

Semantic Publishing: Project Ideas

Below are some project ideas by [Tobias Kuhn](#) and [Michel Dumontier](#) for a future project proposal.

Motivation

Software has boosted scientific productivity, but we cannot use software to reliably interpret or aggregate the resulting scientific findings, because they are published in narrative articles optimized for humans and not for algorithms. This project will make a first step to solve this problem by using recent technologies to build a flexible framework for intuitive web interfaces for scientists to directly publish their findings in a formal computer-interpretable format.

Project Description

This project will use the nanopublication technology to develop and test a framework for enabling quick and easy publishing of formal snippets of scientific results, which we call "nanoblogging", and will apply and evaluate it for two chosen use cases.

With [genuine semantic publishing](#) as outlined in our position paper and unlike most existing approaches on "semantic publishing", formal semantic notations are used to express scientific findings (and not just metadata), which are published by the researchers themselves (and not by curators or text mining pipelines) as primary and self-contained publication units (and not just as annotations). We can then not only automatically interpret and aggregate scientific findings (due to their formal semantics) but also accelerate the dissemination of results (because we don't need to wait for a narrative paper to be published) at highest quality (because they are written by the researchers themselves).

[Nanopublications](#) are a recent concept based on semantic technologies that provide provenance-aware Linked Data containers for small fragments of scientific results. In previous work, we have developed the technology for [strong identifiers](#), reliable [publishing and archiving](#), and flexible [dataset and version management](#) for such nanopublications, which have since been successfully applied to publish [several datasets in the Life Sciences](#).

The framework to be developed will be applied to implement nanoblogging interfaces for two specific use cases: gene-disease relations and biotic interactions. For the first of these, the final web interface could look like this:

Nanoblogging: Gene-Disease Relations

Gene: ... (NCBI Gene ID) Relation: associated with (sio) Disease: ... (UMLS ID)

Source: own research literature meta analysis other

Type: unspecified in vivo in vitro in silico [more...](#)

Publication Status (can be changed later): no paper (yet) pre-print published paper

URL/DOI: ...

Logged in as Tobias Kuhn (orcid.org/0000-0002-1267-0234). Author of this nanoblog: same other [Publish](#)

This interface will allow researchers to directly publish their findings (about gene-disease relations in this case) as formal Linked Data, without requiring any technical skills, by just selecting items from auto-completed drop-down menus and filling in a few menu elements (and login via [ORCID](#)).

First Year Plan

The first year of the project can consist of the following phases: **Months 1-3:** Development of a meta-model to integrate the needed domain and metadata aspects of the use cases, building upon [SIO](#), [OBO](#), [ECO](#), [PROV](#), [Micropublications](#), [ORCA](#), [EXPO](#), and others; **Months 4-6:** Implementation of the framework with the first use case, *gene-disease relations*, with support from Laura Furlong (leader of the [DisGeNET](#) database at Universitat Pompeu Fabra, Spain); **Months 7-9:** Integration of feedback from the first use case and application of the framework to the second use case, *biotic interactions*, with support from Jorrit Poelen (leader of the [Global Biotic Interactions](#) initiative); **Months 10-12:** Evaluation in the form of a field experiment, asking researcher in the given fields to publish some of their findings with our nanoblogging interfaces, manually assessing the quality of these nanoblogs, and asking about the user experience in a questionnaire.