# The Digital Botanical Gardens Initiative

This manuscript (<u>permalink</u>) was automatically generated from <u>digital-botanical-gardens-initiative/dbgi-green-paper@8cdea70</u> on November 7, 2022.

## **Authors**

- - · 🕥 <u>digital-botanical-gardens-initiative</u>

■ — Correspondence possible via <u>GitHub Issues</u> or email to The Digital Botanical Gardens Initiative Consortium <dbgi@protonmail.com>.

### **Abstract**

The Digital Botanical Gardens Initiative (DBGI) ambitions to explore innovative solutions for the collection, management and sharing of digital information acquired on living botanical collections. A particular focus will be placed on the large scale characterization of the chemodiversity of living plants collections through mass spectrometric approaches. The acquired data will be structured, organized and connected with relevant metadata through semantic web technology. The gathered knowledge will then inform ecosystem functioning research and orient biodiversity conservation projects. The DBGI initially aims to take advantage of the readily available living collections of Swiss botanical gardens to establish robust and scalable chemo- and biodiversity digitisation workflows. The ultimate goal is to apply these approaches in the field and at the global scale in wild ecosystems.

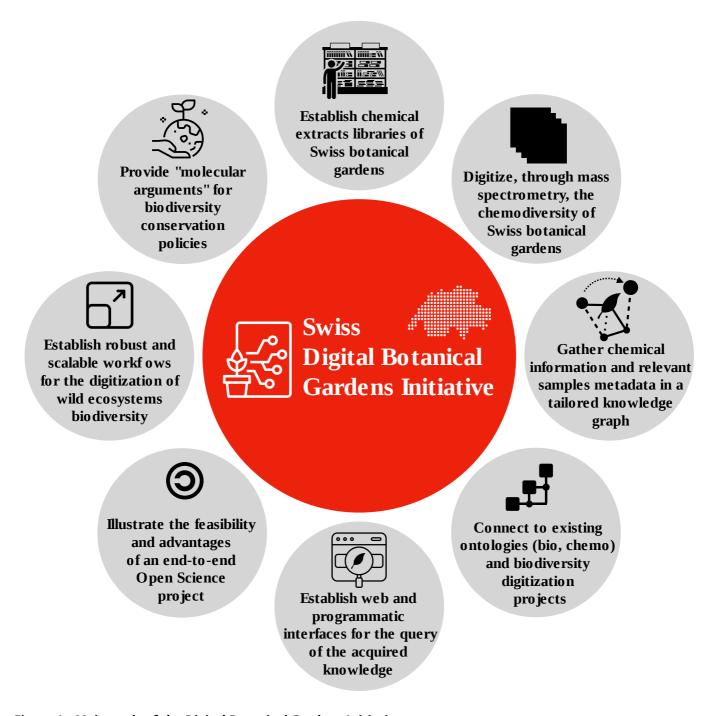


Figure 1: Main goals of the Digital Botanical Gardens Initiative.

The main goals of the DBGI are resumed in Figure 1 Some details are given hereafter:

- 1. Establish chemical extracts libraries of Swiss botanical gardens. These chemical libraries can be considered as complementary to herbarium samples. They are easily conserved over time and in a reasonable space. They represent the chemical diversity of a sample. They can be easily aliquoted. They can be screened for bioassays.
- 2. Digitize, through mass spectrometry, the chemodiversity of Swiss botanical gardens. Here high-resolution mass spectrometry is considered as a digital scanner allowing to capture the chemical fingerprint of the profiled sample. State-of-the-art computational metabolomics solutions are used to organize and annotate the acquired spectra with molecular data.

- 3. Gather chemical information and relevant samples metadata in a tailored knowledge graph. Chemical information acquired at the previous step (spectra and chemical structures) are connected to relevant samples metadata (taxonomy, phenology, geolocalisation, time of collection etc.). For this semantic web technologies (namely the RDF data model) are employed and a tailored knowledge graph is established.
- 4. Connect to existing ontologies (bio, chemo) and biodiversity digitization projects. Chemical and biological objects of the graph are connected to relevant pre-established ontologies (e.g. CHEBI, Plants Ontology) and data graphs (e.g. Wikidata). Connection with complementary biodiversity digitization efforts will be done (e.g BiCIKL https://bicikl-project.eu/, Dissco https://www.dissco.eu/)
- 5. Establish web and programmatic interfaces for the query of the acquired knowledge. A web interface will allow a convenient access to the data acquired within the framework of the project. A dashboard will allow simple visualizations (e.g pie charts, barplots, treemap) to interpret the data. In addition a SPARQL endpoint and an application programming interface (API) will allow retrieval of data programmatically.
- 6. Illustrate the feasibility and advantages of an end-to-end Open Science project. The DBGI will strictly follow the Open Science guidelines by using open-source software and making available the acquired data and scripts under an open license agreement. In addition the DBGI results will be published at the moment they are acquired (previous to formal publication or even pre-prints) thus following the Open Notebook Science concepts.
- 7. Establish robust and scalable workflows for the digitization of wild ecosystems biodiversity. The DBGI, albeit ambitious, is a pilot project. The future objective is to propose digitization workflows for the characterization of the wild ecosystems chemodiversity, on a global scale.
- 8. Provide "molecular arguments" for biodiversity conservation policies. The ultimate goal of DBGI is to use all the gathered metabolic information to support, extend or implement conservation efforts worldwide. We believe that by providing chemical maps of the landscape it will be possible to contribute to the prioritization of conservation and restoration targets. In other words, by establishing large scale chemical maps we expect to provide "molecular arguments" to biodiversity conservation endeavors (e.g. "This piece of land has a high content of antibacterial scaffolds." or "This place might be poor in species diversity but rich in a rare chemodiversity".)

This manuscript is a template (aka "rootstock") for <u>Manubot</u>, a tool for writing scholarly manuscripts. Use this template as a starting point for your manuscript.

The rest of this document is a full list of formatting elements/features supported by Manubot. Compare the input (.md files in the /content directory) to the output you see below.

# **Basic formatting**

| Bold text      |  |  |
|----------------|--|--|
| Semi-bold text |  |  |

Centered text

Right-aligned text

#### Combined italics and bold

#### Strikethrough

- 1. Ordered list item
- 2. Ordered list item
  - a. Sub-item
  - b. Sub-item
    - i. Sub-sub-item
- 3. Ordered list item
  - a. Sub-item
- List item
- · List item
- · List item

subscript: H<sub>2</sub>O is a liquid

superscript: 2<sup>10</sup> is 1024.

unicode superscripts<sup>0123456789</sup>

unicode subscripts<sub>0123456789</sub>

A long paragraph of text. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Putting each sentence on its own line has numerous benefits with regard to <u>editing</u> and <u>version</u> control.

Line break without starting a new paragraph by putting two spaces at end of line.

# **Document organization**

Document section headings:

# **Heading 1**

# **Heading 2**

**Heading 3** 

**Heading 4** 

**Heading 5** 

**Heading 6** 



Horizontal rule:

Heading 1's are recommended to be reserved for the title of the manuscript.

Heading 2's are recommended for broad sections such as Abstract, Methods, Conclusion, etc.

Heading 3's and Heading 4's are recommended for sub-sections.

#### Links

Bare URL link: <a href="https://manubot.org">https://manubot.org</a>

<u>Long link with lots of words and stuff and junk and bleep and blah and stuff and other stuff and more stuff yeah</u>

Link with text

Link with hover text

Link by reference

### **Citations**

Citation by DOI [1].

Citation by PubMed Central ID [2].

Citation by PubMed ID [3].

Citation by Wikidata ID [wikidata:Q56458321?].

Citation by ISBN [4].

Citation by URL [5].

Citation by alias [6].

Multiple citations can be put inside the same set of brackets [1,4,6]. Manubot plugins provide easier, more convenient visualization of and navigation between citations [2,3,6,7].

Citation tags (i.e. aliases) can be defined in their own paragraphs using Markdown's reference link syntax:

# Referencing figures, tables, equations

Figure 2

Figure 3

```
Figure <u>4</u>

Figure <u>5</u>

Table <u>1</u>

Equation <u>1</u>

Equation <u>2</u>
```

# **Quotes and code**

Quoted text

Quoted block of text

Two roads diverged in a wood, and I—I took the one less traveled by, And that has made all the difference.

Code in the middle of normal text, aka inline code.

Code block with Python syntax highlighting:

```
from manubot.cite.doi import expand_short_doi

def test_expand_short_doi():
    doi = expand_short_doi("10/c3bp")
    # a string too long to fit within page:
    assert doi == "10.25313/2524-2695-2018-3-vliyanie-enhansera-copia-i-
        insulyatora-gypsy-na-sintez-ernk-modifikatsii-hromatina-i-
        svyazyvanie-insulyatornyh-belkov-vtransfetsirovannyh-geneticheskih-
        konstruktsiyah"
```

Code block with no syntax highlighting:

```
Exporting HTML manuscript
Exporting DOCX manuscript
Exporting PDF manuscript
```

# **Figures**



**Figure 2:** A square image at actual size and with a bottom caption. Loaded from the latest version of image on GitHub.



**Figure 3: An image too wide to fit within page at full size.** Loaded from a specific (hashed) version of the image on GitHub.



Figure 4: A tall image with a specified height. Loaded from a specific (hashed) version of the image on GitHub.



**Figure 5:** A vector .svg image loaded from GitHub. The parameter sanitize=true is necessary to properly load SVGs hosted via GitHub URLs. White background specified to serve as a backdrop for transparent sections of the image.

# **Tables**

**Table 1:** A table with a top caption and specified relative column widths.

| Bowling Scores | Jane | John | Alice | Bob |
|----------------|------|------|-------|-----|
| Game 1         | 150  | 187  | 210   | 105 |
| Game 2         | 98   | 202  | 197   | 102 |
| Game 3         | 123  | 180  | 238   | 134 |

**Table 2:** A table too wide to fit within page.

|    | Digits 1-33                            | Digits 34-66                          | Digits 67-99                          | Ref.      |
|----|--|---------------------------------------|---------------------------------------|-----------|
| pi | 3.14159265358979323<br>846264338327950 | 28841971693993751<br>0582097494459230 | 78164062862089986<br>2803482534211706 | piday.org |
| e  | 2.71828182845904523<br>536028747135266 | 24977572470936999<br>5957496696762772 | 40766303535475945<br>7138217852516642 | nasa.gov  |

 Table 3: A table with merged cells using the attributes plugin.

|       | Colors     |                  |
|-------|------------|------------------|
| Size  | Text Color | Background Color |
| big   | blue       | orange           |
| small | black      | white            |

# **Equations**

A LaTeX equation:

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \tag{1}$$

An equation too long to fit within page:

$$x = a + b + c + d + e + f + g + h + i + j + k + l + m + n + o + p + q + r + s + t + u + v + w + x + y + z + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9$$
(2)

# **Special**

▲ WARNING The following features are only supported and intended for .html and .pdf exports. Journals are not likely to support them, and they may not display correctly when converted to other formats such as .docx.

LINK STYLED AS A BUTTON

Adding arbitrary HTML attributes to an element using Pandoc's attribute syntax:

Manubot Manubot Manubot Manubot Manubot. Manubot Manubot Manubot Manubot. Manubot. Manubot Manubot. Manubot. Manubot. Manubot. Manubot.

Adding arbitrary HTML attributes to an element with the Manubot attributes plugin (more flexible than Pandoc's method in terms of which elements you can add attributes to):

Manubot Manubot.

Available background colors for text, images, code, banners, etc:

white lightgrey grey darkgrey black lightred lightyellow lightgreen lightblue lightpurple red orange yellow green blue purple

Using the Font Awesome icon set:



Light Grey Banner
useful for general information - manubot.org

# **1** Blue Banner

useful for important information - manubot.org

**♦ Light Red Banner** useful for *warnings* - <u>manubot.org</u>

## References

#### 1. Sci-Hub provides access to nearly all scholarly literature

Daniel S Himmelstein, Ariel Rodriguez Romero, Jacob G Levernier, Thomas Anthony Munro, Stephen Reid McLaughlin, Bastian Greshake Tzovaras, Casey S Greene *eLife* (2018-03-01) <a href="https://doi.org/ckcj">https://doi.org/ckcj</a>

DOI: <u>10.7554/elife.32822</u> · PMID: <u>29424689</u> · PMCID: <u>PMC5832410</u>

### 2. Reproducibility of computational workflows is automated using continuous analysis

Brett K Beaulieu-Jones, Casey S Greene

*Nature biotechnology* (2017-04) <a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6103790/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6103790/</a>
DOI: 10.1038/nbt.3780 · PMID: 28288103 · PMCID: PMC6103790

#### 3. **Bitcoin for the biological literature.**

Douglas Heaven

Nature (2019-02) https://www.ncbi.nlm.nih.gov/pubmed/30718888

DOI: 10.1038/d41586-019-00447-9 · PMID: 30718888

#### 4. Open access

Suber Peter

(2012) https://www.worldcat.org/oclc/795846161

ISBN: 9780262517638

#### 5. <a href="https://greenelab.github.io/meta-review/">https://greenelab.github.io/meta-review/</a>

#### 6. Opportunities and obstacles for deep learning in biology and medicine

Travers Ching, Daniel S Himmelstein, Brett K Beaulieu-Jones, Alexandr A Kalinin, Brian T Do, Gregory P Way, Enrico Ferrero, Paul-Michael Agapow, Michael Zietz, Michael M Hoffman, ... Casey S Greene

Journal of The Royal Society Interface (2018-04) <a href="https://doi.org/gddkhn">https://doi.org/gddkhn</a> DOI: <a href="https://doi.org/gddkhn">10.1098/rsif.2017.0387</a> • PMID: <a href="pyg29618526">29618526</a> • PMCID: <a href="https://doi.org/gddkhn">PMC5938574</a>

#### 7. Open collaborative writing with Manubot

Daniel S Himmelstein, Vincent Rubinetti, David R Slochower, Dongbo Hu, Venkat S Malladi, Casey S Greene, Anthony Gitter

PLOS Computational Biology (2019-06-24) https://doi.org/c7np

DOI: 10.1371/journal.pcbi.1007128 · PMID: 31233491 · PMCID: PMC6611653