# SCIENCE AND TECHNOLOGY TEXT MINING: <u>MANAGEMENT DECISION AIDS</u>

BY

DR. RONALD N. KOSTOFF
OFFICE OF NAVAL RESEARCH
800 N. QUINCY ST.
ARLINGTON, VA 22217
PHONE: 703-696-4198
FAX: 703-696-4274
INTERNET: kostofr@onr.navy.mil

(*The views in this paper are solely those of the author and do not necessarily represent the views of the Office of Naval Research*)

## ABSTRACT

This report addresses issues of quality in the context of management decision aids. It proposes rules for high quality management support procedures using these aids, and presents criteria for more effective implementation of these decision aids into the strategic management of science and technology. Two illustrative examples of quality in the context of management decision aids are presented. The first addresses quality in the context of science and technology roadmaps, and the second addresses quality in the context of information retrieval for science and technology text mining. Finally, the report discusses the major barriers to implementation and integration of these decision aids into the strategic science and technology management process:

- Techniques treated as add-ons
- Techniques treated independently
- Mismatch between performers and users

**KEYWORDS**: Decision aids; information storage; information processing; information retrieval; roadmaps; metrics; peer review; data mining; text mining; bibliometrics; scientometrics; resource allocation; project selection; operations research; management science.

# REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188



1. REPORT DATE (DD-MM-YYYY)
15-07-2003

2. REPORT TYPE
Technical

3. DATES COVERED (FROM - TO)
xx-xx-2001 to xx-xx-2003

4. TITLE AND SUBTITLE
SCIENCE AND TECHNOLOGY TEXT MINING
MANAGEMENT DECISION AIDS
Unclassified

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kostoff, Ronald Neil ;

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Office of Naval Research
800 N. Quincy St.
Arlington, VA22217

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
,

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report addresses issues of quality in the context of management decision aids. It proposes rules for high quality management support procedures using these aids, and presents criteria for more effective implementation of these decision aids into the strategic management of science and technology. Two illustrative examples of quality in the context of management decision aids are presented. The first addresses quality in the context of science and technology roadmaps, and the second addresses quality in the context of information retrieval for science and technology text mining. Finally, the report discusses the major barriers to implementation and integration of these decision aids into the strategic science and technology management process: ?h Techniques treated as add-ons ?h Techniques treated independently ?h Mismatch between performers and users

15. SUBJECT TERMS
Decision aids; information storage; information processing; information retrieval; roadmaps; metrics; peer review; data mining; text mining; bibliometrics; scientometrics; resource allocation; project selection; operations research; management science.

16. SECURITY CLASSIFICATION OF:

| a. REPORT | b. ABSTRACT | c. THIS PAGE |
|---|---|---|
| Unclassified | Unclassified | Unclassified |

17. LIMITATION OF ABSTRACT
Same as Report (SAR)

18. NUMBER OF PAGES
16

19. NAME OF RESPONSIBLE PERSON
Kostoff, Ronald
kostofr@onr.navy.mil

19b. TELEPHONE NUMBER
International Area Code
Area Code Telephone Number
703696-4198
DSN
-

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

## BACKGROUND

The growth in available databases and information storage and processing capabilities has resulted in an attendant proliferation of computer-based management decision aids. These management support techniques include roadmaps, metrics, peer review, data and text mining, information retrieval, bibliometrics, and retrospective studies. The potential benefits to S&T available from use of these techniques may be substantial, but the benefits realized so far have been minimal. There are two central reasons for this deficiency. First, there has been little understanding of, and little attention paid to, the intrinsic quality of these decision aids. Second, these decision aids have not been implemented correctly into the overall S&T management process.

The present report discusses the meaning of quality in the context of management decision support, proposes rules for high-quality management support procedures using these aids, and describes criteria for more effective implementation of these decision aids into the strategic management process. To provide tangible demonstration of the decision aid quality problem, and set the stage for the more universal conclusions that follow, two illustrative examples will be presented in some detail. The first concerns quality issues related to S&T roadmaps, and the second concerns the meaning of quality in the context of information retrieval for text mining.

## QUALITY ISSUES RELATED TO S&T ROADMAPS

The author's major roadmap documents (1, 2) focused on principles of high quality roadmaps, different classifications of roadmaps, and specific examples of many different types of roadmaps. As shown by the bibliography in (1), there are hundreds of documents that come under the broad umbrella of S&T roadmaps. One major problem in interpreting and using these documents is the inability to ascertain their quality. There are no independent tests of roadmap quality. Unlike the physical and engineering sciences, there are no primary physical reference standards against which one can benchmark the roadmap product.

Even the metrics of roadmap quality are unclear. Roadmap (and other decision aids) quality is a very subjective term, and has intrinsic and extrinsic components. Quality depends not only on the technical construction of the roadmap (the intrinsic component), but depends on the objectives of the roadmap application as well (the extrinsic component). If the objective of the application is to attract investor interest in a technology/ system, then the quality metric would relate to dollars invested subsequent to the roadmap. How well the roadmap represented the state or potential of S&T is of little consequence, as long as the major objective of capital attraction was achieved. Alternatively, if the objective of the application is to reflect the state and potential of S&T fully, then this becomes the metric of quality. The latter concept/ application of roadmap quality is the one used in the remainder of this section.

To illustrate the roadmap metrics quality problem further, consider the following example. Suppose a prospective technology-push roadmap had been constructed for high energy-density batteries. Suppose further that fifteen years after the roadmap was developed, an assessment was performed of the roadmap predictions as compared to the battery state-of-the-art. Suppose even further that the assessment showed that the roadmap development plan was followed religiously by the technical community, and the long-range technical goals were achieved exactly as predicted by the roadmap. Does that mean the roadmap was of high quality; i.e., that it reflected the state and potential of battery S&T fully?

Not necessarily. The roadmap developers may have been very conservative in their targets, and did not 'push the envelope' to develop the field as vigorously as technology would have allowed. The developers may also have been very narrow in their outlook, and may not have drawn from other disciplines sufficiently to develop the batteries to the greatest extent. It could be stated that the roadmap was precise (in predicting the goals that were actually achieved), but was not accurate (the most visionary goals were not predicted).

On the other hand, the roadmap in this case may have been of the highest quality. The developers may well have had very ambitious targets, and may have drawn from other disciplines to the maximum extent possible. The point to be made here is that the concepts of roadmap quality, and its associated metrics, are very complex and diffuse, yet very important if roadmaps are to become useful operational tools.

A high quality S&T roadmap that integrates all temporal stages of development requires the following conditions:

1) the retrospective component must be an accurate reflection of the evolution and relation of all critical sciences and technologies that resulted in the technology of present interest;
2) the present time component must be an accurate reflection of all critical science and technology related to the technology of interest; and

3) the prospective component should reflect some degree of vision by the planners and should incorporate all the critical science and technology areas that relate to the technology of interest and to the projected targets.

The roadmap's utility is enhanced substantially if some intrinsic processing capability is present; i.e., if the quantitative relationships between the roadmap's component elements can be incorporated in functional form, and sensitivity or tradeoff studies can then be done. Its utility is enhanced further if critical attributes (cost, schedule, risk, performance targets) can be displayed throughout. Thus, a high quality S&T roadmap is analogous to a high resolution picture of the evolving/ changing relationships among science and technology areas related critically to the focal roadmap technology, and incorporates especially the concepts of awareness, risk, coordination, vision, and completeness.

## QUALITY ISSUES RELATED TO INFORMATION RETRIEVAL FOR TEXT MINING

A 1997 article on information retrieval (3) focused on the use of computational linguistics imbedded in an iterative relevance feedback procedure. In this approach, a database query is expanded by incorporating phrase patterns from relevant documents, and the query is contracted by subtracting phrase patterns obtained from non-relevant documents. The final product is a comprehensive query. Quality in the context of information retrieval, from the author’s perspective, requires three conditions:

1) the query will retrieve the maximum number of relevant documents;
2) the query will retrieve a large ratio of relevant to non-relevant documents; and
3) the desired definition of 'relevant' must be incorporated into the query development process.

As in the previous roadmap example, the operational meaning of 'relevant' depends on the objectives of the query/ study. Is the purpose of the query/ study to retrieve all the papers in:

1) a very narrowly focused target technical field;
2) allied technical fields as well;
3) very disparate technical fields that have the potential to provide innovative new insights to advance the target technical field (4)?

Each of these purposes defines a very different concept of 'relevant', and would result in very different numbers of 'relevant' documents being retrieved. The definition of ‘relevant’ is the major determinant of the size of literature retrieved.

Typical S&T literature surveys have none of these three quality conditions. Extensive evaluation of those Medline and Science Citation Index review articles that contained the queries used in the review showed that most queries consist of a few key words fairly closely associated with the desired narrow target literature, with minimal (if any) iterative steps. The results will either contain substantial noise if the search terms are relatively broad, or will be very limited if the search terms are narrowly focused. Some iterative approaches will provide substantial numbers of records with high signal-to-noise ratio using a constrained definition of relevant; i.e., not accessing the disparate literatures from which innovative ideas could potentially flow. Rarely, if ever, are all three necessary conditions for a high quality information retrieval fulfilled. Why is this?

Probably the main reason is time and cost. The author's recent information retrieval/ text mining efforts (5-11) have shown that an iterative process that incorporates a broad scope of 'relevant' disciplines to the target discipline requires the participation of a technical domain expert(s) and a computational linguistics expert (or at least a documented procedure using computational linguistics tools). There is substantial judgement and interpretation required by at least one expert at each iterativestep, and this effort translates directly into significant resource expenditures. The downside of not expending sufficient resources to obtain a high quality product is that allied and related literatures that could serve as the engines of innovation are not accessed. Money saved on the front end is wasted on the back end!

As an example of the level of effort required for a reasonable quality query, the author, in conjunction with technical domain experts, recently developed a query related to electrochemical power sources. Three iterative steps were required; each step required the technical expert(s) to read many hundreds of the retrieved records in order to identify which were relevant and which were not. Then, computational linguistics analyses (3) were performed on both the relevant and non-relevant records to identify phrase patterns and relationships characteristic of the relevant records and the non-relevant records. Substantial time and judgement were required to select the appropriate phrases unique to the relevant records and the non-relevant records, and then modify the query accordingly. Many terms were contained in the final query. Even then, the process could have continued for moreiterations, but it was not considered cost-effective given the time and resource constraints of the specific study.

## GENERALIZED CONCLUSIONS ON DECISION AIDS QUALITY

Conclusions on quality drawn from the above two specific examples, as well as from myriad examples over many decision aid applications, can be generalized to many other S&T management decision aids. For example, a high quality peer review provides an accurate picture of the intrinsic evolution and status of specific S&T, and its inter-relationships with other S&T and with potential end-use applications. High quality text mining provides an accurate picture of the global trends and status of specific S&T. High quality technology transition metrics provide an accurate picture of 1) the number and potential value of technology transitions that actually occurred compared to 2) the number and potential value of technology transitions that could have occurred had the technology managers taken full advantage of existing and potential opportunities. High quality paper citation metrics provide an accurate picture of 1) citations actually received compared to 2) citations that could have been received if

a) the paper were of extremely high quality, and
b) the paper had been disseminated to all potential users.

Quality applications of all these decision aids:

1) reflect most accurately the history, status, and potential of the S&T area(s) being examined;
2) relate these S&T areas to allied S&T areas;
3) draw insights from disparate S&T disciplines; and
4) incorporate challenges to the frontiers of S&T through a vision of their implementation.

Many of the differences between high and low quality decision aids applications revolve around what could or should have been included as opposed to what was actually included in the application (projects, papers, patents). Most quantifiable metrics focus solely on <u>what was achieved</u> (transitions made, papers published, citations received) for purposes of expediency, and <u>production efficiency</u> is never addressed. Since what could or should have been included is a highly subjective topic, the metrics of evaluating decision aid product quality are very difficult to quantify.

Thus, since decision aid quality cannot be ascertained or measured easily from examination of the final decision aid output product, then the focus for evaluating decision aid quality must be shifted from the decision aid product to the decision aid application process. The next section addresses the process requirements for insuring that the decision aids applications are of high quality.

## REQUIREMENTS FOR HIGH QUALITY DECISION AIDS APPLICATIONS

Each of the major management science components (peer review, retrospective studies, metrics, roadmaps) on the author's web site (12) contains a section outlining the principles/ characteristics of a high quality decision aid application (typically in the context of an S&T assessment) using the specific approach examined. All of these principles sections are specific cases of a unified set of generic principles/ characteristics/ requirements for high quality S&T evaluations and decision aids applications. While the priority order for specific principles may vary slightly for different techniques, the following list represents the fundamental requirements necessary for best practices. The language is stated in terms of an S&T evaluation, but is easily generalized to all decision aid applications.

1) Senior Management Commitment
The most important factor in a high-quality S&T evaluation is the serious commitment to high-quality S&T evaluations of the evaluating organization's most senior management with evaluation decision authority, and the associated emplacement of rewards and incentives to encourage such evaluations. Incorporated in senior management's commitment to quality evaluations is the assurance that a credible need for the evaluation exists, as well as a strong desire that the evaluation be structured to address that need as directly and completely as possible.

2) Evaluation Manager Motivation
The second most important factor is the operational evaluation manager's motivation to perform a technically credible evaluation. The manager:

a) sets the boundary conditions and constraints on the evaluation's scope;

b) selects the final specific evaluation techniques used;

c) selects the methodologies for how these techniques will be combined/ integrated/ interpreted, and

d) selects the experts who will perform the interpretation of the data output from these techniques.

In particular, if the evaluation manager does notfollow, either consciously or subconsciously, the highest standards in selecting these experts, the evaluation's final conclusions could be substantially determined even before the evaluation process begins. Experts are required for all the evaluation processes considered (peer review, retrospective studies, metrics, economic studies, roadmaps, data mining, and text mining), and this conclusion about expert selection transcends any of these specific applications.

3) The third most important factor is the transmission of a clear and unambiguous statement of the review's objectives (and conduct) and potential impact/ consequences to all participants. This statement should occur at the very beginning of the review process.

4) Competency of Technical Evaluators
The fourth most important factor is the role, objectivity, and competency of technical experts in any S&T evaluation. While the requirements for experts in peer review, retrospective studies, roadmaps, and text mining are somewhat obvious, there are equally compelling reasons for using experts in metrics-based evaluations. Metrics should not be used as a stand-alone diagnostic instrument (13). Analogous to a medical exam, even quantitative metrics results from suites of instruments require expert interpretation to be placed into proper context and gain credibility. The metrics results should contribute, and be subordinate, to an effective peer review of the technical area being examined.

Thus, this fourth critical factor consists of the evaluation experts' competence and objectivity. Each expert should be technically competent in his subject area, and the competence of the total evaluation team should cover the multiple science and technology areas critically related to the science or technology area of present interest. In addition, the team's focus should not be limited to disciplines related only to the present technology area (that tends to reinforce the status quo and provide conclusions along very narrow lines). It should be broadened to disciplines and technologies that have the potential to impact the overall evaluation's highest-level objectives (that would be more likely to provide equitable consideration to revolutionary new paradigms).

5) Selection of Evaluation Criteria
The fifth most important factor is selection of evaluation criteria. These criteria will depend on the interests of the audience for the evaluation, the nature of the benefits and impacts, the availability and quality of the underlying data, the accuracy and quality of results desired, the complementary criteria available and suites of

diagnostic techniques desired for the complete analysis, the status of algorithms and analysis techniques, and the capabilities of the evaluation team.

6) Relevance of Evaluation Criteria to Future Action
A factor of equal importance to evaluation criteria selection is one that has been violated in almost every metrics briefing the author has attended, spanning many government agencies, industrial organizations, and academic institutions. In general, this factor tends to be violated for the evaluation criteria used in any of the evaluation approaches under the decision aids umbrella. The factor will be stated in terms of a metrics-based evaluation, but it should be considered as applicable to all evaluation techniques.

EVERY S&T METRIC, AND ASSOCIATED DATA, PRESENTED IN A STUDY OR BRIEFING SHOULD HAVE A DECISION FOCUS; IT SHOULD CONTRIBUTE TO THE ANSWER OF A QUESTION WHICH IN TURN WOULD BE THE BASIS OF A RECOMMENDATION FOR FUTURE ACTION.

Metrics and associated data that do not perform this function become an end in themselves, offer no insight to the central focus of the study or briefing, and provide no contribution to decision-making. They dilute the theme of the study, and, over time, tend to devalue the worth of metrics in credible S&T evaluations. Because of:

1) the political popularity and subsequent proliferation of S&T metrics;
2) the widespread availability of data; and
3) the ease with which this data can be electronically gathered/ aggregated/ displayed,

most S&T metrics briefings and studies are immersed in data geared to impress rather than inform. While metrics studies provide the most obvious examples, this conclusion can be easily generalized to any of the evaluation methods.

7) Reliability of Evaluation
Another factor of equal importance is reliability or repeatability. To what degree would an S&T evaluation be replicated if a completely different team were involved in selection, analysis, and interpretation of the basic data? If each evaluation team were to generate different evaluation criteria, and in particular, generate far different interpretations of these criteria for the same topic, then what meaning or credibility or value can be assigned to any S&T evaluation (14)? To minimize repeatability problems, a diverse and representative segment of the overall competent technical community should be involved in the construction and execution of the evaluation.

8) Evaluation Integration

A fourth factor of equal importance is the seamless integration of evaluation processes in general into the organization's business operations. Evaluation processes should not be incorporated in the management tools as an afterthought, as is the case in practice today, but should be part of the organization's front-end design. This allows optimal matching between data generating/ gathering and evaluation requirements, not the present procedure of force fitting evaluation criteria and processes to whatever data is produced from non-evaluation requirements.

9) Global Data Awareness

A fifth factor of equal importance is data awareness. In all of the decision aids, placement of the technology of interest in the larger context of technology development and availability globally is an absolute necessity. This tends to be a central deficiency of most management decision aids. Lack of S&T documentation, inaccessibility of S&T that is documented, inability to retrieve S&T documents due to poor retrieval methods, inability to extract information from large retrievals, and general lack of interest and will in global data awareness, mitigate against attaining comprehensive global data awareness.

10) Normalization across Technical Disciplines

For evaluations that will be used as a basis for comparison of science and technology programs or projects, the next most important factor is normalization and standardization across different science and technology areas. For science and technology areas that have some similarity, use of common experts (on the evaluation teams) with broad backgrounds that overlap the disciplines can provide some degree of standardization. For very disparate science and technology areas, some allowances need to be made for the relative strategic value of each discipline to the organization, and arbitrary corrections applied for benefit estimation differences and biases. Even in this case of disparate disciplines, some normalization is possible by having some common team members with broad backgrounds contributing to the evaluations for diverse programs and projects (15). However, normalization of the criteria interpretation for each science or technology area's unique characteristics is a fundamental requirement. Because credible normalization requires substantial time and judgement, it tends to be an operational area where quality is sacrificed for expediency.

11) Cost of S&T Evaluations

The next critical factor for quality S&T evaluations is cost. The true total costs of developing a high quality evaluation using sophisticated normalization techniques and diverse experts for analyses and interpretation can be considerable, but tend to be understated. In high quality evaluations, sufficient expertise is represented on the evaluation team, as well as by the presenters. The major contributor to total costs is the time of all the individuals involved in presenting, analyzing, and interpreting the data. With high quality personnel involved in the presentation and evaluation process, time costs are high, and the total evaluation costs can be non-negligible. Especially when suites of diagnostics are combined, as when a metrics-based evaluation is performed in tandem with a qualitative peer-review process (13), the real costs of these experts could be substantial. Costs should not be neglected in designing a high quality S&T evaluation process.

12) Maintenance of High Ethical Standards
The final critical factor, and perhaps the foundational factor, in any high quality S&T evaluation is the maintenance of high ethical standards throughout the process. There is a plethora of potential ethical issues, including technical fraud, technical misconduct, betraying confidential information, and unduly profiting from access to privileged information. This stems from an inherent bias/ conflict of interest in the process when real experts are desired to participate in every aspect of an S&T evaluation. The evaluation managers need to be vigilant for undue signs of distortion aimed at personal gain.

## LIMITATIONS OF PRESENT S&T DECISION AID IMPLEMENTATION APPROACHES

Above and beyond problems with decision aids' quality issues are problems with the implementation and integration of these decision aids into the strategic S&T management process. There are three major implementation-related problems with management decision aids, both in practice and in the published literature. These problems are:

1) the management support techniques tend to be treated as add-ons;
2) the management support techniques tend to be treated independently; and
3) there is a major mismatch between the developers of the (especially literature-based) management support techniques and the users of these techniques.

The first two of these problems stem from the same fundamental cause, namely, that advanced computerized management support techniques are not conceptualized and implemented as an organic component of the management structure and process. The third problem arises from the separation of the contributors to the published literature from the evaluation practitioners. Each of these three problems, and some potential solutions, will now be addressed.

1) Techniques Treated as Add-ons

The various decision aid tools and procedures are not incorporated into the structure of the organization, but are treated as add-ons. For example, management/ technology metrics are generally not imbedded as an integral part of an organization's intrinsic operating structure. They tend to be employed on a fragmented basis in response to external pressures. They tend to make use of whatever data is available as a result of ordinary business practices, and not the desired type of focused data that would address progress toward corporate strategic goals if the use of metrics were an integral organizational component. Thus, in practice, the data obtained from normal business operations determines the metrics that can be credibly employed, and the metrics in turn determine the objectives whose progress can be gauged. Conversely, in a strategic management process, the objectives would determine the metrics used to gauge progress, and the metrics in turn would determine the data required for their quantification. This metrics example can be extrapolated generically to other management science techniques; they all tend to be used on a sporadic basis. This fragmented approach makes little use of the full power available from the existing management science tools.

2) Techniques Treated Independently

Generally, the various management science techniques, if used at all within an organization, are employed independently. One person or group may be doing metrics, another person or group peer review, a third person or group roadmaps, a fourth person or group data mining, and so on. The synergies that can be exploited by employing these tools in a unified approach are never realized. Reference 4 presents an example of promoting and stimulating innovation through a combination of workshop-based and literature-based approaches; this example illustrates some of the synergistic benefits possible from accessing multiple management science tools. In the complex systems of management science, as in the complex systems of physical/ biological/ engineering sciences, the whole is indeed greater than the sum of its parts. In all these complex multi-component systems with highly interactive elements, the intelligence that links the components and allows communication and control provides the benefits from the synergy.

3) Mismatch Between Performers and Users

Over the past few years, the author has conducted a number of literature surveys and subsequent studies in fields that can be loosely called 'management science', including research assessment, peer review, metrics, text mining, information retrieval, resource allocation, project selection, and roadmaps. The specific conclusions from the metrics survey will be described, and then generalized to cover all the areas surveyed.

Most of the documents retrieved in the metrics survey described the generation of a multitude of metrics of large data aggregates, with no indication of the relevance of these metrics to any questions or decisions supporting S&T evaluations. The foundation of this problem is the strong dichotomy between the researchers who publish metrics studies in the literature, and the managers who use metrics to support budgetary allocation and other management decisions. Most of the people who employ metrics for management purposes do not document their experiences and approaches in the literature. Most of the principle and concept and (potential) application papers in the metrics literature are written by people who have never used or applied metrics for management decision-making purposes. In addition, many of the researchers who perform metrics studies focus on single approaches or single approach applications, in order to promote the concepts that they have developed. The managers who use metrics, conversely, have very eclectic requirements. They need suites of metrics, or suites of metrics combined with other evaluation approaches, in order to perform comprehensive multi-faceted S&T evaluations. Thus, there is a serious schism between the incentives and products of the metrics researchers (suppliers) and the incentives and requirements of the metrics users (customers).

Consequently, there are two major gaps in the literature on S&T metrics. First, there are few relevant papers published. Second, most of the concept and principle and (potential) application papers that do exist bear little relation to the reality of what is required to quantitatively support science and technology assessments and evaluations for decision-making. Because of the deficiency of metrics studies relevant to S&T applications, it is difficult to extract the conditions for high quality metrics-based evaluations solely from the open literature. Drastic alterations in this overall situation are required if metrics are going to support future government and industry business requirements in any credible manner.

While there are some minor differences among the diverse management decision aid domains surveyed, the following observation generally appears to transcend disciplines, and can be considered universal and invariant. Most of the people who conduct program evaluations/ assessments/ plans (including practitioners who use the management science tools listed above in their repertoire) do not document their studies and/ or approaches/ techniques in the literature. Most of the management science papers in the literature are written by people who have never conducted program evaluations/ assessments/ plans. Consequently, there is a major gap in the management science literatures, which is reflected as a major split between the theory and the practice of management science.

Consider, for example, the advanced operations research (and other) techniques available in the literature for resource allocation applications, and then observe how resources are allocated in practice. Or, as another example, consider the esoteric literature publications on information retrieval techniques, and contrast those with methods actually used by librarians and other information resource personnel to retrieve information. Many of the papers in the management science literature are very sophisticated, while most of the techniques actually used by the practitioners are very primitive and rudimentary. While the literature papers may have substantial academic merit, many bear little relation to the reality of conducting program evaluations/ assessments/ plans. The practice of management science lags far behind what the technology of management science can offer.

## SUMMARY AND CONCLUSIONS

For management decisions aids to gain wider acceptance, more attention needs to be paid to their quality. This includes intrinsic, extrinsic, and implementation quality. The decision aid quality metrics need to be sharpened for specific applications, the requirements for high quality applications have to be considered carefully, and the decision aids need to be integrated into an organization's overall management processes.

## REFERENCES

1. Kostoff, R. N. "Science and Technology Roadmaps". http://www.onr.navy.mil/sci_tech/special/technowatch/. 1998.

2. Kostoff, R. N., and Schaller, R. R. "Science and Technology Roadmaps". IEEE Transactions on Engineering Management. 48:2. 132-143. May 2001.

3. Kostoff, R. N., Eberhart, H. J., and Toothman, D. R., "Database Tomography for Information Retrieval", Journal of Information Science, 23:4, 1997.

4. Kostoff, R. N. "Science and Technology Innovation". Technovation. 19:10. 593-604. October 1999.

5. Kostoff, R. N. "Database Tomography for Technical Intelligence: Comparative Analysis of the Research Impact Assessment Literature and the Journal of the American Chemical Society. Scientometrics. 40:1. 1997.

6. Kostoff, R. N., Eberhart, H. J., and Toothman, D. R. "Database Tomography for Technical Intelligence: A Roadmap of the Near-Earth Space Science and Technology Literature". Information Processing and Management. 34:1. 1998.

7. Kostoff, R. N., Eberhart, H. J., and Toothman, D. R. "Hypersonic and Supersonic Flow Roadmaps Using Bibliometrics and Database Tomography". Journal of the American Society for Information Science. 50:5. 427-447. 15 April 1999.

8. Kostoff, R. N., Green, K. A., Toothman, D. R., and Humenik, J. "Database Tomography Applied to an Aircraft Science and Technology Investment Strategy". Journal of Aircraft, 37:4. 727-730. July-August 2000. A more detailed description of the study can be found at Kostoff, R. N., Green, K. A., Toothman, D. R., and Humenik, J. A. "Database Tomography Applied to an Aircraft Science and Technology Investment Strategy". TR NAWCAD PAX/RTR-2000/84. Naval Air Warfare Center, Aircraft Division, Patuxent River, MD.

9. Kostoff, R. N., Braun, T., Schubert, A., Toothman, D. R., and Humenik, J. "Fullerene Roadmaps Using Bibliometrics and Database Tomography". Journal of Chemical Information and Computer Science. 40:1. 19-39. Jan-Feb 2000.

10. Kostoff, R. N., Tshiteya, R., Pfeil, K. M., and Humenik, J. A. "Electrochemical Power Source Roadmaps using Bibliometrics and Database Tomography". Journal of Power Sources. 110:1. 163-176. 2002.

11. Kostoff, R. N., Shlesinger, M., and Malpohl, G. "Fractals Roadmaps using Bibliometrics and Database Tomography". Fractals. December 2003.

12. http://www.onr.navy.mil/sci_tech/special/technowatch/

13. Kostoff, R. N., "Peer Review: The Appropriate GPRA Metric for Research", Science, Volume 277, 1 August 1997.

14. Cole, J.R. & Cole, S. “Peer Review in the National Science Foundation: Phase Two of a Study”. Washington, DC. National Academy Press. 1981.

15. Van den Beemt, F.C.H.D. and Le Pair, C., "Grading the Grain: Consistent Evaluation of Research Proposals", Research Evaluation, 1:1, 1991.