



HAL
open science

Madbot

Thomas Denecker, Julien Seiler

► **To cite this version:**

Thomas Denecker, Julien Seiler. Madbot. Journée "Cycle De Vie / Gestion Des Données De La Biologie", F. Bekkouche; X. Borda; F. Brau; W. Demet, Apr 2024, Sophia Antipolis, France. hal-04531785

HAL Id: hal-04531785

<https://hal.science/hal-04531785v1>

Submitted on 4 Apr 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.

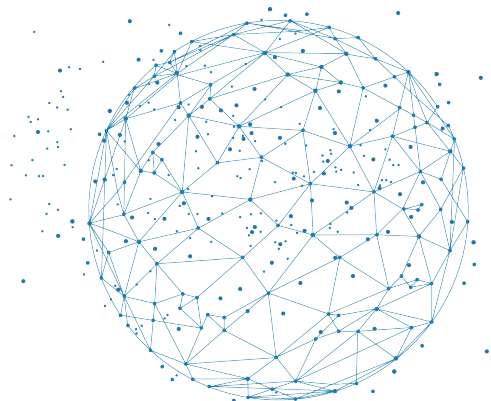
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Speaker
webcam

Lifecycle / Biology Data Management Day - Sophia Antipolis - April 04, 2024



Madbot

A gateway to open science

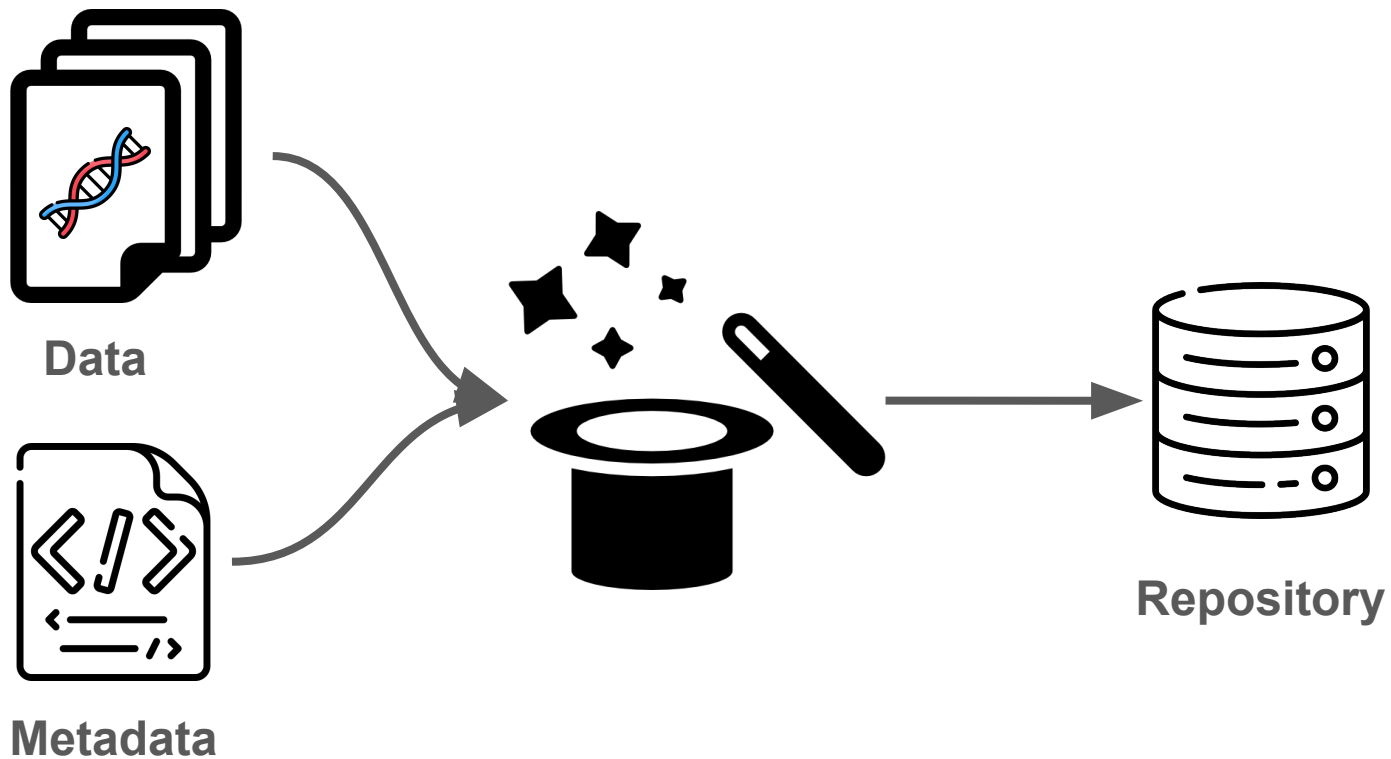
Julien Seiler

IGBMC & IFB - <https://orcid.org/0000-0002-4549-5188>

Thomas Denecker

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





First issue: where are the files?

Speaker
webcam



We're not all organized in the same way...



Metadata, you see, is really a love note – it might be to yourself, but in fact it’s a love note to the person after you, or the machine after you, where you’ve saved someone that amount of time to find something by telling them what this thing is.

Jason Scott

<http://ascii.textfiles.com/archives/3181>



Metadata, you see, is really a love note – it might be to yourself, but in fact it’s a love note to the person after you, or the machine after you, where you’ve saved someone that amount of time to find something by telling them what this thing is.

Jason Scott

<http://ascii.textfiles.com/archives/3181>

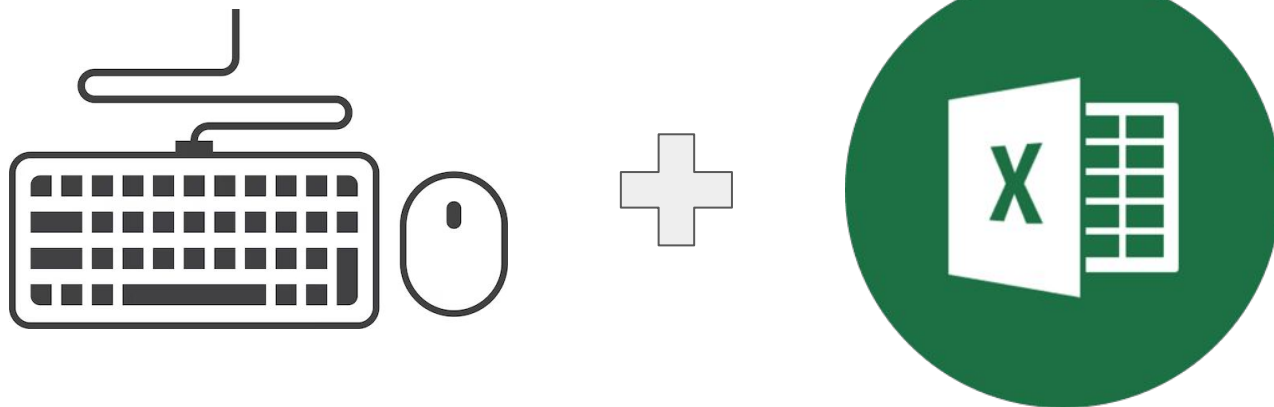
But how do you write this love note?

Speaker
webcam



🤔 How can someone other than me retrieve the information easily?

Speaker
webcam

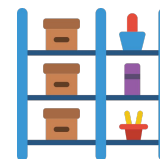


🤔 How do I co-edit metadata? How can I always have the right version?

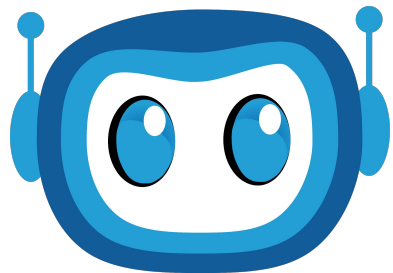
Madbot (Metadata And Data Brokering Online Tool) is a web application that provides a **dashboard for managing research data and metadata**.



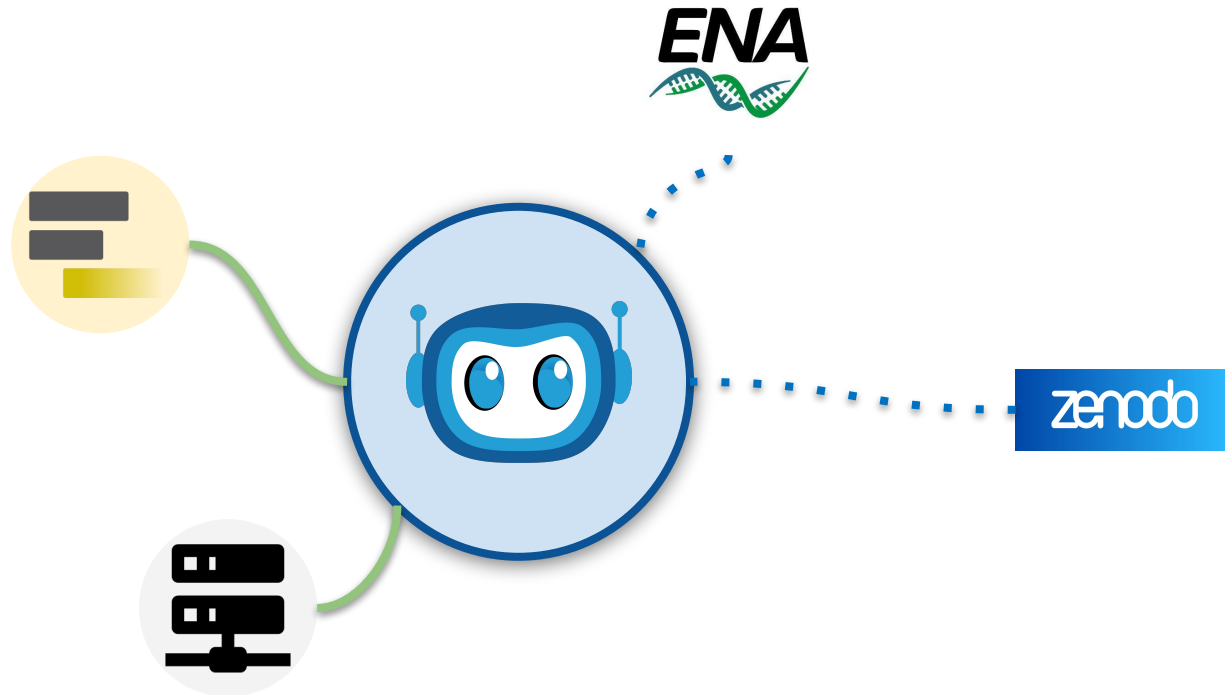
The application can be used to **aggregate and track metadata** for scientific projects, and to **identify exactly where the data is stored** via a number of connectors.



It also supports researchers in **publishing data and metadata** to many general (RGD, Zenodo, etc.) and thematic (ENA, GEO, etc.) data warehouses.

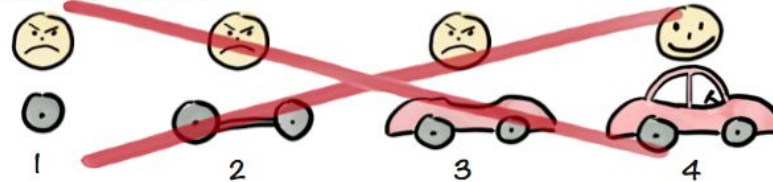


Speaker
webcam

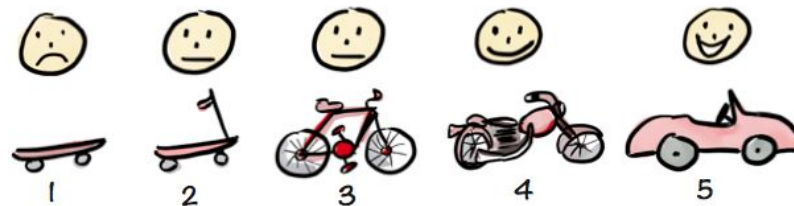


We want to produce a working application as quickly as possible.

Not like this....



Like this!

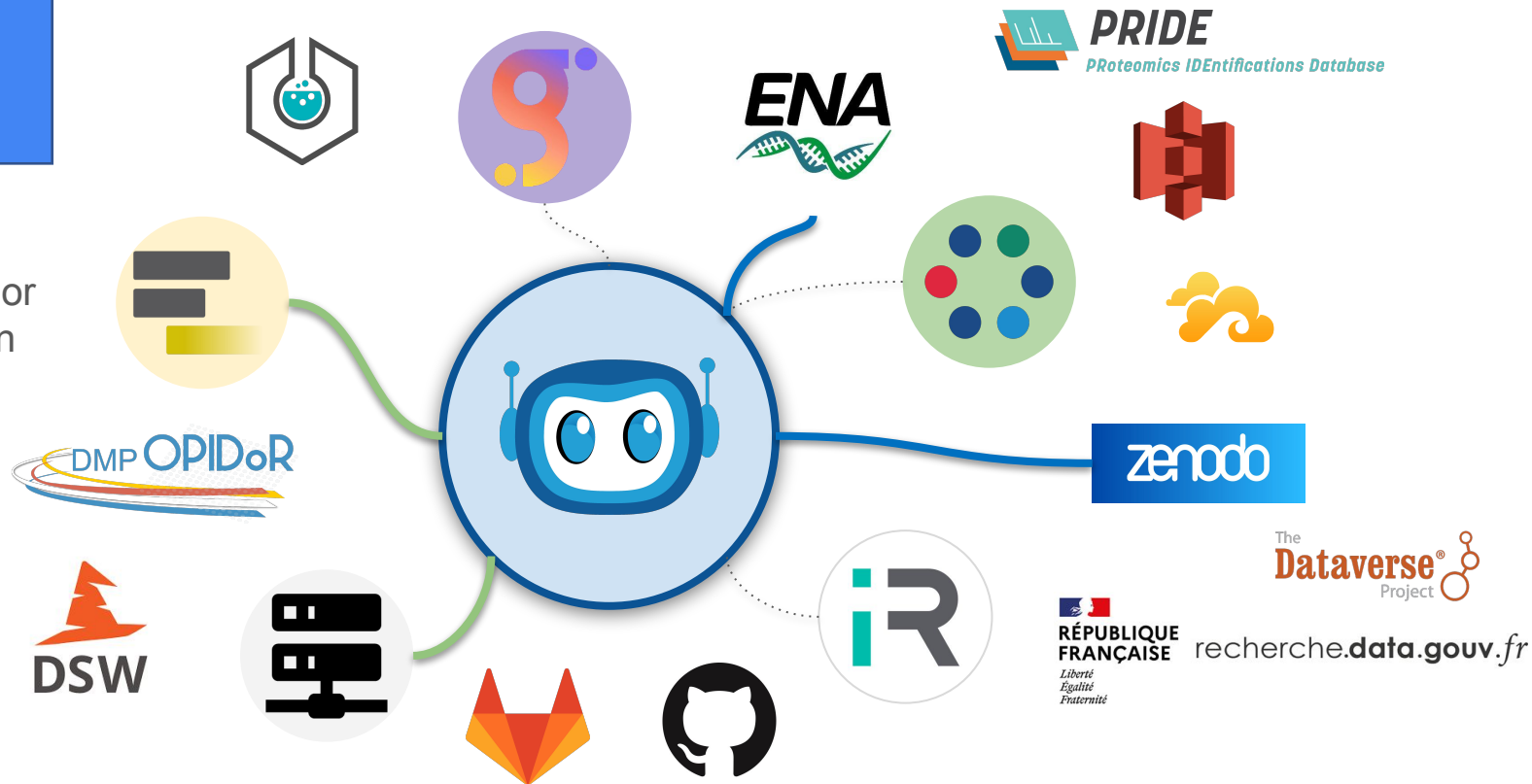


The Madbot application capable of achieving these objectives is our **Minimal Viable Product (MVP)**.

Many connectors are possible for the future

Speaker
webcam

When?
After the MVP, or
by contribution



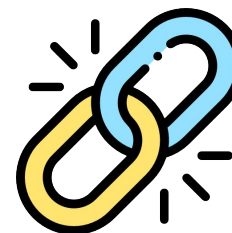
- Step 1 — Provide a REST API and OAuth2 authentication service**
With a developer documentation and an interactive interface
- Step 2 — Provide custom structure to organize data and metadata**
For example, ISA model (Investigation, Study, Assay)
- Step 3 — Link data files from SSHFS & Galaxy**
With a simple, secure and reusable connection system
- Step 4 — View a data dashboard**
Volume, activity, etc.
- Step 5 — Associate metadata**
With input help : automatic curation, pre-selection etc
- Step 6 — Publish data to ENA and Zenodo**
With a collect of external identifiers

Speaker
webcam



Authentication systems

LifeScience AAI (Elixir), DOI,
GitLab (already tested), ORCID



New connectors

Recherche data gov,
DMP(Opidor & DSW), iRODs,...



Hackathon



Training

Speaker
webcam



Thomas DENECKER

Full stack developer & Scrum
master
Paris



Julien SEILER

Full stack developer & scrum
master
Strasbourg



Laurent Bouri

Full stack developer
Strasbourg



Imane MESSAK

Full stack developer
Paris



Baptiste ROUSSEAU

Full stack developer
Paris

+ a working group
throughout France

Speaker
webcam



The **Institut Français de Bioinformatique (IFB)** is funded by the Programme d'Investissements d'Avenir (PIA), grant RENABI-IFB, Agence Nationale de la Recherche number ANR-11-INBS-0013.

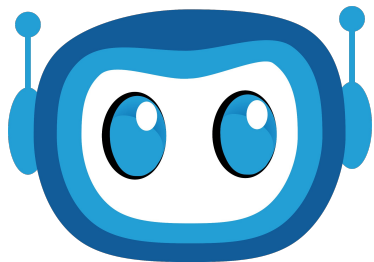


MUDIS4LS is supported by the French government grant from the Agence Nationale de la Recherche under the Investissements d'Avenir Program for structuring research facilities / EQUIPEX+, reference ANR-21-ESRE-0048.

The **maDMP4LS** project (*machine-actionable Data Management Plans for Life Sciences*) has been funded by the French Agence Nationale de la Recherche (grant [ANR-19-DATA-0017-01](#)) as part of the Flash call "[Open Science: research practices and open data](#)".

The **OpenLink** project (A gateway between imaging data management tools to apply FAIR principles) has been funded by the French Agence Nationale de la Recherche (grant [ANR-19-DATA-0011-01](#)) as part of the Flash call "[Open Science: research practices and open data](#)".

Speaker
webcam



Our project is open source under license on Gitlab

<https://gitlab.com/ifb-elixirfr/madbot>



If you want to add a connector, join our working group or develop in Madbot, contact us!

ifb-madbot@groupe.france-bioinformatique.fr

