

WHITE PAPER

An approach to developing the Federation Metrics Working Group

**Points for Discussion at the Federation meeting,
2-4 December 1998
Santa Barbara, California**

**David L. Skole
Michigan State University**

**Howard Burrows
Raytheon ITSS**

with inputs from

**Martha Maiden
NASA**

**Amy Glasmeier
Penn State University**

**Gerald Urquhart
Michigan State University**

**Kevin Curtis
Keystone Group**

Introduction.

This document is meant as a basis for discussion. Nothing herein is etched in stone. It is a very brief summary of an initial meeting of the Co-chairs held in October, as a way to stimulate and organize the discussions at the upcoming Federation-wide meeting.

The document first lays out the background issues, focusing on the important aspect of what expectations are out there in the broad community of stakeholders and how we need to address those issues through evaluation. The necessity of meeting metric and evaluation requirements of the Federation's main stakeholder – NASA – is prominent in this document. Then the document provides a general review of the approach and strategic issues for a Working Group, followed by some first order suggestions on four Components of Metrics which could be further developed by a Working Group. To that end the document then provides a suggested set of Output Metrics. Lastly we review some of the issues for constituting a Working Group and its membership, with a charge to the Federation as a whole.

Background: History of the Federation and the Role of Evaluation

Beginning in the 1980s with the development of new instruments and aerospace capabilities, NASA began generating high-quality remotely sensed data for earth science applications. Given technological limitations and a relatively narrow base of presumed customers, the first-generation EOSDIS system was based around a model emphasizing centralized computing and a hierarchical system of information production, storage, and retrieval.

Several developments led up to the creation of the NASA-ESIP Federation experiment. New technology enabled high speed distributed computing. The Internet and the World Wide Web for the first time made possible the widespread and timely generation and dissemination of information. In turn, an expanding customer base created demands for changes in both the scope and character of science data. Congress reprioritized NASA's mission to include increasing the relevance and accessibility of its products and services. Finally, the emergence of an Information Economy created new opportunities for service industries across the board. From 1959 to 1994, the service sector grew from 49 to 62 percent of U.S. gross domestic product (GDP).¹ These developments together led to calls for a new system that was more user/science-driven, less centralized, and better able to capitalize on the new service economy.

In response to these developments and at the recommendation of an NRC report, NASA established a competitive process for selecting participants as charter members of a working prototype federation. Twenty-four NASA-funded projects were selected; these projects took on responsibility for forming a federation to govern their interaction and operations. Based on a spring 1998 National Research Council report, the federation goals are to "enable optimum ways to develop, produce, and publish/distribute environmental information and to provide associated services to science and society."² In this capacity, NASA expects each ESIP in the federation to "perform selected functions of NASA ESE DIS services as NASA's agents, including the processing of data to higher-level products, the archiving of data in a searchable, retrievable way, providing data and services to user communities, including NASA affiliates, and maintaining metadata and data quality to ensure the archive is readily transferable."³ In accordance with these expectations, cost-effectiveness, increased capacity, greater user involvement, more flexibility, fewer bottlenecks, and well-conceived and smaller projects are the projected outcomes of the ESIP program.

¹ Science and Engineering Indicators 1998
<http://www.nsf.gov/sbe/srs/seind98/access/c8/c8s2.htm>

² National Research Council (1998), *Toward and Earth Science Enterprise Federation: Results from a Workshop*. Washington, DC.

³ Notes from presentation by Dr. Martha Maiden, ESE program manager, May 5-7, 1998.

Challenges and Opportunities

The federation model of information production and dissemination presents a number of exciting and important challenges to the data producer and user communities. The first set of issues relates to the federation and its operation. A major challenge is determining the extent to which the federation model facilitates the development of protocols and uniform operating procedures across the projects. A second challenge relates to the degree of project interaction and support. The model's success depends upon reaching a balance in which the flexibility and local efficiencies achieved through decentralization trade off against economies of scale achieved through the federation. A third challenge is the extent to which the ESIP experiment generates some degree of community acceptance and greater participation in data development and data usage by the broader scientific user community. The effectiveness of the federation model corresponds to the degree of acceptance and validation given to the federation experiment by the existing user community including such as the data centers of NOAA and U.S. Geological Survey.

The second set of issues refers to the success of the individual projects. The challenge in this instance is to develop evaluation metrics that capture project-specific attributes of data development and delivery. Of significant import is the recognition that the selected projects are outcome-based and hence responsible for providing deliverables at their end. To be effective, evaluation requirements must be adapted to projects in an effective, inexpensive way.

In addition to the challenges there are also opportunities. The historic high since 1957 for the Nation's R&D/GDP ratio was reached in 1964 at 2.87 percent; the low was 2.12 percent in 1978.⁴ Clearly science has yet to achieve the economic benefit achieved in other sectors of the new Information Economy. The new technologies and customer base give us the opportunity to generate increased funding for basic research, increased funding that is appropriate given the increased ability to generate quality data for analysis.

There is another sea change that we might address with this Federation. A unit of commerce within the scientific community has been the journal article. The number of articles published per year nearly doubled from 1975 to 1995 (312 thousand to 557 thousand)⁵ while the average subscription cost leapt from \$39 to \$284. In taking advantage of new technology we need to examine ways to move from a journal article economy to one that better benefits science and scientists. We might seek ways to increase the degree of recognition accorded the ESIP's information producers and disseminators.

4 NSF 99-302, October 16, 1998 Division of Science Resources Studies DATA BRIEF, <http://www.nsf.gov/sbe/srs/databrf/sdb99302.htm>

5 Tenopir, C and DW King. Economic and Use Patterns of Scientific Journals. In Proceedings Socioeconomic Dimensions of Electronic Publishing Workshop April 1998, IEEE Technical Activities ISBN 0-7803-5101-0 .

In creating the federation, NASA has established a two-level experiment. The first level consists of the federation itself, a self-governing, autonomous, bottom-up, constituent-driven governance system bearing responsibility for generating relevant data products for the earth science community. The rationale underlying this experimental element is that through the creation of a shared governance system, protocols, standards, rules, and interoperability will be achieved. The second level relates to the highly decentralized nature of the federation model, which is presumed to deliver data elements that embody diversity, relevance, and timeliness of data delivery while being sensitive to science community needs—qualities that often evade a more centralized system akin to current practice.

Approach for Metrics Working Group

The task of metrics for the ESIP federation can be divided into three major categories as shown in Table 1. The success of the ESIP Federation experiment relies on successes at each of these levels. The challenge for the Metrics WG is to evaluate performance and "successes" in each of these areas, assessing the function of the individual ESIPs and the Federation as an on-going process for internal evaluation and planning, and for evaluating the achievements of the ESIPs and Federation for NASA-ESE and external review.

Table 1. The three application areas of Federation metrics and the answers they require.

ESIPs-level	Federation-level	NASA-ESE or Information Economy
Is the individual ESIP achieving its goals?	Is the Federation aiding the ESIPs and their users?	Is the ESIP Federation forwarding ESE's goals?

Metrics at the level of the individual ESIP must address the function of the ESIP—Is the ESIP achieving its goals? In the proposals to NASA, individual ESIPs proposed "internal" metrics on the function of the ESIP itself. Although not all of these measures are applicable to each ESIP, metrics must be performed at the ESIP level and some commonality is needed among the measures. Metrics at the ESIP level can be generally referred to as "Evaluation Metrics."

Additionally, the Federation itself needs to measure its function—satisfaction for the members (ESIPs), satisfaction of the user community, and integration into the information economy. At the level of the information economy, NASA needs to determine whether the ESIP Federation experiment has succeeded.

Internal metrics can be performed within each ESIP to measure the function, volume of data and information, and service that each ESIP provides. During the proposal process, each ESIP suggested metrics to perform. From the 24 funded proposals, a large number of possible measures were suggested. They fall into the following categories, with

examples of each type listed parenthetically: 1) General Metrics (cost savings, market breadth); 2) Data Metrics (web statistics, amount of data transferred); 3) Performance Criteria (user-friendliness, speed of use, bandwidth); 4) Outcome Measures (dollar value of new parameter, dollar value per product, cost/benefit analysis); and 5) Measures of Market Access (number of users, user categorization, user community).

Federation level metrics are more challenging to outline. One of the most important measures of the Federation is in its utility and benefit to the ESIPs themselves. Does participating in the federation aid the individual ESIP? The question of how metrics on the Federation can be measured needs to be addressed in order to determine whether it is the role of the ESIPs to grade the Federation, or whether the Federation should evaluate itself, or whether NASA-ESE should evaluate the Federation. A related question is how well does a federated approach benefit the user in synergistic ways not present if individual ESIPs were not federated? Or, does governance improve the federation and improves user access and requirements, or make it more responsive to technological changes?

Finally, during the ESIP Federation experiment, the Metrics WG or another body needs to measure whether the ESIPs and the Federation are succeeding in distributing "information" within the context of the burgeoning information economy and the direct requirements of NASA. Moreover, the Federation of ESIPs also needs to address the general needs of an environmental information economy, and thus needs to be responsive to needs of the broad external stakeholders beyond NASA, including the NRC, other agencies, etc. But clearly, the primary objective should be to meet NASA's goals and requirements, and demonstrate clear gains in cost, efficiency and effectiveness over previous approaches. In many respects, this is the ultimate question that NASA-ESE needs to answer in evaluating the ESIP Federation experiment.

Table 2. What needs to be measured at each level?

ESIPs	Federation	Information Economy
Evaluation Metrics: <ul style="list-style-type: none"> • Fulfillment of goals • Efficient archiving • Efficient access 	Effectiveness of scale ESIP participation Cooperation and Clustering of ESIPs User satisfaction	ESE Mission Statement ESIP-Federation Vision Get NASA out of old (centralized?) role

Strategic Objectives

The strategic role of metrics and the Working Group can be twofold: 1) to ensure that ESIPs, the Federation, and the NASA-ESE are fulfilling their mission (diagnostic), and 2) to aid the Federation and WG in identifying areas that need to be addressed to improve the ESIPs and the Federation (prognostic). An individual measure can be used in both senses; moreover, it is the role of the WG to determine which emphasis it will choose--to use

metrics in a diagnostic or a prognostic mode. These objectives should be discussed and developed by the WG.

Diagnostic use of metrics will allow the ESIPs, Federation, and NASA-ESE to gauge the function of individual ESIPs and the Federation. Diagnostic metrics should be regularly reported to and analyzed by the WG for release to the community.

Prognostic Metrics should also be a goal of the WG. By identifying potential successes and problems ahead of time, the WG can share this information with the Federation members to help them in ESIP implementation. Through prognostic metrics, good examples can be shared with other ESIPs. Alternatively, potential difficulties can be recognized early and modified for success.

Components of Metrics

The working group could initiate the following types of evaluations: 1) self-evaluation by ESIPs, which would focus on individual ESIP performance and success and conducted internally but made publically available, 2) self-evaluation of the Federation, focusing on a federation-wide assessment rather than individuals and 3) Evaluation of the Federation by External Review. The first two evaluations are called for at the end of the document and require a short report from each ESIP. After convening, the WG should discuss how to proceed on the external review.

The specific application of metrics to the Federation effort could be developed in four broad areas: (a) an external evaluation and report of the federation to be developed cooperatively with the Federation and published as one end product of the “experiment” to be distributed to the various immediate stakeholders, including NASA, NRC, etc., (b) an on-going process of assessing user satisfaction and user feedback to ensure we have a mechanism for improving services and evaluating what users find good and bad about individual ESIPs and the Federation, (c) development of a user services forum as a means for outreach and education, to inform and broaden the user community, and (d) a forward-looking and strategic process of identifying new applications, user groups, and mechanism for product and services generation in the context of the broader vision of a environmental information economy.

External Evaluation Report. First, within an historical context, it may be necessary to develop a kind of Report Card on the Federation – how does it meet the expectations and needs of various stakeholders as illustrated in the precursor documents coming from the NAS, NASA, etc. This would be a useful starting point, and is further addressed by way of background in the opening section of this document. To do this, it would be necessary and useful to develop a protocol for evaluation, as well as a specific instrument. Under contract from the Federation, Drs. Glasmeier and Feller will be developing this component. Copies of the work statement are available. We suggest that the evaluation instrument be developed in consultation with the Federation and then initially tested on one or a few ESIPs to get Federation feedback before it is implemented fully. This

evaluation would emphasize doing what we need to do to answer the basic questions which NASA and other stakeholder will have regarding the basic performance, utility, and success of a federated approach. This would benefit from a strong external review for the sake of objectivity and impartiality, but should be done cooperatively with the Federation. The output of this review would be two reports: (a) an interim report at least one year prior to the end of the “experiment” to be reviewed by the federation and NASA, and (b) a final report at the end of the experiment to be published and distributed widely

Assessing User Satisfaction and User Profiles. The second areas would be to assess user satisfaction, which would be an on-going processes beginning as soon as possible. The purpose of this type of metric would be to institute an internal process for use by individual ESIPs and the Federation as a whole for obtaining user feedback. In essence this would be aimed an providing an on-going assessment of how best to serve our user needs as individual ESIPs and as a Federation of ESIPs. The approach could contain several layers of efforts, including: (a) installation of simple but informative profiling instruments at each ESIP such as web-page reports, (b) installation of user groups at each ESIP which would meet regularly to provide input to the ESIPs, and (c) development of focus groups for ESIP-specific or Federation-wide assessments. Other means for assessing user satisfaction could be developed by the Working Group.

Developing a User Services Forum for Outreach. Currently there is no mechanism for the Federation providing outreach and information dissemination. The Metrics Working Group could provide this function, with the development of a series of initiatives which are aimed at increasing user access, user understanding of our goals, and providing up-to-date information on services and products as they become available. This could be done through regular mechanisms, such as web pages, newsletters, sample CDs etc. Or, it could be done through conferences and workshops which target categories of users, such as K-12, policy makers, etc. Indeed this is not metrics in an ordinary sense. However, this kind of activity should be done in close collaboration with metrics, and we argue as part of the metrics effort itself in order to provide a basis for measuring and improving effectiveness.

Strategic Assessments. It will be important for the Federation to be constantly planning and evaluating a strategic course of action, taking into account new developments in technology and changing priorities. We envision the Working Group taking a role which serves the strategic and forward-looking needs of the Federation, making assessments of where we could alter or change direction. Given the short time horizon of the experiment this may only be marginally practical, but the concept is valid.

Proposed Protocols and Measures

The WG must decide upon measures to be used in evaluation of the ESIPs, the Federation, and the overall contribution to the Information Economy. In the short term, the WG must develop a basic suite of metrics to be employed by the ESIPs and interpreted by the WG and Federation. A possible model for a minimum suite are the formal metrics employed now by EOS-DIS. These include: 1) data volume received, 2) number of

distinct users, 3) number of products delivered, 4) characterization of user by domain name, 5) outcomes, 6) impacts, and 7) budget expenditures. This initial group of metrics would provide an initial measure for ESIP performance and overall Federation performance in comparison with a baseline from the existing approach centered on DAACs. Although these metrics focus on the individual ESIP, they can be summed to provide information on the overall Federation and may provide useful comparisons to EOS-DIS. Nonetheless they are clearly simplistic and inadequate alone. For instance, while they measure total volume, there exists no companion metric to indicate how many “orders” for data were erroneous. For example lessons learned from the Pathfinder program suggest as much as 20% of all orders need to be re-ordered due to errors or mistakes across a wide field of issues, including metadata errors, media faults, ordering errors, etc.

The development of metrics should build upon what is measured now and work to establish an efficient way to implement what can be measured later . With the information from the ESIPs on the current state of their metrics, the WG can proceed to identify common metrics for the ESIPs. Because of the wide variety of ESIPs, developing a single suite of metrics applicable to every ESIP is not possible.

Initially, we could suggest five general Metrics Categories which could be applied to ESIPs individually or to the Federation as a whole. Note that these are metric categories rather than metrics themselves:

- 1) Cost of fulfilling user request (COFUR);
- 2) Expanding user base (not just numbers);
- 3) Expanding product base;
- 4) Significance of resulting science; and
- 5) Significance/success of technical R&D problems solved.

These five categories are all classed as Output Metrics in that they provide information on how well the Federation is providing products and services in response to user requests. What is missing is a suite of Outcome Metrics, which provide information on how well the Federated approach is improving the environmental information economy and how well it is meeting NASA data and information system needs. These could be developed by the Working Group.

Furthermore, the Metrics WG should define a time table for achievement of major steps in the ESIP development. While trying to avoid a "Big Brother" role, the WG can provide guidelines rather than deadlines. Guidelines on progress for the ESIPs will help ESIPs set goals and gauge achievements.

Forming a Metrics Working Group

The chairs of the Metrics Working Group will assemble a Working Group composed of members of the Federation, users, experts, and advisors. The goal of the WG is to have

members who represent the different types of ESIPs as well as the different functions (Products, services, technologies) of the ESIPs. Additionally, users of the ESIPs are an important part of the working group and need to be incorporated in deciding and designing measures. The team of Amy Glasmeier and Irwin provide expert input for the WG on measuring and interpreting metrics.

Additionally, the Metrics WG group may wish to seek an advisory member, someone cognizant of the past role NASA has taken in data distribution and what future changes were envisioned with the formation of the ESIP federation.

Some initial activities for the WG have been suggested, but the convening of the WG itself will further these and other activities. The initial suggestions include 1) design, evaluation, and implementation of a "metrics instrument" for measuring individual ESIPs, 2) measuring user and "potential user" satisfaction with the ESIP, and 3) measuring outreach efforts for user services.

Charge to the Federation Members

To further the objectives of the WG, we request from each ESIP a short self evaluation of its current activities and an evaluation of the Federation. Please include: 1) what metrics are currently in use or being developed (one paragraph), 2) who are your perceived audience/stakeholders (one paragraph), and 3) the name of a person from within your ESIP who can serve as a contact for the WG and possible representative to the WG. Also, please attach output from any metrics currently performed, to be used in providing ideas for the WG to define how measures should be recorded and reported.