# Enterprise Data Modelling: Developing an Ontology-Based Framework for the Shell Downstream Business

Matthew West – Shell

Chris Partridge – BORO Centre

Mark Lycett – Brunel University
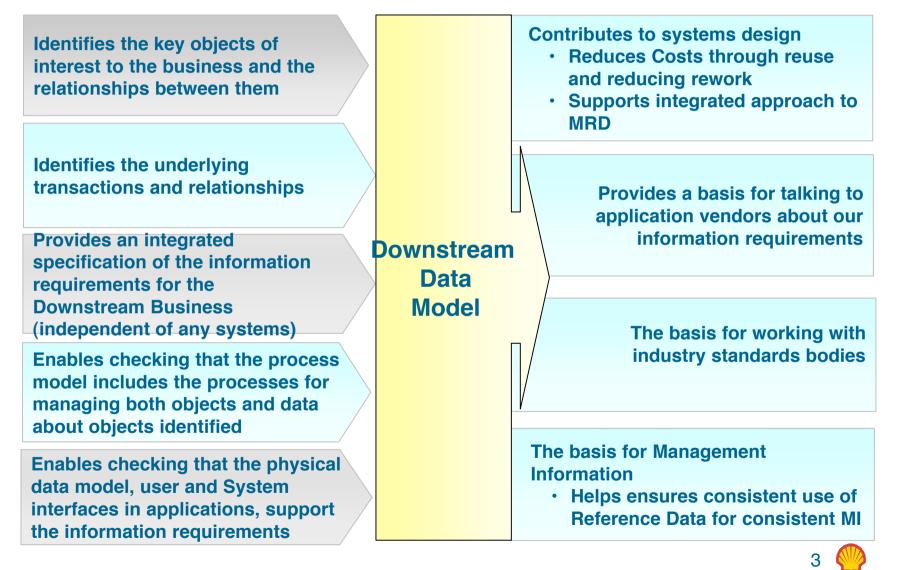
# The Context

- Shell Downstream

    – From Oil Tanker to Petrol Pump

    – c80,000 employees

    – More than 100 countries

    – Downstream One an initiative to globalize Shell's Downstream business around a single set of processes and systems

    – Consistent Reference Data a critical element of business integration

# What is the purpose of Shell's Downstream Data Model (DDM)

**Identifies the key objects of interest to the business and the relationships between them**

**Identifies the underlying transactions and relationships**

**Provides an integrated specification of the information requirements for the Downstream Business (independent of any systems)**

**Enables checking that the process model includes the processes for managing both objects and data about objects identified**

**Enables checking that the physical data model, user and System interfaces in applications, support the information requirements**

**Downstream Data Model**

**Contributes to systems design**
- **Reduces Costs through reuse and reducing rework**
- **Supports integrated approach to MRD**

**Provides a basis for talking to application vendors about our information requirements**

**The basis for working with industry standards bodies**

**The basis for Management Information**
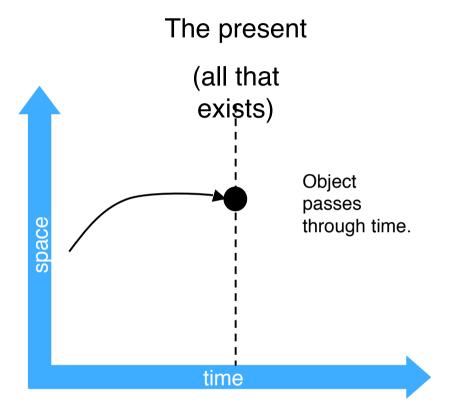- **Helps ensures consistent use of Reference Data for consistent MI**

3

# Foundations

- Computationally Independent Model (CIM in Model Driven Architecture)

- Ontologically based

- Epistemological gloss – to prevent requiring information that is not known or needed

- Starting points and Methodologies

  - EXPRESS

    - ISO standard data modelling language (ISO 10303-11)

  - ISO 15926

    - Abstract data model designed to support large scale integration

    - Based on 4D paradigm

  - Developing High Quality Data Models (HQDM)

    - A Shell developed data modelling methodology with a "middle out" approach

  - The BORO Methodology

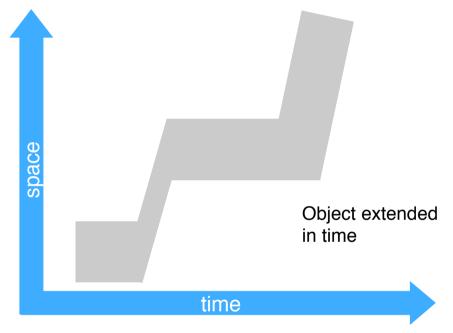    - A reengineering methodology that starts from a clear ontological foundation and re-engineers data.
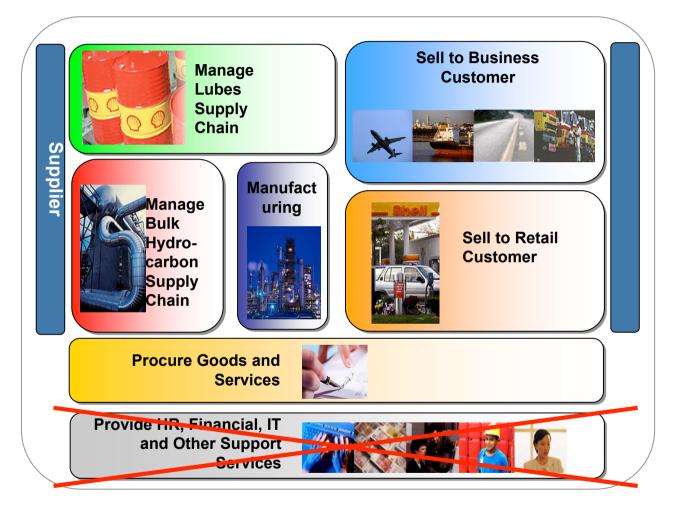
# 3D    vs 4D + Extensionalism

### The present

### (all that exists)



Object passes through time.

space

time

### The past and the future exist as well as the present



Object extended in time

space

time

1. Physical objects do not have temporal parts.

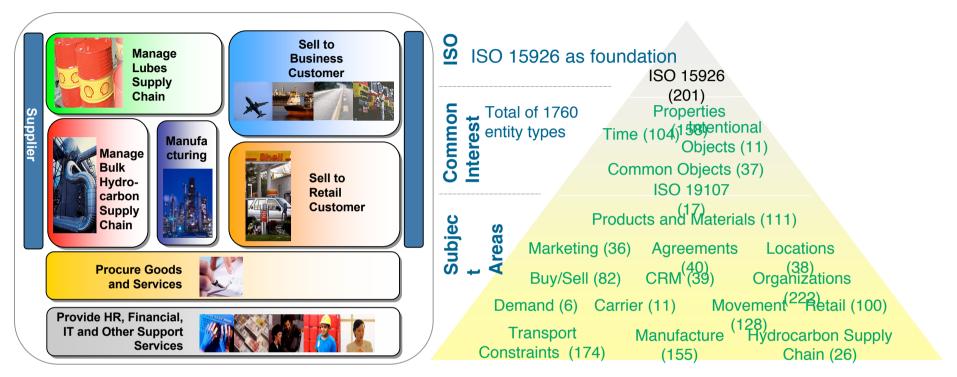2. Different physical objects may coincide (non-extensional).

1. Individuals extend in time as well as space and have both temporal parts and spatial parts.

2. When two individuals have the same spatio-temporal extent they are the same thing (extensionalism).

5

# Scope



- Depth – from metaphysical choices to business concepts

# Changes in Approach



**ISO**
ISO 15926 as foundation

**Common Interest**
Total of 1760 entity types

**Subject Areas**

ISO 15926 (201)

Properties (158)
Time (104)
Intentional Objects (11)
Common Objects (37)

ISO 19107 (17)

Products and Materials (111)

Marketing (36)
Agreements (40)
Locations (38)

Buy/Sell (82)
CRM (39)
Organizations (222)

Demand (6)
Carrier (11)
Movement (128)
Retail (100)

Transport Constraints (174)
Manufacture (155)
Hydrocarbon Supply Chain (26)

Develop Data Model by Process area:
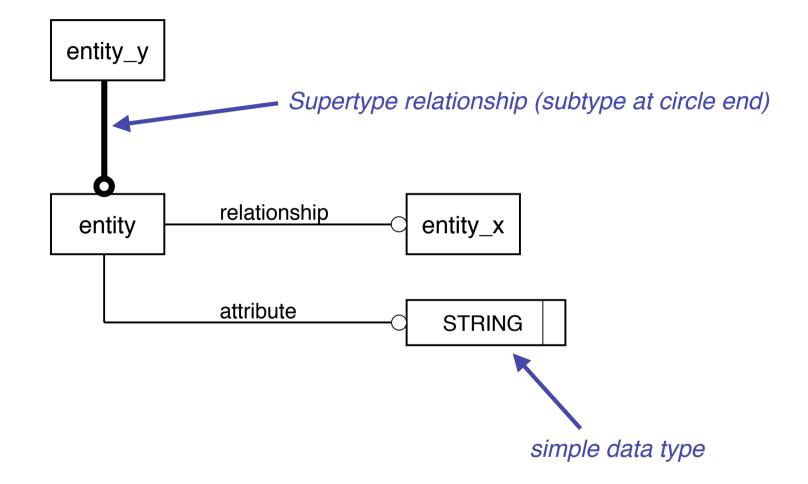- Clear accountability
- How evidence is organized

But
- Misses commonality across process areas

Develop Data Model by Subject area:
- Common elements brought together
- Integration across Process Areas

7

# Model Notation: EXPRESS-G



entity_y

Supertype relationship (subtype at circle end)

entity — relationship — entity_x
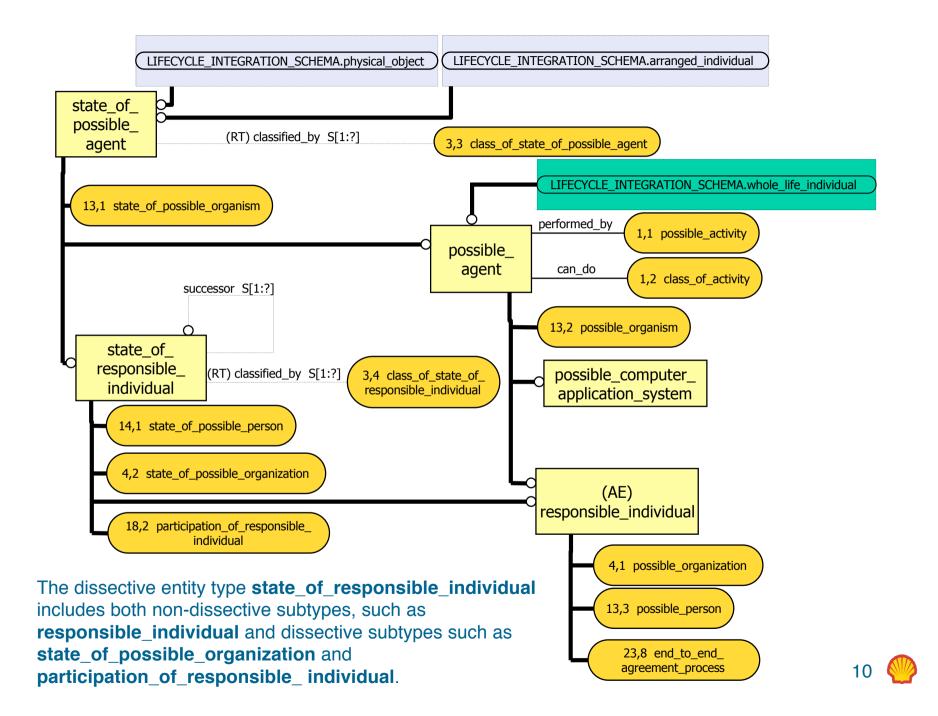
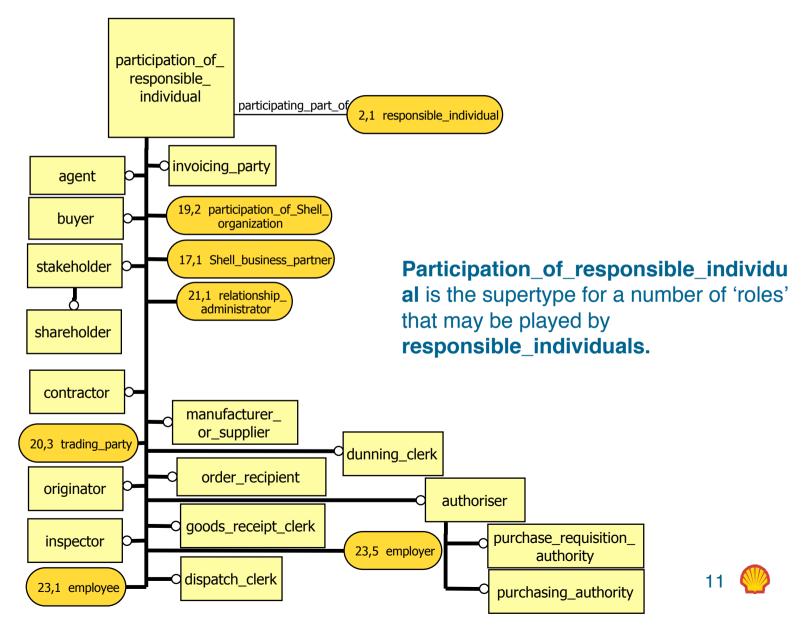entity — attribute — STRING

simple data type
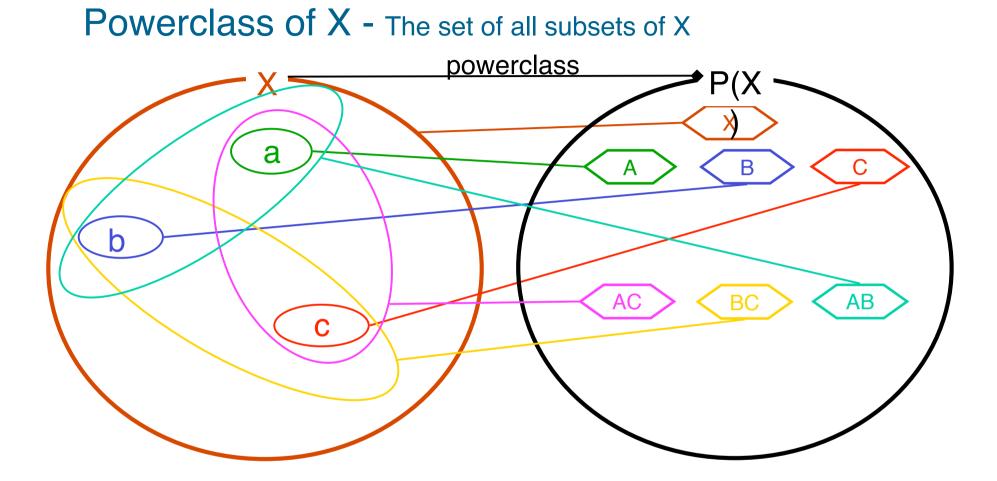
# Dissective and non-dissective

- Spatially

  - Mass vs count objects

    - An amount of oil is dissective (divide it in two you have two amounts of oil)

    - A car is non-dissective (divide it in two, you do not have two cars)

- Temporally

  - States vs whole life objects

    - A temporal part of a pumping activity is a pumping activity

    - A temporal part of a Project Programme is not a Project Programme

- Pragmatically, counting non-dissective things provides a more useful number.

  - If cars were spatially dissective, then if I looked in my garage and asked how many cars I had, I would count my car, my car minus a bonnet, my car minus a wheel, etc. – not normally a useful number.

  - If Project Programmes were temporally dissective, then if asked how many Shell Downstream programmes there are and I would count the Shell programme , the Shell programme minus its first day, the Shell programme minus its last day, etc. – not normally a useful number.

LIFECYCLE_INTEGRATION_SCHEMA.physical_object

LIFECYCLE_INTEGRATION_SCHEMA.arranged_individual

state_of_
possible_
agent

(RT) classified_by  S[1:?]

3,3  class_of_state_of_possible_agent

LIFECYCLE_INTEGRATION_SCHEMA.whole_life_individual

13,1  state_of_possible_organism

possible_
agent

performed_by

1,1  possible_activity

can_do

1,2  class_of_activity

successor  S[1:?]

13,2  possible_organism

state_of_
responsible_
individual

(RT) classified_by  S[1:?]

3,4  class_of_state_of_
responsible_individual

possible_computer_
application_system

14,1  state_of_possible_person

4,2  state_of_possible_organization

(AE)
responsible_individual

18,2  participation_of_responsible_
individual

4,1  possible_organization

13,3  possible_person

23,8  end_to_end_
agreement_process

The dissective entity type **state_of_responsible_individual**
includes both non-dissective subtypes, such as
**responsible_individual** and dissective subtypes such as
**state_of_possible_organization** and
**participation_of_responsible_ individual**.

10

# Roles played by responsible individuals



**Participation_of_responsible_individual** is the supertype for a number of 'roles' that may be played by **responsible_individuals.**

# Powerclass of X - The set of all subsets of X



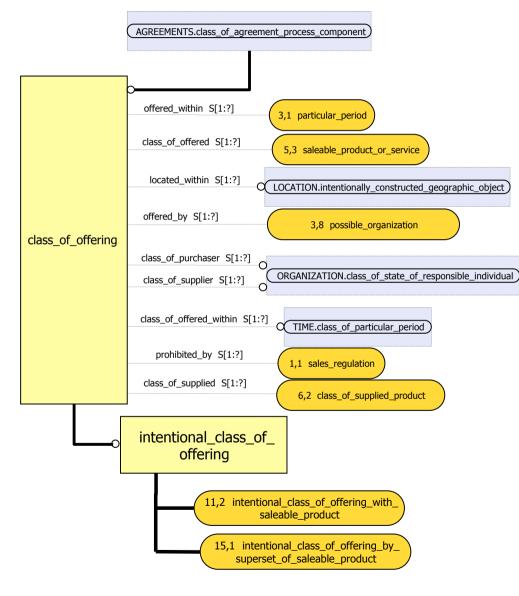In the ISO 15926 and Shell's DDM the powerclass of an entity type **X** is denoted by **class_of_X**.

# Powerclass in the DDM



AGREEMENTS.class_of_agreement_process_component

class_of_offering

offered_within S[1:?] — 3,1 particular_period

class_of_offered S[1:?] — 5,3 saleable_product_or_service

located_within S[1:?] — LOCATION.intentionally_constructed_geographic_object

offered_by S[1:?] — 3,8 possible_organization

class_of_purchaser S[1:?]
class_of_supplier S[1:?] — ORGANIZATION.class_of_state_of_responsible_individual

class_of_offered_within S[1:?] — TIME.class_of_particular_period

prohibited_by S[1:?] — 1,1 sales_regulation

class_of_supplied S[1:?] — 6,2 class_of_supplied_product

intentional_class_of_offering

11,2 intentional_class_of_offering_with_saleable_product

15,1 intentional_class_of_offering_by_superset_of_saleable_product

Here **class_of_offering** is the powerclass of all offers to sell something. On the other hand, **intentional_class_of_offering** is just those classes of offerings that are those that were intentionally made (and not any arbitrary or accidental classes).

13

# Summary and Conclusions

- This paper has examined the development of Shell's Downstream Data Model.

- The solution has its roots in a framework based on the notion of ontological commitment, making a choice for 4 dimensionalism with extensionalism.

- This paper has taken examples from the DDM of the impact of that choice in relation to two patterns.

    - Spatial and temporal dissectiveness in space and time.

    - Powerclasses as a means of providing an enriched classification mechanism.

- As an outcome, we believe that ontological understanding needs to be separated from both the 'epistemological gloss' and the 'implementational gloss'.

- An enhanced understanding of ontology is required in order that commitments can be laid bare and examined.

- Our experience here is that collaboration between those involved in conceptual modelling on the information systems side of the fence and those involved in philosophical ontology is potentially fruitful.

# Questions?