

Institutional Data Guideline – Redundant Data Repositories (RDRs)

Executive Summary

The Data Management Committee (DMC) Policy document contains a policy statement which states “Unnecessary copies of Institutional Data are discouraged”. Many copies of redundant data are in support of “analytical and reporting processes” and for the purpose of this guidance document we will refer to them as RDRs (Redundant Data Repositories) This document provides guidance on creation of RDRs, which must be done in accordance with DMC policies, and Enterprise Data Warehouse (EDW) standards. Data Custodians, as defined by the DMC, have the authority to approve and deny the creation and operation of RDRs, and the access to data in those RDRs. Creation of RDRs involves a big commitment on the part of the responsible parties to adequately resource the development and ongoing maintenance an RDR in order to maintain adequate security and usage controls. This guidance document helps explain the responsibilities of those making the commitments.

This document does not include guidance on copies of redundant data for backup and availability purposes. It is intended solely for guidance on data used primarily for analytical and reporting purposes. It is also not intended to provide guidance on small sets of relatively static data, commonly stored within spreadsheets or in saved report files, but rather on large sets of data that dynamically generate analytics or have broad adhoc user access.

Redundant data can manifest itself in two distinct components, a back end data storage component (via databases, OLAP cubes, etc.) and a front end analytical and report data delivery component. (via Business Intelligence (BI) tools, spreadsheets, pdf documents, etc.) This document provides guidance for both components of an RDR.

Rationale

Creation of any new RDR should be carefully planned in coordination with the Office of Information Management for any data originating in any official “system of record”, as defined in DMC policies, along with the appropriate Data Custodians.

The following list represents best practices and standards for the creation and maintenance of RDRs:

- Unless granted an exception by the DMC, RDRs will be created on central servers and databases approved and supported by UW Technology and the Office of Information management.
- RDRs may be created to store historical data which is at risk of loss, due to purge operations in source systems, and the EDW has not yet accommodated acquisition and storage of that data.
- RDRs may be created to support specialized computational analysis, requiring performance service levels not currently available in the EDW and the general BI reporting environments. This commonly requires specialized aggregation, segmentation and indexing schemes not provided in the EDW or BI environments which are optimized for overall general performance.
- Except for archiving and retention purposes, analytics and reports should be dynamically produced at run time from data in the RDR, and not stored as snapshots in time, using formats like MS Excel spreadsheets or Adobe PDF documents.
- Analytics and reports that are originally produced for individual and unit level access, which subsequently need to be delivered across units or departments, should not be delivered by making copies and distributed. They should be migrated to the central EDW and BI service, or modified to use the same central access and control systems as the central EDW and BI services.

|

Advisory Opinions Sought

This guideline document has been reviewed and approved by the DMC, along with all the Associate Vice Provosts in the Office of Information Management.

Concurrence

The undersigned concur with the proposed change in data access described in this Data Management Guideline.

Executive Officer Signature: Printed: Title: Date:

Chair of DMC Signature: Printed: Title: Date:

|

Document Location: <http://www.washington.edu/provost/oim/dmc/>

Attachments:

Questions regarding this document should be emailed to: dmc-support@u.washington.edu