

**Minutes of the Decision Support/Data Warehousing Constituent Group  
EDUCAUSE 2001  
Indianapolis, Indiana**

Monday, October 29, 2001, 12:45 PM – 2:45 PM

Tuesday, October 30, 2001, 4:55 PM – 6:10 PM

CG-Facilitator: Betsy Blythe, Director, Information Warehousing and Access,  
Virginia Tech, Email: [BBlythe@vt.edu](mailto:BBlythe@vt.edu)

Co-Facilitator: Lore Balkan, Data Administrator, Information Warehousing and Access,  
Virginia Tech, Email: [Lbalkan@vt.edu](mailto:Lbalkan@vt.edu)

*Attendees:* There were two sessions for this constituent group. The first session had 27 attendees from a wide variety of institutions. Most attendees are already involved with decision support/data warehousing projects. The follow-up session had 15 attendees, only 5 of whom did not attend the earlier session. It should be noted that more than 50% of those attending these sessions also attended the Data Administration constituent group held on Monday at 4:55 PM. The fact that there was such a high percentage of repeat attendance for these groups is indicative of the interest level. This, along with the wide-range of topics the group was interested in discussing, probably merits scheduling the same number of sessions next year.

*Format:* The session opened with introductions (name & institution) and desired topics of discussion. Topics were recorded and open discussion proceeded, generally following the list of topics below. There was only time to address the first several topics.

- Politics of information access
  - Critical to establish data stewardship up front,
  - Important to education on appropriate use, and to respond to abuse.
  - Must be attentive to FERPA – therefore establish a security paradigm (access based on need to know to do job) and build trust on it.
  - Institutional policy should override individual overzealous stewards.
  - Emphasis on self-service – access should be user centered and role based
  - Authentication always sticky
  - 24 x 7 access is the expectation
- Creating an analytical culture for use of the warehouse for decision making
  - When high level officials use it, others will follow
  - Demonstrate power of drill-down to identify specifics
  - A data driven culture of accountability helps, e.g., Baldrige Criteria
  - Uncover a hot project or burning question like “why do those students who start in college of engineering not finish in engineering?” and then prove how data can help address the question.
  - Make sure data is correct so you have credibility.
- Executive information/balance scorecards – Senior management’s key performance indicators

- Keep it simple – determine what are the key numbers executives want, then give it to them so they can peddle it.
- An icon on desktop gives executives personal info they can customize...now they want the same functionality relative to their organizations.
- Find out what the hot indicators are and when they are of interest (business cycles).
- Support those who want current counts of resources they manage, e.g., people, dollars, equipment, etc.
- Developing a portal an excellent way to distribute EIS.
- Data mining for academic planning
- Canned reports vs. ad hoc reporting – which works when
- Understanding historical data in terms of environmental factors/context
- Understanding and managing user expectations for the warehouse
- Making sure users use warehouse data properly
  - Need for Data Management and Data Access Policies
  - Need for organizational structure that identifies and assigns roles and responsibilities for managing data and access to data
  - Data access with relation to Legal and Federal issues
- Inclusion of non-traditional or external data sources in the warehouse
- Transformation tool features and pros/cons of various transformation tools
- Obtaining sponsors and educating sponsors
- Factors to consider for schedule of extractions – daily/weekly/monthly; multiple sources; multiple platforms
  - How much data is being extracted – some data marts are completely replaced every night, others are updated with changes
  - How much time does the ETL process take? In some cases, it must be run on a weekend in order for it to finish
  - Why would you even include day old data in a data mart? Wouldn't users use the production system for this? This depends upon how the data is being used. In some cases, even though the users want to look at a summarized view of the data, they want to be able to drill down to the detail. This need depends upon the culture of the University. When you have the detail in the data mart, it does help build credibility and trust in the data.

*Other Comments of Interest:*

There is a desire to develop sustainable warehouse processes that are as tool independent as possible. This is a challenge.

There is interest in strategies for making the warehouse the universal source for official reporting. The timing of reports varies with the user and the purpose and must be considered to provide for the appropriate point-in-time snap-shots.

There is interest in how to best separate the presentation layer from the warehouse so users can use their query tool of choice.

There is the realization that Portals will play a major role in the presentation of the warehouse to users.

The administration of end-user tools is not trivial and needs to be factored into choices.

The users view of the warehouse hinges on the front-end tool.

It is important to distinguish “display tools”, which are designed for creating very tailored reports from “ad hoc tools”, which are designed to simplify data access.

Detail data is usually needed to adequately support ad hoc analysis. Aggregates are generally not sufficient to provide the capability to drill down to detail if there are any questions. Aggregates are useful for simplifying queries when the level of the question is clearly understood and repetitively asked.

Some use an Operational Data Store (ODS) to address the business questions in cases where the warehouse doesn't have everything of interest. This is generally not an end-user solution. The groups definition of an ODS ranged from a read only copy of the production system, to data marts that were updated as soon as changes occurred. There is recognition that soon after building the warehouse comes the very real challenge of maintaining and sustaining the warehouse and also managing the continually growing number of eager users. Carrying transaction level detail definitely increases the complexity of this management.

The data management/data administration efforts are key to the success of the data warehouse. There are significant policy and infrastructure issues related to successful data management.

*Items of Business:*

The failure of some query tool vendors to reasonably price their products for educational institutions (i.e., per user pricing instead of concurrent user pricing) tends to eliminate some good tools as viable alternatives (e.g., Brio). No action item on this issue at this time.

*Informal Survey of Attendees at first session:*

Data Warehouse Databases Used by Attendees:

<b>Database</b>	<b># of Institutions</b>
Oracle	18
Sybase	0
SQL Server	1

Informix	0
DB2	1
Access	1

Source Applications Feeding Data Warehouses Used by Attendees:

<b>Application Vendor</b>	<b># of Institutions</b>
Peoplesoft	5
Banner	7
SCT+	4
Homegrown	5

Query Tools Used by Attendees: \*

<b>Query Tool</b>	<b># of Institutions</b>
Brio	4 (3 using web)
Business Objects	2 (1 using web)
Cognos	2 (1 using web)
SAS	2
Oracle Discoverer	4
Focus	2
Crystal	4

\* 5 using Web; 2 for OLAP (BO/Explorer and Cognos/Powerplay)

ETL Tools Used by Attendees:

<b>ETL Tool</b>	<b># of Institutions</b>
Informatica	1
Ascential Datastage	1
Oracle Warehouse Builder	1
Connectics	1

*Resource Information (compiled by facilitators):*

Books	<u><a href="#">The Data Warehouse Lifecycle Toolkit: Tools and Techniques for Designing, Developing, and Deploying Data Warehouses</a></u> by Ralph Kimball
	<u><a href="#">The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses</a></u> by Ralph Kimball
	<u><a href="#">Data Warehouse Design Solutions</a></u> by Christopher Adamson, Michael Venerable
	<u><a href="#">Data Warehouse and Business Information Quality</a></u> by Larry P. English
	<u><a href="#">Oracle Data Warehousing Unleashed</a></u> by Sams Publishing
Magazines	Intelligent Enterprise <a href="http://www.intelligententerprise.com/">http://www.intelligententerprise.com/</a>
	DM Review <a href="http://dmreview.com/">http://dmreview.com/</a>
	The Data Administration Newsletter <a href="http://www.tdan.com/">http://www.tdan.com/</a>
University Sites	Ron Allan's DW survey site <a href="http://www.georgetown.edu/users/allanr/dwconfig/">http://www.georgetown.edu/users/allanr/dwconfig/</a>
	Arizona State University <a href="http://www.asu.edu/Data_Admin/">http://www.asu.edu/Data_Admin/</a>
	Stanford <a href="http://www-db.stanford.edu/warehousing/warehouse.html">http://www-db.stanford.edu/warehousing/warehouse.html</a>
	University of Pennsylvania <a href="http://www.upenn.edu/computing/da/dw/Contains.html">http://www.upenn.edu/computing/da/dw/Contains.html</a>
	University of Maryland <a href="http://www.oit.umd.edu/DataAdmin/">http://www.oit.umd.edu/DataAdmin/</a>
	Georgetown University <a href="http://georgetown.edu/uis/ia/dw/">http://georgetown.edu/uis/ia/dw/</a>
	MIT <a href="http://web.mit.edu/warehouse/">http://web.mit.edu/warehouse/</a>
	Purdue <a href="http://www.adpc.purdue.edu/WAI/DSS/DSS.htm">http://www.adpc.purdue.edu/WAI/DSS/DSS.htm</a>
	University of New Mexico <a href="http://www.unm.edu/cirt/dwhpage/DWHHome.htm">http://www.unm.edu/cirt/dwhpage/DWHHome.htm</a>
	Oregon State University <a href="http://www.orst.edu/dept/computing/warehouse/">http://www.orst.edu/dept/computing/warehouse/</a>

	<p>Virginia Tech  <a href="http://www.ais.vt.edu/IWA/">http://www.ais.vt.edu/IWA/</a></p>
Web Resources	<p>Collection of articles including Star Schema 101 – Neil Raden  <a href="http://www.archer-decision.com/">http://www.archer-decision.com/</a></p>
	<p>The Data Warehouse Information Center  <a href="http://www.dwinfocenter.org/">http://www.dwinfocenter.org/</a></p>
	<p>Ralph Kimball’s Design Tips  <a href="http://www.rkimball.com/html/kudesigntips.html">http://www.rkimball.com/html/kudesigntips.html</a>          To receive Ralph Kimball’s design Tips via email          Send email to <a href="mailto:julie@ralphkimball.com">julie@ralphkimball.com</a> with subject line of ADD</p>
Conferences	<p>Data Warehouse Institute  <a href="http://www.dw-institute.com/">http://www.dw-institute.com/</a></p>
	<p>DCI’s Datawarehousing Conference &amp; Exposition  <a href="http://www.dci.com/datawhse/">http://www.dci.com/datawhse/</a></p>
Training	<p><a href="http://www.rkimball.com/">http://www.rkimball.com/</a></p>
Discussion Lists	<p>Data Warehouse List (very active list! Can be good resource)          To subscribe to the data warehousing list server:          Send an E-mail to <a href="mailto:dwlist-request@datawarehousing.com">dwlist-request@datawarehousing.com</a>          The first line of your message must be "subscribe".          No subject line is required.</p>