

TITLE: "UNDERSTANDING GROUP DATA ELEMENTS"

by Tim Bryce
Managing Director
M. Bryce & Associates (MBA)
P.O. Box 1637
Palm Harbor, FL 34682-1637
United States
Tel: 727/786-4567

E-Mail: timb001@attglobal.net

WWW: <http://www.phmainstreet.com/mba/>

Since 1971: *"Software for the finest computer - the Mind"*

"Group Data tells us a lot about what objects are important to an enterprise."

- Bryce's Law

INTRODUCTION

In the past I have discussed the need to manage data (and all information resources) as a valuable resource; something to be shared and reused in order to eliminate redundancy and promote system integration. Now, our attention turns to how data should be defined. Well defined data elements are needed in order to properly design the logical data base as well as developing a suitable physical implementation.

First, let's understand how data is used; there are three purposes:

- Indicative Data - to identify the important objects needed to run a business, be it something as tangible as a customer, employee, product, part, etc. or as intangible as an event, such as a shipment, a billing, or transaction.
- Descriptive Data - alphanumeric values that are not strong enough to identify an object, but convey important business facts about an object, such as names, addresses, text, codes, etc.
- Quantitative Data - numeric values that are either calculated or are calculable. Measurements and computations are typical examples: "Net-Pay," "Quantity Ordered," "Elapsed Time," "Percent of Gross," etc.

Basically, there are two types of data elements: Primary and Generated

PRIMARY data refers to data in its virgin state; as introduced to the system from an external source (such as a person or department). "Source" defines who is responsible for entering the data to a system, and who has ultimate authority for the definition of the data element its meaning).

GENERATED data refers to data that relies on other data elements in order to produce the necessary result. This type of data can involve elaborate calculations and algorithms (e.g., $DD-1 + DD-2 = DD-3$). "Net Pay," "Balance Amount" and "Percent Complete" are some examples of calculated data.

GROUP DATA

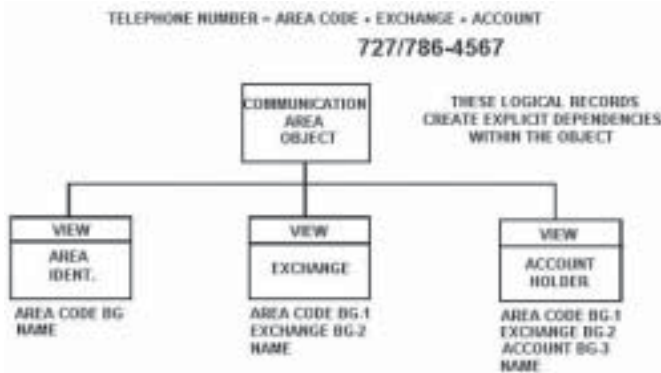
Most data elements easily fall into the two categories of Primary and Generated, but there is a seeming anomaly that often confuses people, namely "Group" Data Elements, such as "Credit Card Number," "Telephone Number," "Check Number," etc. Many companies treat it as a Primary value when, in reality, it is often a Generated value; let me explain.

Group data is actually a concatenation of multiple data elements. For example, "Credit Card Number" typically consists of "Financial Institution ID," "Bank Region Number," "Bank Branch Office ID," and "Account Number." The "Customer Number" on a power bill may represent such things as: "Primary Power Station," "Sub-Station," and "Account." To a communications company, a "Telephone Number" has considerable meaning and represents "Area Code," "Exchange," and "Account." There are many other examples of "group" data: such as product identification codes, bank codes, insurance policy numbers, etc.

It is a common misconception that group data elements should be used for basic groupings (keys) in logical records; THEY SHOULD NOT! Group data is used as a convenient means to describe dependencies between primary data elements. As such, a group data element provides tremendous insight into objects and views. For example, consider the objects and views associated with a "Telephone Number."

(continued on page 2)

(continued from page 1)



Observe the dependencies between the three views. Each has an impact on the others. Should the first view be deleted, the second and third views will also be deleted. From this perspective, the basic grouping defines dependencies and eliminates the problem of multiple occurrences.

Data elements such as "Telephone Number" and "Credit Card Number" should only be defined as group items if they truly represent a concatenation of indicative data elements representing objects. For example, "Telephone Number" is a valid group item to identify a "Communications Area" and its views for a telephone company. But if "Communication Area" is not a pertinent object to your business, there is little point in defining it as a group item. Instead, it is a simple primary value.

Another hint as to whether a data element is primary or generated, is whether the company assigns the values to it, if they do not, then it is probably a primary value.

CONCLUSION

So why do companies have group data elements? Because it is a convenient means to quickly identify the objects and views important to the business. For example, it allows companies to "roll up" data; e.g, the number of accounts within a specific area. Further, a group data element may not be suitable for logical data base design, but has been found to be a useful means in the design of the physical file design (expedites accessing data).

How many Group data elements does a company truly use? Not as many as you might think. For example, Credit Card Numbers, Bank Codes, and Telephone Numbers are primary values in my business. However, a "Customer Number" is a group item consisting of a "Contract" and "Location." It all depends on whether these

are important objects your company manages, hence, "Indicative Data" is devised and assigned to uniquely identify it.

To the average programmer, there is little concern for how data is defined other than to assign a suitable program label. However, data definition is a very important consideration to Systems Engineers and Data Engineers who are charged with building major integrated systems. Think about it.

For more information on Data Definition, see:

"PRIDE"-Data Base Engineering Methodology (DBEM)
<http://www.phmainstreet.com/mba/pride/db.htm>

Also see "The Anatomy of a Data Element" at:
<http://www.phmainstreet.com/mba/pride/dbmeth.htm#anatomy>

END

"PRIDE" Special Subject Bulletins can be found at:

<http://www.phmainstreet.com/mba/mbass.htm>

They are also available through the "PRIDE Methodologies for IRM Discussion Group" at:

<http://groups.yahoo.com/group/mbaprider/>

You are welcome to join this group if you are so inclined.

The "Management Visions" Internet audio broadcast is available at:

<http://www.phmainstreet.com/mba/mv.htm>

Also, be sure to read Tim's Blog at:

<http://blogs.ittoolbox.com/pm/irm/>

"PRIDE" is the registered trademark of M. Bryce & Associates (MBA) and can be found on the Internet at:

<http://www.phmainstreet.com/mba/pride/>

Copyright © MBA 2005. All rights reserved.