

Seven Deadly Sins Of Data Modelling

The discipline of data modelling has been with us for over 3 decades, but common issues prevent many organisations from realising benefit. Chris Bradley explores seven common *non-technical* mistakes.



Background

Data Modelling has been around since the mid 1970's and for the last 30 years I've been lucky (or not) to be involved in this field. When it first emerged it promised so much

"A single consistent definition of data"

"Master data records of reference"

"Reduced development time"

"Improved data quality"

"Impact analysis"

.....and so on.

Who in their right mind would NOT want to have these benefits realised in their organisations? Surely they are no brainers?

Yet why is it, some 3 decades on that in many organisations the benefits of data modelling still need to be "sold" and in others the big benefits simply fail to be delivered? Is there something that WE the Information Architecture community are doing wrong? I'd like to suggest seven key areas where we are committing deadly sins!

1. Lack of focus on benefits

Why is data modelling undertaken? It's not for the benefit of the modellers or the information community themselves. If you have one take away from this section, please take this

"So, just what are the benefits of doing this Data Modelling thing?"

NOBODY OWES YOU A LIVING

We are all (or should be) ambassadors for our organisations and the wider Information Management Community. We need to recognise that this is a business, not an academic exercise. So, cutting to the chase, what are the things we need to be aware of?

Project requirements vs. big picture

Much as though we may like everybody to develop their models with the whole enterprise in mind, it's critically important to recognise that most model consumers will be in a "project" environment. Introducing "enterprise wide" dependencies on them may simply be unpalatable. Thus recognise that there may well be sub-optimal approaches that projects need to employ. However, you as an Enterprise wide Information Architect must have your eyes on the wider picture.

Reward drives behaviour

Leading on from above. If project teams are incentivised to bring their projects in on time and budget, and you start introducing actions that "benefit the Enterprise" but there's no bonus or reward in it for them, then don't be surprised when you're turned down. I once worked in an organisation where project teams were rewarded for developing data components for re-use and for re-using existing data components. This company consistently had the best models I've ever seen

WIIFM

Really the same as above, but I like the acronym. WIIFM =

"What's In It For Me?" Whenever you're asking someone to do something for "the good of the Enterprise" always put yourself in their shoes and apply the WIIFM test!

Metrics

Collect metrics on productivity, performance and the like. You WILL be asked for this at some point

"So just why are we doing this data modelling stuff, exactly what benefits have we realised?"

Space and time don't permit me to fully expand on this within this article; however I will be writing another article about quantifying modelling benefits

Evidence and quotes

In addition to the metrics, make sure you constantly collect evidence and where possible quotes from satisfied senior users in your organisation. Make this a standard part of your engagement with projects – don't wait until 9 months after you've finished – maybe the enthusiastic manager has moved on by then!

"We reduced the number of bespoke messages and re-development in this project by 22% saving almost \$1.5m"

Sustained improvement

Look at applying something like the Information Management Maturity Model. Aim for continual improvement. Even consider professional certification for Information professionals (more on this later).

2. Forgetting the purpose

Why are we producing the model – what’s its purpose?

Must we always follow a top down approach?

What about bottom up or even middle out?

Do we need to incorporate History or the “as-is” position?

These and a whole host of other questions need to be asked when you embark upon a modelling project. And remember that a data model is not JUST for DBMS development. There’s an entire article devoted to this here and here. Look at this survey we undertook at a large client with over 300 modellers, the main reasons for producing a data model make interesting reading...

Why produce a data model?

Top 10 reasons in a large multinational organisation (2007/8)

1. Capturing Business Requirements
2. Promotes Reuse, Consistency, Quality
3. Bridge Between Business and Technology Personnel
4. Assessing Fit of Package Solutions
5. Identify and Manage Redundant Data
6. Sets Context for Project within the Enterprise
7. Interaction Analysis: Compliments Process Model
8. Pictures Communicate Better than Words
9. Avoid Late Discovery of Missed Requirements
10. Critical in Managing Integration Between Systems

“Data modelling is not easy – remember it’s not just about drawing a picture”

3. Language and intellectual snobbery

Past baggage

Unfortunately in many organisations the term “modelling” often has some historical baggage associated with it. Frequently this stems from earlier exercises where lots of workshops were conducted and no apparent benefit for the user realised. In addition to being great Information architects, we must also be great communicators and ensure we give regular feedback on the project, especially to the people who have given us their time. In these days of Intranets and SharePoint, there’s simply no excuse for failing to keep a project blog or SharePoint going.

Inappropriate language

I don’t mean Serena Williams’s outburst to the line judge (although on some projects the “F” word is in too regular use). I’m referring here to discussing inappropriate levels or types of information to different audiences. Don’t show DDL, Tables and Physical models to a Business user. Likewise don’t show a high level conceptual model to a DBA. Please make your message and language relevant to the audience.

Banish bigotry

The most embarrassed I ever felt about being in the Information management profession was in a previous company when I witnessed a colleague telling a business user why he didn’t like the notation we’d adopted (Barker as it happens) and how “in his opinion the UML class model notation was far superior”. It’s difficult enough in many cases to get business colleagues to the table to discuss this Information stuff. I’d rather we produced ANY type of data model than argue with the business over the type and notation. Methodology bigots and their dogma must be banished!

4. Discipline

Dumbing down

It’s not just about drawing a picture! Remember there’s way more to data modelling, particularly the business engagement.

Don’t forget the metadata

Related to the point above, the metadata is key to making a model that is useful. Include descriptions, notes use cases and examples.

Training and the right personnel

I never fail to be surprised at how many staff (and scarily many are from large international consultancies) are placed onto projects as “data modellers” yet really don’t understand the basics of why models are required, the levels of models, yet alone how to effectively engage and develop them. There are many excellent training courses available. And now professional certification (e.g. DAMA Certified Data Management Professional) can help weed out the interlopers too.

Relevant standards and guidelines

Use them! If you don’t have any, develop some. There are plenty of good ideas and standards that you can borrow from.

Communication

The majority of problems in this field stem from poor communication. Give constant feedback. Speak to business users in their language. Don’t try and walk through a 150 entity model in one session, break it up into small chunks. These “soft skills”

are (IMHO) the biggest area where improvement is required in the data architects community.

Honesty

Be honest with yourself and your business colleagues. Most of the time this modelling stuff isn't easy. Ask them to help you understand their business.

5. Inappropriate positioning

Please don't do modelling for its own sake. Understand the purpose of the modelling effort you're embarking on.

Is for communication only?

Is it for development of a physical DBMS?

The purpose and level required of the model will affect the way you approach its development and the end result.

All too often we see data modelling performed in isolation, the so called information silos of DM, PM, DBA etc. Also it's often left until too late in the lifecycle to be useful or the methodology bigots take way too long in completing a "useful" model by giving too much focus on final 20% to be "theoretically perfect"

Regarding the "fight for survival" it's also sadly common to see data modelling considered "an overhead" or cross departmental charging for a data modelling infrastructure. These are sure signs that the organisation hasn't truly bought-in to the concept.

Finally regarding models themselves you may see (or in this case not see actually) hidden or unpublished models. What's the point of developing a model then NOT sharing it! Sadly you'll frequently encounter limited re-use of modelling concepts (just how many definitions of CUSTOMER or PRODUCT do we really need!). This is in many cases due to Projects being left to their own devices "sorry we can't engage with Data Management on THIS project - the train has already departed". It's often a consequence of the DM function not being resourced appropriately thus models not being subject to peer or cross-domain review and / or the DM function not being seen to add value.

6. Failing to adapt

As outlined at the start of this article, modelling's early days focused on the development of DBMS's. But now the application landscape of most businesses is not made up simply of bespoke solutions. They use ERP packages, BI and DW, XML message based systems, SOA and more. Data modelling does and should have a place in these but unfortunately the majority of organisations (and practitioners) haven't caught up. Also modelling needs to be utilised in the "non traditional" data areas including Master Data, Transaction data, MI/BI, and Unstructured data.

There are many useful data modeling tools on the market, however selecting the one that's most appropriate for your

organisation is rarely done. Example criteria for examining data modeling tools suitable for high-level models can be found here. Good usage is more important than choosing the academic "best".

Additionally, I believe the industry has been done a disservice by many ERP package vendors. Just how many provide a Logical Data Model with their packages? This makes the integration of the COTS package within the overall information architecture of your company a more difficult task. It still should be done however.

7. Square pegs and round holes

Many Information Management practitioners have become walking TLA factories (TLA = Three Letter Acronym). We talk of IM, IAM, DM, MDM, EDM, EII, CDI, SOA and so on. But do we take the time to explain to our

colleagues what these really mean and why they are of significance?

We need to have the right people in the right role. In today's business climate simply being a good technical modeller isn't enough. We have to be good communicators. We have to demonstrate genuine professional credibility. Certification is coming at last and can help with this.

But do remember what I said at the outset - nobody owes us a living.

It's up to us to successfully engage with the business. We should communicate our successes, create communities of interest and let people know why this is undertaken and how it will benefit them.

Fundamentally it's up to us!

So what's our part in fixing this? Well, that's the topic of another article 😊



"It's up the Data Architecture community to successfully engage with the business"

Chris Bradley has 30 years of Information Management experience. He is a recognised thought-leader in the Information Management field, is a published author (Data modelling for the Business- A handbook for aligning the Business with IT using high level data models), a regular columnist for the BeyeNETWORK and an International conference speaker. Chris also sits on industry advisory panels. Chris leads IPL's Business Consulting practice and can be contacted at BusinessConsulting@IPL.com.