Data Modelling User Guide

Prepared By: Neville Turbit
Version 1.0
1 Feb 09

Table of Contents

Document Origin

No.	Author	Department
1	Neville Turbit	Project Perfect

Change History

Version	Date	Changes
1.0	1 Feb 09	Initial Version

Data Modelling Basics

Overview	This document is not intended to make you a data modeller. The purpose is to provide just enough information to enable a facilitator to manage a data modelling workshop.
Physical v Logical	There are two types of data model. • Physical Model is what actually exists in the system database • Logical Model is how the users may understand the data is stored
Differences	The physical model may be different in a number of ways. Some of these differences may be: • Inclusion of system data such as transaction time which are not seen by the user • Construction of the tables to optimise performance • Temporary tables created to store information
Entities	Entities are the people, places and things we want to keep track of. Example may be: • Clients • Suppliers • Products • Prices • Orders
Attributes	Attributes are the individual components of the Entities. For example for Client, they may be: • First Name • Last Name • Address • Phone Number • Company
Primary Key	The primary key is the unique identifier for each Entity record. Almost always it is a number. Examples may be: • Client Number • Supplier Number • Product Number • Order Number

Data Modelling Basics, Continued

Relationships   Relationships are how we tie the Entities and Attributes together. For example, Orders have Products. A record for an Order will have an order number.

Order No (PK)	Customer	Date
1	765	2/1/08
2	432	4/1/08
3	111	10/1/08

There will be also a detail order Entity that links the orders and the products.
If the order number was 1, the detail record would have a number of records that relate to order 1. Each of those records would have a product number and quantity.
For example:

Order No (FK)	Product No (FK)	Qty
1	999	2
1	888	4
1	777	1
2	999	7
3	888	1
3	777	2

Foreign Keys   The cross reference fields in the example are the Order No and the Product No. These are foreign keys. In other words they relate to the primary key in the orders and products Entities.
The second table is incomplete as it does not have a primary key. There would need to be a primary key called something like OrderDetail Number.

One to Many   One Order, can have many OrderDetail records. One OrderDetail record can only have one Order record. The relationship is one Order record to many OrderDetail records.
You can have 'one to many', but not 'many to many'. If you had many to many, it just wouldn't work. You need to ask yourself if A can have more than one of B, and can B have more than one A record. If the answer is 'Yes" to both, you need to create an entity between the two entities.
In the example above, one Order can have many Products, and one Product can have many Orders. To resolve this we have an OrderDetail record that means one OrderDetail record can only have one Order and one Product.

Data Modelling Basics, Continued

Entities and Attributes

It is sometimes difficult to decide at the logical level if you are discussing an attribute or entity. To take an example is a Client an entity? If one of the attributes of a Client is a Phone Number, and a client can have 1,10 or 20 phone numbers, do we want to make Phone Numbers an Entity. The alternative is to make 20 slots to hold phone numbers. That almost guarantees that you will immediately get a new client with 21 phone numbers. The solution is to ask if there is any information about the client that will have multiple instances. Addresses may be a candidate for a separate entity if the client has more than one location.

Entity Workshop

Overview   If you need to hold a workshop to identify entities, this section will provide some guidance. It will require the people to have some understanding of the terms in the previous section. An example of the output is included as Appendix A.

Brainstorm   The best way to start is to ask:

> "What are the people, places and things we are interested in tracking?"

Write down as many as you can on a white board without getting into a discussion. Sorting can come later. Just have people throw them at you as quickly as they can think about them.

Some will be attributes of Entities and some will be irrelevant but get them all written down.

List the Entities   You can now get the brainstormed list into some sort of order. A data modelling worksheet is available. The order is not important. As each one is entered check to make sure it is not an attribute of another entity.

Including attributes   It is also useful to put down one or two attributes to identify what information is included. For example if the Entity is called 'Other Information', what does it mean? If we add a few attributes such as 'Source of Document' and 'Previous Contact' it makes it clearer as to what the entity is all about. Don't however try to list all attributes.

You can also use the notes to add some description.

Examples   If you do put in attributes, try to give some examples. For example in the situation above, what does 'Source of Document' really mean? If the example says 'Mail', 'Email', 'Fax' it is much clearer. There is less likely to be confusion from someone who was not present. A person not present may interpret source to mean something like who sent the document in the first place rather than how it was received.

Make sure the examples display the full range of data. For example, if a part number can be either numbers or a mix of numbers and letters, show both. Give examples of size. For example if numbers can be 3 to 10 characters, give examples of both.

Input from functional decomposition   You can also add the input from the functional decomposition. These are entities that were mentioned during the functional decomposition workshop.

Entity Workshop, Continued

Review functionality

By going through the functional decomposition, line by line, some more entities may be discovered. Think about what information each function requires. For example a function to add a client may mean that you have to input a credit classification for the client. That may mean a 'Credit Classification Type' entity.

Attribute Workshop

Overview   We assume we are starting with a listing of entities. These will have some attributes included.   We now want to create a listing of all attributes.

Process   Working through the entity sheet, add rows to list attributes. The process will flesh out the entity and may in fact lead to either the entity being split, or combined with another entity. Look for one to one relationships with other entities. That almost always means it is one entity.

Focus   The focus is not to build a database. The focus is on capturing information that may be needed to build the system. How it is put together is the responsibility of the database designer. They will take your work and manipulate it in a technical sense – i.e. turn the logical design into a physical design. Do not get hung up deciding if you have the information in the right slot. Just focus on getting the information in somewhere.

Relationship Workshop

Overview   The purpose of this workshop is to identify how information is connected, and what other entities may exist.

Process   Starting with the first entity, run through each of the other entities asking the questions:

- "Are these directly entities connected in any way? " Something like 'Clients' and 'Shelf Locations' are not directly connected. 'Clients' buy 'Products'. 'Products' have a 'Shelf Location'. They are indirectly connected so you ignore them.
- "Can one of A have more than one of B? Can one of B have more than one of A?" If the answer is "Yes" and "Yes" you need to have an in-between entity called AB. 'Orders' and 'Products' have a many to many relationship so you need an 'OrdersProducts' entity.

If you have 20 entities and you start with 1, after you get to entity 20, go back and start at entity 2. You do not need to check it with 1 as you have already done this in the first pass. Start checking 2 with 3 and work through to 20 again. Next move to 3 and check 4 to 20.

When you have completed all the entities, you will have completed the relationship workshop. It may be tedious but just remember every time you start a new loop it is one less to check.

Appendix A

Overview   The example below is after an entity workshop. It does however show some attributes and relationships (foreign keys)

Data Modelling Worksheet

Name	Attribute	PK/FK	Ref. Table	Note	Example
Client	Client No	PK	No	Only credit customers, not cash sales	10001, 10002, 10003
	Surname				Smith
	First Name				John
	Credit Rating	FK		FK to Credit Rating Entity	15 day, 30 day, 45 day
Products	Product No	PK	No	Both our own products and ones we sell as distributor	123456
	Distributor's Product No			The number we use to purchase from the manufacturer	J97, 4444444, PHR0964329