

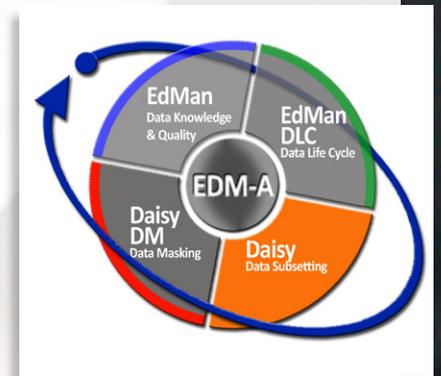


RES[®]

DAISY

Data Subsetting

Subsetting of data, based on the knowledge of the relationships between files and the use of metadata, to ensure time reduction, cost saving and the consistency of the references between data.



The continual transformations that Information Systems organizations currently experience, as a reflection of the significant changes in the organizations over time, require an increasing need for comprehensive new testing environments in terms of the data. Often, to ensure the completeness of the functional testing, it is necessary to create/generate a mirror image of the production data bases which result in significant management issues:

- the need for sufficient space to contain a complete replication
- excessive duration for the batch testing components
- excessive consumption of machine resources
- difficulty identifying the specific data for testing within the large quantity of overall data replicated.

By contrast, ad hoc generated test data typically ends up being insufficient for exhaustive testing. A more appropriate approach is the method of Subsetting, namely, the populating of test environments through copying of selective subsets of data from a reference environment, typically production.

The result of this method is a net reduction in the size of the volumes of data and related problems are eliminated, such as:

- congruency intrinsic to the data
- coverage of the functional cases
- ability to recover to an original base line
- documentation and repeatability of subsetting processes.

The Daisy Solution

Subsetting of data is a challenging issue, especially in complex environments with highly interrelated applications. It can often be misleading when confined to a purely technological level: the knowledge of the application information is in fact a fundamental requirement in order to correctly set the preparation activities of small databases.

Experience, however, shows that adequate technological support is the deciding factor in the success or failure of any proposed data subsetting activity, especially

in relation to the amount of manual tasks the tool can eliminate through automation.

Daisy, a part of the RES EDM-A solution for managing enterprise data, has as its main objective to facilitate subsetting activities and provide support to the various departments and sub-organizations within IT including application development, data and DB Administration and production support.

Daisy provides the IT organization a standardized approach to subsetting activities and supervision of each phase of activity. The web interface and intuitive navigation capability make it easier for the development users by providing an interactive function for the process of data reduction, providing a facility so they can document their own subsetting model and making it ready to be implemented and duplicated as required.

Objectives and Features of Daisy

Like the other RES Data & Data Quality Management solutions, Daisy also relies on the functional knowledge of the context it finds in the EDM-A Catalog, the central repository where the various elements of the archive function are coded along with the information about the definition of the elements for the extraction processes which act on them.

The EDM-A Catalog is also the connecting element between the two basic components of the Daisy functional architecture that is used to implement the basic steps of a proposed subsetting.

The administrative functions of the product, supported by a web-admin component, makes it possible for the user to “educate” Daisy, feeding the EDM-A Catalog, according to the principal of minimizing any manual intervention or activities.

There are numerous “wizards” available with Daisy through the web administration function that make it possible to automate many of the less important repetitive tasks, thus reducing the manual effort by the users to define more complex rules.

The definition discovery of the files and their relationships is one example in addition to the significant generation capabilities for replication or extraction processes based on simple filters.

In particular, the practical application of the concept of metadata, typical of EDM-A, provides for considerable economies of scale, improving the conditions for the selection of an item from a “physical” locationdependent area in the files, to a central “logical” item.

The execution phase, supported by the Process Engine component, enables the automatic generation of batch scripts that reflect the rules stored in the repository and the network scheduling definitions, based on execution priorities that take into account the relationships between the files and ready to be executed according to the required frequency and various strategies, with particular attention given to the adherence to the principles of computing efficiencies.

Lastly, Daisy is ready to finally integrate with the Data Masking function of Daisy DM, which is its natural complement.



SCP - Strategic Consulting Partnership Pty Ltd
PO Box 1529 North Sydney
NSW 2059 Australia
Tel: +61 411 888 390 - Fax: +61 2 9977 4450
www.scpaustralia.com - contact@scpaustralia.com

