

Why is a Corporate Data Model Different?

Dagna Gaythorpe

<http://www.seshat.com>



DAMA UK

Swindon, March 20th 2007

Agenda

- What is a Corporate Data Model?
- How does it fit in with Data Management?
- What are the benefits?
- How different is it?
- Summary
- Questions

Acknowledgement

I would like to thank the British Museum for their help with this presentation.

What is a Corporate Data Model?

A model of the underlying structure of the data that is important to the business.

- Strategic
- Tactical
- Legal/compliance
- Reused

What is a Corporate Data Model?

Database independent - the data modelled in it may not be held in a database!

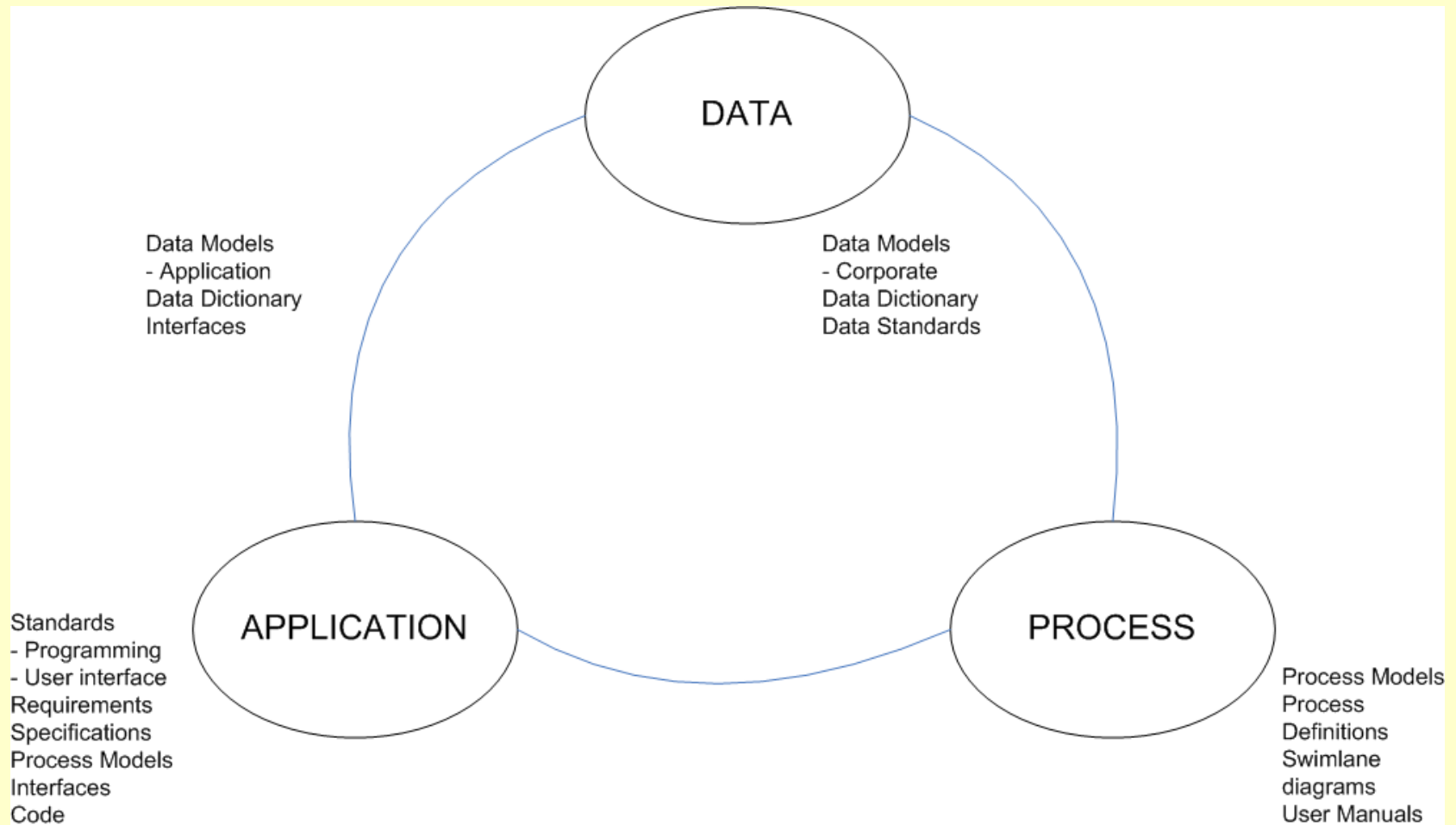
(But usually is).

Has been equated to Town Planning, as distinct from building design. (Simsion and Witt)

What is a Corporate Data Model not?



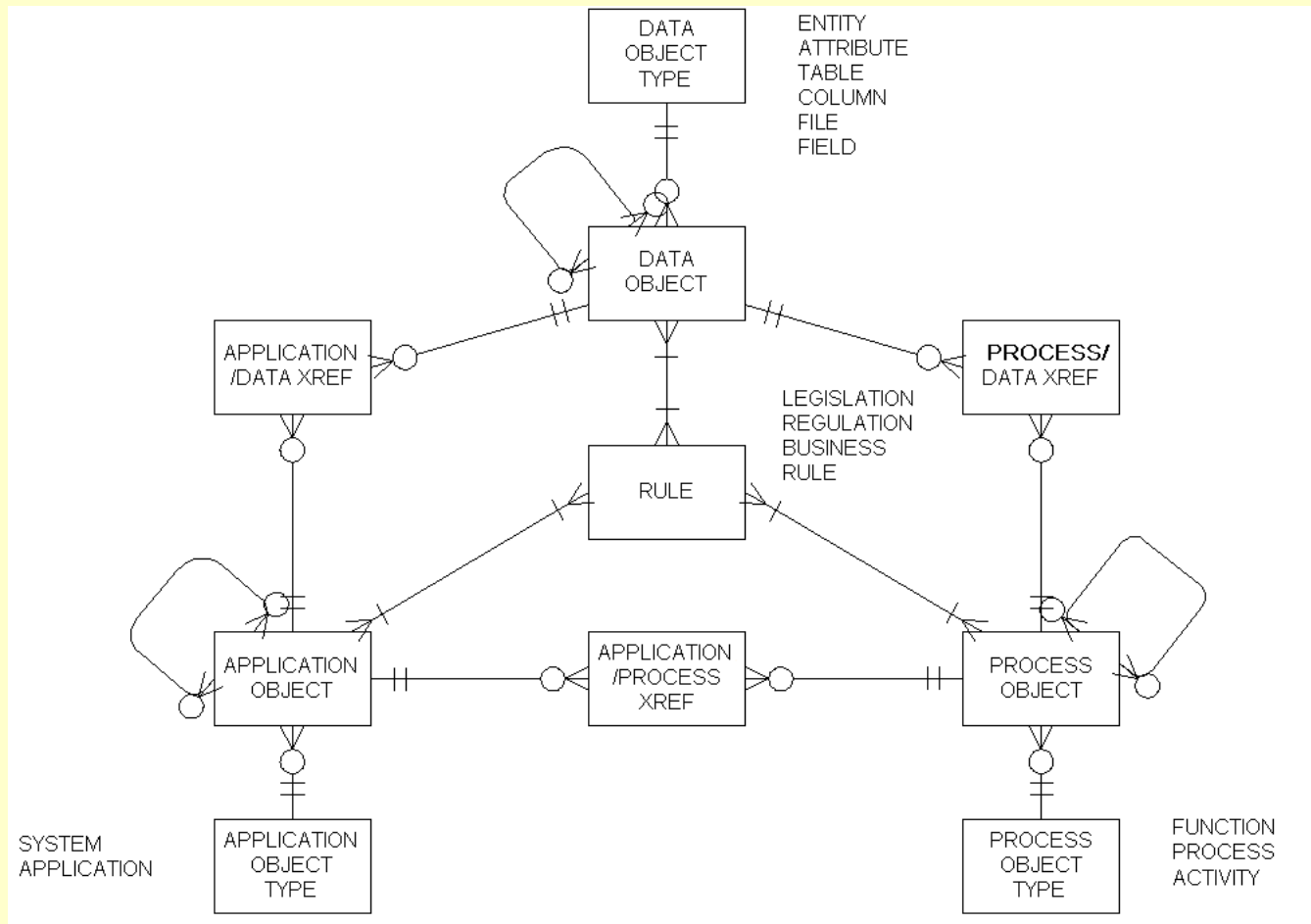
How does it fit in with Data Management?



How does it fit in with Data Management?

- Level 2 in the Zachman Framework
- Modelling the business
- Starting point for Application, Programme and Project models
- Useful for evaluating packages

How does it fit in with Data Management?



How does it fit in with Data Management?



What are the benefits?

- Reuse (after the first few projects)
- Faster (from reuse)
- Better quality (each wheel only gets invented once)
- Cheaper (because you are reusing - so you need less “new” money)
- Integration and interfaces - reduced transformation in ETL
- Comparing with supplier models for package selection
- Compliance!

What are the drawbacks?

- Extra work at first
- Can be hard to convince people that it is useful
 - “My project is special and different”
- Can be hard to convince people to use it
 - “My project is special and different”
 - We hear that a lot...
- Can be seen as an enforcement rather than an enabling function

How different is it?

- More than just a data model - it links to
 - Data flow diagrams
 - CRUD matrix
 - Data dictionary/Corporate Glossary/Object Class Hierarchy

How different is it?

- Subject areas may cover
 - Business areas
 - Core entities
 - Applications

Depending on how much time you have - all these add to the maintenance!

How different is it?

- Not as detailed as a project model
- Generic rather than specific
- Wider scope
- Used to record things that are:
 - Required for compliance
 - Critical to the business
 - Used in more than one place
- Not a database design

How different is it?

Data formats may be missing...

What do you do when you have multiple formats across systems?

- Leave it out (if the tool lets you)
- Put in a default (as long as you remember to tell everyone that CHAR(999) is a placeholder!)
- Put in the format you want everyone to use in the future.

How different is it?

Don't show history.

If you have MDM, record the master system, and the place where history is stored, in the notes about the entity.

Make a note of what history should be held.

Leave it out of the model! (But you could create a subject area containing the things that need history.)

What is it for?

It is a model of the business

(Who probably have a pretty good idea of how they work)

It has subject areas

which can cover a business area (department, function, process, etc)

or an application (or programme)

(this is a useful bit).

More about the useful bit...

When a new project starts, how long does it take to get the basic, 'starter', data model?

- A week?
- A day?
- 2 hours?
- 30 minutes?

More about the useful bit...

Using a Corporate Data Model as a base, I have produced the starter model in 2 hours.

The meeting to decide what entities were relevant typically took 30 minutes, and then the model had to be created from the CDM.

Faster, Better, Cheaper

Pick any two...

Faster, Better, Cheaper

But if you have already done a lot of the basic work,
and can reuse it
and only have to invent the new stuff...

Then you can have it all for new projects.

Faster, Better, Cheaper

Faster - because you have one you did earlier.

Better - because it has already been checked by other projects, and is probably already in use somewhere in the organisation.

Cheaper - because the project only has to pay for the new stuff (plus a bit for getting the old stuff out of the central model).

Faster, Better, Cheaper

And if you are reusing data formats, then you reduce the amount of data conversions, checking, reformatting and converting.

Simpler interfaces, faster development, better quality...

Faster, better and cheaper again!

Summary

- Faster, better cheaper
- Once it gets underway
 - the first few times will be slower and more expensive!
- Integration and interfaces - reduced transformation in ETL
- Compliance!

Summary



Why is a Corporate Data Model Different?

Questions?

Contact Information

Dagna Gaythorpe

<http://www.seshat.com>

dgaythorpe@seshat.com