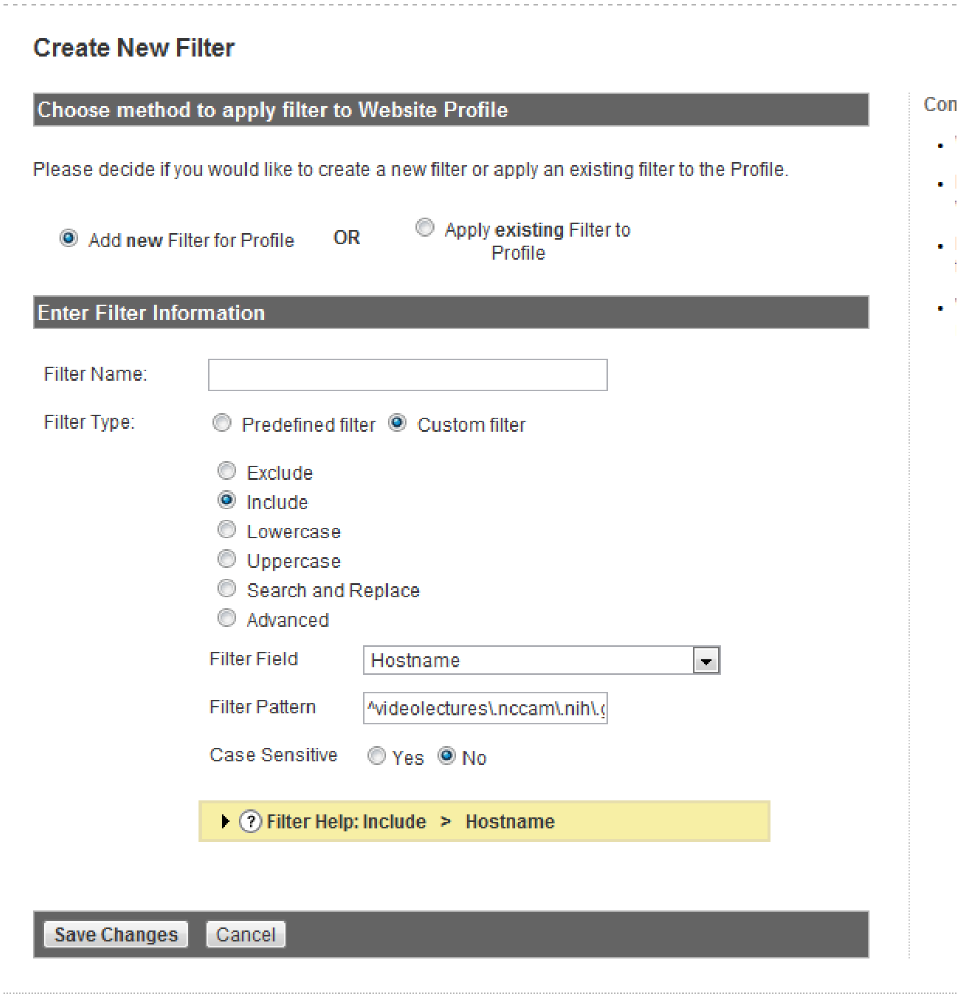
Select "Include," and the "Filter Field" will be "Hostname."

The "Filter Pattern" will be

^videolectures\.nccam\.nih\.gov$ – for videolectures.nccam.nih.gov

^(sub domain)\.nccam\.nih\.gov$ – (subdomain) equals what you are filtering for

***Figure F.8.*** *Screen shot of the Google Analytics "Create New Filter" page.*



Create New Filter

Choose method to apply filter to Website Profile

Please decide if you would like to create a new filter or apply an existing filter to the Profile.

● Add **new** Filter for Profile OR ○ Apply **existing** Filter to Profile

Enter Filter Information

Filter Name:

Filter Type: ○ Predefined filter ● Custom filter

○ Exclude
● Include
○ Lowercase
○ Uppercase
○ Search and Replace
○ Advanced

Filter Field: Hostname

Filter Pattern: ^videolectures\.nccam\.nih\.

Case Sensitive: ○ Yes ● No

You can create filters for populating profiles or separating specific subdomains, such as the site's "training" section. Each profile can have its own set of permissions and users.

## Google Analytics Limitations and Restrictions

Google Analytics is a free solution that has limitations and restrictions not found among fee-based tools.

Below is a list of potential limitations and restrictions:

- No vendor support or help desk – Google does not provide customer support, as do fee-based vendors. To obtain support, NIH will need to train internal staff or purchase technical support services from consultants or third-party organizations that have Google Analytics experience.
- Data storage – Google says it will archive your data for 24 months. We have seen instances where all historical data has been maintained by Google beyond 24 months. This data-retention

service can change at any time that Google decides to trim the amount of archived data it maintains.

- Data sampling – Google provides reports based on data samples when your site has more than a certain number of visits during in a certain date range. For example, if your site has more than 500,000 visits during your selected date range but you were only looking at a single page with about 1,000 page views, your visit counts will be based on sampled data. This is because Google Analytics has to process all of your site’s visit data in the selected date range. In a sampled- data report, some results will be from a statistical sample of site visits.

## Summary

Google Analytics offers NIH a free, user-friendly Web analytics tool. We believe that deploying Google Analytics will enable ICs to gain insights about visitors, visits, task completions, and more. It will also allow for trans-NIH analytics and opportunities for benchmarking. Since Google Analytics is a free tool, it has some limitations and restrictions are not found in fee-based tools.

# Appendix K.

# Piwik Evaluation: A Free, Open-Source Web Analytics Tool

## Contents

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix K. Piwik Evaluation: A Free, Open-Source Web Analytics Tool

## Introduction

As part of its work with NIH, Semphonic has made an effort to identify, outline, and make recommendations about tools that provide different types of analytics, including Web and social media analytics. A  complete list of these tools is in **Appendix I.**

While preparing Appendix I, we discovered one tool of particular note called Piwik (www.piwik.org).  It is an open-source Web analytics tool with several characteristics that make it an interesting option for NIH. This appendix provides information about the Piwik tool.

## Tool Capabilities

### *Overview*

Piwik was designed to be an open-source alternative to Google Analytics. This means that it's free and can handle enterprise-level implementations. It also has many of the same reporting capabilities—like page views, visits, visitors, referrers, campaigns, and Geolocation—and even a similar look and feel. Like most Web analytics tools, it can be customized, particularly through setting goals and creating custom variables.

Unlike most other Web analytics tools in use today, Piwik is not a software-as-a-service (SaaS) product. It is an "on premise" solution, which means that NIH ICs can install it on their own servers, manage it, and control the data.

### *Installation*

The installation process is quick and, like other Web analytics tools, it involves setting up the database, setting up tables in the database, identifying the administrative users and the Web site to be measured, and generating the JavaScript code for tracking. The software that goes on the server is built with PHP, a general-purpose, server-side scripting language similar to the Microsoft Active Server Pages server-side script engine. Using Piwik requires a developer with knowledge of both PHP and Structured Query Language (SQL). SQL is a language used for adding, removing, or requesting information from databases populated with Web analytics data.

Configuring Piwik occurs in two different places. Piwik has an administrative interface for some settings, while other settings are configured in the source code.

### *Tracking*

Piwik tracks using JavaScript page tagging, as do Google Analytics, Omniture, and most popular Web analytics tools. Standard page code is placed on all the pages on a site to provide out-of-the-box tracking items and can be customized for additional reporting.

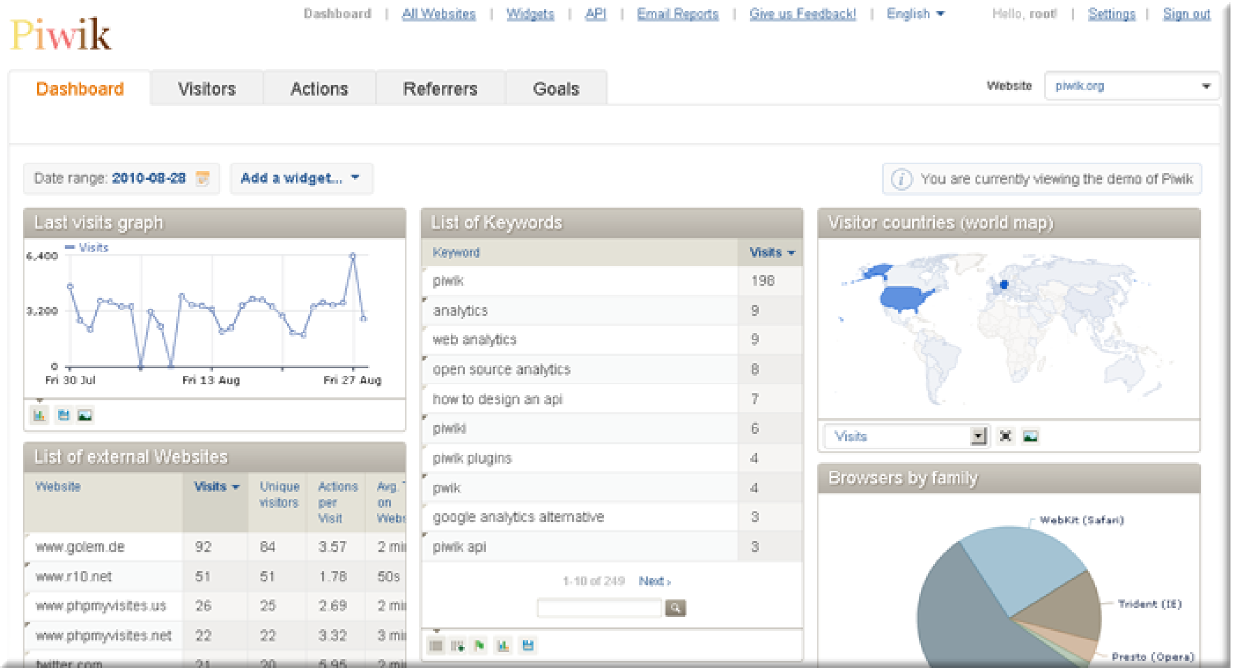
From the standard page code, Piwik populates a page name, which defaults to the URL. It can also be customized to include a subdomain specification so that you can track multiple subdomains.

The Piwik JavaScript code includes additional functions, such as those that facilitate the tracking of custom variables, custom goals, and links. While there are limitations to all three of these tracking features, Piwik's capabilities are still within the range of the expected capabilities of an analytics tool. Goals can be tied to a specific page URL, (customized) title, download, or link click. Goals can be set up to be counted only once—or multiple times—per visit. A maximum of 10 custom variables can be assigned per page, with 5 pertaining to characteristics of the specific page and 5 pertaining to characteristics of the entire visit.

### *Reporting*

The Piwik reporting interface offers a clean visual presentation that displays overview reports in a manner very similar to Google Analytics.

***Figure K.1.*** *Piwik home screen.*



Date range: 2010-08-28

You are currently viewing the demo of Piwik

List of Keywords

| Keyword | Visits |
|---|---|
| piwik | 198 |
| analytics | 9 |
| web analytics | 9 |
| open source analytics | 8 |
| how to design an api | 7 |
| piwiki | 6 |
| piwik plugins | 4 |
| pwik | 4 |
| google analytics alternative | 3 |
| piwik api | 3 |

List of external Websites

| Website | Visits | Unique visitors | Actions per Visit | Avg. on Webs |
|---|---|---|---|---|
| www.golem.de | 92 | 84 | 3.57 | 2 mi |
| www.r10.net | 51 | 51 | 1.78 | 50s |
| www.phpmyvisites.us | 26 | 25 | 2.69 | 2 mi |
| www.phpmyvisites.net | 22 | 22 | 3.32 | 3 mi |
| twitter.com | 21 | 20 | 5.95 | 2 mi |

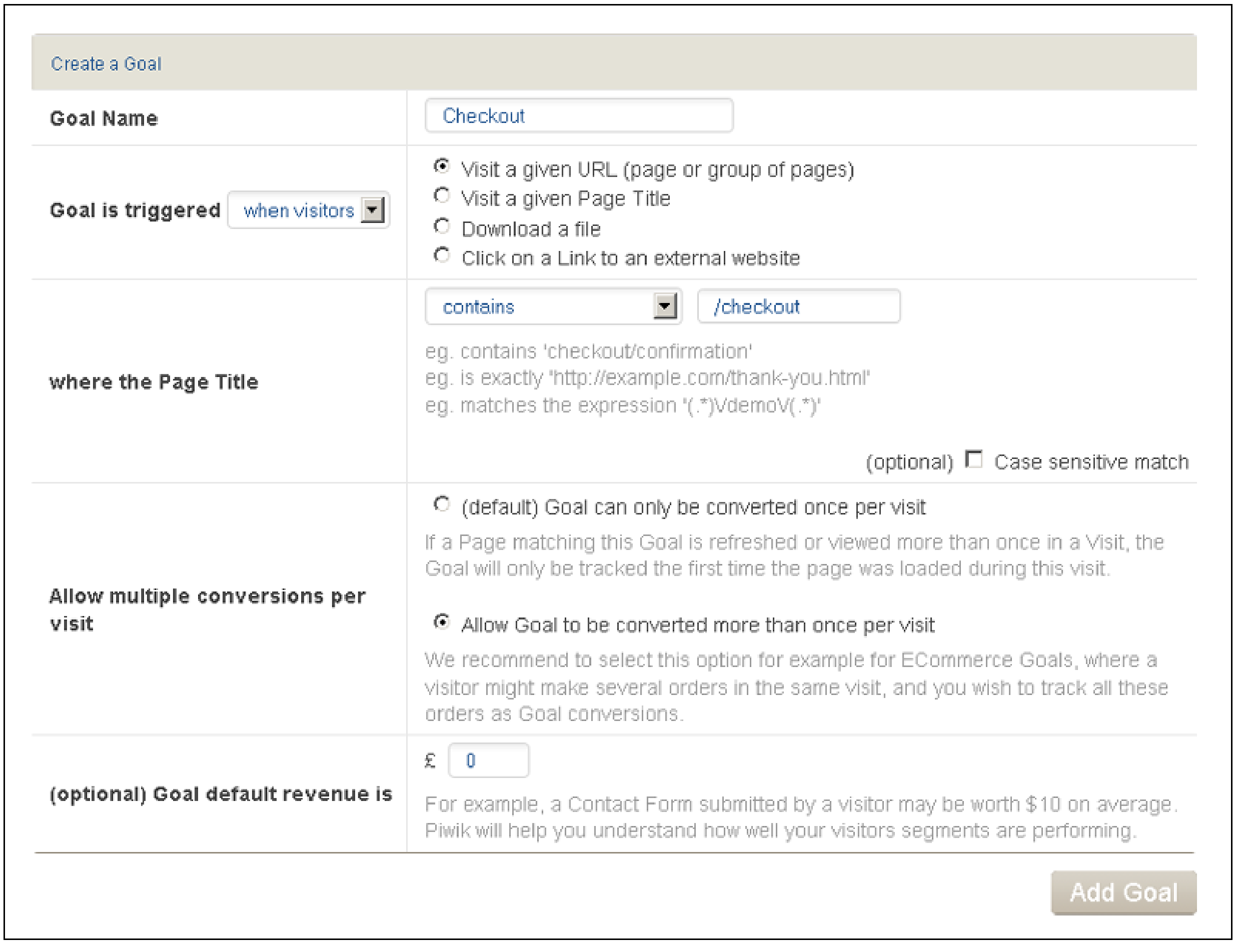
Piwik has a user-friendly interface that allows for running most basic reports. The reports available out of the box in the Piwik interface are page views, visits, unique visitors, time on site, entry and exit pages, downloads, operating system, visitors' locations, referrers, and keywords. The interface is intuitive and provides an easy way to view and export data. Piwik's ability to generate custom dashboards with multiple reports in the interface and to easily export graphs as images makes it a very convenient tool. It can also handle ecommerce sites, so it can report on functions such as fee-for-service activities.

### Goals

Goals are used to track any kind of conversion or success event on the site. They can be triggered on a page or by an action and are used to evaluate the effectiveness of a channel or keyword.

The Piwik goal-configuration screen allows users to manually configure the goals that are relevant to their sites (**Figure K.2**).

***Figure K.2.*** *Screen shot of the Piwik Goal Wizard.*



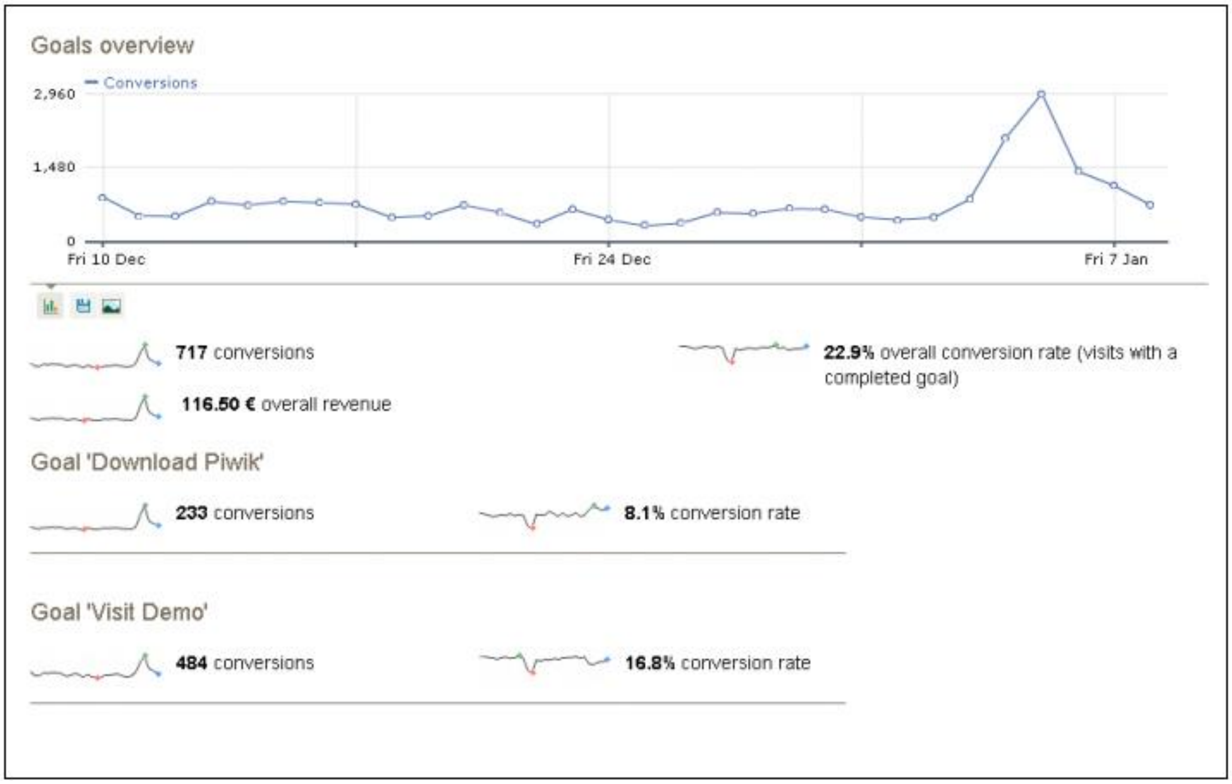
Create a Goal

| | |
|---|---|
| **Goal Name** | Checkout |
| **Goal is triggered** when visitors | ● Visit a given URL (page or group of pages)<br>○ Visit a given Page Title<br>○ Download a file<br>○ Click on a Link to an external website |
| **where the Page Title** | contains /checkout<br>eg. contains 'checkout/confirmation'<br>eg. is exactly 'http://example.com/thank-you.html'<br>eg. matches the expression '(.*)\/demo\/(.*)'<br>(optional) ☐ Case sensitive match |
| **Allow multiple conversions per visit** | ○ (default) Goal can only be converted once per visit<br>If a Page matching this Goal is refreshed or viewed more than once in a Visit, the Goal will only be tracked the first time the page was loaded during this visit.<br>● Allow Goal to be converted more than once per visit<br>We recommend to select this option for example for ECommerce Goals, where a visitor might make several orders in the same visit, and you wish to track all these orders as Goal conversions. |
| **(optional) Goal default revenue is** | £ 0<br>For example, a Contact Form submitted by a visitor may be worth $10 on average. Piwik will help you understand how well your visitors segments are performing. |

Visits, conversions, and conversion rate are available by visitor location, custom variable, server time, keyword, search engine, Web site, campaign, and referrer-type reports. By default, each type of conversion can be triggered only once per visit, but this setting can be changed. Piwik users can define and track an unlimited number of goals using the main Piwik menu. Goals are essentially used as metrics across different reports to see what values in those reports led visitors to take required actions or complete conversions (see **Figure K.3**). For example, Piwik can show that certain keywords had much higher conversion rates for newsletter signups than others.

Goal reporting in Piwik is straightforward and, again very similar to goal reporting in Google Analytics.

*Figure K.3.* Sample of a Piwik goal-overview report.



**717** conversions

**116.50 €** overall revenue

**22.9%** overall conversion rate (visits with a completed goal)

Goal 'Download Piwik'

**233** conversions

**8.1%** conversion rate

Goal 'Visit Demo'

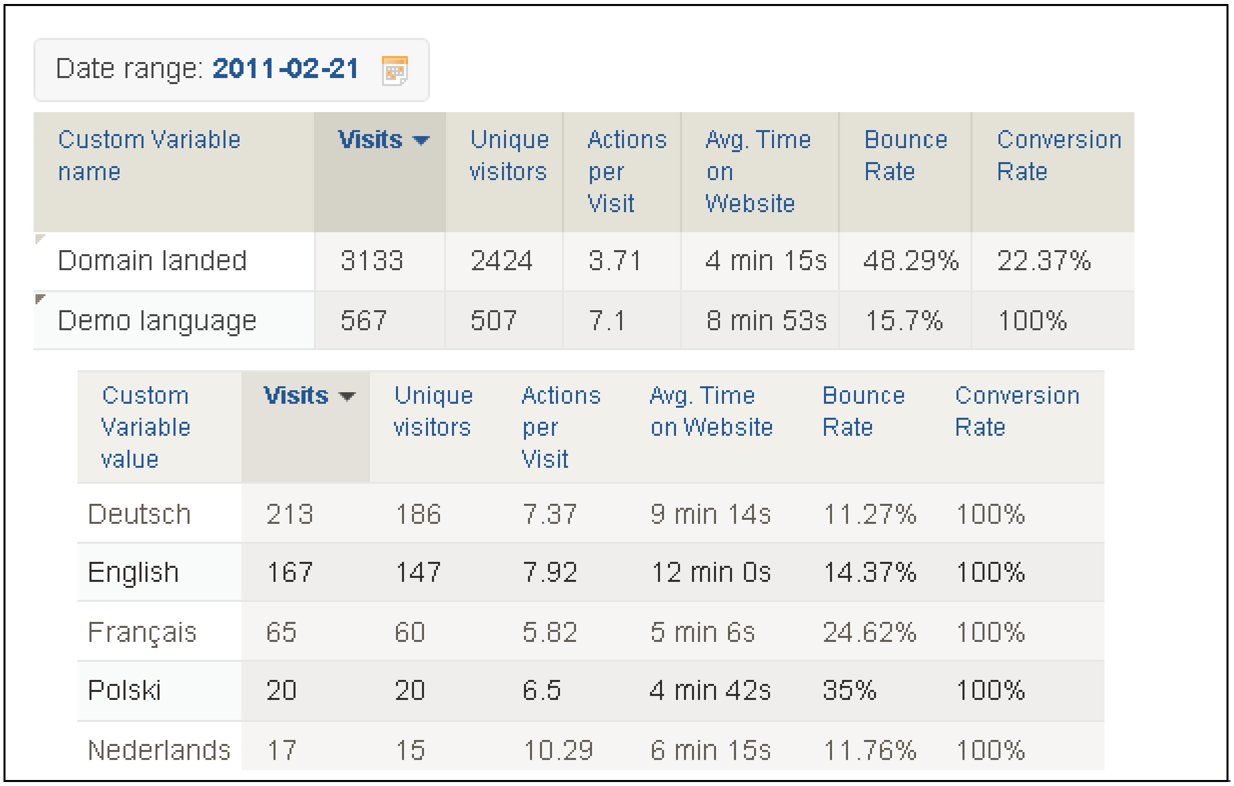
**484** conversions

**16.8%** conversion rate

### Custom Variables

Piwik allows up to five custom variables per visit and per page view (see sample, **Figure K.4**). In practice, this means that users can set up to 10 custom variables on a page (5 at the visit level and 5 at the page level). Piwik's option is very similar to Google Analytics' approach to custom variables. Customization has long been considered a strength of Omniture (which offers more than 100 custom variables), as compared with the number of custom variables available from a free solution like Google Analytics. In theory, a large number of available custom variables should allow better reporting flexibility and more complex, granular measurement. In practice, however, depending on the site, 10 custom variables are probably adequate if they are used carefully.

***Figure K.4.*** *A sample Piwik custom-variable report.*



Date range: **2011-02-21**

| Custom Variable name | Visits | Unique visitors | Actions per Visit | Avg. Time on Website | Bounce Rate | Conversion Rate |
|---|---|---|---|---|---|---|
| Domain landed | 3133 | 2424 | 3.71 | 4 min 15s | 48.29% | 22.37% |
| Demo language | 567 | 507 | 7.1 | 8 min 53s | 15.7% | 100% |

| Custom Variable value | Visits | Unique visitors | Actions per Visit | Avg. Time on Website | Bounce Rate | Conversion Rate |
|---|---|---|---|---|---|---|
| Deutsch | 213 | 186 | 7.37 | 9 min 14s | 11.27% | 100% |
| English | 167 | 147 | 7.92 | 12 min 0s | 14.37% | 100% |
| Français | 65 | 60 | 5.82 | 5 min 6s | 24.62% | 100% |
| Polski | 20 | 20 | 6.5 | 4 min 42s | 35% | 100% |
| Nederlands | 17 | 15 | 10.29 | 6 min 15s | 11.76% | 100% |

### Campaign Tracking

Campaign tracking, a fundamental part of any site's Web analytics initiative, is used to measure how well the visitors from external marketing or outreach campaigns are performing once they get to a site. Like most other major analytics vendors, Piwik is set up to pull campaign parameters out of the campaign URLs and report on them. Piwik defaults to using the "pk_campaign" or the "pk_kwd" parameter, and it has a link-builder tool to help generate appropriately named campaigns for your URLs (http://piwik.org/docs/tracking-campaigns/url-builder/). It also supports Google Analytics campaign parameters.

### Ecommerce

In its most current release (1.5), Piwik launched a variety of ecommerce tracking capabilities. It can capture the most critical ecommerce information and pass it into metrics reports. These metrics are then available separately for the various reports within Piwik. Tracking is also available for shopping carts, product page views, and category page views. In addition to reporting on metrics such as orders, revenue, and conversion rate, Piwik reports include an overview of shopping activities; best product stock-keeping units (SKU), names, and categories; and an ecommerce log report that tracks visitors' click paths. All these reports and metrics are available in the Piwik reporting interface.

## Support

Because Piwik is an open-source, free, community-developed tool, the level of support offered for it is not the same as for a paid solution or a centrally developed, but free, tool such as Google Analytics. The Piwik Web site offers fairly comprehensive documentation, but it is highly technical.

The Piwik Web site does host forums where users can go to find answers to their questions. Obviously, these forums are user-driven and are only as professional and authoritative as the individuals participating in them.

There are "Piwik consultants," third-party developers who become certified by either patching the existing code or developing new plug-ins to sit on top of Piwik’s code. Certification involves working with Piwik developers and getting code patches or plug-ins approved. Piwik has a development road map for expanding its current functionality and relies on its users to address the road map’s next steps.

## Tool-Use Case

Piwik is best suited to small-to-medium-sized Web sites with standard reporting needs and supported by at least one strong technical person. Because the data would be housed on NIH servers and the complexity of Piwik’s data processing protocols, installing it on a large, traffic-heavy site probably would not work well. As Web sites become larger and more complex, reporting needs may outgrow the amount of customization available from the Piwik tool.

Because the data would be housed on NIH servers and Piwik’s data processing protocols are so complex, installing it on a large, traffic-heavy site probably would not work well. As Web sites become larger and more complex, reporting needs may outgrow the amount of customization available from the Piwik tool.

Since Piwik requires far more technical expertise than typical SaaS Web analytics tools—including knowledge of JavaScript, PHP, and SQL—it is imperative that any organization planning to use Piwik have at least one strong technical person who can devote a significant amount of time to maintaining, upgrading, and customizing the tool.

From a reporting standpoint, Piwik is best suited for sites that have modest reporting needs, but with sufficient technical support, it can be scaled for more advanced applications. Piwik easily produces reports on basic, standard metrics. These data can be exported in spreadsheet formats and accessed via Piwik widgets. The widgets can be embedded into pages or, in more advanced situations, using the Piwik application program interface (API). As with the tool in general, technical support is required to use the widgets or API-based reporting.

## Tool Pros and Cons

### *Pros*

- Free to use, with all costs coming as technical resources and staff hours.
- NIH can own and control the data.
- Nice interface for basic reporting. Out-of-the-box tracking covers major important data points.
- Availability of custom variables with multiple expiration settings allows for a thorough understanding of how successfully users convert.
- Goal tracking is easy to set up and allows a couple of different methods for triggering the desired actions or goal conversions.
- It is possible to get in-depth, custom reporting if you have a database administrator who is proficient in PHP and SQL.
- Piwik's Web site features free documentation and user forums for troubleshooting help within the larger Piwik user community.

### *Cons*

- Since Piwik is *not* a tool built on a SaaS model, the upkeep and maintenance burden would fall on the IC. The IC must troubleshoot issues, develop new uses, and keep the tool up-to-date by installing new releases.
- Customization in Piwik requires knowledge of PHP and server-side technologies.
- Piwik is scalable, but only to a point. Because the reports are run in real time, larger sites may experience server delays and run reports slowly. Piwik recommends staggering schedules for running reports. However, it offers some configuration options that can help mitigate problems with high data volumes.
- The number of custom variables available is limited.
- Advanced reporting on segmentation, multiple variables, and third-party integration can be difficult.

## Recommendations

Piwik is a tool that can provide the types of data and configuration options that should be expected from an enterprise-wide Web analytics solution. However, acquiring more-complex data and customization from Piwik would take a lot of work by NIH employees.  The increased work required by organizations using Piwik is the tradeoff for free access to this open-source solution that is fully contained on the organization’s own servers.  For these reasons, Semphonic recommends that ICs seriously consider Piwik when they meet most or all of the following conditions:

- They place a premium on the ability to own their own data.

- They have *at least* one dedicated person with technical proficiency in JavaScript, SQL, and PHP (or other comparable server-side technologies).
- They plan to measure small to medium traffic volumes and page counts.
- The number of conversion or success events ("goals" in the parlance of Piwik), with straightforward attribution models, is limited.
- Their reporting needs are relatively simple and straightforward (that is, the needs would be considered “out of the box” by other major analytics tools).

# Appendix L.
# Glossary for NIH Web Analytics Best Practices

Semphonic
16 Digital Drive Suite 220
Novato, CA 94949
www.semphonic.com
December 15 , 2011

# Appendix L. Glossary for NIH Web Analytics Best Practices

**A/B testing** – The process of exposing randomly selected site visitors to two different versions of a page or section in order to determine which is more effective.

**analytics tool** – A software package that allows data to be collected and analyzed. Typically, analytics tools are classified based on the type of data they are designed to work with, such as Web, video, mobile, search engine, and social media data.

**application program interface (API)** – A set of routines, protocols, instructions and tools for accessing data collected by an analytics solution for export into a different system. By contrast, analytics tools contain a graphical user interface that provides for a display of the data in a report format.

**asynchronous –** A type of Web analytics code that allows the analytics software to run while the page loads without interrupting the user experience.

**campaign codes** – Currently, the most common mechanism for tracking marketing, outreach, and internal campaigns. When a campaign is created and links to a designated site, campaign codes are placed into the link URLs. This allows analytics tools to record traffic and activity originating from the campaign.

**Content Distribution Index** – A success metric that indicates what percentage of visitors share particular content using content-sharing tools, like AddThis. It is calculated by dividing the total number of content-share events by the total number of page views.

**content relevance** – A success metric that conveys the number of site pages that are not viewed compared with the number viewed. It is calculated by dividing the number of pages that received a set number of page views in a given time period by the total number of pages on the site. Content relevance results can highlight the need to revise, retire, or relocate pages that are not viewed as often as others are.

**content management system (CMS)** – Used to enter, serve, and administer content on a Web site. Often, the CMS allows users to define content characteristics in ways that are convenient for generating analytics data and reports.

**conversion** – Describes a completed success event or desired action on a Web site or online application. Since success events are specific to individual sites, site owners define conversion events. Common conversion events include newsletter signups, content downloads, and membership registrations.

**cookie** – A piece of non-executable text that a Web server stores on a computer user's hard disk. Using cookies, a Web site can store information on a visitor's computer for a specified time and later retrieve it. Cookies contain pieces of information that differentiate a visitor's computer and browser during the next visit to the Web site. Placing cookie text allows Web sites to "remember" visitors' preferences, surfing patterns, and behavior. Cookies can be "set" for different time spans. "Session" cookies expire when a visitor leaves the Web site, and "persistent" cookies remain on the visitor's computer until their expiration dates or until the visitor deletes them. Both cookie types track pages visited during an individual visit. Only persistent cookies can track visitors who visit the site and return later.

**cost per million (CPMM)** – A typical pricing schedule for Web analytics software: clients are charged per million server calls.

**cross–selling, or content cross sell** – Refers to the ability of a piece of Web content to lead users to additional content on the site. This term is derived from the ability of content to “drive” users across different content categories.

**cross-selling platform** – A tool that highlights additional content and helps visitors get to other parts of the site.

**dashboard** – A collection of Web analytics reports compiled for distribution to Web site managers, staff, and stakeholders. Typical formats include Excel, PowerPoint, HTML, and PDF.

**deep-dive analysis** – A focused effort to answer a *specific* research question using analytics data.

**engagement** – A measure of the extent to which a user is actively participating in a Web site experience.

**first-party cookie** – A browser cookie issued to a visitor's computer by the site the person is visiting. In other words, if the visitor is browsing on myorg.gov, first-party cookies are cookies issued by the domain myorg.gov.

**fixed Web** – Designates a site designed for browsing by personal computers. Alternatively, "mobile Web" designates a site designed for access by mobile devices, such as smart phones or tablet computers.

**functional analysis** – A type of analysis built on the premise that different pages on a site serve different purposes and should be evaluated differently. In this type of analysis, analysts assign function-related titles to individual pages and then analyze the site using the page functions as a basis for the analysis.

**funnel analysis** – A type of analysis that examines the effectiveness of multistep processes on Web sites. Typical funnels include newsletter signups and registration processes. A funnel analysis should go beyond a simple examination of the number of visitors who leave the process at each step. It can help Web teams identify problems with the funnel processes or recognize which audience groups have the most success with them.

**ICs** – Institutes and centers (at NIH).

**integration** – Refers to the ways that various data-collection systems, analytics tools, and decision-making processes are combined to enhance their overall effectiveness. It includes both technical integration – that is, finding ways to combine and centralize data provided by different tools – and codifying decision-making processes that incorporate the data.

**internal search** – Refers to the process used by visitors to find content on a Web site using the onsite search function.  The onsite search function is usual displayed as text box and labeled “search.”

**internal search effectiveness** – A type of analysis that studies how effective the site's internal search tool is from a variety of perspectives, including how search is incorporated into the site design, which activities lead users to search, and how well the internal search generates relevant results and leads users to content that further engages them.

**key performance indicators (KPIs)** – Key performance indicators are metrics calculated from common data points that when put in context with goals and objectives can help indicate the health and effectiveness of a site.

**latency –** The amount of time between when data are collected by a Web analytics solution and when reports from those data are viewable in a Web analytics solution.

**metrics** – Systems of measurement that assess the performance of Web sites and other online initiatives.

**mission-critical content effectiveness** – A success metric that evaluates visitor interest in content that an organization (for example, NIH or an IC) considers important. Mission-critical content effectiveness is calculated by dividing mission-critical page-view events by total page views.

**mission-success score** – A success metric that evaluates visitor engagement based on the visitor’s reading of mission-critical content and completing mission-critical tasks. The mission-success score is calculated by dividing the number of mission-critical page views and task-completion events by the total visits.

**online measurement –** See "Web measurement."

**OMB** – Office of Management and Budget.

**page-value analysis** – A type of analysis that involves developing  a scoring system for pages based on their functions and the site’s goals and objectives. The analysis involves evaluating the pages and given them numerical scores.

**pathing** – A study of how users move through the site, especially the order in which they access various pages during their visit.

**persistent cookie** – A piece of nonexecutable text placed on a visitor’s computer to capture and store descriptive data when the visitor is on the site. Persistent cookies are placed on visitors’ computers by the Web sites they visit and remain on the user's computer for a set amount of time and during multiple site visits. Most government Web sites do not capture or store personally identifiable information (PII) using cookies.

**Piwik** – A free, open-source Web analytics tool hosted on an organization's own servers, so the data analyzed by Piwik can be owned by that organization and not by the company offering analytics services.

**return on investment** (ROI) – A measure of the amount of value or benefit derived from costs or labor associated with the deployment of digital assets, such as a Web site, a Web site application, new content on a site, or a software solution. It is calculated by multiplying a monetary value given to each type of task completion and subtracting the costs of creating the content or function that allows site visitors to complete tasks. ROI may also be based on costs avoided when visitors complete site-based tasks.

**RSS** – Often said to stand for "really simple syndication," a data format that allows for easy publishing and updating of content over time. In particular, it is used for distributing types of content that are regularly updated, such as blogs or news services. Using applications called "RSS readers," users can receive fresh, updated content that has been "pushed" to them using RSS.

**segmentation** – The process of dividing a site's audience into distinct groups based on specific parameters; the most common parameters are "visit type" and "visitor type." Segmentation is used to obtain a deeper understanding of a site's audience and to support targeted analyses.

**service-level agreement (SLA) –** A contract that determines the terms of the service a vendor is providing.

**session cookie** – A piece of non-executable text placed on a computer to capture and store descriptive data during the time a visitor is on the site. It is placed on the visitor's computers by the Web site being visited. When the visitor leaves the site, the session cookie and data are deleted. Most government Web sites do not capture or store personally identifiable information (PII) using cookies.

**software as a service (Saas)** – A model for distributing software in which the applications are hosted by the service provider and made available to customers via the Internet or another network.

**solutions (analytics)** – A generic term used to refer to software tools that allow for collection and analysis of particular types of data (see **Appendix D**).

**success metrics** – Measures or indicators designed to assess a site's success based on specific goals, objectives, and benchmarks.

**system-development kit (SDK) –** A set of tools that allows developers to build applications for a software platform. SDKs are similar to APIs.

**task-completion effectiveness** – A success metric that indicates visitors' interest in, and successful completion of, tasks important to the site owner (such as NIH or an IC). It is a percentage based on the total number of task completion page views divided by total number of site page views

**third-party cookie** – A cookie issued to a user's computer by a domain that is not on the Web site they are browsing. In other words, if the user is browsing myorg.gov, a third-party cookie is a cookie *not* issued by myorg.gov.

**triangulation** – The process of taking multiple sources of Web site evaluation data, such as Web analytics reports, user-experience testing, and voice-of-customer surveys, and using them as complementary methods for building a well-rounded foundation from which to make decisions about Web site design.

**use case** – A profile or projection of how a user could interact with the site. Use cases are useful for defining audience segments and planning the layout and navigation for a site.

**user authentication** – The process through which a user's identity is verified. It must be carried out before a user can access secure services or content on a site.

**visitor-segment analysis** – The process of examining the characteristics or actions of distinct *visitor* groups. Visitor analyses may focus on how visitor groups interact with the site, use content, and conversion-event rates.

**Web analytics –** The practice of collecting, reporting on, and analyzing data that describe how a Web site is used.

**Web, or digital, analytics data –** Information about how an organization's Web site and other computerized resources are used.

**Web measurement –** The practice of recording data that describe the online presence of an organization.

**Web site optimization** – The process of revising a site to improve its performance and/or best meet the needs of visitors. Optimizing a site involves examining the site as a whole and in component parts and making modifications to improve site functions and content.