

Underspecified Scientific Claims in Nanopublications

Tobias Kuhn and Michael Krauthammer

Department of Pathology
Yale University School of Medicine

Workshop on Web of Linked Entities (WoLE) at ISWC
Boston

11 November 2012

Introduction

Nanopublications are **minimal portions of scientific results** in RDF format.

Nanopublications have been proposed to ...

- ... make it easier to **find, connect and curate** core scientific statements
- ... determine the **attribution, quality and provenance** of scientific statements

See <http://nanopub.org> for more information on nanopublications

Vision: Changing Scholarly Communication

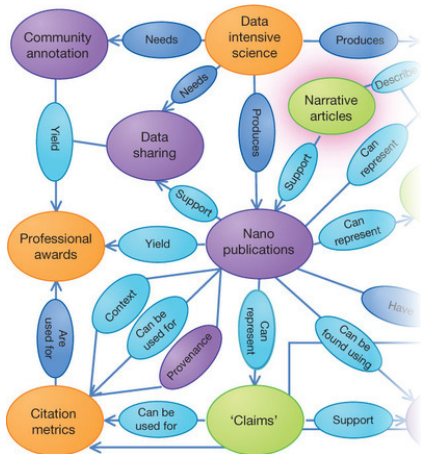
Now

Narrative articles at the center



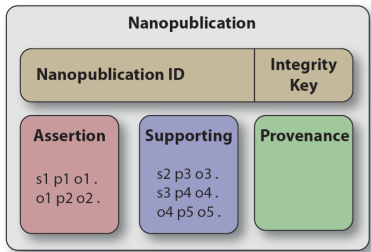
Future

Nanopublications at the center



Images from Mons et al. The value of data. *Nature genetics*, 43(4):281–283, 2011

Broadening the Scope

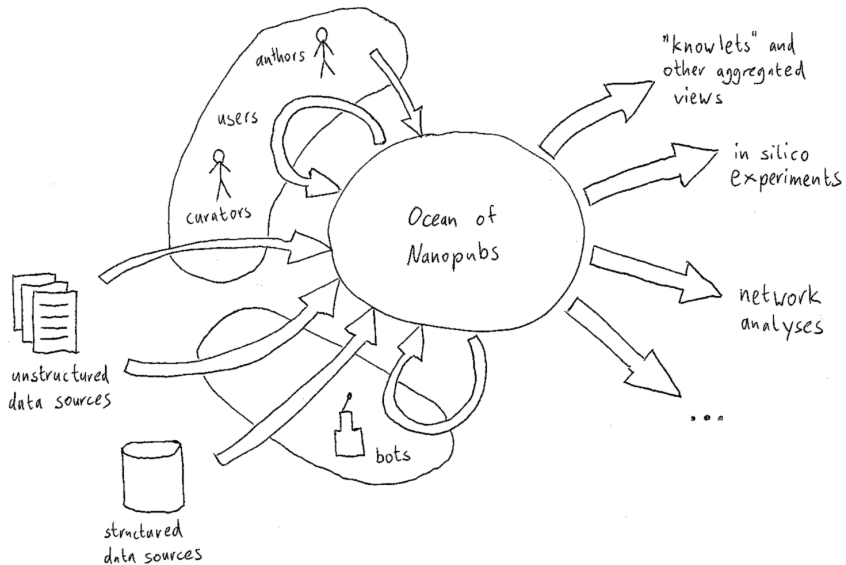


Assertions

- “Classical nanopublications”: Scientific claims in formal notation, e.g. *transmits(mosquito, malaria)*
- **Extension 1:** Support for informal and underspecified claims, e.g. “Malaria is transmitted by mosquitoes.”
- **Extension 2:** Include non-scientific assertions, e.g. *disagrees(tkuhn, “Malaria is transmitted by mosquitoes.”)*

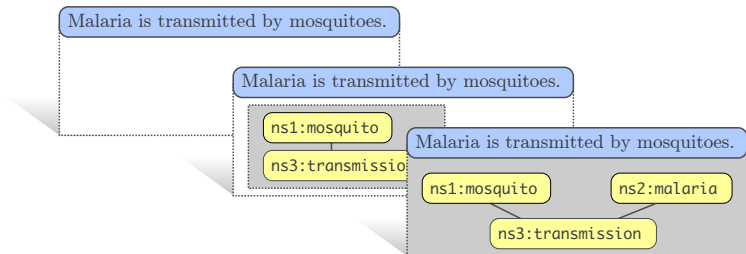
Image from The Open PHACTS Nanopublication Guidelines, V1.81, 2012

Ocean of Nanopublications



Levels of Formalization

- Informal (only an AIDA sentence)
- Underspecified (formal representation for part of the sentence)
- Fully formal (formal representation for the complete sentence)



AIDA Sentences

The sentences of our approach should be **AIDA**:

- **Atomic**: a sentence describing one thought that cannot be further broken down (in a practical way)
- **Independent**: a sentence that can stand on its own, without external references like “this effect” or “we”
- **Declarative**: a complete sentence ending with a full stop that could (in theory) be either true or false
- **Absolute**: a sentence describing the core of a claim ignoring the (un)certainty about its truth (no “may” or “probably”) and ignoring how it was discovered (no “evaluation showed that”)

Example: “The majority of patients with idiopathic REM sleep behavior disorder who develop a neurodegenerative disease develop Parkinson disease and Lewy body dementia.”

But wait a moment...

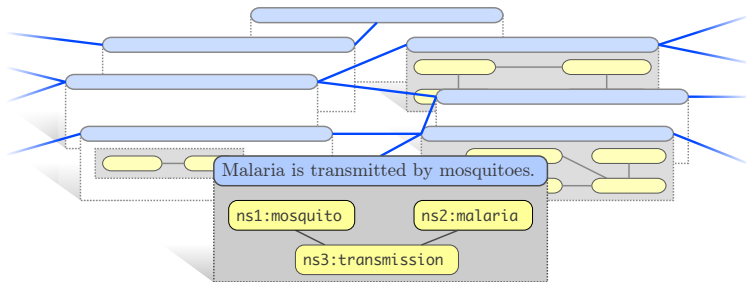
... isn't that just "returning back to the Web as it was before RDF, OWL and Linked Data"? (*anonymous reviewer*)

That is a very good question, and the answer is **no**.

AIDA sentences can be **interlinked and related** to other entities and become part of the formal Semantic Web.

We can be **formal on a higher level** even if we are forced to be **informal on the lower level!**

Linking Scientific Claims




Possible relations:

- [CLAIM] is equivalent to / contradicts / is similar to [CLAIM]
- [PERSON] agrees with / disagrees with / challenges [CLAIM]
- [STUDY] provides (counter-)evidence for [CLAIM]


These relations can be published as nanopublications too!

Nanobrowser: “Scientific/Social/Distributed/Semantic Wiki”



nanobrowser



 **Malaria is transmitted by mosquitoes.**
[http://krauthammerlab.med.yale.edu/nanopub/claims/en/Malaria+is+transmitted+by+mosquitoes.](http://krauthammerlab.med.yale.edu/nanopub/claims/en/Malaria+is+transmitted+by+mosquitoes)

Nanopublications



 MalariaPub 2012-10-12 14:37

Opinions

 Tobias Kuhn **agrees.** [

 somebody **disagrees.** [

Same Meaning

 Mosquitoes transmit malaria. [

Evaluation (ongoing...)

There are a number of issues that need to be evaluated:

- What amount of scientific results that can be represented as AIDA-style nanopublications but not as classical ones?
- How easy or hard is it for authors/curators to write AIDA sentences for scientific results?
- Can we automatically extract AIDA-style nanopublications from existing articles at a satisfactory level of quality?
- Can we automatically cluster similar AIDA sentences?
- Do AIDA sentences make it easier to collaborate on scientific results?
- ...

Thank you for your Attention!

Questions?