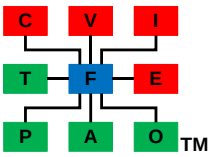

The Data Organization

Organizing Master Data

By Rainer Schoenrank

Data Warehouse Consultant

April 2023

Biography

Rainer Schoenrank is the senior data warehouse consultant for The Data Organization. He has degrees in physics from the University of Victoria and computer science from the University of Victoria and California State University Hayward. He has built data warehouses for clients such as Pacific Bell, Genentech, GE Leasing, SGI, PPFA, Brobeck, BofA, Clorox, Leapfrog and Intuitive Surgical. He can be reached at rschoenrank@computer.org.

TABLE OF CONTENTS

1. INTRODUCTION

Moving in the DDLC from the conceptual model to the semantic data model, the conceptual entities become master data tables.

The master data tables are the foundation of the enterprise record keeping and the data is required to be organized according to GAAP.

Organizing the data in the master data tables enables businesses to analyze and understand their data in a more meaningful and useful manner. The organized data can provide businesses valuable insights into their operations, customers, and markets, and make informed decisions to drive growth and success.

There are two ways organize the data in a master data table. First, the data can be organized in terms of its attributes and attribute values, such as by descriptive attributes (customer name) of the master data (customer). Secondly, the data in a master data table can be organized into a hierarchy, which is a classification of data elements based on their characteristics and relationships to each other. A hierarchy can provide a clear and standardized categorization of the data in a master data table, making it easier to understand and navigate the data, and to ensure consistency and accuracy of the data across the organization.

2. ORGANIZING MASTER DATA BY ATTRIBUTE VALUES

The data within a master data table (conceptual entity) can be organized in terms of its characteristics defined by the values of the table attributes, such as customer list by zip code, customer list by name, employee list by address, etc.

The DBMS in which the master data table resides provides tools for creating reports that provide different views of the master data as required by the business user.

3. ORGANIZING MASTER DATA BY HIERARCHY

The data in a master data table can be organized into a hierarchy, which is a classification of data items based on their characteristics and relationships to each other. A hierarchy can provide a clear and standardized categorization of the records in a master data table, making it easier to understand and navigate the data, and ensure consistency and accuracy of the data across the organization.

The table below shows the master data tables and the GAAP hierarchy associated with the master data.

	Master Data Table	Possible GAAP Hierarchy
Organizational	Position	Organization Chart
	Task	Processing System
	Product	Product Catalog
	Financial Calendar	
Operational	Customer	Sales Territory
	Worker	Sales Team
	Inventory Item	
	Supplier	
Measurement	Sales	Chart of Accounts
	Labor Time	
	Inventory Time	
	Buy	Chart of Accounts

Figure 1. Possible Master Data Organizations

For example, from the table above, positions are organized into organization charts which categorizes positions by department and function as required by APQC. The positions can be organized into many hierarchies for different purposes: a chart for the business management, another chart for the board of directors, another chart for the SEC, etc., each chart serving a different purpose.

The Sales and Buy Measurements are classified into the general ledger's chart of accounts. The general ledger chart of accounts categorizes sales by the hierarchy required by GAAP. This type of organization allows for effective data analysis and reporting, and supports better data-driven decision making.

The chart of accounts' name gives the purpose for the organization, e.g., a chart of accounts for the board of directors, a chart of accounts for the SEC, a chart of accounts for the investors, a chart of accounts for the IRS, etc.

To allow the master data to be linked into hierarchies required database tables that are not part of the master data table structure but the hierarchy tables have a many to many relationship to the master data table.

4. MASTER DATA HIERARCHY DATA MODEL

The hierarchy data model is a data structure that allows master data rows to be related to a single node in a hierarchy tree. The data model has two tables that capture the hierarchy, one table for the place in the hierarchy (hierarchy node) and one table for the relationship between the nodes (hierarchy relationship). This model follows the structure of the mathematical tree graph and assumes that there can be more than one hierarchy required to organize the hierarchy nodes (hierarchy purpose).

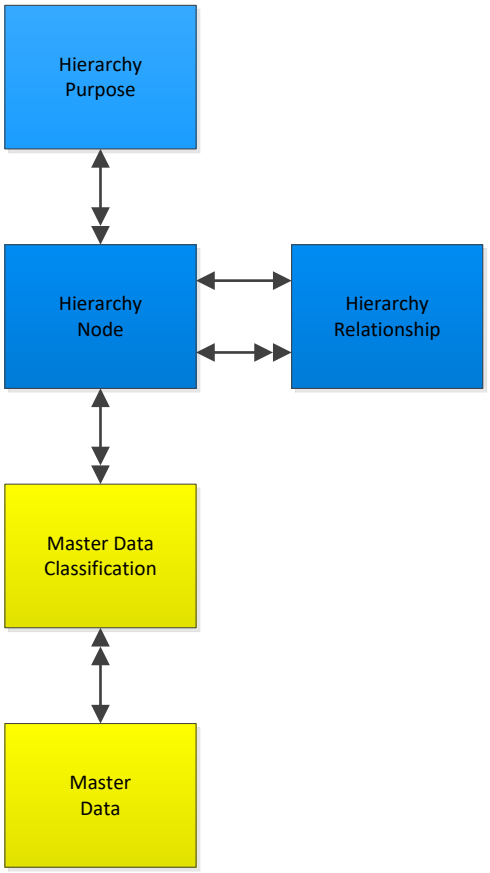

Figure 2: Taxonomy Conceptual Data Model

In Figure 2, each line represents a relationship between the entities. One arrowhead represents one item and two arrowheads represent many items.

5. ORGANIZATION CHART EXAMPLE

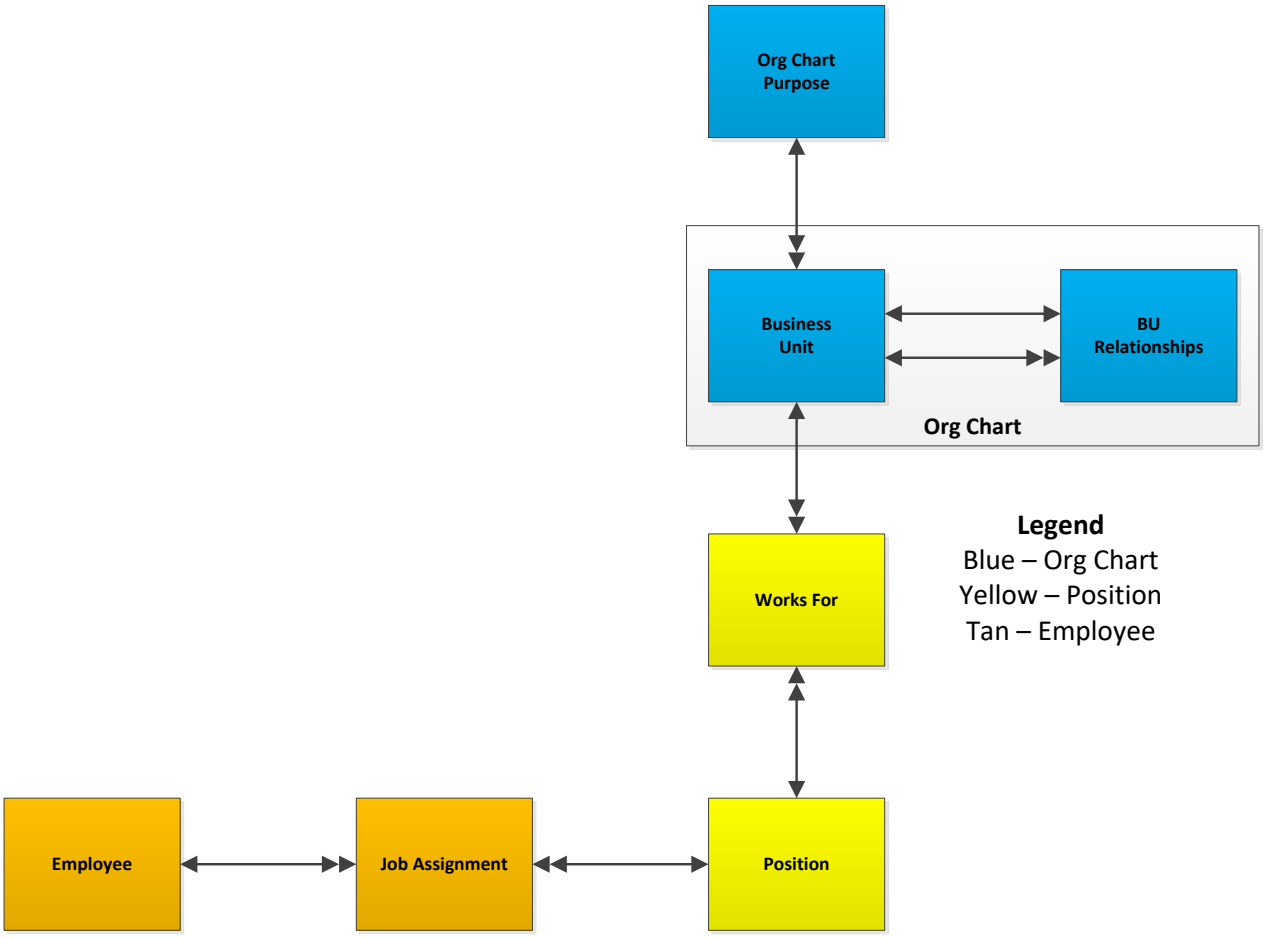

Figure 3. Organization Chart Context

6. GENERAL LEDGER EXAMPLE

The G/L chart of account is a hierarchy that organizes the money that flows into and out of the business. In most Financial Applications, the data is copied from the Accounts Receivable and Accounts Payable Application into the General Ledger. The money is transferred to the G/L once a month when we close the books for that month. Then the amount in a G/L account can be adjusted using journal vouchers that transfer money between accounts.

This process has several drawbacks. One, the G/L is only up to date at the beginning of each month. Two, the money in each G/L account has multiple sources which are hidden by the transfer process. Three, the use of journal vouchers to move money between G/L accounts obscures the corrections required to the transfer process. Four, there is no data lineage or traceability for the individual amounts recorded in the Accounts Receivable and Accounts Payable applications.

These drawbacks can all be addressed by changing the data models used for the General Ledger. Instead of having a separate database for the General Ledger, it becomes an integral part of the Enterprise Data Warehouse (EDW).

Building the conceptual data model from the Sales portion (Accounts Receivable) of the EDW, the Purchasing portion (Accounts Payable) of the EDW and organizing the recorded dollar amounts into a taxonomy (Chart of Accounts), we get the following conceptual data model:

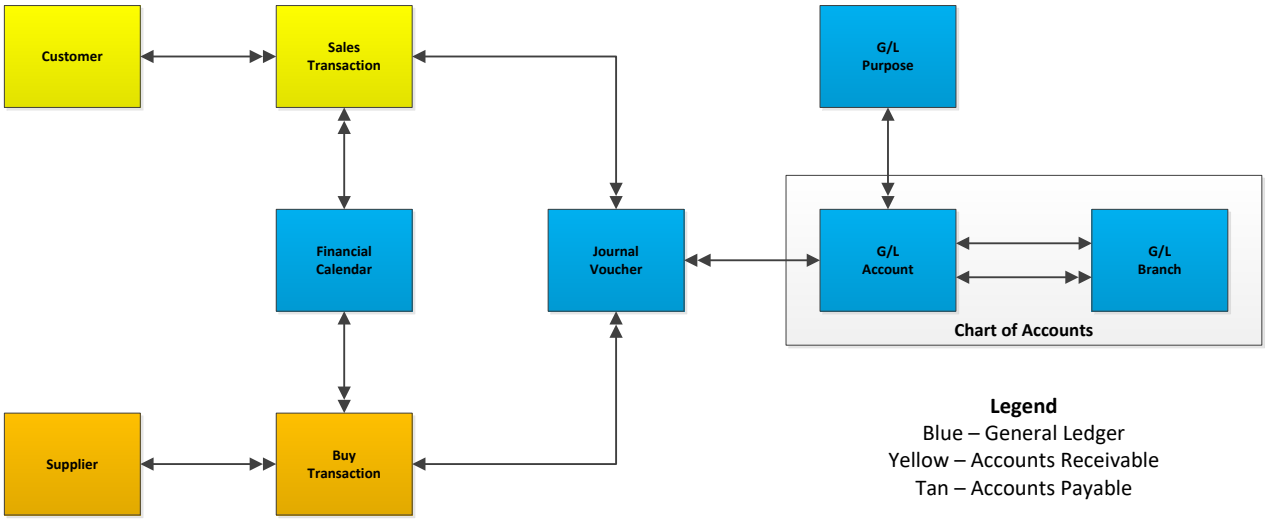

Figure 4. General Ledger Context

This data model has the advantage of making the data lineage transparent, the G/L is always up to date, the mechanism to assign a transaction to a G/L Account is explicitly given in the Journal Voucher, and there is no duplication of data into another application database.

7. REFERENCES

APQC Helping Organizations Work Smarter, Faster, and with Greater Confidence | APQC

"Classification Theory", Encyclopedia Britannica, 15th Edition, 1974, Vol. 4, pp. 691-694

Date, C.J., *The Relational Data Dictionary*, O'Reilly Media, Inc., Sebastopol, CA, 2006

GAAP Generally Accepted Accounting Principles (United States) - Wikipedia

William Kent, "Data and Reality", North Holland Publishing Co., Amsterdam, 1978

W. S. Jevons, "The Principles of Science", Chapter 30, Dover Publishing Inc., New York, 1958

Robert L. Glass and Iris Vessey, "Contemporary Application-Domain Taxonomies", IEEE Software, July 1995, pp. 63-78

