



Industry Advisory Council
eGovernment Shared Interest Group

White Paper on:

The Use of Metrics
in
Electronic Records Management (ERM)
Systems

- August 2004 -

Disclaimer

While the American Council for Technology/Industry Advisory Council (ACT/IAC) has made every effort to present accurate and reliable information in this report, ACT/IAC does not endorse, approve or certify such information, nor does it guarantee the accuracy, completeness, efficacy, timeliness, or correct sequencing of such information. Use of such information is voluntary, and reliance on it should only be undertaken after an independent review of its accuracy, completeness, efficacy, and timeliness. Reference herein to any specific commercial product, process, or service by trade name, trademark, service mark, manufacturer, or otherwise does not constitute or imply endorsement, recommendation, or favoring by ACT/IAC.

ACT/IAC (including its employees and agents) assumes no responsibility for consequences resulting from the use of the information herein, or from use of the information obtained from any source referenced herein, or in any respect for the content of such information, including (but not limited to) errors or omissions, the accuracy or reasonableness of factual or scientific assumptions, studies or conclusions, the defamatory nature of statements, ownership of copyright or other intellectual property rights, and the violation of property, privacy or personal rights of others. ACT/IAC is not responsible for, and expressly disclaims all liability for, damages of any kind arising out of use, reference to or reliance on such information. No guarantees or warranties, including (but not limited to) any express or implied warranties of merchantability or fitness for a particular use or purpose, are made by ACT/IAC with respect to such information.

Table of Contents

Executive Summary	1
1. Background and Introduction	3
2. IAC Study Group Challenges from OMB and NARA.....	6
2.1 Challenge from OMB	6
2.2 Challenge from NARA	7
3. Organization of Study	8
3.1 Study Leadership.....	8
3.2 Study Group Membership.....	8
3.3 Acknowledgements	10
3.3.1 Non-IAC Firms.....	10
3.3.2 Reviewers.....	10
3.3.3 Primary Authors of White Paper	10
3.4 IAC Study Guidelines	11
3.5 Study Methodology	12
3.6 Starting Point Assumptions and Constraints.....	16
4. Major Facts, Findings and Conclusions	18
5. Categories of Metrics and Specific Metrics.....	22
5.1 Access to ERM Services	23
5.2 Accuracy.....	24
5.3 Capacity.....	25
5.4 Efficiency	26
5.5 Participation.....	27
5.6 Productivity	28
5.7 Search and Retrieval	29
5.8 System.....	31
5.9 User Satisfaction	32
5.10 Utilization.....	35
5.11 Legal.....	35
6. Aggregating Metrics and the Balanced Scorecard Approach	41
6.1 Introduction.....	41
6.2 Roll-up Approaches	41
6.2.1 Return-on-Investment.....	41
6.2.2 Key Performance Indicators	41
6.2.3 Balanced Scorecards	42
6.3 Metrics System Evaluation – Achieving ERM Excellence.....	43

Executive Summary

In July 2003 the Industry Advisory Council (IAC) convened a Study Group at the request of the Office of Management and Budget (OMB) and the National Archives and Records Administration (NARA). The Study Group was asked to examine industry best practices in the area of metrics relating to electronic records management (ERM). Specifically, the Group was challenged to determine appropriate metrics that are meaningful without being unduly burdensome.

The Study Group was composed of volunteers from IAC member companies and was co-chaired by representatives from IAC and from NARA, whose role was to keep the participants focused on the specific challenges at hand. Non-member companies and other government agencies were also invited to participate. All activities were guided by IAC Study Guideline, detailed in Section 3.4 of this report. The primary methodology used was one of identifying and examining case studies or exemplars of candidate metrics in use today in organizations that have implemented ERM systems. The primary constraints for the study were time, and also the lack of direct, broad knowledge of the exemplars on the part of the participants. Thus, much of the study hinged on the use of anecdotal evidence.

The single most important conclusion reached by the Study Group is that there is no silver bullet. There is no one metric that seems to capture the success of an ERM system and relate unambiguously to the mission of an agency; nor is there a single, universal metrics capture and reporting tool (or product) that can be adopted for widespread use in ERM systems. Rather, there appears to be eleven (11) common categories of metrics in use today—some of which are less burdensome to capture than others, some of which reflect a measure of IT system performance, and some of which reflect mission success more directly than others. In fact, another of the Study Group's conclusions is that the measurement of ERM performance is currently immature, and that most measurements tend to be IT-related rather than related to records management itself. In addition, valid comparisons of ERM practices across agencies are difficult to make, and probably should not be made. Finally, the Study Group concludes that NARA and OMB should stimulate industry to develop new tools, products, and services that can help standardize and harmonize the efficient capture and reporting of ERM metrics across government.

The bulk of this document, as of the Study itself, deals with eleven specific categories of metrics identified through the exemplars. The Study Group identified metrics in the categories of access to services, accuracy, capacity, efficiency, participation, productivity, search and retrieval, system, user satisfaction, utilization, and legal. The legal category, in particular, touches on the need for identification of, and incorporation of, "best practices" in ERM

systems. The legal category also raises important issues as well as opportunities for potentially measuring or assessing the degree of compliance of an agency's ERM system to meeting legal and regulatory requirements.

Section 5 discusses many specific metrics under these categories, their ease of capture and how meaningful each might be. The discussion culminates with a chart, which summarizes these findings. Section 6 looks at the possibility of aggregating these metrics using a balanced scorecard approach. The inescapable conclusion, as noted above, is that there is no simple, single answer; tradeoffs must be made to arrive at metrics that are both meaningful to measure ERM success and not too burdensome to capture on an enterprise-wide basis.

1. Background and Introduction

In July 2003, at the request of the Office of Management and Budget (OMB), the National Archives and Records Administration (NARA) approached the Industry Advisory Council (IAC) about conducting a study focused on identifying best practices for measuring bottom-line, business-related performance enterprise-wide in modern electronic records management (ERM) systems.

The IAC was organized in 1989 as a non-profit Advisory Group within the American Council for Technology (ACT). See <http://www.iaconline.org>. The mission of the IAC is to bring industry and government executives together to exchange information, support professional development, improve communications and understanding, solve issues and build partnership and trust, thereby enhancing the government's ability to serve the nation's citizenry. This is accomplished by:

- Providing a forum for studies and analysis of public sector management issues, arriving at IAC positions on these issues, and presenting the results to the Chief Information Officers Council (CIO), the Procurement Executives Council (PEC), the Chief Financial Officers Council (CFO), the Office of Management and Budget (OMB), ACT, and other interested groups.
- Advising the ACT, the above noted councils, government managers and policy officials on the possible impacts of industry trends on government technology issues, serving as a sounding board for changes to Federal regulations, and assisting the ACT and the above noted councils in public relations and public affairs programs aimed at improving the health of government.
- Assisting the ACT and the above noted councils in identifying professional speakers to deliver individual presentations on technical, administrative, or managerial topics, selected by the ACT or its councils.
- Providing education and training.

In response to NARA's request and consistent with IAC's mission and charter, the IAC established the Electronic Records Management (ERM) Metrics Study initiative as a forum to explore ways of measuring modern electronic records management practices of potential interest to various government managers and policy officials.

This initiative falls under the aegis of the Best Practices Committee within the Industry Advisory Council's (IAC) eGovernment Shared Interest Group (SIG). The study effort commenced with a Kickoff Meeting of over 30 interested industry

representatives on July 31, 2003. At that time, the scope and objectives of the study, as well as the necessary time frame for task accomplishment, were discussed and refined:

- The scope of the study was defined to identify and discuss “best practices” associated with measuring performance of electronic records management systems (ERMs). Specifically, the Study Group was challenged to:
 - Review the current state-of-the-practice in measuring performance in various ERMs in industry and government,
 - Identify and characterize various categories of metrics,
 - Identify specific performance metrics as industry “best practices”.

While NARA expressed interest in identification of, and elucidation of, overall best practices and principles of operation for modern ERMs in the federal environment, NARA emphasized that the primary focus of the study needed to be on identification of best practices associated with metrics. NARA was particularly interested in learning best practices in applying such metrics on an enterprise-wide basis.

- NARA representatives specifically requested that the study focus primarily on identifying metrics relating to efficiency and effectiveness of a federal agency’s mission (and its service to its constituents). In addition, the NARA representatives also asked the Study Group to identify candidate metrics relating to the efficiency of the ERM system itself in support of the agency’s mission. The agreed-upon approach was to identify various ERM Projects across both government and industry, which could serve as “exemplars” and as sources of lessons learned on metrics. The NARA representatives also agreed that time constraints for the study would serve as a limiting factor to the scope of the project.
- NARA officials emphasized that the study was not directly related to, nor chartered by, the Electronic Records Archives (ERA) Project Office. NARA also clarified that the scope of the study more broadly related to overall NARA and federal agency e-government initiatives.
- Within the scope of the study, NARA stated that the terminology “electronic records” broadly covers electronic documents as well as fielded and relational data (e.g., structured records). The scope of the study covers official agency records as defined by the Federal Records Act and functionally includes: e-mail, (Geographical Information System (GIS) data, fielded data in IT systems, documents, memos,

images and photographs, multimedia, etc. For purposes of the study, NARA also stipulated that the essential criterion is that the records be in an electronic format either via being “born” digitally or being converted from paper and/or analog forms to digital format.

Consistent with the mission and charter of the IAC, this White Paper presents the major facts, findings, and conclusions of the Study Group that was formed at NARA’s request.

IAC wishes to ensure that its White Papers reflect a broad, comprehensive industry view of the topic. Accordingly, this White Paper is presented in a reasoned and objective factual manner with maximum attention and respect paid to minority and dissenting opinions of Study Team members. Indeed, the joint industry-government Study Group that was formed was open and collegial in its fact-finding activities, its analysis, and its discussions.

While the White Paper presents a consensus viewpoint of the Study Group members, it does not seek to provide advice, opinions, or recommendations from the group acting in a collective mode. In other words, no formal votes were taken on content. Additionally, the final version of this White Paper incorporates the beneficial review comments, suggestions, and recommendations of an independent peer-review panel of industry and government reviewers. Lastly, notwithstanding NARA officials serving as co-chair of the study and providing the ERM Initiative perspective, the White Paper also does not reflect an official NARA position.

2. IAC Study Group Challenges from OMB and NARA

2.1 Challenge from OMB

As part of the E-Government Strategy issued in April 2003¹, performance metrics were developed for each of the 24 E-Government (E-Gov) initiatives. In keeping with the Government Performance and Results Act and Office of Management and Budget (OMB) guidance, these metrics are intended to be outcome-based; i.e., they measure performance in terms of how well they accomplish the desired outcome or goal of the project (e.g., more citizens can file their tax returns electronically).

The E-Records Management (ERM) Initiative is part of the Internal Efficiency and Effectiveness E-Gov Portfolio. The Initiative will provide the tools that agencies will need to manage their records in electronic form, addressing specific areas of electronic records management where agencies are having major difficulties. This project will provide guidance on electronic records management applicable government-wide and will enable agencies to transfer electronic records to NARA in a variety of data types and formats so that they may be preserved for future use by the government and citizens.

OMB asked NARA, as managing partner of the ERM Initiative, to develop performance measurement metrics and reporting instructions to enable Federal agencies to measure their progress in implementing electronic records management. OMB defined the desired metrics as:

- Document search/retrieval burden
- Document recovery burden

Since these metrics are not currently captured across the Federal Government, NARA and OMB agreed that a baseline would be developed in FY 2004. The challenge was two-fold:

- To determine the appropriate unit of measurement. How should the broad metrics be defined? Should the burden metric be cast in terms of hours, dollars, or some other unit?
- To develop a reporting methodology that would accurately capture the data with the least reporting burden on the agencies.

¹ See http://www.whitehouse.gov/omb/egov/2003egov_strat.pdf, and particularly page 38 for the ERM Initiative.

2.2 Challenge from NARA

With the agreement of the Portfolio Manager, the NARA ERM Initiative co-managers approached the Industry Advisory Council (IAC)'s Committee on Best Practices for E-Gov for assistance in identifying best practices in capturing metrics for ERM. NARA hoped to find measurement practices in place in the private sector or at other government organizations that could be used as a model for capturing the OMB-specified metrics at Federal agencies.

NARA was also very interested in identification of any metrics that were not workable, viable, or appropriate for measuring the progress in implementation of electronic records management that could serve as lessons learned.

Because NARA was charged with developing a baseline for the fiscal year beginning October 1, 2003, NARA set an extremely ambitious schedule for the IAC Study Group to produce its White Paper. NARA also set the scope and objectives of the study effort as discussed in more detail in the Background and Introduction above. In requesting the study, NARA recognized that the tight time frame might affect the ability of the group to accomplish a broad and encompassing data collection activity. The essential work of the Study Group in consulting with NARA on best practices for metrics was completed in the October 2003 timeframe. This white paper documents the major findings and conclusions of the Study Group including beneficial comments and suggestions made from independent peer reviewers.

3. Organization of Study

3.1 Study Leadership

The Study Group was composed of NARA representatives, industry representatives, and representatives from other government activities. This section of the White Paper identifies the organization and composition of the Study Group:

- NARA Representatives:
 - Ms Nancy Allard
ERM Initiative Co-Manager
Government Study Chair
E-mail: nancy.allard@nara.gov
 - Mr. Mark Giguere
ERM Initiative Co-Manager,
E-Mail: mark.giguere@nara.gov

- IAC Representatives:
 - Ms. Tricia Reneau Iveson
SAIC
IAC eGovernment SIG Best Practices Committee Chair
E-mail: ivesonp@saic.com
 - Dr. Rick Klobuchar
SAIC Corporate VP and Chief Scientist
IAC Study Chair
Richard.L.Klobuchar@saic.com

3.2 Study Group Membership

The Study Group consisted of volunteers from IAC member firms as well as *ex officio* representatives from several government agencies, who also attended meetings.² All industry members of the Study Group were volunteers. Table 1 identifies members of the Study Group, who attended at least one (1) Study Group meeting, and their respective IAC member affiliations.

² Note: The Study Group warmly welcomed the contributions and beneficial discussions of several ERM experts and professionals, whose firms are not current members of IAC. Acknowledgement of their specific contributions is provided below.

Table 1. Members of the Study Group and Their IAC Member Affiliation
(presented alphabetically)

Study Group Member	Affiliation
Asma, Jake	Program Analyst, GSA OGP Interagency Management Division
Alatis, Cathy	IMC
Allard, Nancy	ERM Initiative, Co-Manager/Governmentt Study Chair, NARA
Dodge, Catherine	ECM Practice Director, Impact Innovations
Duong, Eric	ECS Technology, Inc.
Farish, Rob	Unisys
Galata Vitalie, Donna	Associate, Booz Allen Hamilton
Graham, Jim	VP Technology Management Services, LEADS Corp.
Giguere, Mark	ERM Initiative, Co-Manager, NARA
Gutierrez, Ana	Strategic Initiatives Office, CSC
Harrity, Carol	CSC
Hugney, Joanna	Principal, Booz Allen Hamilton
Iveson, Tricia	IAC eGov SIG Best Practices Committee Chair, SAIC
Jacobson, Peter	PS4C Partners, LLC
Khan, Zafar	PS4C Partners, LLC
Klobuchar, Rick	Corporate VP and Chief Scientist (SAIC) / Industry Study Chair
Le Floch, Patrick	PS4C Partners, LLC
Linden, Alan	ERM Specialist, SAIC
Malik, Adnan	EDS
Miller, Steve	QA Manager, Government Solutions, Treasury Division (PIDS), Northrop Grumman IT
Nemeth, Frank	Systems Engineer, EDS
O'Donnell, Mike	Director, Systems Management, CSC
Ostrum, Scott	Technical Consultant, HP
Patterson, Carla	Project Manager, E-Records Management Initiative, NARA
Plows, Sally	Product Specialist, Records Management Group, Open Text Corporation
Queen, Cheryl	CSC
Rice, Jim	Technical Director, East AmberPoint, Inc.
Roberts, Craig	Northrop Grumman IT
Rogers, Carolyn	HP
Sagar, Ajay	Vice President, Pixl, Inc.
Schock, Sandra	Project Manager, HP
Seraf, Neeran	Saraf Software Solutions

Study Group Member	Affiliation
Sheahan, Robert	BEA Systems
Simms, Marcus	Federal Sector, EzGov, Inc.
Taylor, Ted	HUD
Terkowitz, Roberta	Client Manager, IBM
Walker, Oliver	Records Management Officer, Office of Housing, HUD
Wing, Rob	Director, Systems and Technology Division, HUD

3.3 Acknowledgements

3.3.1 Non-IAC Firms

The Study Group wishes to recognize the contributions of the following non-IAC-affiliated individuals who provided beneficial consultation and guidance as well as assisting the Study Group in identifying ERM system exemplars across government and industry.

- Tim Sprehe, Sprehe Information Management Associates
- Bruce Schinkelwitz, re:er, Inc.

Their support and assistance across the study effort is most appreciated.

3.3.2 Reviewers

The IAC wishes to thank and recognize the following independent peer reviewers, whose beneficial suggestions and recommendations are incorporated in the final White Paper

- Robert Williams, President, Cohasset Associates
- Sherrie Smith, Office of the Comptroller of the Currency Records Officer
- Susan Sallaway, Office of the Comptroller of the Currency

3.3.3 Primary Authors of White Paper

In any study effort of this type, there is always a dedicated “core team” of members, whose service and contributions deserve to be recognized. The Industry Advisory Council wishes to specially recognize the contributions of the following individuals, who endured through multiple team meetings, phone calls, and e-mails; and who functioned as the major authors of this White Paper:

- Dr. Rick Klobuchar, SAIC
- Ms. Cathy Dodge, Impact Innovations
- Mr. Jake Asma, GSA
- Ms. Bobbi Terkowitz, IBM
- Ms. Cheryl Queen, CSC
- Mr. Frank Nemeth, EDS
- Ms. Cathy Alatis, IMC

3.4 IAC Study Guidelines

This section of the White Paper identifies the high-level rules and guidelines that were followed throughout the course of the study effort. The guidelines are derived from the following:

*Industry Advisory Council
Shared Interest Group
Operating Policies and Practices
(Approved 6/4/03)*

The Operating Policies and Practices recognize that on occasion, SIGs will author White Papers on topics of interest. IAC wishes to ensure that such papers reflect a broad, comprehensive view of the topic. As such, all white papers must be reviewed by at least three peers who are knowledgeable of the topic, and are not representatives of the same company or organization. After the peer review process is complete and the peer comments are incorporated, the white paper, along with the names of the peer reviewers, is submitted to the IAC Board, Vice Chair for Industry Liaison and the IAC Executive Director for final review and publication. All White Papers incorporate the IAC disclaimer statement.

IAC is a non-profit association, and must avoid certain lobbying activities in working with Governmental entities. Specifically, non-profit associations must avoid certain lobbying activities in order not to jeopardize their tax-exempt status. Also, associations that focus on governmental issues must avoid activities that could violate the Federal Advisory Committee Act or similar state or local laws. The following guidelines were applicable to the members of the Study Group convened to address NARA and OMB's request:

- Must not be "established" or "utilized" by an executive agency of the federal government. A group is "utilized" if an agency looks to it as a preferred source from which to obtain advice or recommendations on a specific issue or policy. Note: By requesting this study, NARA and OMB stipulate that the IAC Study Group and its members are not a "preferred source".

- Should not provide advice, opinions, or recommendations from the group acting in a collective mode. Note: This White Paper presents facts, finding, and conclusions regarding “best practices” for measuring performance in modern ERM systems. It does not seek to provide advice, opinions, or recommendations. Consistent with the notion of the IAC providing a forum for discussion, no formal vote on content was cast. Where appropriate, however, the White Paper does indicate where a general industry consensus was achieved.
- May provide information and viewpoints from individual participants.
- Should assure that every issue represents a wide range of viewpoints.
- Should assure that a group’s composition is not static over time.
- Should not contact or urge the public to contact members of a legislative body for the purpose of proposing, supporting, or opposing legislation or any other legislative action by Congress or by any state or local legislative body.
- Should not advocate the adoption or rejection of legislation or any other legislative action.
- Should not participate in or intervene in any political campaign on behalf of or in opposition to any candidate for public office.
- May participate in the nonpartisan analysis, study, or research of matters pertaining to legislation if it is strictly educational and does not constitute an attempt to influence legislation. The association must not advocate the adoption or rejection of legislation. All materials must be available to the public.
- May not appear before a legislative committee except in response to an official request for testimony.

3.5 Study Methodology

This section of the White Paper briefly identifies the overall methodology that was followed as part of the data collection effort and the ensuing technical discussions.

Identification of Specific ERM Projects and Systems as Exemplars.

Members of the Study Group were requested to submit data and information about organizations (private industry and government) where ERM performance metrics are captured or employed for current electronic records management systems, projects, and applications.

The Study Group desired that the projects, systems, and applications that were identified and surveyed should serve as “exemplars” for future consideration across the electronic records management community, where NARA could benefit by not having to “re-invent the wheel”. Specifically, the Study Team was interested in learning what metrics (if any) are in current use, how the metrics are utilized, what are their burdens or costs to capture, and what are their overall effects on the ERM project or the enterprise.

Types of Metrics Solicited. In attempting to identify exemplars and consistent with the challenges posed by OMB and NARA, the Study Group was particularly interested in identifying well-defined metrics associated with the following:

- Record search burden (e.g., finding electronic records)
- Record recovery burden (e.g., accessing and retrieving records)

Metrics for Public Services Relating to ERM. At the beginning of the Study effort, Study Group members were challenged to identify metrics relating to the bottom-line mission performance of federal agencies. The spirit of the eGovernment initiative is to provide a Government that “works better and costs less.” Thus, identification of quantifiable and well-defined ERM metrics relating to capacity, throughput, security (especially data and records integrity), assured service availability, ubiquitous access, lower cost, improved turnaround times, etc. were also of interest.

The Study Group was also concerned about particular metrics within a system that might have proven to be unreliable, non-specific, intractable to interpret, or too burdensome or onerous to collect. For these types of potentially “bad” metrics, the Study Group wanted to learn the specifics as potential lessons-learned.

Who is the Consumer? In studying and analyzing the metrics and best practices, the Study Group recognized that the nature of the “consumer” is also expected to be an important factor. In that regard, the Study Team was interested in learning about “who” and/or “what” the metrics are sampling. For example, in analyzing and evaluating the exemplars, the Study Group desired to know whether the metrics addressed the “public at large”, specific customers, agency/company employees, federal agencies, other government agencies, corporations, or foreign users, etc.

What is the ERM Business Practice? Lastly, in identifying and evaluating the exemplars, the Study Team wanted to know what specific “bottom-line” agency and/or industry business practices the metrics supported. The Study Group sought to identify exemplars with records search and recovery burden related to:

- Servicing FOIA requests

- Support for legal discovery
- Historical research
- Genealogy
- Auditing and controls
- Legal and regulatory compliance
- Public information dissemination
- Statistical analysis
- Archival records management
- Grants management
- ERM systems operations and management
- Specific agency mission support (e.g., medical, environmental, emergency and disaster, defense), etc.

In most cases, it was anticipated that the specific mission support that the highlighted ERM system/project provided as an exemplar would be readily apparent by merely noting the cognizant agency or industry activity involved.

Data Collection Protocol. After the kickoff meeting and within the scope of the study as mentioned above, members of the Study Group were requested to provide data and information on various ERM systems and/or projects that they were cognizant of, and knowledgeable of, and that they believed could serve as exemplars for NARA to provide guidance to other federal agencies on “best practices” for measuring performance in ERM systems.

As discussed at the kickoff meeting, this data collection effort was expected to be mostly an “essay test” to initially scope and bound (e.g., survey) the metrics that are currently in vogue across the ERM community. As the project progressed, the NARA representatives continually reiterated their desire for objective and thoughtful answers on the part of industry to some very tough questions to meet the mandate of the OMB with regard to metrics. Accordingly, the members of the Study team were encouraged to be forthright, objective, and forthcoming in their responses and inputs.

The following identifies the type of data and information that was solicited on the candidate exemplars:

**Requested Information About the Relevant ERM System, Project,
Corporation, Agency, or Application
(e.g., the “Exemplar”)**

Name of the Exemplar?

Your role and/or involvement with the Exemplar?

Does this Exemplar capture the specific metrics of document retrieval and record recovery burden?
If not, what metrics are captured?

How are those metrics defined?

What are the techniques employed to measure those metrics?

For whichever specific metrics you provide, anecdotally quantify the burden of capture on a continuum from ‘easy’ to ‘hard’.

Are there any restrictions on the use of the data and information provided herein?

Can this information be posted on the IAC Best Practices Web Site for this Study? If not, please advise on any caveats or restrictions, including any permissions and timing considerations. For example, the “Name of the Exemplar” needs to be kept anonymous.

Metrics. Please list and briefly describe the metrics in use within this Exemplar. In the discussion, please address the following for each metric:

- Definition of the metric including any units of measure (e.g., hours/minutes saved, customer response time, numbers of customers served per unit time, cost to service a customer request, etc.)?
- What services relating to ERM (such as e-Government services) are being measured?
- Who or what is getting measured? (e.g., Who is the Consumer?)
- What agency and/or industry business practices or missions are impacted or assessed with this metric?
- How is the metric collected?
- Frequency of collection of the metric?
- Estimated burden for collecting the metric? Initial burden? Recurring burden?
- Estimated return on investment (ROI) for the metric (if any)
- Do you have any “before” and “after” statistics or anecdotal evidence relating to the metric?
- Are there any electronic vs. paper paradigm issues associated with the metric?
- To what “use” is this metric applied? Who uses it? How is it used?
- Your assessment (Scale of 1 – 10 with 10 being the best) on how useful and viable this metric might be for future NARA and federal agency use?
- What is your estimate of the cost and/or level of effort that might be involved in “cloning” or duplicating the process and/or technology used to capture this metric for a federal agency? How hard do you think it would be to gain “re-use”?

General Lessons Learned, Discussion, and Recommendations. Please feel free to provide additional discussion, lessons learned, and recommendations relating to metrics and best practices for electronic records management – as evidenced within this Exemplar.

Deliberations and Discussions. To disseminate information among the Study Group, a collaborative Web site was established. Over the period from August – October, 2003, the Study group met on five (5) separate occasions to discuss facts, findings, and conclusions; as well as to develop and review sections of this White Paper.

3.6 Starting Point Assumptions and Constraints

In pursuing this study effort, the Study Group was relatively unimpeded with regards to starting point assumptions and constraints.

Assumptions. The major assumptions that were made included the following:

1. We assumed that identifying exemplars would be a good and viable approach to elucidate categories of metrics and specific performance metrics. Overall, this turned out to be a good assumption. While the Study Group noted some differences in terminology across the systems, it was generally possible to factor, group and categorize the metrics. Tacit in this assumption was that the Study Group also identified enough exemplars to properly characterize the state-of-the-practice for metrics capture, management, and reporting in modern ERM systems.
2. We initially assumed that metrics which reflect the contribution of the ERM system to the performance of the agency's mission are of greater value than metrics which relate solely to the performance of the ERM system itself. In the analysis of the exemplars, the Study Group determined that both type of metrics are desirable and necessary. For example, a system with poor performance (measured as a system response time) can lead to user frustration, which in turn can jeopardize the mission acceptance of the ERM system.
3. We initially assumed that electronic document search and retrieval times were potentially important in assessing and measuring the impact of an ERM system on an agency's bottom-line mission. As the Study Group identified and evaluated the exemplars, it started becoming apparent that this type of metric was intractable and subject to interpretation. In that regard, a shorter session time for the public searching for records does not necessarily equate to improved public perception of an agency's mission. In fact, the opposite may be true. In a poorly implemented system, the public at-large may become frustrated and abandon the system and its search capability even more quickly. By the same token, if there is keen public interest, members of the public may actually stay longer and make even more difficult and time-consuming searches. This phenomenon is analogous to "hooking the public on Web surfing."

Constraints. There were two (2) major constraints to this study:

1. The requested time frame for producing the study was short (e.g., a few months) and all of the labor was volunteer labor. Availability of personnel to go out and identify and characterize various ERM systems to an adequate level of detail was problematic. The problem was exacerbated to a degree by Hurricane Isabel which had a significant footprint on study participants in Virginia and Maryland.
2. The other major constraint was that various Study Group members frequently only had anecdotal knowledge about performance metrics being captured in ERM systems. For some systems, it proved difficult to secure an “official” position or a published and releasable source of information. That is, many of the study participants reported that they were working on, or aware of, proprietary or privileged ERM systems, that they were not at liberty to disclose without proper permission of the agency or the industry client. The short time frame for the study also exacerbated this problem. A related factor is that many of the Study Group members reported that their customers and clients did not want to disclose what they were measuring for competitive, privacy, and/or security reasons.

4. Major Facts, Findings and Conclusions

“What gets measured is what gets done.”

This section of the White Paper provides the major facts, findings, and conclusions of the Study Group in identifying, assessing, discussing, and analyzing the Exemplars. The next section of the White Paper identifies and discusses major categories of metrics that were observed in the various systems.

The major facts, findings, and conclusions are:

1. **There is no “silver bullet” category of metrics that appears to “capture it all” and relate unambiguously to an agency’s or company’s bottom-line mission.** Furthermore, there is no single, universal metrics capture and reporting tool (or product) that can be adopted for widespread use in ERM systems. In considering the exemplars, the Study Group identified eleven (11) common categories of metrics that can serve as indicators of value or goodness of an ERM system in impacting an agency’s mission and business functions. On the other hand, how the metrics can be aggregated into a single measure or viewpoint is problematic, and can depend heavily on an administrator’s point-of-view.

It is widely known that points-of-view can differ across branches of government, as well as across administrators and managers in individual agencies. Any approach to aggregate metrics to get a universal bottom-line number requires trade-offs to be made and arbitrary weighting factors to be assigned. The danger in assigning the weights is that critical faults or deficiencies in one ERM metric area can be easily masked or averaged out. Such individual faults and deficiencies may serve as valuable warnings and indicators that can help isolate and correct hidden or latent problems with the ERM system. When aggregated, the fault, error, or problem can be easily masked and/or missed.

2. **In the past, both government and industry have used a Balanced Scorecard (BSC) approach to aggregate metrics and measures for IT systems.** The BSC approach has both its detractors and proponents. The BSC approach does not appear to be as popular as it once was. See Section 6 of this White Paper for a discussion of aggregation methods.
3. **Measuring ERM performance (across federal agencies and private industry) is still a somewhat immature practice, especially as regards relating the measures to the bottom-line mission of the organization.** With most ERM systems, the metrics that are captured tend to be IT oriented rather than RM oriented. The Study Group observed that metrics

are routinely being used in the system performance area; less so relating to the overall mission of the agency or firm. This observation reflects the fact that metrics implementation tends not to be high priority in systems design and is frequently added as an afterthought. Indeed, with new systems development, the Study Group is aware of several ERM and IT systems, where business-related performance metrics can now be specified and reported as part of an overall Service-Level Management (SLM) agreement.

- 4. In assessing the Exemplars and with broad knowledge of industry and government practices resident on the Study Team, there is unanimity of opinion that no two agencies do or measure ERM alike.** At the current state-of-the-practice, valid comparisons of ERM business practices across agencies are difficult to make and probably should not be made. The general belief on the Study Team is that agency-wide standards and guidelines need to be established and adopted in this area of measuring ERM performance. Additional work is also needed in the development of automated tools and products for ERM metrics capture and reporting for both legacy and new ERM systems. Ideally, such tools can be adopted across government and provide a *de facto* vehicle for standardization and harmonization.

The good news is that there are examples of ERM systems, which automatically capture and report metrics in the eleven (11) categories identified in this study with minimal recurring burden on the agency. Some systems even facilitate the capture of metrics in a Web environment. The bad news is that many legacy ERM systems have no such automated provisions, or they don't work as desired. In one instance, the Study Group identified an ERM system, which is prone to crash if the Administrator turns on the metrics capture mechanism. The sense of the Study Group is that metrics capture in the eleven (11) categories identified need not be burdensome or onerous. Through well-conceived systems architecture and engineering, automated metrics capture and reporting can be readily incorporated into ERM systems, especially newer systems. Of course, most of the metrics that can be captured automatically by the system are only data points, and not the more-desirable outcome-based metrics.

- 5. Within an ERM system, not everything that can be measured needs to be measured nor should it be. Metrics should have a purpose for continuing improvement, and it is best to design the capture and management of metrics into a system upfront or provide for an SLM approach.** The value of the metric number should not be compared across organizations and agencies. The metrics should serve as the baseline for future ERM systems development for that specific agency. The important factor is that the ERM metrics should bear on the efficiency

and effectiveness of the agency's mission and its service to its constituents; not solely to the efficiency of the ERM system itself. Metrics need to relate to an agency's mission performance.

6. **In comparing paper-based records management systems with electronic records management systems, there are important “paper vs. electronic” paradigm issues to be understood when attempting to derive ERM metrics and to compare them “pre vs. post” deployment of the ERM system.** In the next section of this White Paper, the Study Group identified a number of metrics categories that are significantly improved by the switch from a paper-based environment to an electronic environment. Notwithstanding that observation, the Study Group generally believes that there are important paper vs. electronic paradigm issues that need to be understood in deriving worthwhile metrics including:
 - a. Paper records management is still far more mature with millennia of experience base; ERM is at best 30 – 40 years old.
 - b. Many legal issues of ERM are still unsettled as regards data integrity and preservation of at-risk digital records.
 - c. Authenticity of electronic records can be problematic and is very difficult to assure, much less measure.
 - d. The sheer definition of an electronic record is still a concern, so it is even difficult to count electronic records. (Example: relational databases)
 - e. Paper-based records are physical; electronic records are logical and can support hyperlinks as complex objects.
 - f. There are still important look-and-feel issues associated with electronic records. Quite frequently, the printed output does not look like the screen output, raising potential issues about what is the authentic record.
 - g. In a Web environment, public access to electronic records can be 24/365. In a paper environment, physical presence during business hours is typically required.

7. **“Good” vs. “Bad” Metrics.** As part of the scope of this project NARA asked the Study Group to consider lessons learned with industry and government ERM systems and to also identify any metrics that were potentially ambiguous, intractable, unreliable, or burdensome to capture. After considerable discussion among the Study Group members, a consensus emerged about certain metrics. Among these problematic metrics are record search time, record retrieval time, number of seats (or licenses), session time, and the raw number of records. While these all can be captured, the interpretation of each of these is quite controversial, even to the point where a long session time, for example, could be indicative of great success or utter failure. There is consensus across the Study Group that potential value can come with how these metrics are

considered in combination with other metrics. Taken in isolation, however, many metrics can be problematic, as discussed in the following section on Categories of Metrics and Specific Metrics.

8. **Legal and Regulatory Compliance.** Increasingly, ERM systems are being called upon to provide storage of, and access to, electronic records as an essential business function or process for the agency or organization. For example, in the financial sector, the *Sarbanes-Oxley Act of 2002* (SOX) fundamentally establishes electronic recordkeeping as a critical business process (in and of its own right) and provides a framework for defining legal and regulatory requirements in such areas as accuracy/quality, integrity, and security that enterprise ERM systems must meet. In that regard, the Study Group sought exemplars to help answer the question: “Can the degree (or ability) of an agency’s ERM system, itself, to support an agency’s business processes in the legal and regulatory arena be measured?”

In looking at the exemplars, the Study Group concluded that the legal category of metrics is perhaps the least mature category. In that regard, it is very difficult to quantitatively measure the quality or adequacy of electronic records management processes such as dispositioning. Nonetheless, the Study Group determined that there are some reasonable indicators and metrics that can be captured and reported to help measure the extent to which an ERM system is in conformance with overarching legal and regulatory requirements. Some bottom-line, ERM process-oriented performance parameters can be measured that have potential legal significance, bearing, and liability for an enterprise. However, the *caveat* is that it is very difficult to establish a threshold for “goodness” and “badness” for these metrics. Possible legal indicators and metrics are discussed in more detail in the next section.

In adjudging whether an ERM system can properly support an organization’s legal and regulatory requirements, it is also worthwhile to note that many agencies and organization are now seeking formal certification and accreditation (C&A) for their enterprise ERM systems. At the current time, the Study Group notes that formal C&A testing represents more of a one-time, checklist of acceptable and best practices for an ERM system, than a recurring metric or measure of effectiveness (MOE), *per se*.

5. Categories of Metrics and Specific Metrics

As the Study Group examined the available literature and exemplars of metrics for ERM, we found it useful to group the possible metrics into broad categories:

- Access to Services
- Accuracy
- Capacity
- Efficiency
- Participation
- Productivity
- Search and Retrieval
- System
- User Satisfaction
- Utilization
- Legal

This taxonomy provided us an opportunity to discuss specific metrics and sub-categories which in turn can be “rolled up” to provide an indication of the ERM system’s performance overall with regard to a particular category.

One of the aspects that the members of the Study Group kept in mind while examining possible metrics was the ultimate use to which these metrics might be put. The challenge from OMB (and, hence, from NARA) is to find useful metrics for responding to the Government Performance and Results Act (GPRA). Therefore, we attempted to focus on metrics that are indicative of “results” rather than just “activities.” Furthermore, the metrics that we present reflect issues of legal and regulatory compliance and good business practices as well.

Throughout the following discussion, it is critical to bear in mind that the source of the measurement and the capture frequency for any metric will inevitably increase the burden of that measure. Thus, the impact of frequent measurement will be greater for manually collected data rather than automated measurements. Similarly, the enterprise-wide scope of a measurement may make it prohibitively burdensome. While it may be reasonable to collect a metric for a particular system or process (e.g., Freedom of Information Act or FOIA requests – refer to Section 5.6 for more detail), it may not be reasonable to do so for all requests for all records across an entire department of the federal government. Finally, collecting metrics from manual processes and legacy systems is likely to be far more burdensome than collecting metrics from systems which have been designed with the need for these metrics as part of the design criteria.

The study group examined exemplars from both the public and private sector, including:

- Allina Health System, Minneapolis, MN, Ambulatory Electronic Records Implementation Cost Benefit: An Enterprise Case Study
- COPS Office, Electronic Records Management Study
- Department of Defense, Office of Inspector General, Electronic Records Management System Process Analysis
- Department of Energy, Office of Records and Business Management, Office of CIO, Evaluation Report on Electronic Records Management Email Pilot Project
- Environmental Protection Agency, Draft Functional Requirements for Electronic Records Management
- Food Conglomerate, Electronic Records Management System
- General Accounting Office Best Practices Study for Electronic Records Management
- Nuclear Regulatory Commission, Agency-wide Document Access and Management System (ADAMS)
- Office of the Comptroller of the Currency Records Management Program
- Pharmaceutical Company, Electronic Records Management System
- NARA Access to Archival Databases (AAD) Project

Each of the metrics discussed below is in fact collected in at least one system the Study Group examined. However, as noted previously, much of the evidence is anecdotal and/or proprietary, and therefore we are unable to provide a specific reference to where it is being used (or, in fact, if it is being used as opposed to being merely collected!) No one exemplar had metrics in all of the categories below.

In addition, the members of the study group noted that many of the metrics relate more to the computer industry than to records management *per se*—and those tend to be the easiest metrics to collect, though not necessarily the most meaningful in the context of this study. The Study Group was particularly interested in finding metrics that have an analog in paper-based systems, in order to facilitate a “before-and-after” comparison. With the increasing shift from paper-based records to electronic records, the Study Group was also interested in identifying any recurring metrics that could serve as indicators for adjudging the bottom-line efficacy of ERM systems in a legal and regulatory framework.

5.1 Access to ERM Services

With the emergence of the Web, one of the great benefits of an ERM system may be the wide access it can provide to the records of an agency. This is true for all user communities, both those internal to an agency and the external users, often the general public.

For internal users, the hours of access to the system (for declaring and classifying, as well as retrieving documents for internal use) are probably quite similar between a paper-based pre-automation system and an automated ERM system. Generally speaking, such internal use of the system occurs on the agency premises and during normal working hours, although certain functions (such as declaring an email message to be a record) may well take place off-site and off-hours. In fact, there is no reason why access to an automated system may not occur anywhere at any time, other than agency policy and the generally limited security and access tools in use in agencies for workforce mobility.

The situation is quite different for the general public. Automated ERM systems can allow access to public documents from virtually any computer at any location at any time of day or night. They also allow for self-service, which is becoming a more and more essential part of e-government efforts aimed at service to the citizen. This is a significant improvement over previous records management access capabilities, where one frequently had to go to an agency's reading room and request that a record be brought out for review.

Metrics can be rather easily derived which reflect the hours of operation (for internal and external users) and access points to the ERM systems—and will almost always show a great increase in both after implementation of an automated system for e-records. The increased access to records services can easily be several orders of magnitude for “virtual visitors.”

5.2 Accuracy

Accuracy, like the Participation category discussed in Section 5.5, is an important factor in determining the success of an ERM system. It requires an examination of records to determine whether those that should have been declared were, and whether those that were declared should have been. One should also measure the accuracy of the classification of those records. Another aspect of measuring accuracy, that has potential legal significance (See Section 5.11), also relates to the accuracy and quality of the processes that are used by the agency to accomplish its ERM mission requirements.

The degree to which record classifications have been completely defined can also lend itself as a metric when indirectly observed through the number of re-classifications that occur as a result of misclassification. An ERM system with an adequate number of defined classifications will allow a greater degree of accurate assignment whereas a system with too few or too many classification options will create a greater number of misclassifications, owing to confusion and potential ambiguity in distinguishing between the classifications.

Like participation, accuracy is not easily measured automatically. In general, live oversight and periodic audits are necessary to determine whether the declarations and classifications have been done correctly.

5.3 Capacity

One of the more commonly measured aspects of an ERM system, in particular, (as opposed to automated systems, in general) is the number of records stored and/or the amount of storage—in other words, the size of the holdings. Many companies further break down the size of their holdings by record type (created by Word 2000, for example), or by collection. These metrics are kept automatically by most systems in existence today. In fact, in some systems, one can report the number of records by any metadata field. So, these are easy metrics to capture in an ERM system, and they offer some indication of how much the system is being used to manage the records of an organization.

Counting the raw number of records retrieved by an ERM system or contained within an ERM system is not a good metric in and of itself. Taken in isolation, the general tendency is to think that “more is better” from an electronic records management perspective. This is not necessarily the case. The user may be after a single “smoking gun” record. Then again, the user may be after a large collection of records. The user may not even know that the record exists within the system. At that point, the user would like some measure of assurance that the record (or records) does not exist, in which case returning a zero (0) results set is the right answer. Thus, counting the raw number of records returned can be highly misleading.

One factor to be considered in “counting the raw number of records” as a metric in an ERM system has to do with potential user interest that might be manifested in the number of virtual visitors accessing the system. In general, the richer the content in an ERM system, the more the public will have an interest in using the system. The problem is that “richness of content” is determined in part by the uniqueness of the records material as well as the volume of records material in the system. It is axiomatic that a popular records set, such as a genealogical records set with lots of records, will be more popular than a site with an arcane records collection. Then again, a site with a small but unique and popular collection will likely be of major interest to an agency’s bottom-line mission of serving the public. Thus, the raw number of records is not necessarily a good metric

Additional information, such as the total number of records (inside and outside the system) or the expected size of the holdings is necessary to determine whether the system is being used to manage all or most of the target records. It should also be noted that bigger is not always better, as it is just as harmful to “over declare” records as it is to fail to declare records. Size of holdings,

especially when used in conjunction with the accuracy of declaration and/or an assessment if the records are properly dispositioned (See Legal discussion in Section 5.11), can be a useful measure of an ERM system.

Another storage measure is cost per gigabyte of storage. With this metric, one can determine how efficiently one is storing the holdings in an ERM system—and can certainly compare the storage costs of an ERM system with the paper-based storage mechanisms it replaces. It is to be expected that this type of metric will demonstrate great savings in storage costs, especially as the costs of disk and tape storage continue to decrease. In general, this metric is fairly easily calculated and poses little burden on an agency to collect.

One reviewer of this document noted that with regard to computing and reporting cost per gigabyte of storage, certain media have a limitation in that they can not intrinsically support “event-based records retention policies”. In that regard, to the extent that an organization’s mission and essential business functions mandate an event-based electronic records retention policy, any metric which calculates the cost per gigabyte of storage for certain media needs to properly factor this limitation into the bottom-line, reported metric result.

5.4 Efficiency

The efficiency metrics attempt to measure how well an organizational entity or individual can perform an ERM-related, day-to-day task. Factors such as work environment, policies and procedures, user training, computing resources, and information dissemination will determine the degree of efficiency that can be achieved and measured. However, the measurement of these factors is largely subjective. Certainly, an organization should have policies and procedures in place for the administration of records, but the degree of completeness tends to be subjective. Another potential measure is the degree and frequency of user training, but here again this tends to be subjective based on the effectiveness of the training for the individual’s job functions.

The metric, session time, is similar to the metric, system search time discussed in Section 5.7, in being ambiguous and intractable. In that regard, motivation and the user experience are major contributing factors to the amount of time that a user will spend in a session using an ERM system. In addition, there are other external factors, such as time-of-day and competing demands (on users) that also contribute to the amount of time that a user will spend in a session. While it is fair to say that users, in general, would like to spend their session times efficiently, the mere measurement of their session time taken in isolation, does not adequately provide an indicator of efficiency. Then again, training on the part of the user as well as knowledge of the records domain may be important factors. In measuring public and consumer “efficiency”, which is surely of interest as part

of the e-government initiative, the contributing factors and parameters need to be factored and considered.

5.5 Participation

The participation metrics address the use that is being made of the system by the owner of the system (as distinct from customers, e.g., the public). The owner would be anyone who declares, classifies, and manages records within the ERM system, and not those who simply refer to, or make use of, those records.

Participation is often measured by the number of seats/licenses although this is misleading. Counting the number of seats/licenses is not necessarily a good metric. While this metric could be an indicator for “participation” (see the next section of this White Paper), in isolation the metric could be intractable to interpret. In general, this metric is more applicable to thick-client ERM applications where software needs to be installed on an individual’s PC or workstations. The process for installation and operations and maintenance of such applications can be high. In contrast, a thin-client application can be readily made available and accessible to users via a Web-based client-server application. The Study Group does not wish to convey that thin-client applications are always better than thick-client applications. The point is that the metric of counting the number of seats/licenses is subject to interpretation and potentially intractable without defining the overall ERM system architecture.

Rather than counting the number of seats/licenses, the level of participation in an ERM system is crucial and is a critical measure of its success, although it is less often and less easily measured than many of the other metrics discussed in this section. The issue is how many people are declaring records vs. how many should be declaring them? (The accuracy metric, discussed above, is closely linked to participation and mere participation, without accurate declaration and classification, is not a good measure of success.) If the level of participation is low, then the records may not be properly managed, as there is a high risk of undeclared records or records with poor integrity.

There may be many reasons for a lack of participation in an ERM system. It may be an issue of training, of course, but it may also be an issue of burden and ease-of-use. One of the specific issues related to burden identified by the study group is that of thin-client vs. thick-client applications. In other words, if the application is a “thick client” requiring a substantial amount of code to be installed on a user’s workstation to facilitate his/her participation in declaring and/or classifying records, then barriers to participation may simply be too high.

Participation is measured primarily through periodic live oversight or through audits. It should be looked at as often as once a month in the initial months after implementation of a system, to identify where additional training may be needed.

Once the ERM system has been in place for several months, and the participation levels are acceptably high, measurements need only be taken annually—depending, of course, on the results of the reviews and the severity of impact of a failure.

5.6 Productivity

In contrast to efficiency described in Section 5.4, productivity metrics attempt to quantify the value of combined technical and organizational efficiencies realized by organizational entities or individuals in performing ERM related tasks. They, therefore, relate to the business processes, which are supported by the ERM system. A common business process example for ERM systems is the number of invoices per hour that a clerk can process, which is often compared to the same measure before implementation of ERM. Another example is the productivity of a clerk who needs to visit the archives to search for, and retrieve paper-based records. Which business processes should be measured will vary from organization to organization, and the criterion for selection of the processes to measure should be their importance to the organization. In other words, they should reflect the critical lines of business, and should be processes that are important to those lines of business.

The productivity improvements resulting from implementation of ERM systems are primarily achieved through improved workflow. This was reported anecdotally, with several companies commenting that the greatest impact of such implementation was the rationalization and improvement of their workflow. This was, however, reported anecdotally, and not measured in a clearly quantified or standardized way. The quantification of such benefits must be done through the system, examining the time level of effort: how many records were declared, how many scheduled, in a given increment of time.

Within the Federal Government, an important and ubiquitous ERM function is that of processing Freedom of Information Act (FOIA) requests. Since these requests involve searching for relevant records, they are very dependent on the records management practices of each agency. A metric of the number of FOIA requests processed in a day (or its converse, the average time to process a FOIA request) is a measure of productivity that reflects the effectiveness of records management. It also enables a comparison of before-and-after implementation of ERM, as agencies are required to keep track of the FOIA requests they process. Although this measures only one of many business processes, its universal nature may make it a good (though not necessarily an exhaustive) indicator of the value of ERM throughout the Federal Government.

Utilization of resources and their rate of change is an important and valuable productivity metric that can assist with tracking the efficacy of the ERM. Any increase or decrease in the productivity of staff workers can be attributed to a

change in the system environment or to the volume of record processing. Such a metric would usually be an important indicator in any organization and can alert upper management to changing conditions.

It is important to note that any measure of productivity is not especially meaningful without reference to that same measure at another point in time. In this case, the most critical aspect is the productivity “pre vs. post” implementation of ERM. Another measure that could be of great interest would be a comparison of the rate of service for those records that remain outside the electronic system vs. electronic records remaining within the system.

Another productivity metric, which is of great interest in the corporate world, is turnaround time, or time to market. In the e-government context, this might translate into time to service a request. As with the FOIA requests above, this type of metric may be readily captured and depends on the type of request involved.

Productivity metrics, except for the raw counts of records processed, are usually not automatically captured within a system, but need to be calculated, and may be more or less difficult to capture depending on whether the productivity of the business process is measured in some way during the normal course of business. A company knows how many invoices it issues in a month, which makes it easy to calculate the rate of invoicing; it may be more difficult to determine the rate at which it provides another, less essential service.

5.7 Search and Retrieval

One of the metrics that seems quite obviously useful in judging an ERM system is the number of successful records searches. This can be based on a combination of factors, including system search time, system retrieval time, quality of the metadata, and even the popularity of the search domain. Upon discussion, however, it appears that it is quite difficult to interpret the results as it relates to the success of the search and retrieval function.

System search time is highly variable and also depends in large measure on the degree of sophistication and knowledge of the records domain on the part of the user. Additionally, the amount of time a person will spend on search depends in large measure on motivation. In that regard, a shorter system search time does not necessarily equate to better performance. User persistence, with allocation of more time and effort, may in fact provide for better results.

Also, in searching for records, there are two fundamental ways of searching, namely: goal-driven and data-driven searching. In the goal-driven approach, the user has a concept or idea of what they are looking for. It is a matter of finding the record(s) and proving its/their existence. In a data-driven search, the user

does not necessarily know that the records exist. This is a discovery mode of operation, analogous to data mining. In practice, most researchers switch back and forth seamlessly and iteratively between goal-driven and data-driven modes of searching. This phenomenon makes any simple metric of search time within an ERM system highly variable and intractable to interpret.

System retrieval time is similarly highly variable and intractable to interpret without a lot of other ancillary information. In most modern ERM systems, it is axiomatic that the hard part is finding and searching for the record(s). If the record(s) can be found, generally it tends to be a relatively simple matter to retrieve the records, especially if they are on-line. Then again, if only the metadata and indices are on-line, physical retrieval of the record(s) as paper-based copy, audio-video materials, etc. will be much slower, and not necessarily a function of the ERM system.

System retrieval time is also highly variable for a variety of technical reasons outside of the control of the agency or system. Specifically, in a Web environment, user connection speed is a factor as well as overall Internet network latencies. Also, the search results set may be rather large, or the individual files, voluminous. Variable transport time for the records would then be an issue. Time-of-day and user demand on the system could also be a factor in system retrieval time. Lastly, on the Internet, Web sites may be subject to a "flash crowd" phenomenon, such as occurs with news and media events where a Web site may be inundated. Under such high stress conditions that can exceed the system design parameters, system retrieval time can readily grow by orders of magnitude, or the system itself may in fact fail catastrophically.

The ability to retrieve records is also directly tied to system architecture and performance. System performance metrics such as retrieval speed, as mentioned above, fall in this category.

Of more importance is the quality of a search. The quality of a search is predicated upon the choice of metadata used for indexing the record. If the right indexes are present, then a user, either internal or external, will have no difficulty finding a record of interest. However, if insufficient, inappropriate, or incomplete indexes were assigned, an ambiguous or excessive or incomplete result set may be returned confounding the user and reducing the usefulness of the ERM system. Metrics for gauging the effectiveness of searches can be addressed by the quality of the metadata. Roughly, the number of search indexes and the number of classification categories for records can provide a measure of simplicity vs. complexity and thus ease vs. difficulty in formulating search criteria.

In assessing the utility of metrics focused on search and retrieval, another phenomenon to consider is the popularity of the domain. The experience of Web surfers is that they surf and search all day, because of their interest in the media and the content they are retrieving. Metrics in this area are very difficult to define

and are virtually intractable to interpret. Persistence of the researcher is also a related factor. In that regard, some researchers will search all day to try to find a “golden nugget” or a single “smoking gun” record. Of course, they may not know whether the record is in the system or not. They may not even have a semblance of what they are looking for, and they are in a discovery mode of operation. Nonetheless, they persist in their search endeavors, and any metric relating to search time would surely be misleading.

5.8 System

Compared to other categories of metrics, system metrics which measure the performance etc. of the automated system itself are quite readily available and easy to capture. We found that they are commonly reported metrics for ERM systems. Although they are quite valuable in the operation and tuning of the system and quite easy to understand and to graph, system metrics are not necessarily good measures of the value or success of the management of electronic records. In fact, as one member of the study group commented, they are more valuable for managing the performance of the contractor who may be providing and/or managing the system than they are for determining the value of the system.

Several of these system metrics might be called Performance and Stress Test metrics. These metrics are quite useful in reporting the performance of the system, and are almost universally used in stress testing, or testing the system under very high loads. The throughput of the system is commonly measured; this is the number of transactions that a system can perform in a unit of time. For an ERM system, the transactions might be record declarations, record classifications, searches, and/or retrievals. This is a valuable metric for determining whether the system is performing in accordance with its specifications, or its Service Level Agreements (SLAs), if any, but is more a measure of its size than of its success or usefulness.

Another Performance and Stress Test metric is response time. This is commonly measured for queries of a records management application, and is a measure of how long it takes to receive a reply to one’s query. As discussed above for throughput, it is commonly measured, can be measured with automated tools, and is more useful in determining whether the system is performing properly than in determining whether it is successful or useful.

Similarly, Time to First Byte and/or Time To Last Byte (TTFB/TTLB) are fairly common metrics of Web-based systems (whether records applications or not). They are measures of the performance of the system, and will vary with the load on the system and, especially for TTLB, with the load on the network. These metrics are relatively easy to capture with automated tools, and are quite valuable in discovering bottlenecks and tuning the system. As with all of the

Performance and Stress Test category metrics, these are less meaningful in measuring the success and usefulness of the ERM system. On the other hand, if response time deteriorates at high user demand due to system scalability limitations, users may become frustrated and abandon the system. At a minimum, they will have a poor user experience that might reflect on the agency's bottom-line mission. An example of this situation occurred with the Public Records Office of the UK National Archives when they first put the records of the 1901 census online. Demand was so high that performance suffered. As a result, the UK National Archives census application was down for over eight (8) months.

Another sub-category of metrics is Availability. This is a measure of the percent of time that the system is up and operational. It is a System metric, often referred to as "uptime," and is usually measured for any Information Technology system. As a system metric, it is quite useful for telling us something about the system itself, but less useful in determining its value to the agency or its users; unless of course, the image of the agency is tarnished.

Since systems are generally sized to accommodate a specific volume of processing, overrunning the suspected volume could significantly reduce productivity as computing resource demands are overloaded and system performance is degraded. On the flip side, under-running volumes could reduce productivity as excessive resources are applied to the ERM tasks at hand with workloads diminished and below those normally experienced by staff workers. Again, this metric would be best suited as an indicator to upper management.

Operations and Maintenance (O&M) costs is another System metric that can be calculated for ERM systems. These are ongoing costs, as opposed to the one-time costs for design, development, and implementation. These O&M costs can be further broken down into the costs for ingestion, indexing, systems maintenance, and user training. This metric is an excellent indicator of the steady state (as opposed to startup) costs for an ERM system, and are not especially burdensome to capture.

It is quite common within the private sector to measure the return on investment (ROI) of a system, or to do a cost/benefit analysis. The O&M costs of the system are a necessary input to this sort of metric, which also require the quantification of the benefits (and cost avoidance) of the ERM implementation.

5.9 User Satisfaction

User satisfaction is one of the most commonly measured aspects of ERM systems. It is a nearly universal metric in the corporate systems we studied. It is, unfortunately, not a metric that is frequently captured automatically by the system and it is, of course, a "softer" metric than many.

User satisfaction is normally captured via surveys performed periodically. As noted above in the discussion on major findings, there are some issues with the use of surveys, particularly in the area of the burden of collection and the subjective nature of the results. Even so, the satisfaction of the users of a system is an important measure of the success of that system, and is well worth considering.

Because there are different categories of users of every ERM system, including both internal and external users, multiple surveys may need to be created. In fact, in the case of one of the study group exemplars (a health care group), the satisfaction of administrative staff was measured separately from the satisfaction of the patients and the physicians. Some of the satisfaction issues that should be addressed are ease of use, quality of searching, and speed of service.

After due discussion and deliberations, the Study Group has concluded that user surveys are a useful category of metrics for assessing the relationship of an ERM system to an agency's bottom-line business process and mission. Indeed, in evaluating this category with the exemplars, user surveys were the most prevalent category of metrics observed. At a high level, the following factors need to be considered in adopting a user survey metrics schema:

- a. There is a fundamental need to define the customers (both internal and external) for an organization that will use a survey methodology for its metrics. Properly executed, a survey can be very useful to derive the baseline for setting expectations of the ERM system.
- b. User Surveys tend to be highly subjective and heavily depend on who is the customer that is being surveyed. Managers, developers, and integrators need to be cognizant of what is being measured and only attach an appropriate level of significance to the results.
- c. Hard numbers don't always produce the best results from the survey. Interpretation and follow-up is frequently required.
- d. Conducting and analyzing the results of a user satisfaction survey tends to be somewhat burdensome for the agency. This situation is particularly the case where a paper-based survey is used. On the other hand, the Study Group identified several ERM systems where an on-line user survey can be conducted. For several Web-enabled systems at NARA, such an approach is now accomplished via random popup (9 – 20%) of a Java application. The survey data is then collected automatically. Such an approach is also widely used on many e-commerce sites to provide consumer feedback to the electronic store owners.

It is the opinion of the Study Group that the most effective way to implement a user satisfaction metric for ERM systems within the Federal Government would be to create short survey templates which can easily and quickly be customized to address the specific environment within each agency. Guidance also needs to be provided regarding the frequency of surveys, the number of users surveyed,

and the length of the survey questionnaires, to manage the burden associated with this method of collecting metrics. The following figure is an example of a customer satisfaction survey for a public electronic records access system at NARA that “pops up” on a random basis as a user ends a session. The responses to the survey questionnaire are captured automatically and rolled up to provide aggregate results.

ForeSee Results Survey - Microsoft Internet Explorer

NARA **Customer Satisfaction Survey**

Thank you for visiting our site. You have been randomly selected to take part in this survey to let us know what we are doing well and where we need to do better. Please take a minute or two to give us your opinions. The feedback you provide will help us enhance our site and serve you better in the future. All responses are strictly confidential.

1: Please rate the accuracy of information on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

2: Please rate the quality of information on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

3: Please rate the usefulness of the services provided on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

4: Please rate the convenience of the services on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

5: Please rate the ease of navigation on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

6: Please rate the speed of loading the page on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

7: Please rate the reliability of site performance on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

8: Please rate the ability to limit sharing of your personal information on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

9: Please rate the organization of search results on this site.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

10: Please rate how the search feature helps you to narrow the results to find the information you want.
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10 Don't Know

11: What is your overall satisfaction with this site?
 1=Poor 10=Excellent
 1 2 3 4 5 6 7 8 9 10

5.10 Utilization

Unlike Participation above, the metrics within the Utilization category deal with the use being made of the system by its customers—those who refer to records, not those who declare and/or classify them. There are two categories of customers – internal to the organization and external to the organization (e.g. public). For internal users, an audit trail can be used to monitor the individual’s activity or use of the system and for external users or “virtual users”, it is possible to track the remote access usage of the system. However, this tends to be more a system performance metric as opposed to a true measurement of utilization because the intent of the access may not be known.

5.11 Legal

As noted in Section 4, ERM systems are increasingly being called upon to provide storage of, and access to, electronic records as an essential business function or process for the agency or organization. The Sarbanes-Oxley Act of 2002 (SOX), for example, establishes electronic recordkeeping as a critical business process. SOX also provides a framework for defining legal and regulatory requirements in such areas as accuracy/quality, integrity, and security that enterprise ERM systems must meet. Thus, capturing metrics in a number of metrics categories as identified above can be important if an organization’s ERM system is ever challenged in a court of law or regulatory proceedings.

With regard to the potential use of metrics in a legal and regulatory environment, one peer reviewer for this white paper noted the frequent “...*need to prove the negative in management evidence – the evidence by which the process of managing records is documented.*” The reviewer also noted: “*Increasingly, regulatory agencies demand the ability to prove the negative – that you did not do it wrong... Proving the negative means generating much more management evidence information.*” The reviewer also suggested that performance-oriented metrics for ERM systems could be valuable in providing the necessary “management evidence information” to help sustain an organization’s position in either a legal or regulatory action. Lastly, the reviewer raised the specter of performance-oriented metrics relating to legal and regulatory compliance thus “*assuring that compliance was achieved.*” The reviewer also pointed out “...*the need of organizations to have established policies and procedures in this regard.*”

In looking at the exemplars, the Study Group concludes that the legal category of metrics is the least mature category of metrics possible in today’s ERM systems. Indeed, it is very difficult to quantitatively measure the quality or adequacy of electronic records management processes such as dispositioning. Nonetheless, the Study Group determined that there are some reasonable indicators and metrics that can be captured and reported to help measure the extent to which

an ERM system is in conformance with overarching legal and regulatory requirements. Some bottom-line, ERM process-oriented performance parameters can be measured that have potential legal significance, bearing, and liability for an enterprise. However, the *caveat* is that it is very difficult to establish a threshold for “goodness” and “badness” for these metrics much in the same sense that a jury is called upon to weigh the evidence.

In adjudging whether an ERM system can properly support an organization’s legal and regulatory requirements, it is also worthwhile to note that many agencies and organization are now seeking formal certification and accreditation (C&A) for their enterprise ERM systems. ERM systems with their embedded information and process flows can be certified and accredited to meet defined and established policy. At the current time, the Study Group notes that formal C&A testing represents more of a one-time, checklist of acceptable and best practices for an ERM system, than a recurring metric or measure of effectiveness (MOE), *per se*.

On the other hand, modern ERM systems can have their policies, procedures, and processes embedded in software as a workflow mechanism. The workflow mechanism in such systems can be monitored and observed; and statistics that relate to ERM process accuracy and integrity can readily be derived. For example, counting and reporting any violations of policies and procedures can provide both qualitative and quantitative indicators of the accuracy of the RM/EM processes by which records are managed. Such a record of the numbers and types of violations that are caught, missed, and/or are attempted could have legal weight, significance, and bearing on an organization’s bottom-line records management business functions; as well as the overall security of the system. In general, assessing the accuracy of processes and procedures requires constant surveillance and monitoring of the system. Some aspects of counting and reporting on violations of policies and procedures can be automated via intrusion detection technologies and some specialized software now being developed and marketed to ensure Sarbanes-Oxley (SOX) compliance.

Another potential approach to assessing the bottom-line effectiveness of an ERM system to meet intrinsic legal and regulatory requirements is to look at the electronic records within the system from a “materials management” perspective. One reviewer drew an analogy with the legal liabilities of food management in a freezer noting that *“If the food is spoiled or out-of-date, you need to get ‘rid of it’.”* Serving spoiled food to a customer can certainly have very significant legal liabilities, and most restaurant owners have a very active materials management policy as regards perishable foods and sanitary standards. It is just too easy to be sued or shut down by the health authorities. By the same token, electronic records can also be perishable and can pose legal liabilities for an organization if the records are not properly managed and dispositioned in a timely and responsible manner. Indeed, the entire ERM system, together with senior

management, can be held liable if critical processes and standards for ERM are violated or for that matter, neglected. This is the essence of Sarbanes-Oxley.

Thus, the materials management analogy for food suggests a potential performance-oriented metric for electronic records management in the legal and regulatory realm. By analogy to “food in a freezer”, at any given point in time, the records within an ERM system are in various states and/or they may be queued for disposition. The records may be: active, dormant, processible, draft, searchable and accessible to various classes of users, signed, scheduled for destruction, etc. From a legal and regulatory perspective, the real issue is what is the state of the records in the inventory? And, what ought to be the state of the records? An ERM system that has an appreciable amount of its inventory of records queued for destruction (with *prima facie* evidence of neglect or inaction in actually doing so) could pose undo bottom-line legal risk for the organization.

Similar to the materials management problem confronting a restaurant, the Study Group suggests a possible metric to measure the fraction of the inventory of electronic records within an ERM system that is in the wrong state. At the current state-of-the-practice, this type of metric can be captured via reviews and audits with the potential for future automation as tools in response to Sarbanes-Oxley are improved and refined.

Sample Candidate Metrics for ERM Systems

Measurement Category	Metric	Capture Method	Capture Medium	Capture Burden	Comments
Access to Services	Hours of operation	Manual	Periodic audit	Low	Almost certainly greatly improved with automation
	Access Points	Automated	System	Low	Almost certainly greatly improved with automation
Accuracy	Percentage of records correctly declared	Manual	Periodic audit	High	Measure of quality
	Percentage of records correctly classified	Manual	Periodic audit	High	Measure of quality
Capacity	Size of holdings, i.e., number of records (possibly by record type)	Automated	System	Low	No indication of quality
Efficiency	Ease of performing daily tasks	Manual	Survey	High	Purely subjective but indicative of success and acceptance of ERM
Participation	Number of seats	Automated	System	Low	No indication of quality
	Number of people declaring records	Manual	Live Oversight	Medium	Indicative of acceptance of system
	Number of people classifying records	Manual	Live Oversight	Medium	Indicative of acceptance of system
	Number of people retrieving records	Manual	Live Oversight	Medium	Indicative of acceptance of system

Measurement Category	Metric	Capture Method	Capture Medium	Capture Burden	Comments
Productivity	Number of requests processed per week	Automated	System	Low for one system; High across enterprise	Difficult to measure enterprise-wide across multiple processes; may only be useful as a sampling metric, e.g., for FOIA requests only
Search and Retrieval	System search time	Automated	System	Low	No indication of quality
	System retrieval time	Automated	System	Low	No indication of quality
	Number of successful searches	Automated	System	Low	Difficult to interpret; returned result is not necessarily desired result
	Number of search indexes	Automated	System	Low	Indicator of complexity and therefore ease of use
	Number of classification categories	Automated	System	Low	Indicator of complexity and therefore ease of use
System	Throughput, i.e., transactions per hour or per unit of time	Automated	System	Low	Measures IT performance, not success of ERM
	Response time, i.e., time to retrieve a record	Automated	System	Low	Measures IT performance, not success of ERM
	Availability, i.e., system uptime	Automated	System	Low	Measures IT performance, not success of ERM

Measurement Category	Metric	Capture Method	Capture Medium	Capture Burden	Comments
User Satisfaction	User satisfaction rating	Manual	Survey	High	Nearly universal metric for ERM exemplars
Utilization	Number of people retrieving records	Automated	System	Low	Indicative of acceptance of system; no indication of success or satisfaction
	Virtual Visitors	Automated	System	Low	Indicative of acceptance of system; no indication of success or satisfaction
Legal	Numbers and types of process violations that are caught, missed, and/or are attempted	Semi-Automatic	System	Medium	Measure of accuracy and quality of the ERM processes with potential legal weight, significance, and bearing
	Fraction of the inventory of electronic records within an ERM system that is in the wrong state	Semi-Automatic	System	Medium-High	Indicative of the quality of the processes and services provided within an ERM system

Note: Any of these metrics should be used to measure improvement over time relative to a baseline. The numbers are not meaningful in and of themselves. Additionally, the Study Group determined that there is no universal, “silver bullet” metric.

6. Aggregating Metrics and the Balanced Scorecard Approach

6.1 Introduction

As a major finding and conclusion, the Study Team observed that there is no universal “silver bullet” metric or measure that adequately captures or measures the success of an ERM system on an agency’s bottom line mission. Another major finding is that there is no single, universal metrics capture and reporting tool (or product) that can be adopted for widespread use in ERM systems. Accordingly, using a family of metrics is currently a recommended industry “best practice.” This section of the White Paper briefly explores the use of the balanced scorecard approach as a potential vehicle for agencies to aggregate and roll-up metrics.

The Study Group concluded that performance measurements and quantifiable indicators are essential for the development, execution, and monitoring of the desired ERM strategy. Metric systems will be fundamentally different depending on the strategic position they tend to support. No one metric category can be seen as the only factor in the determination of the success of a program, but needs to be put in perspective with other metrics categories. The question is how or even whether to aggregate metrics results.

6.2 Roll-up Approaches

6.2.1 Return-on-Investment

To date, industry performance metrics have focused mostly on how a company can tie traditional measures of financial performance to strategic goals. Companies perform investment planning by identifying the projects to be considered in their portfolio and the full consequences of alternative projects. Values are assigned to various inputs and outputs, and cost benefits are then typically added up to estimate profits of the project. This is an example of a brand of roll-up or aggregation of component costs and benefits to achieve a projected return-on-investment (ROI). Such investment analysis is common practice across industry, and has been applied at some level within government circles.

6.2.2 Key Performance Indicators

In contrast, a good example of a metrics-focused approach is the use of key performance indicators (KPI) scorecards. In developing KPIs, a user or developer defines target performance levels for a number or family of metrics and then decides the best way to represent variance from each of the metrics targets. Scoring for these KPIs is limited to the metrics themselves and is not rolled into a measure or score indicating actual performance in relation to strategy. KPIs allow

business executives to see the health of their organization at a glance, similar to a doctor individually checking out each of your vital organs. While taking a person's temperature might be an indicator that a person is ill, more detailed and specific diagnostics are generally required. By keeping individual metrics discrete in KPIs, the manager can check out all of the key performance indicators and take the necessary actions.

6.2.3 Balanced Scorecards

As an extension of the KPI concept, the Balanced Scorecard (BSC) approach was developed to provide for a broader methodology for aggregation of metrics. The Balance Scorecard was originally designed by Robert Kaplan and David Norton (1992, 1993, and 1996) for private sector businesses. BSC is a popular strategy-focused management tool that allows companies to drive their business performance based on measurement and follow-up. Measures are divided into four distinct perspectives: financial, customer orientation/satisfaction, internal processes, and employee learning. To succeed, organizations need to supplement traditional financial measures like return on investment (ROI) with the three other perspectives, namely: customer orientation, internal processes, and the ability to innovate and retain employees. The BSC concept has evolved from a measurement system to a core management system with a process to implement and obtain overall feedback about strategy.

Overall, BSCs can provide an approach to roll-up and aggregate metrics to give the "big picture". However, BSCs fundamentally require the assignment of weights to the various constituent measures to achieve the required "balance". To the extent that the weighting factors are arbitrary, the BSC approach has its detractors. To the extent that a manager is satisfied with the overall aggregation of metrics and the "balance" that is achieved, the BSC approach also has its proponents.

In implementing a formal BSC approach, the following measurement perspectives need to be considered:

- **Employees:** User satisfaction surveys are conducted to measure attitudes towards implementation of system, project, or application. How do employees continue to learn and grow? In the IT arena, e-businesses consider and weigh staff retention in critical business applications development and network operations as important. The Balanced Scorecard goal is to align everyone within an organization so that all employees understand how what they do supports the strategy. In theory, productivity goes up as employees "buy into" the metrics. Employee job satisfaction is likely to transfer to customer/client satisfaction. This perspective on the Balanced Scorecard relates most to the "Participation" metrics category identified in this White Paper.

- **Customers:** Customers are the recipients of a product or service. In measuring customer satisfaction in a BSC environment in industry, the following questions are relevant. How do you decide whom to please? How do we become our target customers' most valued supplier? How do we gain customer loyalty and increase market share? Do we determine value in terms of ability to meet customers' requirements at the point of need? This perspective on the Balanced Scorecard relates most to the "User Survey" metrics category in this White Paper.
- **Financial:** In private industry, financial perspectives, especially profit, tend to be paramount. In contrast, public sector perspectives tend to view financial issues as constraints and put the emphasis on efficient and effective accomplishment of agency mission and goals. Setting clear priorities and leveraging resources to meet objectives are common drivers in deriving financial oriented metrics in a BSC approach. On the assumption that "Time = Money", this perspective on the Balanced Scorecard relates most to the "Efficiency" metrics category in this White Paper.
- **Integration/Operation Excellence.** Within the BSC framework, one of the goals of effective performance metrics is to integrate organizational goals and act as an umbrella for a variety of often disconnected programs such as quality, re-engineering, process re-design, and customer service. The BSC approach recognizes that successful performance systems have direct linkage to organizational assessment and compensation processes. They are decentralized and have a formal integrated process to "roll up" local data into the organization performance measurement system. In that regard, BSC approaches are consistent with this White Paper in attempting to relate metrics to the Agency's bottom-line. Operation excellence considers what internal processes [both short and long term] must be optimized to achieve financial and customer objectives. It ensures that quality is the fundamental component of each and every part of an organization's mission and objectives.

6.3 Metrics System Evaluation – Achieving ERM Excellence

Management teams need a performance measurement system that best links performance metrics to electronic records management (ERM). Within this section of the White Paper, the BSC and KPIs have been discussed as possible approaches to considering and aggregating metrics for modern ERM systems.

The Study Group concludes that once a performance measurement system is adopted and implemented, the real task at hand becomes how to integrate the system throughout the functional and business-units of the organization.

In addition, measures and goals of the system must be continually reviewed, revised, and raised to keep the ERM system current and effective. In short, managers must continually ask, “Is the ERM system doing what it is supposed to do?”

Once an infrastructure for performance metrics and monitoring is in place, an organization feedback loop from metrics to the strategic plan can be used to identify areas and functions that need continuing improvement. For individual agencies planning on embarking on a program for measurement of ERM practices, the Study Group suggests that the evaluation effort consider investigating the tools and technology that leading companies are currently using.