

# BORO Analysis Tools

---

C-FORS Summer School in Foundational Ontology  
(C-FORS 2025)

23 May 2025, University of Oslo, Norway

Chris Partridge  
Andrew Mitchell  
Oscar Xiberta Soto  
Diego Alberto Zabala Betancur  
Jonathan Eyre



Morning Sessions		
	9:00 - 9:05	Session 0 – Introduction
	9:05 - 9:45	Session 1 – Context
	10:00 - 10:45	Session 2 – BORO Ontology
	<b>11:00 - 12:00</b>	<b>Session 3 – Analysis Tools</b>
Afternoon Sessions		
	1:15 - 3:30	Session 1 – Practical Examples
	3:30 - 5:00	Session 2 – Examples Discussion / Presentation

- ④ Overview
- ④ LOAD: Structured Data Table Migration
  - BORO Top Level Tables
- ④ EVOLVE
  - Support Tools
    - Space Time Maps
    - Ontological Euler Diagrams
    - BORO UML
    - BORO eXcel Table (Manual) Pipeline
  - BORO KNIME Data Pipeline
- ④ Project planning



# Overview

- To provide some 'practice' with ontologization interoperability pipelines
  - by recreating the first stage of an ontologization interoperability pipeline
    - With (simple) examples
  - directly experience some of the challenges (in an attenuated form)
- Probably significantly more 'practice' needed before one becomes expert

# Two practical problems

Practical problems
lump of clay
ISO Countries

The focus this afternoon will be on the first problem.  
If you manage to finish the first problem, the second problem is available.





# Six BORO Summer School tools

Abbreviation	Name	Stage
SDTM	Structured Data Table Migration	LOAD
STM	Space Time Maps	EVOLVE
OED	Ontological Euler Diagrams	EVOLVE
BUML	BORO UML	EVOLVE
BXTP	BORO eXcel Table (Manual) Pipeline	EVOLVE
BKnDP	BORO KNIME Data Pipeline	EVOLVE

NB: This is a list developed for the Summer School. There is a large range of evolving tools, which we have selected from.  
Note also, our default language/tool is Python not KNIME. KNIME has been selected for its no-code credentials – making it more appropriate here.

# **LOAD: Structured Data Table Migration**

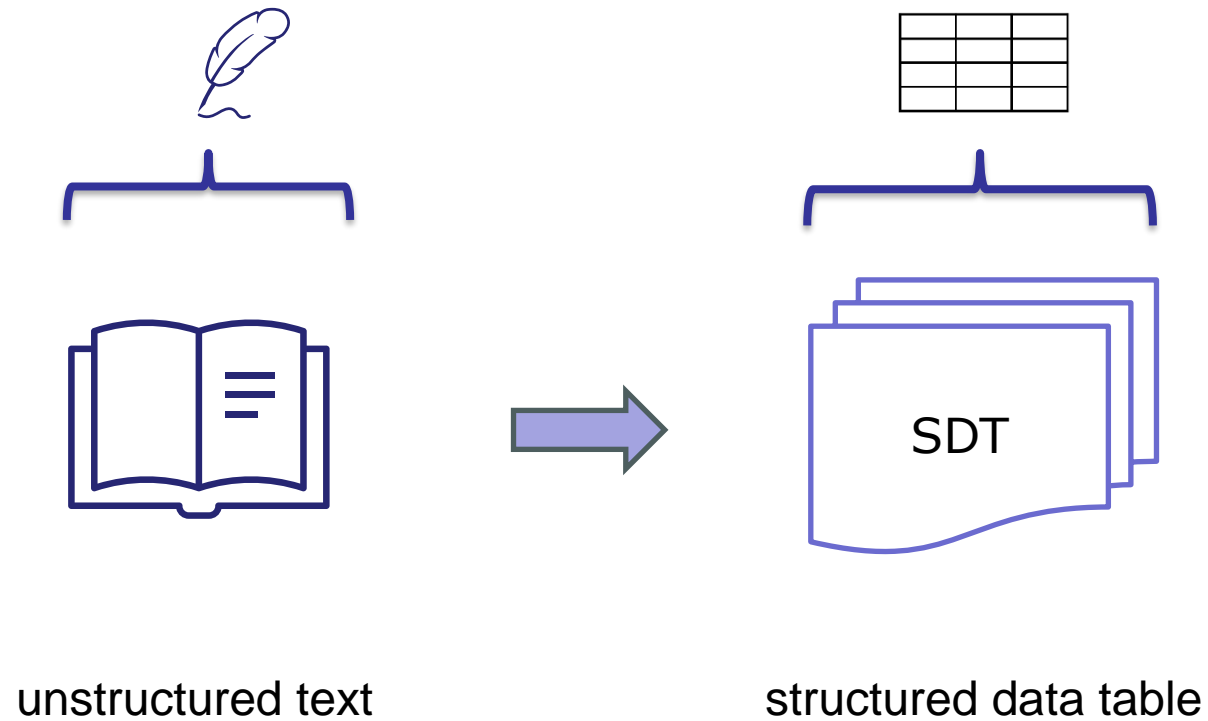
---



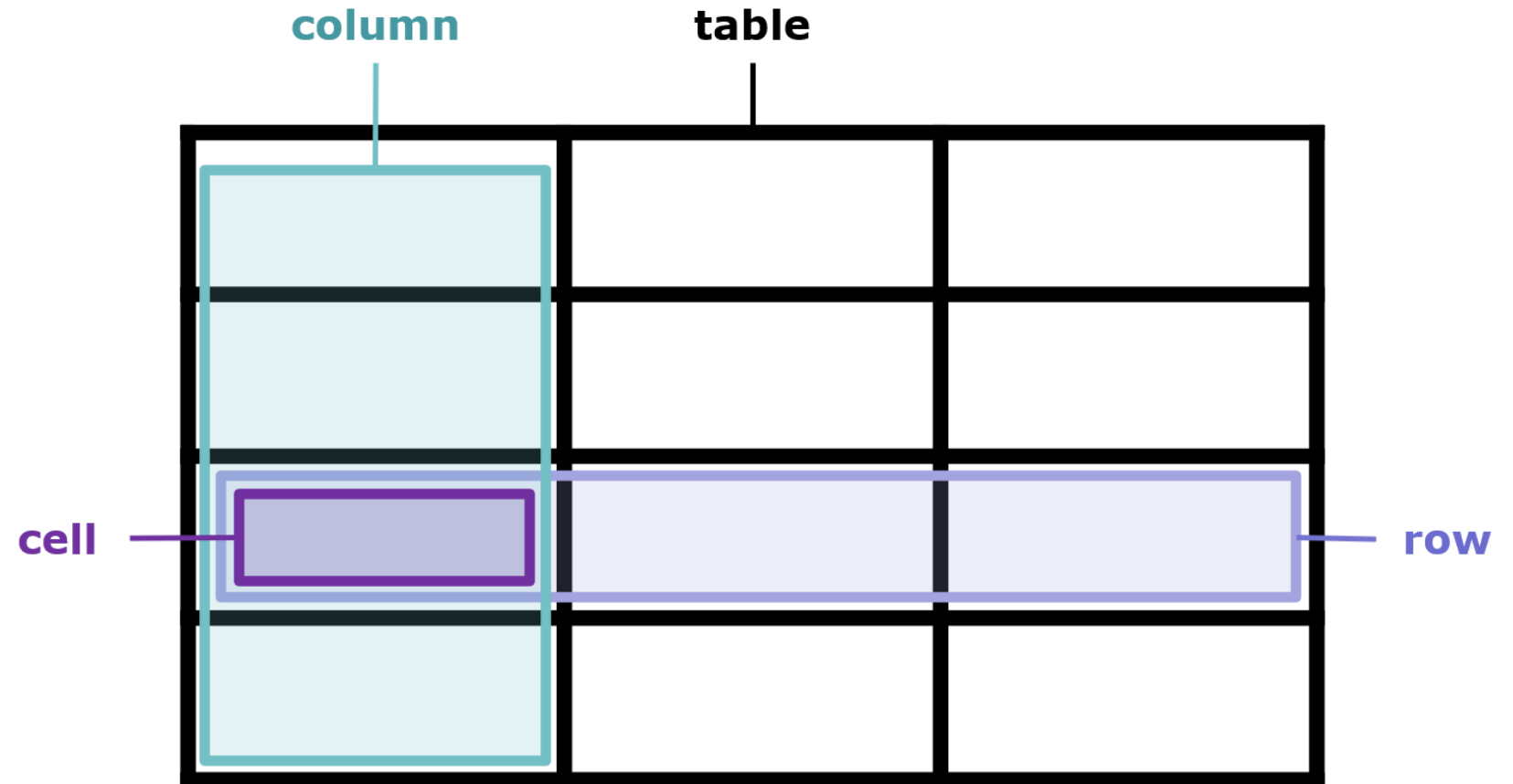
## Principle: shift (data) structure left

- ⦿ A goal of the bCLEARer pipeline process is to shift left, as far a possible, the move to structured data
  - enable us to work with machines talking to machines as soon as possible
- ⦿ If one starts with unstructured text
  - one needs to structure it
  - with the minimum intervention
  - preserving as much of the explicit structure as possible
- ⦿ A good first step is a shift to simple tables
  - the simplicity of their structure helps to keep the structuring clear
- ⦿ One goal of Load is to ensure one is working with structured data

# Structured data table migration



# Base data table structure



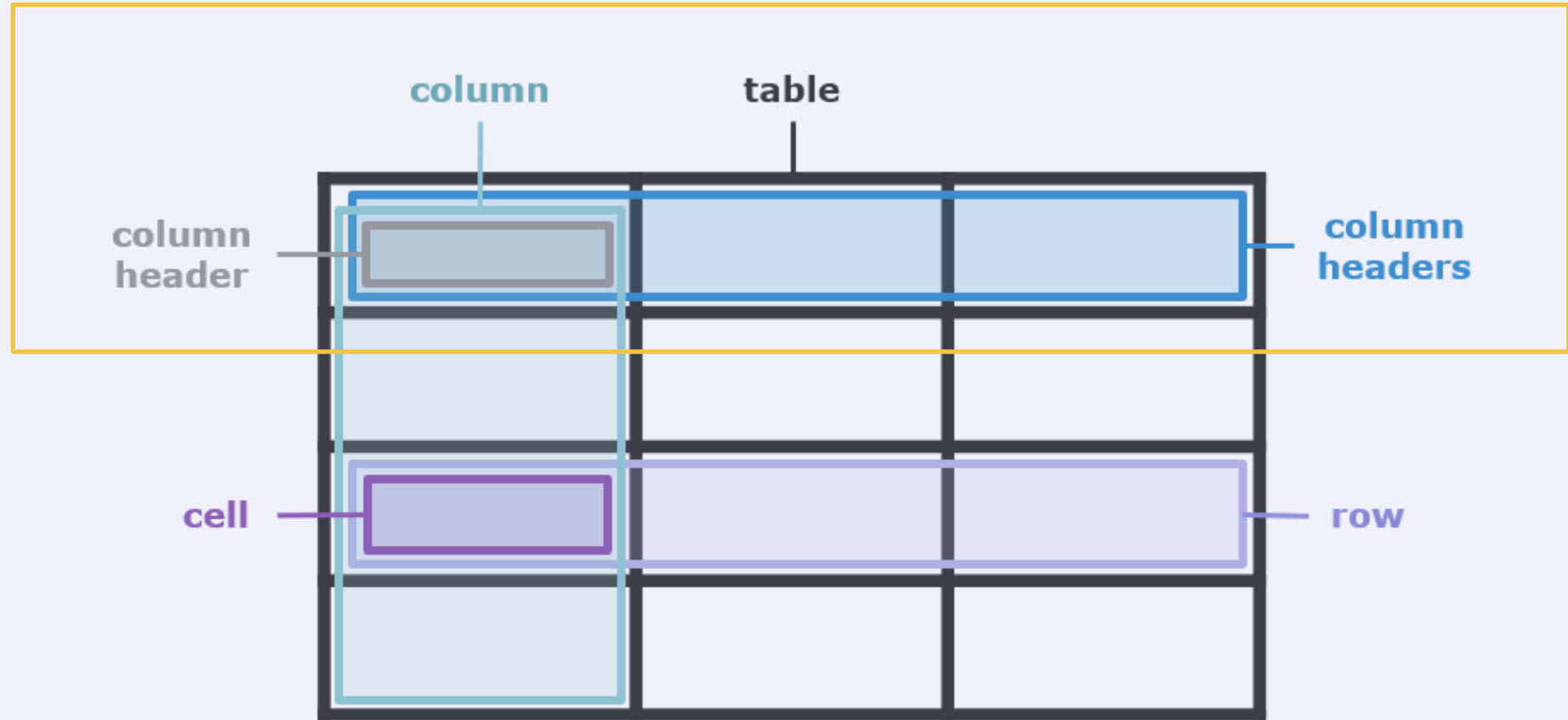
# Base table composition structure

## *base table composition structure*

<b>data type</b>	<b>structure type</b>	<b>table structure element</b>	<b>description</b>
data item	table	row	composed of cells organised by column
data column	table	column	composed of cells organised by row - has a single column heading cell.
data cell	table	cell	the intersection of columns and rows - has content



# Extended data table structure





## Lump of clay - starting point

---

- ④ We start with unstructured text:
  - “There is a lump of clay that, at time  $t_1$ , is used to make an aesthetically valuable statue. At time  $t_2$ , the statue is destroyed. At time  $t_3$ , the same lump of clay is reshaped to make a different statue. This statue is aesthetically valuable too.”
  
- ④ NOTE: We do this here
  - and provide you with the results for the practical example.

# Lump of clay – migration to table 1

“There is a lump of clay that,  
at time t1, is used to make an  
aesthetically valuable statue.

At time t2,  
the statue is destroyed.

At time t3,  
the same lump of clay is  
reshaped to make a  
different statue.

This statue is aesthetically  
valuable too.”

Excel interface showing a table with the following data:

A	B	C	D	E	F	G
lump of clay names						
lump of clay #1						

The table is located in the 'lumps\_of\_clay' workbook, and the 'statues' sheet is active. The status bar at the bottom indicates 'Ready' and 'Accessibility: Good to go'.

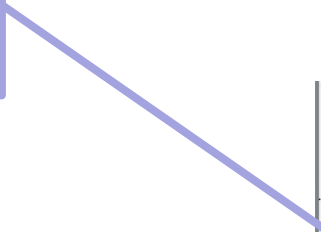
## Lump of clay – migration to table 2

“There is a lump of clay that, at time t1, is used to make an aesthetically valuable statue.

At time t2,  
the statue is destroyed.

At time t3,  
the same lump of clay is  
reshaped to make a  
different statue.

This statue is aesthetically  
valuable too.”



	A	B	C	D	E
1	statue_names	aesthetically valuable	made from	made on	destroyed on
2	statue #1	Yes	lump of clay #1	t1	t2
3	statue #2	Yes	lump of clay #1	t3	
4					

lumps\_of\_clay   statues   +   ⋮

# Lump of clay – migration to table 3

“There is a lump of clay that, at time t1, is used to make an aesthetically valuable statue.

At time t2,  
the statue is destroyed.

At time t3,  
the same lump of clay is  
reshaped to make a  
different statue.

This statue is aesthetically  
valuable too.”

	A	B	C	D	E
1	statue_names	aesthetically_valuable	made from	made on	destroyed on
2	statue #1	Yes	lump of clay #1	t1	t2
3	statue #2	Yes	lump of clay #1	t3	
4					

lumps\_of\_clay   statues   +   ⋮

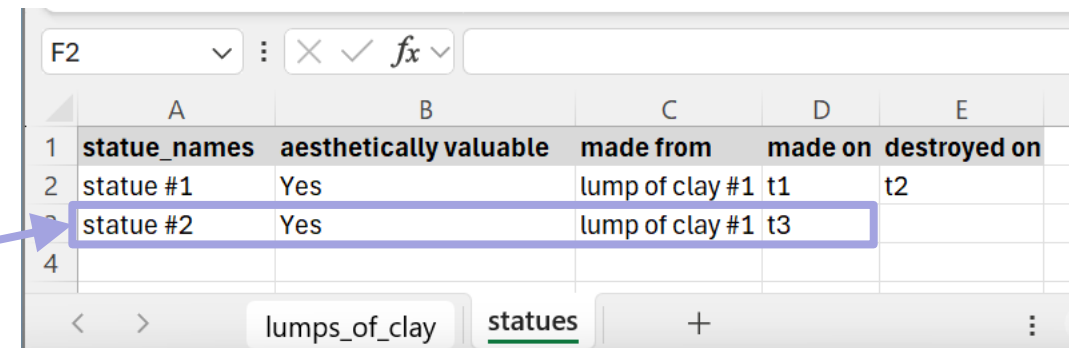
# Lump of clay – migration to table 4

“There is a lump of clay that, at time t1, is used to make an aesthetically valuable statue.

At time t2, the statue is destroyed.

At time t3, the same lump of clay is reshaped to make a different statue.

This statue is aesthetically valuable too.”



	A	B	C	D	E
1	statue_names	aesthetically valuable	made from	made on	destroyed on
2	statue #1	Yes	lump of clay #1	t1	t2
3	statue #2	Yes	lump of clay #1	t3	
4					

lumps\_of\_clay   statues   +   ⋮



# Lump of clay - structured data tables – proto time

not typical time data

	A	B	C	D	E	F	G	H	I	J
1	statue_names	aesthetically valuable	made from	made on	destroyed on					
2	statue #1	Yes	lump of clay #1	t1	t2					
3	statue #2	Yes	lump of clay #1	t3						
4										

	A	B	C	D	E	F	G	H	I	J	K
1	lump_of_clay_names										
2	lump of clay #1										
3											

a\_collect - Example - lump of clay - stage 1 - proto time.xlsx

# Lump of clay - structured data tables – dataise time

	A	B	C	D	E
1	statue_names	aesthetically valuable	made from	made on	destroyed on
2	statue #1	Yes	lump of clay #1	t1	t2
3	statue #2	Yes	lump of clay #1	t3	
4					

a\_collect - Example - lump of clay - stage 1 - proto time.xlsx

A

B

C

D

E

F

G

1

lump\_of\_clay\_names

2

lump of clay #1

3

<

>


lumps\_of\_clay

statues

times




+

Ready

 Accessibility: Good to go

A1

:

statue\_names

A

B

C

D

E

1

statue\_names

aesthetically valuable

made from

made on

destroyed on

2

statue #1

Yes

lump of clay #1

2023-06-20

2023-10-24

3

statue #2

Yes

lump of clay #1

2024-07-06

<

>


lumps\_of\_clay

statues

times




+

Ready

 Accessibility: Good to go

A1

:

time\_names

A

B

C

D

E

F

G

1

time names

YYYY-MM-DD

2

t1

2023-06-20

3

t2

2023-10-24

4

t3

2024-07-06

5

<

>


lumps\_of\_clay

statues

times

+

Ready

 Accessibility: Good to go

</

NB: added domain knowledge about time granularity

a\_collect - Example - lump of clay - stage 2 - dataise time.xlsx

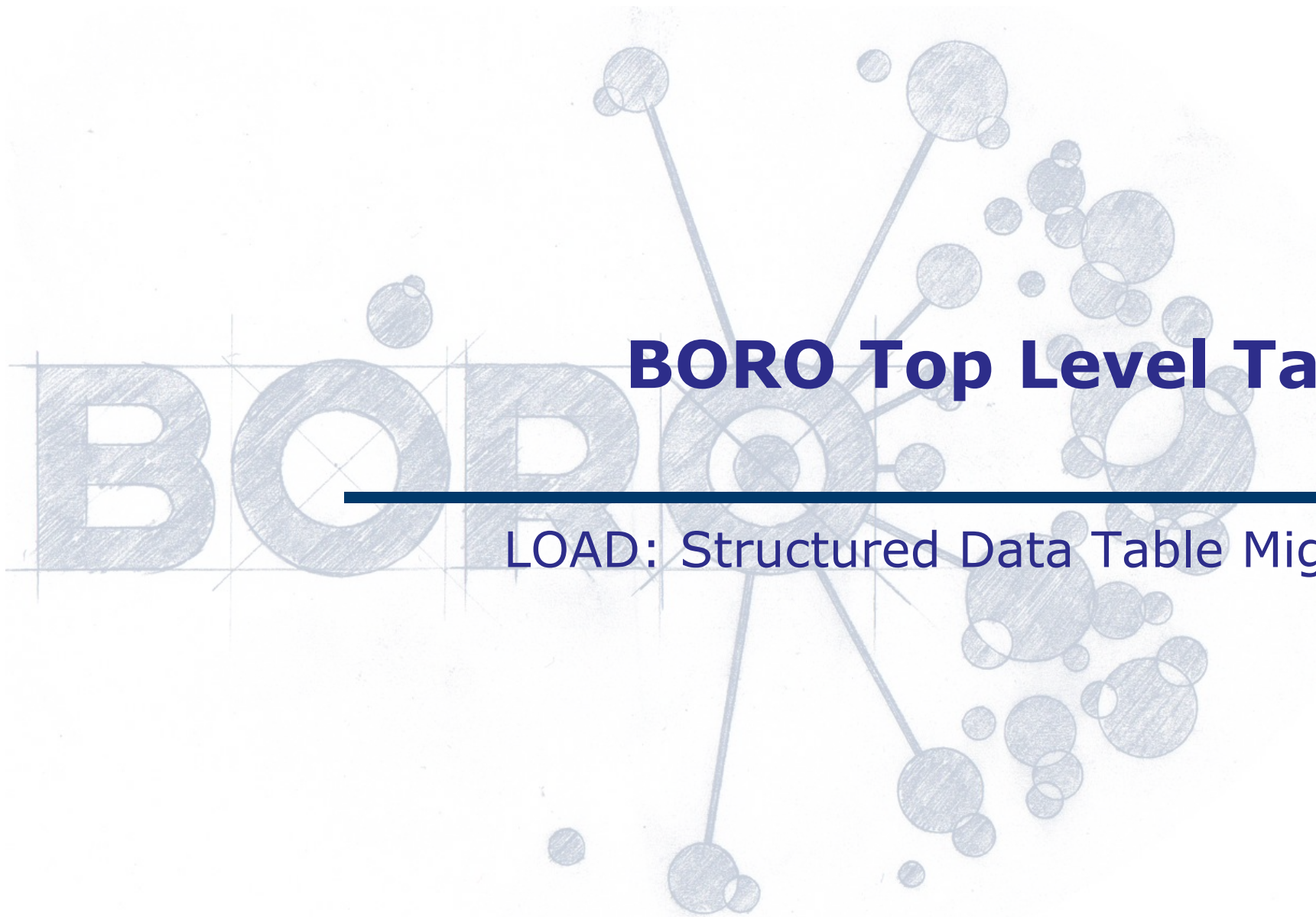
# ISO Countries

Already structured data (table)

	A	B	C	D	E	F
1	<b>English short name</b>	<b>French short name</b>	<b>Alpha-2 code</b>	<b>Alpha-3 code</b>	<b>Numeric</b>	
2	Afghanistan	Afghanistan (l')	AF	AFG	4	
3	Albania	Albanie (l')	AL	ALB	8	
4	Algeria	Algérie (l')	DZ	DZA	12	
5	American Samoa	Samoa américaines (les)	AS	ASM	16	
6	Andorra	Andorre (l')	AD	AND	20	
7	Angola	Angola (l')	AO	AGO	24	
8	Anguilla	Anguilla	AI	AIA	660	
9	Antarctica	Antarctique (l')	AQ	ATA	10	
10	Antigua and Barbuda	Antigua-et-Barbuda	AG	ATG	28	
11	Argentina	Argentine (l')	AR	ARG	32	
12	Armenia	Arménie (l')	AM	ARM	51	
13	Aruba	Aruba	AW	ABW	533	
14	Australia	Australie (l')	AU	AUS	36	
15	Austria	Autriche (l')	AT	AUT	40	
16	Azerbaijan	Azerbaïdjan (l')	AZ	AZE	31	
17	Bahamas (the)	Bahamas (les)	BS	BHS	44	

< > with formatting without formatting +

a\_collect - Example - ISO 3166 - Part 1 -- Country code.xlsx



# **BORO Top Level Tables**

---

**LOAD: Structured Data Table Migration**



# BORO top level tables

---

- ⦿ The top level needs to be introduced to the pipeline
- ⦿ Various ways of doing this
- ⦿ Here, we add it in directly at LOAD
  - this is done for you in the starter pack
- ⦿ We add a simplified table version
  - suitable for the worked examples





# BORO top level tables - pure

table_names	column_1_names	column_2_names	column_3_names	column_4_names
Sets	bie_ids	names		
Individuals	bie_ids	names		
Tuples	bie_ids			
elements-sets	bie_ids	element_bie_ids	set_bie_ids	
parts-wholes	bie_ids	part_bie_ids	whole_bie_ids	
sub-super-sets	bie_ids	sub_set_bie_ids	super_set_bie_ids	
tuple-places	bie_ids	placing_bie_ids	placed_bie_ids	place_numbers

# BORO top level tables - pragmatic

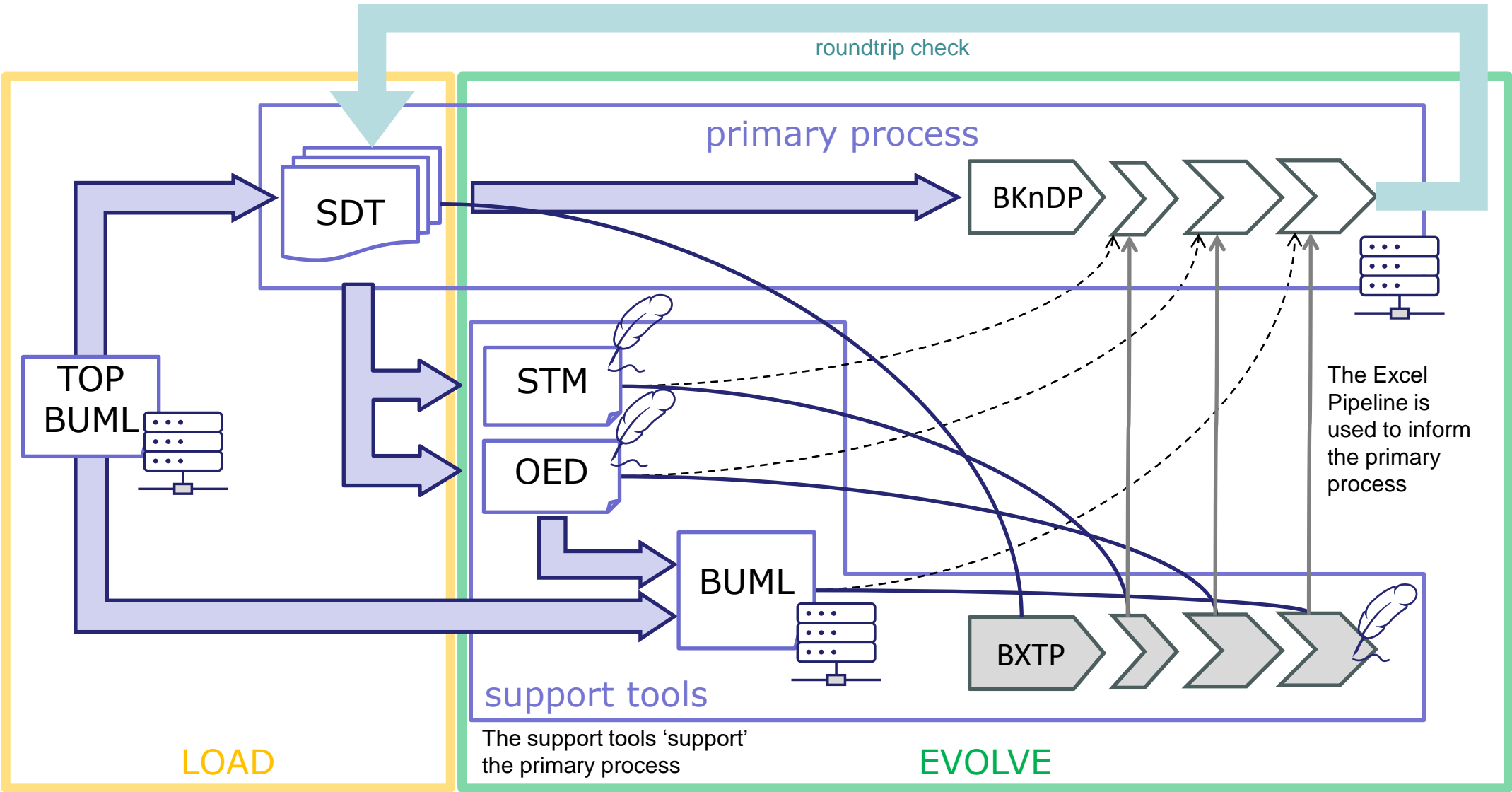
table_names	column_1_names	column_2_names	column_3_names	column_4_names
<b>Sets</b>	bie_ids	names		
<b>Individuals</b>	bie_ids	names		
<b>Tuples</b>	bie_ids	place1_bie_ids	place2_bie_ids	place3_bie_ids
<b>grounding-relation-places</b>	bie_ids	relation-place-type_names	placing_bie_ids	placed_bie_ids
<b>elements-sets</b>	bie_ids	element_bie_ids	set_bie_ids	
<b>parts-wholes</b>	bie_ids	part_bie_ids	whole_bie_ids	
<b>sub-super-sets</b>	bie_ids	sub_set_bie_ids	super_set_bie_ids	



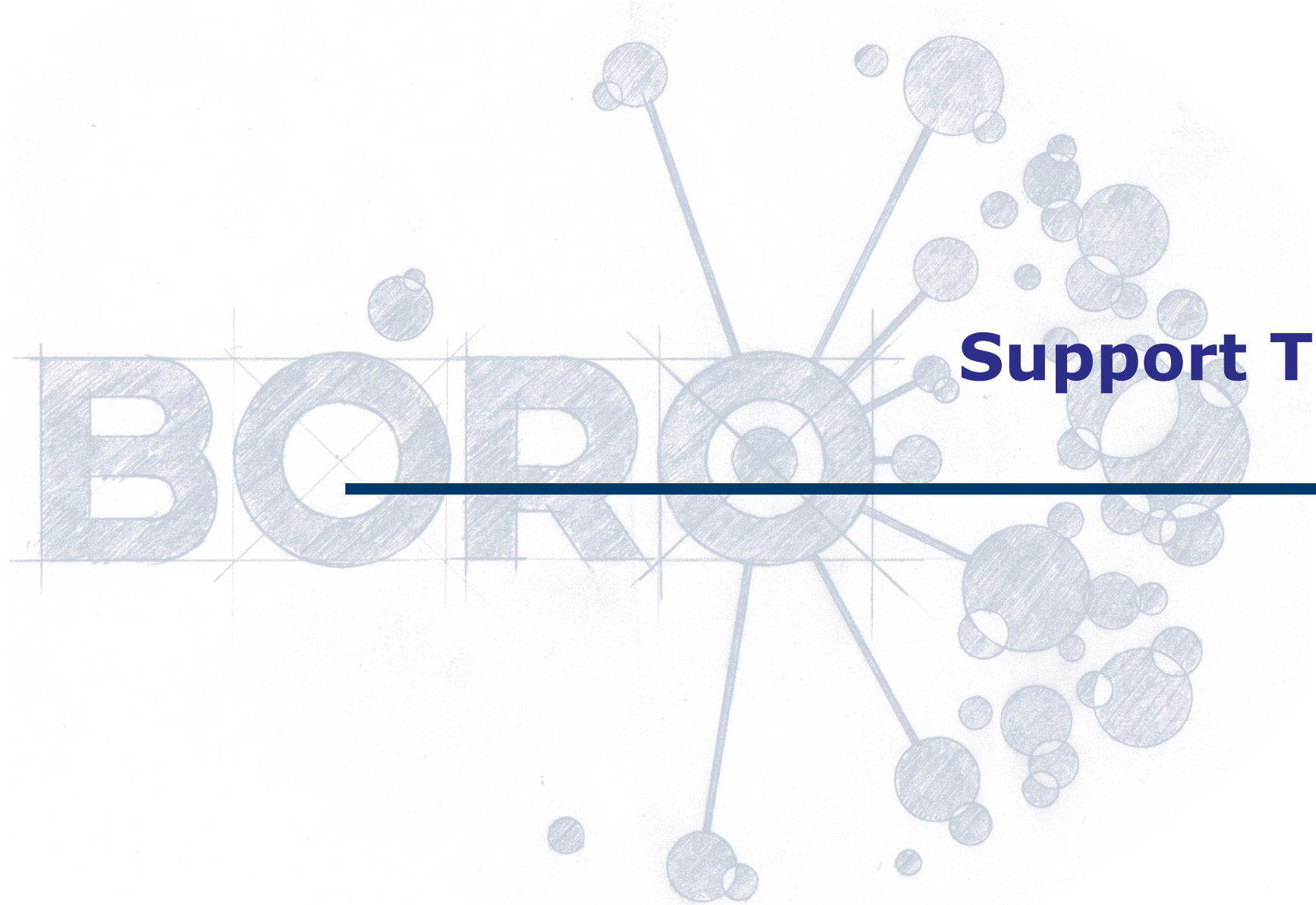
**EVOLVE**

- ⑥ Same process stages for both practical problems
- ⑥ Broadly,
  - there is
    - a support process that 'designs' the pipeline
      - this has human aspects
    - a primary process that IS the pipeline
      - this is a 'pure' machine process
        - the process is executable, repeatable and inspectable
- ⑥ Hopefully,
  - building the primary process will help you experience the practice on ontologising a pipeline

# EVOLVE - practical problem process





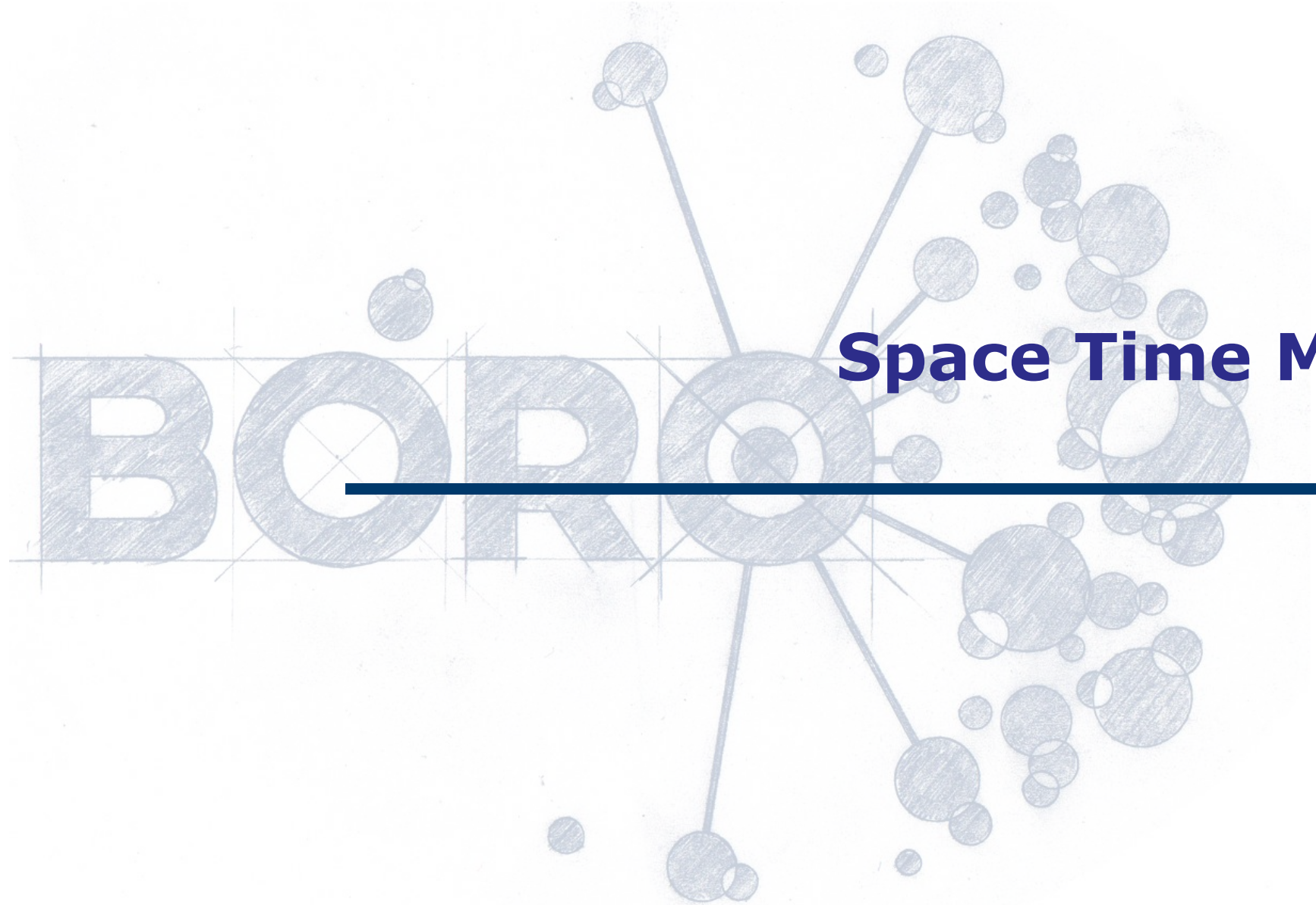


## Support Tools

# Four BORO Summer School support tools

Abbreviation	Name
STM	Space Time Maps
OED	Ontological Euler Diagrams
BUML	BORO UML
BXTP	BORO eXcel Table (Manual) Pipeline

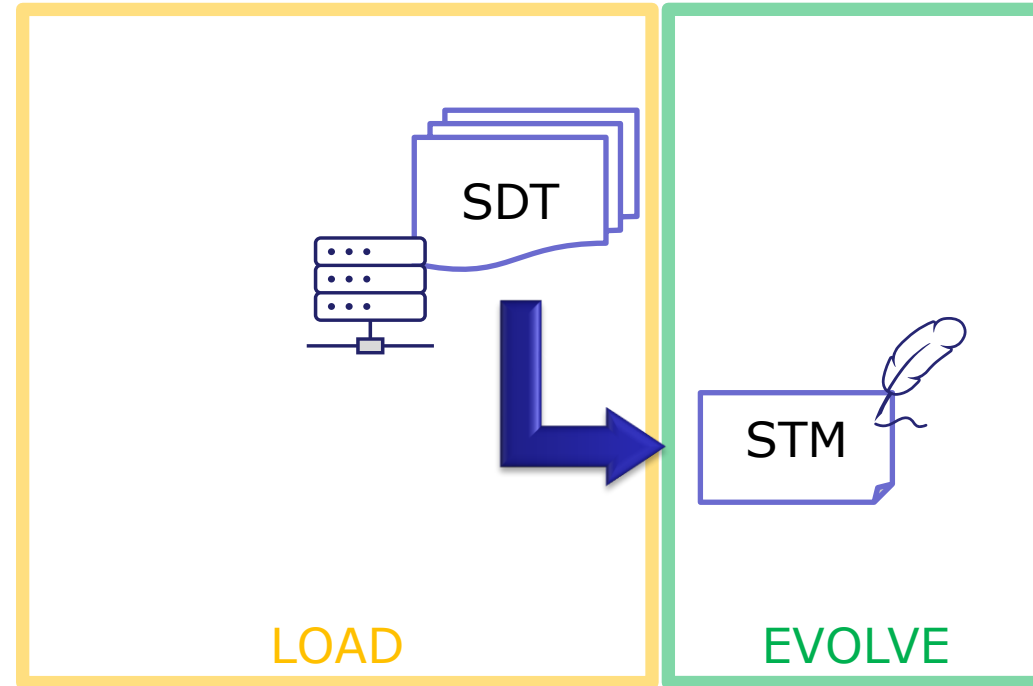
These Summer School ‘tools’ are part of a much wider, richer toolkit  
They are useful pedagogical devices  
As often, with practices, once one gains competence with these, one  
deploys them in innovative radically different ways



## Space Time Maps

---

# Practical problem process

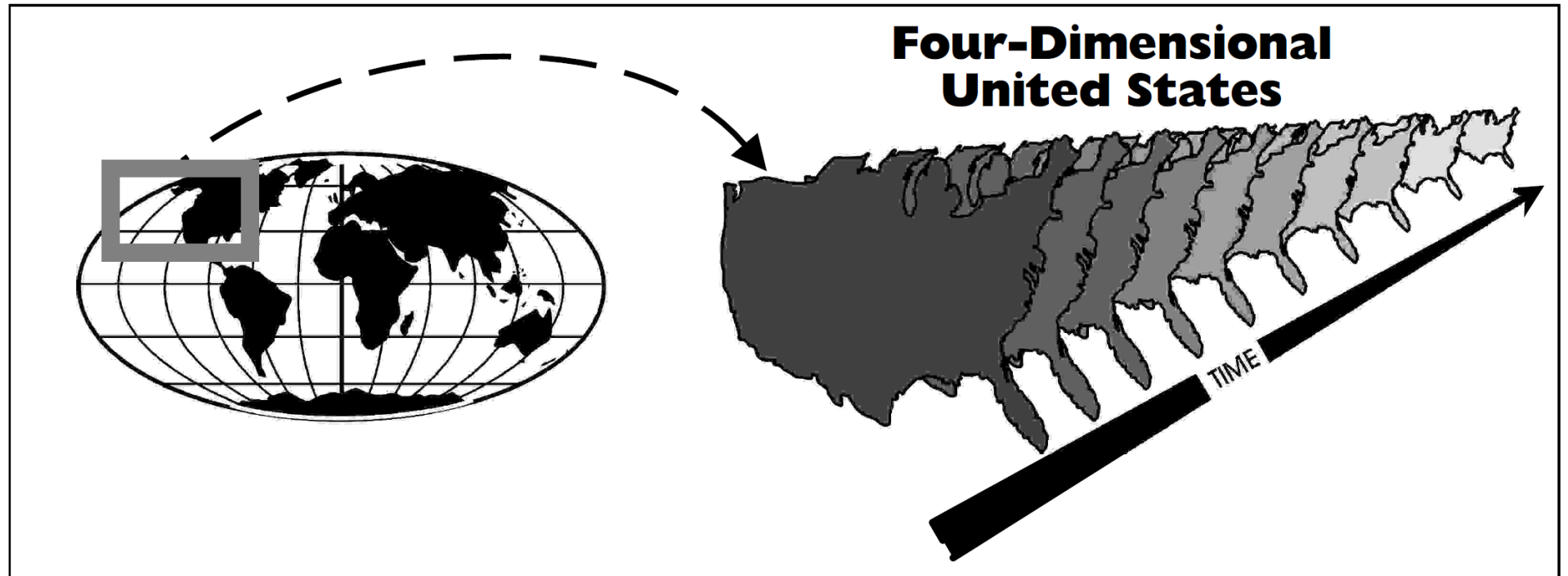


- STMs Focus on Individuals
  - other BORO objects are constructed from (grounded in) Individuals
- STMs are an intuitive way of visualising the mereology of Individuals
  - in particular the overlap and part hood relations
- As Individuals (in BORO) have a mereological extensional criteria of identity
  - the 'extensional' two-dimensional surface of a page (or screen) has sufficiently similar characteristics to represent 'directly' (visually) the mereology
  - humans find this a useful tool

- STMs are typically used as a first stage of analysis, where in the second stage these are translated into a BORO model
- STMs provide an easy way to visualise 4D elements in a domain
  - and as such are a useful way to start understanding them
- The STM visualisation aims to abstract the individuals to their mereological essentials
- Too time consuming to draw for the whole domain
  - so, collect representative examples of important 'typical' mereological relations
  - these are candidates for STMs



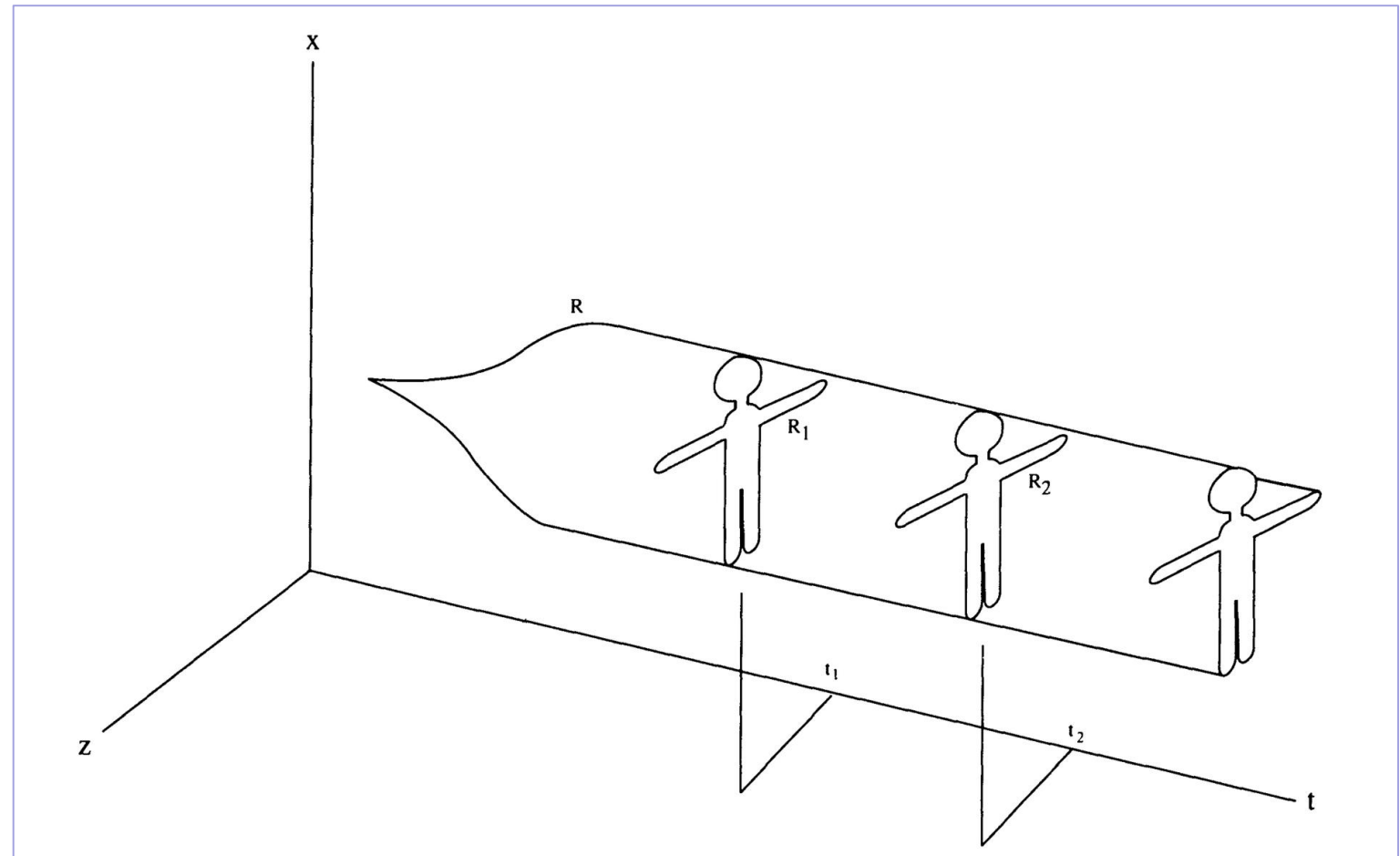
- BORO is an extensional ontology
  - the criterion of identity for Individuals is mereological (spatio-temporal) extension
  - or more colloquially four-dimensional (4D)










Not that usual, but STMs appear in various places.  
In this case, not quite the same, less focus on exact mereology.

# Van Inwagen's space-time diagrams



Van Inwagen, P. (1990). Four-Dimensional Objects. *Noûs*, 24(2), 245.

# Three components of an STM Individual


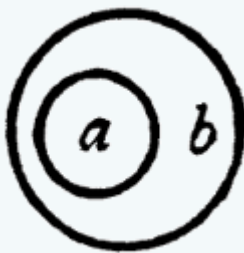
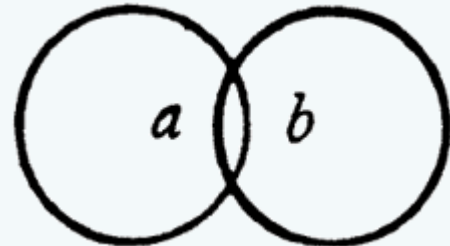
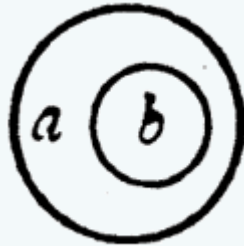
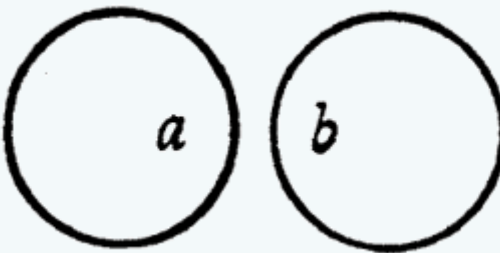
A region representing a spatio-temporal extent	A border around the region – the <i>contour</i> – marking the region as representing an Individual	A string of letters labelling the Individual
		A
		



# The Gergonne five possible extensional relations

- The Gergonne relations give the five possible ways in which two objects can simply share extension



coincident	contained	overlap	contains	disjoint
"if and only if every a is a b and every b is an a"	"if and only if every a is a b and not every b is an a"	"if and only if it is not the case that either every a is a b or every b is an a or no a is a b"	"if and only if every b is an a and not every a is a b"	"if and only if no a is a b"
				

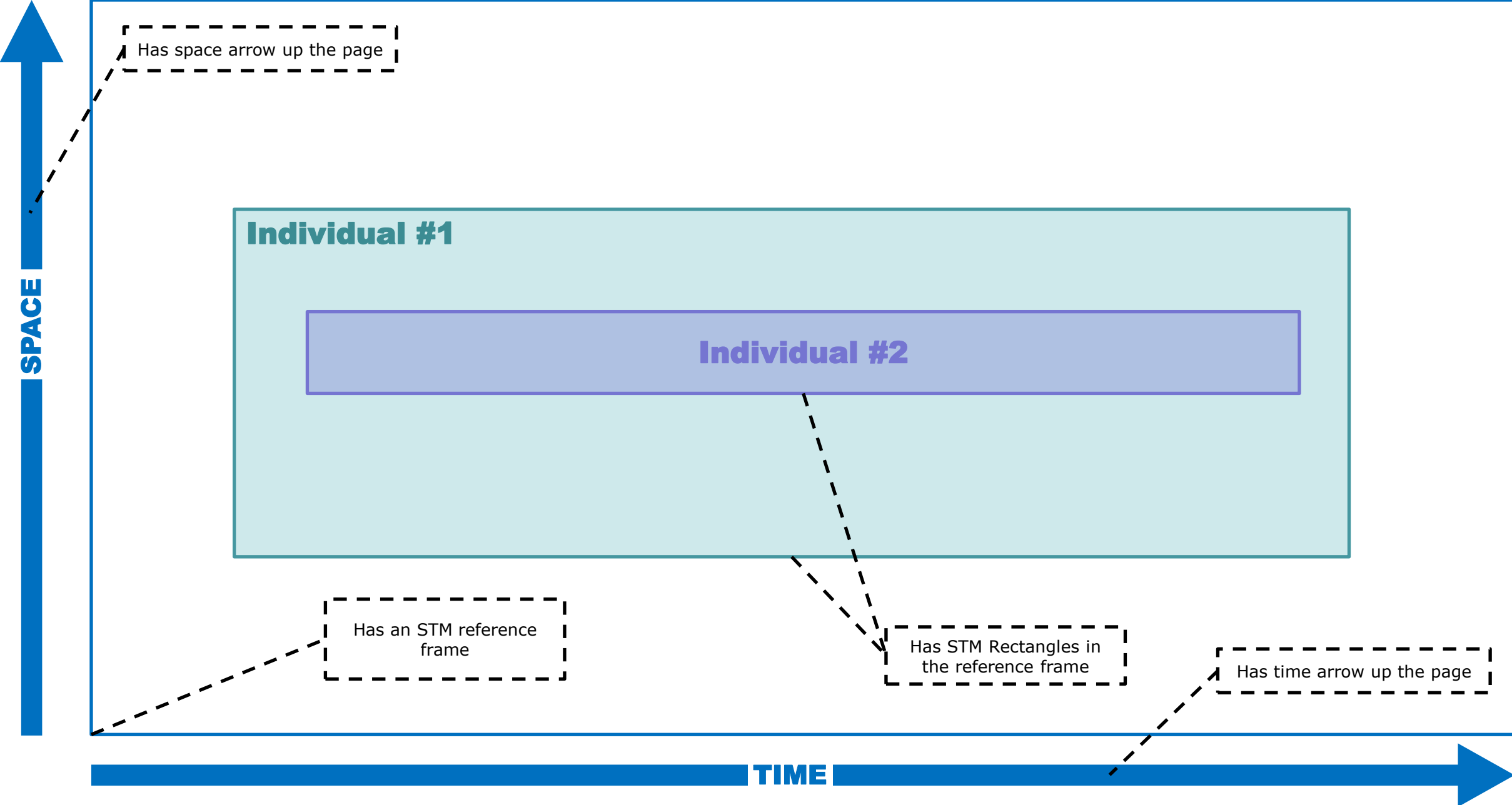
Based upon: J. A. Faris - The Gergonne Relations - The Journal of Symbolic Logic, Vol. 20, No. 3 (Sep., 1955), pp. 207-231  
 based upon: J. D. Gergonne, Essai de dialectique rationelle, Annales des mathematiques pures et appliquees, Vol. 7, (1817).

## Clues to look for

- space axis along the left side of the page
- time axis across the bottom of the page
- reference frame contained within the axes
- Individual rectangles (shapes) in the reference frame

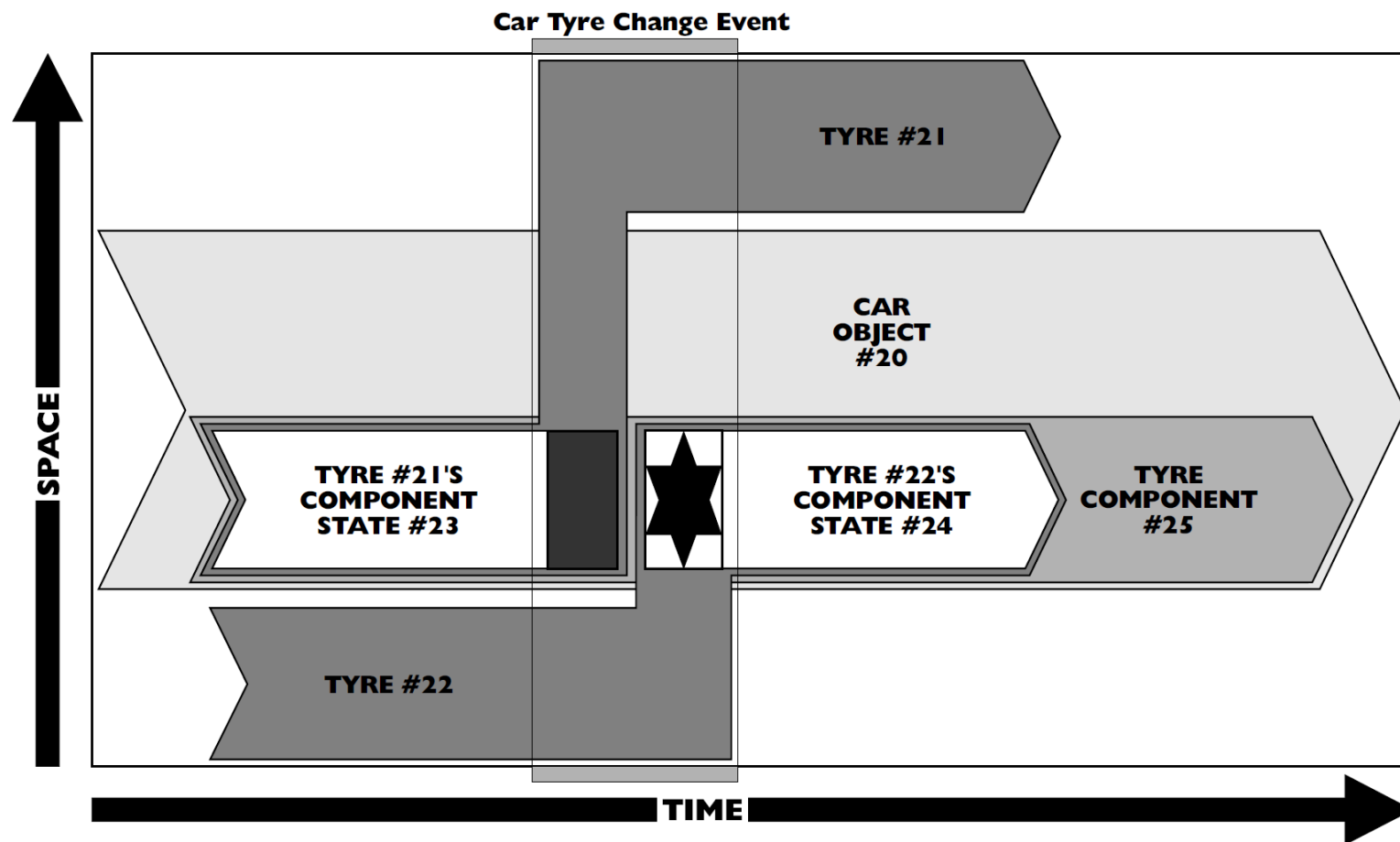


# Lore ipsum



## Patterns: overlap

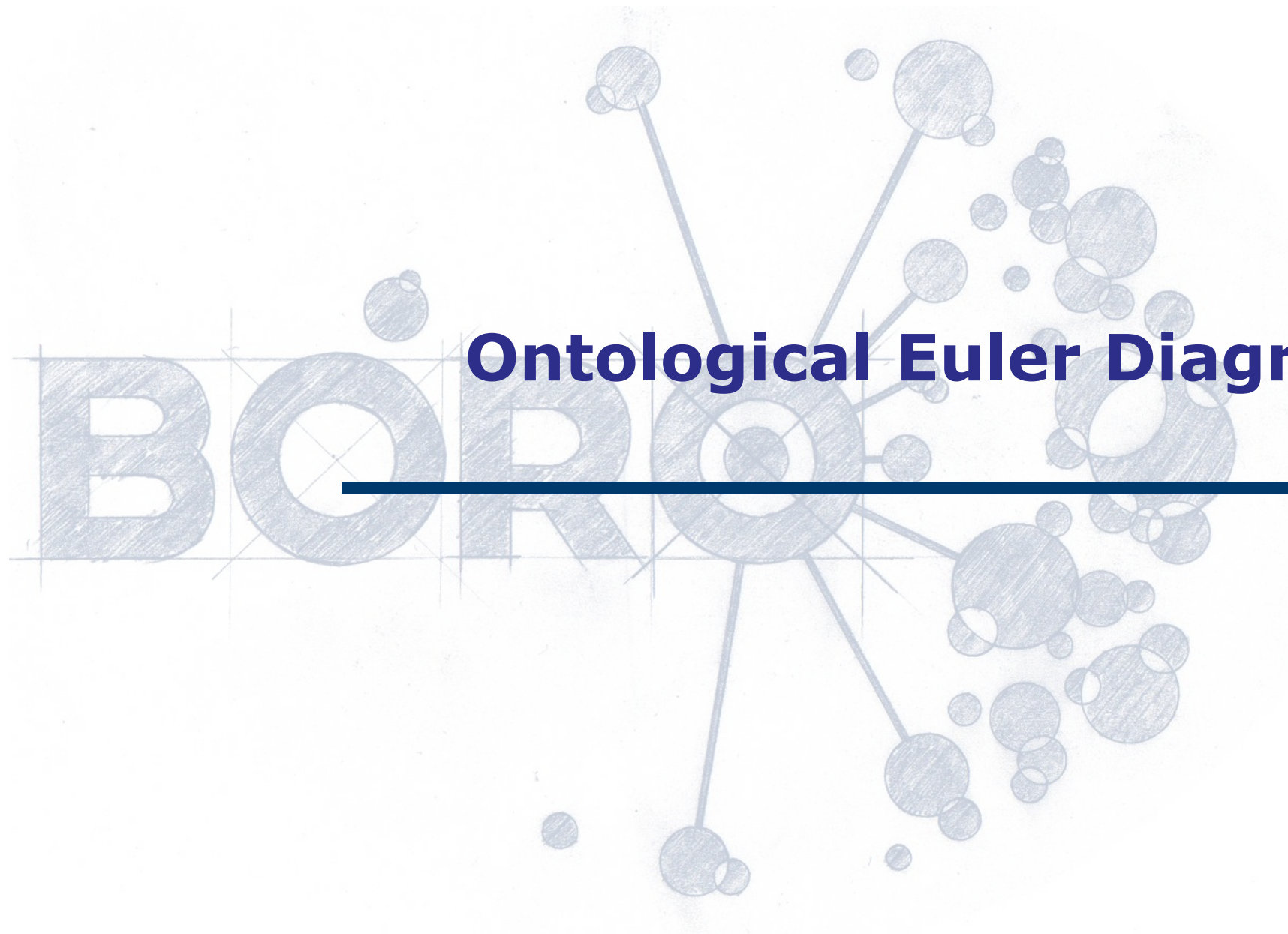
**Figure 8.14:**  
Car tyre change  
space-time map



From *Business Objects: Re-engineering for Reuse*

- The 2D reference frame picks out the relevant fragment of the 4D domain
- The domain's objects are shown as boxes
  - a box's boundary represents the Individual's real boundary
  - boxes' containment, overlap, etc. show the mereological relationships between Individuals
  - a box's name helps to further identify the Individual

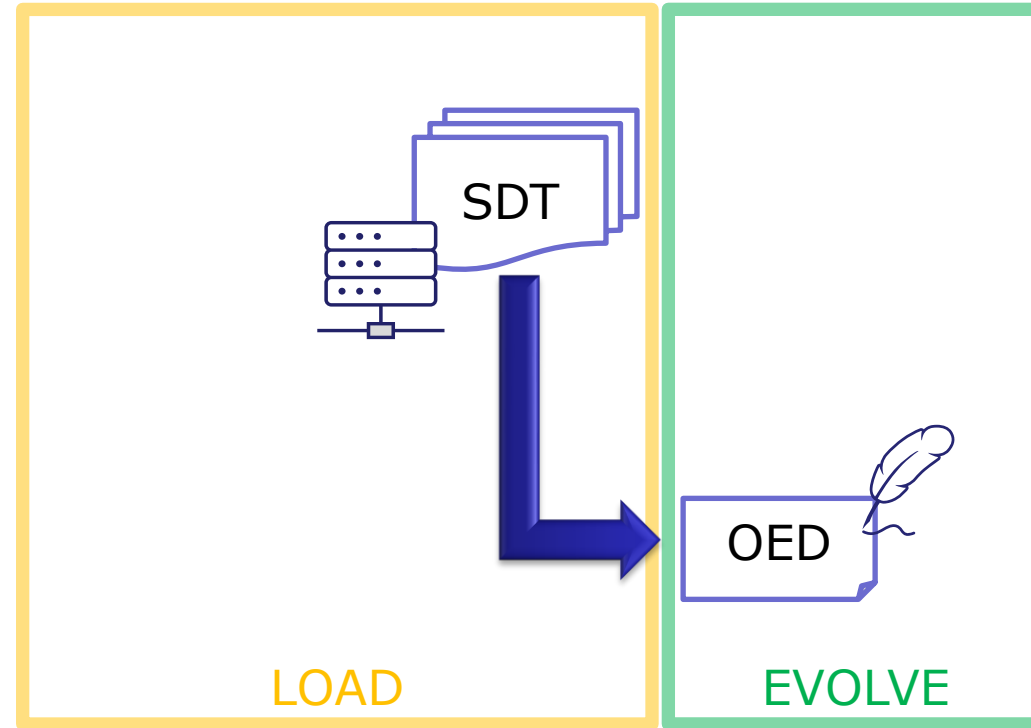




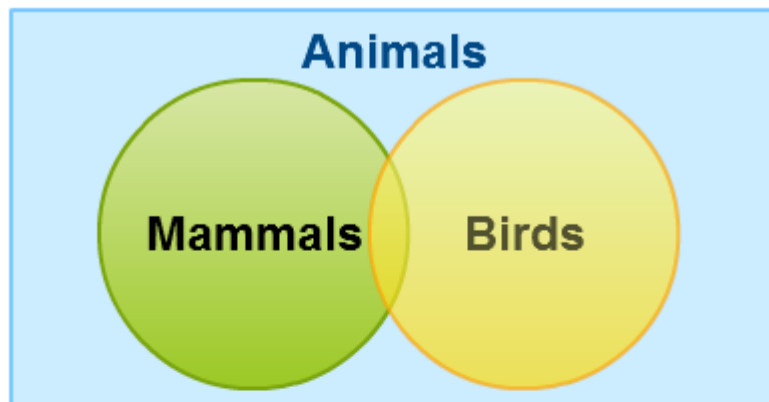
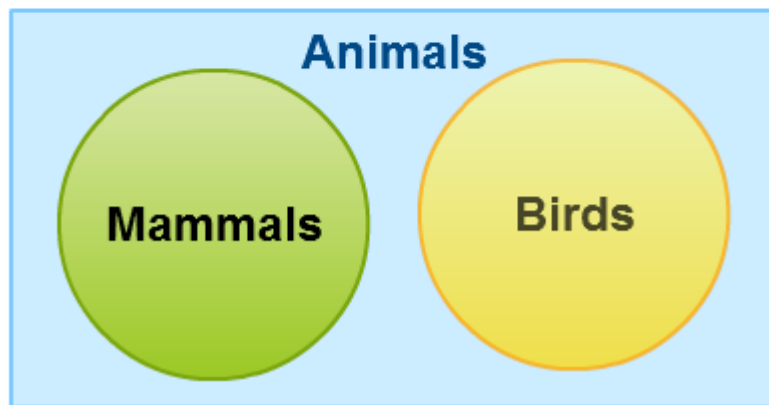
# Ontological Euler Diagrams

---

# Practical problem process



# Euler not Venn diagrams

V  
E  
N  
NE  
U  
L  
E  
R

A Venn diagrams show all combinations (even impossible ones).

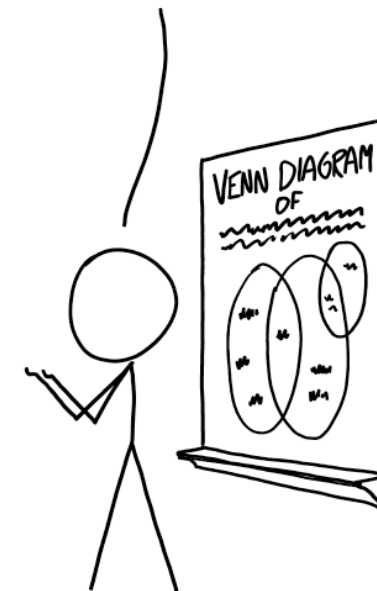
ACTUALLY, THAT'S AN *EULER* DIAGRAM, BECAUSE-

COME *ONNNNN*!

*EVERYTHING* IS NAMED AFTER EULER.  
EULER'S CONSTANT, EULER'S FUNCTION.  
CAN'T WE LET JOHN VENN HAVE THIS?

NO.

ALSO, NUMBERS ARE  
NOW "EULER LETTERS."



<https://xkcd.com/2721/>

# Types in a BORO foundational ontology

---

## ⦿ Sets' identity relations

- from an ontological perspective, these will exist simpliciter
- from an epistemic perspective, we may or may not know whether these exist
- an ontological picture should not contain epistemological elements
- to handle this issue, for the topic of the Euler diagram
  - we make a set identity relation omniscience assumption
    - in other words, all the identity relations between sets are known

## ⦿ With this assumption, it makes sense to adopt

- the 'Existential Import Convention'
  - where adding a region to the Euler diagram implies that it exists
    - in other words, it has members

# Sets in a BORO foundational ontology

---

- In BORO's extensional ontology
  - given a Set's identity is grounded in its members
  - then the way in which two Sets' identity can be related are broadly
    - with the same members are identical
    - with different members are not identical
    - that share members are partially identical
    - with no members in common are disjoint

## Three basic extensional set patterns – for two sets

- There are only three ways two sets can be 'identity' related

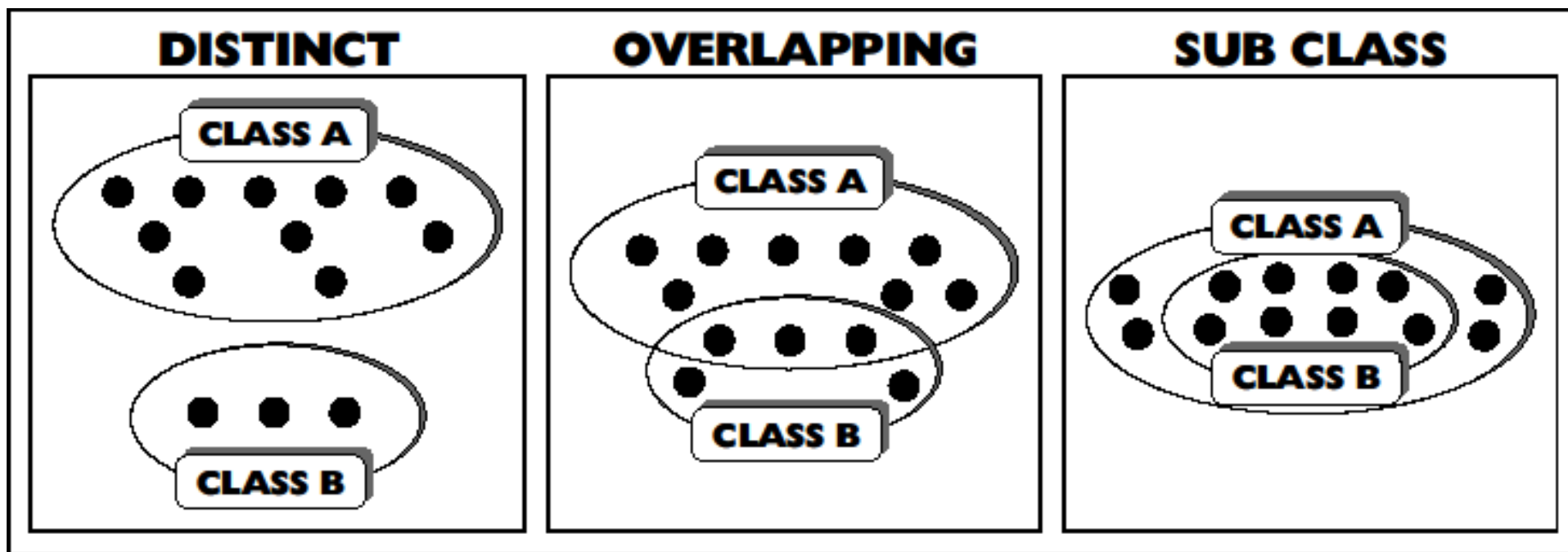
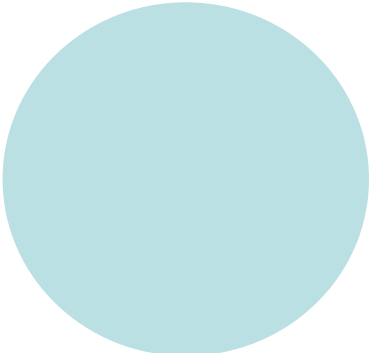
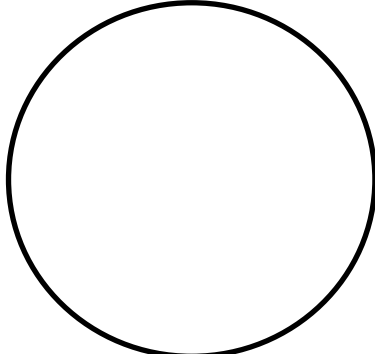
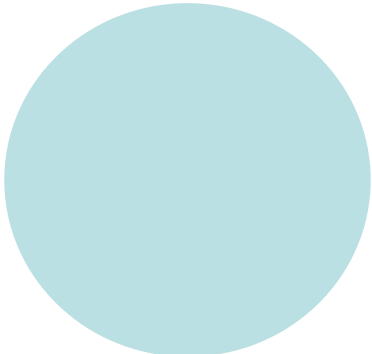
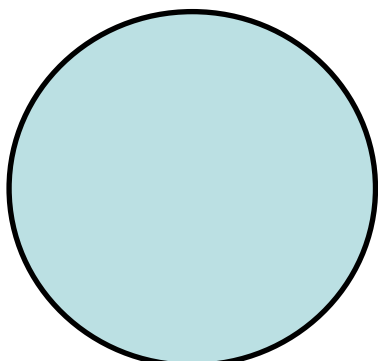
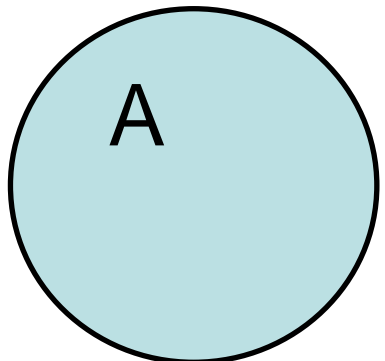


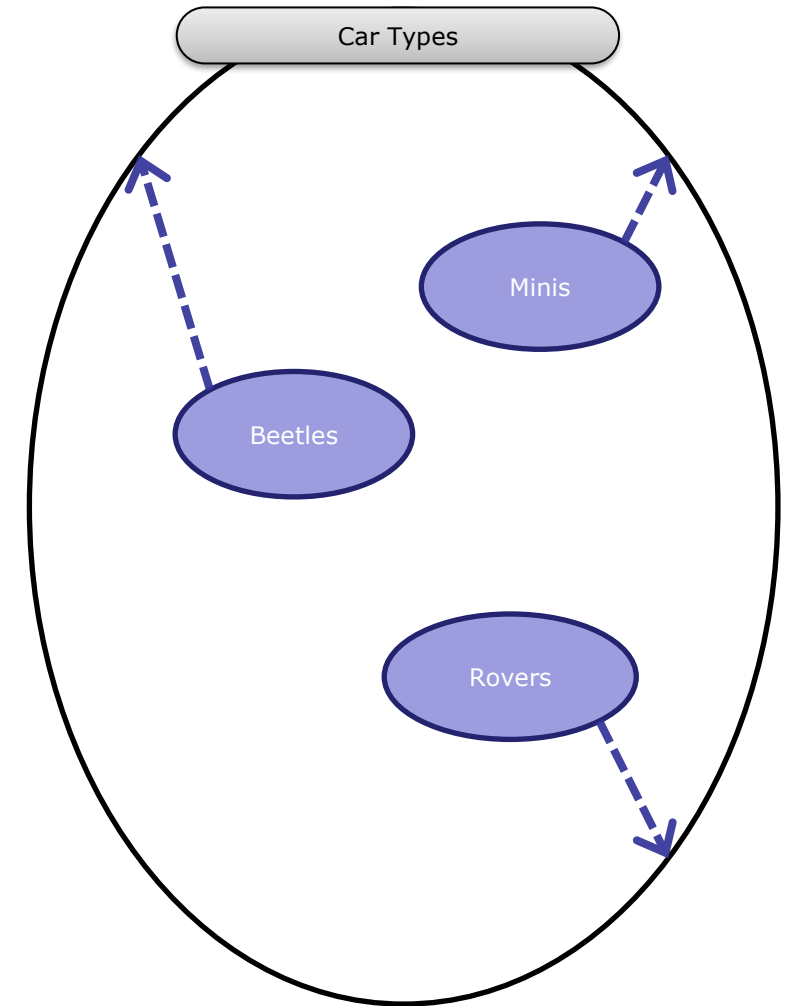
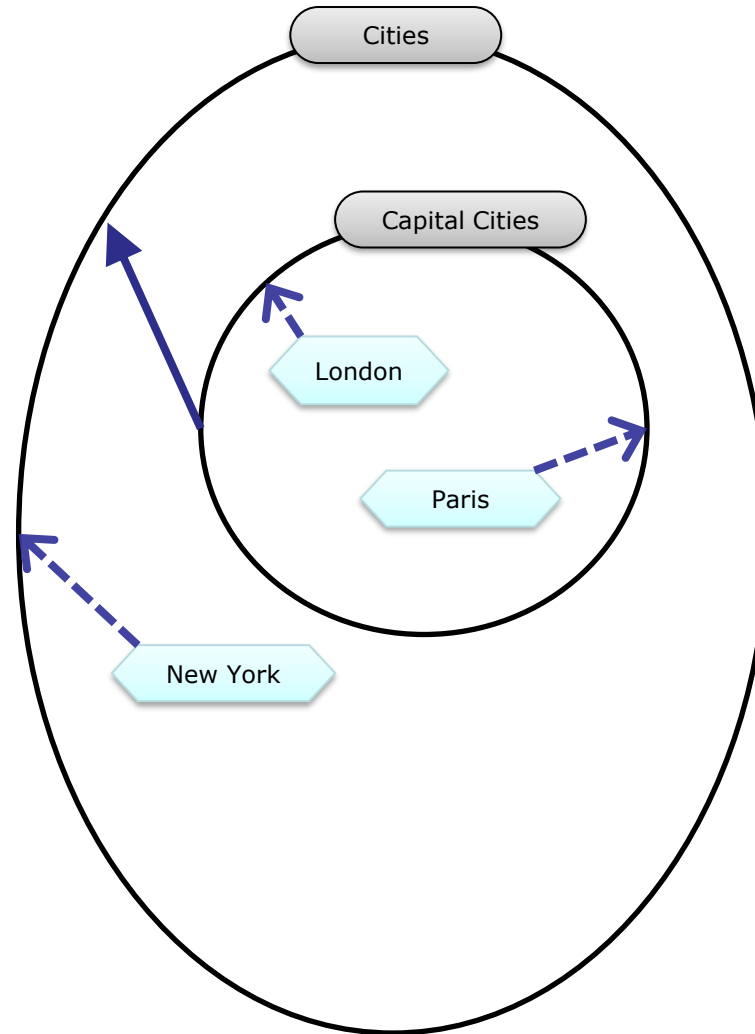
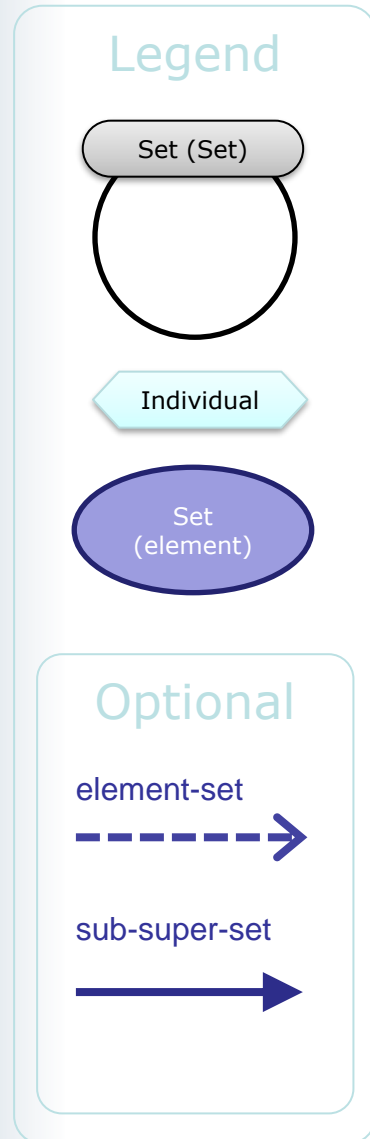
Figure 10.15 - Pattern for classes  
From *Business Objects: Re-engineering for Reuse*

# Three components of an Euler set

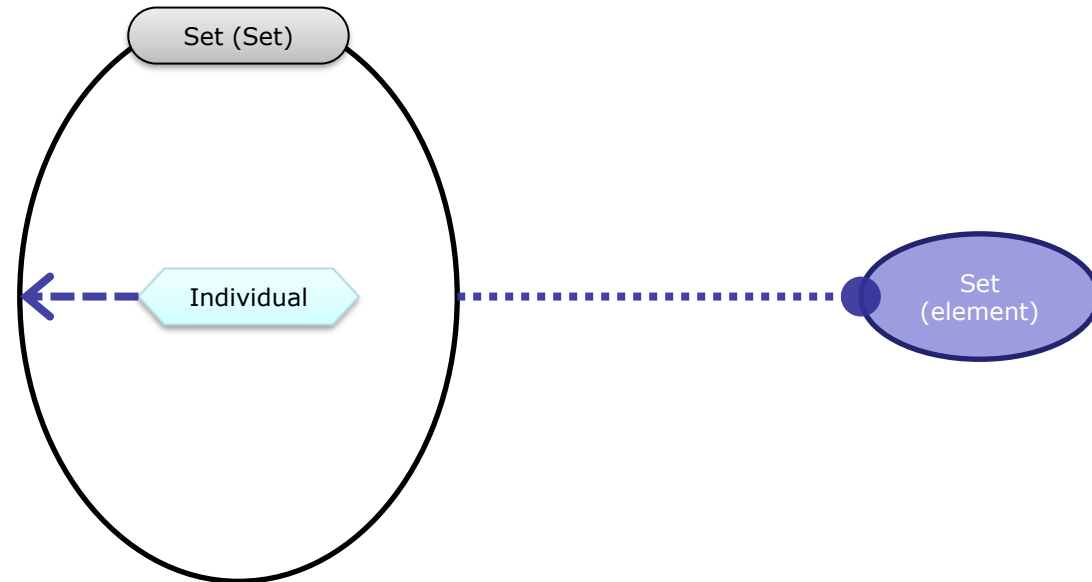
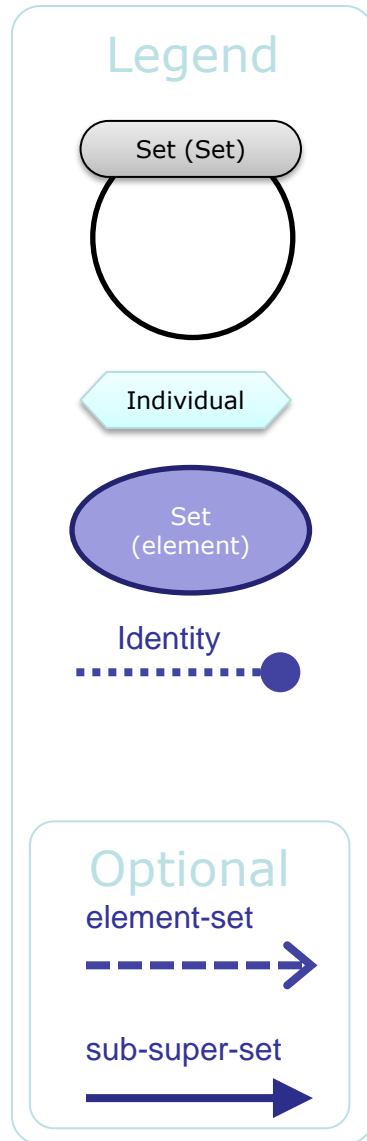
A region representing a collection of objects	A circle around the region marking the region as representing a type – the <i>contour</i>	A string of letters labelling the type
		A
		



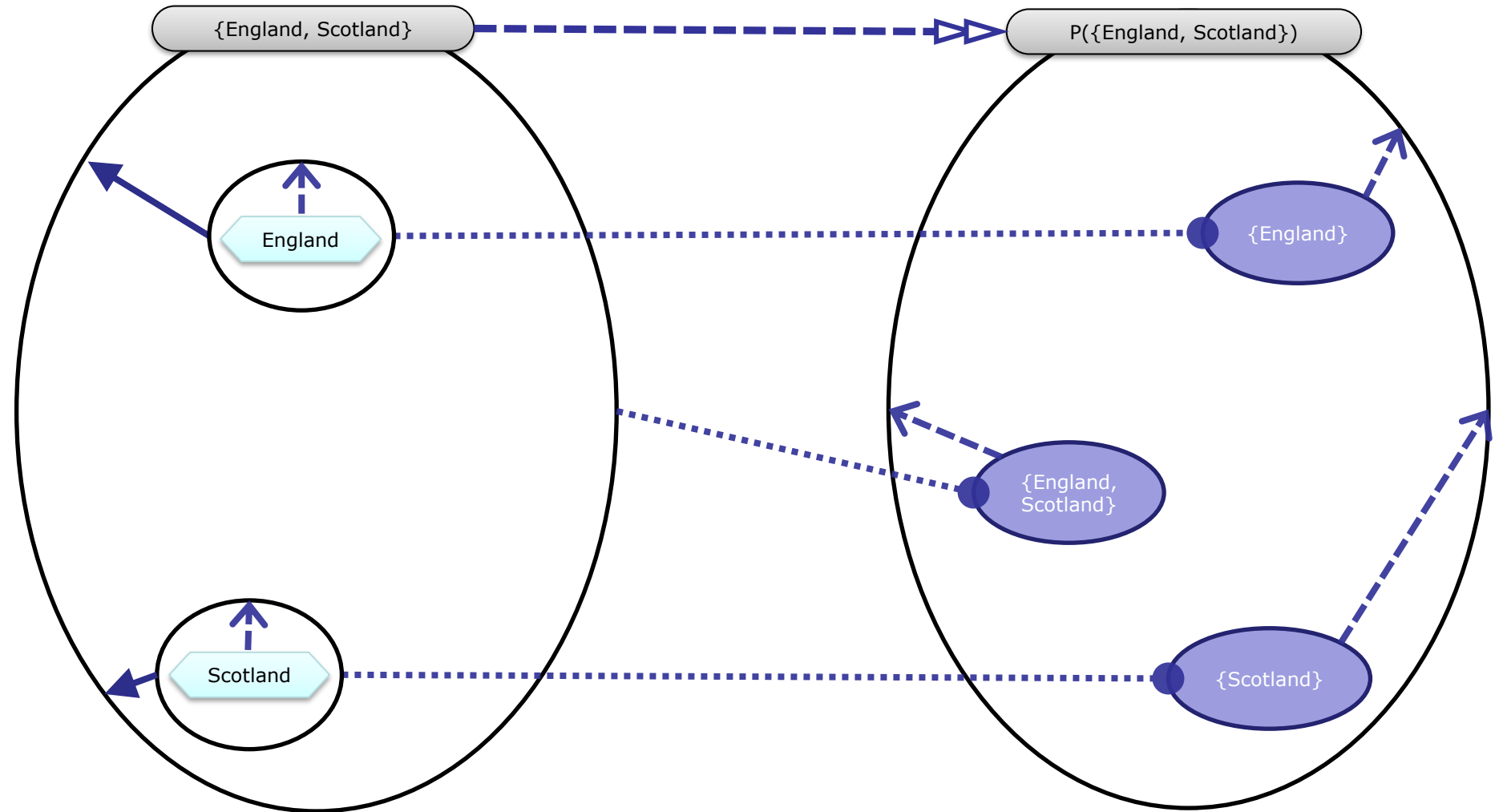
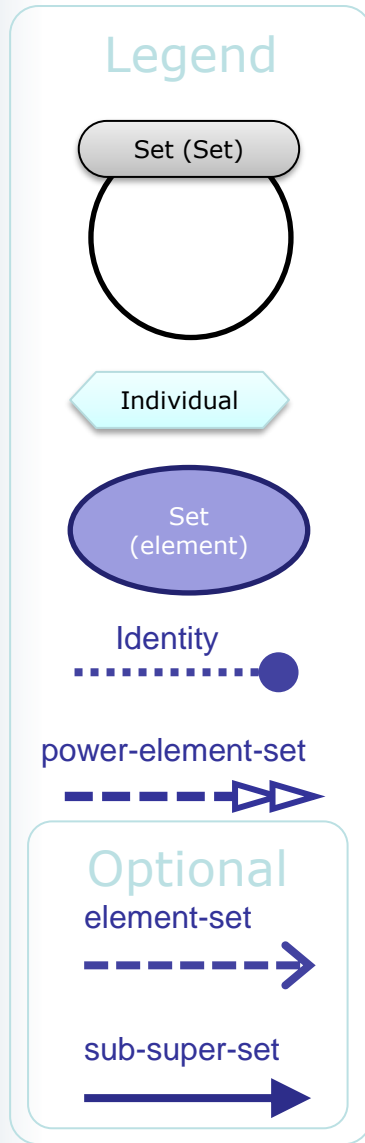
# Example simple order type OEDs



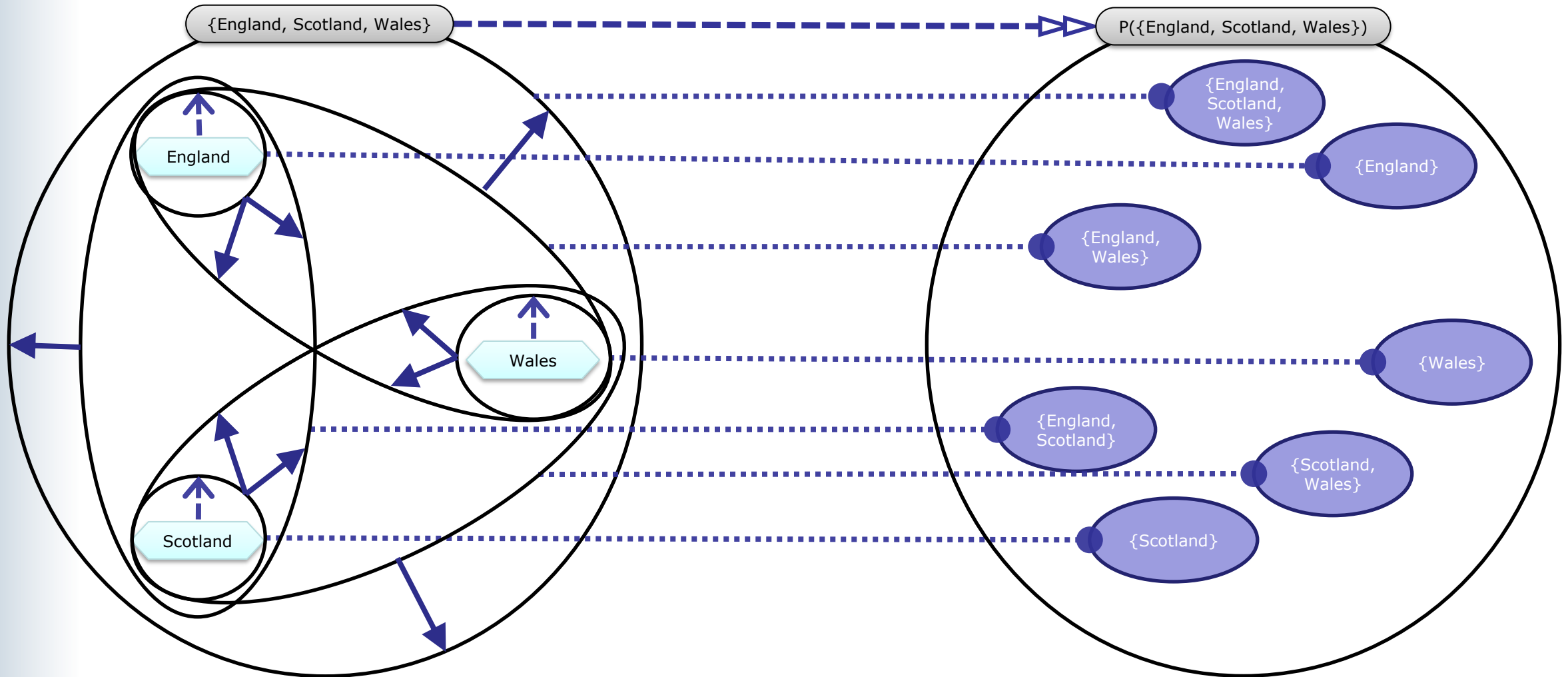
# Type ascending



# OED powerset pattern: two member types



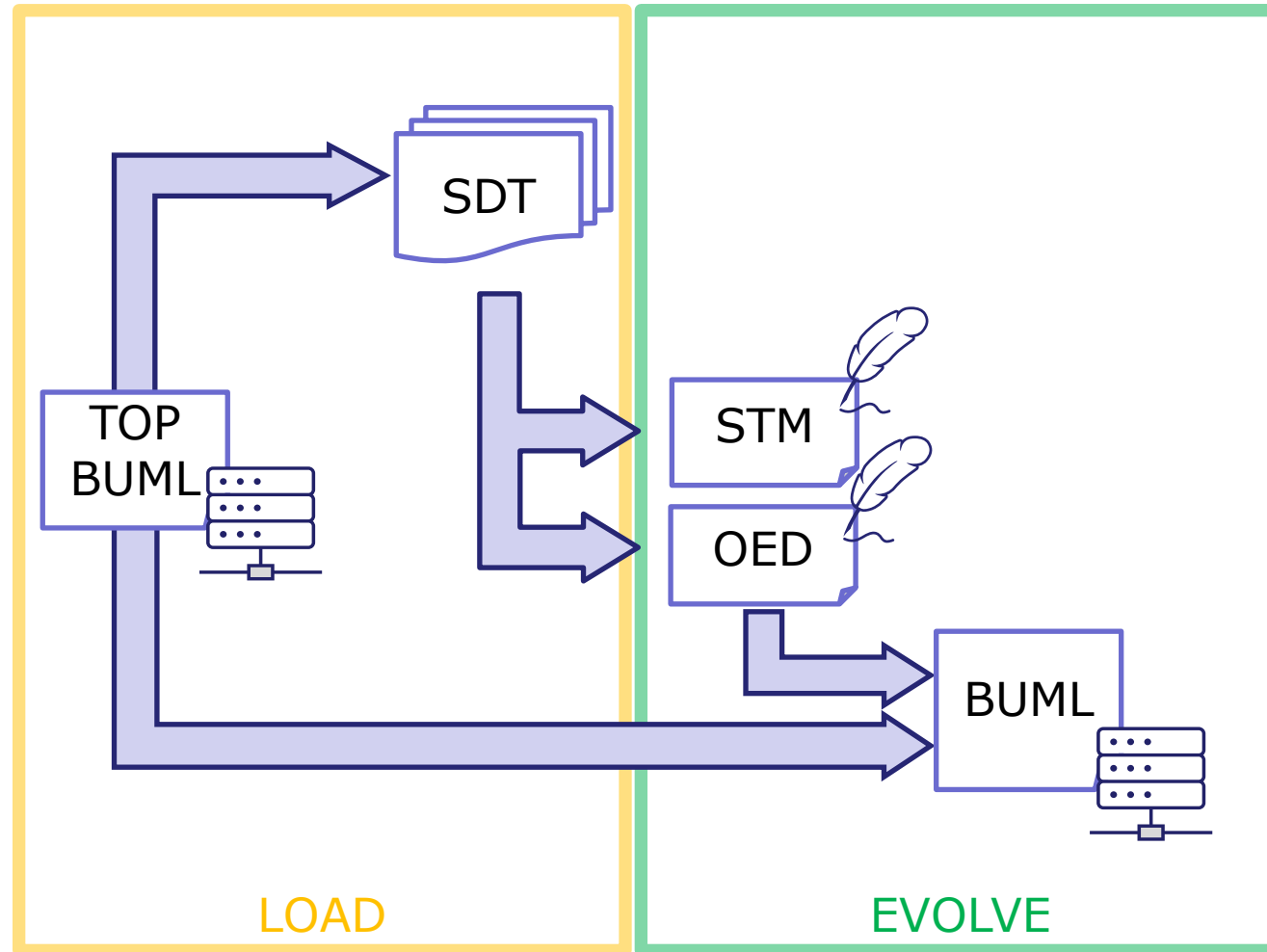
# Multiple order patterns: OED powerset pattern





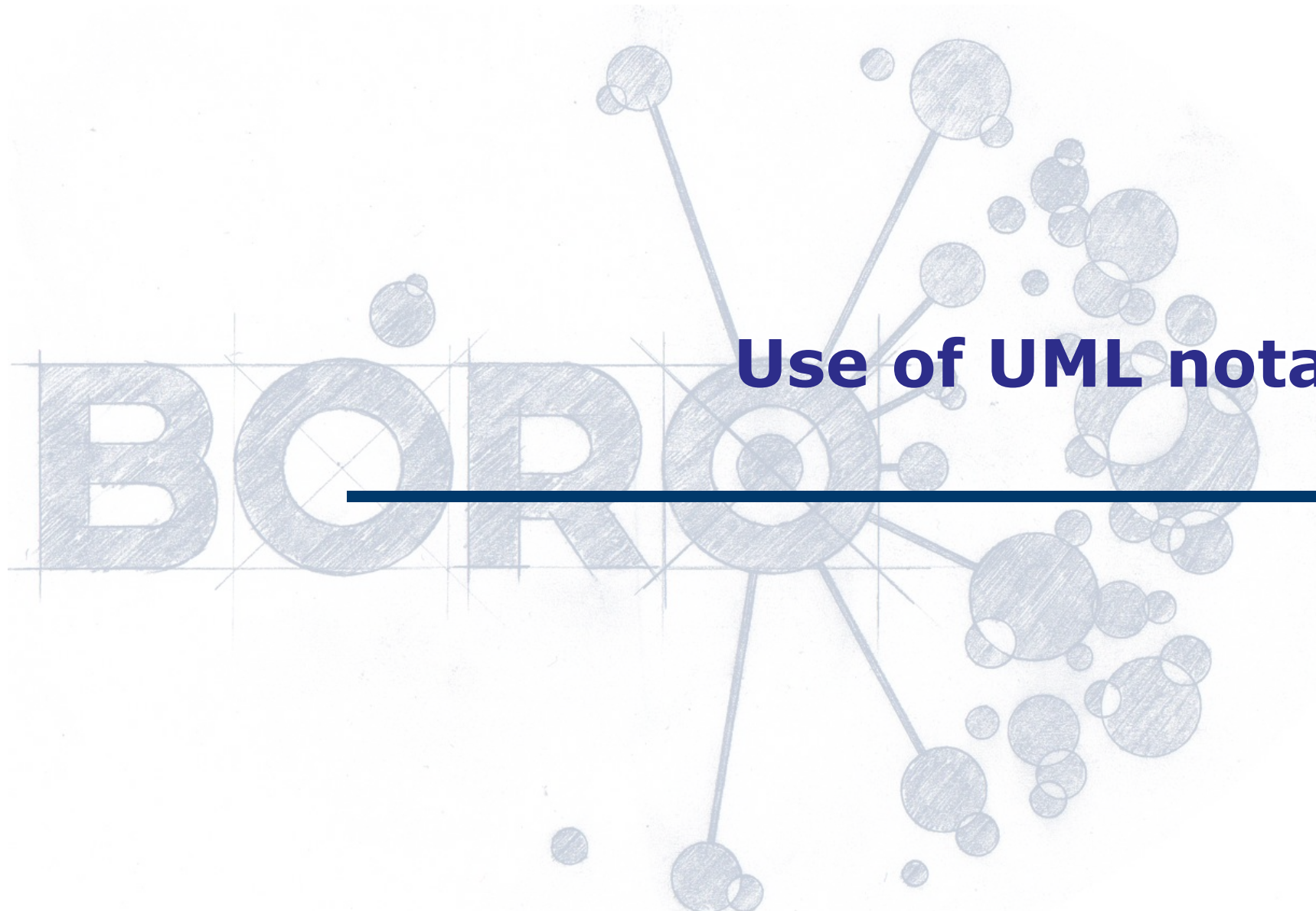
**BORO UML**

# Practical problem process



- BUML is the BORO variant of UML
- We use UML notation to visualise BORO ontological structure
  - UML classes
  - UML generalizations, dependencies, n-ary associations and associations
- In addition, we make use of stereotypes, for various purposes, e.g., to indicate the ontological category of the represented objects
- All of these are used with a specific sense, which sometimes may deviate from the UML conventions. The BUML sense is defined by the underlying constructional semantics
- We use Enterprise Architect to draw diagrams, but other software options are available



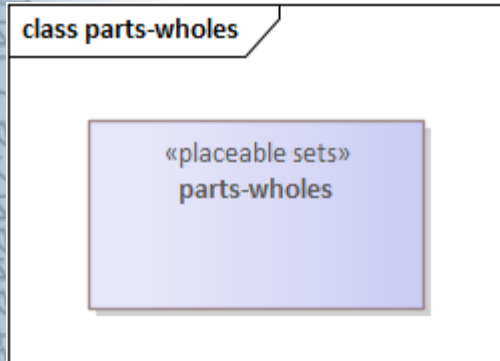


## Use of UML notation

# Use of UML classes

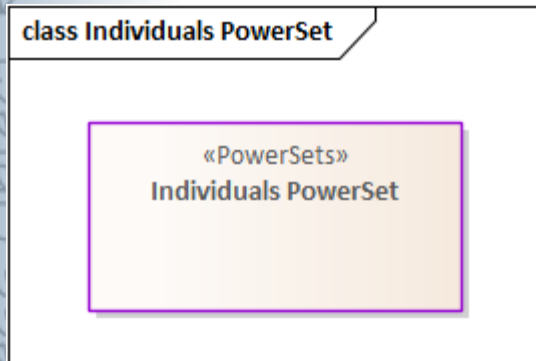


We use UML classes to represent sets, e.g., sets of Countries.



Sometimes, sets contain higher-order objects, like tuples.

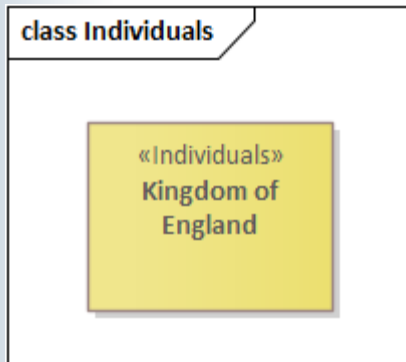
*Stereotype <<placeable sets>> refers to the ontic category of the represented objects,  
But we do not expect you to follow this convention in your work study.*



Sets may also contain sets.

*Stereotype <<PowerSets>> refers to the ontic category of the represented objects,  
But we do not expect you to follow this convention in your work study.*

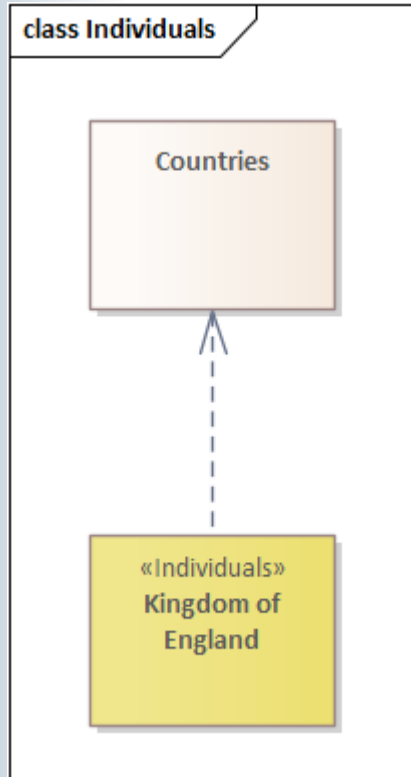
# Use of UML classes



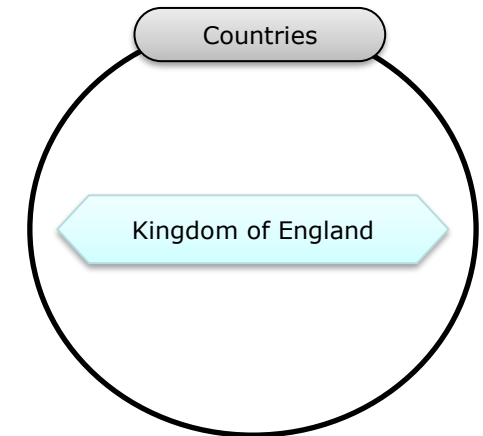
Also we use UML classes to represent Individuals.

*Stereotype <<Individuals>> shows the ontic category of the represented object, we do not expect you to follow this convention in your work study.*

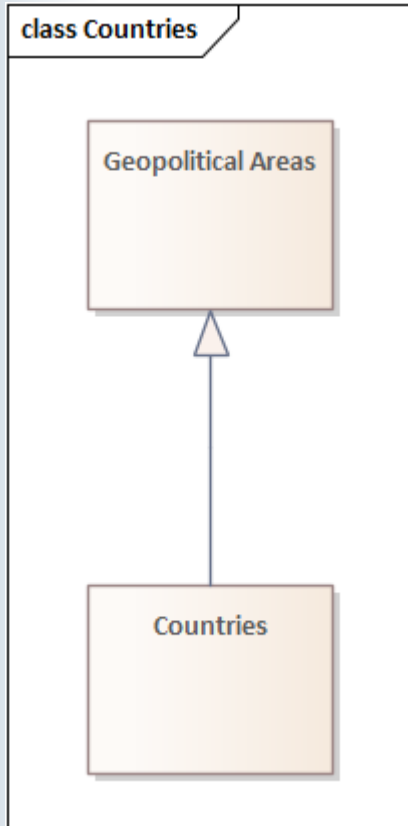
# Use of UML dependencies



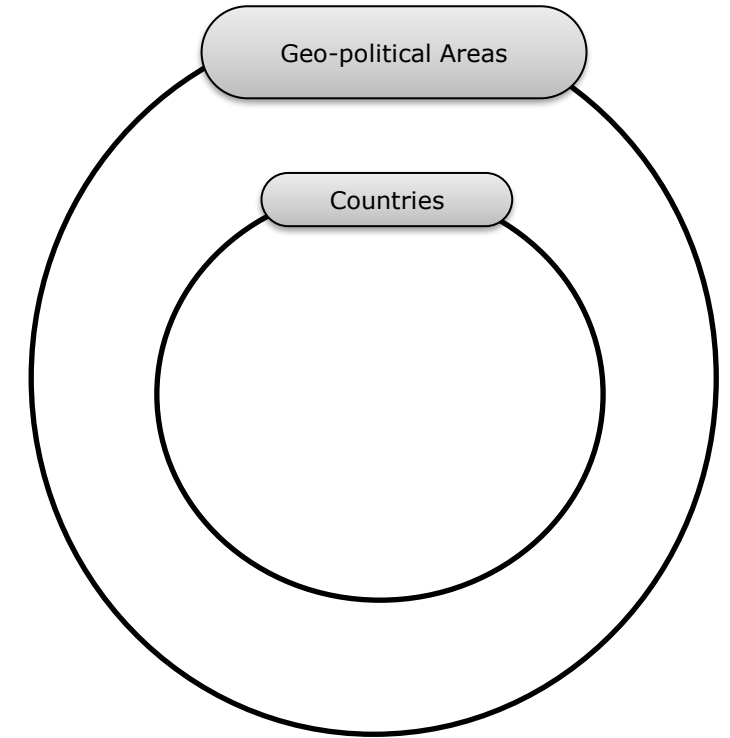
Dependency link shows element-sets (tuple).



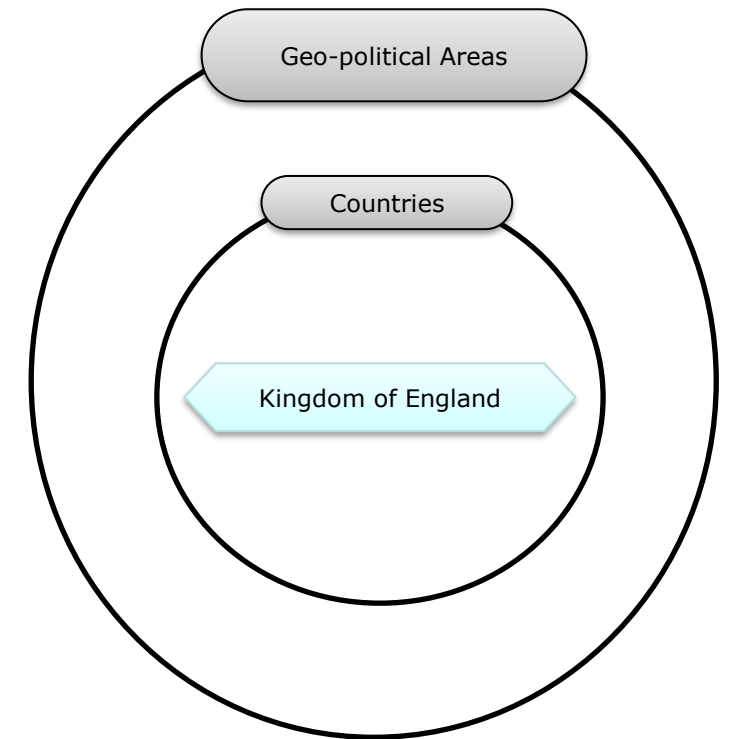
# Use of UML generalizations



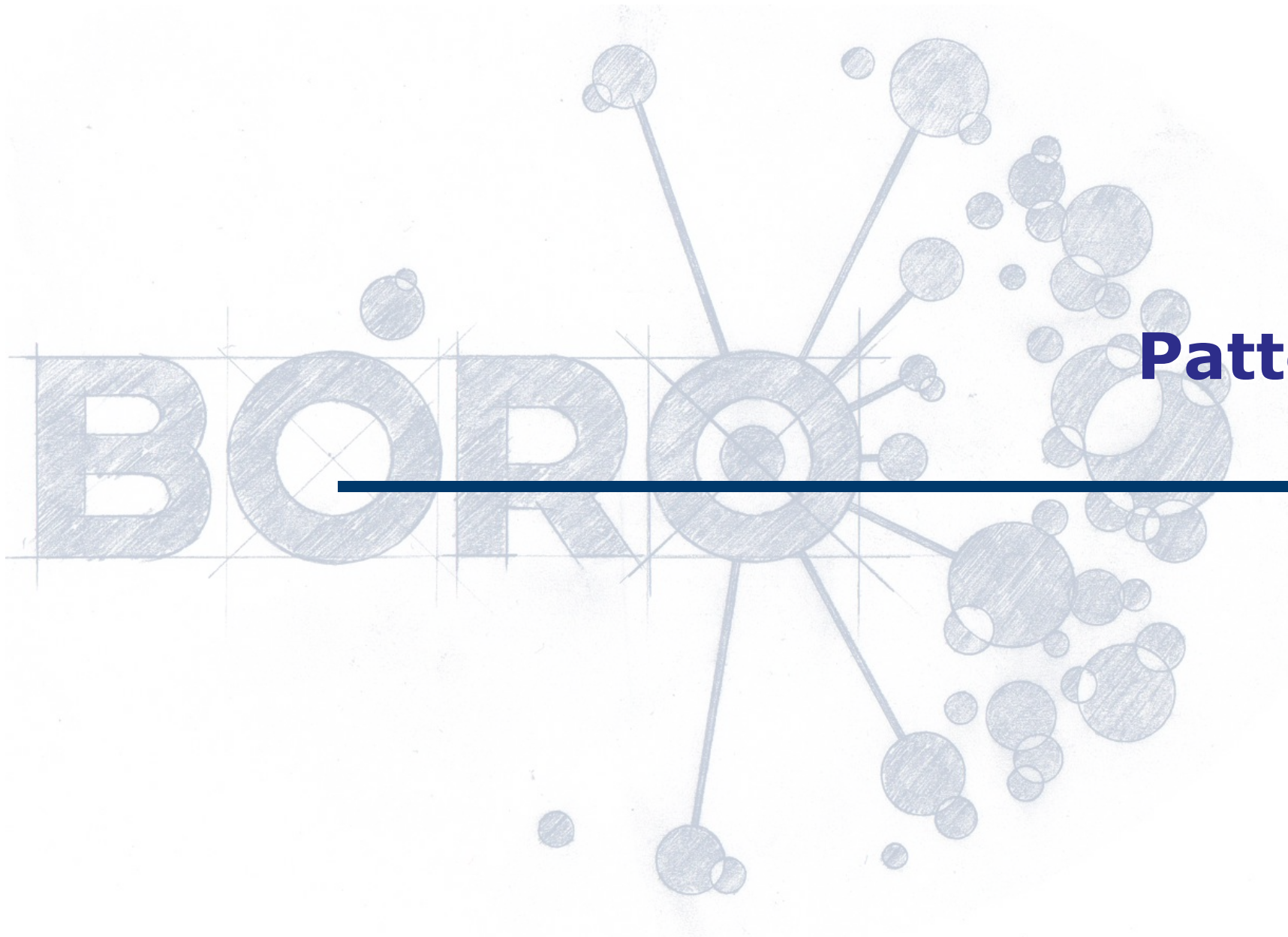
Generalization link shows sub-super-sets (tuple).



# Use of UML dependencies and generalizations





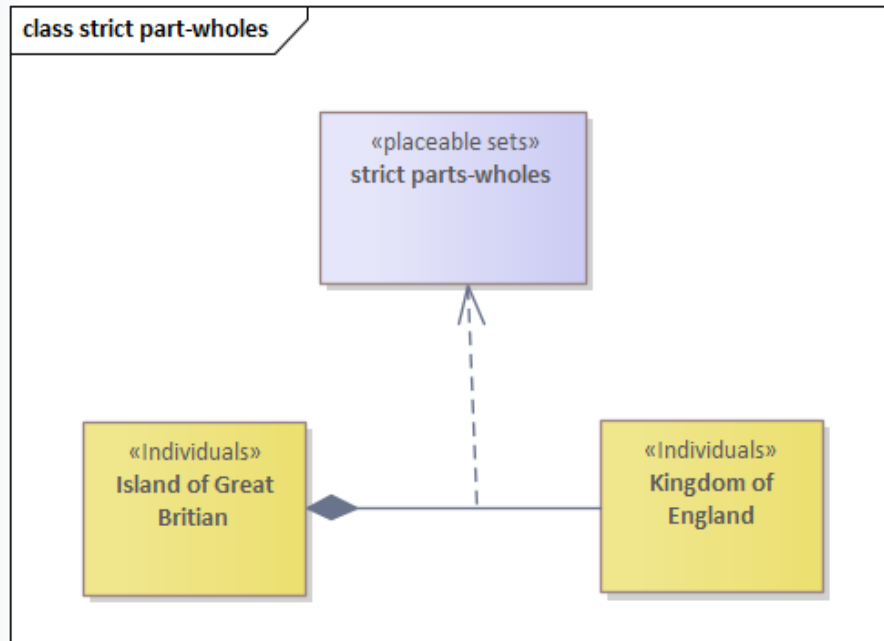


## Patterns



# Parts-wholes - basic pattern

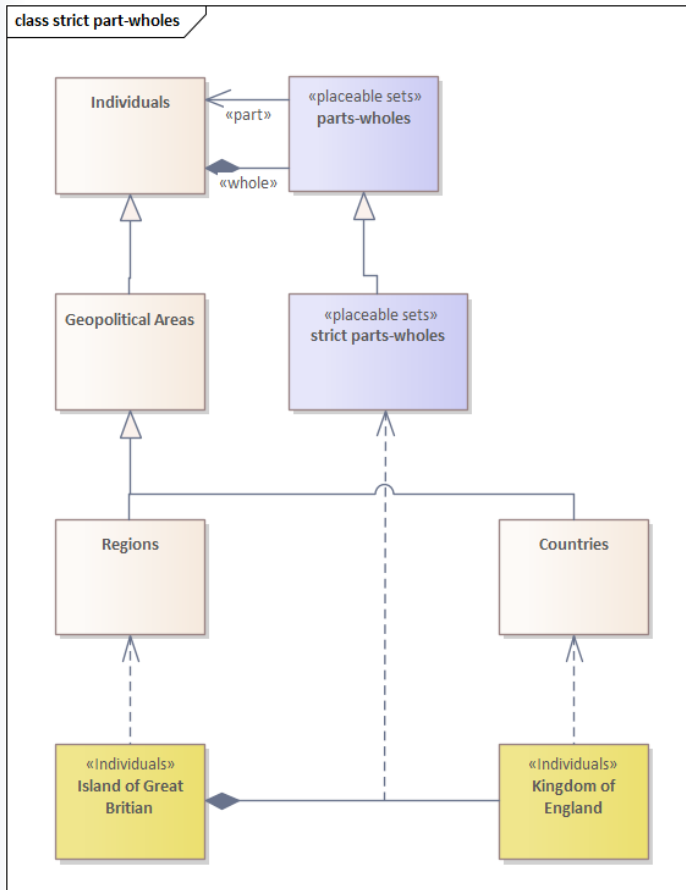
strict = proper (excl. identity)



UML composition shows part-whole (tuple).

<<place1>> and <<place2>> association stereotypes show, respectively, the whole and the part.

# Parts-wholes - full pattern



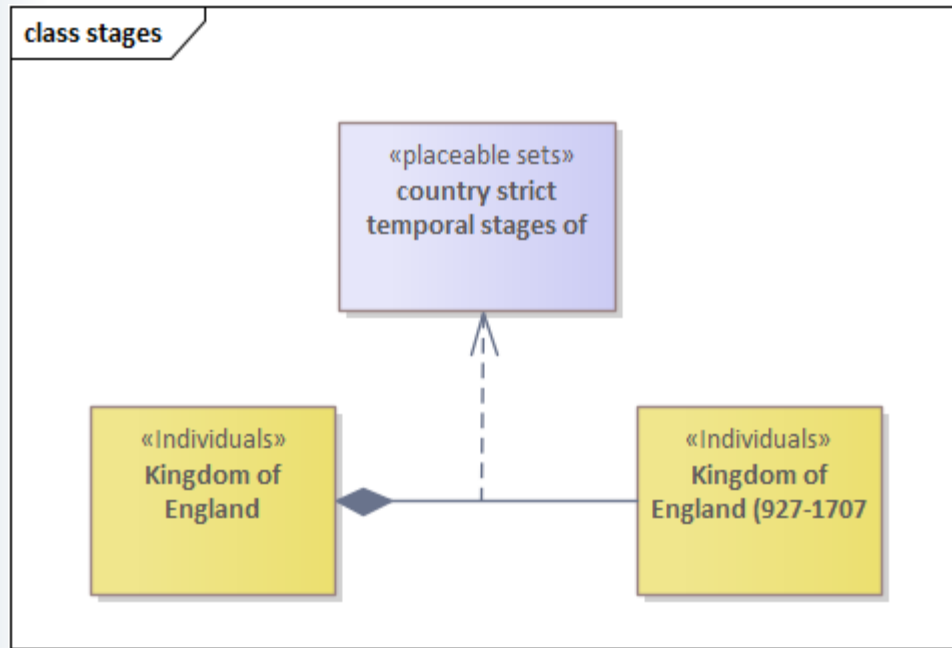
parts-wholes includes identity tuples, e.g., <London, London>.



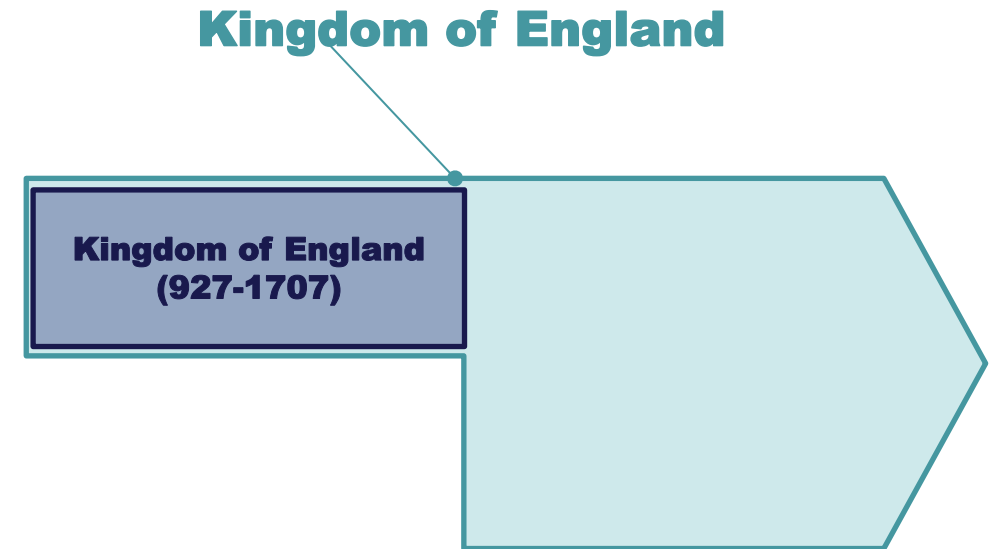
strict parts-wholes is a **set** of tuples, so it can have supersets (and subsets).

<<part>> and <<whole>> association stereotypes show, respectively, the whole and the part.

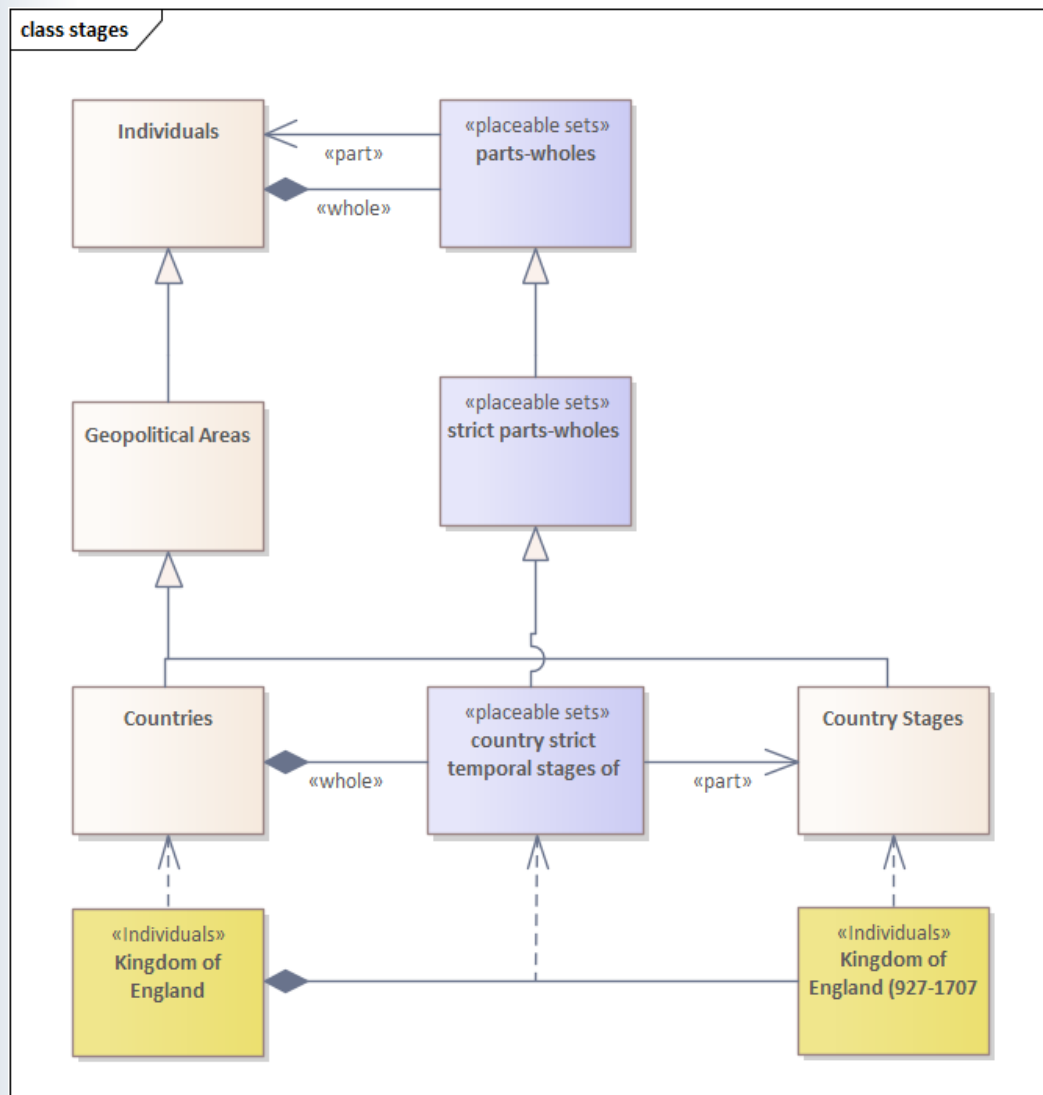
# Temporal stages – basic pattern



UML composition shows the temporal-stages (tuple).



# Temporal stages – full pattern

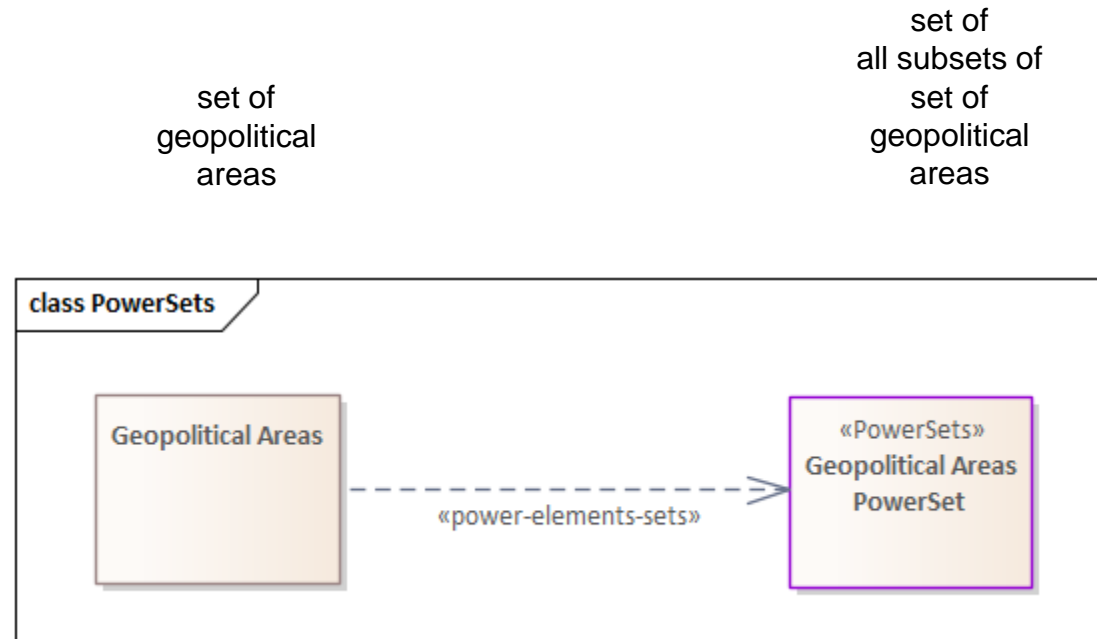


**Kingdom of England**



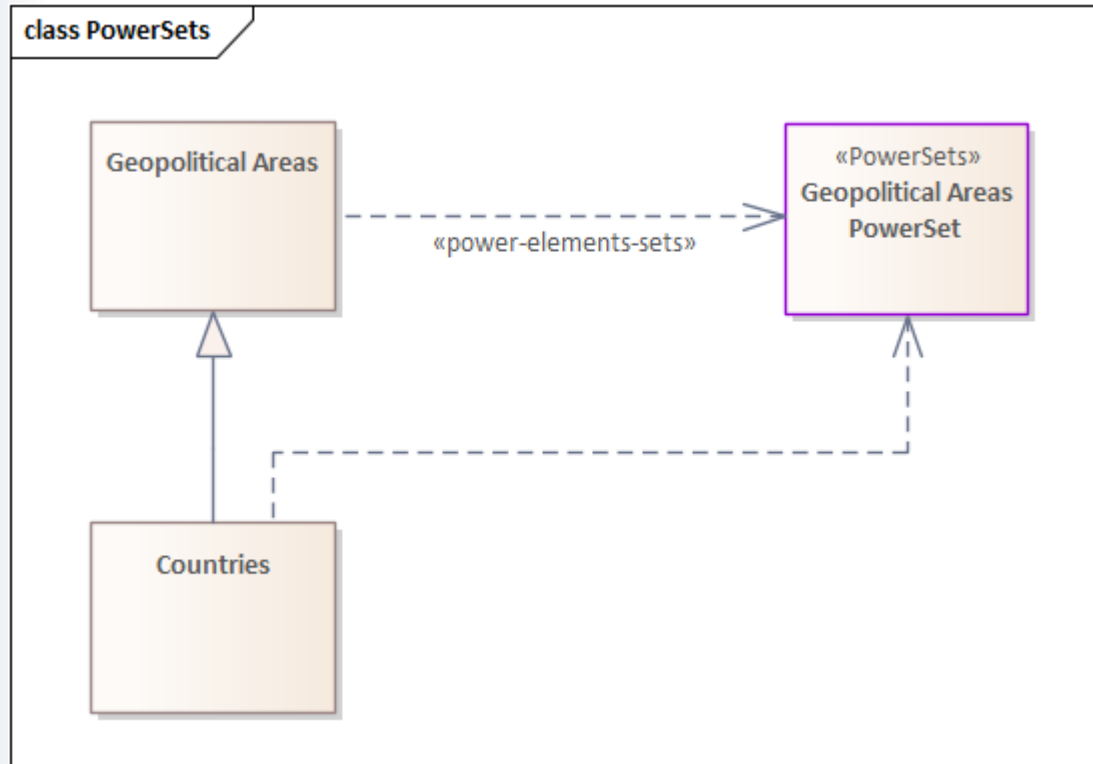
country strict temporal stages of is a **set** of tuples, so it can have supersets (and subsets)

# Powerset pattern



dependency stereotype shows the set and its powerset

# Powerset pattern – implicit tuples

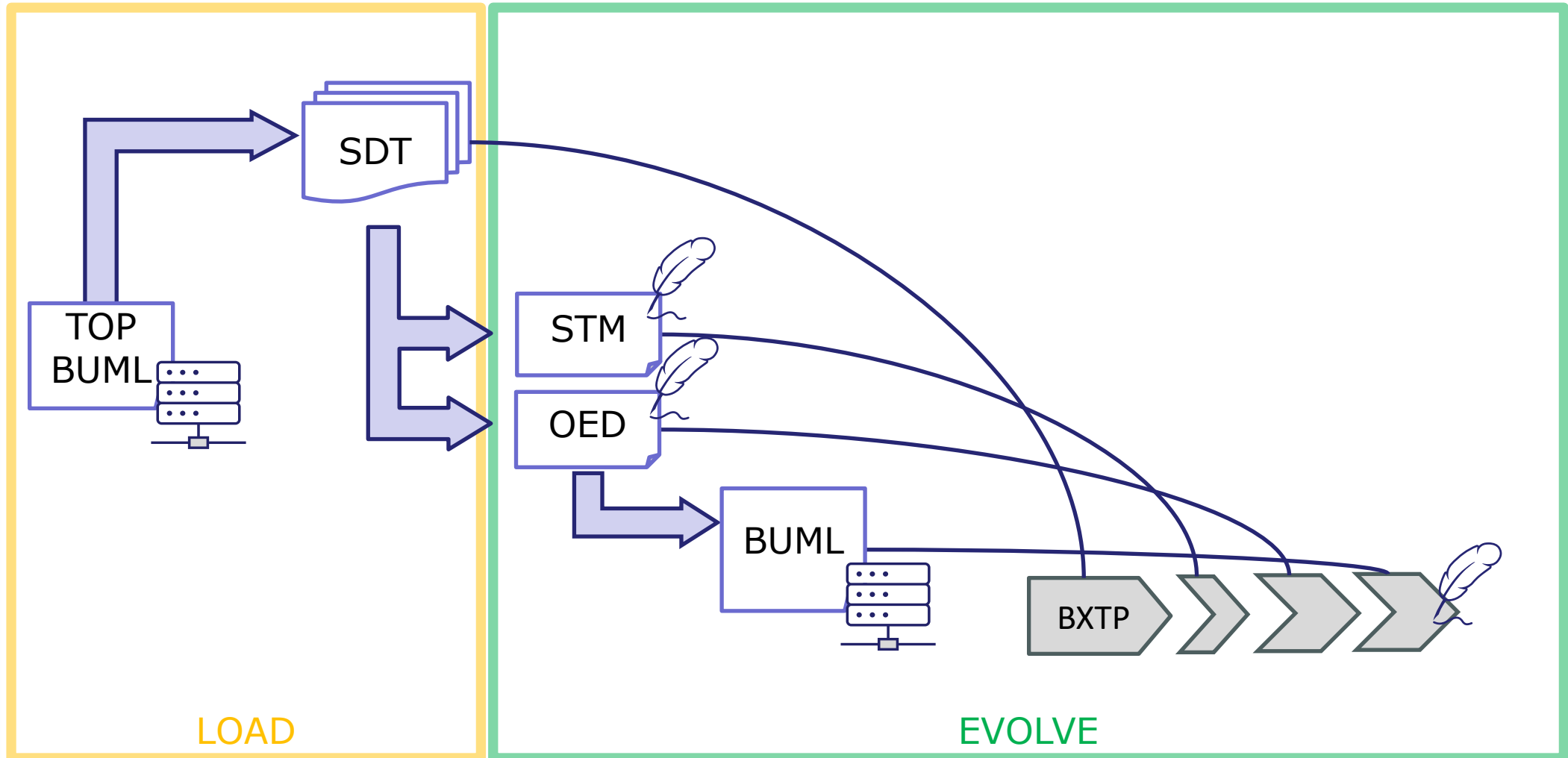


Dependency shows implicit element-set (tuple).

# **BORO eXcel Table (Manual) Pipeline**

---

# Practical problem process





# Micro transformation pipeline architecture

---

- ④ Define by contrast
  - not a single monolithic structure
    - for example, a single complex transformation between two schemas
- ④ Pipeline consists of a network of simple micro transformations
  - the micro transformations are:
    - algorithmic – they can be coded
    - simple to allow for maximal inspectability
    - are repeatable
    - aim to be reusable
      - there is an 'art' to finding high levels of reusability

# 'Formal' transformation patterns

---

- Repeatable, reusable 'formal' patterns
  - can use same computer code
    - with enough work

# LOAD – bie-ise identifier pattern

	A	B
1	path_names	fso_types
2	D:\folder_1	folder
3	D:\folder_1\file1.txt	file
4	D:\folder_1\file2.txt	file
5	D:\folder_1\file3.txt	file
6	D:\folder_2	folder
7	D:\folder_2\file4.txt	file
8	D:\folder_3	folder
9	D:\folder_3\file5.txt	file
10	D:\folder_3\file6.txt	file
11		

< > a\_file\_system\_objects

	A	B	C
1	bie_ids	path_names	fso_types
2	K7TGBEHQT763L	D:\folder_1	folder
3	NIPIIQHJWF46B	D:\folder_1\file1.txt	file
4	G3MKPAJIN4E5O	D:\folder_1\file2.txt	file
5	2XIEV/LZZPSVZ	D:\folder_1\file3.txt	file
6	LUT4KTDY4Q3AE	D:\folder_2	folder
7	7BV66YAOOAPSA	D:\folder_2\file4.txt	file
8	YMQMY6AZ4A3SO	D:\folder_3	folder
9	3RIIKMPJMO2W4	D:\folder_3\file5.txt	file
10	X6KVOFQU63LAQ	D:\folder_3\file6.txt	file
11			

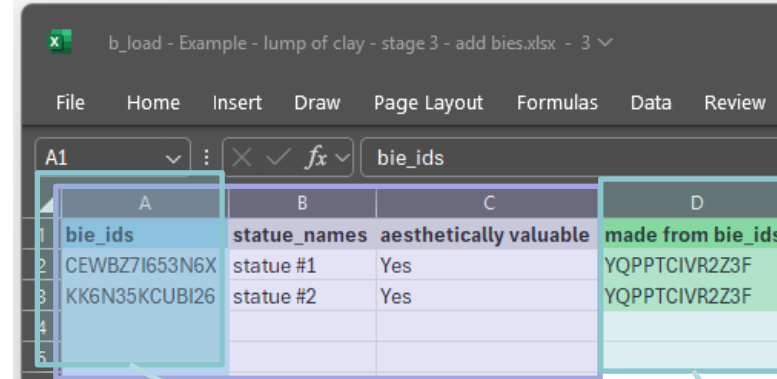
< > file\_system\_objects files folders

- A column is added to the beginning of the table
- Its name is 'bie\_ids'
  - The cells uniquely identify the row of the table
- No changes are made to the other columns

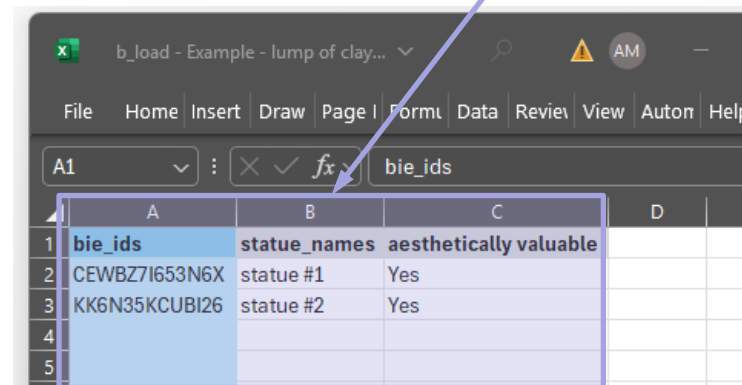
Note: the identifying `bie_ids` column is 'inherited' by both split tables.  
The pattern of inheritance can be of two types.

1. Identifier preserving
2. Identifier introducing.

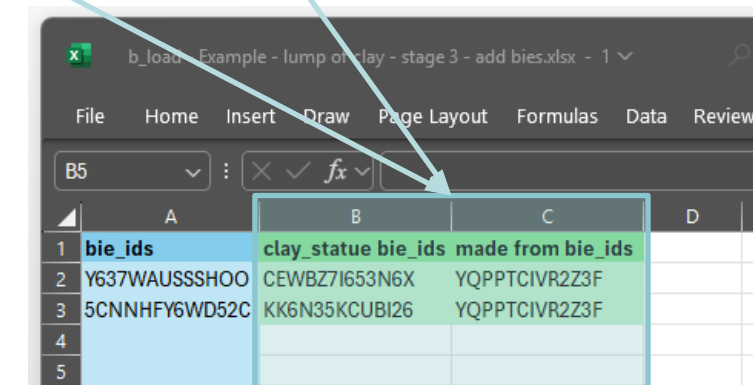
This is driven by the intuition of the intended identity.



bie_ids	statue_names	aesthetically valuable	made from bie_ids
CEWBZ7I653N6X	statue #1	Yes	YQPPTCIVR2Z3F
KK6N35KCUBI26	statue #2	Yes	YQPPTCIVR2Z3F



bie_ids	statue_names	aesthetically valuable
CEWBZ7I653N6X	statue #1	Yes
KK6N35KCUBI26	statue #2	Yes



clay_statue	bie_ids	made from bie_ids
Y637WAUSSSHOO	CEWBZ7I653N6X	YQPPTCIVR2Z3F
5CNNHFY6WD52C	KK6N35KCUBI26	YQPPTCIVR2Z3F

# Row split pattern

	A	B	C
1	<b>bie ids</b>	<b>path names</b>	<b>fso types</b>
2	K7TGBEHQT763L	D:\folder_1	folder
3	NIPIIQHJWF46B	D:\folder_1\file1.txt	file
4	G3MKPAJIN4E5O	D:\folder_1\file2.txt	file
5	2XIEV7LZZPSVZ	D:\folder_1\file3.txt	file
6	LUT4KTDY4Q3AE	D:\folder_2	folder
7	7BV66YAOOAPSA	D:\folder_2\file4.txt	file
8	YMOMY6AZ4A3SO	D:\folder_3	folder
9	3RIIKMPJMO2W4	D:\folder_3\file5.txt	file
10	X6KVOFQU63LAQ	D:\folder_3\file6.txt	file
11			

file\_system\_objects | files | folders

	A	B	C
1	<b>bie_ids</b>	<b>path_names</b>	<b>fso_types</b>
2	NIPIIQHJWF46B	D:\folder_1\file1.txt	file
3	G3MKPAJIN4E5O	D:\folder_1\file2.txt	file
4	2XIEV7LZZPSVZ	D:\folder_1\file3.txt	file
5	7BV66YAOOAPSA	D:\folder_2\file4.txt	file
6	3RIIKMPJMO2W4	D:\folder_3\file5.txt	file
7	X6KVOFQU63LAQ	D:\folder_3\file6.txt	file
8			

file\_system\_objects | **files** | folders

	A	B	C
1	<b>bie_ids</b>	<b>path_names</b>	<b>fso_types</b>
2	K7TGBEHQT763L	D:\folder_1	folder
3	LUT4KTDY4Q3AE	D:\folder_2	folder
4	YMOMY6AZ4A3SO	D:\folder_3	folder
5			

file\_system\_objects | files | **folders**

Typically:

- The column heading row is in ALL output tables.
  - The output tables all have the same format
- All rows are in one or other of the split tables.

# Implicit foreign key bie-ise pattern – stage 1 – add bie\_ids

	A	B	C	D	E	F
1	<b>bie_ids</b>	<b>statue_names</b>	<b>aesthetically valuable</b>	<b>made from</b>	<b>made on</b>	<b>destroyed on</b>
2	CEWBZ7I653N6X	statue #1	Yes	lump of clay #1	2023-06-20	2023-10-24
3	KK6N35KCUBI26	statue #2	Yes	lump of clay #1	2024-07-06	
4						
5						
6						

Match the value in the Foreign table

	A	B	C
1	<b>bie_ids</b>	<b>time_names</b>	<b>YYYY-MM-DD</b>
2	R7E3FV2GKUN5F	t1	2023-06-20
3	Y2IXUKCPJUS6Y	t2	2023-10-24
4	24J67PH4E27UG	t3	2024-07-06
5			
6			

Some columns will be foreign keys  
This needs to be made explicit

Add bie id to new column

	A	B	C	D	E	F	G	H	I
1	<b>bie_ids</b>	<b>statue_names</b>	<b>aesthetically valuable</b>	<b>made from</b>	<b>made from bie_ids</b>	<b>made on</b>	<b>made on bie_ids</b>	<b>destroyed on</b>	<b>destroyed on bie_ids</b>
2	CEWBZ7I653N6X	statue #1	Yes	lump of clay #1	YQPPTCIVR2Z3F	2023-06-20	R7E3FV2GKUN5F	2023-10-24	Y2IXUKCPJUS6Y
3	KK6N35KCUBI26	statue #2	Yes	lump of clay #1	YQPPTCIVR2Z3F	2024-07-06	24J67PH4E27UG		
4									
5									
6									

# Implicit foreign key bie-ise pattern – stage 2 – remove original columns

	A	B	C	<del>D</del>	E	<del>F</del>	G	<del>H</del>	I
1	<b>bie_ids</b>	statue_names	aesthetically valuable	<del>made from</del>	made from bie_ids	<del>made on</del>	made on bie_ids	<del>destroyed on</del>	destroyed on bie_ids
2	CEWBZ7I653N6X	statue #1	Yes	<del>lump of clay #1</del>	YQPPTCIVR2Z3F	<del>2023-06-20</del>	R7E3FV2GKUN5F	<del>2023-10-24</del>	Y2IXUKCPJUS6Y
3	KK6N35KCUBI26	statue #2	Yes	<del>lump of clay #1</del>	YQPPTCIVR2Z3F	2024-07-06	24J67PH4E27UG		
4									
5									
6									

b1\_clay\_statues | b1\_times | c1\_clay\_statues | c2\_clay\_statues | c3\_c ... +



	A	B	C	D	E	F
1	<b>bie_ids</b>	statue_names	aesthetically valuable	made from bie_ids	made on bie_ids	destroyed on bie_ids
2	CEWBZ7I653N6X	Clay Statue #1	Yes	YQPPTCIVR2Z3F	R7E3FV2GKUN5F	Y2IXUKCPJUS6Y
3	KK6N35KCUBI26	Clay Statue #2	Yes	YQPPTCIVR2Z3F	24J67PH4E27UG	
4						
5						

b\_sub-super-sets | c1\_clay\_statues | c2\_clay\_statues | c3\_clay\_statues | c3\_statues\_temporally

# Table-Column heading pushdown pattern

	A	B
1	bie_ids	table_names
2	M5HLU3UUJ727D	lumps_of_clay
3	CBFGCPVTIUQ7R	clay_statues
4		
5		
6		
7		

stages tables

Add table to Objects

	A	B
1	bie_ids	object names
2	M5HLU3UUJ727D	lumps_of_clay
3	YQPPTCIVR2Z3F	
4	CBFGCPVTIUQ7R	clay_statues
5	CEWBZ7I653N6X	
6	KK6N35KCUBI26	
7		

c5\_objects c5\_ta

Add row to Objects

	A	B	C
1	bie_ids		
2	CEWBZ7I653N6X		
3	KK6N35KCUBI26		
4			

c4 clay statues

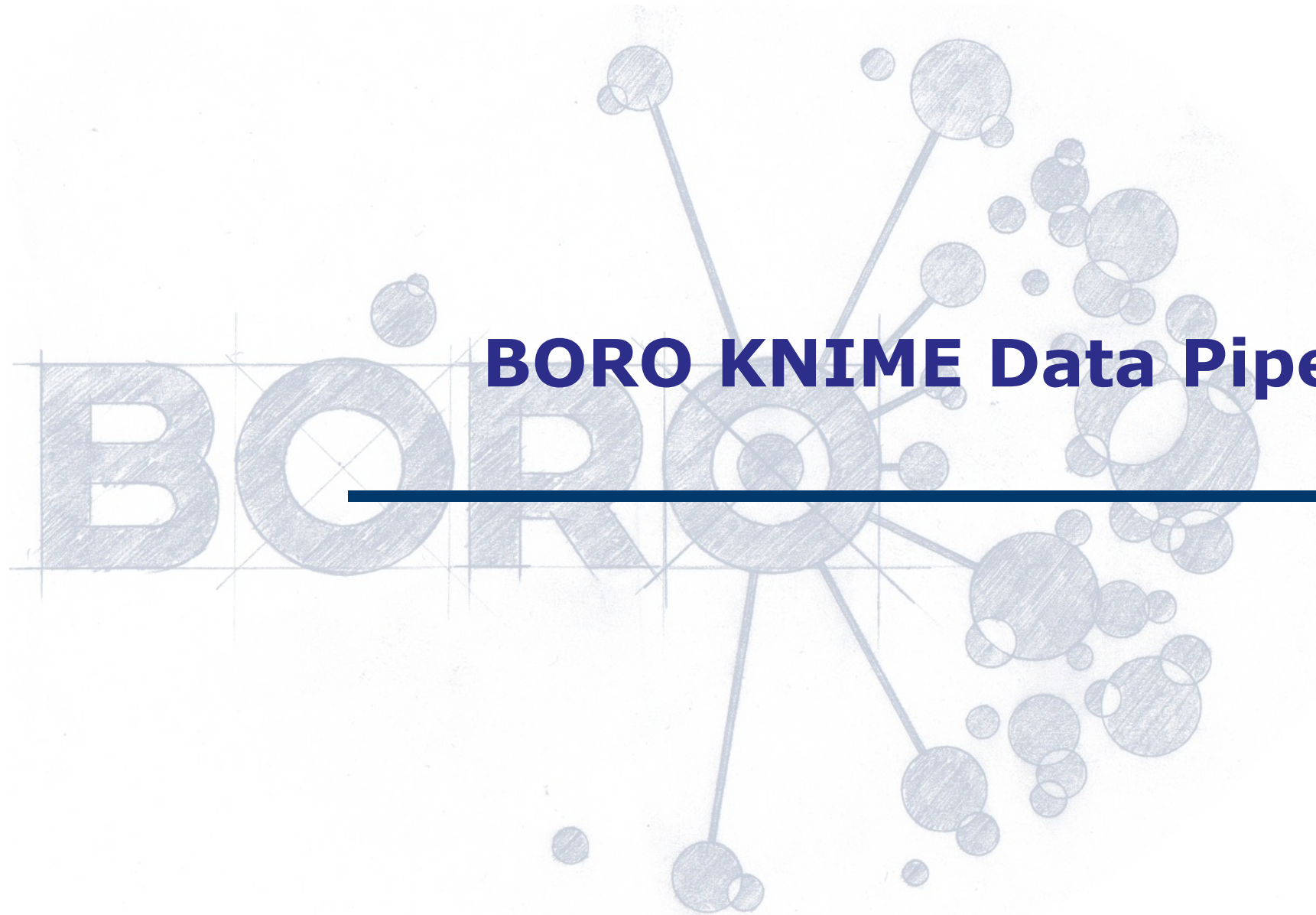
And add the relations

	A	B	C
1	table_bie_ids	row_bie_ids	
2	M5HLU3UUJ727D	YQPPTCIVR2Z3F	
3	CBFGCPVTIUQ7R	CEWBZ7I653N6X	
4	CBFGCPVTIUQ7R	KK6N35KCUBI26	
5			
6			
7			

c5\_table-row\_relations

Tables and column headings can represent objects  
Useful to have these as rows in the domain.

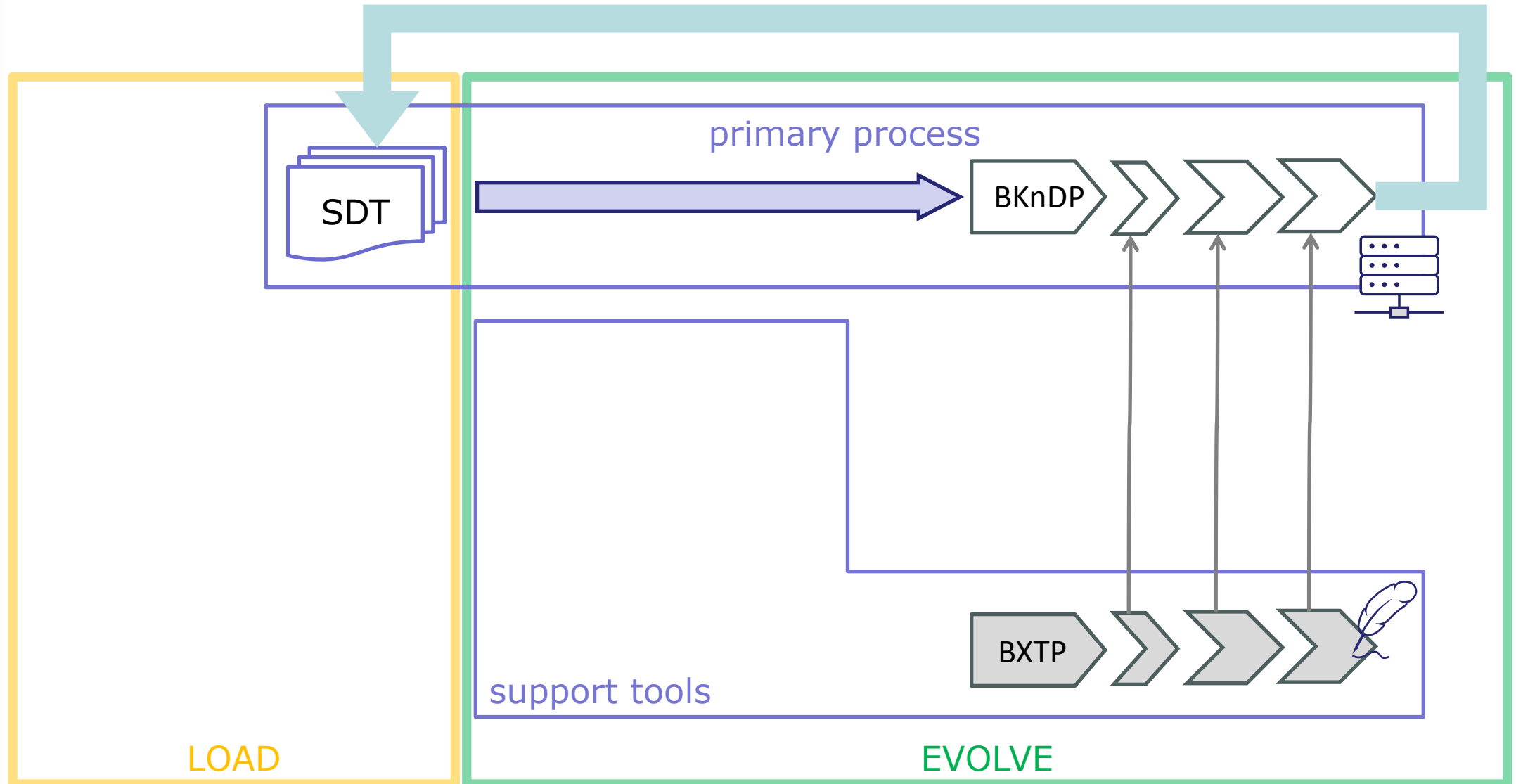




# **BORO KNIME Data Pipeline**

---

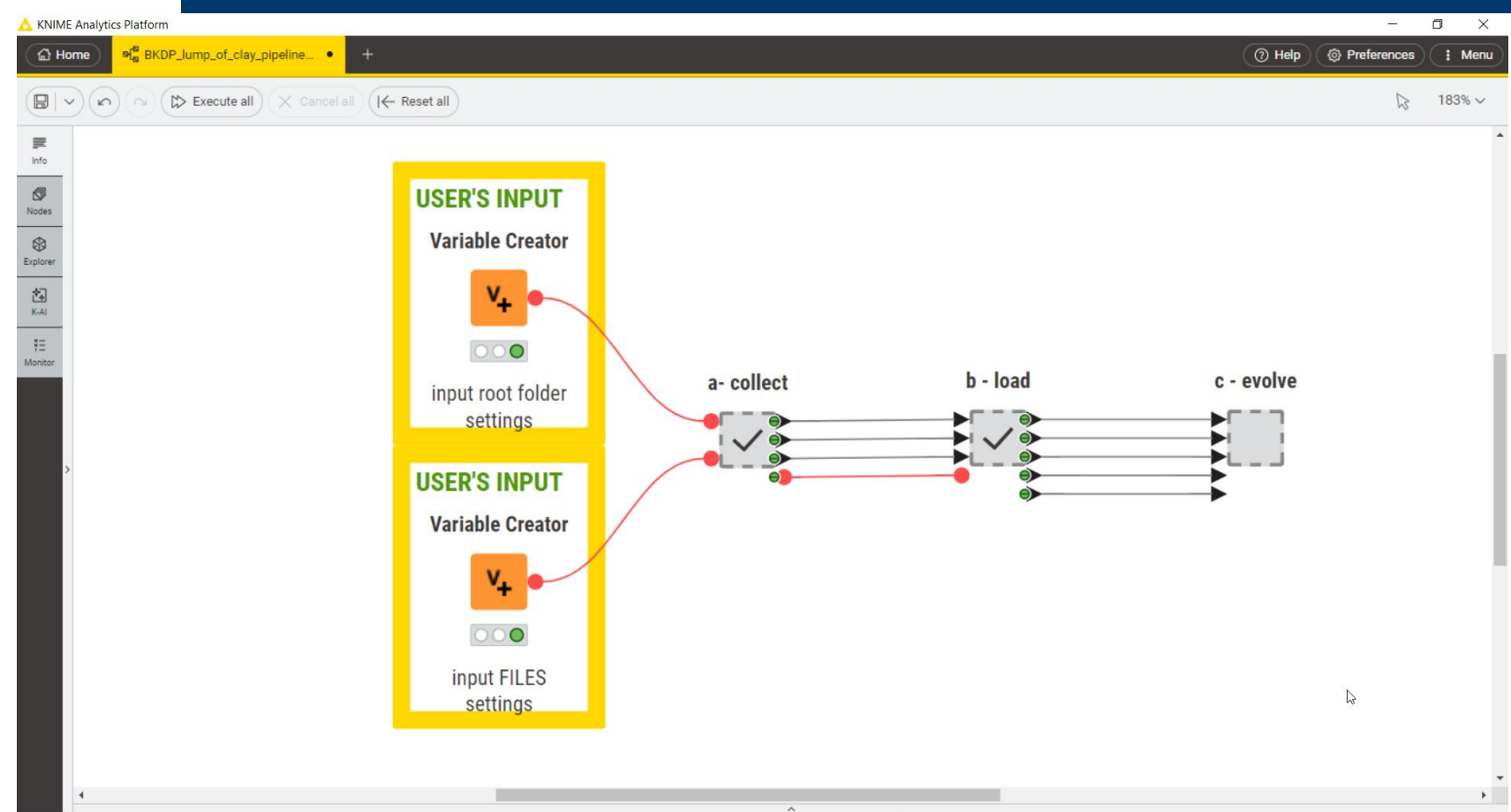
# Practical problem process



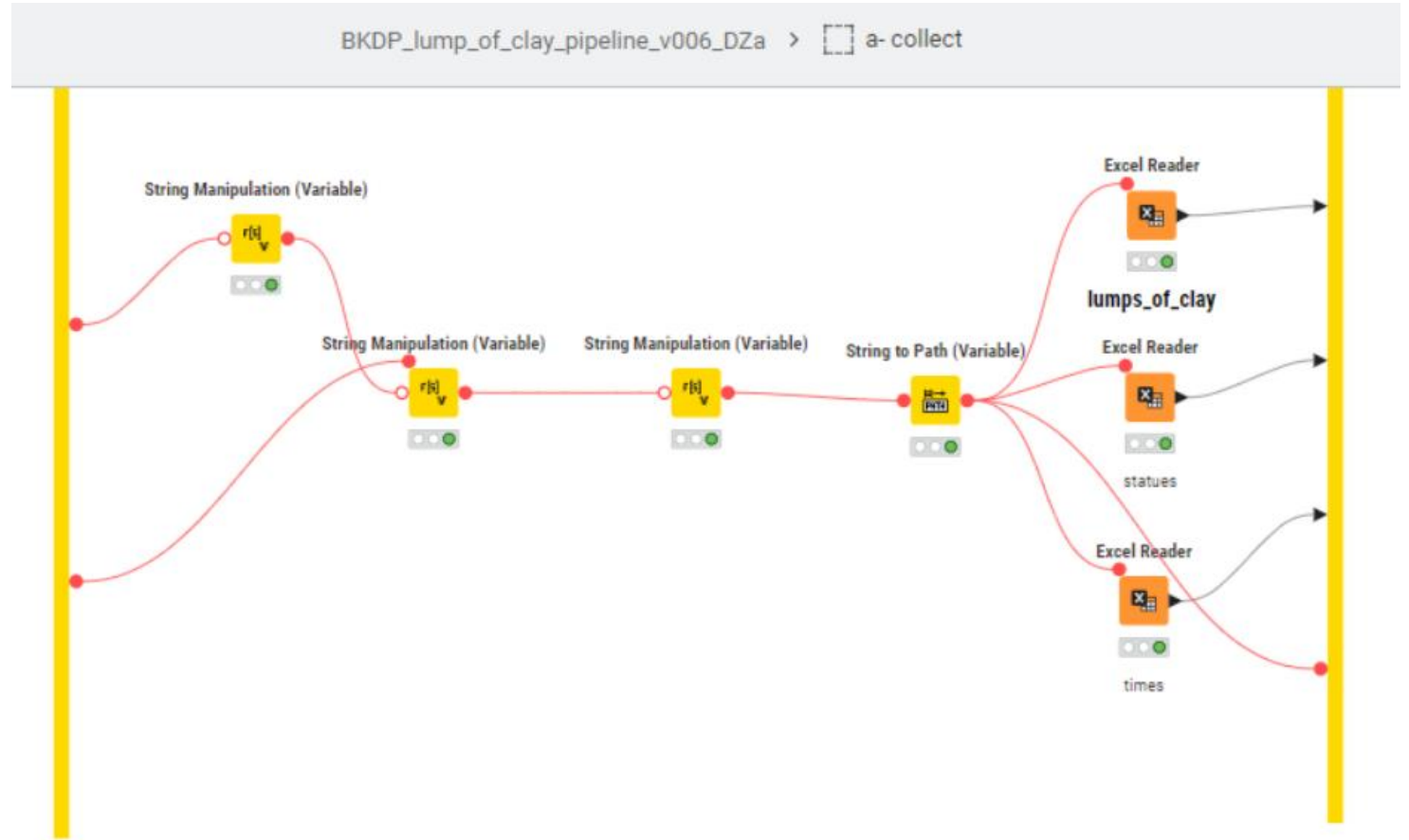
- ⑥ To attempt to automatically (using a machine/computer) transform the input into the designed output
  - acquire a feel for
    - the constraints upon these kinds of machine pipeline transformations
    - the nature of these kinds of machine pipeline transformations
- ⑥ To attempt to build reusable micro transformations
  - acquire a feel for the reuse economics of transformation patterns

- The 'BORO eXcel Table (Manual) Pipeline' should provide a sufficiently detailed specification of most of the micro transformations
  - the KNIME is an implementation of the specification
  - Developing the code to execute the transformation automatically should clarify the formal moves that need to be made

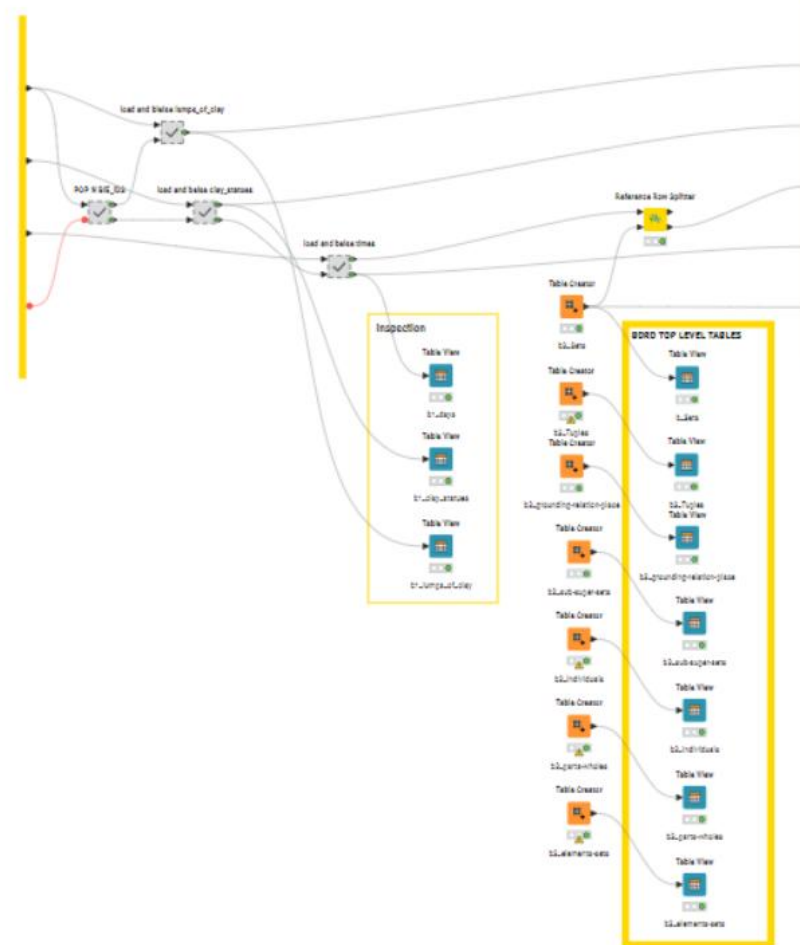
# Knime top level – bCLEARer stages



# Knime COLLECT – done



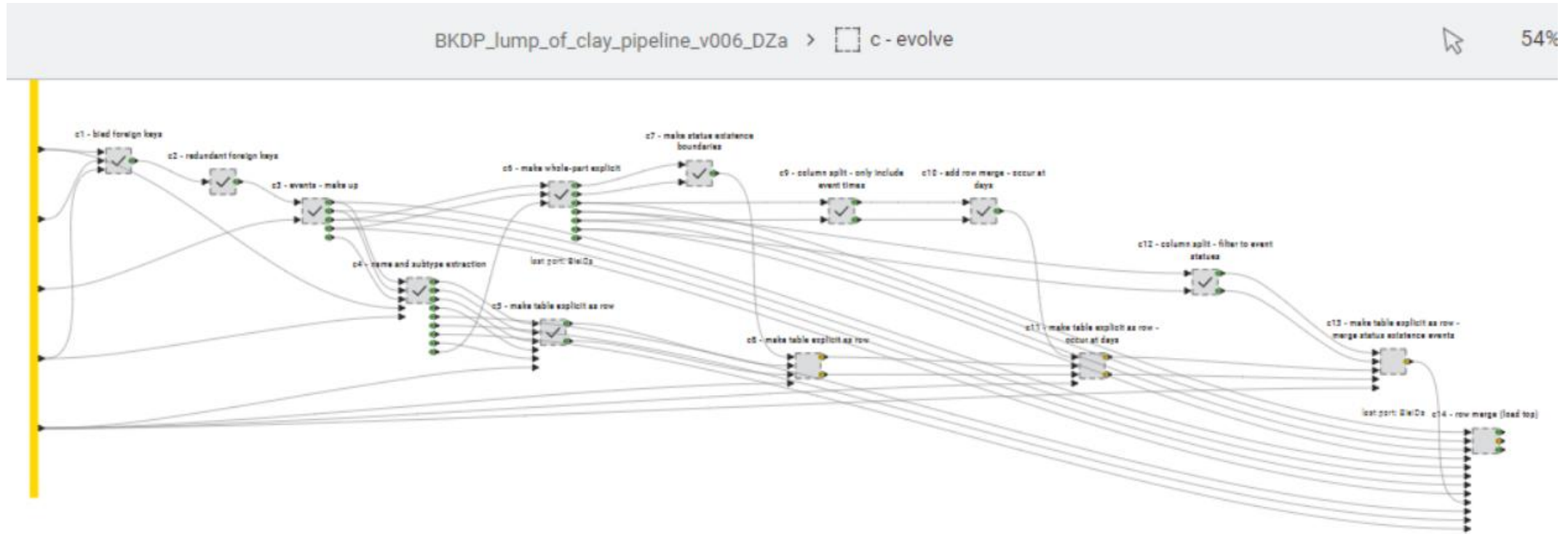
## Knime LOAD – done

BKDP\_lump\_of\_clay\_pipeline\_v006\_DZa > ☐ b - load





# Knime Evolve – work in here





## **Project planning**

---

# Suggested project planning

- ④ Remember everything is on GitHub
- ④ Suggest you plan your 'project'
  - probably better to do a little of each task,
  - rather than just complete the first few
- ④ Task order – earlier tasks feed into later tasks
  1. STM
  2. OED
  3. BUML
  4. BXTP
  5. BKnDP
- ④ Try to do at least 1, 2 and 4
  - if you can, try a little 5 (block out the last half hour to try)
- ④ Remember task 5 has starter code



# Questions







THE END.