

# ***Semantic workflows: design & provenance on the Virtual Imaging Platform (VIP project)***

---

Nadia Cerezo, Alban Gaignard, Johan Montagnat

Modalis team - I3S laboratory - CNRS / UNSA  
Sophia Antipolis – 17/10/2012

# Addressing **information overload** in e-science activities

---

- **e-Science platforms**

- Data (variety)
- Processing tools
- Workflows
- Domain ontologies

- **Problem**

- **Relevant information** specific to each user (developers, scientists, etc.)
- Several **intertwined granularities** and **representation** layers
- ➔ Information overload

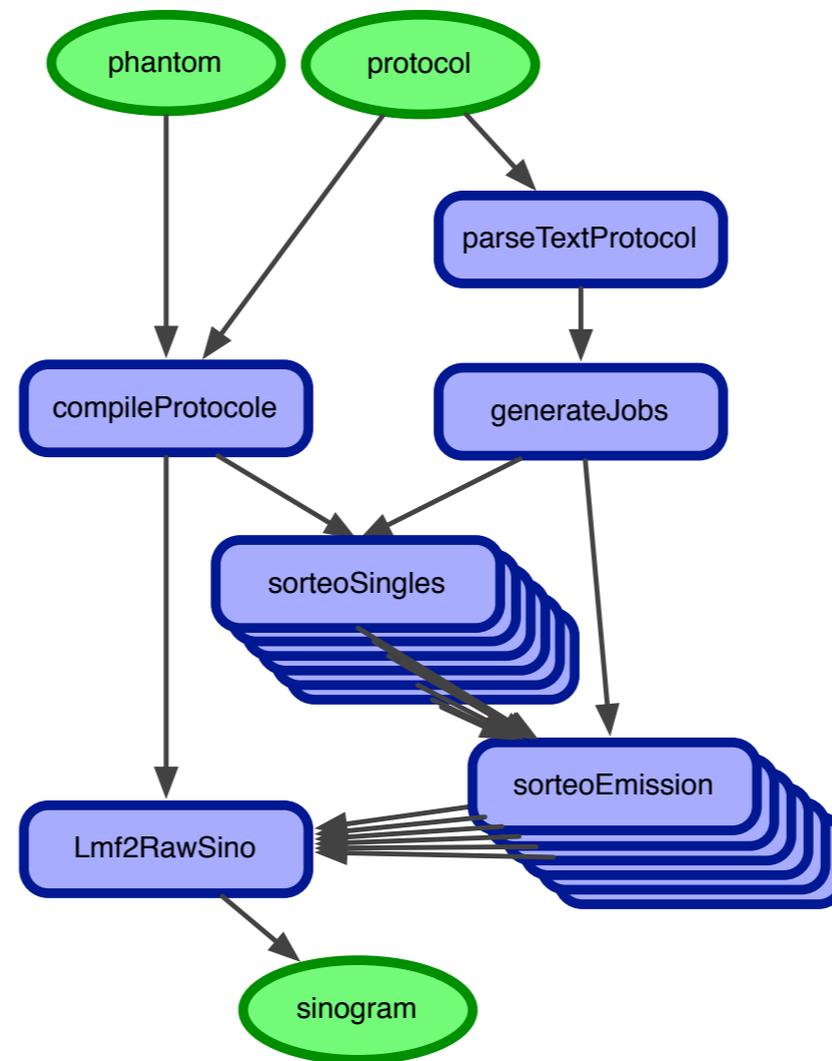
- Help e-scientists to **focus** on **meaningful** information

- exploiting **data** in e-science experiments
- exploiting e-science **workflows**

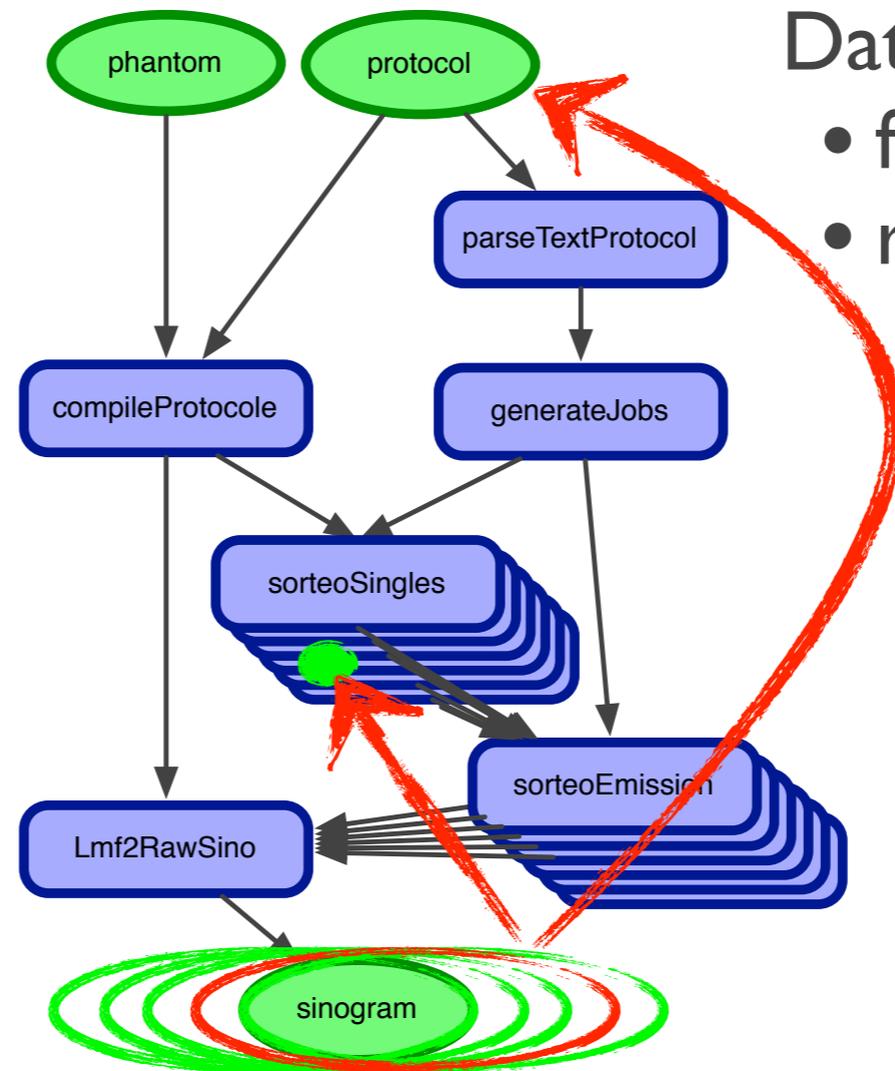
Exploiting **data** and domain **knowledge**  
through e-science workflows

# Semantic Provenance 1/3: Why annotating workflow data production ?

---



# Semantic Provenance 1/3: Why annotating workflow data production ?

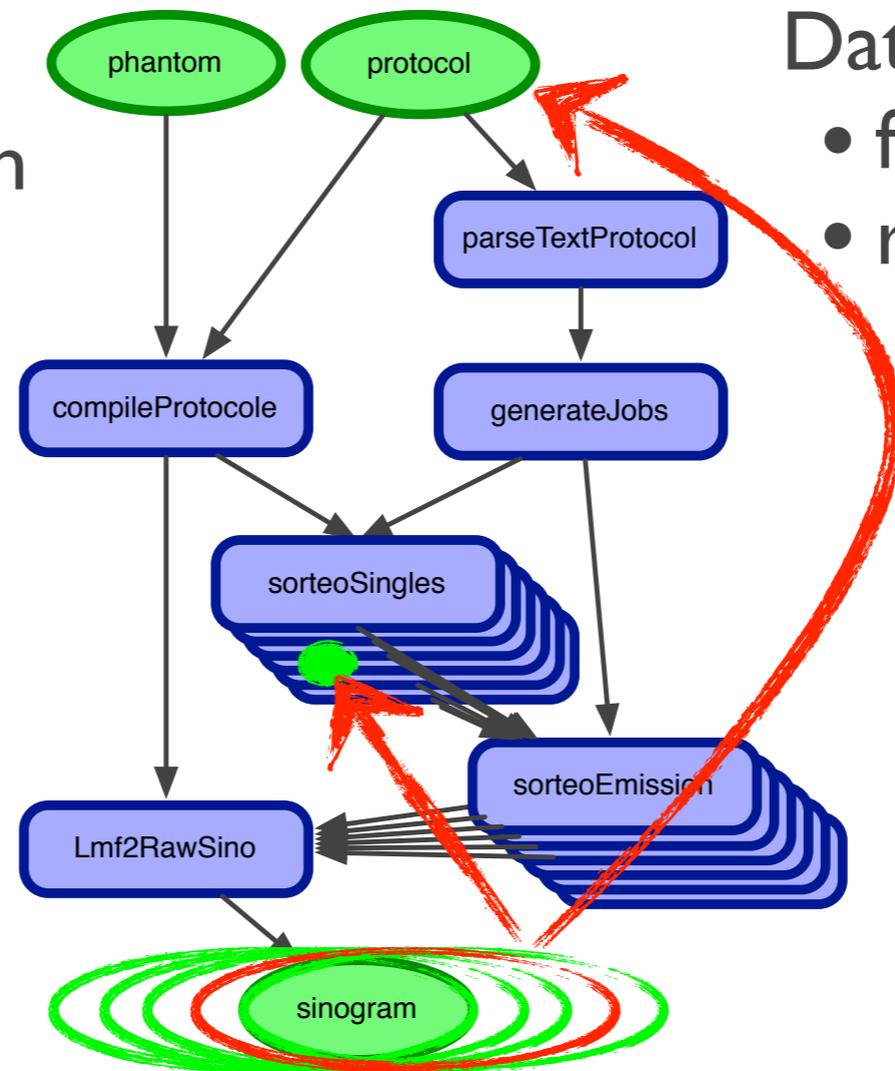


## Data **causality**:

- failures / abnormalities ?
- reproducibility ?

# Semantic Provenance 1/3: Why annotating workflow data production ?

Propagate **knowledge** on data & processing tools ?

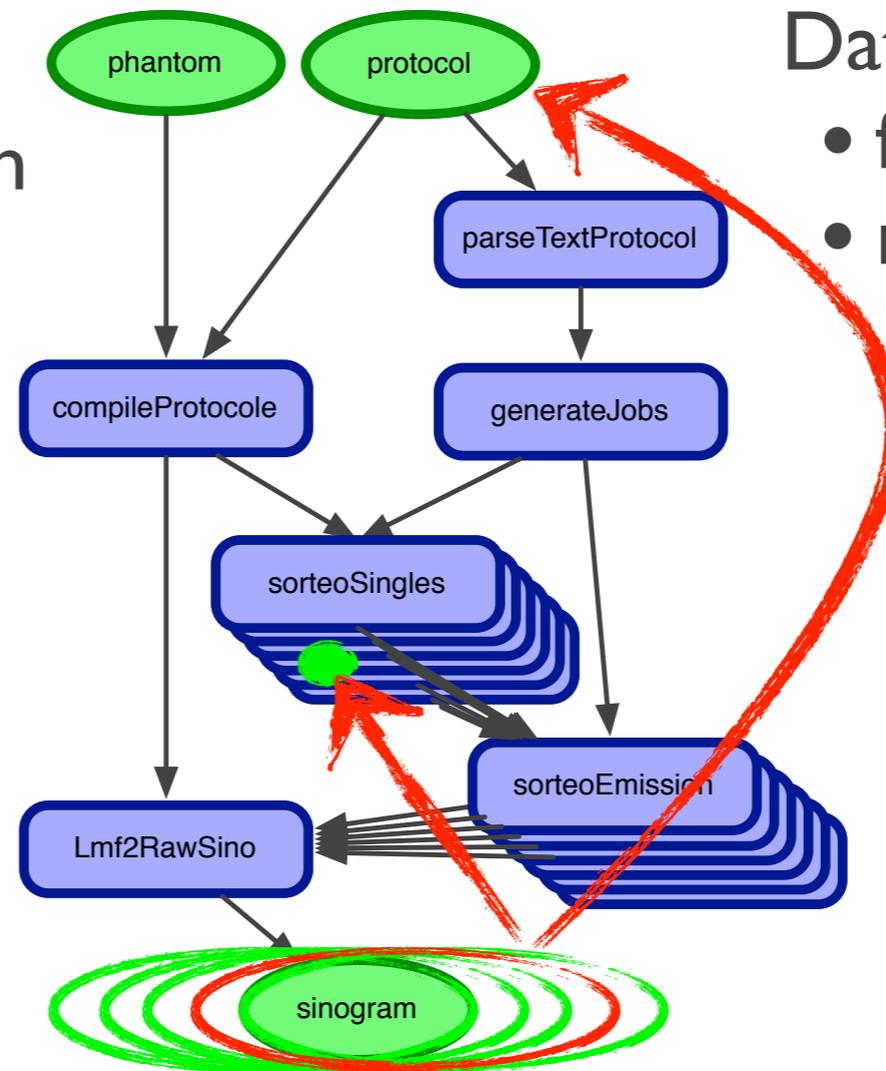


Data **causality**:

- failures / abnormalities ?
- reproducibility ?

# Semantic Provenance 1/3: Why annotating workflow data production ?

Propagate **knowledge** on data & processing tools ?



Data **causality**:

- failures / abnormalities ?
- reproducibility ?

## Objectives

1. Track **data provenance** to enhance **quality**
2. Infer **new domain-specific statements** (from WF runs)

➔ *meaningful experiment summaries*

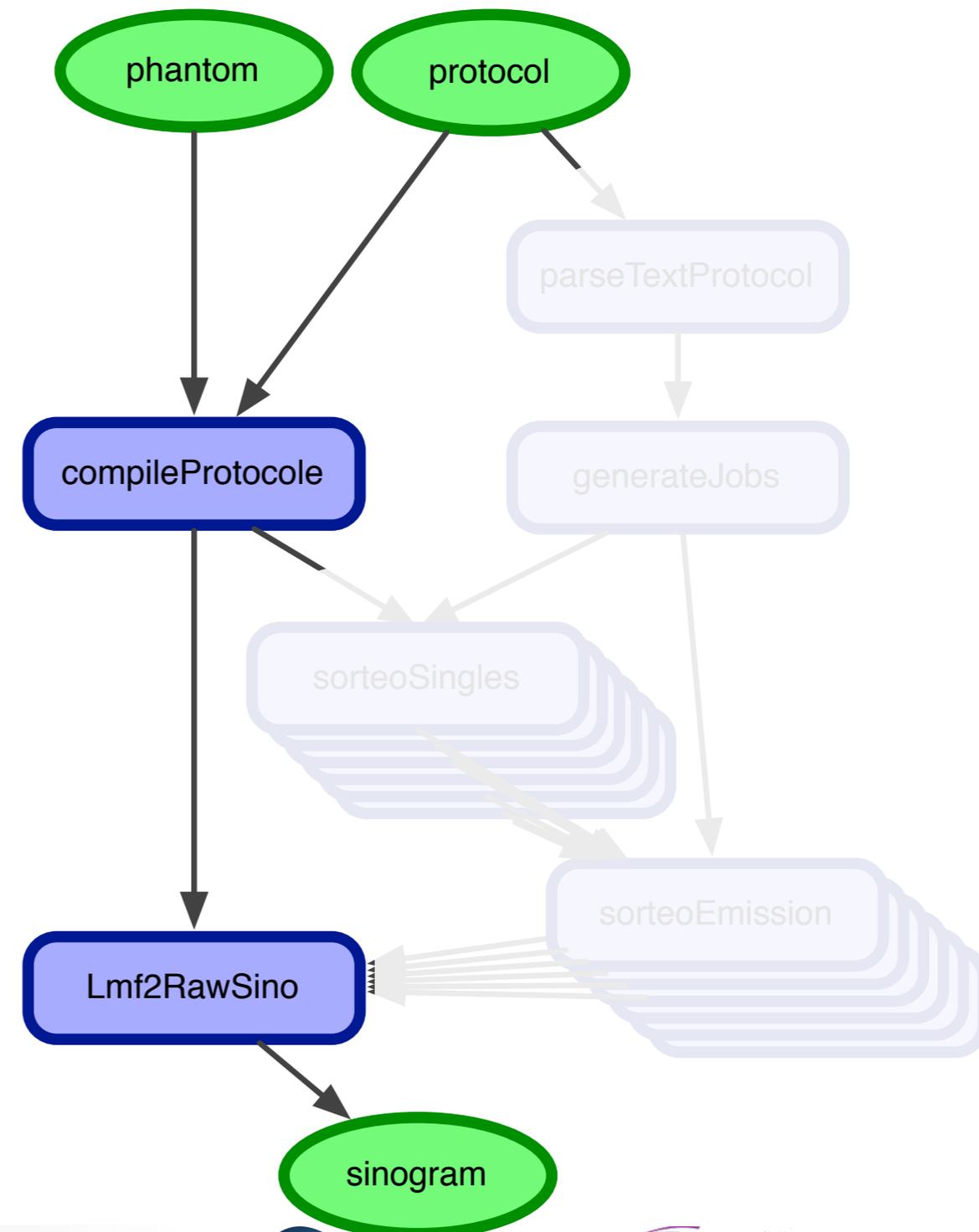
# Semantic Provenance 2/3: Methods

- Semantic annotations for WF components + OPM
- Rules as SPARQL CONSTRUCT queries

```

1 PREFIX opmo: <http://openprovenance.org/model/opmo#>
2 PREFIX opmv: <http://purl.org/net/opmv/ns#>
3 CONSTRUCT {
4   ?in rdf:type <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
5   ?out rdf:type <http://vip.cosinus.anr.fr/simulated-data.owl#pet-sinogram>
6   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#has-for-phantom ?in
7   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#results-from ?x
8 } WHERE {
9   ?agent (ie:refers-to/rdf:type)
10    <http://vip.cosinus.anr.fr/vip-simulation.owl#image-reconstruction-simulator-component> .
11   ?wcb opmo:cause ?agent .
12   ?wcb opmo:effect ?x .
13   ?x rdf:type opmv:Process .
14
15   ?wgb opmo:cause ?x .
16   ?wgb opmo:effect ?out .
17   ?out (opmo:avalue/opmo:content) ?cOut
18
19   ?agent2 (ie:refers-to/rdf:type)
20    <http://vip.cosinus.anr.fr/vip-simulation.owl#parameters-generation-simulator-component> .
21   ?wcb2 opmo:cause ?agent2 .
22   ?wcb2 opmo:effect ?y .
23   ?y rdf:type opmv:Process .
24
25   ?used opmo:cause ?in .
26   ?used opmo:effect ?y .
27   ?in (opmo:avalue/opmo:content) ?cln .
28   ?used (opmo:role/rdfs:label) ?role .
29
30   ?agent2 ws:has-input ?inPort .
31   ?inPort (ie:refers-to/rdf:type)
32    <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
33   ?inPort rdfs:comment ?role .
34 }

```

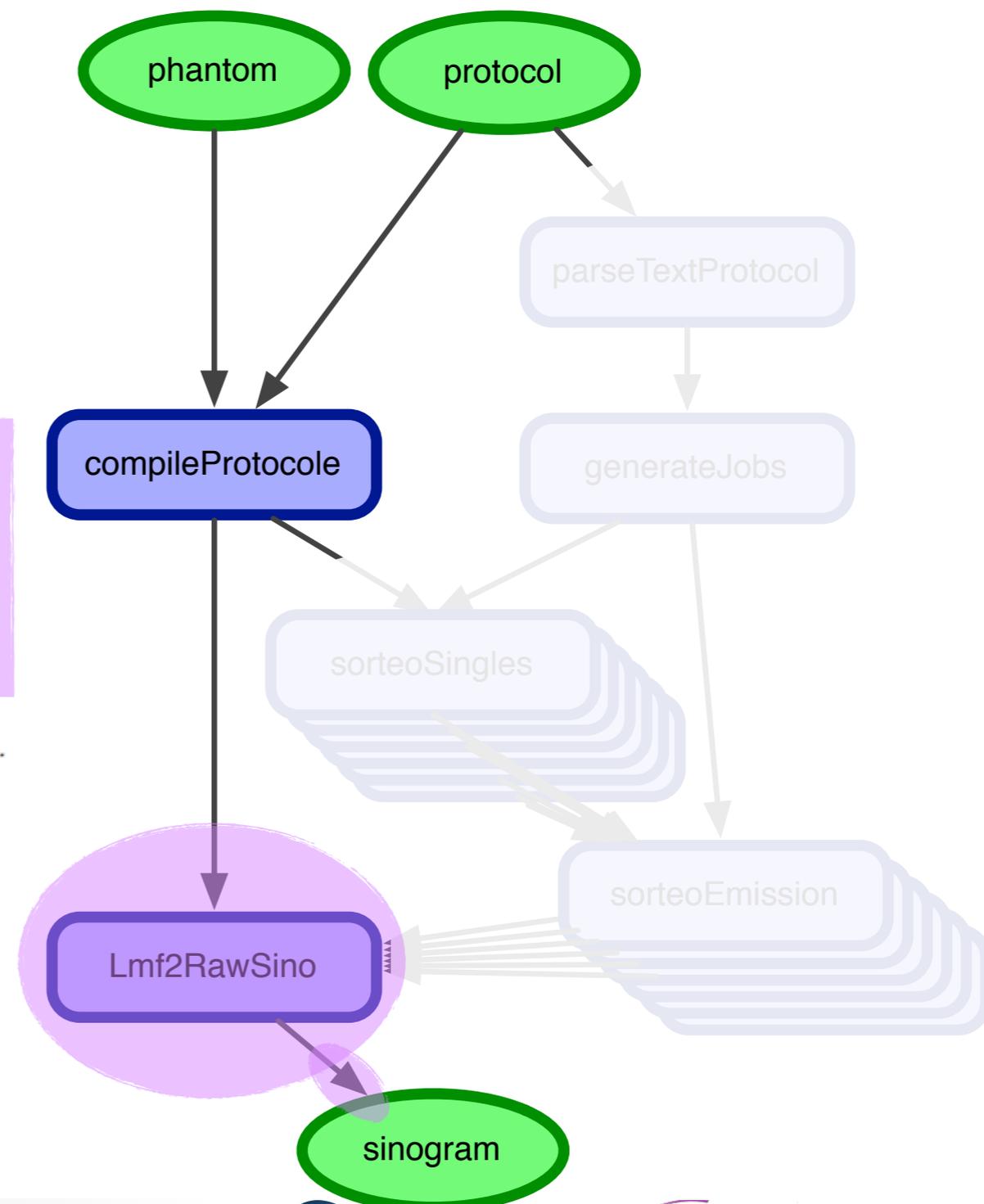


# Semantic Provenance 2/3: Methods

- Semantic annotations for WF components + OPM
- Rules as SPARQL CONSTRUCT queries

```

1 PREFIX opmo: <http://openprovenance.org/model/opmo#>
2 PREFIX opmv: <http://purl.org/net/opmv/ns#>
3 CONSTRUCT {
4   ?in rdf:type <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
5   ?out rdf:type <http://vip.cosinus.anr.fr/simulated-data.owl#pet-sinogram> .
6   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#has-for-phantom ?in> .
7   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#results-from ?x> .
8 } WHERE {
9   ?agent (ie:refers-to/rdf:type)
10    <http://vip.cosinus.anr.fr/vip-simulation.owl#image-reconstruction-simulator-component> .
11   ?wcb opmo:cause ?agent .
12   ?wcb opmo:effect ?x .
13   ?x rdf:type opmv:Process .
14
15   ?wgb opmo:cause ?x .
16   ?wgb opmo:effect ?out .
17   ?out (opmo:avalue/opmo:content) ?cOut
18
19   ?agent2 (ie:refers-to/rdf:type)
20    <http://vip.cosinus.anr.fr/vip-simulation.owl#parameters-generation-simulator-component> .
21   ?wcb2 opmo:cause ?agent2 .
22   ?wcb2 opmo:effect ?y .
23   ?y rdf:type opmv:Process .
24
25   ?used opmo:cause ?in .
26   ?used opmo:effect ?y .
27   ?in (opmo:avalue/opmo:content) ?cIn .
28   ?used (opmo:role/rdfs:label) ?role .
29
30   ?agent2 ws:has-input ?inPort .
31   ?inPort (ie:refers-to/rdf:type)
32    <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
33   ?inPort rdfs:comment ?role .
34 }
  
```



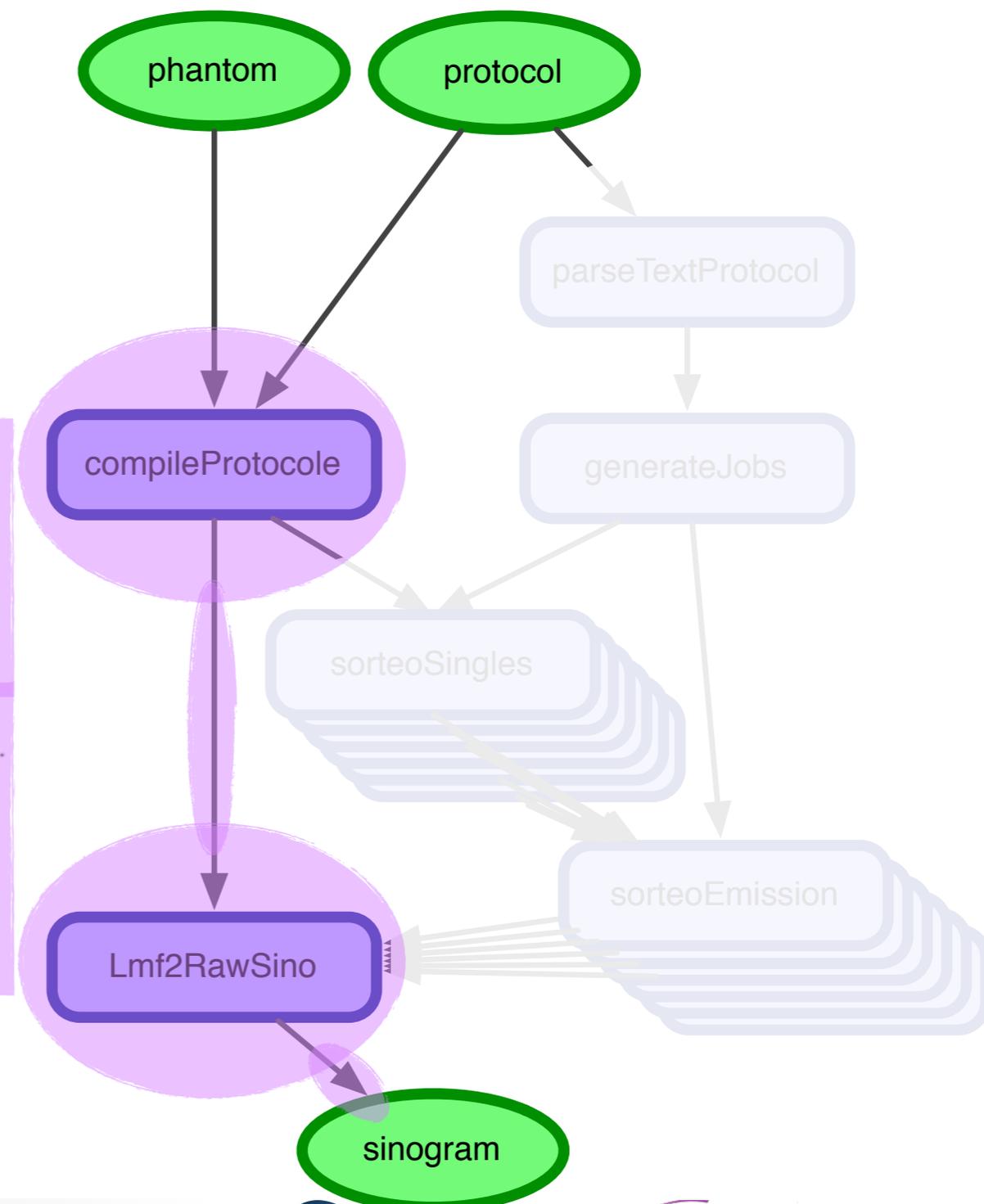
# Semantic Provenance 2/3: Methods

- Semantic annotations for WF components + OPM
- Rules as SPARQL CONSTRUCT queries

```

1 PREFIX opmo: <http://openprovenance.org/model/opmo#>
2 PREFIX opmv: <http://purl.org/net/opmv/ns#>
3 CONSTRUCT {
4   ?in rdf:type <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
5   ?out rdf:type <http://vip.cosinus.anr.fr/simulated-data.owl#pet-sinogram>
6   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#has-for-phantom ?in
7   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#results-from ?x
8 } WHERE {
9   ?agent (ie:refers-to/rdf:type)
10    <http://vip.cosinus.anr.fr/vip-simulation.owl#image-reconstruction-simulator-component> .
11   ?wcb opmo:cause ?agent .
12   ?wcb opmo:effect ?x .
13   ?x rdf:type opmv:Process .
14
15   ?wgb opmo:cause ?x .
16   ?wgb opmo:effect ?out .
17   ?out (opmo:avalue/opmo:content) ?cOut
18
19   ?agent2 (ie:refers-to/rdf:type)
20    <http://vip.cosinus.anr.fr/vip-simulation.owl#parameters-generation-simulator-component> .
21   ?wcb2 opmo:cause ?agent2 .
22   ?wcb2 opmo:effect ?y .
23   ?y rdf:type opmv:Process .
24
25   ?used opmo:cause ?in .
26   ?used opmo:effect ?y .
27   ?in (opmo:avalue/opmo:content) ?cln .
28   ?used (opmo:role/rdfs:label) ?role .
29
30   ?agent2 ws:has-input ?inPort .
31   ?inPort (ie:refers-to/rdf:type)
32    <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
33   ?inPort rdfs:comment ?role .
34 }

```



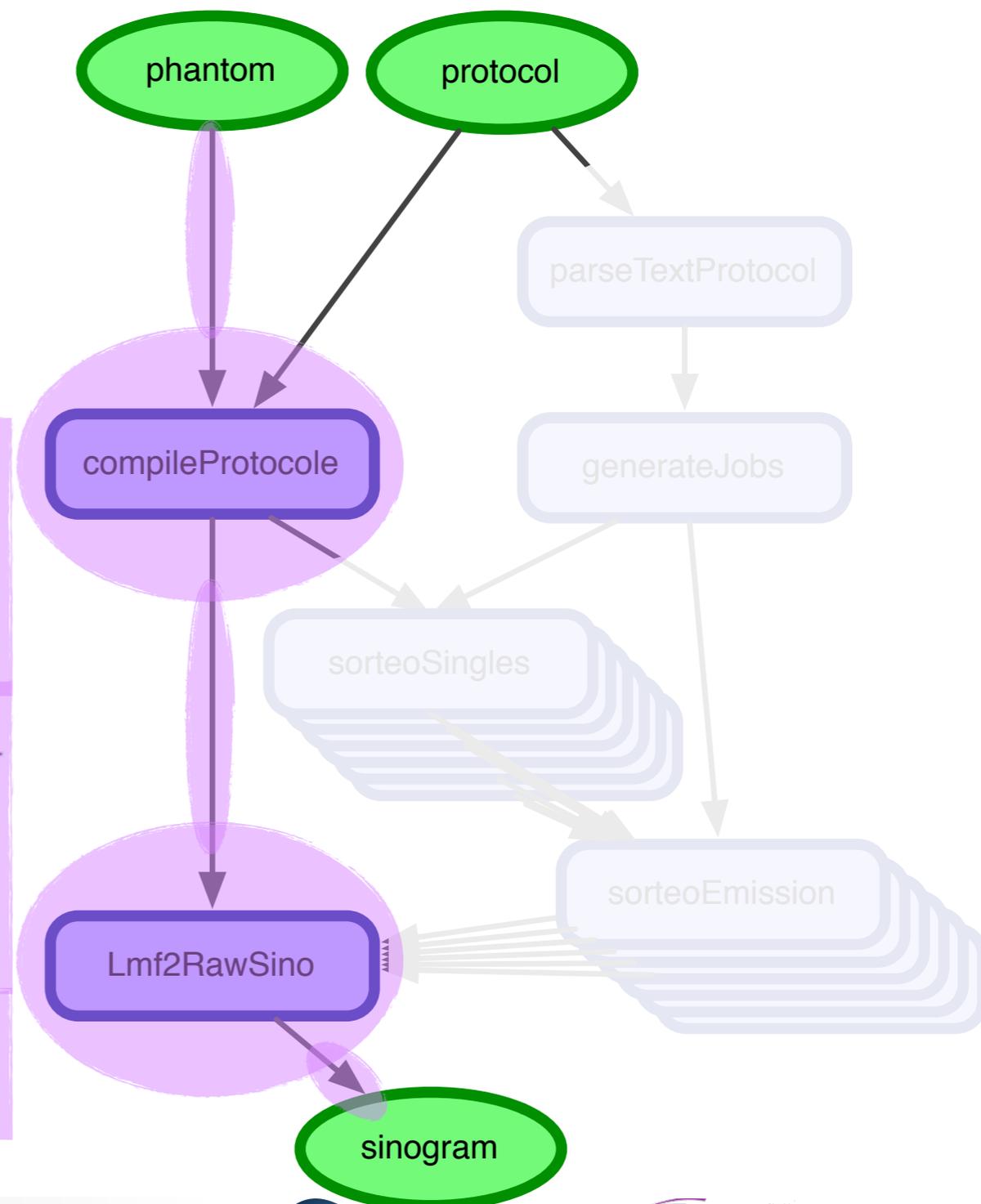
# Semantic Provenance 2/3: Methods

- Semantic annotations for WF components + OPM
- Rules as SPARQL CONSTRUCT queries

```

1 PREFIX opmo: <http://openprovenance.org/model/opmo#>
2 PREFIX opmv: <http://purl.org/net/opmv/ns#>
3 CONSTRUCT {
4   ?in rdf:type <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
5   ?out rdf:type <http://vip.cosinus.anr.fr/simulated-data.owl#pet-sinogram> .
6   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#has-for-phantom ?in
7   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#results-from ?x
8 } WHERE {
9   ?agent (ie:refers-to/rdf:type)
10    <http://vip.cosinus.anr.fr/vip-simulation.owl#image-reconstruction-simulator-component> .
11   ?wcb opmo:cause ?agent .
12   ?wcb opmo:effect ?x .
13   ?x rdf:type opmv:Process .
14
15   ?wgb opmo:cause ?x .
16   ?wgb opmo:effect ?out .
17   ?out (opmo:avalue/opmo:content) ?cOut
18
19   ?agent2 (ie:refers-to/rdf:type)
20    <http://vip.cosinus.anr.fr/vip-simulation.owl#parameters-generation-simulator-component> .
21   ?wcb2 opmo:cause ?agent2 .
22   ?wcb2 opmo:effect ?y .
23   ?y rdf:type opmv:Process .
24
25   ?used opmo:cause ?in .
26   ?used opmo:effect ?y .
27   ?in (opmo:avalue/opmo:content) ?cln .
28   ?used (opmo:role/rdfs:label) ?role .
29
30   ?agent2 ws:has-input ?inPort .
31   ?inPort (ie:refers-to/rdf:type)
32    <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
33   ?inPort rdfs:comment ?role .
34 }

```



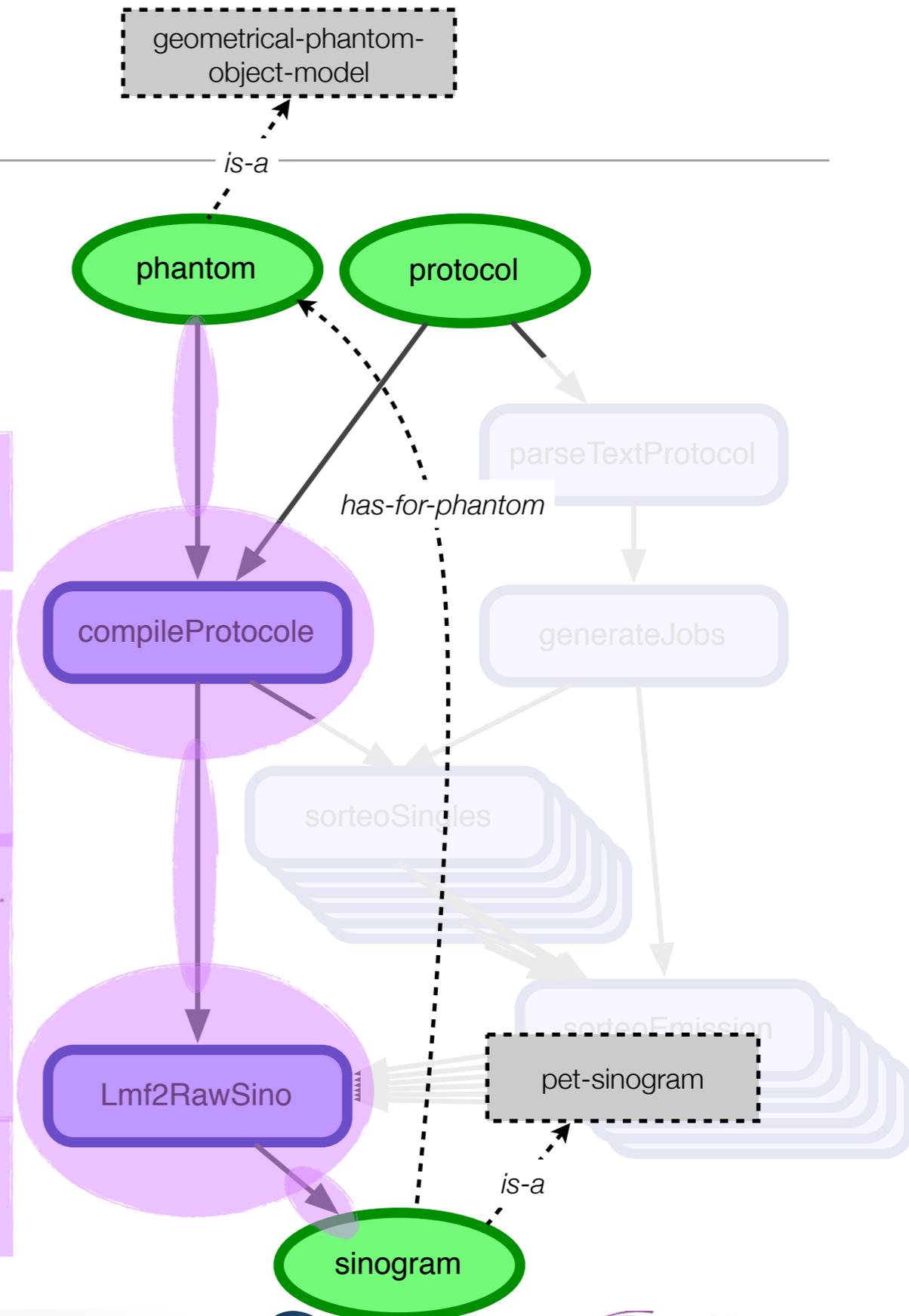
# Semantic Provenance 2/3: Methods

- Semantic annotations for WF components + OPM
- Rules as SPARQL CONSTRUCT queries

```

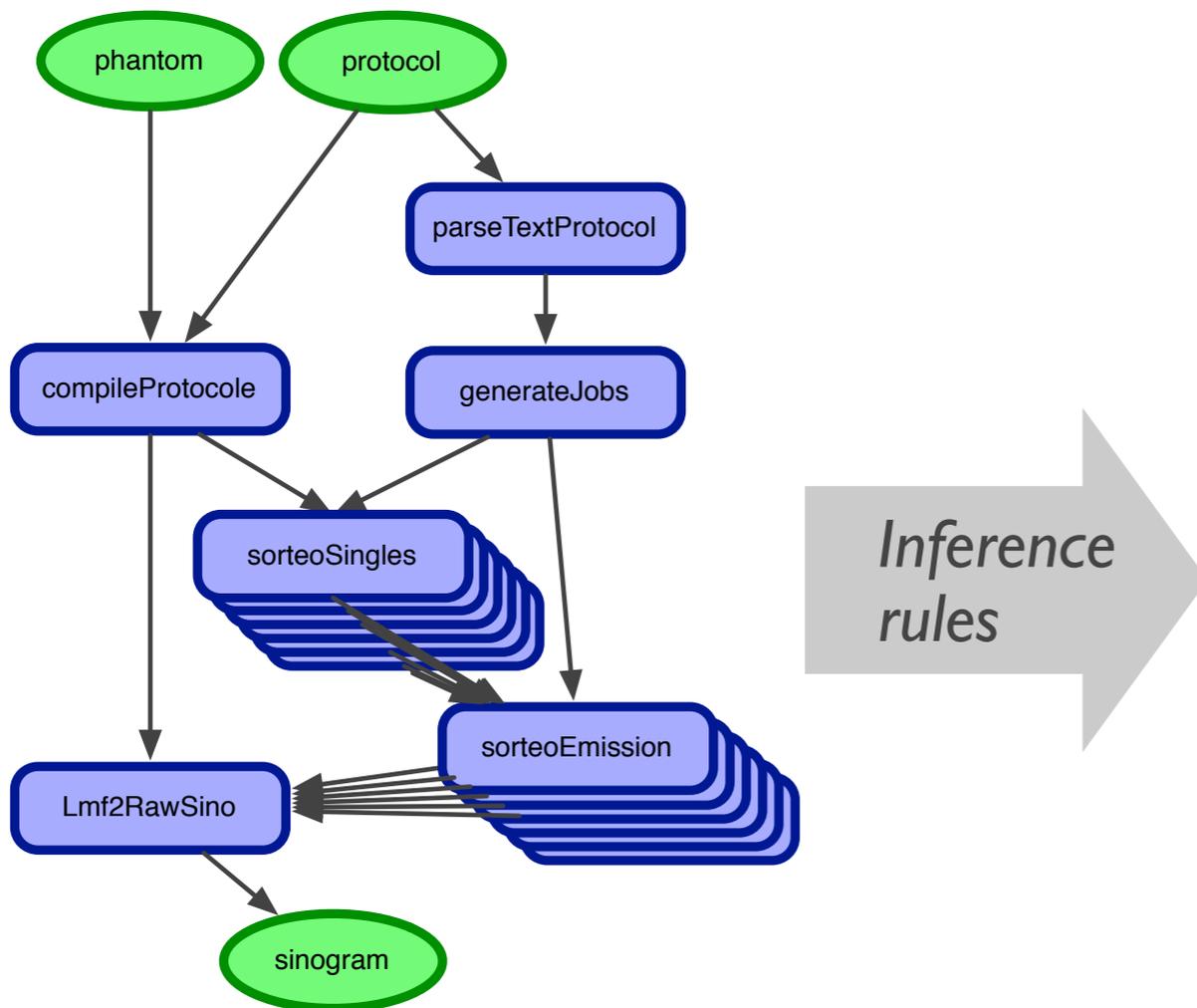
1 PREFIX opmo: <http://openprovenance.org/model/opmo#>
2 PREFIX opmv: <http://purl.org/net/opmv/ns#>
3 CONSTRUCT {
4   ?in rdf:type <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
5   ?out rdf:type <http://vip.cosinus.anr.fr/simulated-data.owl#pet-sinogram> .
6   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#has-for-phantom ?in> .
7   ?out <http://vip.cosinus.anr.fr/simulated-data.owl#results-from ?x> .
8 } WHERE {
9   ?agent (ie:refers-to/rdf:type)
10    <http://vip.cosinus.anr.fr/vip-simulation.owl#image-reconstruction-simulator-component> .
11   ?wcb opmo:cause ?agent .
12   ?wcb opmo:effect ?x .
13   ?x rdf:type opmv:Process .
14
15   ?wgb opmo:cause ?x .
16   ?wgb opmo:effect ?out .
17   ?out (opmo:avalue/opmo:content) ?cOut
18
19   ?agent2 (ie:refers-to/rdf:type)
20    <http://vip.cosinus.anr.fr/vip-simulation.owl#parameters-generation-simulator-component> .
21   ?wcb2 opmo:cause ?agent2 .
22   ?wcb2 opmo:effect ?y .
23   ?y rdf:type opmv:Process .
24
25   ?used opmo:cause ?in .
26   ?used opmo:effect ?y .
27   ?in (opmo:avalue/opmo:content) ?cIn .
28   ?used (opmo:role/rdfs:label) ?role .
29
30   ?agent2 ws:has-input ?inPort .
31   ?inPort (ie:refers-to/rdf:type)
32    <http://vip.cosinus.anr.fr/vip-model.owl#geometrical-phantom-object-model> .
33   ?inPort rdfs:comment ?role .
34 }

```



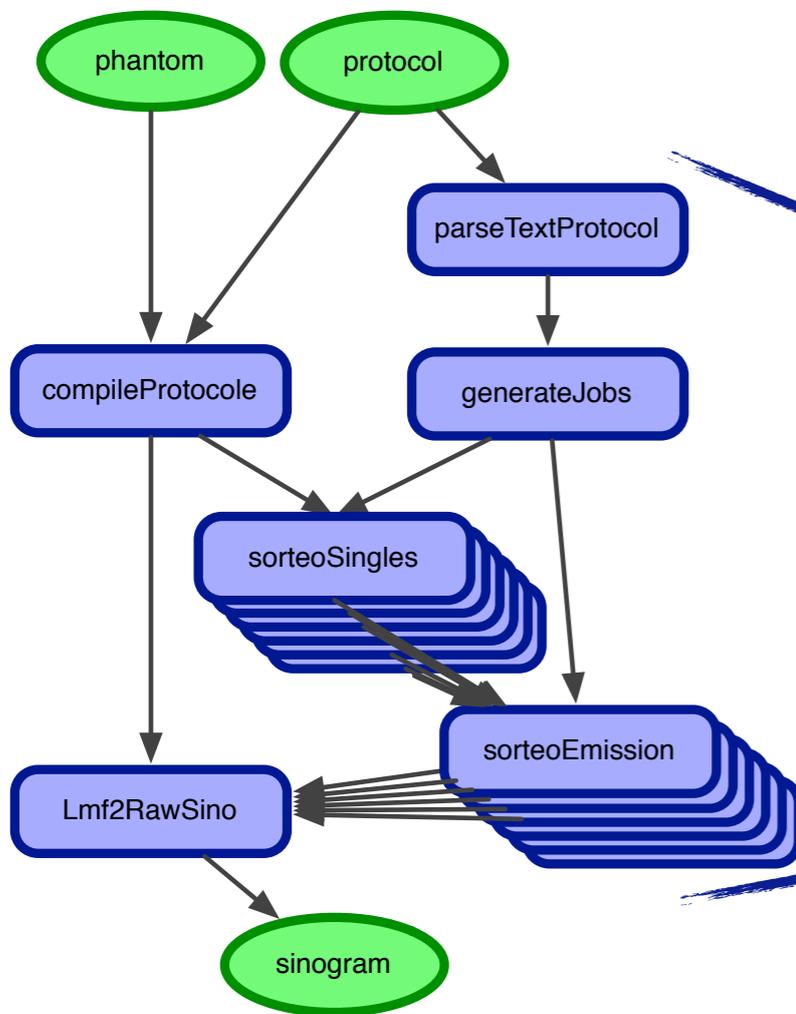
# Semantic Provenance 3/3: Status

Fine-grained & technical provenance :



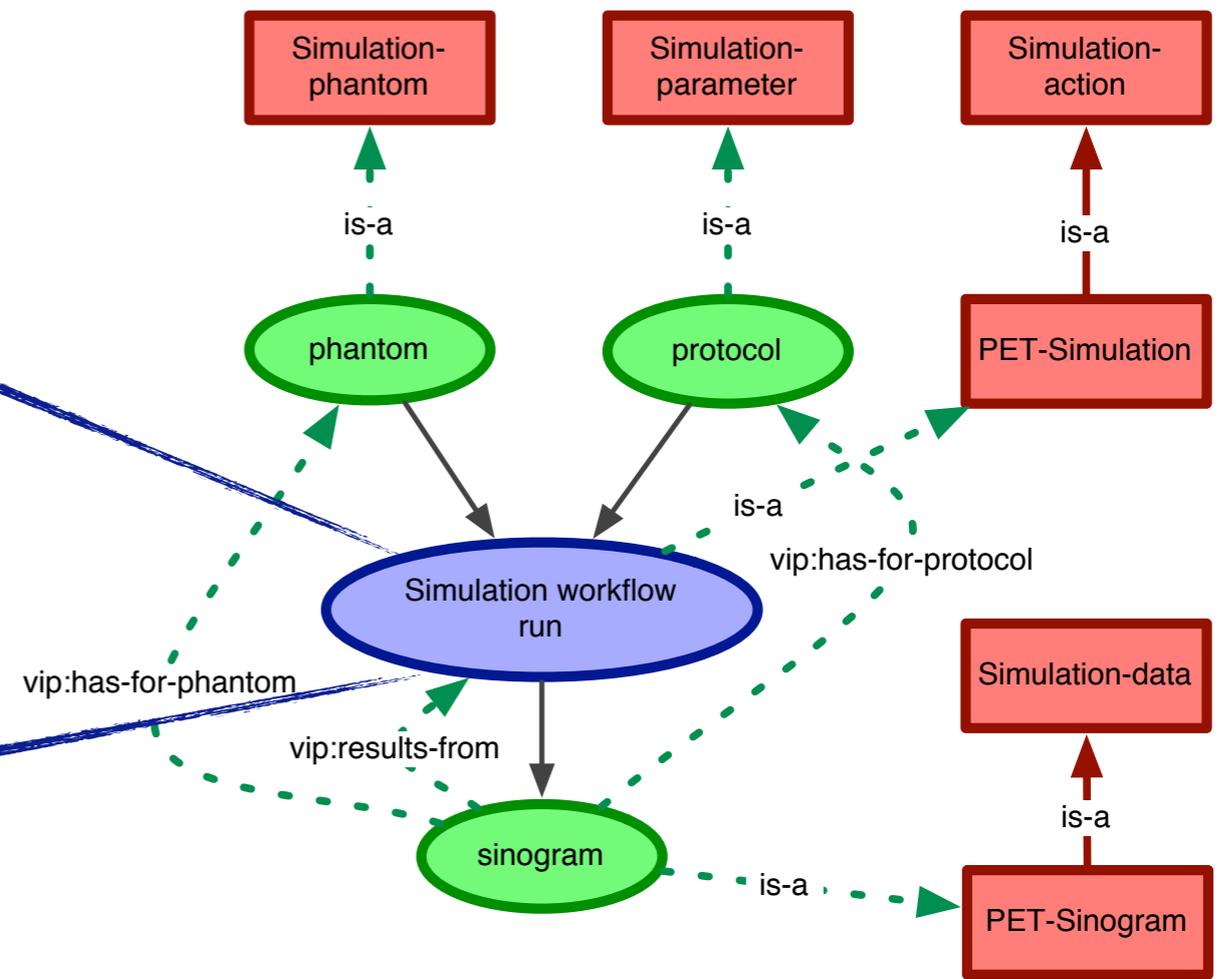
# Semantic Provenance 3/3: Status

Fine-grained & technical provenance :



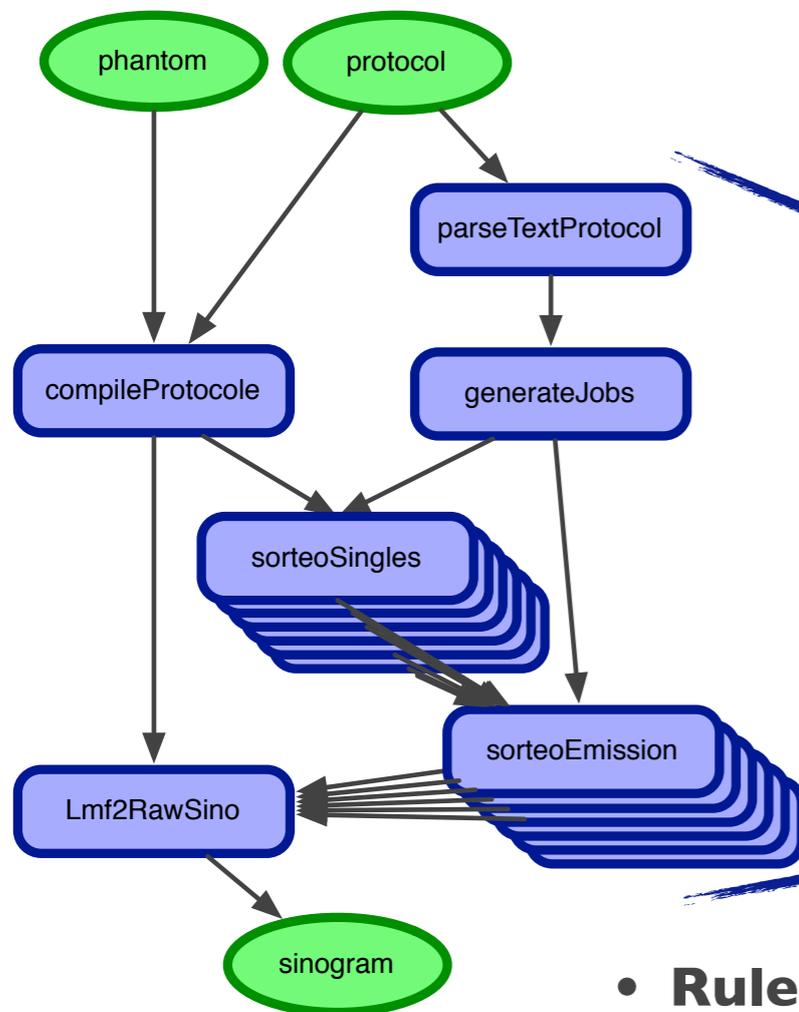
*Inference rules*

Coarse-grained & meaningful provenance :

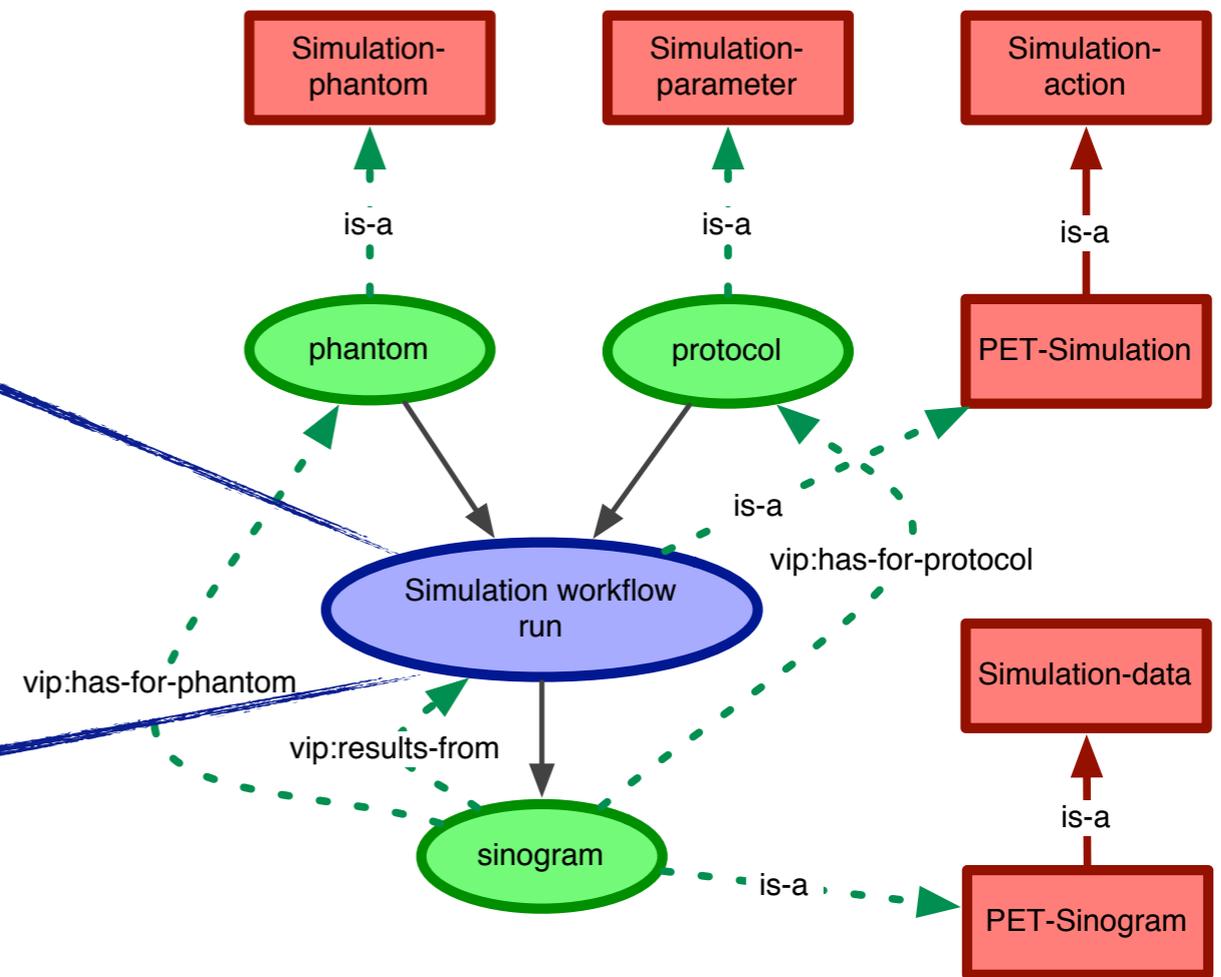


# Semantic Provenance 3/3: Status

Fine-grained & technical provenance :



Coarse-grained & meaningful provenance :



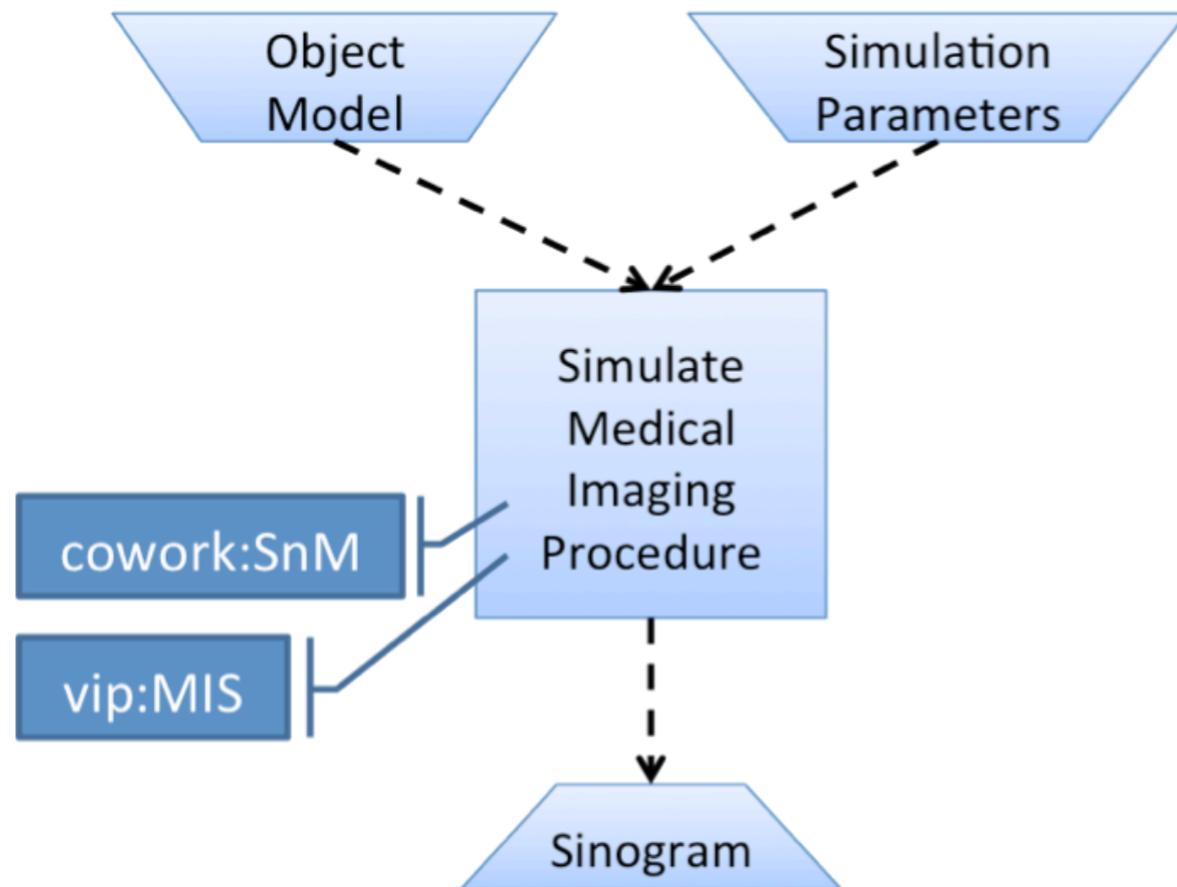
Inference rules

- **Rules genericity ?**
  - ✓ adapt to simulator component evolutions
  - do not adapt to workflow structure evolutions
- Inferences from **Distributed provenance ?**

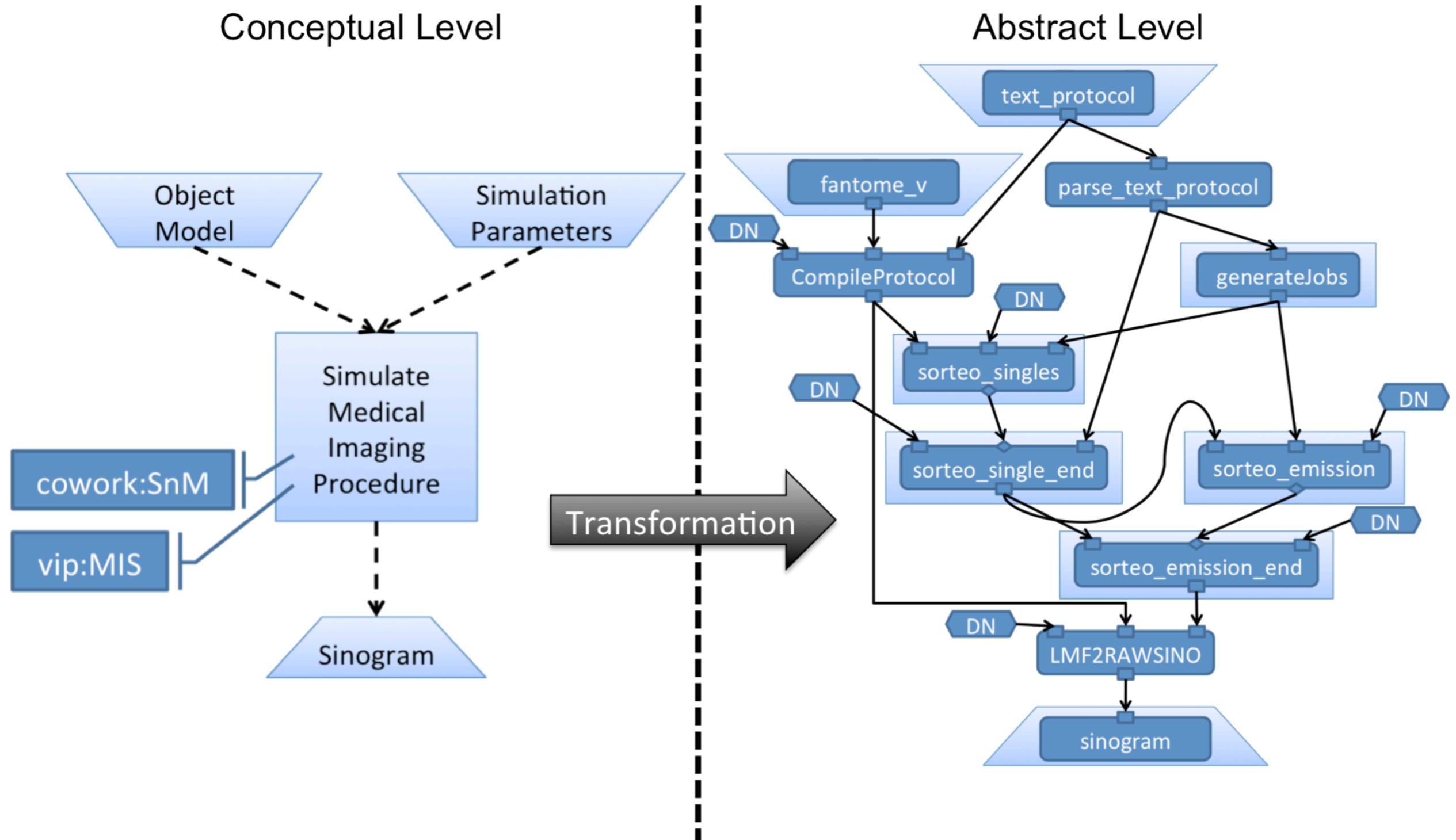
Exploiting domain **knowledge**  
for e-Science **workflow** design

# Conceptual workflows 1/3: Objectives

## Conceptual Level

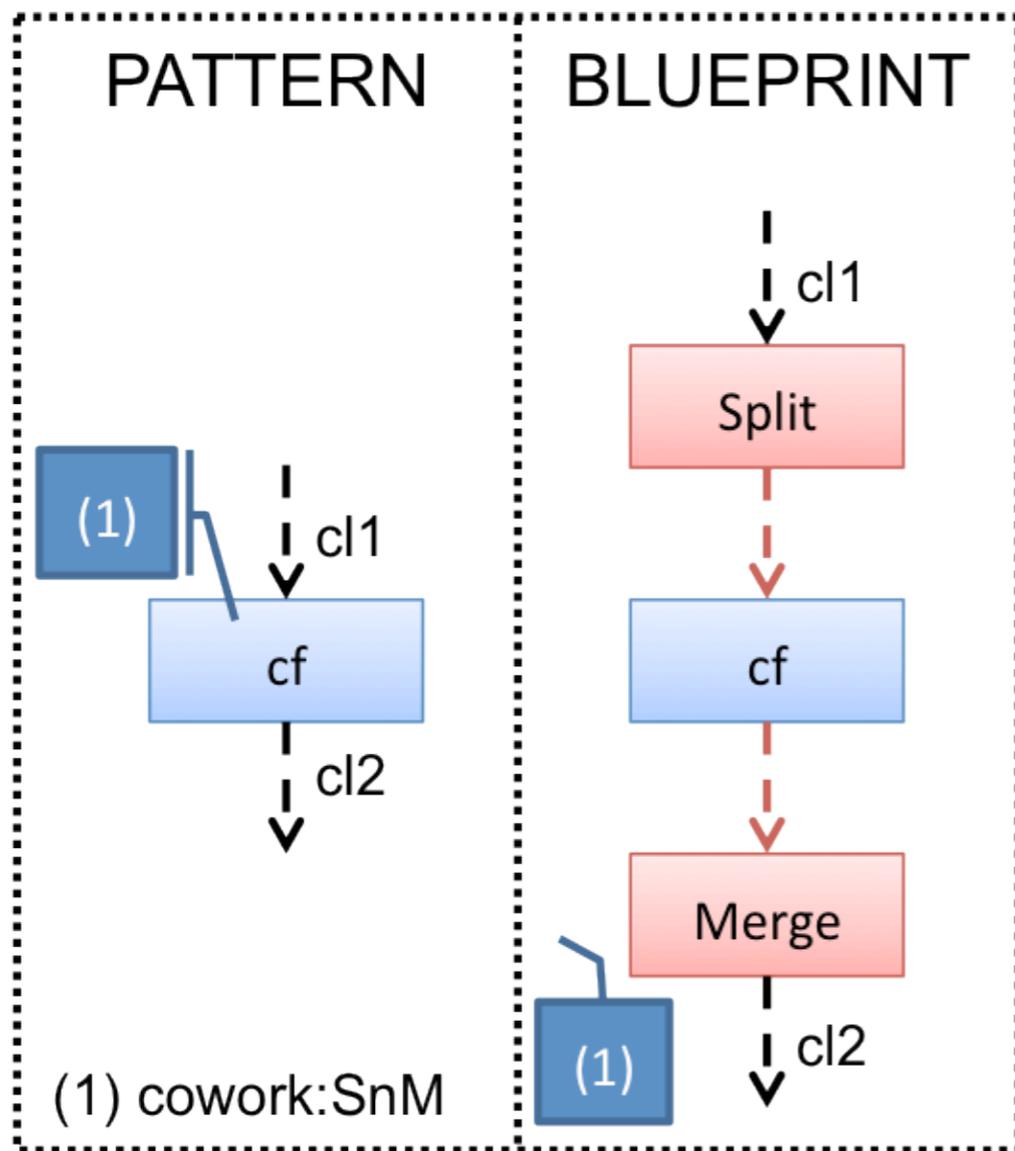


# Conceptual workflows 1/3: Objectives



# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

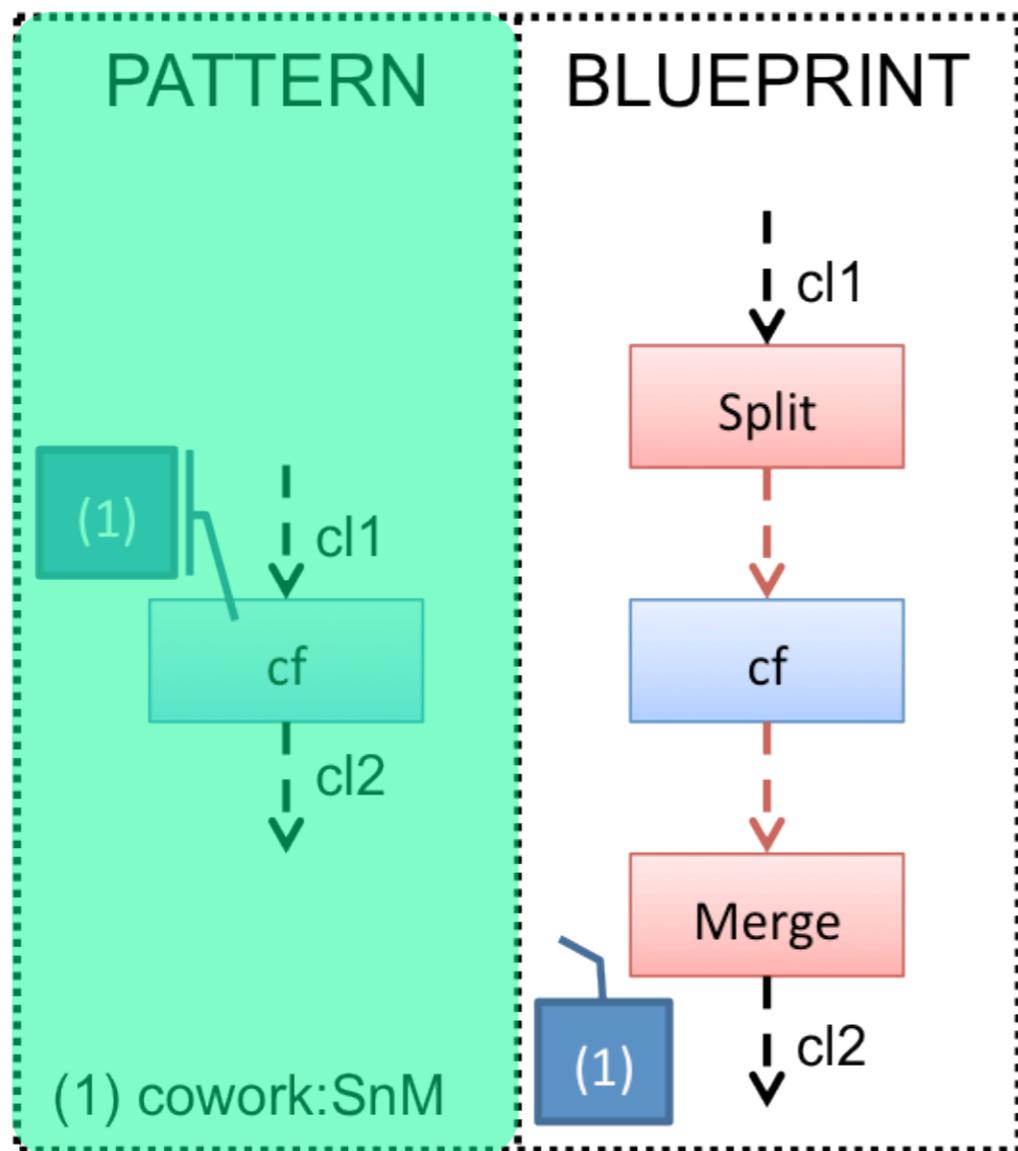
```
PREFIX : <(cowork-url)>
CONSTRUCT {
  ?cl1 :hasConceptualTarget _:Split .
  _:l1 :hasConceptualSource _:Split .
  _:l1 :hasConceptualTarget ?cf .
  _:l2 :hasConceptualSource ?cf .
  _:l2 :hasConceptualTarget _:Merge .
  ?cl2 :hasConceptualSource _:Merge .
  ?par :hasSpecification :SnM .
}

WHERE {
  ?cl1 :hasConceptualTarget ?cf .
  ?cl2 :hasConceptualSource ?cf .
  ?cf :hasRequirement :SnM .
}
```

(cowork-url) <http://www.i3s.unice.fr/~cerezo/cowork/latest/cowork.rdfs>

# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

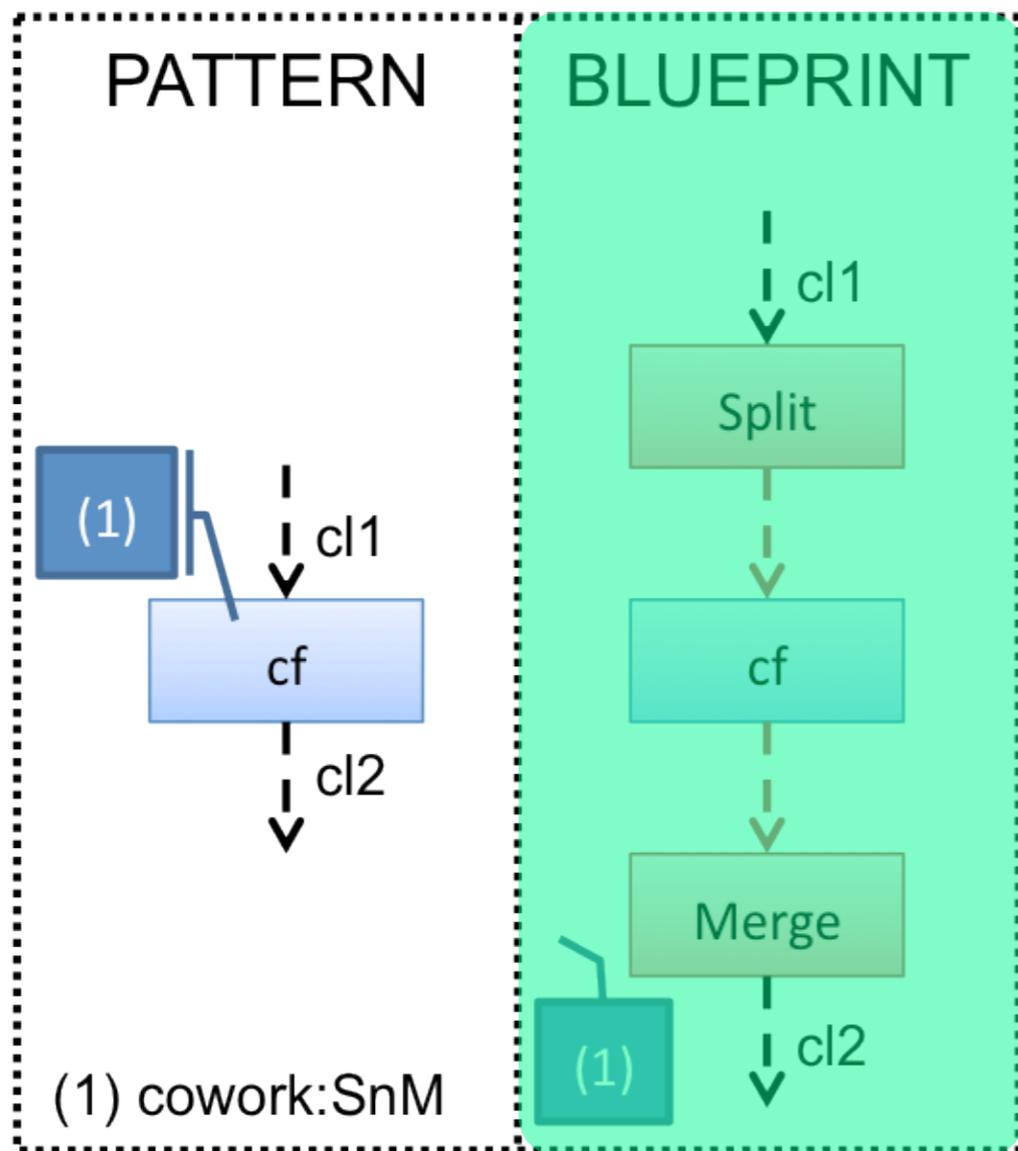
```
PREFIX : <(cwork-url)>
CONSTRUCT {
  ?cl1 :hasConceptualTarget _:Split .
  _:l1 :hasConceptualSource _:Split .
  _:l1 :hasConceptualTarget ?cf .
  _:l2 :hasConceptualSource ?cf .
  _:l2 :hasConceptualTarget _:Merge .
  ?cl2 :hasConceptualSource _:Merge .
  ?par :hasSpecification :SnM .
}
```

```
WHERE {
  ?cl1 :hasConceptualTarget ?cf .
  ?cl2 :hasConceptualSource ?cf .
  ?cf :hasRequirement :SnM .
}
```

(cwork-url) <http://www.i3s.unice.fr/~cerezo/cwork/latest/cwork.rdfs>

# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

PREFIX : <(cowork-url)>

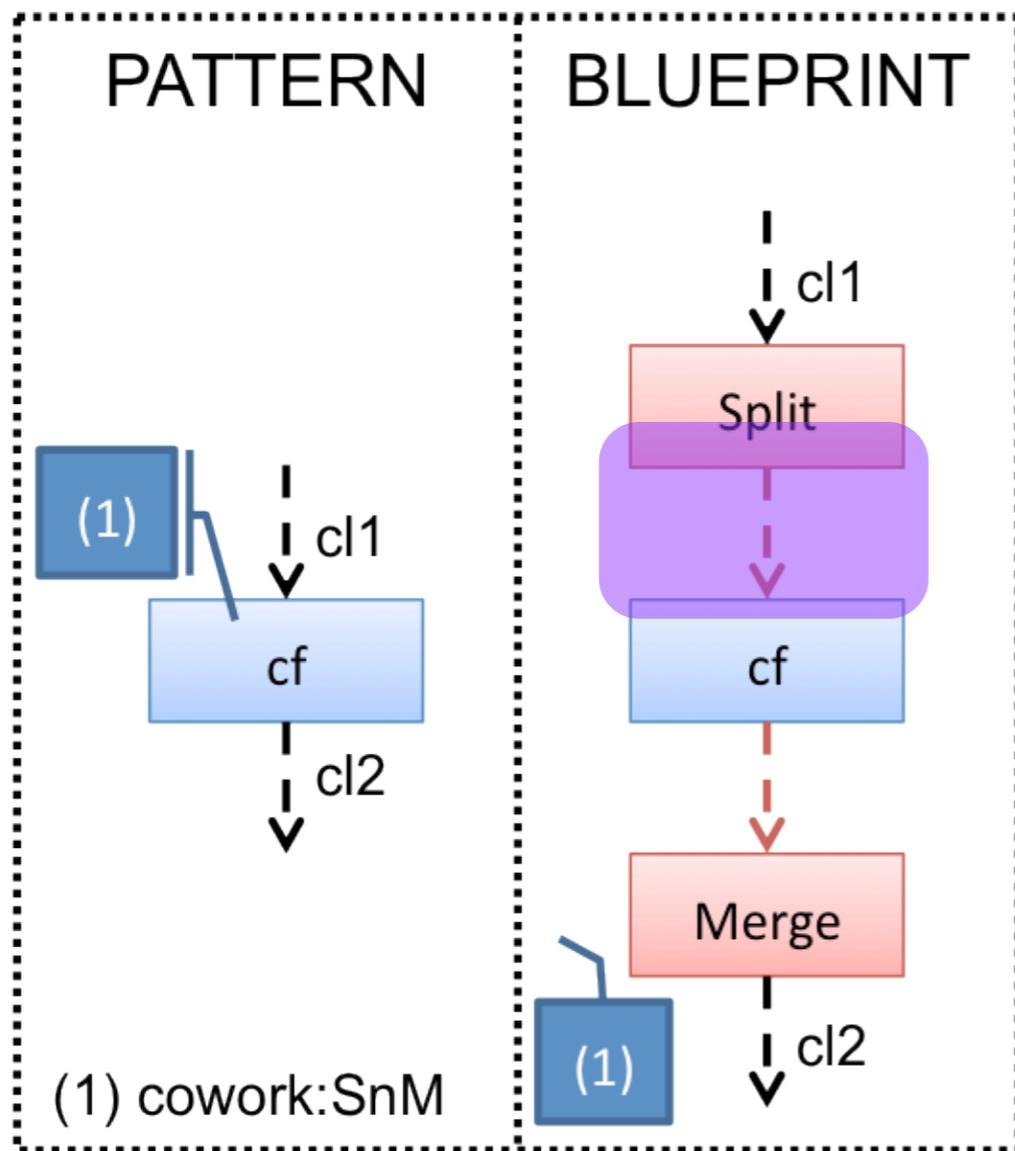
```
CONSTRUCT {  
  ?cl1 :hasConceptualTarget _:Split .  
  _:l1 :hasConceptualSource _:Split .  
  _:l1 :hasConceptualTarget ?cf .  
  _:l2 :hasConceptualSource ?cf .  
  _:l2 :hasConceptualTarget _:Merge .  
  ?cl2 :hasConceptualSource _:Merge .  
  ?par :hasSpecification :SnM .  
}
```

```
WHERE {  
  ?cl1 :hasConceptualTarget ?cf .  
  ?cl2 :hasConceptualSource ?cf .  
  ?cf :hasRequirement :SnM .  
}
```

(cowork-url) <http://www.i3s.unice.fr/~cerezo/cowork/latest/cowork.rdfs>

# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

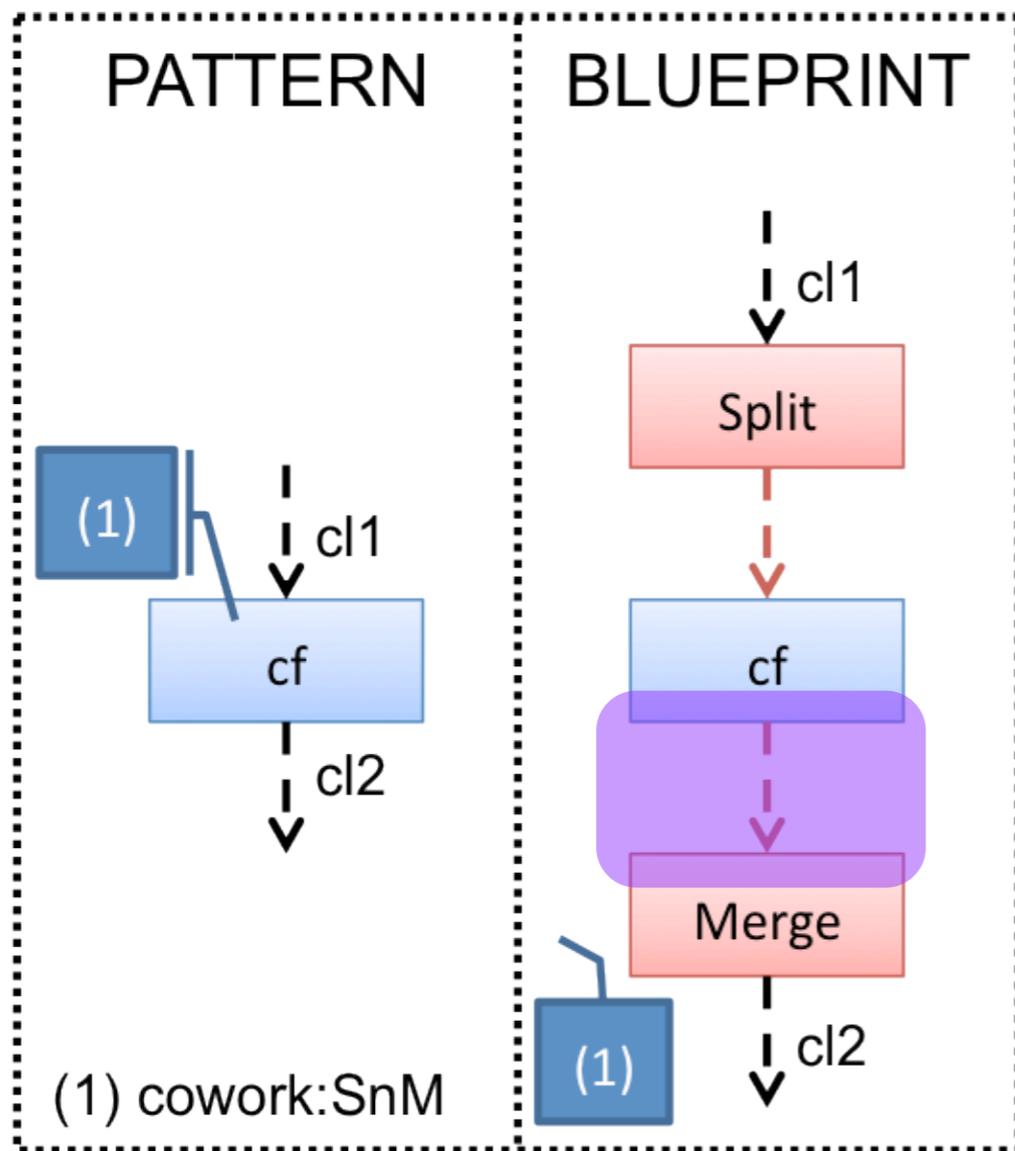
```
PREFIX : <(cowork-url)>
CONSTRUCT {
  ?cl1 :hasConceptualTarget _:Split .
  _:l1 :hasConceptualSource _:Split .
  _:l1 :hasConceptualTarget ?cf .
  _:l2 :hasConceptualSource ?cf .
  _:l2 :hasConceptualTarget _:Merge .
  ?cl2 :hasConceptualSource _:Merge .
  ?par :hasSpecification :SnM .
}

WHERE {
  ?cl1 :hasConceptualTarget ?cf .
  ?cl2 :hasConceptualSource ?cf .
  ?cf :hasRequirement :SnM .
}
```

(cowork-url) <http://www.i3s.unice.fr/~cerezo/cowork/latest/cowork.rdfs>

# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

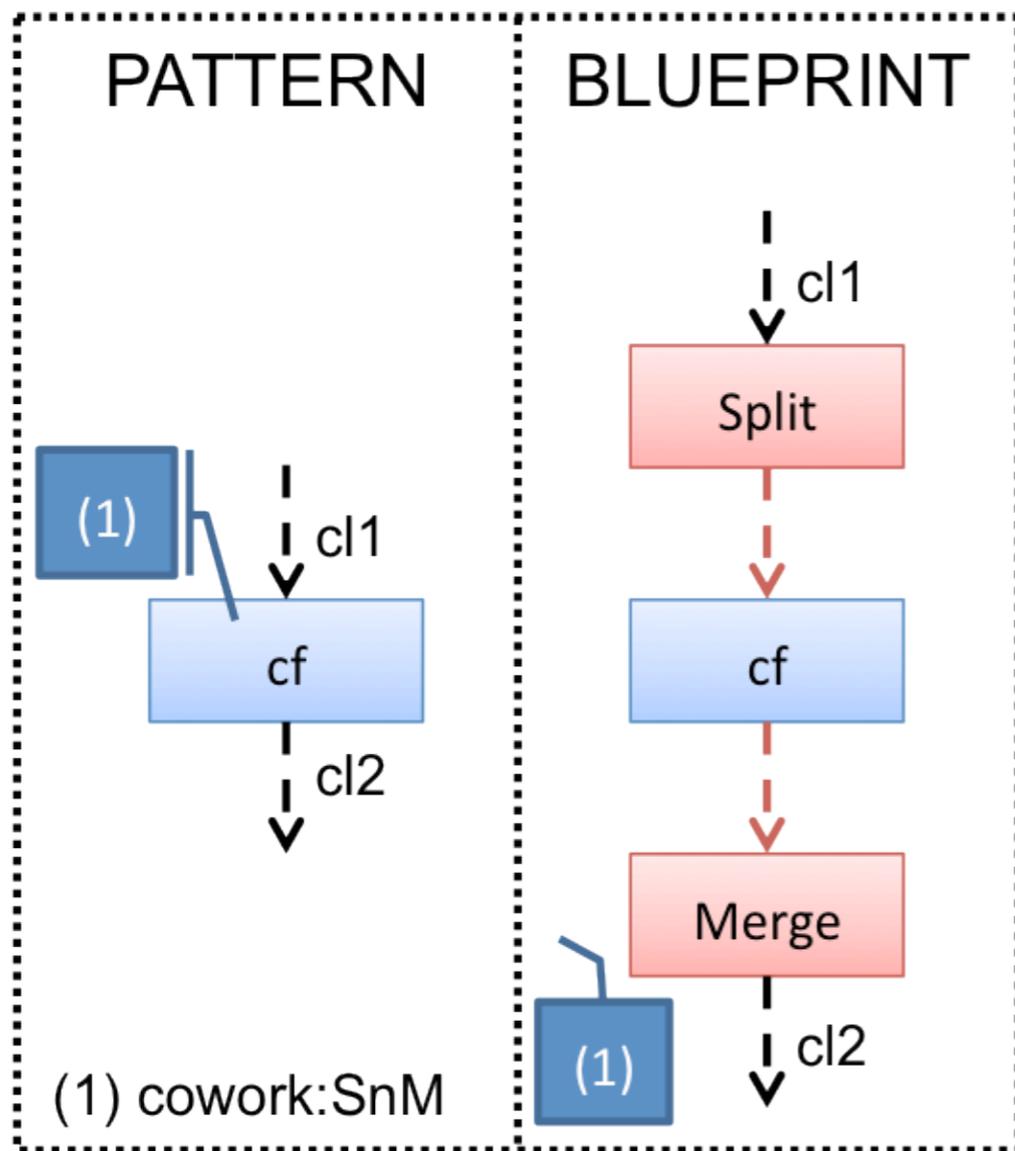
```
PREFIX : <(cowork-url)>
CONSTRUCT {
  ?cl1 :hasConceptualTarget _:Split .
  _:l1 :hasConceptualSource _:Split .
  _:l1 :hasConceptualTarget ?cf .
  _:l2 :hasConceptualSource ?cf .
  _:l2 :hasConceptualTarget _:Merge .
  ?cl2 :hasConceptualSource _:Merge .
  ?par :hasSpecification :SnM .
}

WHERE {
  ?cl1 :hasConceptualTarget ?cf .
  ?cl2 :hasConceptualSource ?cf .
  ?cf :hasRequirement :SnM .
}
```

(cowork-url) <http://www.i3s.unice.fr/~cerezo/cowork/latest/cowork.rdfs>

# Conceptual workflows 2/3: Methods

## Fragment:



## SPARQL query:

```
PREFIX : <(cowork-url)>
CONSTRUCT {
  ?cl1 :hasConceptualTarget _:Split .
  _:l1 :hasConceptualSource _:Split .
  _:l1 :hasConceptualTarget ?cf .
  _:l2 :hasConceptualSource ?cf .
  _:l2 :hasConceptualTarget _:Merge .
  ?cl2 :hasConceptualSource _:Merge .
  ?par :hasSpecification :SnM .
}

WHERE {
  ?cl1 :hasConceptualTarget ?cf .
  ?cl2 :hasConceptualSource ?cf .
  ?cf :hasRequirement :SnM .
}
```

(cowork-url) <http://www.i3s.unice.fr/~cerezo/cowork/latest/cowork.rdfs>

# Conceptual workflows 3/3: Status

---

## **Done**

- Conceptual Workflow model
- Weaving mechanism
- Discovery feature
- Conversion to GWENDIA

## **In development**

- Workflow designer
- Composition assistance
- VIP platform integration

## **Future directions**

- Conversion to other languages (e.g. IWIR, SCUFL)
- Interoperability
- Provenance integration

# Future works (integration)

---

## **Conceptual workflows → provenance**

- Assist the design of provenance-based inference rules
- Enhance the genericity of rules (support WF structure evolutions)
- Exploiting implicit workflow parameters

## **Provenance → conceptual workflows**

- Design new workflows based on produced and annotated data
- Sample annotated data to help in WF design
  - link data to conceptual workflow components
  - suggest annotations at conceptual WF design-time

Thank you !

[cerezo@i3s.unice.fr](mailto:cerezo@i3s.unice.fr)

[alban.gaignard@i3s.unice.fr](mailto:alban.gaignard@i3s.unice.fr)

[johan@i3s.unice.fr](mailto:johan@i3s.unice.fr)