



**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. JOBIM 2024 - Journées ouvertes en biologie, informatique, et mathématiques, Jun 2024, Toulouse, France. , 2024. hal-04618277

**HAL Id: hal-04618277**

**<https://hal.science/hal-04618277v1>**

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 BOURHY<sup>1</sup>, Laurent BOURI<sup>1</sup>, Christophe BRULEY<sup>2</sup>, Olivier COLLIN<sup>3</sup>, Frédéric de LAMOTTE<sup>1,4</sup>, Thomas DENECKER<sup>1</sup>, Marie-Dominique DEVIGNES<sup>5</sup>, Jean-François DUFAYARD<sup>1,6</sup>, Alban GAINARD<sup>7</sup>, Nadia GOUÉ<sup>8</sup>, Gildas LE CORGUILLE<sup>1,9</sup>, Paulette LIEBY<sup>1</sup>, Hervé MENAGER<sup>1,10</sup>, Imane MESSAK<sup>1</sup>, Hamid OUAHIOUNE<sup>1</sup>, Claire RIOUALEN<sup>1</sup>, Thomas ROSNET<sup>1</sup>, Baptiste ROUSSEAU<sup>1</sup>, Julien SEILER<sup>1,11</sup>, Jacques VAN HELDEN<sup>1</sup>

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

Corresponding Authors: [julien.seiler@cnrs.fr](mailto:julien.seiler@cnrs.fr), [frederic.de-lamotte@france-bioinformatique.fr](mailto: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

[recherche.data.gouv.fr](http://recherche.data.gouv.fr)

Partners in success: data workshops  
 Tools:

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



Tools:

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



The Institut Français de Bioinformatique (IFB) is funded by the Programme d'Investissements d'Avenir (PIA), grants Agence Nationale de la Recherche ANR-11-INBS-0013 (PIA1 INBS RENABI-IFB) and ANR-21-ESRE-0048 (PIA3 Equipex+ MUDIS4LS).