Guidelines to annotate experimental protocols by using Nanotate tool

SCOPE

We are manually annotating experimental protocols in life sciences. We focus on the annotation of protocol steps and materials participating in each one of them, like i) the Sample(s) tested, ii) Equipment and iii) Reagents. This document is about how to install and use the Nanotate tool; give examples about what is a sample, equipment, and reagents and how this information should be annotated by using our tool.

STARTING TO USE NANOTATE TOOL

Before You Begin:

- Use Chrome, Firefox or Safari browser.
- Create and be logged into your Hypothes.is account.

1) Go to the homepage of Nanotate tool (see fig.1): https://nanotate.bitsfetch.com/

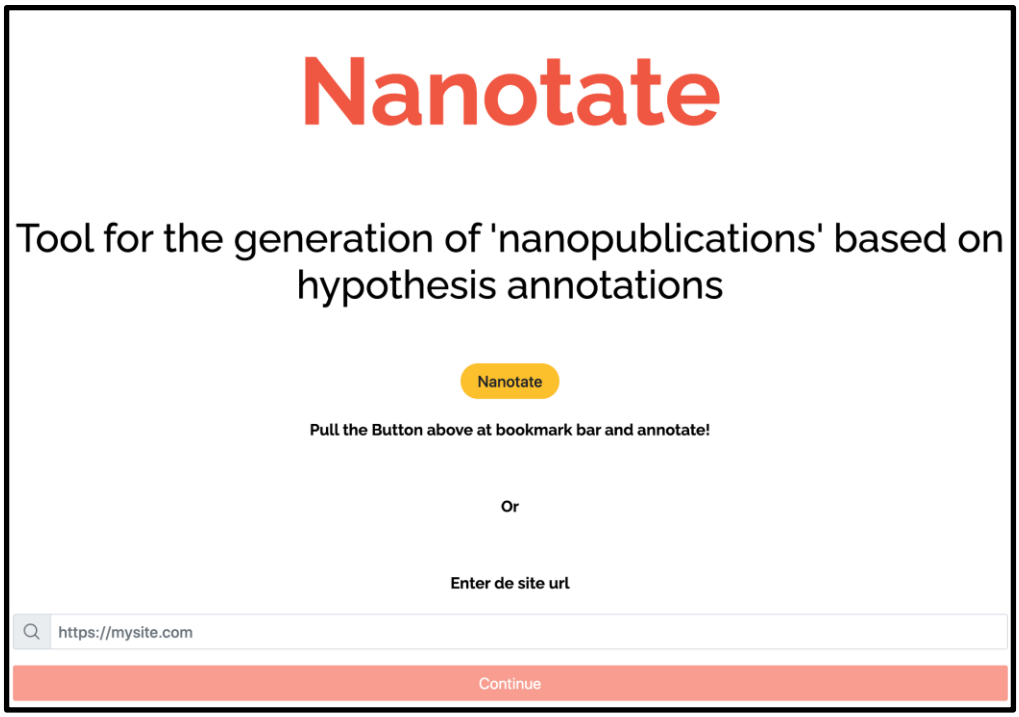

Figure 1

2) From the Chrome browser, drag the yellow button “Nanotate” to the bookmark bar to start to annotate.

3) After that, choose an open access protocol that you want to annotate. (An example is available here: DOI: 10.21769/BioProtoc.47)

4) To start to annotate, first select the text and then click on the Nanotate bookmark. See Fig. 2.

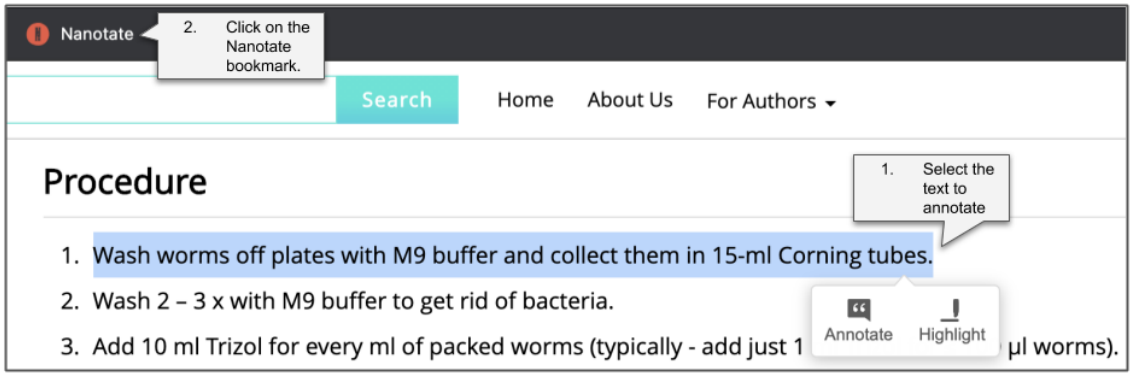

Figure 2

5) The Nanotate UI will redirect you to a page where you need to insert your hypothesis credentials. See Fig. 3.
    a) The Hypothesis username and
    b) Token.

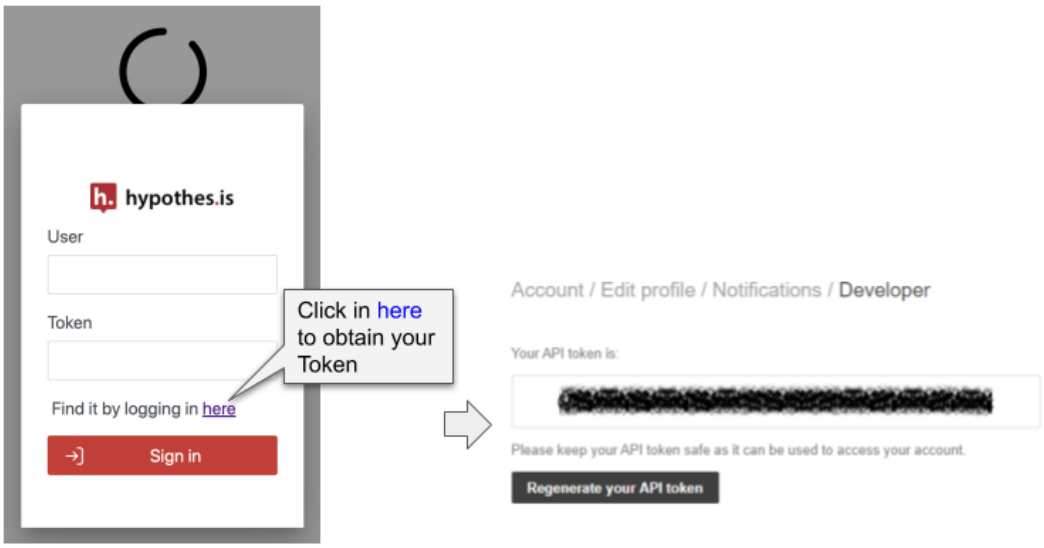

Figure 3

6) After that, you can start the annotation process!

WHAT SHOULD BE ANNOTATED?

Words or phrases related to:

- Sample(s), specimen(s), or organism(s) to be tested.
- Equipment, instruments, and consumables.
- Reagents, chemical compounds, solutions, or mixtures used.
- Steps.

Some Examples

The sample tested in a protocol may be an organism or a part of it. Some examples include:

SAMPLE	Whole organism	Scientific name*: Arabidopsis thaliana*, *Oriza sativa*, *mangifera indica, Mus musculus.* Common name: Mouse Ear Cress, rice, mango, mouse.
	Anatomical part	leaf, stem, cells, tissues, membranes, organs, skeletal system, muscular system, nervous system, reproductive system, cardiovascular system, etc.
	Biomolecules	Nucleic acids: Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Proteins: enzymes, structural or support proteins (keratin, elastin, collagen), antibodies, hormones, etc.
	Body fluids	Blood serum, saliva, semen, amniotic fluid, cerebrospinal fluid, gastric acid, etc.

The equipment used in protocols include high-throughput equipment, instruments and consumables. Some examples:

EQUIPMENT	High-throughput equipment	Liquid Handling Platforms, Real-Time PCR Detection System, Microplate Reader, etc.
	Instruments	Goggles, Bunsen burner, spot plate, pipet, forceps, test tube rack, mortar and pestle, etc.
	Laboratory glassware	Beaker, Erlenmeyer flask, graduated cylinder, volumetric flask, etc.
	Standard equipment	Balances, shakers, centrifuges, refrigerators, incubators, thermocyclers, fume hood, etc.
	Consumables	Weighing dishes, pipette tips, gloves, syringes, petri dishes, test tubes, micro centrifuge tubes, glass slides, filter paper, etc.

The reagents used in protocols include buffers, solutions and culture media. Some examples:

REAGENTS	Chemical compound/ Substance	Glucose, ethanol, glycerol, chloroform, acetic acid, isopropyl alcohol, etc.
	Solutions / buffers	70% ethanol, 10X PCR buffer, phenol:chloroform:isoamyl alcohol, etc.
	Cell culture media	Nutrient media, minimal media, selective media, differential media, etc.

The step of a protocol is a formal statement describing a task that should be done or operated. An example:

Part of the speech	**Source**
“Wash worms off plates with M9 buffer and collect them in 15-ml Corning tubes.”	doi: 10.21769/BioProtoc.47

HOW TO ANNOTATE

The annotators should focus on annotating the “procedure section” where the steps are listed in numerical order.

The annotation UI includes:

- tag box: add one or multiple tags from the available options (sample, equipment, reagent, software, input, Output, Step).

- **Note**: this is the only required field to publish an annotation. If the selected text does not include any tag the following message will be shown (see fig. 5).

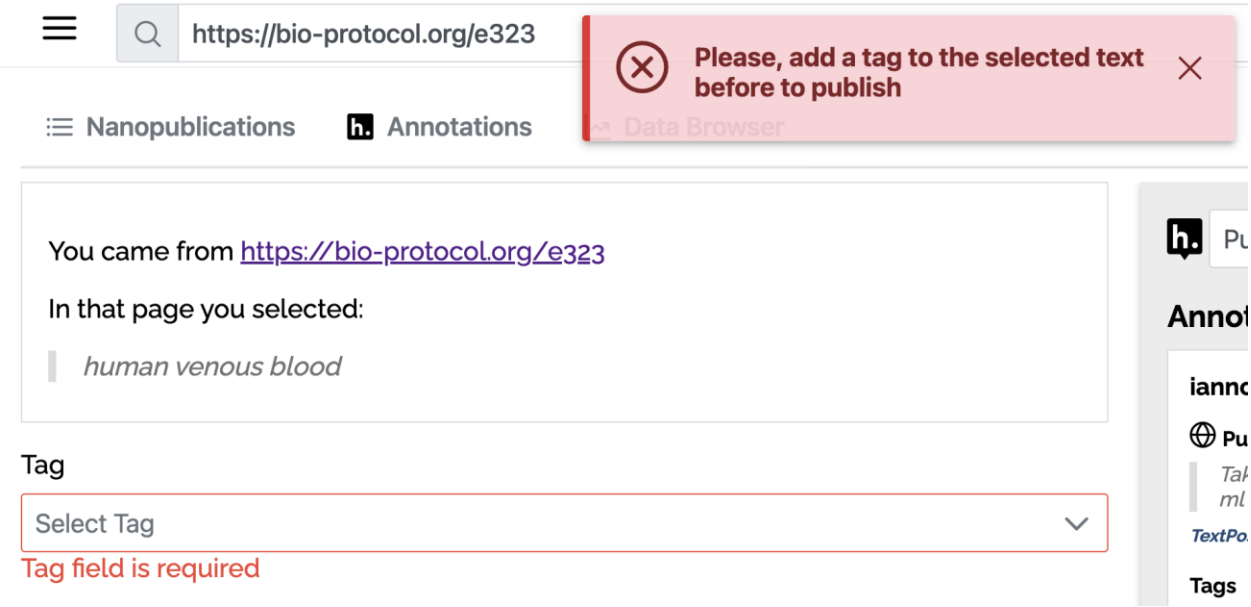

Figure 5

- ontology box: where annotators can link the annotated text with classes from 8 ontologies available in Bioportal.
- comment box: where annotators can add any comment indicating when an annotation or decision was hard to make
- publish button: press this button to publish the annotation.
- annotation panel. In this panel the annotators can visualize the annotations done. The annotations can be edited or deleted at any moment.
    - **Note**: If the annotation is not displayed automatically, then you click on the reload button to refresh the page.

Annotation of steps

Each step should be selected and annotated individually.

- **Note:** Please, avoid selecting the entire list of steps for adding a unique tag; the database of annotations will consider that the protocol “x” has only one Step entry.

- **Note:** A step is a message describing how something should be done or operated. The steps are not linked to ontology terms.

A video about how to annotate steps is available at: https://youtu.be/ZONgbha5AIA

Annotation of samples

The annotators should annotate only the sample(s) tested in the protocol.

Tagging samples: In steps where a sample is mentioned, this could play an “input or “output” role. An example is presented in Fig. 6. There, steps 1 and 2, the samples highlighted in pink are also the input in their respective steps. In both cases the selected text should be tagged as “sample” and as “input”. In cases where the sample plays an output role, the selected text should be tagged as “sample” and as “output” (See Fig.6, step 16).

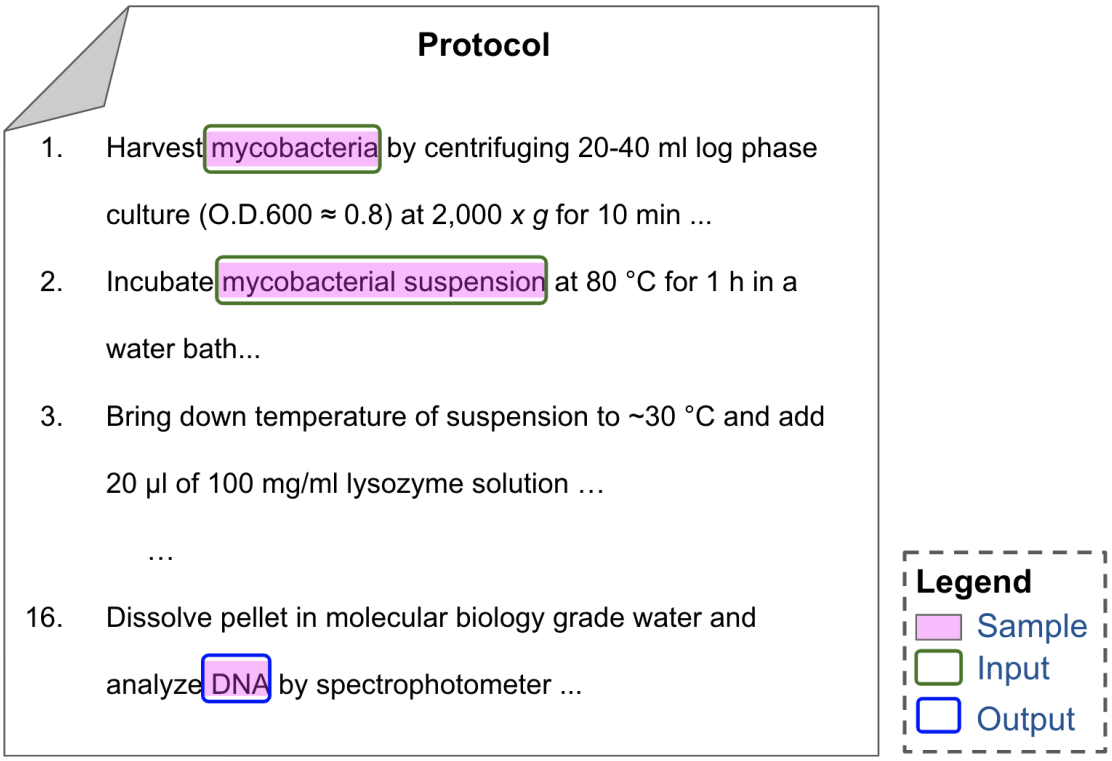

Figure 6

Mapping samples to ontology terms: When possible, the samples should be linked to ontology terms that come from NCBI Taxonomy or UBERON; see video available at: https://youtu.be/9hEJu8YCCto

Otherwise, please, choose an ontology term from another ontology suggested by our system; see video available at: https://youtu.be/05G7ZuLMP2U

Annotation of equipment

all equipment, instruments or consumables used in steps should be annotated. In steps sometimes is mentioned the capacity of the containers used. When this information is available, the annotation should include the storage capacity of the containers used and described in the protocol. Some examples are:

Part of the speech	Source
“Wash worms off plates with M9 buffer and collect them in 15-ml Corning tubes.”	doi: 10.21769/BioProtoc.47
“The supernatant from step 7 should be clear. Carefully transfer the supernatant to a 12 x 75 mm polypropylene culture tube.”	doi: 10.21769/BioProtoc.213
“Transfer the sample into a 2.0 ml beat beater tube.”	doi: 10.21769/BioProtoc.3634
“While DNA is digesting, make your 0.35(w/v) agarose gel mix: 400 ml 1x TBE + 1.4 g Agarose in a 1 L glass bottle.”	doi: 10.21769/BioProtoc.213

Mapping equipment to ontology terms: When possible, the equipment should be linked to ontology terms that come from OBI (first option), BAO, EFO or SP; see video available at: https://youtu.be/QyWbAavGEsQ

Annotation of reagents

In steps are used chemical, biochemical compounds and solutions (e.g., buffers, Cell culture media). Sometimes, chemical compounds are used as a solute in a solution. Please annotate the different concentration grades of the solutions or mixtures (e.g. 70% alcohol, 100mg/ml lysozyme solution, 1x TE buffer, 3M sodium acetate)

- **Note**: please do not annotate the quantity used for reagent. some examples:

    - 20 g of peptone `instead of` 20 g of peptone
    - 1 liter of dH2O `instead of` 1 liter of dH2O

Mapping reagents to ontology terms: When possible, the reagents should be linked to ontology terms that come from CHEBI (first option), ERO (second option) and SP (third option); see scenario 3. otherwise, please, choose an ontology term from another ontology suggested by our system; see video available at: https://youtu.be/N28yaYmM068