

Laboratory Notebooks

Thomas Denecker

IFB, <https://orcid.org/0000-0003-1421-7641>

BiGEst



What is this?



Année : _____

Numéro : _____

Cahier de laboratoire

Cahier de laboratoire

Laboratory notebook



Numéro : _____
Number: _____

Unité de recherche : _____
Research Unit: _____

Prénom et nom du ou des utilisateur(s) : _____
First name and last name of the user(s): _____

Signature(s) : _____
Signature(s): _____

Date de début : _____ Date de fin : _____
Start date: _____ End date: _____



Année : _____

Numéro : _____

Cahier de laboratoire
Laboratory notebook

Numéro : _____
Number: _____

Unité de recherche : _____
Research unit: _____

Prénom et nom du ou des utilisateur(s) : _____
First name and last name of the user(s): _____

Signature(s) : _____
Signature(s): _____

Date de début : _____ Date de fin : _____
Start date: _____ End date: _____

Ministère de l'Enseignement Supérieur et de la Recherche
Institut Français de Bioinformatique

The laboratory notebook allows :

- day-to-day recording of the details of the work
- to report on the progress and scientific experimentation, from the idea to the conclusion
- to keep knowledge in a lab

Also very useful for drafting a patent or for proving anteriority.

What is it for ?



A legal tool :

- Each notebook and the pages are numbered.
- On the cover page, we find on each notebook the mentions of the owner of the results.
- Each page has two parts at the bottom intended to be dated and signed: two signatures = two people, user and third party (witness),
 - ideally a third party not involved in the research work but capable of understanding it

<https://www.curie.asso.fr/-Cahier-de-laboratoire-national-.html>

<https://slideplayer.fr/slide/3817405/>

What is it for ?



For all those who carry out research work :

- researchers,
- engineers,
- technicians,
- doctoral students,
- trainees,
- etc.

Année : _____

Numéro : _____

Cahier de laboratoire
Laboratory notebook

Numéro : _____
Number: _____

Unité de recherche : _____
Research unit: _____

Prénom et nom du ou des utilisateur(s) : _____
First name and last name of the user(s): _____

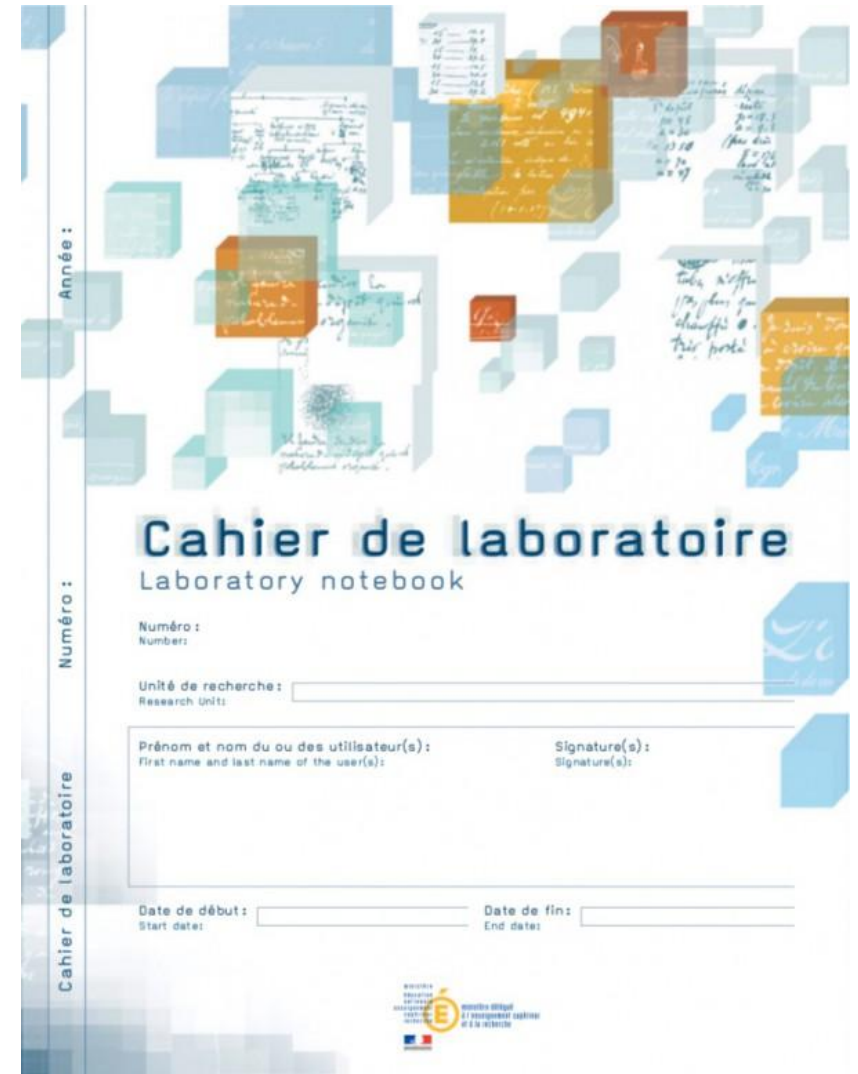
Signature(s) : _____
Signature(s): _____

Date de début : _____ Date de fin : _____
Start date: _____ End date: _____

Ministère de l'Enseignement Supérieur et de la Recherche



Are you using it ?





Modern LN since 2009 (C.U.R.I.E. Network)

But less and less adapted to recent evolutions of our work

- Increased data quantity
- Change in the nature of data
- Dematerialization
- Security

We need an electronic tool for individual traceability.



Le rapport du groupe de travail « Cahier de laboratoire électronique » (ELN) présente une vision partagée sur la définition, le cadrage, les usages et le périmètre fonctionnel de l'ELN, qui doit pouvoir s'intégrer dans les environnements informatiques et institutionnels existants. Il émet un ensemble de recommandations sur les critères de choix d'un outil et intègre une liste comparative d'outils existants.



Calendrier du projet



Note de synthèse

- Caractéristiques techniques souhaitées
- Inventaire des solutions existantes sur le marché

Janvier 2021

Mise en œuvre de la stratégie adoptée

- Mise en œuvre de la procédure d'achat (licences, etc.)
- Cahier des charges de la solution à développer

Mars-Juillet 2021

Mise à disposition de ou des solutions sélectionnées

- Outils déployés sous licences CNRS
- Solution électronique développée en interne

Janvier 2022 – Juin 2023



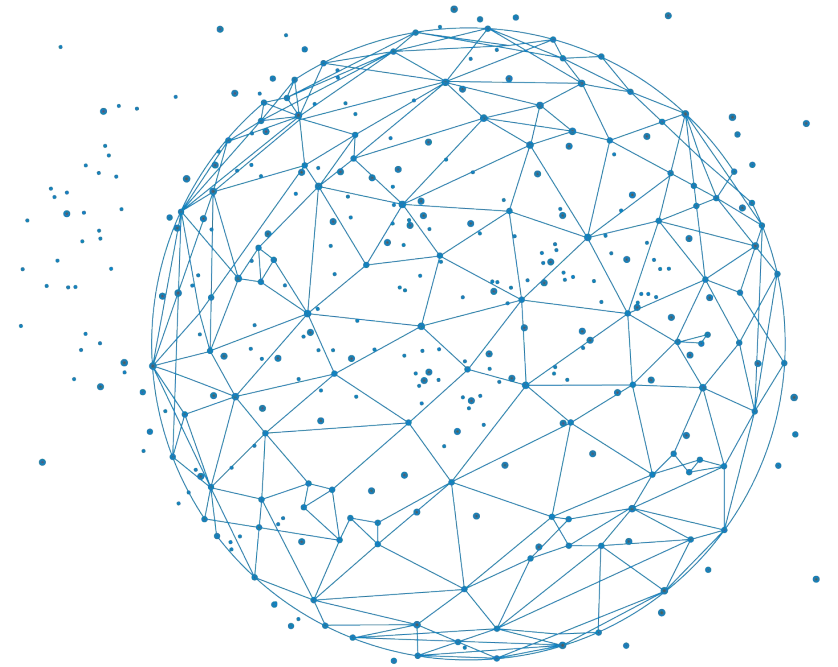
<https://jso-cnrs-2020.sciencesconf.org/333438/document>

Cahier de laboratoire électronique 17.11.20 P 19/18



- Introduction to Laboratory Notebooks ✓
- Literate programming
- Markdown
- Notebooks for bioinformatics
 - R Notebooks (Rmarkdown)
 - Quarto
 - Jupyter(Lab)
- Practical session (JupyterLab)

Literate programming





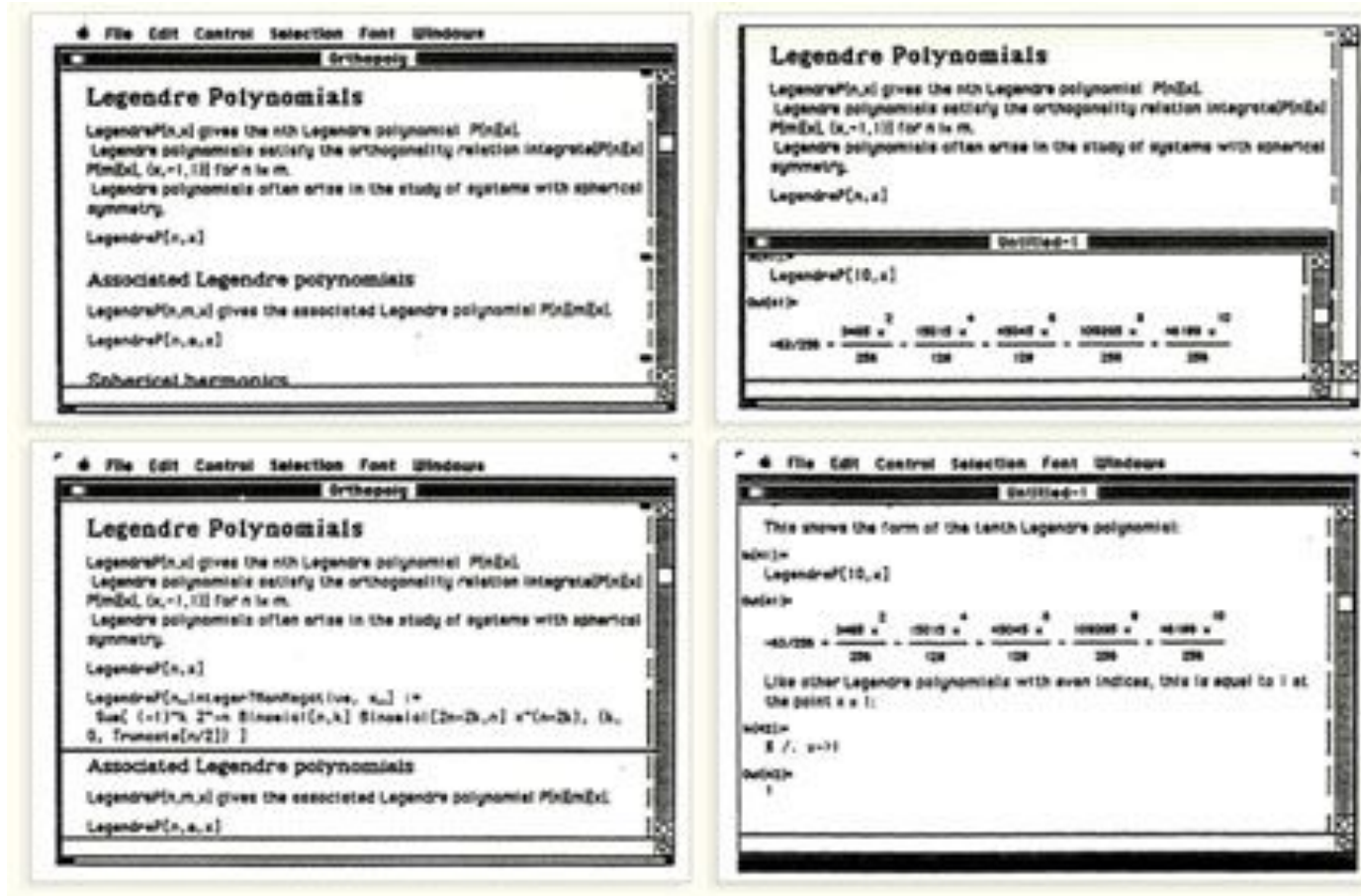
Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

— Donald E. Knuth, *Literate Programming*, 1984



A literate computing environment is one that allows users not only to execute commands interactively, but also to store in a literate document the results of these commands along with figures and free-form text.

- Millman KJ and Perez F (2014)



Wolfram Mathematica notebook (1987)



What does it look like ?

jupyter Lorenz Differential Equations (autosaved)

File Edit View Insert Cell Kernel Help Python 3

Code Cell Toolbar: None

Exploring the Lorenz System

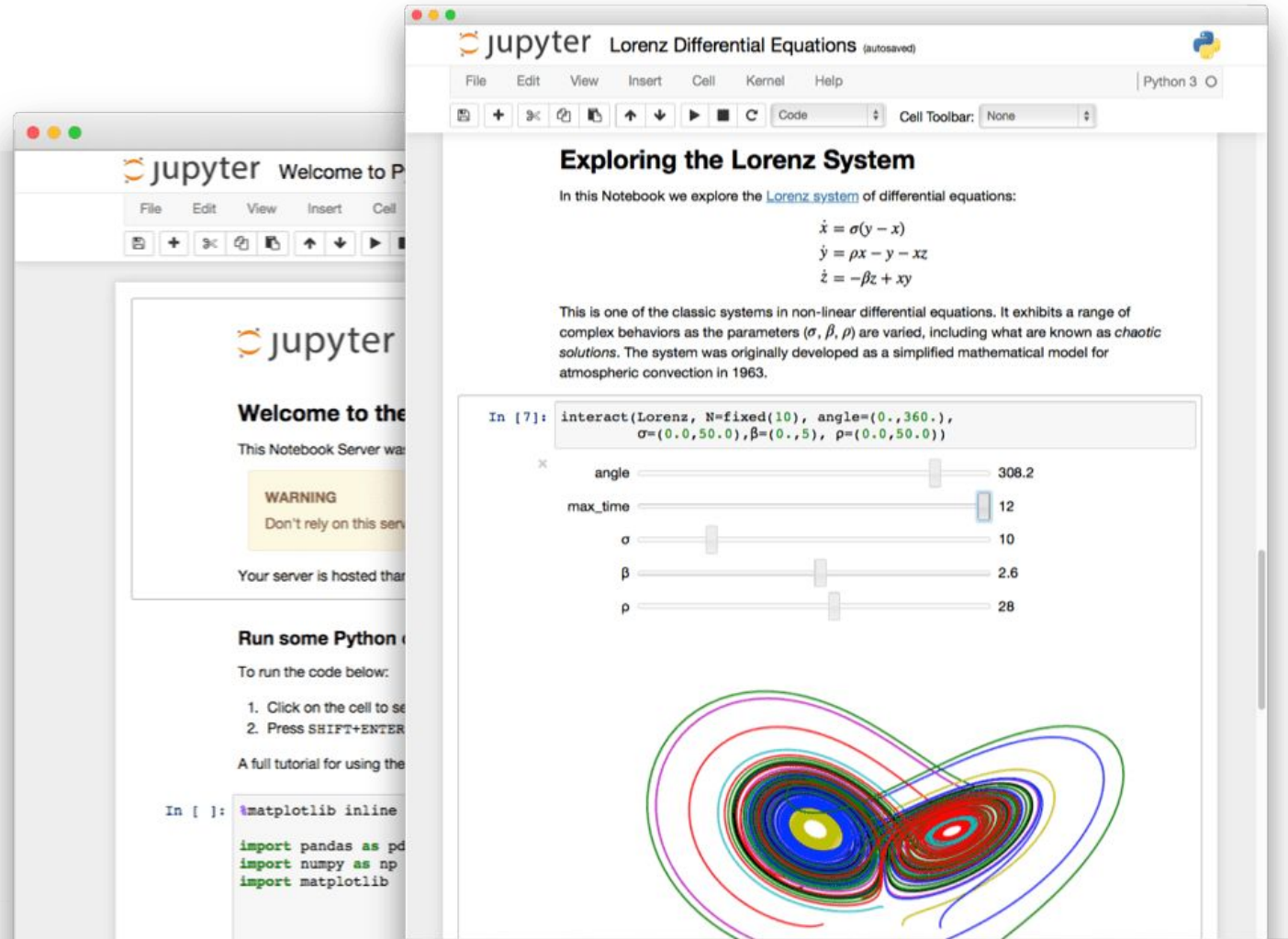
In this Notebook we explore the [Lorenz system](#) of differential equations:

$$\begin{aligned}\dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy\end{aligned}$$

This is one of the classic systems in non-linear differential equations. It exhibits a range of complex behaviors as the parameters (σ, β, ρ) are varied, including what are known as *chaotic solutions*. The system was originally developed as a simplified mathematical model for atmospheric convection in 1963.

```
In [7]: interact(Lorenz, N=fixed(10), angle=(0.,360.),
                sigma=(0.0,50.0), beta=(0.,5), rho=(0.0,50.0))
```

angle 308.2
max_time 12
 σ 10
 β 2.6
 ρ 28



Interactive programming interface

allowing to combine both natural and computer languages.

In one file:

- Explanations
- Code
- Results
- Graphs and plots



Why using literate programming frameworks ?

Use cases:

- Labbook
- Day to day analyses
- Analysis reports
- Writing scientific article

Example of an article entirely written using a notebook



File (on a repository)

The screenshot shows a GitHub repository for 'colomoto-docker'. The file 'invasion.ipynb' is selected, showing its commit history and content. The file content is the title and author information of the article: 'Prediction of Mutations to Control Pathways Enabling Tumour Cell Invasion with the CoLoMoTo Interactive Notebook (Tutorial)'. The authors listed are Nicolas Levy, Aurélien Naldi, Céline Hernandez, Gautier Stoll, Denis Thieffry, Andrei Zinovyev, Laurence Calzone, and Loïc Paulevé.

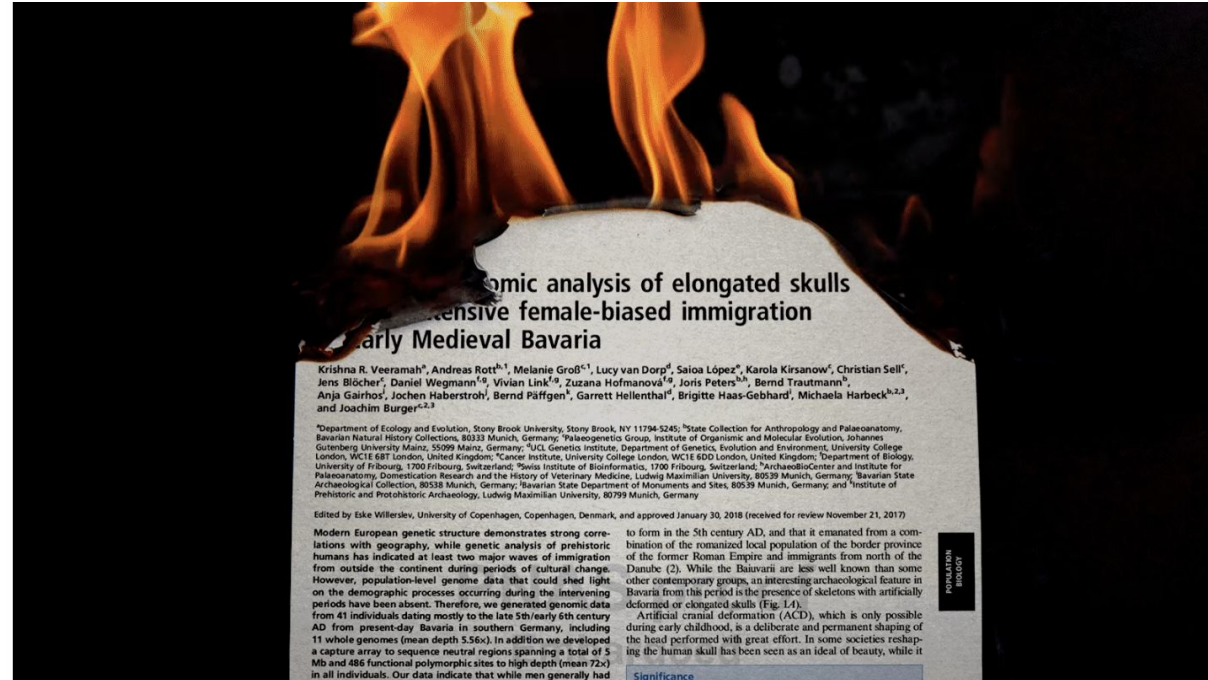
Published article

The screenshot shows the published article on the Frontiers website. The article title is 'Prediction of Mutations to Control Pathways Enabling Tumor Cell Invasion with the CoLoMoTo Interactive Notebook (Tutorial)'. The authors are listed as Nicolas Levy, Aurélien Naldi, Céline Hernandez, Gautier Stoll, Denis Thieffry, Andrei Zinovyev, Laurence Calzone, and Loïc Paulevé. The article has 1,866 total views and is categorized as a 'PROTOCOLS ARTICLE'.

Executable file

The screenshot shows the executable file page on nbviewer. The article title is 'Prediction of Mutations to Control Pathways Enabling Tumour Cell Invasion with the CoLoMoTo Interactive Notebook (Tutorial)'. The authors are listed as Nicolas Levy, Aurélien Naldi, Céline Hernandez, Gautier Stoll, Denis Thieffry, Andrei Zinovyev, Laurence Calzone, and Loïc Paulevé. The abstract is visible, starting with 'Boolean and multi-valued logical formalisms are increasingly used to model complex cellular networks...'

DOI:10.3389/fphys.2018.00787



PNAS / Richard Goerg / Getty / The Atlantic

SCIENCE

THE SCIENTIFIC PAPER IS OBSOLETE

Here's what's next.

By James Somers

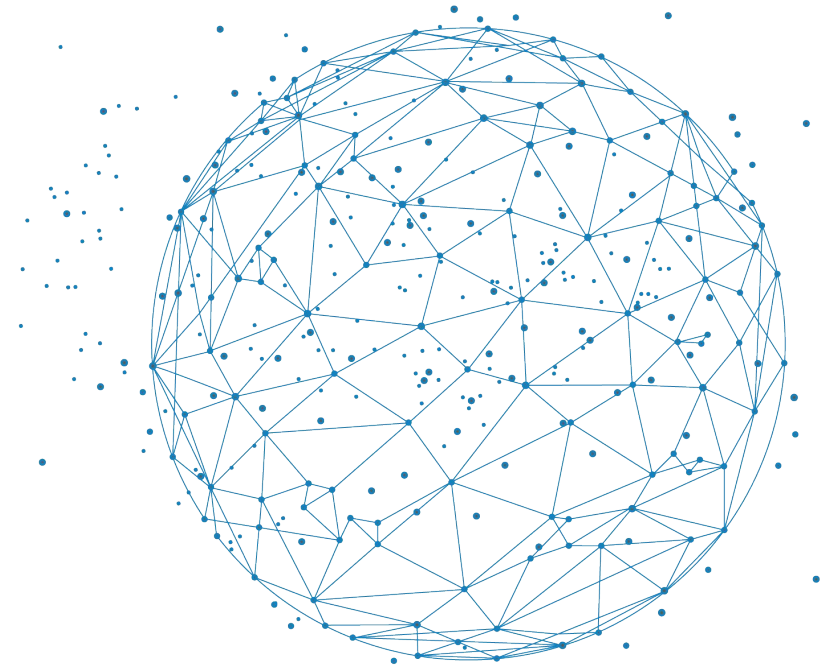
APRIL 5, 2018

SHARE SAVE

THE SCIENTIFIC paper—the actual form of it—was one of the enabling inventions of modernity. Before it was developed in the 1600s, results

We need to be careful

Markup / Markdown





Definition

A markup language uses tags to define elements within a document.

Three different types and usage

- Presentational (used by traditional word-processing systems)
 - Markup is invisible
- Procedural, provides instructions to process the text (e.g. TeX, PostScript)
 - Markup is visible and can be directly manipulated by the author.
- Descriptive, to label documents parts (e.g. LaTeX, HTML, XML...)
 - Emphasizes the document structure.



Example in HTML

```
<h1>Heading</h1>
<h2>Sub-heading</h2>
<a href="www.webpage.com">Link</a>
<ul>
  <li>List-item1</li>
  <li>List-item2</li>
  <li>List-item3</li>
</ul>
```



Heading

Sub-heading

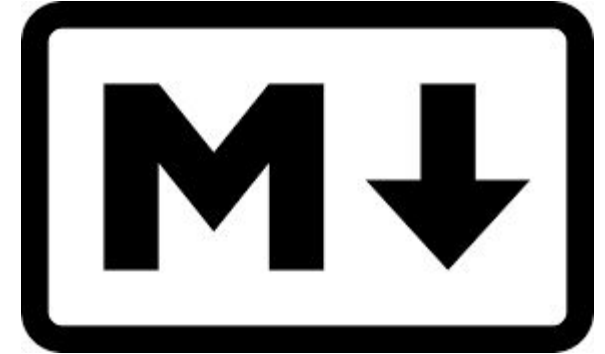
Link

- List-item1
- List-item2
- List-item3

Markdown is a Lightweight markup language

Designed to be :

- easy to write using any generic text editor (plain-text-formatting syntax)
- easy to read in its raw form



From GitHub's help page

<https://docs.github.com/en/get-started/writing-on-github/getting-started-with-writing-and-formatting-on-github/basic-writing-and-formatting-syntax>

Example in markdown

```
# Heading
```

```
## Sub-heading
```

```
### Another deeper heading
```

```
A [link](http://example.com).
```

```
Text attributes _italic_, *italic*, **bold**, `monospace`.
```

```
Bullet list:
```

- * apples
- * oranges
- * pears

But how is this useful for literate programming?

When you want to weave both code (to be interpreted) and formatting information, you precisely need a lightweight language for the formatting part.

```
1- example.Rmd
1 ----
2 title: "Viridis Demo"
3 output: html_document
4 ----
5
6 ```{r include = FALSE}
7 library(viridis)
8 ```
9
10 The code below demonstrates two color palettes in the
11 [viridis](https://github.com/sjmgarnier/viridis) package. Each plot
12 displays a contour map of the Maunga Whau volcano in Auckland, New
13 Zealand.
14
15 ## Viridis colors
16 ```{r}
17 image(volcano, col = viridis(200))
18 ```
19
20 ## Magma colors
21 ```{r}
22 image(volcano, col = viridis(200, option = "A"))
23 ```
```

Viridis Demo

The code below demonstrates two color palettes in the `viridis` package. Each plot displays a contour map of the Maunga Whau volcano in Auckland, New Zealand.

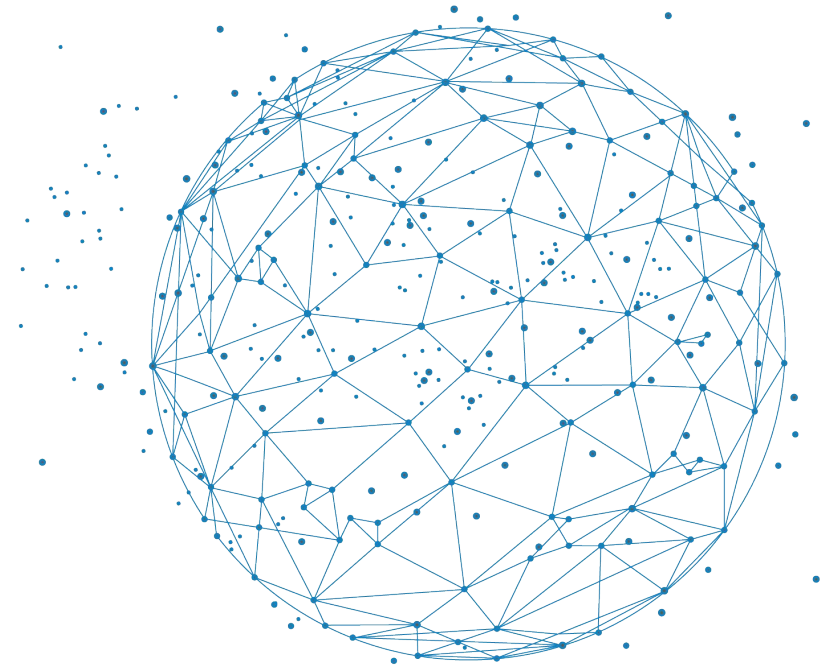
Viridis colors

`image(volcano, col = viridis(200))`

Magma colors

`image(volcano, col = viridis(200, option = "A"))`

Best practices





Notebooks et science ouverte : FAIR mieux

Mariannig Le Béhec, Célya Gruson-Daniel, Clémence Lascombes, Émilien Schultz

► To cite this version:

Mariannig Le Béhec, Célya Gruson-Daniel, Clémence Lascombes, Émilien Schultz. Notebooks et science ouverte : FAIR mieux. 2024. hal-04485968

HAL Id: hal-04485968

<https://hal.science/hal-04485968>

Preprint submitted on 1 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Catégorie générale Rendre votre analyse traçable et reproductible

Bonnes pratiques

- Utiliser un gestionnaire de révision
- Gérer les dépendances du projet
- Fournir des applications sans dépendances tierces
- Placer les imports au début du fichier
- Assurer le bon fonctionnement du code en complet et pas uniquement la partie modifiée

Écrire un code de qualité (i.e. un code qui peut être facilement partagé et réutilisé)

- Structurer votre code en modules (abstraire le code en fonctions et les placer dans un module dédié ; placer les dépendances au début du carnet)
- Tester votre code
- Nommer vos carnets de manière cohérente
- Respecter les standards
- Utilisez les chemins relatifs
- Définir les *requirements*
- Documenter votre code pour vous et les autres
- Utiliser les titres Markdown pour structurer votre carnet
- Garder votre carnet clair
- Garder votre carnet concis
- Faire la différence entre les artefacts produits pendant le développement et la production

Exploiter le paradigme de la programmation lettrée

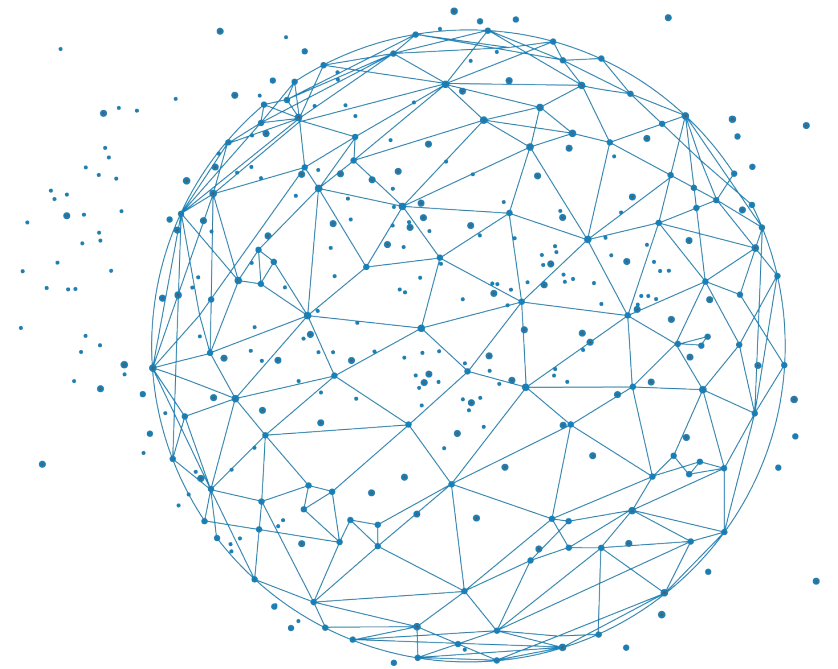
Garder votre carnet clair et concis

Faire la différence entre les artefacts produits pendant le développement et la production

Adopter une diffusion ouverte

- Mettre à disposition votre carnet
- Mettre à disposition vos données

Limites





In this article, the authors highlight several limitations:

Reproducibility

- Lack of documentation
- Lack of consistency in cell execution
- Version tracking
- The quality of the code is often low

Interface dependent

Limited for large data sets

And now we try?

