Best Practices for Metrics Collection and Analysis

August 2, 2010

Report of the NSDL Metrics Working Group

A report of the
National STEM Distributed Learning (NSDL)
Resource Center
University Corporation for Atmospheric Research (UCAR)
Boulder, Colorado

This report is a publication of the National STEM Distributed Learning program's National Science Digital Library (NSDL) Resource Center, hosted at the University Corporation for Atmospheric Research (UCAR).

Mailing address:
National Science Digital Library
UCAR
3300 Mitchell Lane
Boulder, Colorado 80301
Contact: 303-497-2943
nsdlsupport@nsdl.ucar.edu

Editor: Eileen McIlvain, NSDL Resource Center

Citation:
Lightle, K., Almasy, E., Barbato, L., Clark, S., George, Y., Hsi, S., Lowe, C., MacKinney, P., McIlvain, E. (2010). Best Practices for Metrics Collection and Analysis: Report of the NSDL Metrics Working Group. Boulder, Colorado: National Science Digital Library, University Corporation for Atmospheric Research, 39 p.

Acknowledgement:
This material is based upon work supported by the National Science Foundation under Grants No. 0840858 and 0840744. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Table of Contents

1 Overview

1.1 Introduction

This report provides recommendations and best practices suggestions for the use of metrics in digital library project evaluation, formulated by the National Science Digital Library (NSDL) Metrics Working Group.

In this document, we provide an overview of which metrics are useful for assessing digital library activities, make recommendations on how to collect those metrics, and provide examples of how to use collected metrics in larger evaluation efforts.

The document is split into five sections:

1. Section 1 provides an overview of the document's audience and its creation
2. Section 2 provides an introduction to metrics for new NSDL projects and why they are important
3. Section 3 provides recommendations for determining which metrics are useful for evaluating particular activities, describes activities and data and the relationships between them, and provides a list of possible project impacts
4. Section 4 provides a set of example vignettes to illustrate evaluation efforts in a number of NSDL projects and other contexts
5. Section 5 provides information on common metrics packages, a glossary, and references that are of use to NSDL project evaluators and technical staff as they assess common library services

The projects targeted by this document include those funded under the National Science Foundation's (NSF, 2009a) National Science, Technology, Engineering, and Mathematics (STEM) Education Distributed Learning program (NSF, 2009b). The program includes projects funded under the Pathways I and II tracks, Targeted Research, Services, and small grants, and any future tracks created as part of the NSDL funding program.

Because NSDL partnerships and projects are diverse in nature, scope, and target audience, parts of this document may not be applicable to every project. However, we have endeavored to provide recommendations that are of use to the full range of digital library projects and services. Digital library projects beyond the NSDL program are also encouraged to make use of these guidelines, as suits their needs.

1.2 Background

The need for a set of best practice recommendations emerged from discussions at multiple levels within the NSDL community:

- 2008 NSDL Pathways meeting and Pathways monthly teleconferences[1]
- 2008 NSDL Annual Meeting of project Principal Investigators (PIs)
- NSDL Resource Center and NSDL Technical Network Services collaboration for administration and project management of the NSDL
- Emphasis within the National Science Foundation's NSDL program solicitation, (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5487) on the need for effective evaluation practices that demonstrate or illustrate impact on learning in STEM education

In an effort to help improve the evaluation efforts of the NSDL community, the NSDL Metrics Working Group was formed in January 2009 and charged with developing a set of best practices recommendations to the NSDL community for metrics collection and reporting. Composed of volunteers from the NSDL Pathways projects and representatives of the NSDL Resource Center and Technical Network Services, this document emerged from discussion within the Metrics Working Group between February 2009 and April 2010.

These recommendations apply to metrics collection across all NSDL project activities: from technical development to portal use, from outreach and professional development activities to overall evaluation plans and activities. By outlining key areas where building in metrics-gathering techniques ahead of time can enhance project evaluation, we hope to help projects improve their assessment efforts and enhance reporting across NSDL projects.

1.3 Recommendations Summary

Given the wide range of NSDL project types and their unique characteristics, the application and recommendation of a minimal set of "common metrics" that can be reported to NSDL on a regular basis has proven to be quite challenging. Quantitative data are susceptible to multiple interpretations, and variability and divergence among projects almost immediately becomes evident. In addition, projects may employ different web metrics collection and analysis tools, from Google Analytics, to Webalizer, to Omniture Site Catalyst, or others. Web metrics packages can be opaque about how their measures are calculated and/or rendered—another complicating factor to aggregation of data across multiple projects.

Utilizing automated means for "roll-up" reporting of multiple sites via a metrics collection service is the most feasible, efficient, and least human resource-intensive method for gaining a more comprehensive picture of NSDL usage and reach. NSDL uses Google Analytics (GA) for recording metrics on NSDL web sites, as do a number of other NSDL projects.

---

[1] NSDL *Pathways* projects are special projects funded under one track of NSF's National STEM Education Distributed Learning program. Pathways provide audience-specific views of digital resources and services for unique communities. They may be discipline-focused (e.g. ChemEd DL), audience-focused (e.g. Middle School Portal Math and Science Pathways (MSP^2), or specialize in the delivery of particular resource types (e.g. WGBH's Teachers' Domain, providing multimedia resources for STEM education). Other funding tracks of NSDL include Services projects, Targeted Research projects, and small grants supporting the continuation of work funded by NSDL.

Roll-up reporting

The Metrics Working Group recommends that NSDL projects install an NSDL code snippet into their Google Analytics accounts that will automatically report to the NSDL roll-up account for their site's activity.

Advantages of this method include:

1. Enabling of multiple views on a variety of metrics at any given point ("slicing and dicing").
2. Enabling NSDL to promote specific events and report on user activity across multiple sites for those unique events, such as multi-partner workshop presentations at national conferences, like the National Science Teachers Association, or American Association for the Advancement of Science (AAAS).
3. Enabling more effective trends monitoring and reporting for NSDL collectively, without highly intensive labor required aggregating multiple sources of information by any one individual.

The Metrics Working Group encourages NSDL projects to send data to the NSDL Google Analytics account for the purposes of roll-up reporting for NSDL, even if projects currently use a different metrics collection and analysis package as their preferred, in-house metrics service. Projects may continue to use other favored metrics services for their own needs, analysis, and reporting for their own project or program goals. Projects do not need to create a separate Google Analytics account to send data to the NSDL Rollup account.

A reporting area for NSDL Google Analytics roll-up data will be available on the NSDL Community site, Evaluation and Metrics resource area (http://nsdlnetwork.org/evaluation), and will allow community contribution and discussion of these important issues for project management and sustainability.

NSDL Roll-Up Reporting

NSDL-funded projects may request the NSDL Google Analytics code by sending an email to: nsdlsupport@nsdl.ucar.edu, with subject line:

NSDL CODE SNIPPET FOR GA

Please identify project name, NSF award number, and principle investigator (PI) in your request.

Individual project reporting of metrics

Individual NSDL projects will find it valuable to include metrics information on their project activities in their annual and final reports to the National Science Foundation, and/or other funders.

While individual projects rightly must decide for themselves what the most important metrics for their unique projects are, the Metrics Working Group suggests the following range of metrics might be helpful in identifying trends of usage for web sites. These are not intended to be comprehensive or exhaustive; they are a small subset of possible metrics that may prove helpful when analyzing from one time frame to another, or over a particular time range (e.g. quarterly, six months, one year):

- Percentage of new visitors
- Average time on site
- Visitor locations (geographic: country, state, zip code)
- Traffic sources (direct, referring, search engines). Analysis of types of domains: K12 schools; colleges/universities; museums/science centers, etc.
- Search engine terms

- Non-library downloads (i.e. downloads exclusive from collection search functions, e.g. for outreach materials)

Annual Survey

In addition to reporting on common web metrics that can reveal quantitative measures of NSDL sites usage, the Metrics Working Group recommends that the NSDL Resource Center conduct an annual survey of all NSDL projects, soliciting unique success stories in the overall categories of activity described in the Activities Matrices (found in Section 3.2), and report these findings to the NSDL community and to the National Science Foundation.

Combined with quantitative measures, such vignettes help to reveal a more comprehensive picture of NSDL's collective reach, value, and impact on teaching and learning across STEM disciplines.

Categories include:

- Web portal or web site / General technical development
- Community building: Online applications and social media
- Outreach and marketing activities
- Professional development and training activities

Feedback

This report is offered as a living document—one intended to grow and develop with commentary and response from the digital library community.

To provide comments, examples, or additional best practices, readers may:

1. Contribute to the NSDL Community Network *Got Metrics? Forum* (http://nsdlnetwork.org/forums/general-discussions/got-metrics). (Requires NSDL Community Network user account creation)

2. Join the *Metrics Interest Group* (https://www.nsdlnetwork.org/content/group/633/metrics-interest-group.

3. Send questions, comments, suggestions, or submit additional vignettes to nsdlsupport@nsdl.ucar.org, with a subject line such as "Metrics report comment/contribution, etc."

1.4 The NSDL Metrics Working Group Members

Kimberly Lightle, Chair; Director of Digital Libraries, College of Education and Human Ecology, School of Teaching and Learning, The Ohio State University
- PI: Middle School Portal Math and Science Pathways (MSP2) (NSDL)
- PI: Beyond Penguins and Polar Bears: Integrating Science and Literacy in K-5 Classrooms (DR K-12)

Ed Almasy, Internet Scout projects, University of Wisconsin, Madison
- Co-PI: Applied Math and Science Education Repository (AMSER) Pathway
- Co-PI: ATE Resource Center (ATE)

Lyle Barbato – American Association of Physics Teachers (AAPT)
- Technical Lead, ComPADRE Physics and Astronomy Pathway (NSDL)

Yolanda George, American Association for the Advancement of Science (AAAS)
- PI: Biological Sciences EdNet (BEN) Pathway

Sherry Hsi, Lawrence Hall of Science at the University of California, Berkeley
- Co-PI, Science and Math Informal Learning Educators (SMILE) Pathway

Cathy Lowe, Kent State University
- Project Manager, Materials Digital Library Pathway (MatDL)

Paul Mackinney, University of California, Berkeley
- Web Developer, Engineering Pathway

Sharon Clark – Digital Learning Sciences, University Corporation for Atmospheric Research
- Web Developer, NSDL Technical Network Services

Eileen McIlvain – NSDL Resource Center, University Corporation for Atmospheric Research
- Communications Manage, Pathways Liaison, NSDL Resource Center

2 What are metrics?

2.1 Definition of Metrics

Metrics are defined processes for quantitatively measuring sets of criteria. In evaluation efforts, metrics are often used to help assess functions, tools, or services. Examples of common metrics might include the number of attendees at a workshop or the average number of visitors to a website in a week. In ambitious cases, metrics may be defined to help measure user satisfaction or educational impact. In such cases, a variety of factors may be taken into account, but a number or set of numbers will be yielded. Because metrics generate numbers, they are represented visually more readily than qualitative information.

2.2 Digital Libraries and Web Metrics

Web metrics are a category of metrics of particular interest to digital libraries and related projects. The compiled usage data and transactions of visitors to a web site, web metrics are often collected automatically by a website's server or from scripts embedded on a site's web pages. Common values tracked include the number of unique visitors to a site, the number of pages viewed by visitors, and a user's path through a site (see matrices in Section 3.2). Such data may be used to determine how a site is being used and the audience a site is reaching. This information may then be used to help improve digital library functions, tools, and services.

Standard web metrics are difficult to define for digital libraries due to the variety of project audiences and goals (Manduca, Iverson, Fox, & McMartin, 2005). For example, a project building a collection for a particular audience (e.g., field biologists teaching undergraduates) will have a different set of benchmarks than a project with a different goal and audience (e.g., professional development for afterschool science leaders). Such variety has resulted in few direct comparisons of digital libraries in evaluation efforts. Instead, internal trends within single projects are often emphasized.

However, even internal trends in a single digital library may be difficult to interpret. This is because projects rarely have expectations for what web metrics values should be. Because of this lack of expected values, digital libraries often report growth trends over counts. However, for services targeting captive audiences or with market saturation, growth trends are of little use. Because of such limitations, web metrics should be used in conjunction with other methods in digital library evaluation.

2.3 Metrics and Evaluation

Although metrics provide useful data, reporting their values is not enough for assessing a program's broader impact. Because projects often impact entire communities, a repertoire of evaluation strategies and methods must be employed beyond those that focus on individuals. For a rich program evaluation, data from metrics must be triangulated with interviews, focus groups, ethnography, and other approaches.

2.4 Why Should You Care About Metrics?

Identifying and collecting appropriate metrics is a key component in program or product development and sustainability, including planning, implementation, and evaluation, both formative, and summative (Evaluation Resource Institute, 2009; Reeves, Apedoe, & Woo, 2003). During program or product planning, it is important to define performance metrics that can verify that overall goals and objectives are being achieved and that a product or service is effective and useful. Internally, metrics are a part of the business plan, and externally, this information is often a component of proposals for the government, foundation, and business sectors.

Establishing a metrics collection plan at the beginning of a project or activity is a key best practice, and is a cyclical process, as metrics should contribute to both short-term decision-making as well as long-term strategic planning. During implementation, diagnostic metrics can be continually used to make decisions about the development of a product or service. Appropriate diagnostic metrics provide needed feedback to improve or revise services, including "what, whom, when" and under what conditions.

Metrics help project managers know and understand their users, clarifying dimensions of engagement and interest. They can serve as a springboard for deeper inquiries, and when combined with other kinds of reporting, such as user profiles, surveys, anecdotal vignettes, and full-scale evaluation efforts, metrics can help projects enhance their utility to end-users. Metrics can help to reveal systemic barriers and areas worthy of reform, and help funders fully understand the dimensions of the many challenges for educators, learners, and service providers in STEM education.

Demonstrating impact through innovations in digital resource utilization, educational practice, or in technological adaptation is a desired outcome, but often difficult to accomplish. Impact studies take time, are complex, and agreement—even among evaluation experts—is often elusive (Khoo, Buchanan, & Cunningham, 2009). Building in measures in project activities and evaluation plans that reflect the "dailyness" of educational practice, of usage that reveals patterns and trends, can help to uncover educational impact in an incremental fashion, while confirming the value of the effort or activity, and the strategic choices and decisions made. These patterns, in turn, contribute to new development and new collaborations that advance STEM education.

In general, an overall well-designed, systematic evaluation should include clearly defined metrics that:

- Help to determine and refine products and services
- Track number of users by demographics
- Provide evidence that users are satisfied with the quality of a product or service
- Provide evidence of impact on user knowledge, attitudes, or behavior
- Compare products and services to others (benchmarking)

Regularly reporting such clearly presented metrics is an essential element of communications to stakeholders.

In terms of sustainability, appropriate metrics are critical in competing for and maintaining funding via grants, investments or revenue generation. Metrics can provide the verification needed for internal executive officers and boards, users, policy makers, funders, individual contributors, and investors.

3 Recommendations

3.1 Starting with Key questions

Before collecting any data, generating a well-thought-out question is an important first step in designing a good evaluation that uses metrics and data as evidence of impact. Here are some suggestions for generating driving questions:

Pose a question:

- About a particular group, audience, or population
- That makes a meaningful comparison
- That examines (changes to) user knowledge, skills, attitudes, participation, or behavior
- About how a collection, resource, tool, or service is working, changed, or improved
- That asks if your project goals, program objectives, and targets are being reached
- That tests a prediction about an expected result or outcome

Then, examine your questions to see if they are investigable with metrics data or require other data sources.

To get started, we provide a basic set of initial questions. While these questions are mostly about collections, these can be adapted to ask questions about individual resources, tools, and services.

When designing a question to ask, it is useful to consider the *report user*—the recipient of the final report. Reflect on who the report users are, taking note of what they will value and consider valid evidence, and how best to structure your report to communicate this information.

Key Questions

- Does <*collection, resource, tool, service*> have name recognition within the community?
    - If the *collection* is recognized:
        - Does *collection* have recognition as a STEM teaching and learning resource?
        - Do users report that they value *collection* as a web resource?
        - Do users have positive attitudes about using *collection* to locate high quality content?
    - If not:
        - Are users unaware of *collection* but in need of its resources able to find it?
- Who is using *collection* resources? (e.g., students, faculty: K-12 educators, higher education faculty and instructors, etc.)
    - Note: Privacy issues may exist for students
- What are the most popular types of *collection* resources as reported by users? (images, simulations, lesson plans, activities, etc.)
    - Are these the resources the *collection* sees in their logs as being the most used?
- For what purpose(s) are *collection* users using the resources that they find through *collection provider*? (e.g., teaching, professional development, class work, personal interest…)
    - Are *collection* users finding useful resources for these information needs?
- Are users incorporating *collection* tools/services into their own workflow?
    - Are the tools and services that the *Collection* provides those most desired by users for incorporation into their own workflow?
- Are users requesting *collection* content?
    - How many users request additional information?
    - How many users subscribe to *collection* services? (RSS/Atom, Facebook, Twitter, Federated search, OAI, etc.)
- How many new users find *collection* and how often?
    - How many users "join" the community?
    - Are users returning?
        - For the same material repeatedly or different materials?
- Do individual users participate or contribute content to *collection* communities?
    - How many login or regularly participate in the community?
- Are users reporting that they are doing STEM programs in their setting because of *collection*?
    - Do users report that *collection* supports their existing programs?
- Is *collection* having a national reach?
    - Is *collection* having a global reach?
- How satisfied are *collection* users with their experience? (e.g., search process, resources, etc.)
    - Do users have positive attitudes about using digital libraries to locate high quality content?
    - Do users report changed attitudes about using web technology and digital resources for learning, teaching, and professional development, following use of *collection*?

3.2 Identifying Activities, Data, and Impacts

A variety of activities occur around online content and services. Users might log in to a site, search a collection, and download a document, video, or audio podcast. Users can update profile information and respond to a blog post. They can subscribe to an RSS feed or bookmark the site in Diigo. They can come back repeatedly to a certain page or site area.

Project staff may email users or tweet reminders about upcoming webinars and share information about the latest content. Project staff might facilitate face-to-face or online professional development experiences, collect artifacts and survey data from participants, or catalog resources and build online collections.

These activities are relatively easy to keep track of—much of the data can be collected via your web metrics package, but not all of it. Sometimes it is a matter of coding the responses to survey questions and counting the number of bodies that attend a face-to-face presentation at a conference. The harder questions to answer are if and how a video downloaded from your site was actually used in a classroom, what learning happened as a result of a user listening to a podcast, and what teaching strategies were actually used in a classroom, and what were the outcomes from that lesson.

To help classify the activities that project staff and users engage in for evaluative purposes, we have created a set of **Activities Matrices** that link stakeholder activities with metrics that may be collected to help measure them. Activities are distributed into four categories:

- Web portal or web site, or general infrastructure-based activities
- Community building/Web 2.0 social media activities
- Outreach and marketing
- Professional development

For each activity there are corresponding probes and metrics. For instance, for the activity *Online Seminars* the associated probes and metrics are 1) number of participants, 2) number of returning participants, 3) post-session surveys, 4) comments and emails, 5) personal conversations, and 6) number of downloads of the archived recording. Some of these metrics are quantitative, but some are qualitative. Activities in these categories are not exclusive; there may be overlap across categories. Think of the **Activities Matrices** as a brainstorming checklist that you can use to identify venues and corresponding metrics.

Activities that may involve or rely on *personally identifiable information* of users (e.g. email addresses) are highlighted in *italic text*. Care should always be exercised in the use and analysis of such data to ensure that personally identifiable information of users is protected, in accordance with your project's Terms of Use and Privacy Policies.

Matrices and tables appear in this order:

1. Activities Matrices
    a. Web Portal or web site / General technical development
    b. Community Building: Online Applications and Social Media
    c. Outreach and Marketing
    d. Professional Development
2. Supplemental Metrics Data Table
3. Impacts Table

3.3 Activities Matrices

Matrix 1 – Web Portal or Web Site / General Technical Development	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
Web metrics	• Number of registered users • Number of visits • Pages per visit • Unique visitors • Percentage of new visitors • Bounce rate • Average time on site • Visitor locations (geographic: country, state, zip code) • Most popular pages • Traffic sources (direct, referring, search engines). Analysis of types of domains: K12 schools; colleges/universities; museums/science centers • Search engine terms • Non-library downloads (i.e. downloads exclusive from collection search functions, such as for outreach materials)
Digital library collections – search and browse functions	• Resource access counts – the number of times a particular resource is accessed by users (in the case of a digital library that holds or offers direct access to actual content/resources themselves) • Resource record access counts – the number of times a metadata record about a resource is accessed by users (in the case of a digital library that does not hold content itself, but offers access to information about resources and provides links to those resources) • Number of searches performed over a set amount of time (days, weeks, months, etc) • Number of failed searches • Most frequent searches (search terms) • Most browsed classifications
Digital library collections and collection building	• Number of unique collections in a digital library • Number of items in distinct collections • Number of organizational partners (and who) on a project • Number of resources and/or collections accessioned • Number of resources and/or collections deaccessioned within review/curation processes

Matrix 1 – Web Portal or Web Site / General Technical Development	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
Number of resources/collections harvested via OAI/PMH (Open Archives Initiative/Protocol for Metadata Harvesting)	• Harvest count (overall) • Harvest counts (by metadata format: oai_dc, nsdl_dc, other) • Harvest frequency • Harvester IPs or host names
Calendar function	• Number of events • Most popular events (by click) • Most popular event types (by click) – (e.g. workshops, web seminars, conference, meeting, presentation, reception, etc.) • Click-through rate
Highlighted resource feature	• Number of highlighted resources • Number of access/downloads of highlighted resources
Foldering/Bookmarking systems	• Number of folders/bookmarks • Average entries per folder • Average folders per user
Notification systems to resource creators (feedback loop)	• Number of times a resource is accessed/downloaded • Domain analysis of accesses/downloads
Recommender service 1 User-recommended resources to add to a library	• Number of recommendations received per set time period (monthly, yearly)
Recommender service 2 Recommend–to-a-Friend service	• Number of recommendations sent to other users
Request Tracker / Support systems	• Number of inquiries submitted per set time period
Micro-survey; 'intercept' surveys (web-based mini-surveys of users)	• Number of responses to a limited set of questions about a particular page, site, or service • Report on survey results
Infrastructure provision to other projects	• Number and type of application or service provided • Number of customers • Training provided or conducted

Matrix 2 - Community Building: Online Applications and Social Media	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
Wikis	• Number of readers/users • Number of contributors • Number of pages • Most popular pages (by views) • Most active pages (by edits) • Number of discussions and participants
Weblogs	• Number of blogs • Average posts per blog • Average comments per post • Most active blogs (by posts) • Most active blogs (by readers) • Most active blogs (by comments)
Online Discussion Forums	• Number of posts • Most popular forums (by thread clicks) • Number of active participants • Length of discussion, over time (forums are often tied to singular events, such as a meeting or conference. Discussions generated from these events may continue over a length of time and then dissipate. Reporting on how long active discussions continue – their lifetime – may be informative) • Differential reporting for moderated vs. non-moderated forums
RSS feeds or other notification services	• Number of subscriptions • Polling frequency • Polling time distribution (most active polling time) • Click-through rate • Most popular items (by click)
Podcasts	• Most popular resources (by download) • Most popular resources (by page views) • Most popular resources (by number of comments) • Most popular resources (by number of recent comments) • Most popular subject areas (by download) • Most popular subject areas (by page views) • Number of comments per resource
Specific Social Media Services/Tools:	
Facebook	• Number of friends • Number of comments • Number of 'likes' (e.g. for special announcements,

Matrix 2 - Community Building: Online Applications and Social Media	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
	releases, collections, etc.)
Twitter	• Number of followers • Number of comments from followers • Number of retweets per post • Number of click-throughs on links • Number of conversations with followers or potential followers • Number of posts (per week or month)
MySpace	• Number of friends
Diigo	• Number of members
YouTube, TeacherTube, SchoolTube	• Number of downloads • Number of groups
NING	• Number of registered members • Number of groups and whether they were started by members • Number of members of individual groups • Number of blog posts and comments • Number of discussions and comments

Matrix 3 - Outreach and Marketing Activities	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
Materials distribution	• Number of items distributed or print materials used; annually; semi-annually • Detail which groups or kinds of audience received materials (e.g., teachers, general public, researchers, etc.)
Web seminars	• Number of participants • Number of returning participants • Results of post-session surveys, evaluations • Comments / emails from participants • Personal conversations
Publications	• Number articles produced per year • Number of paid subscribers to journals which carry the article; readership
Conference presentations	• Number per year • Number of attendees in session
Email newsletters	• Number of email subscriptions and/or subscribers
Podcasts	• Number of downloads (from iTunes or site tracking)
Website pages	• Number of visits • Number of subscriptions • Logfile click-through data to specific information pages
Web-based micro-surveys or "intercept" surveys (Short-term and limited set of questions to users about features/services on a webpage, a site, etc.)	• Number of responses • Comments received • Results of survey report
User surveys, either online or face-to-face	• Number of surveys distributed • Number of surveys returned • Results of survey report • Comments received

Matrix 4 -Professional Development & Training	
Activities in *italic text* may involve personally identifiable information of users. Exercise care in use of data to ensure user protection per your project's Terms of Use and Privacy Policies.	
Project Activity or Function	**Probes & Metrics**
Conferences, exhibit booths, information sessions	• Number and type of events per year • Number of attendees per session; demographic types of attendees • Number of contacts from exhibit booths (estimated if not specifically tracked) • Number of follow-up contacts made • Number of event-oriented support system requests • Analysis of session evaluations + delayed session evaluations • Testimonials, kudos – positive feedback received from participants; quotes
Face-to-face Workshops	• Number per year • Number of attendees in the session; demographic types of attendees • Post-session surveys, evaluations • Post-session longer-term follow-up (e.g. six month; 12 month)
Online seminars; netcourses	• Number of participants • Number of returning participants, enrollments • Post-session surveys, evaluations • Comments, emails • Personal conversations, communications
Training videos	• Number of downloads • Number of views (e.g. YouTube) • Number of clicks to recommend to a friend

Another View

The **Metrics Data Table** below is provided as a supplement—a possible alternate, more compact view or gathering point for information, with a narrower focus.

Matrix 5 – Metrics Data Table

Context	Data
Web Site	
General	• Number of registered users • Number of visits • Pages per visit • Unique visitors • Percentage of new visitors • Bounce rate • Average time on site • Visitor locations (geographic: country, state, zip) • Most popular pages • Traffic sources (direct, referring, search engines) • Search engine terms
Digital Library	• Resource access counts • Resource record access counts • Number of searches performed • Failed searches • Most frequent searches • Most browsed classifications
Calendar	• Number of events • Most popular events (by click) • Most popular event types (by click) • Click-through rate
Forums	• Number of posts • Most popular forums (by thread clicks) • Number of active participants
Foldering/Bookmarking	• Number of folders/bookmarks • Average entries per folder • Average folders per user
Wiki	• Number of readers/users • Number of contributors • Number of pages • Most popular pages (by views) • Most active pages (by edits)
Weblog	• Number of blogs • Average posts per blog • Average comments per post • Most active blogs (by posts) • Most active blogs (by readers) • Most active blogs (by comments)

Matrix 5 – Metrics Data Table	
Context	**Data**
RSS Feed	• Number of subscriptions • Polling frequency • Polling time distribution • Click-through rate • Most popular items (by click)
OAI-PMH Server	• Harvest count (overall) • Harvest counts (by metadata format) • Harvest frequency • Harvester IPs or host names
Web 2.0	
Facebook	• Number of friends
NING	• Number of registered members • Number of groups • Number of members in each group
Twitter	• Number of followers • Comments from followers • Number of retweets per post • Number of click-throughs on links • Number of conversations with followers • Number of posts (per week or month)
iTunes U	• Most popular resources (by download) • Most popular resources (by page views) • Most popular resources (by number of comments) • Most popular resources (by number of comments) • Most popular subject areas (by download) • Most popular subject areas (by page views) • Number of comments per resource
YouTube, TeacherTube, SchoolTube	• Number of downloads
Diigo	• Number of groups • Number of lists
Outreach	
Exhibit Booth	• Number of follow-up contacts • Number of promotional materials distributed
Workshop	• Number of attendees • Attendee evaluations
Webinar	• Number of participants • Percentage of returning participants • Participant evaluations • Number of downloads of recorded presentation

Impacts Table

The **Impacts Table** provides a list of potential impacts that activities and their measures might be associated with. These impacts are more qualitative than quantitative in nature, and it is by no means an exhaustive list. It simply suggests possible areas of impact that might be supported by data, or measures collected about the activities shown in the Activities Matrices.

Impacts
• Usage • Audience reach • Information seeking trends • Geographic impact indicators • Awareness/Interest building • Engagement • Capacity building among users via integration of practices into user workflow • Attitude alteration; knowledge building • User satisfaction; customer service • Reputation; brand monitoring • Contribution to marketing plan • Partnership or relationship building • Project decision-making • Service provision • Sustainability

3.4 The Fourth Dimension: Examining Trends Over Time

Measuring project and end-user activities helps with summative evaluation, providing hard numbers to help assess success in achieving desired outcomes, but metrics can be used for more than quantifying results after the fact. They can also prove to be invaluable tools for formative evaluation, helping to guide a project underway along the best route to achieving those results. This is where the fourth dimension—time—takes center stage.

Successive Approximation

Most project proposals outline a cohesive set of goals and a plan to achieve those goals, however, in any field involving unknown quantities and evolving conditions, project goals are likely to shift. In the case of NSDL, where the needs of the target population—STEM teachers and students—are not always well defined, and the operating environment—the Internet—is constantly and rapidly evolving, those goals become, effectively, a moving target.

As that target moves, the optimal path required to hit it varies as well. Staying on that path requires a continuous series of course corrections, reallocating resources and adjusting the focus of project efforts. Some of those adjustments are obvious, brought to light by external entities or events. However, most should ideally originate internally, a result of an internal feedback loop between formative evaluation and project development efforts. That feedback loop can be driven in part by metrics data collection and analysis.

Moving Pictures

An individual set of project metrics data measurements can be thought of as a *snapshot*, an instant caught in time. It can be examined to determine how closely project results match the intended goals. Metrics snapshots covering several different aspects of a project can be pasted together, providing a higher-level overview of project efficacy.

When these snapshots are taken periodically, in a series, they provide a dynamic overview, in effect, a moving picture, illustrating ongoing progress. When analysis of the snapshots is interspersed with adjustments to the activities being measured, a feedback loop results, with ongoing progress being continually guided along the optimal path toward the goal. The result is not only greater project impact—because available time and resources are being used more effectively—but also avoidance of project "drift", where activities become a focus and end unto themselves, rather than a means toward a goal.

Loops in Practice

It's easy to consider feedback loops for formative evaluation in the abstract, but it can be harder to implement them in the context of an ongoing project without a concrete approach toward implementation in mind. A prototypical circumstance for the application of a metrics feedback loop in a project with a significant web component would be new user registration.

Having users register is valuable for a number of reasons—it supports personalization, helps with user buy-in, and of course, greatly aids in demonstrating project impact. Unfortunately, most users are reluctant to provide *personally identifiable information* or to spend the time required to register, so making the registration process easy and painless is essential to building a significant base of registered users. Complicating this is the fact that new user registration is an area of functionality that is rarely used by those closely involved with a project.

For these reasons, making changes to streamline the new user registration process is often a focus for web-based projects, but which changes really make a positive difference? The most accurate way to answer this question is to set up and follow a metrics feedback loop based on new user registrations:

1. Make changes to the new user registration process
2. Measure the number of new user registrations over a period of X length and compare that to the number from the previous period (taking into account other factors such as promotional efforts)
3. If the number has gone up, then the direction of changes was effective and should continue to be followed
4. If the number has gone down, then the direction of changes was not effective and a new direction should be explored
5. Return to step 1

More specifically, beyond measuring the number of new user registrations, the point at which users who start but do not complete registration can be examined for indications of steps that might be particularly worthy of streamlining efforts. This type of specific examination can often reveal problems of which the project team was completely unaware because of their inevitable difficulty in considering the viewpoint of new users.

4 Putting It All Together: Vignettes

So what is the best way to deal with these different types of data and what could they possibly tell us? If you have multiple online seminars should you aggregate the data or look at each separately? How many people is the project or activity reaching? Or, is it better to look at the trends as described in *Section 3.4 Putting it All Together -Trends over Time*? Are there increasing numbers of participants in general? Another way to look at the data would be to determine which online seminars are best attended; perhaps certain topics are of greater interest to your users.

The answers to these questions are best determined by the staff working on individual projects, and by the circumstances and purposes for which the data are being collected. Synthesizing data collection into effective project reporting can be challenging, but rewarding, for project managers and evaluators. The best of all possible worlds is to be able to employ a data/evaluation specialist skilled in this kind of analysis and synthesis. Absent a skilled evaluator or specialist, this report is intended to provide some helpful suggestions for projects in developing their data collection and reporting skills.

The following vignettes of metrics collection and analysis are contributed from a variety of NSDL and NSDL-related projects and provide real examples that relate to the activities, data, and impacts described in Section 3. These examples of metrics implementation and analysis show a range of complexity and conclusions drawn to help in either formative or summative evaluation.

To suggest or contribute additional activities, measures, impacts, vignettes, and to join the discussion, go to the Got Metrics? forum or join the Metrics Working Group on the NSDL Community Network site: http://nsdlnetwork.org.

4.1 Vignette 1: Engineering Pathway

Contributor: Lora Oehlberg, University of California at Berkeley

Project Activity: Web Portal or Web Site

Impact: User Satisfaction, Engagement

Project Description

The Engineering Pathway is a portal to high-quality teaching and learning resources in engineering, applied science and math, computer science/information technology, and engineering technology and is designed for use by K-12 and university educators and students. The K-12 engineering curriculum uses engineering as a vehicle for the integration of hands-on science and mathematics through real-world designs and applications that inspire the creativity of youth.

A Needs and Usability assessment was performed in Fall 2006 by a group of independent evaluators.

Questions/Project Goals

The purpose of the needs and usability assessment study ultimately was to support Engineering Pathway in achieving its goals. More specifically, the evaluation was intended to:

- Evaluate the needs of Engineering Pathway's users, specifically in Higher Education
- Evaluate the usability of Engineering Pathway
- Provide suggestions for future development, based on the results of our evaluations

Data & How Gathered

Competitive Analysis: Comparing EngineeringPathway.com with Merlot.org, NSDL.org, Google.com, DigitalEngineeringLibrary.com. Criteria for Evaluation included Home Page, Navigation, Site Organization, Links/Labels, Search/Search Results, Readability, Performance, and Content.

Heuristic Analysis: three evaluators enacted three difference scenarios (Home Page Awareness, Search for a Resource, Submit resource). Heuristics included visibility of system status, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help and documentation, standard query box, relevance of results, scannable text, display of searched terms, display size of result set, support iterative search, optimize search results.

Survey: It was difficult to get a solid response rate; therefore we resorted to Focus Groups and Interviews.

Graduate Student Instructor (GSI) and Teaching Assistant (TA) Focus Groups: Topics included the specifics of GSI/TA work, especially how they create, find, and share learning resources with other GSIs. We also showed the focus group EngineeringPathway.com to get their feedback.

Interviews with Faculty Members: Topics included working with teaching teams, preparing coursework, finding teaching resources, evaluating teaching resources, sharing resources with different audiences, and a walk-through of an example search and submit on EngineeringPathway.com.

Data Analysis

The Competitive Analysis and Heuristic Evaluation were both quantified qualitative approaches that used a benchmark system or heuristic in order to evaluate the effectiveness of the site. The responses from interviews and focus groups were clustered into themes to demonstrate the range of needs these users had, as well as some of their initial feedback and impressions of EngineeringPathway.com

Conclusions

At the end of the Needs and Usability Assessment, the evaluation team provided recommendations for improvement. These included:

- *Help First-Time Users Navigate and Understand the Site*: after the evaluation, Engineering Pathway added a "First Time User? Questions? Get Help and Answers Here!" link, available on the main page on the left navigation menu.
- *Provide the Ability to Browse the Collection*: after the evaluation, Engineering Pathway implemented a "Browse Resources" feature, available from their main left navigation menu.
- *Enable and Build Engineering Education Community*: after the evaluation, Engineering Pathway implemented community pages to allow specific disciplinary communities to get involve and share relevant learning resources.

4.2 Vignette 2: MatDL Pathway and Virtual Labs

Contributor: Cathy Lowe, Kent State University

Project Activity: Professional Development and Training

Impact: Knowledge Building

Project Description

With support from the Course, Curriculum, and Laboratory Improvement (CCLI) Program, funded by the National Science Foundation (NSF), a set of Virtual Labs (VLs) was designed to help students connect microscopic structure and principles to macroscopic outcomes. Formative learning assessment and community-focused dissemination were integrated into the development of these interdisciplinary digital learning resources. The design team included researchers in materials science, chemistry, biophysics, information science, and learning science from MIT, Carnegie Mellon (CMU), Kent State University (KSU), and Pittsburgh Science of Learning Center (PSLC). While the VL simulations are freely available for anyone to use on the NSDL Materials Digital Library Pathway (MatDL) Virtual Labs Wiki (http://matdl.org/virtuallabs), faculty teaching undergraduate science courses and their students are the primary intended audience for these resources.

The VLs have been used in three introductory undergraduate courses; evaluated by students enrolled in those courses; and disseminated through the NSDL (http://nsdl.org) and MatDL Repository (http://matdl.org/repository).

Question to Answer

The design team was ultimately interested in finding out if the VLs positively influenced student learning of target concepts, but also interested in using student feedback to improve the VL experience.

Data and How Gathered

Data was collected online from students enrolled in specific introductory undergraduate science courses at the participating institutions. Initially, the VLs were introduced to small classes and the design team was most interested in conducting formative evaluations to inform improvements to the VLs. While the formative component continues to be important, the resources have reached a more mature stage of development and student learning can be better assessed. Introduction of the VLs was timed to coincide with lecture coverage of the target concepts. Students were asked to take a pre-test of understanding, then to complete the VL activities followed by a post-test, and finally to complete a brief feedback survey. Online assessments and surveys were created and administered using the popular SurveyMonkey online survey tool.

Data Analysis

To date, the largest group of students to use these simulations has been the fall 2009 class of MIT's 3.091 Introduction to Solid State Chemistry. In order to gauge student learning from the VL activity, MIT students were asked to complete a series of questions both before and after using the entropy simulation. The assessment consisted of multiple choice and

free response questions that were designed to tap in to student understanding of free energy diagrams and the influence of temperature on chemical systems.

The analysis included data from 334 MIT students who completed both pre- and post-tests. Our analysis suggested that students improved their understanding of the target concepts by completing the VL activity. A 2-tailed paired sample t-test revealed that students performed significantly better on the post-test ($M = .77$) than on the pre-test ($M = .66$), $t(333) = 12.285$, $p < .001$. In addition to the pre- and post-test assessments, a survey was administered to measure student perceptions about the simulations. 334 MIT students completed the survey, and the results suggested that students found the virtual labs to be a positive learning experience. Students responded to prompts on a 5-point scale where 5 indicated Strongly Agree and 1 indicated Strongly Disagree. Overall, the majority of students agreed that the simulation helped them to connect concepts in new ways (75%), helped them to see how the same principles applied to different topics (70%), was well organized (69%), was easy to understand (70%), gave them a deeper understanding of principles they already knew (65%), and gave them a deeper understanding of principles they did not know before (58%). In an opened ended comment section students most frequently indicated that they liked the interactivity and visual nature of the applets. Some comments combined both aspects e.g., "the interactive applets that made visualization easy." All of these results are consistent with those obtained from smaller classes at the other participating institutions.

Conclusions

Results from the VL activities conducted with MIT's 3.091 Introduction to Solid State Chemistry during fall 2009 suggest that the simulations promote learning and engage students. These results are consistent with those obtained beginning in 2007 from smaller classes at the other participating institutions. Results have been used formatively to improve the VLs as well as to report progress to funding agencies. Future evaluation efforts will include a more rigorous design to allow for stronger conclusions to be drawn regarding student learning.

4.3 Vignette 3: MSP2 Social Network: Analyzing Member Profile Data

Contributor: Kimberly Lightle, The Ohio State University

Project Activity: Community Building

Impacts: Audience Reach, Project Decision-Making

Project Description

The Middle School Portal 2: Math and Science Pathways (MSP2) project supports middle grades educators with high-quality, standards-based resources and promotes collaboration and knowledge sharing among its users. Educators use MSP2 to increase content knowledge in science, mathematics, and appropriate pedagogy for youth ages 10 to 15. MSP2 employs social networking and digital tools to foster dynamic experiences that promote creation, modification, and sharing of resources, facilitate professional development, and support the integration of technology into practice. The MSP2 social network was built using NING (http://ning.com) and can be found at http://msteacher2.org. The MSP2 social network was made public on February 1, 2009.

Questions to Answer

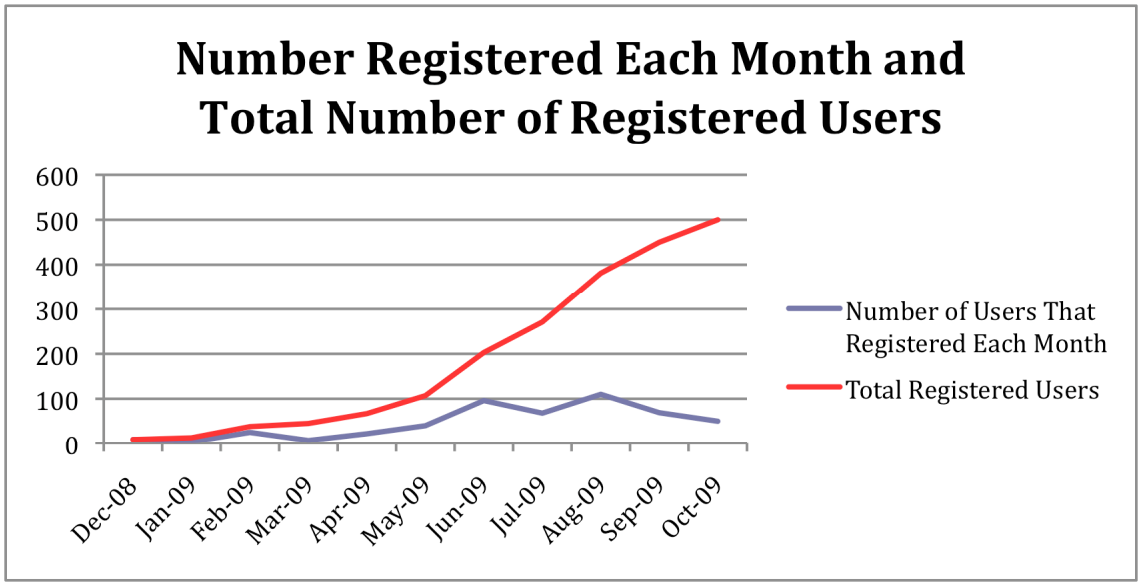
- How many visitors "join" the community by becoming registered members?
- How did registered members discover the MSP2 social network?
- How many users fill out their data profile when joining and to what extent?
- What is the make-up of registered users based on data profile information?

Data & How Gathered

NING allows for an unlimited number of questions to be included as part of the registration process. Questions can be marked as Private or Public. Public questions are viewable on members personal My Pages and can be searched on in the Member section of the social network. MSP2 asks members about their position, grades taught (private), years taught (private), disciplines taught, and how they discovered the MSP2 social network (private). In addition to those demographic questions, two public questions provide a list of educational topics/issues where members can select the ones they "Need help with" and "Can help with." As of November 13, 2009, MSP2 had 565 registered members. 86% of those members provided an answer to at least one of the profile questions. The responses to all questions are exported to an Excel spreadsheet via a Manage tab as part of the NING system.

Data Analysis

Member profile data are downloaded each month, counts are made, and comparisons with previous months conducted. The following data were downloaded November 13, 2009.



How did you find out about MSP2? (400 responses)

Friend or colleague – 137; Email notice – 64; Web search – 63; Print publication – 3; Conference – 35; Attended an online event/conference – 15; Other – 77

Years Taught (330 responses) – Average of 15 years

Position (489 responses)

Classroom Teacher (preK-12) – 319; Technology Integration Specialist – 11; Student – 21; Researcher – 11; Principal/Assistant Principal – 4; Preservice Teacher – 16; Librarian/Media Specialist – 4; IT Support Person – 4; Informal Educator – 7; Homeschool Educator – 2; Curriculum Coordinator – 10; College Faculty – 22; Agency/Association Professional – 3; Other – 63

Grade Level Taught (370 responses)

Preschool – 5; Primary (K-2) – 13; Elementary (3-5) – 36; Middle (6-8) – 324; High – 40; Other – 19

What discipline(s) do you teach? (414 responses)

Mathematics – 195; Science – 215; English/Language Arts – 22; Social Studies – 17; Technology – 20; Music – 3; Physical Education – 1; Foreign Language – 1; Special Education – 5; Enrichment – 4; Other – 118

I need help with… (106 responses)

Assessment/Testing – 38; Being a new teacher – 8; Classroom management – 25; Content knowledge – 22; ESL/ELL – 13; Gifted and talented – 30; Instructional strategies – 53; Integrating technology – 59; Involving parents/caregivers – 15; Literacy – 9; Managing Labs (including safety) – 14; Science fairs/competitions – 16; Special needs students – 25; Standards – 9; Student misconceptions – 22

I can help with… (91 responses)

Assessment/Testing – 19; Being a new teacher – 22; Classroom management – 29; Content knowledge – 45; ESL/ELL – 4; Gifted and talented – 4 ; Instructional strategies – 36; Integrating technology – 30; Involving parents/caregivers – 9; Literacy – 7; Managing Labs (including safety) – 6; Science fairs/competitions – 6; Special needs students – 12; Standards – 22; Student misconceptions – 13

Conclusions

The profile data paint a picture of the membership of the MSP2 social network, how they found out about MSP2, and to a certain extent, their needs. Word of mouth seems to be the best way of getting the word out about the site. We will continue to encourage members to invite their colleagues to join. We will also continue to send out emails through email lists and organizations that work with science and math teachers. We continue to provide resources and services directed at middle school science and math teachers and continue to focus on professional development experiences in integrating technology, instructional strategies, and different populations of students. There are many more questions that need to be asked including what members do while they are in the site and why only a small percentage of MSP2 visitors actually join the network. We also need to unpack what "Other" is – many users are choosing that term.

4.4 Vignette 4: Beyond Penguins and Polar Bears

Contributor: Jessica Fries-Gaither, Ohio State University

Project Activity: Web Portal or Web Site

Impact: Information Seeking Trends; Project Decision-Making

Project Description

Beyond Penguins and Polar Bears (http://beyondpenguins.nsdl.org) is an online magazine (cyberzine) designed to help K-5 teachers learn science content and integrate science and literacy instruction. The magazine connects the elementary science curriculum and literacy strategies to the real-world context of the polar regions. Funded jointly by the **Discovery-Research K12** and the **International Polar Year** programs, the project is a collaboration with the NSDL.

Questions

The *Beyond Penguins and Polar Bears* project team would like to know the following about **information seeking trends** with respect to the cyberzine:

- What kinds of science and literacy content are elementary teachers most interested in?
- What kinds of resources are teachers searching for on the web?
- Does the desired content change over the course of the school year?
- What kinds of content should we provide to best support our target audience and project goal?

Data and How Gathered

The *Beyond Penguins and Polar Bears* cyberzine and blog are configured to collect data with Google Analytics (GA) since the project launch in October 2007. Metrics are collected on a monthly basis to coincide with the cyberzine's publishing schedule. Monthly web metrics reports for the project are available at http://onramp.nsdl.org/collection/onramp:1062.

While many metrics are collected on a monthly basis, the following **general web metrics** are useful in answering the questions listed above:

- Top content (most popular pages)
- Keyword searches (search engine terms)

Each month, the 30 pages from the cyberzine with the most page views (shown as *Top Content* in GA) are recorded and categorized according to resource type: lesson plans, teacher content knowledge, nonfiction reference article, etc. The top content feature shows pages in order of views and includes all archived material, not just the most recent issue of the cyberzine.

In addition, use of the *Traffic Sources* feature of GA provides an analysis of the keyword searches performed across all search engines. There are thousands of keyword searches every month. The top 15 are recorded and cross-searched in Google to determine their

current ranking in a search engine return. The remaining keyword searches are browsed and a representative sample of the results is recorded without specific metrics.

Number of visits or page views and time on page is also recorded for both top content and keyword searches. For keyword searches, additional metrics (pages per visit, percentage of new visitors) are also collected.

Data Analysis

The two metrics discussed in the prior section (top content and keyword searches) are used to make inferences about the needs of elementary teachers and information seeking trends for the site's users. Given the online nature of the cyberzine, it is not possible to know with certainty who the visitors are or if they even match the target audience (K-5 teachers). A more traditional needs assessment of the users is also not possible as it would be for a traditional print magazine.

The two metrics paint slightly different pictures of information seeking trends. Keyword searches provide insight into what topics and types of resources are of interest to users who end up at the *Beyond Penguins and Polar Bears* site. The time on page metric is compared to the overall average of time on site to get a rough sense of how engaging the content was to that particular user (or group of users). If a user found the site through a keyword search and stayed for a significant amount of time, it can be inferred that the page met their expectations or provided enough engaging content to compel them to stay.

In contrast, the top content metric provides insight into what specific cyberzine articles were of interest to visitors to the magazine, as they often visit more than one page on the site. While the topics of these pages are important, so is their general resource type (lesson plans, teacher content knowledge, etc). The information gained from this metric helps us infer what parts of the cyberzine are most compelling to our users.

Since data has been collected for well over a year, long-term trends in the data can be discerned by comparing monthly reports. We've found that while some topics are in high-demand year round, others come and go on an annual cycle according to the school calendar. This type of understanding of the changes in needs will help us improve the project in the future and support the needs of elementary teachers with future projects.

Conclusions

Monitoring the needs of elementary teachers and the use of a complex site such as a cyberzine are difficult given complexities such as search engine traffic and multiple modes of discovery and site navigation. Tracking information such as top content and keyword searches provides a way to make inferences about the needs and wants of our actual users, whether they match our target audience or not.

4.5 Vignette 5: Instructional Architect

Contributors: Beijie Xu & Mimi Recker, Utah State University

Project Activity: Web portal or Web site

Impact: Usage

Project Description

The Instructional Architect (IA) (http://ia.usu.edu) is an educational digital library service to help K12 teachers search, select, sequence, annotate, and reuse online learning resources from the NSDL and Web to create instructional web pages for their students called IA projects. The project was funded under the Services track of the NSDL program, of the National Science Foundation.

System Description

Resource collection – teachers can search for and store links to NSDL resources within the IA context, as well as provide a name and add non-NSDL resource links to their own collection.

Project creation – teachers can create an IA project web page and share it with the public (public view) or only with their own students (student view). The teacher's resource collections are listed on the left, and can be added to the project with a single click.

Questions

- Is IA being used by the audiences we designed it to have?
- How many teachers compared to students are using IA?
- Are more teachers using IA in regions where professional development workshops were conducted?
- Did the outreach and PD workshops work towards increasing users of IA?

Data and How Gathered

The IA has been engineered to collect data with Google Analytics (GA) since early 2006. Teacher pre/post workshop surveys were also collected. IA also uses a customized database that captures many low-level actions on a per-user basis. Specifically, the database collected the following information:

Collections, Collection building	*User participation*
• number of resources • number of projects	• number of registered users • number of resources collected • number of teacher projects created

See book chapter at http://edm.usu.edu/publications/EDM_chapter_final.pdf for references to specific Figures illustrating the following discussion.

Data Analysis

Several different lenses (variables) were used to view the data, and in some cases, were combined: users, usage traffic source, bounce rate, location, content, and time.

Users - A plot of the number of users against the number of IA projects created displays the characteristic long-tail distribution, except in the middle (Xu & Hsi, 2009). This means that approximately 10% of the users account for 90% of the created projects.

Usage - Aggregated usage over a 12-month period can show trends, in this case, a dramatic drop-off during the summer reflecting its school-based usage (See Figure 6 in book chapter).

Traffic source – Examining the direct traffic (high - 57.45%), and concurrently looking at bounce rate (low - 33.93%) showed that the IA has more repeated users and purposeful usage.

Location - After an analysis of the geographic location of IA visitors over a 6-month period, we found out the IA attracts visits from Utah and other places where outreach activities have been conducted. This shows the success of local dissemination (See Figure 7a in book chapter).

Content – Using two sources of data, database and Google Analytics, we can show that the ratio of student login paths to teacher login paths was nearly constant, at about 4:1. This suggests the site is used as intended: teachers create IA projects for their students to use.

Conclusions

We are interested in understanding the Instructional Architect users and their usage patterns. Web usage data could be collected through a variety of sources:

- *Web server log.* It entails no extra cost for the web developers; it is automatically updated whenever a user transaction occurs. But its format is fixed, and thus limited in its capability of providing user information.

- *Google Analytics.* It is free too. Due to the privacy issue, GA does not provide user-level data; however, it provides numerous high-level reports. So, it is a good starting point for understanding website usage.

- *Customized database.* Researchers could configure a database to capture user information of different perspectives. Customized database is highly recommended for analyzing web usage.

- *Data Cleaning* is extremely time consuming. In our case, usage by robots and project developers was removed. Time zone differences due to the increasing global audience were also fixed.

- *Data Transformation* is complex. Decisions about how to aggregate a sequence of user activity into meaningful measures are important. For example, based on user traces, what counts as a successful use by either a teacher or a student? How do we define site "stickiness"?

Because of noise and differences in definitions, different methods for tracking usage data can report different values. In our case, it was important to note that GA and IARD (customized database) measures correlated highly, and that overall deviations were small (less than 10%).

GA shows that the update of the IA is highest in parts of the U.S. with teacher professional development activities. Yearly comparison in visits also showed a large increase in traffic from the NSDL once it began to automatically harvest IA projects.

Because of the large volume of data collected, it is critical to have questions drive the research. Of course, as data analysis and mining progresses, these questions should be repeatedly revisited.

4.6 Vignette 6: NSDL Science Literacy Maps (Service/Tool)

Contributor: Sharon Clark, NSDL Technical Network Services

Project Activity: Web portal or web site

Impact: Usage, project decision-making

Project Description

The NSDL Science Literacy Maps (http://strandmaps.nsdl.org) are an interactive graphical interface that helps K-12 educators and learners understand the relationships between science concepts, and find associated educational resources supporting the teaching of those concepts. The interactive maps are generated through a Web 2.0 Application Programming Interface (API) that lets developers embed the maps in their own web sites and display educational resources and other information in the maps. The maps illustrate learning goals for different grades, and the relationships between goals, for K-12 students across a range of science, technology, engineering, and mathematics (STEM) disciplines.

Questions

The developers of the NSDL Science Literacy Maps would like to know the following about user behavior:

- What maps and benchmarks are used?
- Which tabs within the *information bubble* are viewed?
- Which educational resources are viewed?
- Are users using the aligned resources from the *Top Picks* tab or are they using the programmatically generated resources from the *Related Resources* tab?
- Do users look at the common student misconceptions?
- What other dynamic user interface (UI) features do users use?

Since the NSDL Science Literacy Maps heavily rely on dynamic technologies such as AJAX, specific user activity is difficult to discern from traditional page access analysis. While access to each map can be determined by any analytics program that can understand URL parameters, gathering data for the other questions requires advanced instrumentation of the NSDL Science Literacy Maps UI because most user actions do not correspond to a separate page load that can be tracked via typical web metrics means.

For example, when a user clicks on a benchmark within a map, a dynamic popup, called the *information bubble*, is displayed and is entirely controlled by AJAX. The webpage does not need to reload in order to display the information bubble. Within each bubble is a variety of content sorted by tabs, similar to tab-based navigation on traditional websites. As the user clicks the different content tabs, AJAX is used to re-populate the content of the bubble and does not require a page reload. A typical analytics program would record a single page view for this user even if the user viewed several benchmark information bubbles and clicked on any of the tabs within each bubble.

Data and How Gathered

To gather the necessary data to answer the usage questions, the NSDL Science Literacy Maps have been instrumented with Google Analytics to track certain actions using the trackPageView() function, pageTracker, trackPageview(path), where *path* is a fake URL that

represents the user action. To track which benchmarks were opened within a map, NSDL sends the following path to Google Analytics:

/content_click/bubble/MAP_ID/BENCHMARK_ID/

Similarly, to track when the user views the *Top Picks* tab within a benchmark bubble, we send the following path to Google Analytics:

/content_click/tab/top_picks/MAP_ID/BENCHMARK_ID/

The structure of the fake URL enables the developer to more accurately follow user click paths with Google Analytics' navigation summary feature. It also enables the developer to use Google Analytics' content drilldown feature to determine which UI features are the most used. For example, to track which educational resource a user visited, and include other relevant information such as the corresponding tab, map ID, and benchmark ID the resource was found in, the collection ID of the collection the resource is a member of, as well as which position in the results the resource was located, we send the following path to Google Analytics:

/open_resource/tab/TAB_NAME/MAP_ID/BENCHMARK_ID/COLLECTION_ID/RESOURCE_URL/2

This fake URL enables the developers to use the *content drilldown* feature of Google Analytics to determine which resources are often used from the NSDL Science Literacy Maps display.

In order to provide an aggregated picture of which resources are viewed by users, it is necessary to provide an alternative means to view/access resources that removes the context of which tab was accessed by users. To do this, a new profile was created within Google Analytics containing rewrite filters that changed

/open_resource/tab/related_resources/ and */open_resource/tab/top_picks/*

to */open_resource/tab/*

To create a generalized view of resource usage, another profile could be set up within Google Analytics containing rewrite filters that eliminate the map ID, benchmark ID and tab names from the path, displaying only the collection ID, resource URL, and resource position within search results.

Unlike the benchmark *information bubble* and content tabs, a click on an education resource does not correspond to a legitimate page view. To avoid inflating page view counts in reports, a separate profile has been created to collect only educational resource clicks and a filter has been set up to not include this data in the main website profile.

Conclusions

Results could be used to determine if subject areas that are popular with end-users are lacking sufficient educational resources by seeing if any heavily accessed maps and benchmarks do not correlate to many, or any, opened resources within those benchmarks. This data could influence focused resource recruitment efforts. The generalized view of the resource clicks, combined with similar information from the NSDL.org search results, can be used to determine which collections and resources are the most popular overall, and this in turn, provides essential information toward effectively meeting user needs.

5 Resources

5.1 Metrics Tools

No web metrics package is 100% accurate and metrics are best used to identify trends, as opposed to reporting raw numbers. Many projects utilize more than one web metrics package, and compare data between the two. A selection of web metrics packages are shown below. Additions to the list are sought, and may be contributed via:

1. NSDL Contact form at http://nsdl.org/about/contactus/ (enter: web metrics suggestion" in the subject line of email)

2. Got Metrics? Forum on the NSDL Community site: https://www.nsdlnetwork.org/forums/general-discussions/got-metrics

Selected Web Metrics Packages

- Analog — http://www.analog.cx/
  Open source log analysis software, comparable to Webalizer or AWStats

- AW Stats — http://awstats.sourceforge.net
  Open source log analyzer that generates advanced web, streaming, ftp or mail server statistics, graphically

- Alexa — http://www.alexa.com
  A subsidiary of Amazon.com. Collects and analyzes browsing behavior and web traffic reporting by installation of the Alexa toolbar. Ranks sites based on tracking information of users of its toolbar.

- Google Analytics — http://www.google.com/analytics/
  Open source, generally easy to install and configure; no convenient online Help desk, other than forums that take time to crawl for the issue you may be having

- Mint — http://haveamint.com

- Omniture — http://www.omniture.com/en/
  Very customizable and advanced reporting; generally speaking more than a typical project may need, especially when starting out; expensive.

- Piwik — http://piwik.org/
  Piwik aims to be an open source alternative to Google Analytics. Piwik is a PHP MySQL software program that you download and install on your own webserver.

- Urchin — http://www.google.com/urchin/

Also from Google, but not free - designed for use on content behind a security firewall where GA is not applicable

- Webalizer — http://www.webalizer.org/
  Open source web log analysis software

- WebTrends — http://www.webtrends.com/
  An enterprise level web analytics service offering multiple products. Requires some expertise.

5.2 Quick References

The following sites or articles provide observation or discussion on how selected packages compare or other helpful guidance and discussion:

- Bart Gibby – Omniture Sitecatalyst vs Google Analytics: http://www.bartgibby.com/2006/10/14/omniture-sitecatalyst-vs-google-analytics/

- IdealWare "A few good web analytics tools": http://www.idealware.org/articles/fgt_web_analytics.php

- Kaushik, Avanash: Occam's Razor, Analytics blog: http://www.kaushik.net/avinash/

- Mashable: Internet news blog on social networks use and issues: http://mashable.com

- SEO'Brien Analytics vendor comparison: Omniture, Coremetrics, Hitbox and Google Analytics: http://www.seobrien.com/2007/08/analytics/analytics-vendor-comparison-omniture-coremetrics-hitbox-and-google/

- Trending Upwards' Omniture vs. Google Analytics: http://www.trendingupward.net/2008/11/omniture-google-analytics/

- NSDL Brown Bag, August 6, 2009: Google Analytics: Making Metrics Make Sense, presented by Jonathan Weber, Lunametrics, LLC: https://cc.readytalk.com/play?id=67fqhb

5.3 Glossary

Bounce rate – The percentage of visitors who "bounce" away to a different site from the entrance page; the percentage of entrances on a web page that result in an immediate exit from the web site.

Conversion rate - The relationship between visitors to a web site and actions considered to be a "conversion," such as a sale or request to receive more information; often expressed as a percentage.

Diigo – A social bookmarking service enabling organization, tagging, highlighting, and sharing of web pages and sites; http://www.diigo.com. "Diigo" is an abbreviation for "Digest of Internet Information, Groups and Other stuff"

Logfile – A server log is a log file (one or more) automatically created and maintained by a server, of activity performed by it.

Logfile analysis – the parsing of a log file from a web server, and the derivation of indicators about who, when, and how a web server is visited, based on the values contained in a log file.

Long Tail: First coined by Chris Anderson in an October 2004 Wired magazine article to describe the niche strategy of certain businesses, such as Amazon.com or Netflix. In relation to search engine marketing (SEM) the Long Tail refers to the keyword phrases that are highly detailed and specific and may generate low volumes of searches and traffic, but add up to generate a majority of traffic for sites with deep content or product SKUs (*Stock Keeping Unit*, pronounced "skew." A SKU is a number or string of alpha and numeric characters that uniquely identify a product. For this reason, SKUs are often called part numbers, product numbers, and product identifiers.)

Page tag - A piece of JavaScript code embedded on a web page and executed by the browser when the page is viewed.

Page view – Number of times a web page is accessed during a single visit, including repeated viewing of the same page. A request for a file whose type is defined as a *page* in log analysis. An occurrence of the script being run in page tagging. In log analysis, a single page view may generate multiple hits as all the resources required to view the page (images, .js and .css files) are also requested from the web server.

SEO: Search Engine Optimization is the improvement of rankings for relevant keywords in search results by adjusting website structure and content to make them more easily read and understood by a search engine's software programs.

Session length – Time spent by a visitor during each single visit.

Site 'stickiness' - The amount of time spent at a site over a given time period. Often measured in the average minutes per month visitors spend at a site or network. Sometimes measured in terms of page views. When defined as minutes per month, site stickiness is a function of number of visits (repeat usage) and time spent per visit (session stickiness).

Unique visitors – Refers to a measure captured by some web analytics solutions that track the interaction a single user has with a website over time. The uniquely identified client generating requests on the web server (log analysis) or viewing pages (page tagging) within a defined time period (i.e. day, week or month). A Unique Visitor counts once within the timescale. A visitor can make multiple visits. Identification is made to the visitor's computer, not the person, usually via cookie and/or IP+User Agent. Thus the same person visiting from two different computers will count as two Unique Visitors.

Visitor: Similar to unique visitor, visitor refers to an individual that visits a website. A visitor or unique visitor can have multiple visits.

Visit/Session: A series of requests from the same uniquely identified client with a set timeout, often 30 minutes. A visit contains one or more page views. Interaction by a site visitor; the session ends when the visitor leaves the site.

5.4 References

Evaluation Resource Institute (2009). Evaluation Definition: What is Evaluation? Retrieved 13 November, 2009, from http://www.evaluationwiki.org/index.php/Evaluation_Definition

Khoo, M., Buchanan, G., & Cunningham, S. J. (2009). JCDL Workshop Report: Lightweight User-Friendly Evaluation Knowledge for Digital Librarians. *D-Lib Magazine, 15*(7/8).

Manduca, C. A., Iverson, E. R., Fox, S., & McMartin, F. P. (2005). Influencing User Behavior through Digital Library Design. *D-Lib Magazine, 11*(5). Retrieved 15 November 2009 from http://www.dlib.org/dlib/may05/fox/05fox.html

NSF (2009a). National Science Foundation. Retrieved 13 November 2009, from http://nsf.gov

NSF (2009b). National STEM Education Distributed Learning program. Retrieved 13 November, 2009, from http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5487&org=DUE&from=home

Reeves, T. C., Apedoe, X., & Woo, Y. H. (2003). *Evaluating Digital Libraries: A User-Friendly Guide* (Prepared for the National Science Digital Library (NSDL)). Atlanta, GA: University of Georgia.

Xu, B., & Hsi, S. (2009). The Data Deluge: Opportunities for Research in Educational Digital Libraries. *Digital Libraries: Web Development, Software Technologies and Data Management*.