

Infection Transmission Ontology: Standardization of Infection Transmission Data

Elena Slavco, *Faculty of Veterinary Medicine, UU*

Martine de Vos, *Research and Data Management Services, UU*

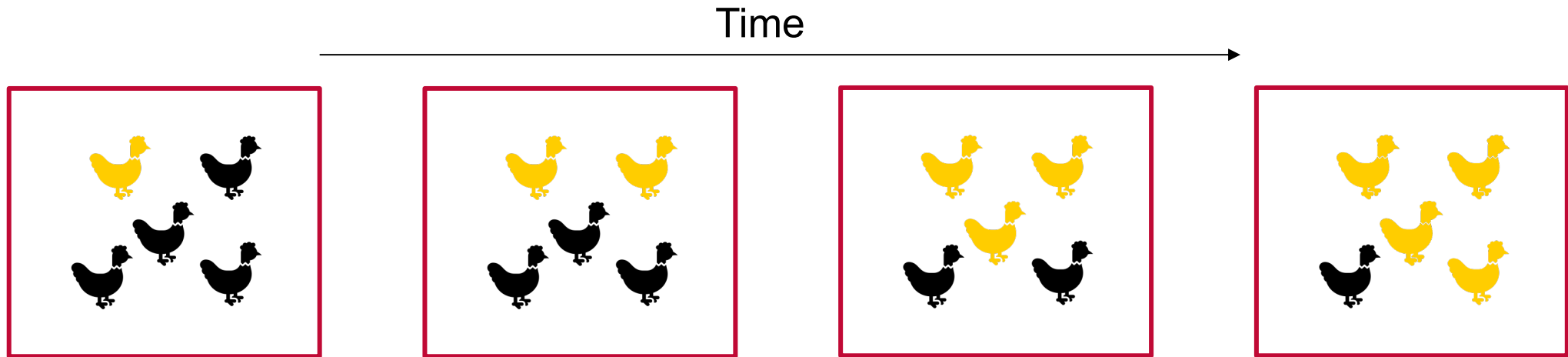
Miel Hostens, *Faculty of Veterinary Medicine, UU*

Jan Top, *Wageningen Food and Biobased Research, WUR*

Egil Fischer, *Faculty of Veterinary Medicine, UU*

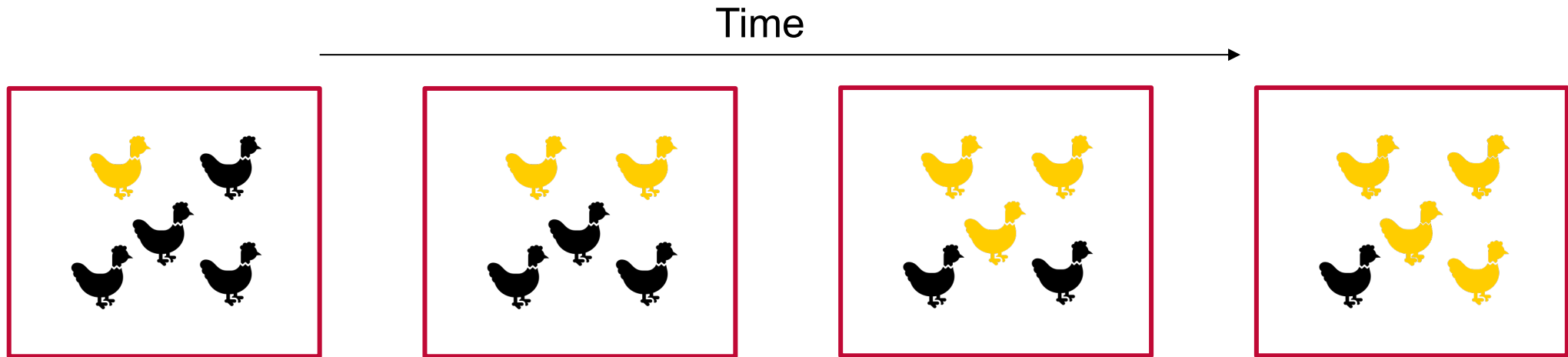
Goal of SUMMERFAIR project

- **Re-using and combining in meta-analysis existing datasets on antimicrobial transmission experiments**



Goal of SUMMERFAIR project

- **Re-using and combining in meta-analysis existing datasets on antimicrobial transmission experiments**



- Transmission experiments are expensive
- No domain specific metadata standards available
- Reusing datasets can help to understand and mitigate infectious diseases

Challenges of SUMMERFAIR project

Dataset A

Isolator	animal	S_I	d1_count_ESBL
1	103	I	100

Dataset B

House	Pen	animalnr_col	I, S1, S2	swab5	swab5_pathogen
H1	I9	21Ro	S1	+	E.Coli

Dataset C

Room	Pen	Pig tag	I/C	1	2H4
H	1	6178	C	N	P

Dataset D

Room	Pen	Pig	Group	D1	D2H0	D2H4
R4	1	6445	NTE2	X	X	P

Challenges of SUMMERFAIR project

Dataset A

Isolator	animal	S_I	d1_count_ESBL
1	103	I	100

Dataset B

House	Pen	animalnr_col	I, S1, S2	swab5	swab5_pathogen
H1	I9	21Ro	S1	+	E.Coli

Dataset C

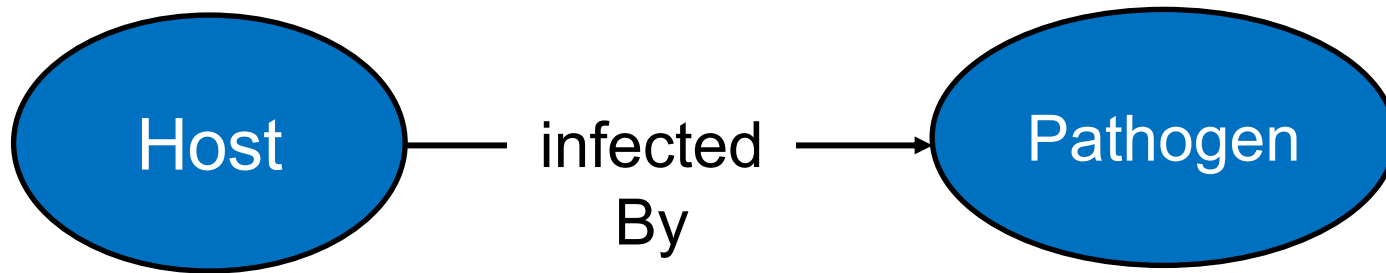
Room	Pen	Pig tag	I/C	1	2H4
H	1	6178	C	N	P

Dataset D

Room	Pen	Pig	Group	D1	D2H0	D2H4
R4	1	6445	NTE2	X	X	P

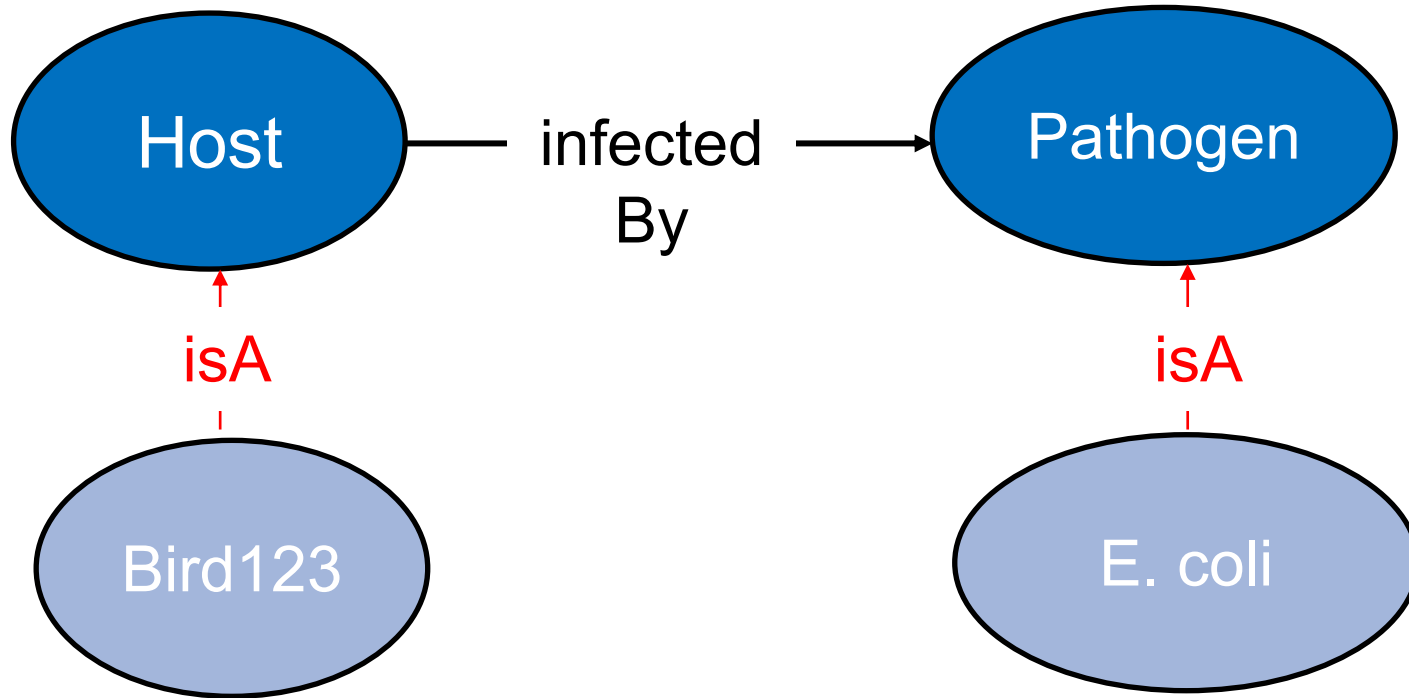
Ontology

a representation scheme describing a formal conceptualization of a domain of interest



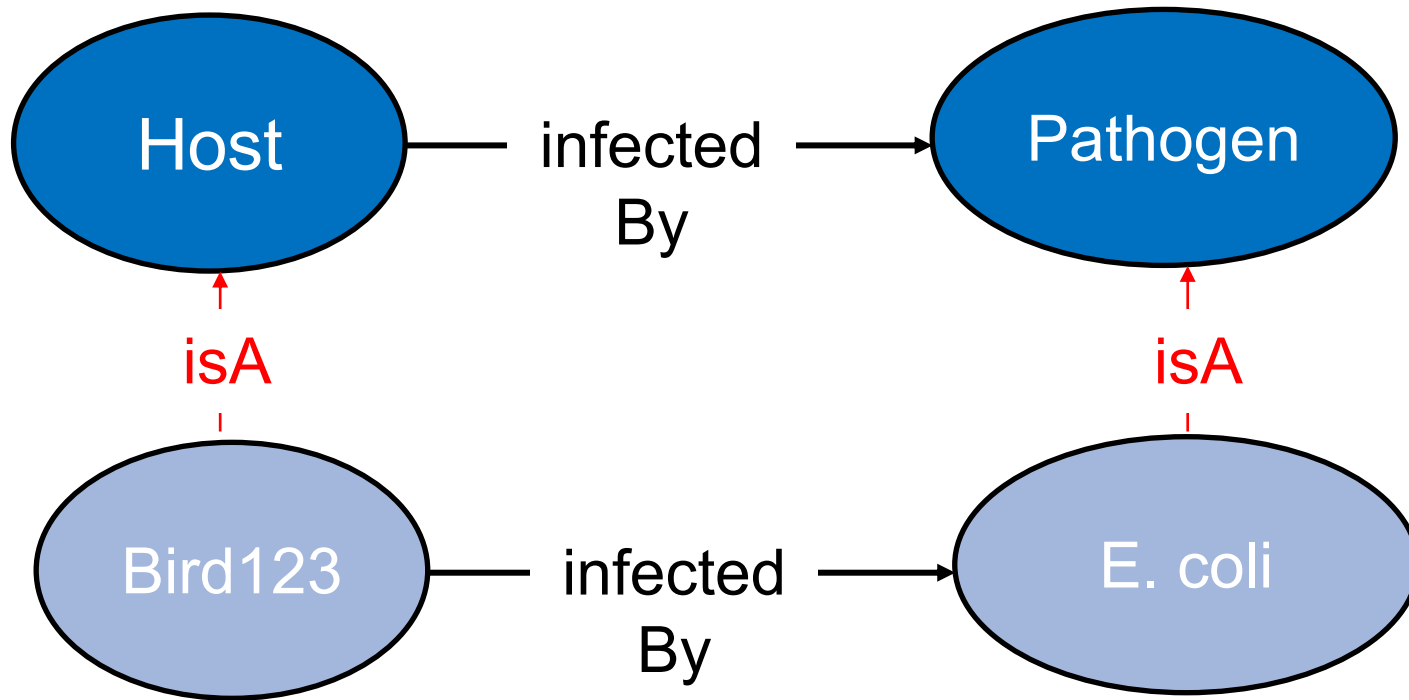
Ontology

a representation scheme describing a formal conceptualization of a domain of interest



Ontology

a representation scheme describing a formal conceptualization of a domain of interest



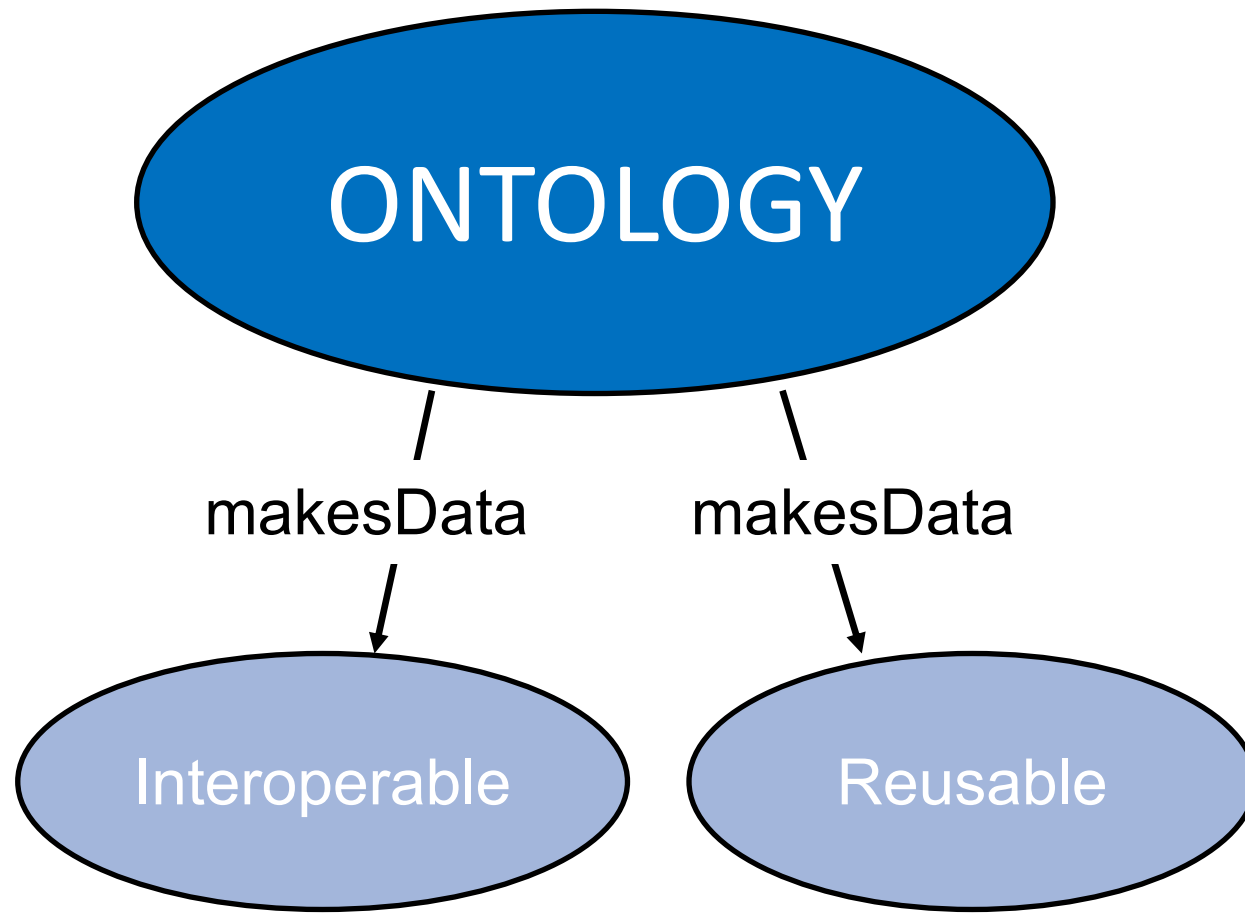
Linked data triples:

Bird123 a Host.

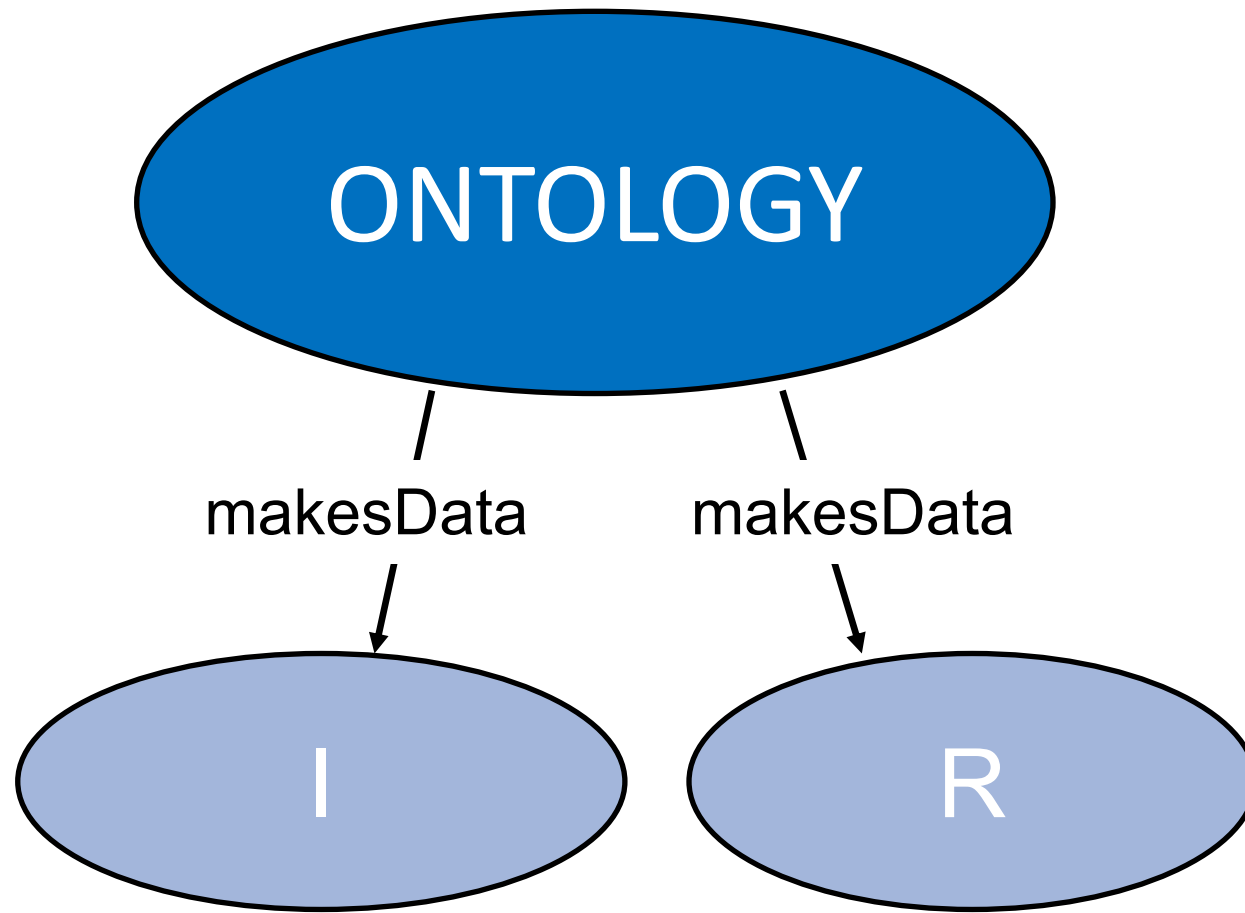
E.Coli a Pathogen.

Bird123 infectedBy E.Coli.

Ontology



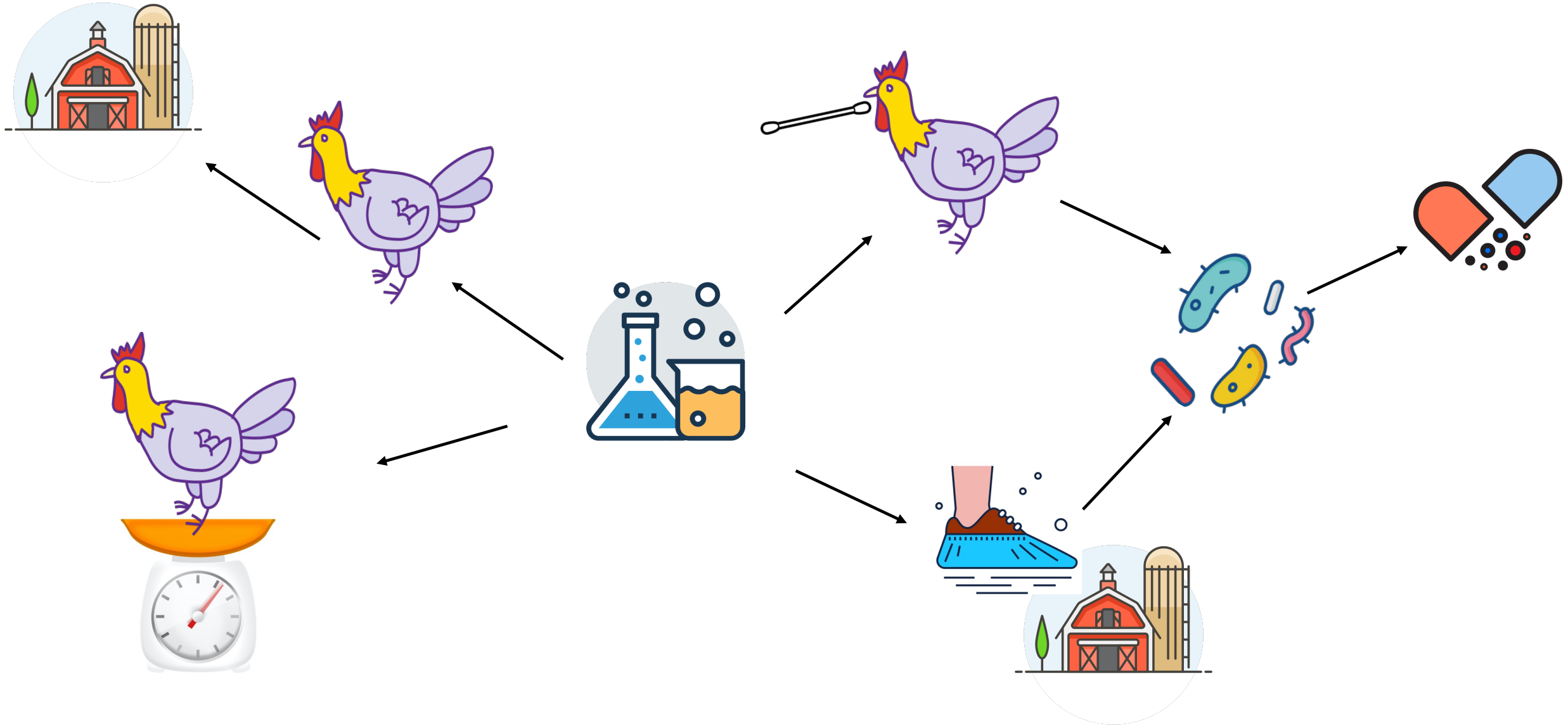
Ontology



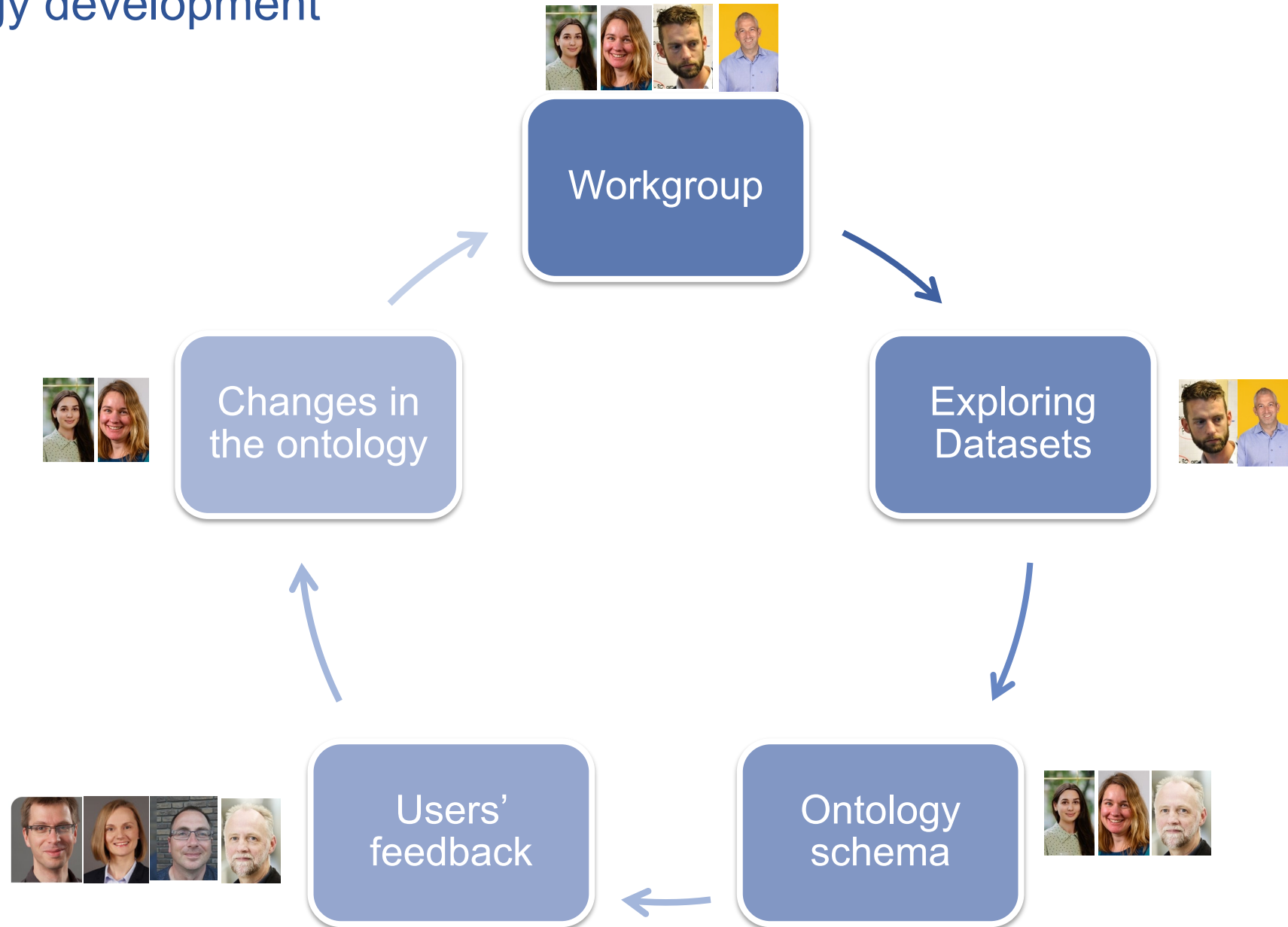
F

A

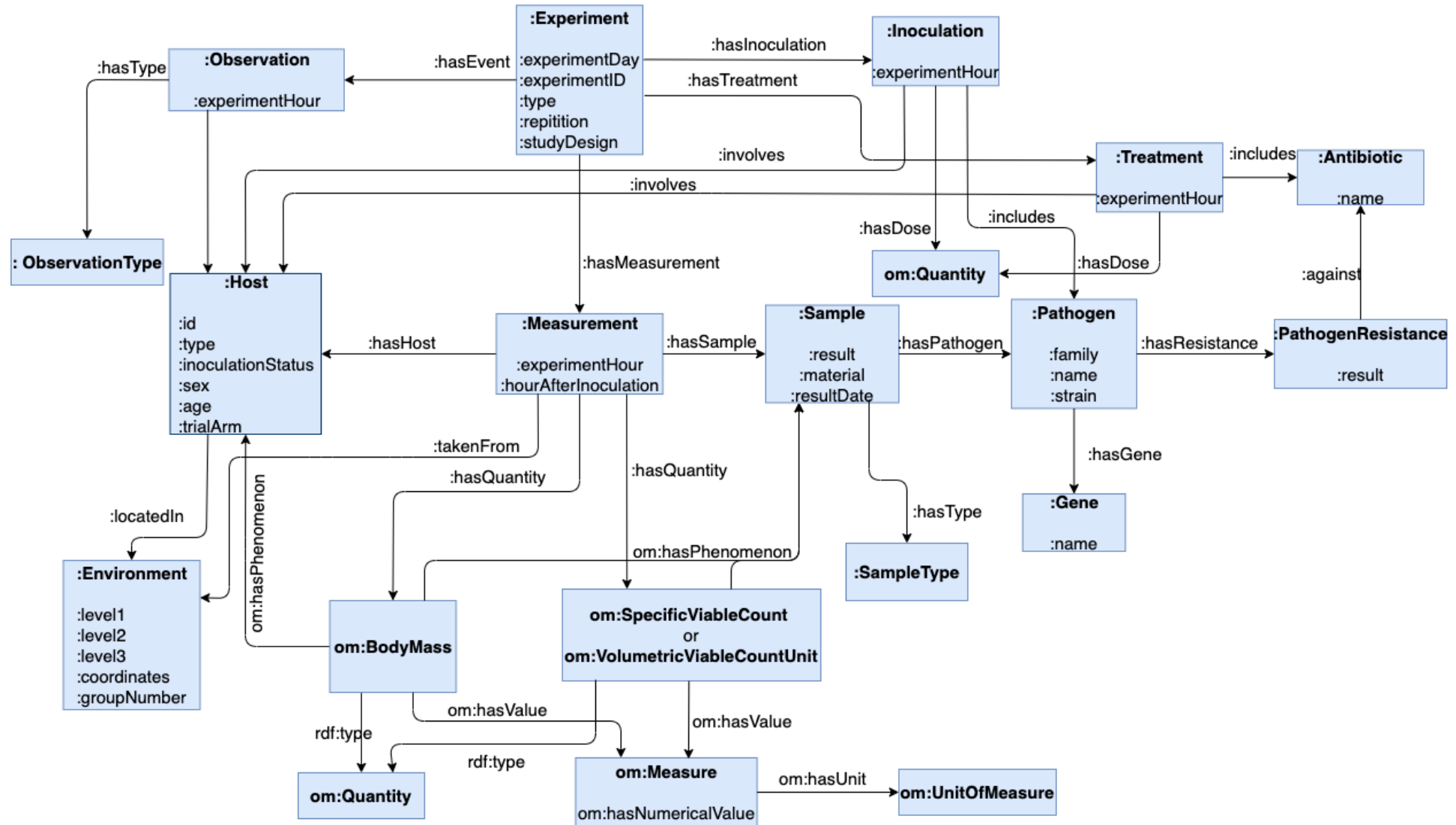
Ontology domain



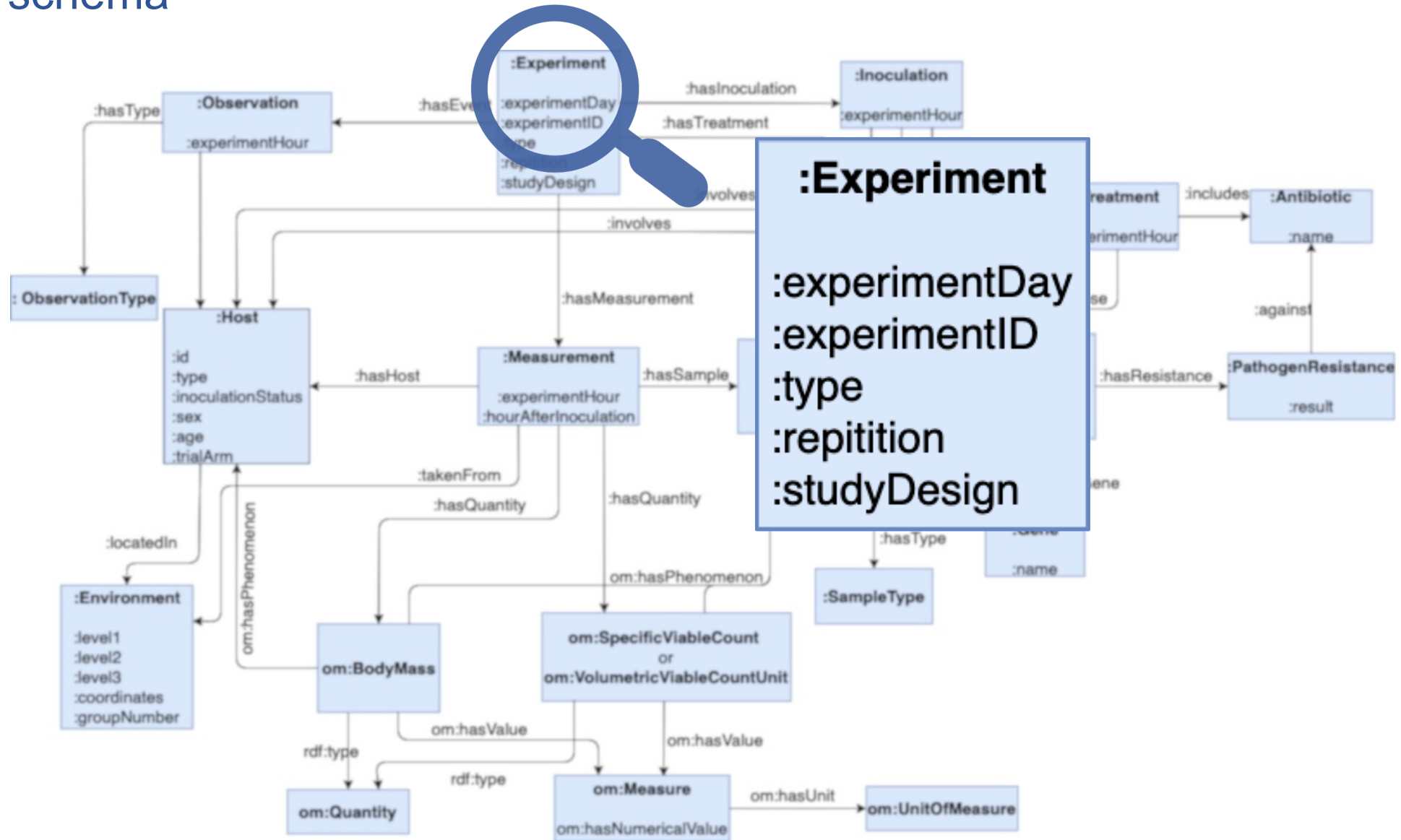
Ontology development



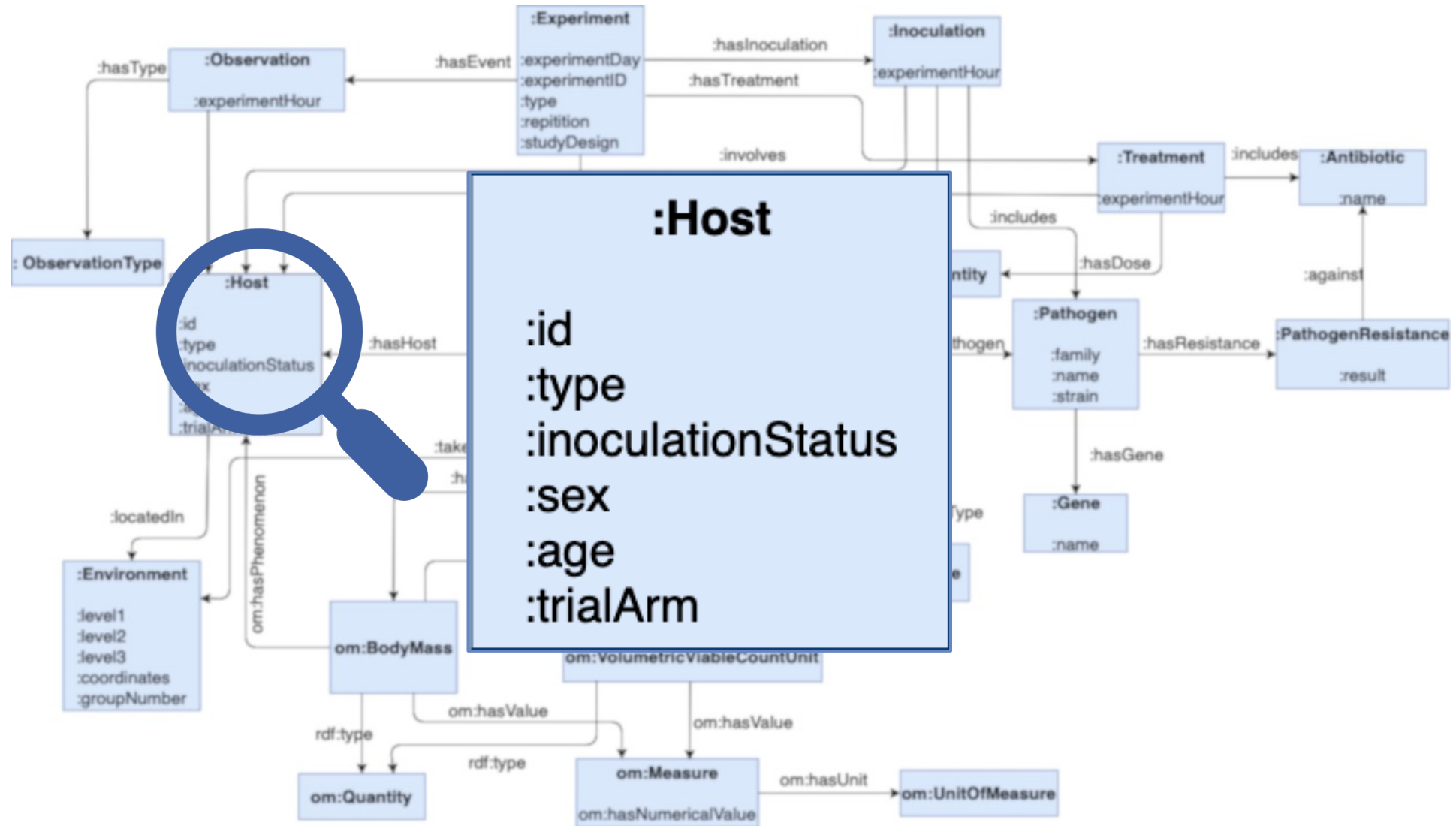
Ontology schema



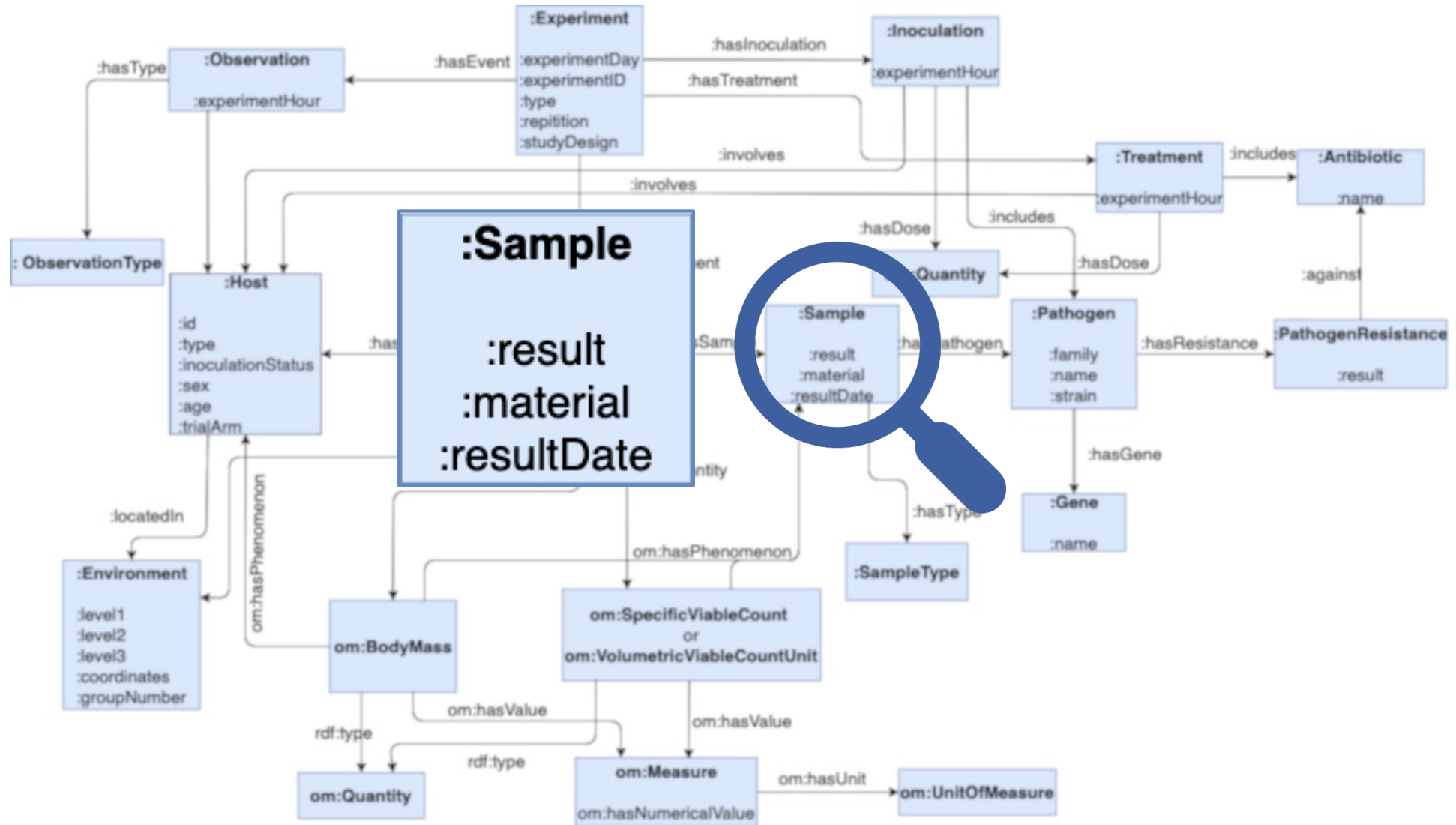
Ontology schema



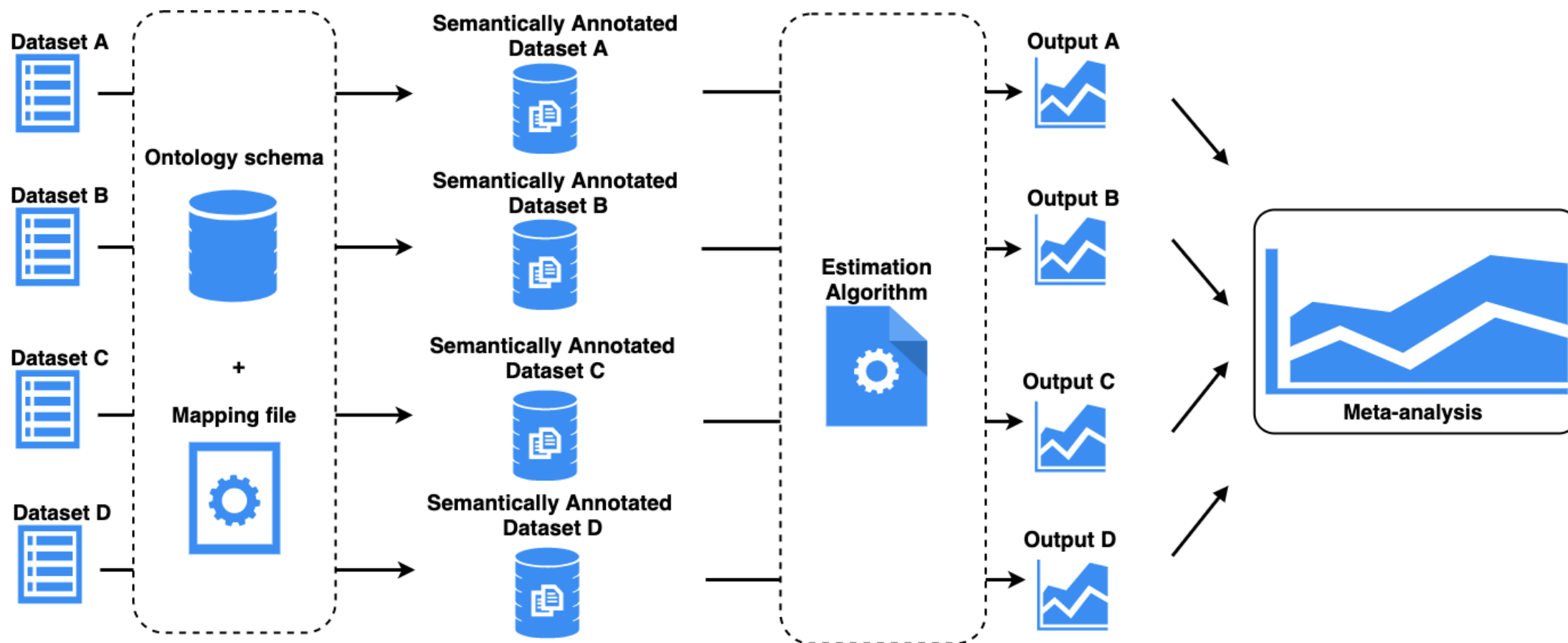
Ontology schema



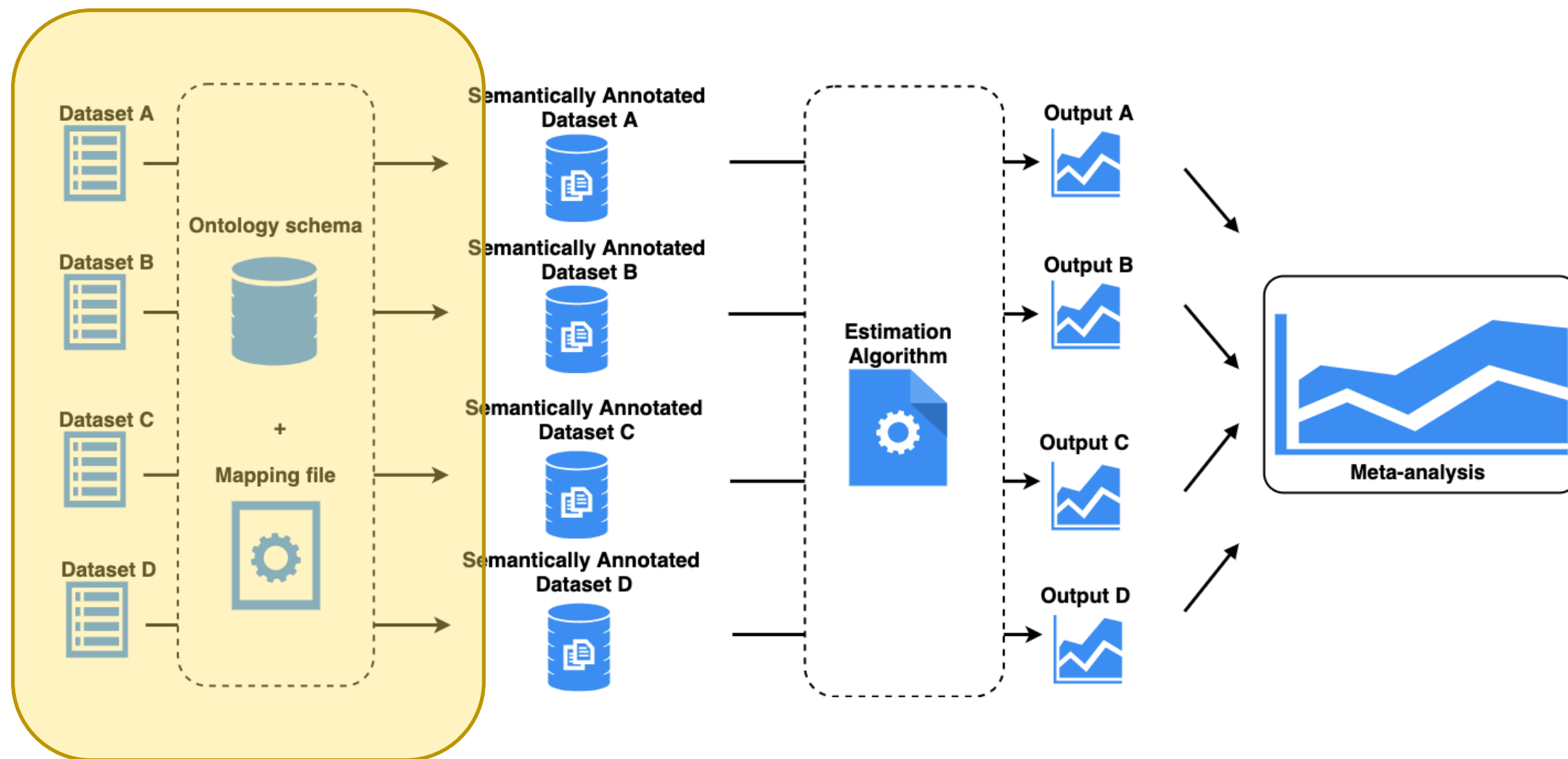
Ontology schema



Workflow



Workflow



Ontology mapping file

ontology_schema:

Experiment:

experimentID: 'DatasetA'
experimentDay:

Environment:

level1:

Host:

id:
locatedIn: Environment

Measurement:

Sample:

result:
hasType: Swab

Pathogen:

name:

Pen	Animal_id	Swab_1_day	Swab_1	Swab_1_value
S1	3_Ge	1	+	E.coli

Ontology mapping file

ontology_schema:

Experiment:

experimentID: 'DatasetA'
experimentDay: Swab_1_day

Environment:

level1: Pen

Host:

id: Animal_id
locatedIn: Environment

Measurement:

Sample:

result: Swab_1
hasType: Swab

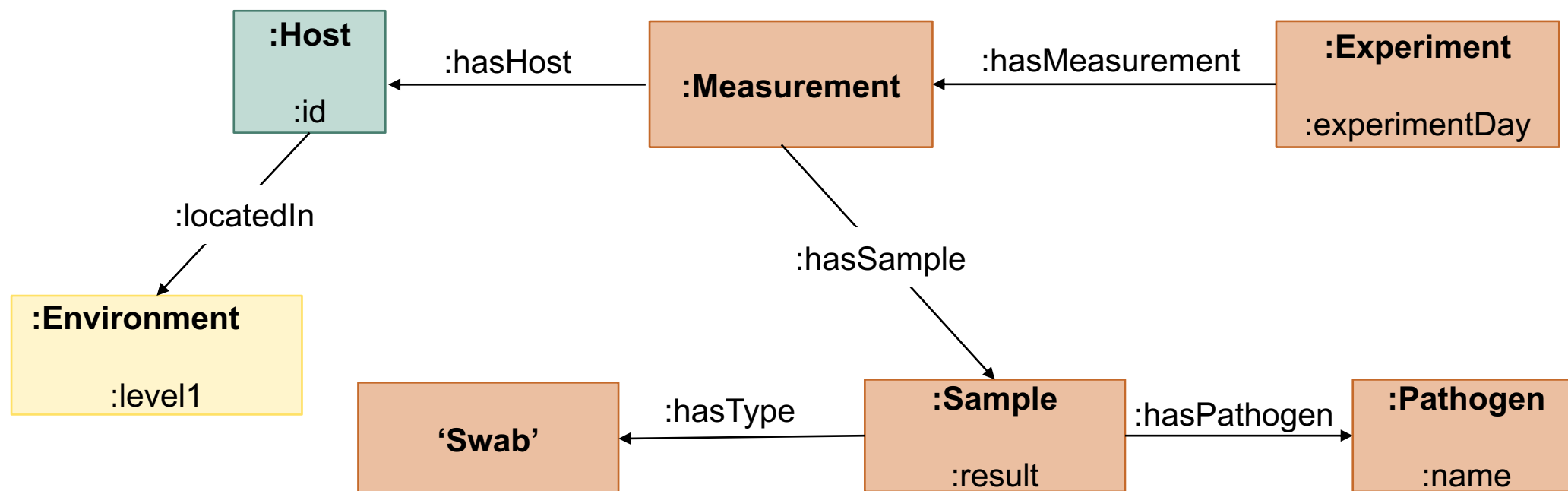
Pathogen:

name: Swab_1_value

Pen	Animal_id	Swab_1_day	Swab_1	Swab_1_value
S1	3_Ge	1	+	E.coli

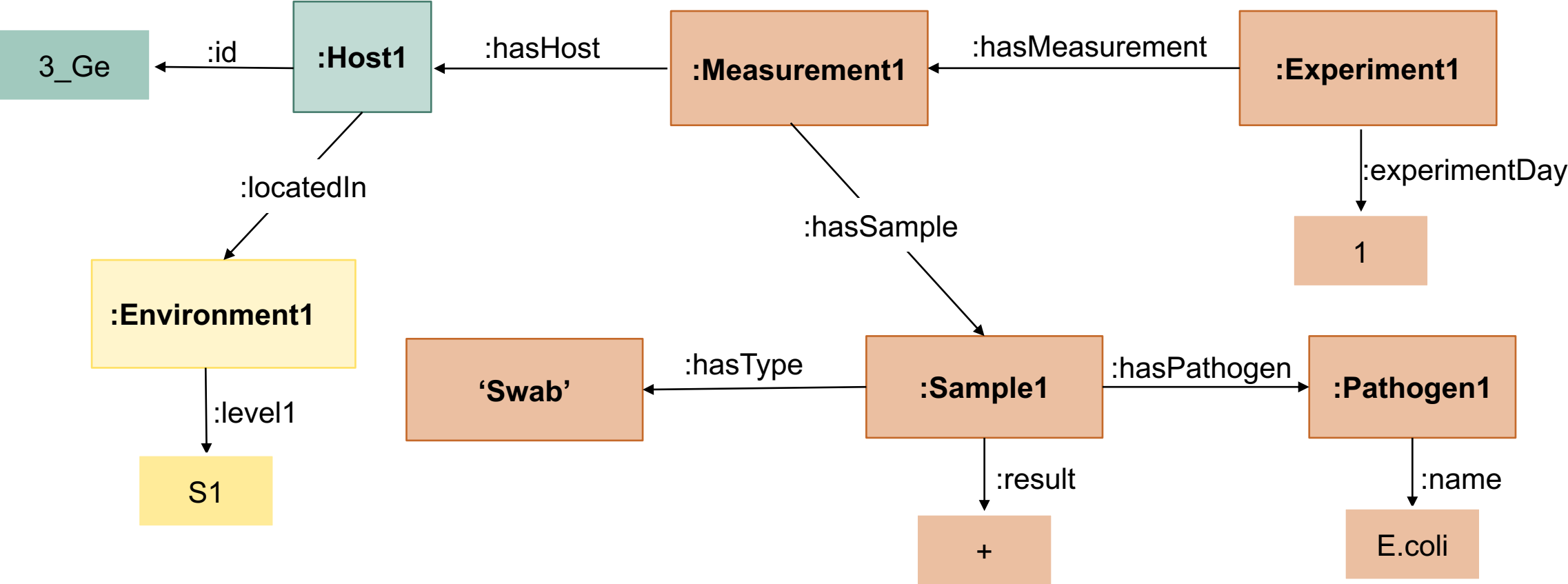
Ontology mapping

Pen	Animal_id	Swab_1_day	Swab_1	Swab_1_value
S1	3_Ge	1	+	E.coli

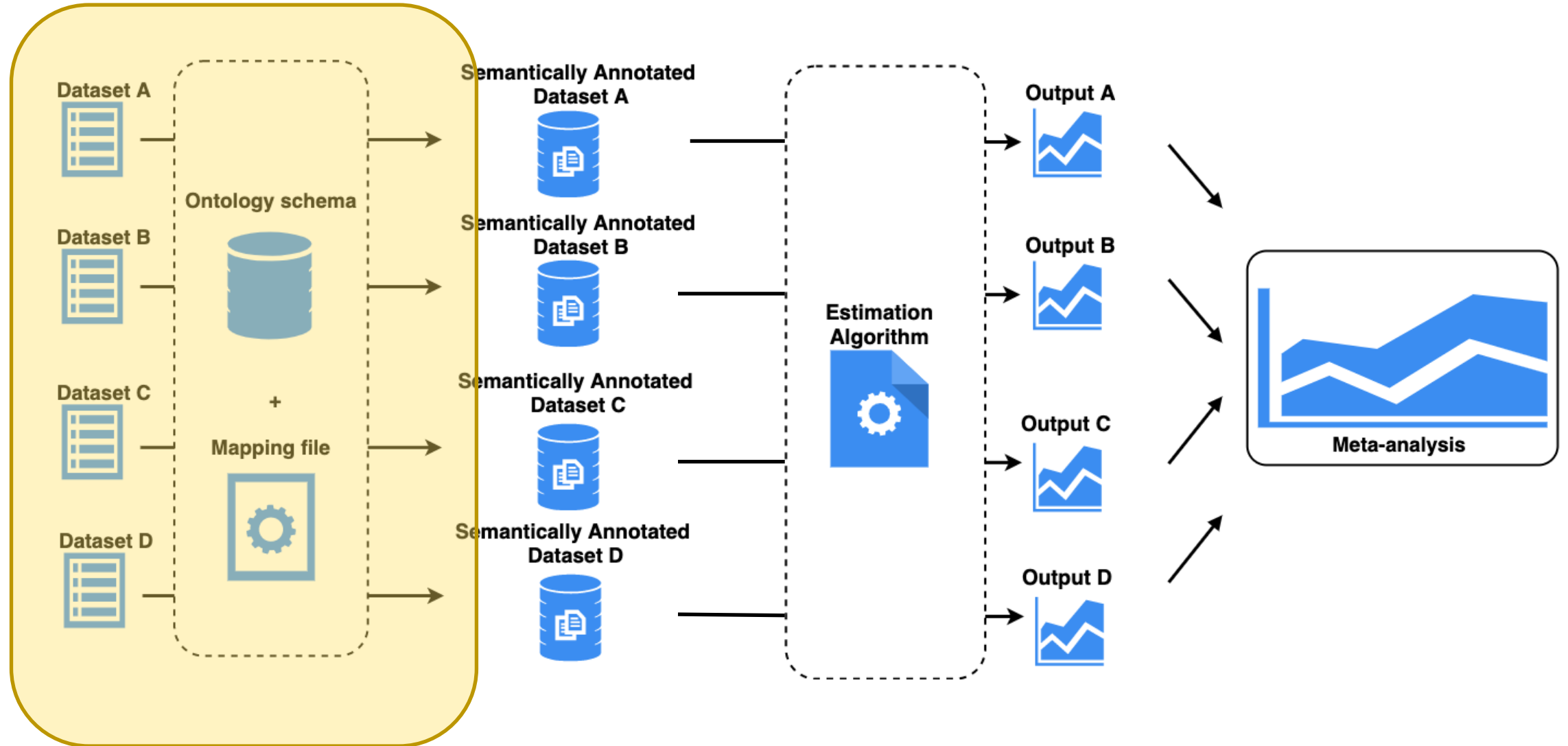


Ontology mapping

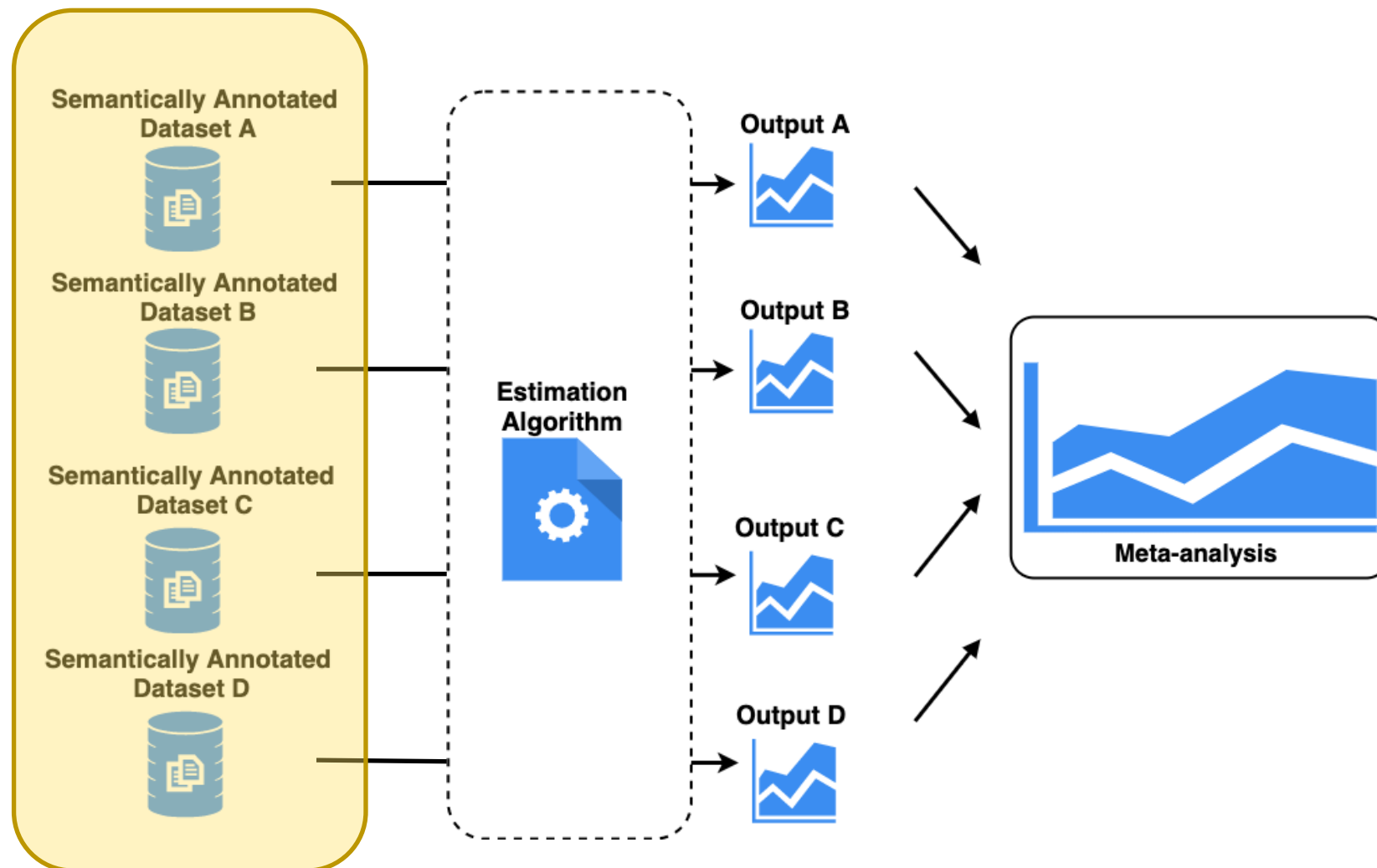
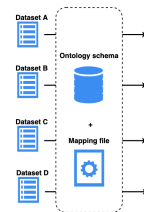
Pen	Animal_id	Swab_1_day	Swab_1	Swab_1_value
S1	3_Ge	1	+	E.coli



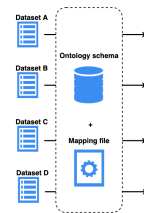
Creating populated ontology



Ontology



Triplestore



**Semantically Annotated
Dataset A**



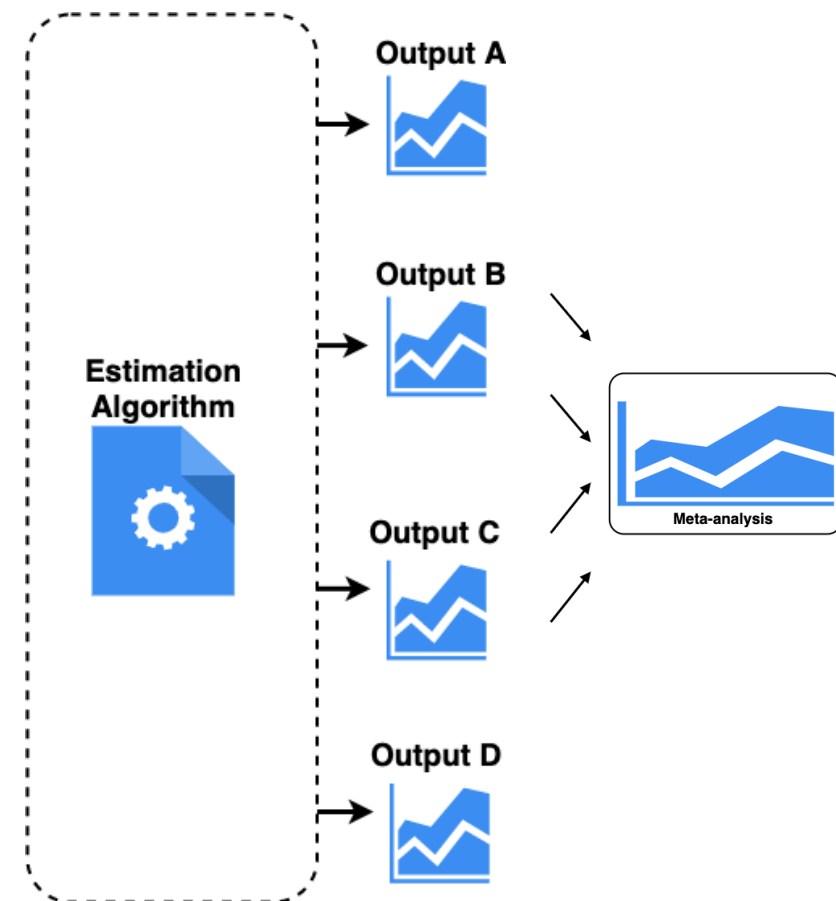
**Semantically Annotated
Dataset B**



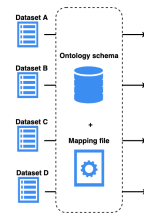
**Semantically Annotated
Dataset C**



**Semantically Annotated
Dataset D**

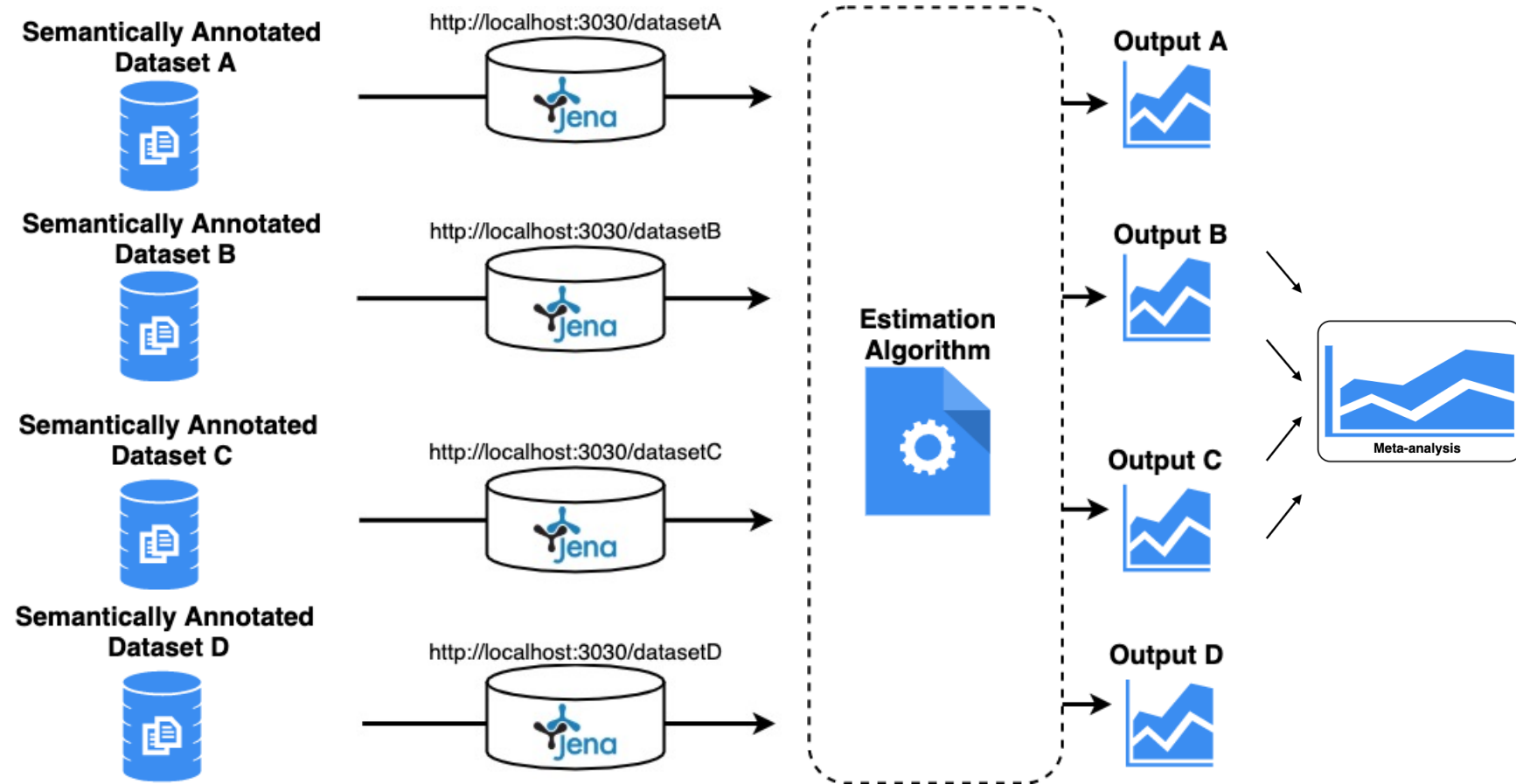


Triplestore

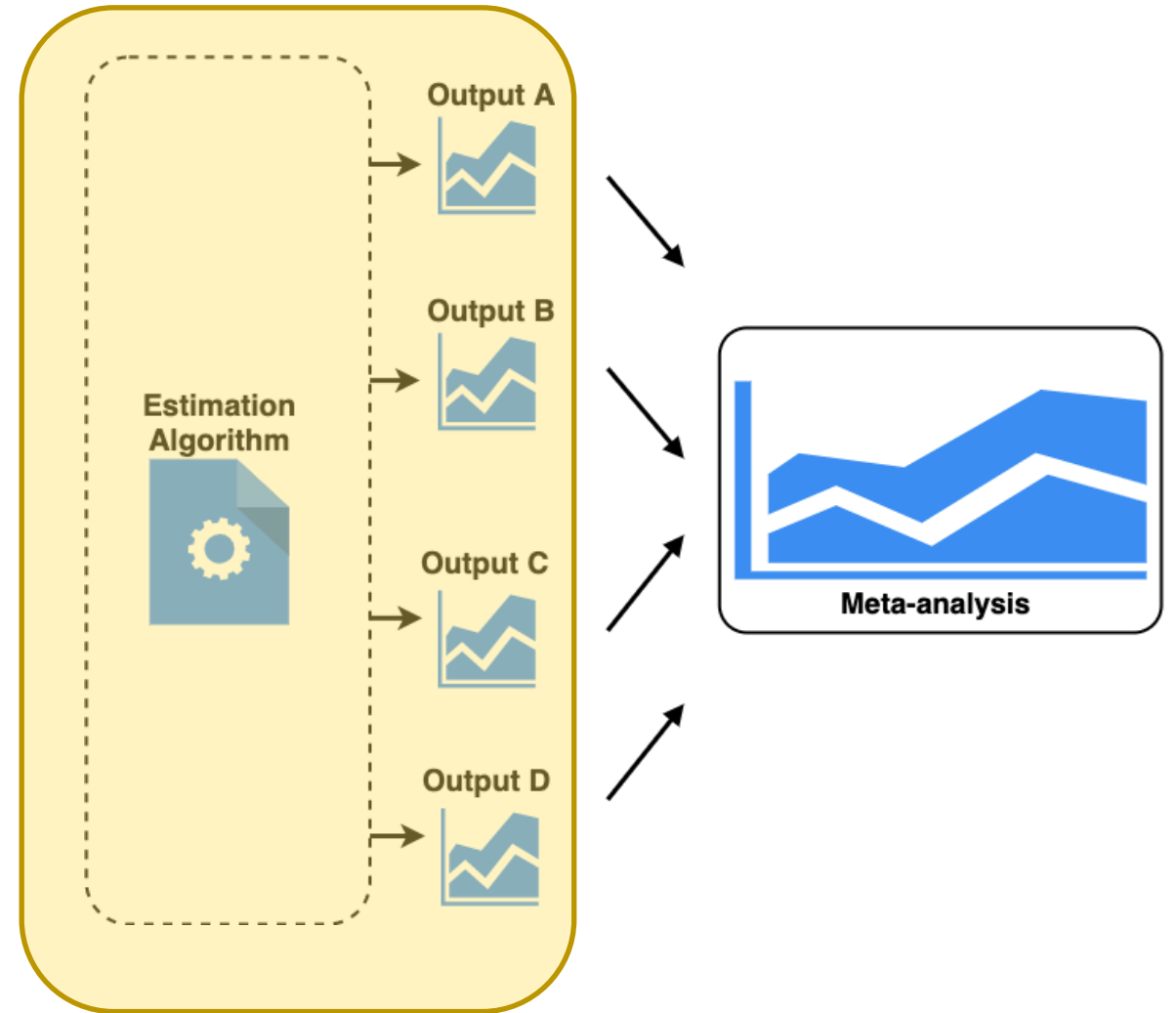
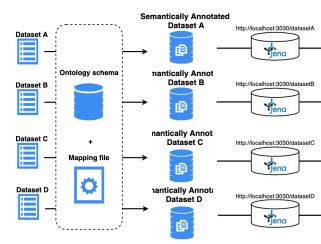


- Graph database that stores semantic facts (triples)
- to store and query data represented in triples
- Each dataset has its endpoint

For our project we are using **Fuseki Jena** that can be run through Docker



Estimation algorithm



Analysis of local dataset

Interpretation

Table 3. Detection (+/-) and quantification (log₁₀ cfu/g feces) of CTX-M-1-*Escherichia coli* in broilers in experiments 2 (dose 10² cfu/mL) and 3, determined at n hours p.i. at days 5 to 7 (8:00 am and 4:00 pm), 8 to 14, 16, 19, and 21 (8:00 am).

Exp	Iso	Dose	Intervention	Bird ID	Seeder or contact	D5 8:00	D5 4:00	D6 8:00	D6 4:00	D7 8:00	D7 4:00	D8 8:00	D9 8:00	D10 8:00	D11 8:00	D12 8:00	D13 8:00	D14 8:00	D16 8:00	D19 8:00	D21 8:00
hours p.i.						0	8	24	32	48	56	72	96	120	144	168	192	216	264	336	384
2	5	10 ²	None	205	Seeder	-	+	3.63	+	5.48	+	5.78	3.57	6.52	6.63	7.57	5.44	6.40	3.88	5.57	5.25
	5	10 ²	None	212	Seeder	-	-	2.57	+	4.63	+	5.57	4.52	4.48	4.63	2.57	4.57	5.36	4.52	4.63	4.70
	5	10 ²	None	275	Seeder	-	-	3.52	+	4.63	+	2.57	5.78	3.63	3.52	4.57	4.30	5.05	3.52	6.78	4.30
	5	10 ²	None	289	Seeder	-	-	2.63	+	4.52	+	3.48	3.44	3.78	2.48	3.00	2.44	2.52	4.70	4.78	6.36
	5	10 ²	None	296	Seeder	+	-	-	+	3.63	+	3.63	5.40	6.52	3.57	4.57	5.52	3.48	6.63	4.70	5.22
	5	10 ²	None	226	Contact	-	-	-	+	2.48	+	3.57	5.88	5.70	4.63	7.88	4.36	4.27	4.57	4.48	5.12
	5	10 ²	None	233	Contact	-	-	-	-	2.33	+	2.52	2.40	2.78	2.52	5.52	3.40	4.70	4.63	6.57	6.13
	5	10 ²	None	240	Contact	-	-	-	+	2.70	+	3.52	2.70	2.48	2.70	3.27	2.63	2.57	3.57	4.57	4.27
	5	10 ²	None	261	Contact	-	-	2.63	-	2.44	+	2.33	2.48	3.52	3.44	2.27	2.63	3.44	6.57	5.44	4.20
	5	10 ²	None	282	Contact	-	+	2.63	+	4.30	-	3.63	2.44	6.48	4.70	4.78	5.10	3.63	4.78	4.44	7.52
	6	10 ²	CEP	220	Seeder	-	-	-	-	2.88	2.48	2.88	2.78	3.70	5.57	4.78	2.52	4.63	4.57	3.63	5.40
	6	10 ²	CEP	234	Seeder	-	-	2.27	+	2.40	+	3.36	4.40	5.52	3.78	3.30	3.40	5.52	4.48	5.40	5.40
	6	10 ²	CEP	241	Seeder	-	-	-	-	2.44	-	2.48	2.40	3.70	4.63	5.63	4.48	2.57	2.63	2.52	3.40
	6	10 ²	CEP	269	Seeder	-	-	-	-	-	-	-	-	4.52	3.63	2.70	3.44	2.48	2.57	2.48	2.48
	6	10 ²	CEP	290	Seeder	-	-	-	-	-	-	-	-	4.8	3.52	2.52	2.63	2.48	3.63	3.33	3.33
	6	10 ²	CEP	206	Contact	-	-	-	-	-	-	-	-	3.00	5.7	2.52	-	-	-	-	-
	6	10 ²	CEP	227	Contact	-	-	-	-	-	-	-	-	3.88	3.63	3.48	4.52	3.57	4.52	4.27	4.27
	6	10 ²	CEP	248	Contact	-	-	-	-	-	-	-	-	3.70	5.63	4.78	6.57	3.57	4.48	5.70	5.70
	6	10 ²	CEP	255	Contact	-	-	-	-	-	-	-	-	3.63	2.57	3.95	3.22	3.57	3.48	3.18	3.18
	6	10 ²	CEP	276	Contact	-	+	-	-	-	-	-	-	3.52	5.7	2.70	3.78	5.52	5.48	5.57	5.57
	7	10 ²	CEP	207	Seeder	-	-	-	-	-	-	-	-	3.53	3.70	3.63	5.88	4.78	4.57	4.33	4.33
	7	10 ²	CEP	242	Seeder	-	+	-	-	-	-	-	-	4.00	2.78	3.78	4.63	3.63	5.00	4.36	3.44
	7	10 ²	CEP	277	Seeder	-	+	-	-	-	-	-	-	-	2.78	2.70	2.57	3.63	3.57	3.48	4.30
	7	10 ²	CEP	291	Seeder	-	-	-	-	-	-	-	-	-	2.57	2.78	2.40	2.40	3.63	3.63	4.33
	7	10 ²	CEP	298	Seeder	-	-	-	-	-	-	-	-	-	2.63	5.57	4.88	4.63	5.70	3.63	4.44
	7	10 ²	CEP	214	Contact	-	-	-	-	-	-	-	-	3.44	4.63	4.36	3.57	3.48	4.57	4.27	4.27
	7	10 ²	CEP	235	Contact	-	-	-	-	-	-	-	-	2.48	2.78	2.88	2.88	2.48	3.57	4.48	4.44
	7	10 ²	CEP	256	Contact	-	-	-	-	-	-	-	-	2.48	2.63	2.52	3.63	3.52	2.30	4.52	4.44
	7	10 ²	CEP	263	Contact	-	-	-	-	-	-	-	-	2.57	3.52	2.70	4.33	3.30	3.70	2.44	2.40
	7	10 ²	CEP	284	Contact	-	-	-	-	-	-	-	-	2.88	3.48	3.57	5.57	2.70	3.78	3.88	3.25
3	2	10 ²	None	309	Seeder	-	-	-	-	-	-	-	-	4.48	3.36	4.27	4.25	3.57	3.22	5.27	5.27
	2	10 ²	None	333	Seeder	-	-	-	+	2.30	+	2.52	2.52	4.40	4.48	3.36	4.27	4.25	3.57	3.22	5.27
	2	10 ²	None	341	Seeder	-	-	3.70	+	4.22	+	5.48	2.48	5.57	3.33	4.27	4.36	5.57	4.52	5.18	6.05
	2	10 ²	None	349	Seeder	-	-	-	+	2.70	+	3.57	3.48	5.48	4.15	3.25	5.15	3.36	4.48	5.18	5.33
	2	10 ²	None	373	Seeder	-	-	-	-	-	+	2.36	4.52	4.15	5.48	5.50	3.57	5.20	5.52	5.25	5.33
	2	10 ²	None	317	Contact	-	-	-	-	-	+	3.57	4.40	4.70	7.30	1.33	5.52	4.63	4.63	5.12	6.06
	2	10 ²	None	325	Contact	-	-	-	-	2.13	+	2.52	3.44	5.44	6.13	5.36	5.27	5.25	4.57	4.27	7.20
	2	10 ²	None	357	Contact	-	-	-	-	-	+	3.48	4.57	4.57	5.33	6.18	7.20	6.33	4.88	5.36	6.33
	2	10 ²	None	365	Contact	-	-	-	-	-	+	4.40	3.52	4.52	7.40	4.30	6.00	6.22	4.57	4.25	5.84
	2	10 ²	None	381	Contact	-	-	-	-	-	+	2.57	4.52	4.70	4.22	5.18	3.57	5.18	5.18	5.40	5.40
	6	10 ²	SYN	305	Seeder	-	-	-	-	-	-	2.33	2.70	2.40	2.13	5.15	5.33	6.40	3.52	4.36	5.48
	6	10 ²	SYN	329	Seeder	-	-	-	-	-	-	-	-	-	-	-	2.48	2.27	2.57	3.06	5.40
	6	10 ²	SYN	337	Seeder	-	-	-	-	-	-	-	-	-	-	-	-	2.48	-	2.3	4.48
	6	10 ²	SYN	353	Seeder	-	-	-	-	-	-	-	-	-	-	2.01	2.12	2.36	2.52	2.20	4.57
	6	10 ²	SYN	361	Seeder	-	-	-	-	-	+	2.36	-	2.52	-	2.57	2.25	2.13	2.25	4.15	5.36
	6	10 ²	SYN	313	Contact	-	-	-	-	-	-	-	-	-	1.85	2.15	2.40	2.15	2.70	2.15	3.36
	6	10 ²	SYN	321	Contact	-	-	-	-	-	-	-	-	-	-	-	3.00	2.06	4.00	4.13	3.48
	6	10 ²	SYN	369	Contact	-	-	-	-	-	-	-	-	-	-	-	2.36	3.00	3.57	3.13	3.48

Analysis of local dataset

Algorithm uses 'rules'

- predefined
- user writable

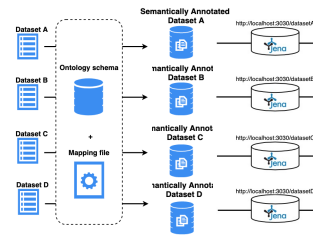
Example rule:

- use cut-off for quantitative values
- animal is positive *since first positive*

Table 3. Detection (+/-) and quantification (log₁₀ cfu/g feces) of CTX-M-1-*Escherichia coli* in broilers in experiments 2 (dose 10² cfu/mL) and 3, determined at n hours p.i. at days 5 to 7 (8:00 am and 4:00 pm), 8 to 14, 16, 19, and 21 (8:00 am).

Exp	Iso	Dose	Intervention	Bird ID	Seeder or contact	D5 8:00	D5 4:00	D6 8:00	D6 4:00	D7 8:00	D7 4:00	D8 8:00	D9 8:00	D10 8:00	D11 8:00	D12 8:00	D13 8:00	D14 8:00	D16 8:00	D19 8:00	D21 8:00		
hours p.i.						0	8	24	32	48	56	72	96	120	144	168	192	216	264	336	384		
2	5	10 ²	None	205	Seeder	-	+	3.63	+	5.48	+	5.78	3.57	6.52	6.63	7.57	5.44	6.40	3.88	5.57	5.25		
	5	10 ²	None	212	Seeder	-	-	2.57	+	4.63	+	5.57	4.52	4.48	4.63	2.57	4.57	5.36	4.52	4.63	4.70		
	5	10 ²	None	275	Seeder	-	-	3.52	+	4.63	+	2.57	5.78	3.63	3.52	4.57	4.30	5.05	3.52	6.78	4.30		
	5	10 ²	None	289	Seeder	-	-	2.63	+	4.52	+	3.48	3.44	3.78	2.48	3.00	2.44	2.52	4.70	4.78	6.36		
	5	10 ²	None	296	Seeder	+	-	-	-	3.63	+	3.63	5.40	6.52	3.57	4.57	5.52	3.48	6.63	4.70	5.22	4.82	
	5	10 ²	None	226	Contact	-	-	-	+	2.48	+	3.57	5.88	5.70	4.63	7.88	4.36	4.27	4.57	4.48	5.12		
	5	10 ²	None	233	Contact	-	-	-	-	2.33	+	2.52	2.40	2.78	2.52	5.52	3.40	4.70	4.63	6.57	6.13		
	5	10 ²	None	240	Contact	-	-	-	+	2.70	+	3.52	2.70	2.48	2.70	3.27	2.63	2.57	3.57	4.57	4.27		
	5	10 ²	None	261	Contact	-	-	2.63	-	2.44	+	2.33	2.48	3.52	3.44	2.27	2.63	3.44	6.57	5.44	4.20		
	5	10 ²	None	282	Contact	-	+	2.63	+	4.30	-	3.63	2.44	6.48	4.70	4.78	5.10	3.63	4.78	4.44	7.52		
	6	10 ²	CEP	220	Seeder	-	-	-	-	-	-	2.88	2.48	2.88	3.70	5.57	4.78	2.52	4.63	4.57	3.63		
	6	10 ²	CEP	234	Seeder	-	-	2.27	+	2.40	+	3.36	4.40	5.52	3.78	3.30	3.40	5.52	4.48	5.40			
	6	10 ²	CEP	241	Seeder	-	-	-	-	2.44	-	2.48	2.40	3.70	4.63	5.63	4.48	2.57	2.63	3.40			
	6	10 ²	CEP	269	Seeder	-	-	-	-	-	-	-	-	2.52	3.63	2.70	3.44	2.48	2.57	2.48			
	6	10 ²	CEP	290	Seeder	-	-	-	-	-	-	-	-	3.52	2.52	2.63	2.48	3.63	3.33				
	6	10 ²	CEP	206	Contact	-	-	-	-	-	-	-	-	3.63	3.63	2.52	2.52	-	-	-			
	6	10 ²	CEP	227	Contact	-	-	-	-	-	-	-	-	3.63	3.63	3.48	4.52	3.57	4.52	4.27			
	6	10 ²	CEP	248	Contact	-	-	-	-	-	-	-	-	3.63	3.63	3.48	4.78	6.57	3.57	4.48	5.70		
	6	10 ²	CEP	255	Contact	-	-	-	-	-	-	-	-	3.63	3.63	2.57	3.95	3.22	3.57	3.48	3.18		
	6	10 ²	CEP	276	Contact	-	+	-	-	-	-	-	-	3.63	3.63	2.57	5.74	2.70	3.78	5.52	5.48	5.57	
	7	10 ²	CEP	207	Seeder	-	-	-	-	-	-	-	-	3.63	3.63	2.70	3.63	5.88	4.78	4.57	4.33		
	7	10 ²	CEP	242	Seeder	-	+	-	-	-	-	-	-	4.00	2.78	3.78	4.63	3.63	5.00	4.36	3.44		
	7	10 ²	CEP	277	Seeder	-	-	-	-	-	-	-	-	2.78	2.70	2.57	3.63	3.57	3.48	4.30			
	7	10 ²	CEP	291	Seeder	-	-	-	-	-	-	-	-	-	2.57	2.78	2.40	2.40	3.63	3.63	4.33		
	7	10 ²	CEP	298	Seeder	-	-	-	-	-	-	-	-	-	2.63	5.57	4.88	4.63	5.70	3.63	4.44		
	7	10 ²	CEP	214	Contact	-	-	-	-	-	-	-	-	-	3.44	3.63	4.63	4.36	3.57	3.48	4.57	4.27	
	7	10 ²	CEP	235	Contact	-	-	-	-	-	-	-	-	-	2.48	2.78	2.88	2.48	3.57	3.52	4.48	4.44	
	7	10 ²	CEP	256	Contact	-	-	-	-	-	-	-	-	-	2.48	2.63	2.52	3.63	3.52	2.30	4.52	4.44	5.25
	7	10 ²	CEP	263	Contact	-	-	-	-	-	-	-	-	-	2.57	2.52	2.40	4.33	3.30	3.70	2.44	2.44	
	7	10 ²	CEP	284	Contact	-	-	-	-	-	-	-	-	-	2.63	3.48	3.57	5.57	2.70	3.78	3.88	3.25	
3	2	10 ²	None	309	Seeder	-	-	-	+	2.30	+	2.52	2.52	4.40	4.48	3.36	4.27	4.25	3.57	3.22	5.27		
	2	10 ²	None	333	Seeder	-	-	-	+	3.25	+	4.57	4.48	3.57	5.44	5.25	5.40	7.25	5.48	6.03	5.06		
	2	10 ²	None	341	Seeder	-	-	3.70	+	4.22	+	5.48	2.48	5.57	3.33	4.27	4.36	5.57	4.52	5.18	6.05		
	2	10 ²	None	349	Seeder	-	-	-	+	2.70	+	3.57	3.48	5.48	4.15	6.25	5.15	3.36	4.48	5.18	5.33		
	2	10 ²	None	373	Seeder	-	-	-	-	-	+	2.36	4.52	4.15	5.48	5.30	3.57	5.20	5.52	5.25	5.33		
	2	10 ²	None	317	Contact	-	-	-	-	-	+	3.57	4.40	4.70	7.30	1.33	5.52	4.63	4.63	5.12	6.06		
	2	10 ²	None	325	Contact	-	-	-	-	-	+	2.52	3.44	5.44	6.13	5.36	5.27	5.25	4.57	4.27	7.20		
	2	10 ²	None	357	Contact	-	-	-	-	-	+	3.48	4.57	4.57	5.33	6.18	7.20	6.33	4.88	5.36	6.33		
	2	10 ²	None	365	Contact	-	-	-	-	-	+	4.40	3.52	4.52	7.40	4.30	6.00	6.22	4.57	4.25	5.84		
	2	10 ²	None	381	Contact	-	-	-	-	-	+	2.57	4.52	4.70	4.22	5.18	5.57	5.63	4.44	5.40			
	6	10 ²	SYN	305	Seeder	-	-	-	-	-	-	2.33	2.70	2.40	2.13	5.15	5.33	6.40	3.52	4.36	5.48		
	6	10 ²	SYN	329	Seeder	-	-	-	-	-	-	-	-	-	-	-	-	2.48	2.27	2.57	3.06	5.40	
6	10 ²	SYN	337	Seeder	-	-	-	-	-	-	-	-	-	-	-	-	-	2.48	-	2.43	4.48		
6	10 ²	SYN	353	Seeder	-	-	-	-	-	-	-	-	-	-	-	-	2.01	2.12	2.36	2.52	4.20	4.57	
6	10 ²	SYN	361	Seeder	-	-	-	-	-	-	+	2.36	-	2.52	-	2.57	2.13	2.27	2.48	4.15	5.36		
6	10 ²	SYN	313	Contact	-	-	-	-	-	-	-	-	-	-	-	2.15	2.40	2.15	2.70	2.15	3.36		
6	10 ²	SYN	321	Contact	-	-	-	-	-	-	-	-	2.52	-	-	-	-	3.00	2.06	4.00	2.13	3.48	
6	10 ²	SYN	369	Contact	-	-	-	-	-	-	-	-	-	-	-	-	-	2.36	3.00	3.57	3.13	3.48	

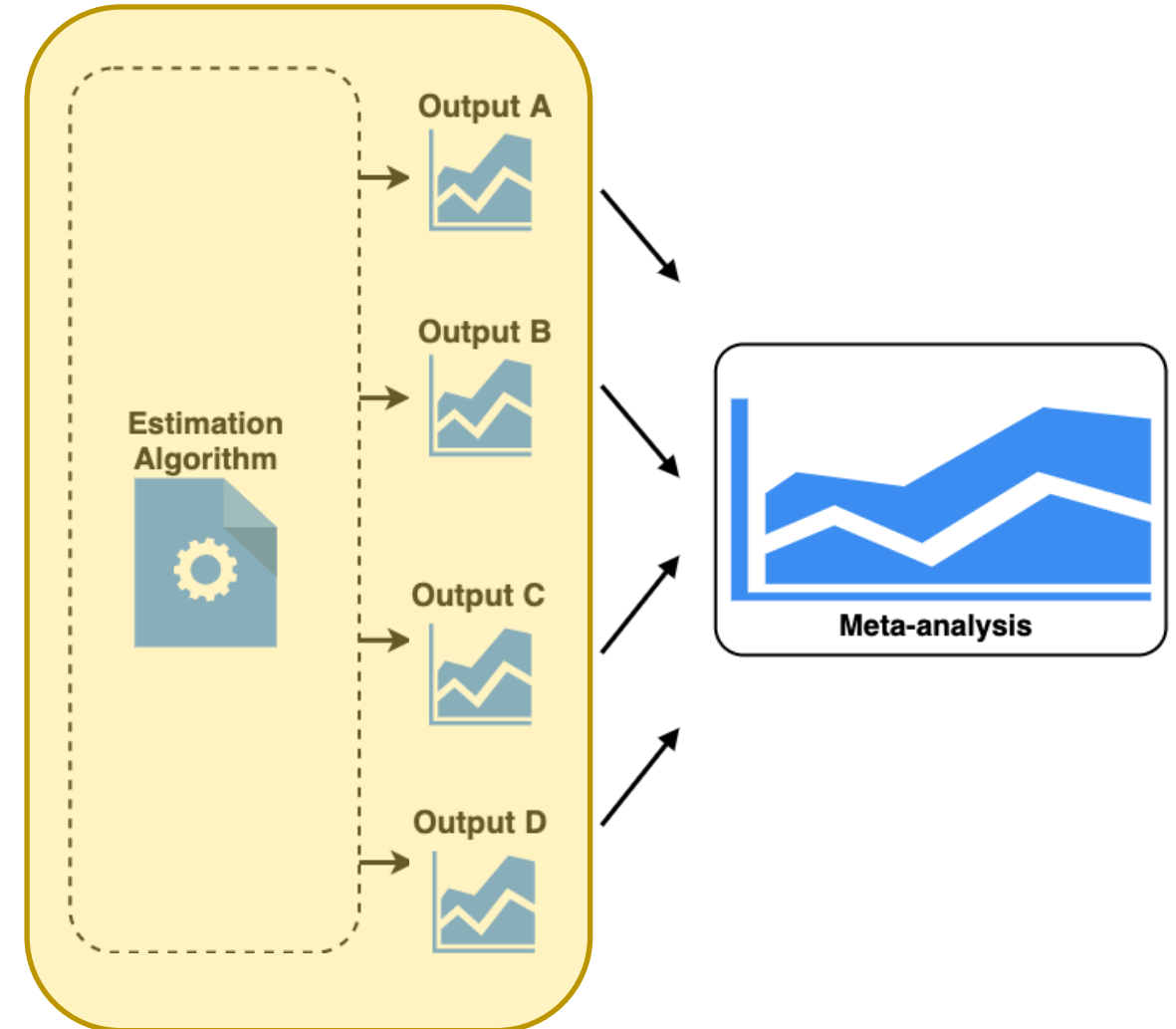
Estimation algorithm



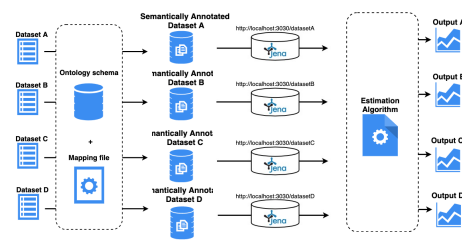
- Query the triplestore to retrieve the data
- Algorithm rules / setting:
 - Quantitative measures \Rightarrow +/- results
 - | | | |
|---|---|---|
| + | - | - |
|---|---|---|

 \Rightarrow

+	+	+
---	---	---
 - Estimation procedure: *GLM*



Meta-analysis



Output A



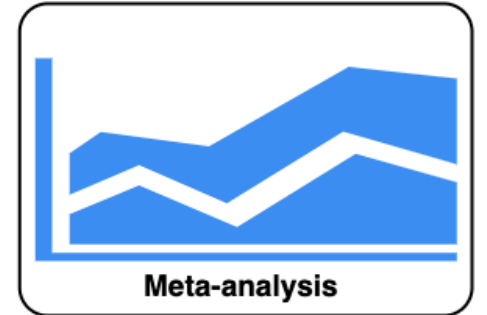
Output B



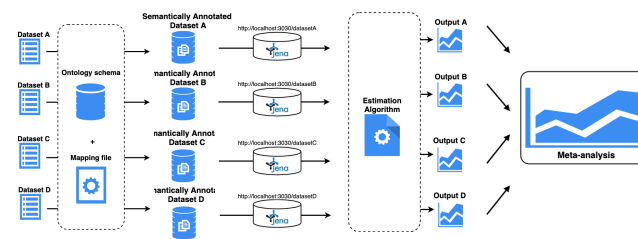
Output C



Output D



Meta-analysis



Study

treatment = (Intercept) broiler

DatasetA

DatasetB

$I^2 = 99\%$, $\tau^2 = 9.4815$, $p < 0.01$

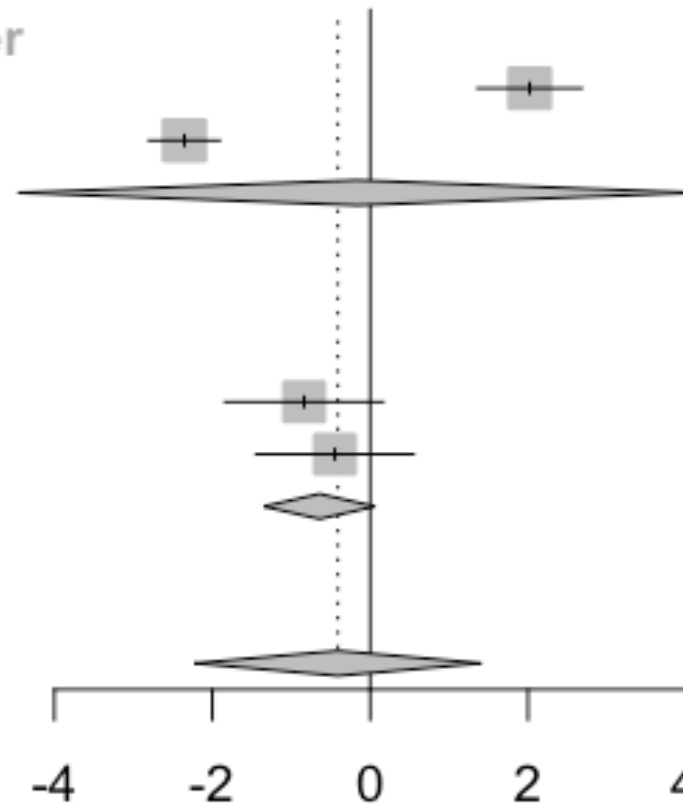
treatment = (Intercept) pig

DatasetC

DatasetD

$I^2 = 0\%$, $\tau^2 = 0$, $p = 0.59$

Transmission
Rate ($\log(\beta)$)



Benefits for Researchers and Domain experts

- Data has context
- Enables reuse of algorithms (estimation scripts)
- Enables collaboration without sharing of data (no DTA!)

Future work

- Apply our approach to examples in the human medical domain
- Enable access to the data without transferring it, i.e., to allow federated data analysis

Acknowledgements



ZonMw

Project is supported by ZonMW (project-number 541003004)

Collaborators:



Jan Top



Ewa Pacholewicz



Peter Hobbelen



Martin Bootsma



UMC Utrecht



**Utrecht
University**



Martine de Vos



Miel Hostens



Egil Fischer



<https://github.com/UtrechtUniversity/summer-fair/wiki>