

A middle-grained ontology
should bridge between coarse &
fine-grained structures

[http://esw.w3.org/HCLSIG/SWANSIOC/
Actions/RhetoricalStructure/alignment/
mediumgrain](http://esw.w3.org/HCLSIG/SWANSIOC/Actions/RhetoricalStructure/alignment/mediumgrain)

Jodi Schneider 2011-04-11

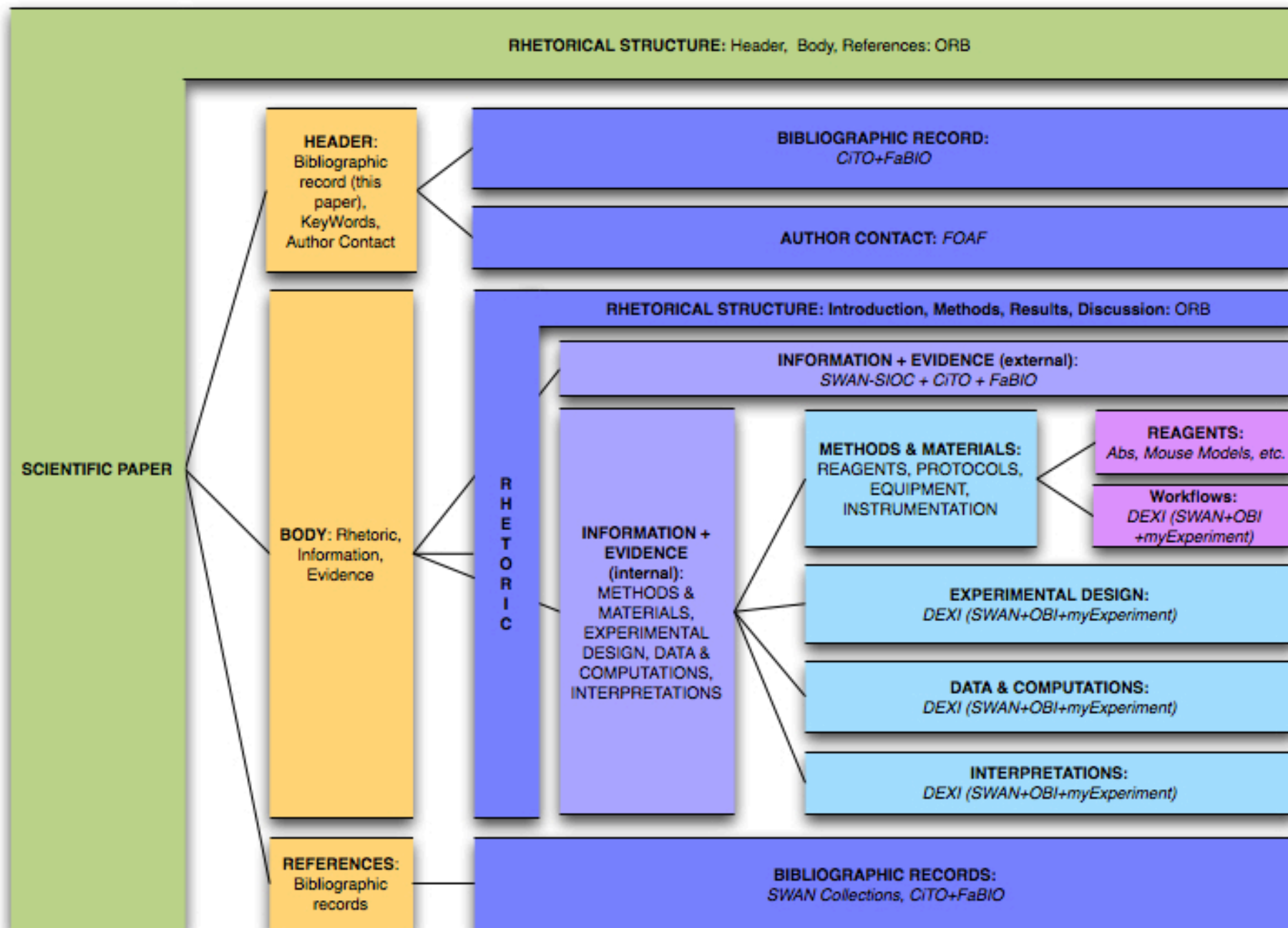
Aligning multiple perspectives

STRUCTURAL

- Block Structure (Entities worth Annotating?)
- Annotation (Relationships between Blocks?)

SEMANTIC FOCUS

- Document
- Process
- Purpose



Slide credit: Tim Clark

Coarse-grained ontology (ORB)

3.1. Header

<i>Class</i>	<i>Definition</i>
<i>orb:Header</i>	The part of the publication that models, or captures, meta-information about the publication, including fields such as title, authors, affiliations, publishing venue or abstract.

3.2. Body

<i>Class</i>	<i>Definition</i>
<i>orb:Introduction</i>	The section describing why was the study in the publication undertaken, what was the tested hypothesis or what was the purpose of the research. It lays down the rationale behind the existence of the publication.
<i>orb:Methods</i>	The section describing when, where, and how was the study done. It includes the materials used as part of the study and who was included in the study groups (patients, etc.).
<i>orb:Results</i>	The part synthesizing the results of the study presented in the paper.
<i>orb:Discussion</i>	The section analyzing whether the tested hypothesis was confirmed. It also interprets the results to understand their consequences and importance. And finally, it shows how the approach and results fit to what other researchers in the field have discovered, including possible perspectives of future research.

3.3. Tail

<i>Class</i>	<i>Definition</i>
<i>orb:Acknowledgments</i>	List pointing to funding bodies or individuals that contributed in a way or the other to enabling or supporting the work presented in the publication.
<i>orb:References</i>	External references to other works (e.g., scientific publications, websites, software) that are relevant for the content of the current publication.

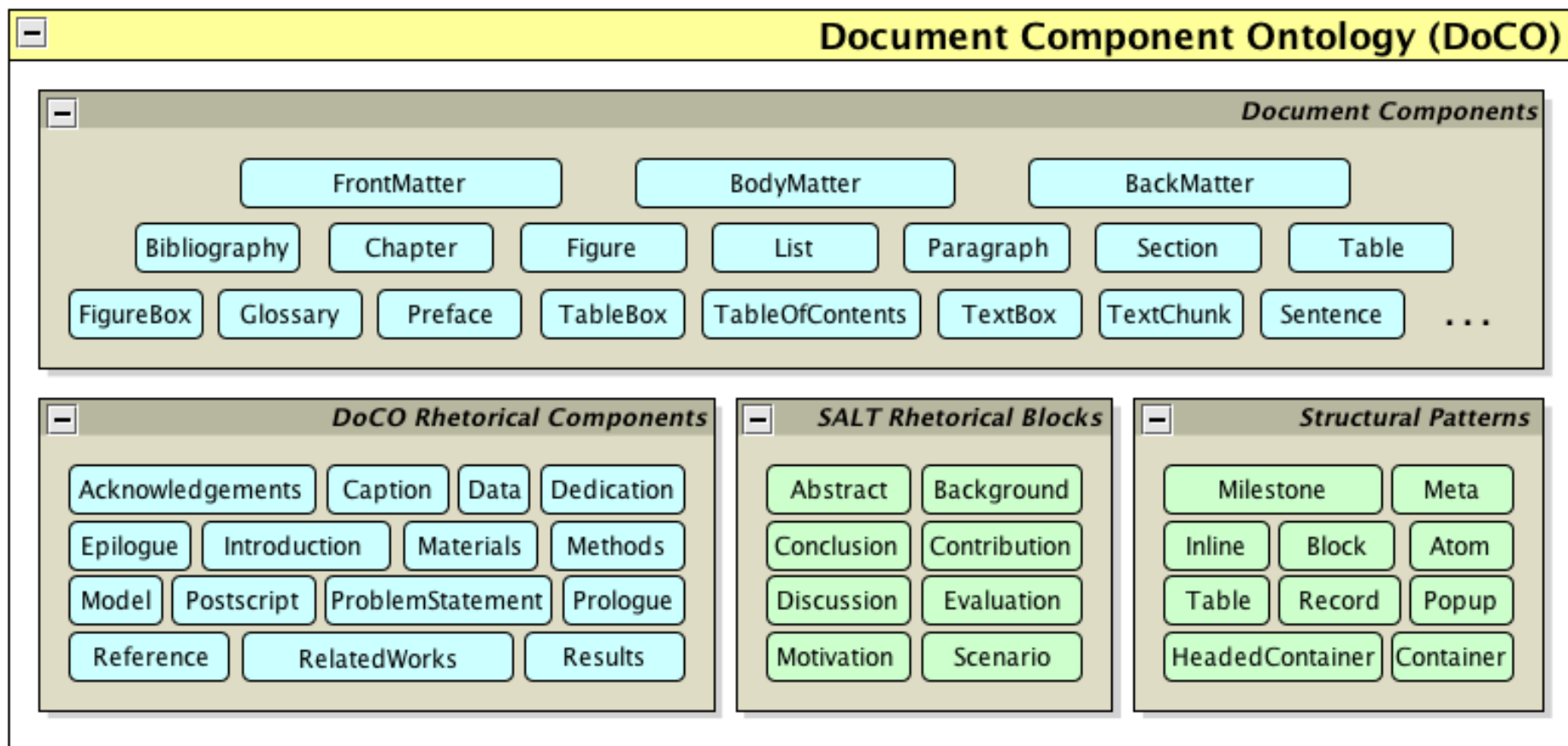
Fine-grained ontologies & approaches

- Document-focused
 - DoCO
 - Anita's block model
- Process-focused
 - DExI
 - myExperiment
- Purpose-focused
 - KeFED

Fine-grained: Document-focused

- Block level
 - Document order
 - Narrative structure (beginning, middle, end)
 - Examples:
 - DoCO
 - Anita's block structure models ("syntagmatic")
- Phrase level
 - Interpretation of claims, hedges
 - Purpose of each clause in the document
 - Examples:
 - Anita: "paradigmatic", Automatic from verb tense
 - Sandor: Automatic detection of novelty from metadiscourse

Rhetorical Blocks



DoCO overview, from
David Shotton
(slightly out-of-date)

Block structure

Medium-grained rhetorical blocks		
Medium-grained (Bio-oriented)		Medium-grained (general)
Block	Sub-block	Block
Introduction	Positioning	Context
	Central problem	Motivation
	Hypothesis	
	Summary of results	-
-	-	Background (Related work)
-	-	Contribution
Experiments	Purpose	Evaluation
	Objects of study	
	Methods	
	Results	Results
	Interpretation	
Discussion	Evaluation of experiments	Discussion
	Comparison with other work	
	Implications	
	Further research	<u>FutureWork</u>

Slide credit: Anita deWaard (Syntagmatic)

Clause structure

Both seminomas and the EC component of nonseminomas share features with ES cells.

Fact

To exclude that

the detection of miR-371-3 merely reflects its expression pattern in ES cells,

Hypothesis

we tested by RPA miR-302a-d, another ES cells-specific miRNA cluster (Suh et al, 2004).

Method

In many of the miR-371-3 expressing seminomas and nonseminomas, miR-302a-d was undetectable (Figs S7 and S8),

Result

suggesting that

miR-371-3 expression is a selective event during tumorigenesis.

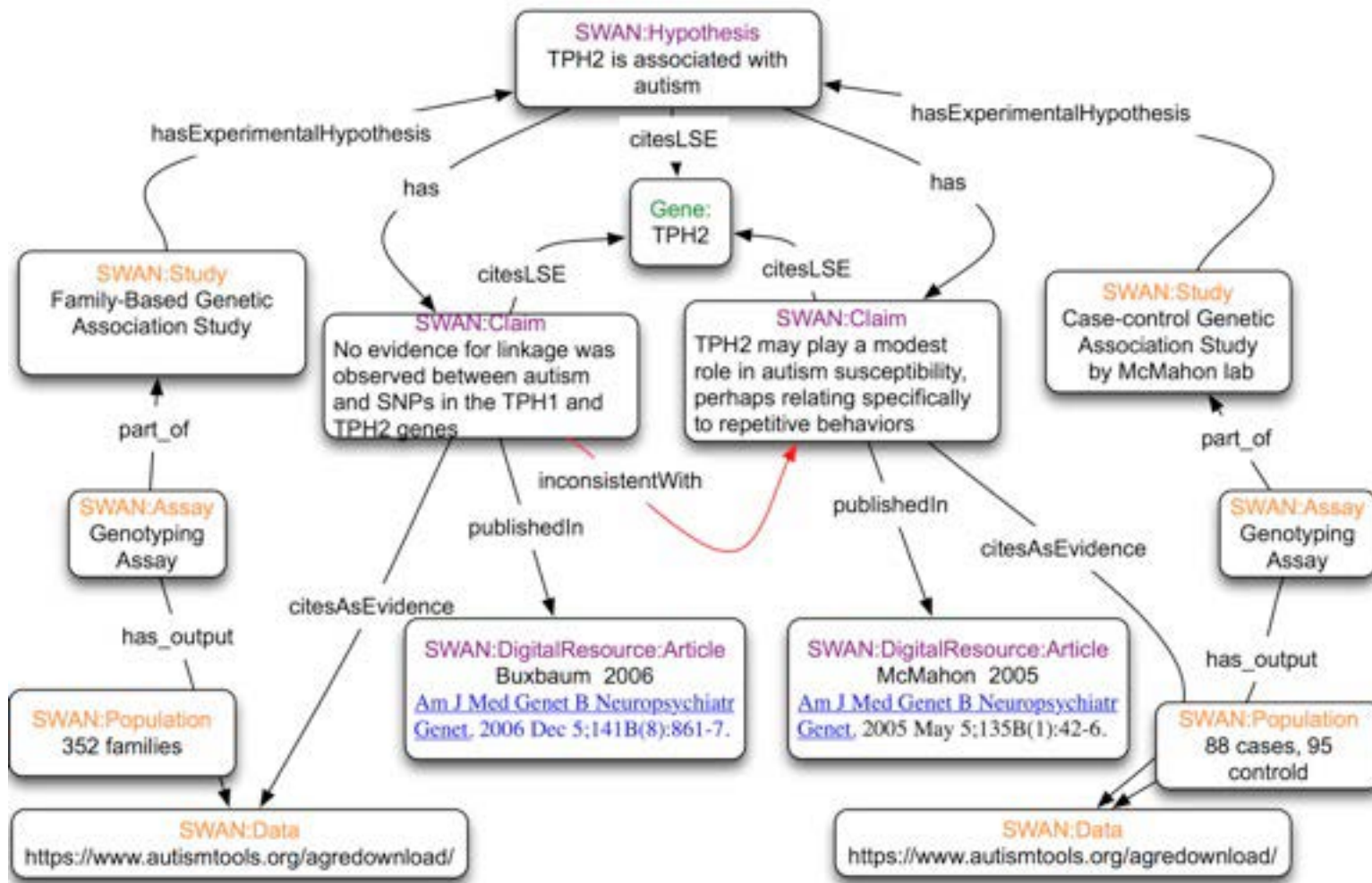
Implication

Slide credit: Anita deWaard,

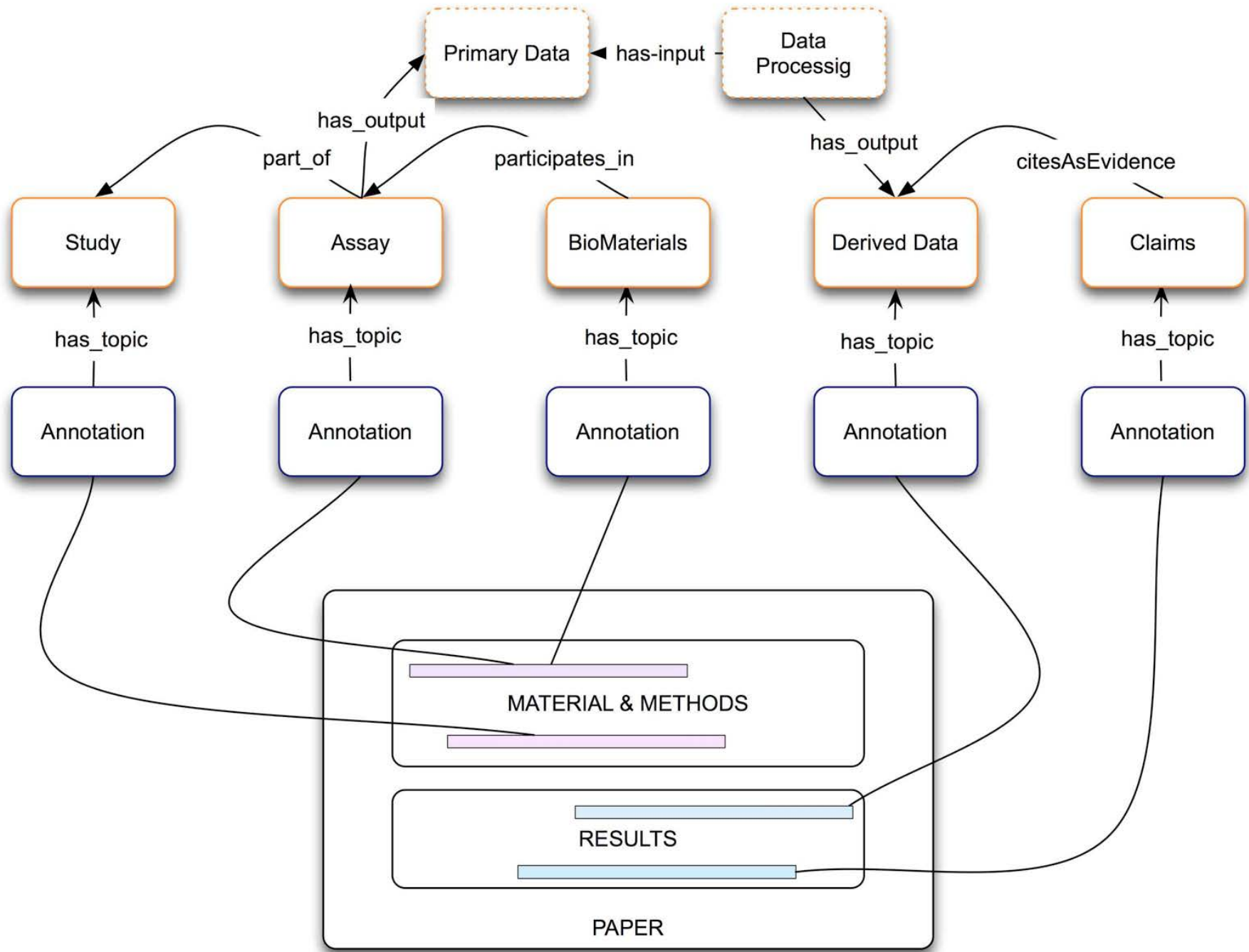
Paradigmatic: verb tense & sentence structure analysis

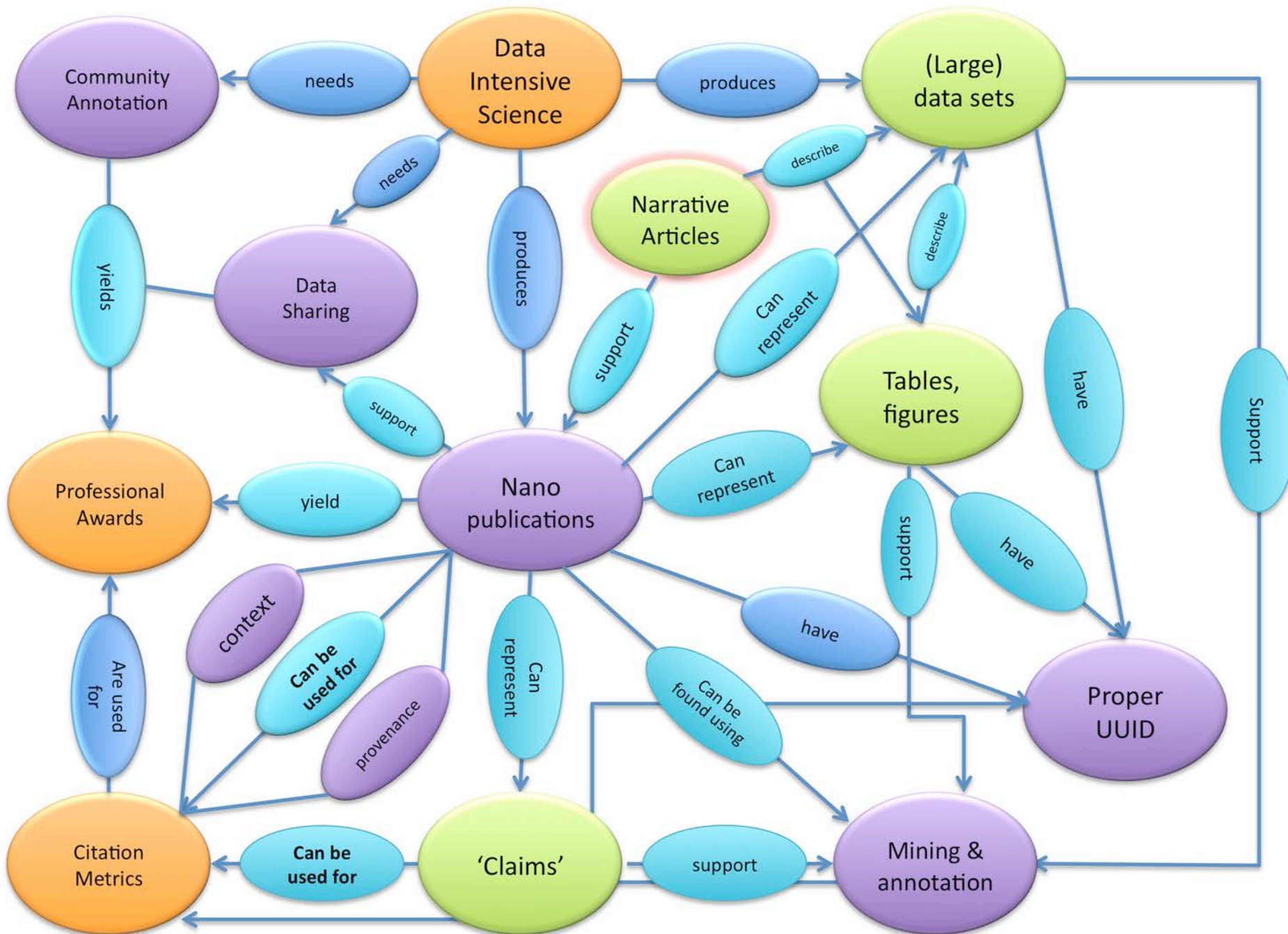
Fine-grained: Process-focused

- “Doing science” perspective
 - Focus on methods, materials, protocols, equipment
 - Examples:
 - DExI
 - myExperiment
 - Focus on domain-level overview
 - e.g. Barend Mons’ view of nanopublications



SWAN/DexI





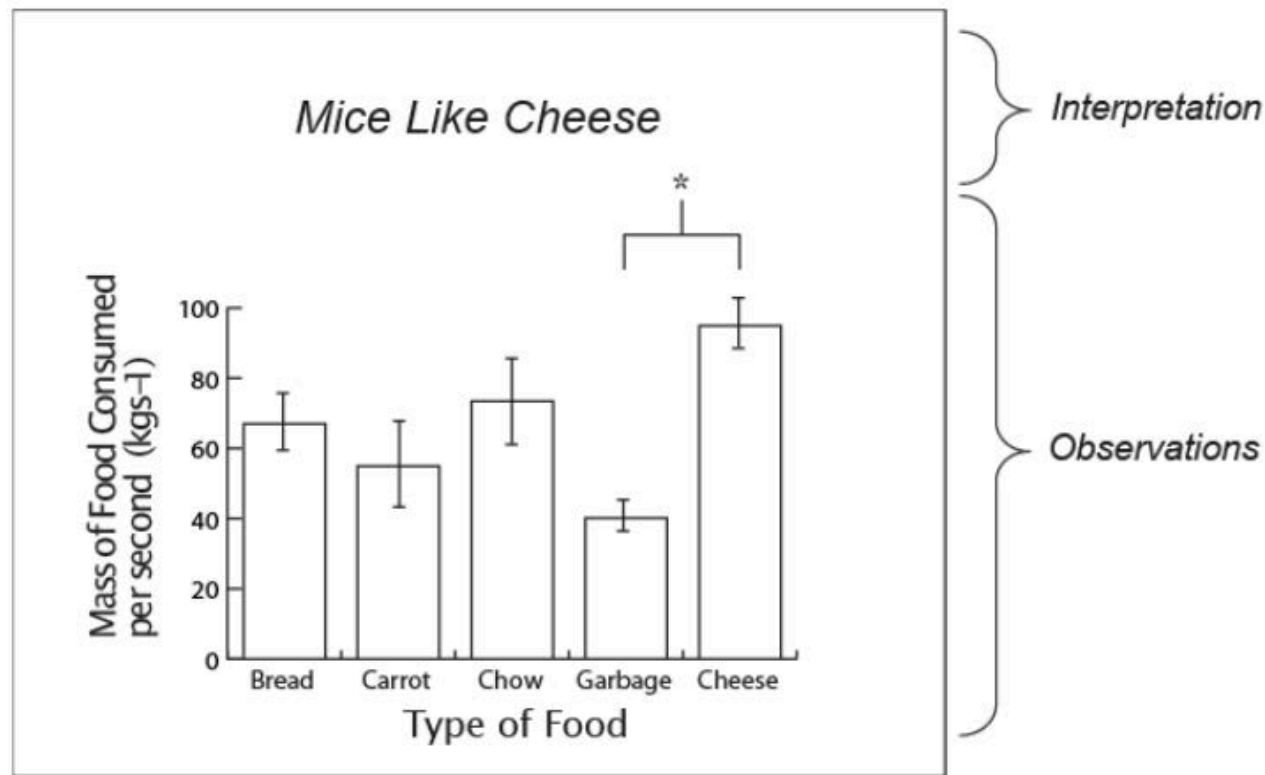
Slide credit: Barend Mons on Nanopublications

Fine-grained: purpose-focused

- Motivational perspective
 - WHY
 - Connects hypothesis, design, interpretation
 - Examples:
 - KeFED

KEFeD connects interpretation & observation

For example...



Questions

- HOW to connect the document discourse to the process of science
- What fine-grained model(s) are we aiming at?
- WHAT are we modeling?
 - the document?
 - the discourse?
 - the process?
 - the underlying reasoning?