

# Metrics – University of Maryland ESIP Evaluation Team

Contact: Catherine Plaisant [plaisant@cs.umd.edu](mailto:plaisant@cs.umd.edu)

December, 1999

We would like to hear about problems you might have collecting such metrics, or additional metrics that may be more relevant.

Listed here is a set of evaluation metrics to be progressively monitored as the ESIP becomes operational.

You are welcome to use it as a template for your own ESIP evaluation

This html file was kept simple in purpose so that it can be edited directly in MS Word (97 and up) or Netscape Composer.

Note: log analysis needs to ignore access from local machines and from repeat crawlers.

## Acquisition of Data

Metric	# for this month	UMD ESIP Comments	Generic "federation" comments
Number of datasets acquired			New category of metric added Oct 4th
Volume of data acquired			

## Availability of data (What is ACTUALLY available to users)

Number of datasets available to users			
Volume of data available (e.g. number of items, volume in Gbytes)			Can do done manually or calculated at ingestion
Number processing services available (e.g. download, burn CD, subset, atmospheric correction) - provided by staff - self service on the web			Probably a more qualitative measure. Useful to show trends over time of the whole federation
<i>Number of algorithms, interface components or software modules available to others</i>			

## Access to data

Number of users visiting the main home page			Can be done with standard web logs + possibly export to report and visualization tools At first, number of visits will be useful.
Number of users visiting other Key pages (e.g. the search page, the dataset description page, the feedback page)			Same
Number of users acquiring data - by download - by ordering on the web - by mail/phone/email contact			Requires keeping FTP logs and maintaining ordering log file or database.
For each dataset (and can be summed for the whole ESIP) - Number of users - Number of searches with the search engine - Number of previews viewed - Number of data download requests - Number of order requests - Number of items retrieved + total volume			This will allow us to be more specific about what data is used more often.  Requires keeping a log of queries submitted to database. This allows us to understand what users are looking for, not just what they order.

## User characteristics

Number or percentage of identified users			Those are the registered people you can interview or survey to understand user characteristics and data usage. Only this portion of users can help with your users' needs assessment.
Number of identified repeat users and new users			Even if only a small portion of users register, the trend of those numbers will be informative
Categorizing by user type e.g. scientist, business user, students, farmers, fishermen...			i.e., are you reaching the right people? This might influence your advertisement strategy or choice of future features or data. We might want to standardize those categories for the whole federation.
Usage patterns			Requires that web, ftp and search logs can be matched with the user ID (i.e. special programming of the search engine and log analysis) This is hard to do... So it has to be complemented by interviews if you really want to know what people are doing with your interface and data!

## Collecting user characteristics requires a sign-in procedure

If you wish to keep the registration procedure voluntary you will still be able to record changes in usage patterns over time. You can mail users who gave their email address a short questionnaire to collect user characteristics. For GLCF we are considering simply asking users to enter their email address. That will still allow us to contact users later on.

At ordering time users are required to provide complete contact info.

Later on as special services are added, a sign-in procedure will require name and password.

## Development of users community

Number of users actively sending messages or calling the office			A phone or email log is important to estimate impact, request more staff etc.
Classification of questions/problems/comments (+ or -) about: <ul style="list-style-type: none"> <li>• Data itself (availability, future plans, quality etc.)</li> <li>• Access to data (format, processing)</li> <li>• Technical problems (bug, server down, and platform problem)</li> <li>• User interface ("I can't figure out how to...")</li> </ul>			This allows you to identify areas of services that need more attention and help justify additional work. Keep a log and analyze it on a regular basis.
Number of references in the GLCF to related products or people			This measures how open or closed your site is. Can users reach scientists who generated the data, can they be referred to other ESIPs or services?
(possibly) Number of participants in the user's group meeting			
(possibly) Number of messages exchanged in online bulletin board			In brief, record all activity among users... and review it regularly.

### Dissemination of results by ESIP team

Number of papers published about GLFC - technical  - earth science			Are you telling your story?
Number of presentation made to the EOSDIS community, or at conferences			
Number of stories or mention in newspapers.			

### Outcomes

Success stories (e.g. Scientific discovery, impact on businesses or local government, etc.)			Qualitative indicators should supplement quantitative data...
Data products generated by users using your data (either kept at users' site or re-ingested in ESIP)			Requires follow-up, and offering users something in return for their information (e.g. making links to their website, data or papers, a reference in your reports)
Papers written by data users (published elsewhere, or available online at ESIP site)			same
Adoption of technology by other data centers (e.g. data processing, interfaces)			

**Technical performance**

Down time percentage			Most likely estimated
Time to load applets			
Time to perform typical searches			
Availability on different platforms			
Time to process order (average and range) - self serve /automatic - staff provided			

Last revised 01/06/00