

LOTUS Discussions

March 16, 2023

Participants

Donat, Patrick, Adriano, Pierre-Marie, Emmanuel

Agenda

To do

- [x] AR: Upload a list of triples we would like to extract from PMC articles with identifiers (Done 2023-03-16, see: <https://github.com/orgs/digital-botanical-gardens-initiative/discussions/12>)

March 2, 2023

Participants:

Pierre-Marie Allard, Patrick Ruch, Donat Agosti, Alexandre Flament, Adriano Rutz

=====

NB: The project is using an OpenNotebook Science
(https://en.wikipedia.org/wiki/Open-notebook_science)

[Project description](#)

Agenda:

- I. project presentation,
- II. Todo
- III. [...]

- The [Digital Botanical Gardens Initiative](#)
- From earth metabolome to botanic species
- Mass spec is used to characterize fingerprints with localization
- All exposed with a RDF endpoints
- [LOTUS](#) provides the link between (natural) small molecules and their occurrence
- All data in LOTUS is also in Wikidata
- Global Metabolome Initiative

Challenge: Create a link between the treatments in TB, the material citation and the small molecules

Gap: some knowledge is still found only in PDF, here comes the TNA project !

Possible extension of the first tasks: A possible extension would be to move towards nano-publications

Key: the level of evidence (silver level of evidence → Evidence and Code Ontology [ECO](#)) for the inferred statements is relatively weak as it uses mass spectrometry

annotate articles in SIBiLS using an extended SIBiLS ontology

how to share the extracted knowledge:

- talking nonopubs

get a list of journals (the most important ones)

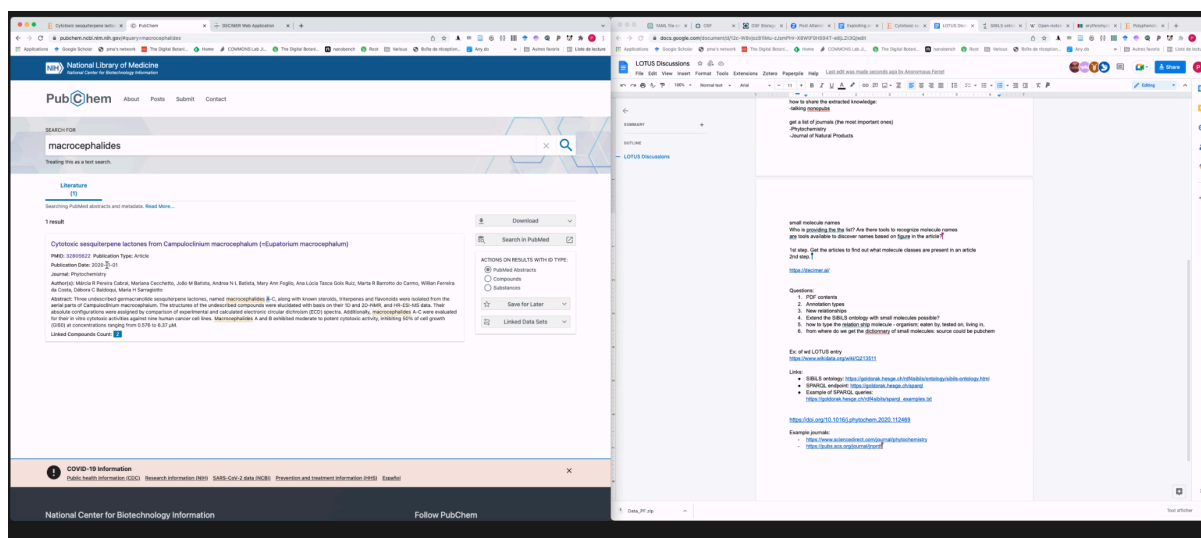
- Phytochemistry
- Journal of Natural Products

small molecule names

Who is providing the the list? Are there tools to recognize molecule names
are tools available to discover names based on figure in the article?

1st step. Get the articles to find out what molecule classes are present in an article

2nd step.



<https://decimer.ai/>

get a

16	19.8	CH3	1.23	d (7.1)	C-6, C-15, C-17
17	19.8	CH3	1.23	d (7.1)	C-6, C-15, C-16

a Chemical shifts of ^{13}C were determined by **HSQC** and HMBC. b 600 MHz. c Overlapping signal.

<<
<
1
>
>>

get a vocabulary that includes all the methods, such as also HMBC

Questions:

1. PDF contents
2. Annotation types
3. New relationships
4. Extend the SiBiLS ontology with small molecules possible?
5. how to type the relation ship molecule - organism; eaten by, tested on, living in,
6. from where do we get the dictionary of small molecules: source could be pubchem
7. can the provenance of users be monitored to give him project related access

Ex: of wd LOTUS entry

<https://www.wikidata.org/wiki/Q213511>

Links:

- SIBiLS ontology: <https://goldorak.hesge.ch/rdf4sibils/ontology/sibils-ontology.html>
- SPARQL endpoint: <https://goldorak.hesge.ch/sparql>
- Example of SPARQL queries:
https://goldorak.hesge.ch/rdf4sibils/sparql_examples.txt

<https://doi.org/10.1016/j.phytochem.2020.112469>

Example journals:

- <https://www.sciencedirect.com/journal/phytochemistry>
- <https://pubs.acs.org/journal/jnprdf>

PubChem RDF for compounds dictionary:

<https://ftp.ncbi.nlm.nih.gov/pubchem/RDF/>

To Do

PM & Adriano provide a list of the articles

PM & Adriano provide ca 20 articles where the elements are highlighted that need to be extracted

DA get Journal of Natural Products and Phytochemistry articles

DA talk to GS/CN to scrape

PM & Adriano defines the annotation schema: molecules, taxon, predicate, methods, specimen, organs ([PO](#)), habitats/location

PM & Adriano provides a list of vocabularies/ontologies for each entities

PM & Adriano provide annotated articles

PR/AF check how to hide access to closed access into SIBiLS

Next meeting /

recurring meeting March 16, 2023, 2 - 3 pm