



HAL
open science

Orchestrating data flows throughout their whole life cycle

Konogan Bourhy, Laurent Bouri, Christophe Bruley, Olivier Collin, Frédéric De Lamotte, Thomas Denecker, Marie-Dominique Devignes, Jean-François Dufayard, Alban Gaignard, Nadia Goué, et al.

► To cite this version:

Konogan Bourhy, Laurent Bouri, Christophe Bruley, Olivier Collin, Frédéric De Lamotte, et al.. Orchestrating data flows throughout their whole life cycle: Key stages and partners for success. JOBIM 2024 - Journées ouvertes en biologie, informatique, et mathématiques, Jun 2024, Toulouse, France. , 2024. hal-04618277

HAL Id: hal-04618277

<https://hal.science/hal-04618277>

Submitted on 20 Jun 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

Orchestrating data flows throughout their whole life cycle

Key stages and partners for success

Konogan BOURHY¹, Laurent BOURI¹, Christophe BRULEY², Olivier COLLIN³, Frédéric de LAMOTTE^{1,4}, Thomas DENECKER¹, Marie-Dominique DEVIGNES⁵, Jean-François DUFAYARD^{1,6}, Alban GAINARD⁷, Nadia GOUÉ⁸, Gildas LE CORGUILLE^{1,9}, Paulette LIEBY¹, Hervé MENAGER^{1,10}, Imane MESSAK¹, Hamid OUAHIOUNE¹, Claire RIOUALEN¹, Thomas ROSNET¹, Baptiste ROUSSEAU¹, Julien SEILER^{1,11}, Jacques VAN HELDEN¹

¹ Institut Français de Bioinformatique, IFB-core, UAR CNRS 3601, Paris, France; ² Laboratoire Biosciences et bioingénierie pour la santé, UA13 INSERM/CEA/UGA, 17 Avenue des martyrs, 38054 Grenoble Cedex 9, France; ³ GenOuest, University of Rennes, INRIA, CNRS, IRISA, Rennes, France; ⁴ UMR AGAP Institut (1334) Avenue Agropolis - TA A-108/03 34398 Montpellier Cedex 5 France; ⁵ Université de Lorraine, CNRS, Inria, LORIA, Nancy F-54000, France; ⁶ UMR AGAP Institut (1334) Univ Montpellier, CIRAD, INRAE, Institut Agro 34398 Montpellier Cedex 5 France; ⁷ Nantes Université, CNRS, INSERM, l'Institut du thorax, Nantes F-44000, France; ⁸ Université Clermont Auvergne, Plateforme AuBi and Mésocentre Clermont-Auvergne, F-63000 Clermont-Ferrand, France; ⁹ Sorbonne Université, CNRS, FR2424, ABiMS, Station Biologique, 29680, Roscoff, France; ¹⁰ Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, 75015, Paris, France; ¹¹ IGBMC, 1 Rue Laurent Fries, 93800, Illkirch, France, CNRS UMS7104

Corresponding Authors: julien.seiler@cnrs.fr, frederic.de-lamotte@france-bioinformatique.fr

Scientific project conception and planning

Metadata : Data :

Success criteria: planning through a Data Mgmt Plan
 Partner for success: data stewards
 Tools:

Sample collection and preparation

Metadata : Data :

Success criteria: adoption of standards
 Partners in success: platforms
 Tools:

Primary data production

Metadata : Data :

Success criteria: description of the production context, enrichment of metadata

Data integration and analysis

Metadata : Data :

Success criteria: long-term storage, collaborative access to data, identification of analysis tools, adoption of standards (data format).
 Partners in success: experts in scientific computing, experts in data analysis

Tools:

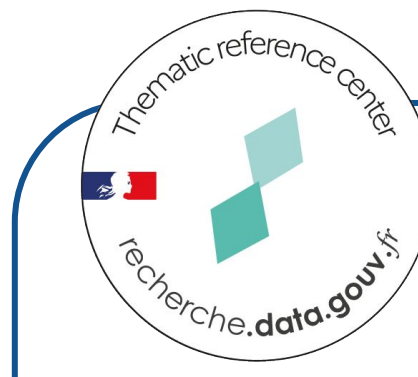
Partners in success: platforms, research teams
 Tools: companion files, Excel spreadsheets

Curation, Preservation, Publication, indexation

Metadata : Data :

Success criteria: identification of relevant and sustainable warehouses

Partners in success: data workshops
 Tools:



A successful data life cycle on the road to FAIR

To manage your research project successfully, you need to optimise every stage. IFB, as a national Thematic Reference Center for Biology and Health, identified key stages you should pay attention to, spanning from the original idea to sharing data in a repository. Additionally, we highlighted partners and tools that shall assist you in **filling the Data and Metadata progress bars** :

Metadata : Data :

FAIR

IFB ecosystem
INSTITUT FRANÇAIS DE BIOINFORMATIQUE
 supported by **MUDIS4LS**
Mutualized Digital Space for FAIR Data in Life and Health Science

National Network of Computing Resources

Clouds and clusters for all bioinformatics needs

Coming out in Autumn 2024

MADBOT

A web application to help research teams **track** their **data** and **metadata**, and **publish** to numerous repositories

FAIR checker

Assess the **FAIRness** of your data and advance in the **FAIRification** journey