ifb
INSTITUT FRANÇAIS DE BIOINFORMATIQUE

METABOHUB

4 Wm

Workflow4metabolomics

cnrs

SCIENCES SORBONNE UNIVERSITÉ

INRAE
la science pour la vie, l'humain, la terre

cea
DE LA RECHERCHE À L'INDUSTRIE

LABERCA

ULB UNIVERSITÉ LIBRE DE BRUXELLES

UNIVERSITY OF BIRMINGHAM

W4M

DUBii - 2021

Mélanie Pétéra

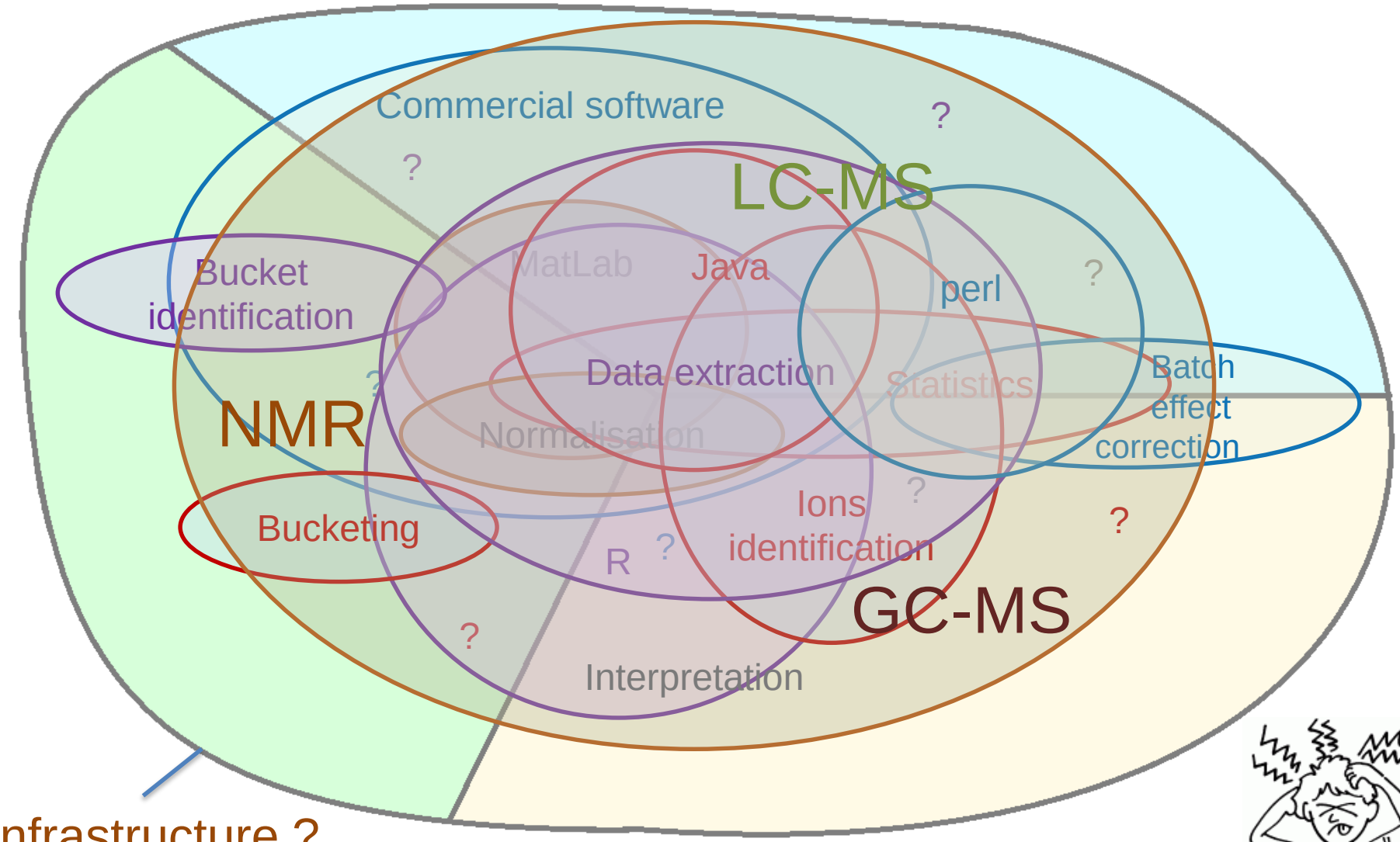
22/03/2021

v 1.0.0

Workflow4Metabolomics – W4M

POURQUOI ?

Une multitude d'outils



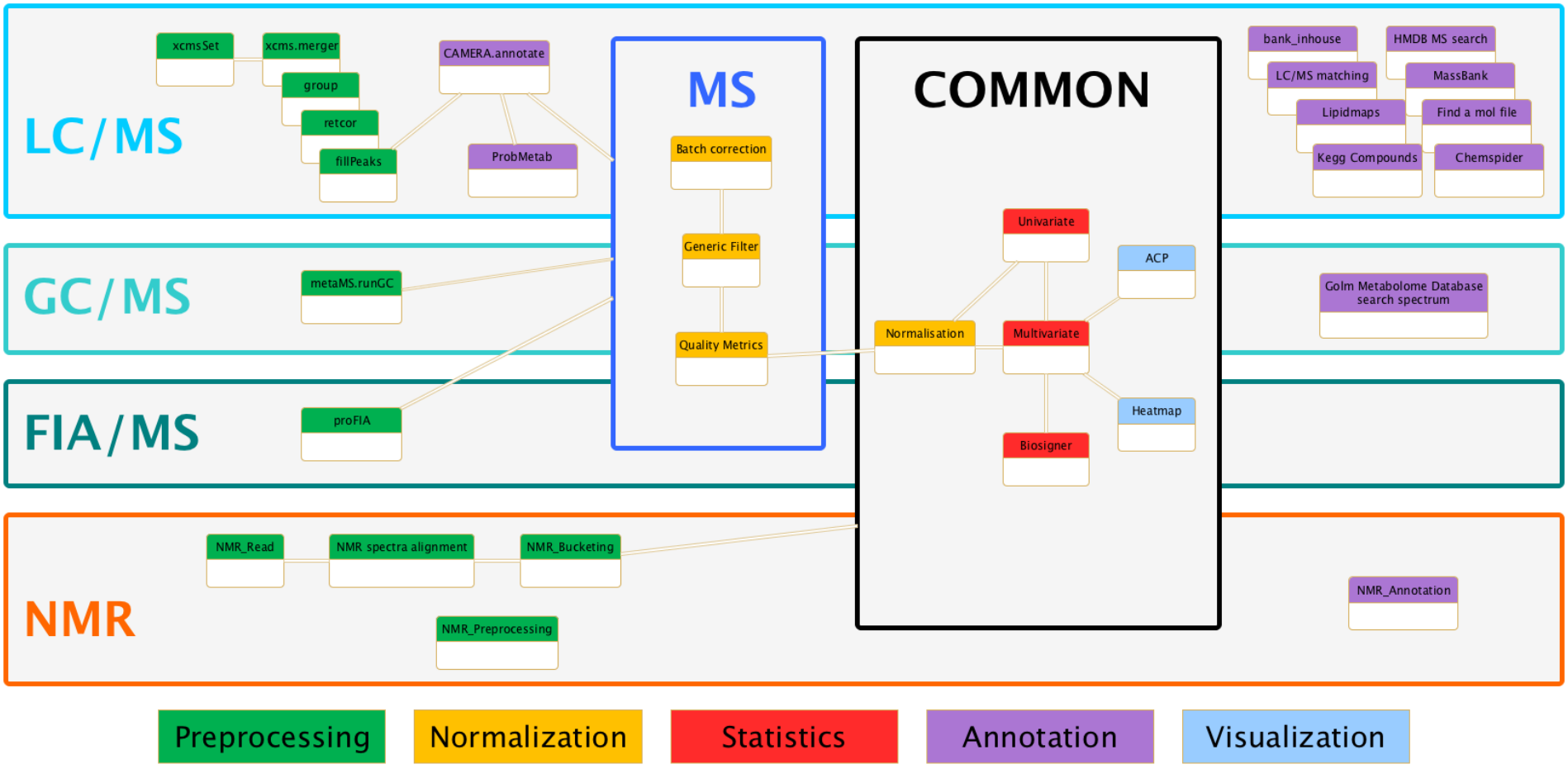
Infrastructure ?

COMMENT ?

W4M – un travail collaboratif

- Un partenariat entre plusieurs structures dont MetaboHub et l'IFB (infrastructures d'avenir)
- Développement d'outils autour de la métabolomique : RMN, MS
- Mise à disposition de ces outils pour pouvoir traiter les données « de A à Z »

Des outils mis en commun pour couvrir l'essentiel



Le choix de Galaxy

https://galaxyproject.org/

Gestionnaire de workflow

Galaxy

Data Intensive *analysis* for everyone

- Versatile and reproducible workflows
- **Web** platform
- **Open source** under Academic Free License
- Developed at Penn State, Johns Hopkins, OHSU and Cleveland Clinic with substantial outside contributions

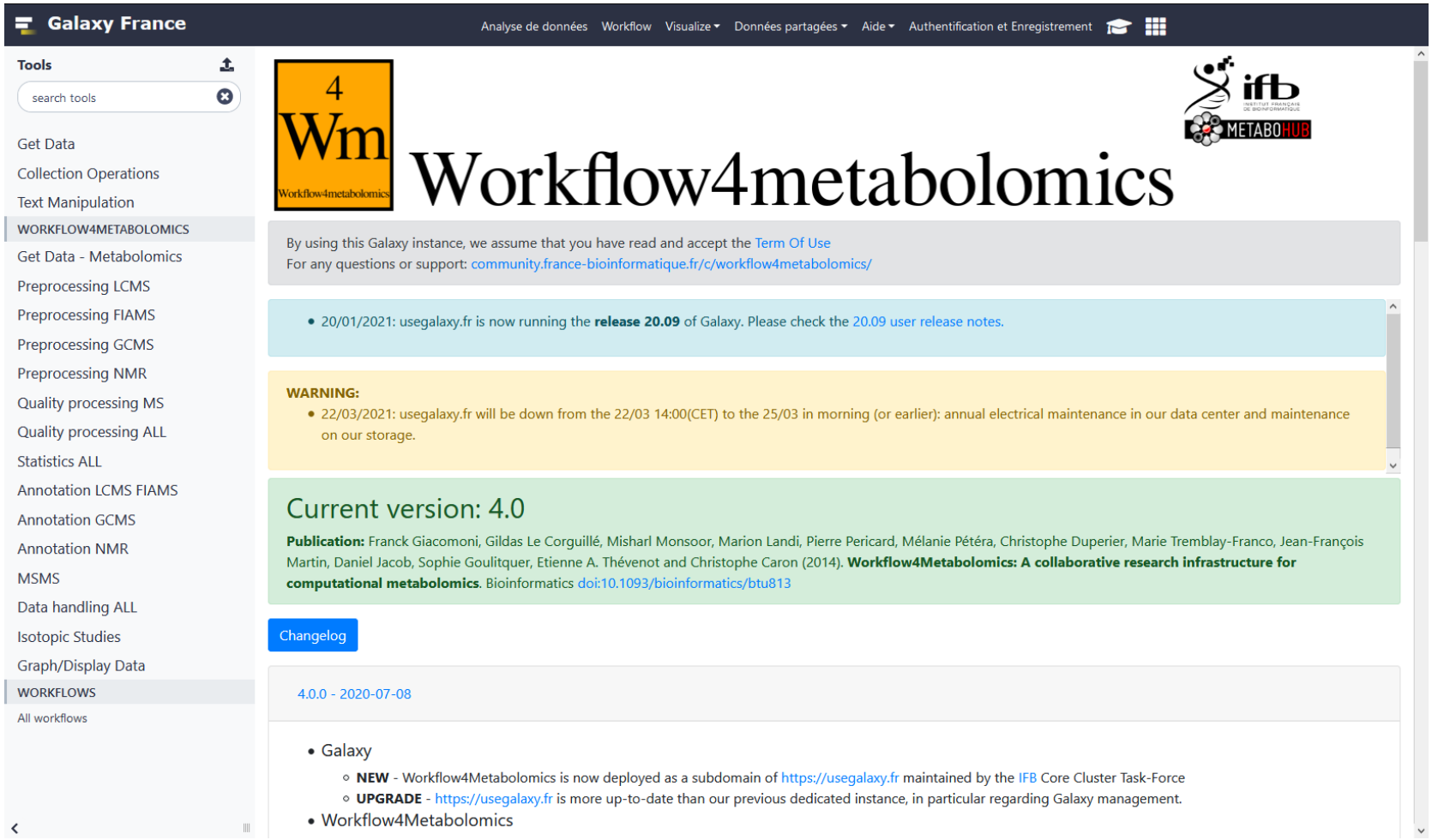
https://training.galaxyproject.org/training-material/topics/introduction/slides/introduction.html#3

Le choix de Galaxy

Un équilibre entre besoins des utilisateurs et des pourvoyeurs d'outils

User advantages	Developer advantages
√ GUI	√ Easy integration of tools
√ Ergonomics	√ Multi-language
√ Modularity	√ Collaborative development
√ Reproducibility	√ Open-source
√ Data & workflow sharing	√ Possibility of link with external ressources
√ Possibility of new tools' integration	

Un sous-domaine dédié sur l'instance nationale usegalaxy.fr



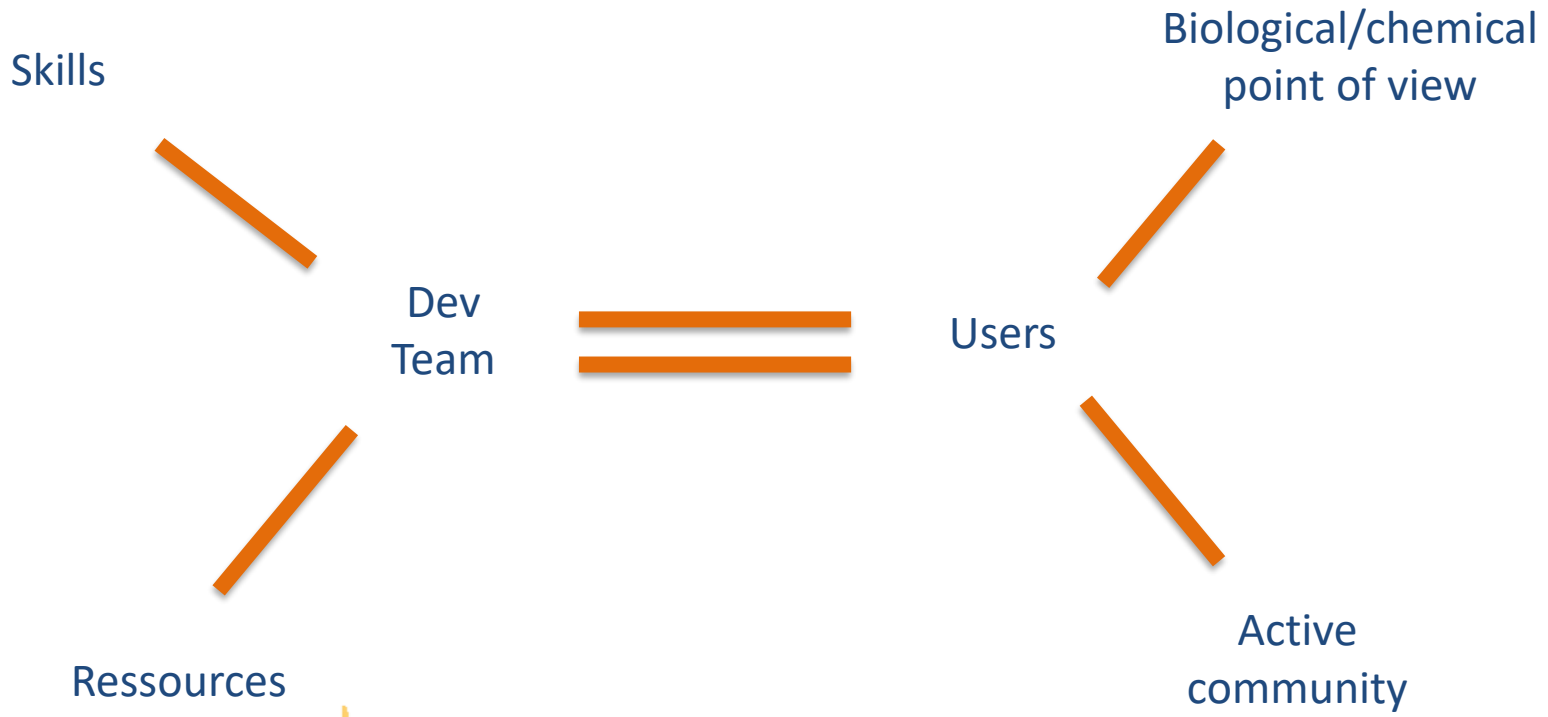
https://workflow4metabolomics.usegalaxy.fr/

UNE COMBINAISON GAGNANTE !

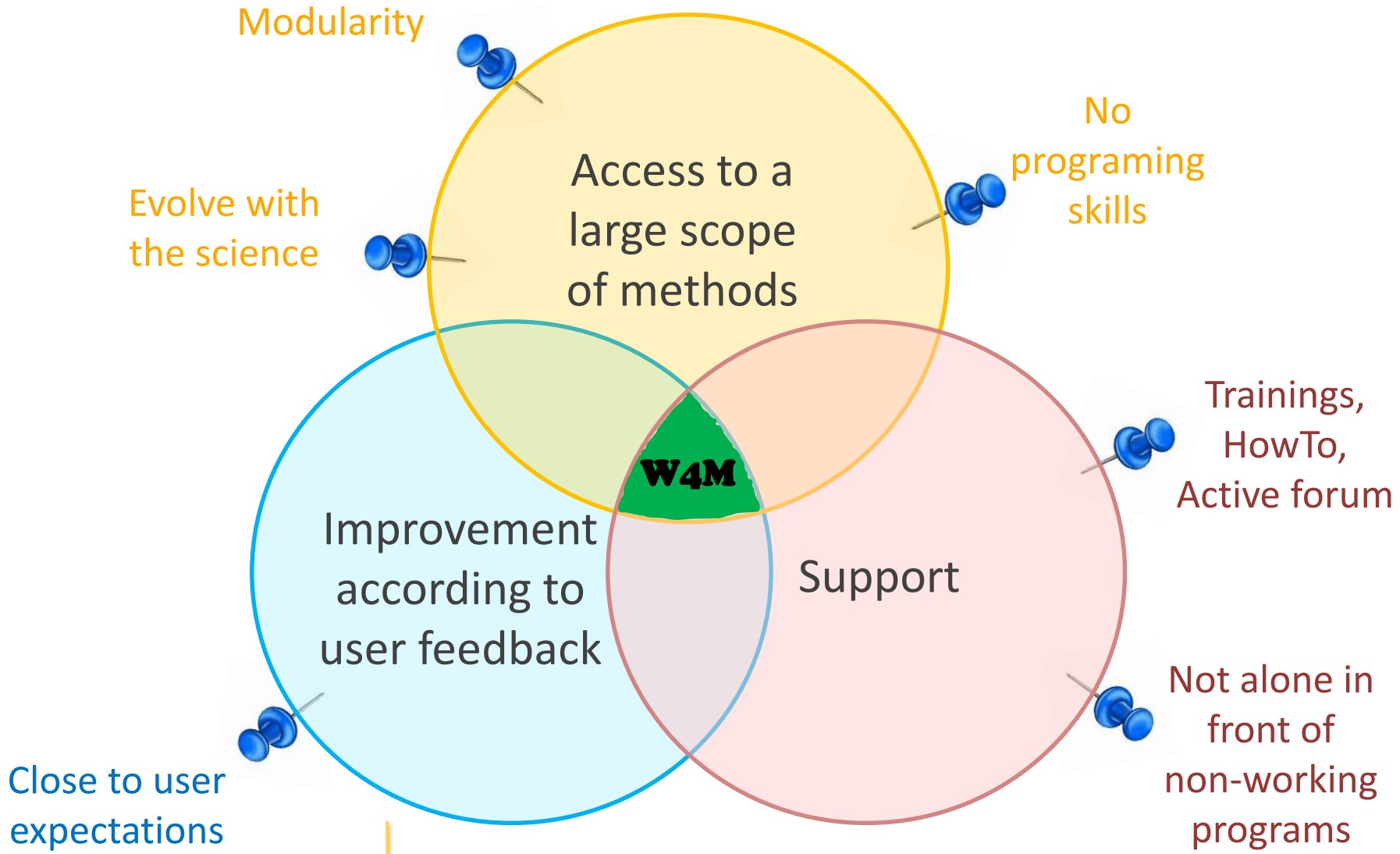
La chimie de W4M

Our equation

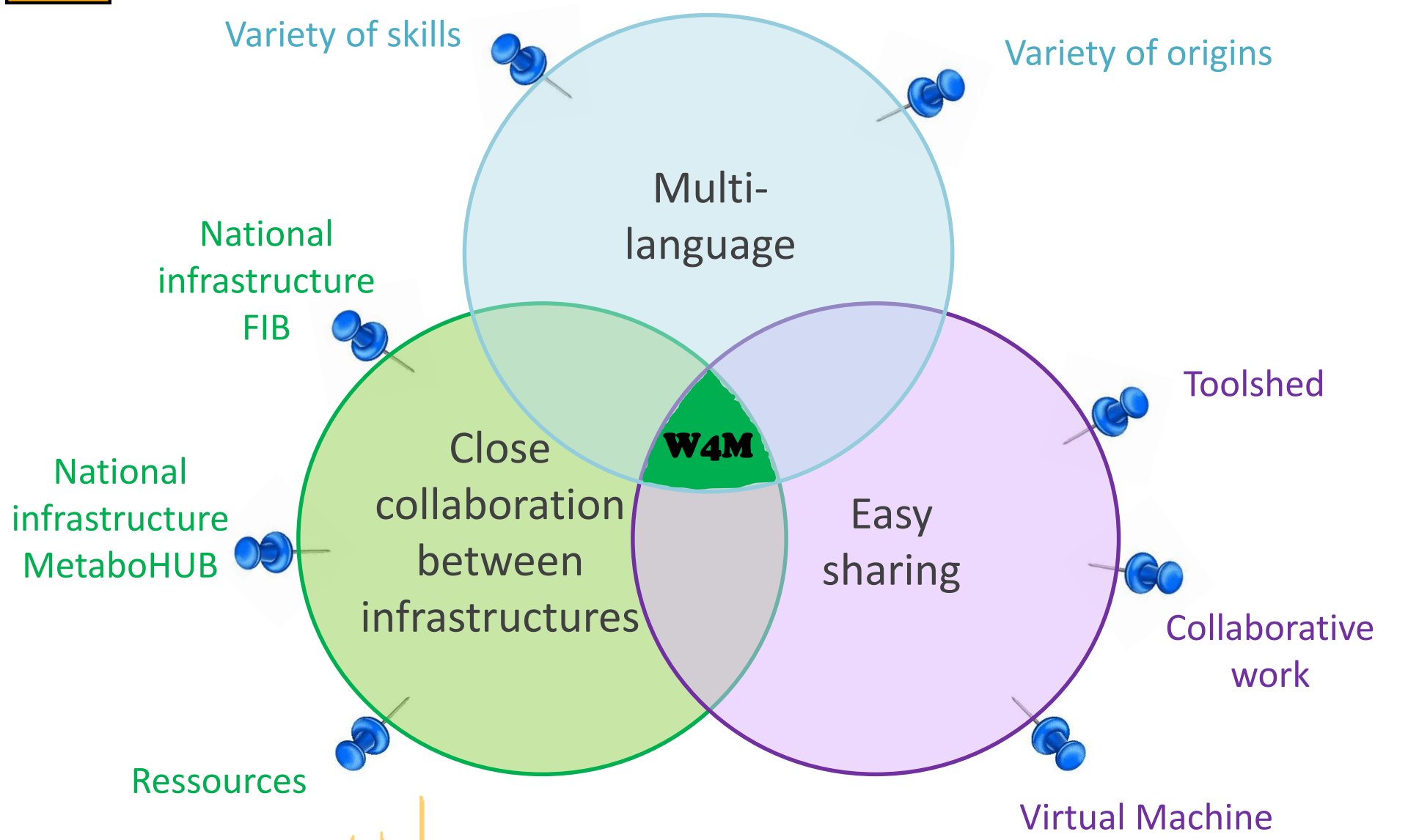
Easy access to existing tools + Development of new tools + Easy handling for users → W4M



Efficace pour les utilisateurs



Efficace pour les développeurs



Pour aller plus loin

https://workflow4metabolomics.org/

Bioinformatics Advance Access published January 16, 2015

Bioinformatics, 2015, 1–3
doi: 10.1093/bioinformatics/btu813
Advance Access Publication Date: 19 December 2014
Applications Note

Gene expression

Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics

Franck Giacomoni[1,†], Gildas Le Corguillé[2,†], Misharl Monsoor[2], Marion Landi[1], Pierre Pericard[2], Mélanie Pétéra[1], Christophe Duperier[1], Marie Tremblay-Franco[3], Jean-François Martin[3], Daniel Jacob[4], Sophie Goulitquer[2], Etienne A. Thévenot[5,*] and Christophe Caron[2,*]

[1]INRA, UMR 1019, PFEM, 63122 Saint Genes Champanelle, [2]CNRS, UPMC, FR2424, ABiMS, Station Biologique, 29680 Roscoff, [3]INRA, UMR 1331, PF MetaToul-AXIOM, Toxalim, F-31027 Toulouse, [4]INRA, Metabolome Facility of Bordeaux Functional Genomics Center, IBVM, 33140 Villenave d'Ornon and [5]CEA, LIST, Laboratory for Data Analysis and Smart Systems (LADIS), MetaboHUB Paris, F-91191 Gif-sur-Yvette, France

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Associate Editor: Ziv Bar-Joseph

Received on August 22, 2014; revised on November 2, 2014; accepted on December 4, 2014

Abstract

Summary: The complex, rapidly evolving field of computational metabolomics calls for collaborative infrastructures where the large volume of new algorithms for data pre-processing, statistical analysis and annotation can be readily integrated whatever the language, evaluated on reference datasets and chained to build ad hoc workflows for users. We have developed Workflow4Metabolomics (W4M), the first fully open-source and collaborative online platform for computational metabolomics. W4M is a virtual research environment built upon the Galaxy web-based platform technology. It enables ergonomic integration, exchange and running of individual modules and workflows. Alternatively, the whole W4M framework and computational tools can be downloaded as a virtual machine for local installation.

Availability and implementation: http://workflow4metabolomics.org homepage enables users to open a private account and access the infrastructure.

W4M is developed and maintained by the French Bioinformatics Institute (IFB) and the French Metabolomics and Fluxomics Infrastructure (MetaboHUB).

Contact: contact@workflow4metabolomics.org

1 Introduction

Metabolomics, the high throughput analysis of small molecules in biological samples, heavily depends on data pre-processing, statistical analysis and chemical and biological annotation, which are complex, transdisciplinary processes involving both computation volume of proposed algorithms written in various languages, making their evaluation by the bioinformatics community (including reviewers) and their chaining within ad hoc workflows by experimenters difficult (Smith *et al.*, 2013).

A few user-oriented online platforms for metabolomics data pre-

Contents lists available at ScienceDirect

International Journal of Biochemistry and Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Create, run, share, publish, and reference your LC–MS, FIA–MS, GC–MS, and NMR data analysis workflows with the Workflow4Metabolomics 3.0 Galaxy online infrastructure for metabolomics

Yann Guitton[a,1], Marie Tremblay-Franco[b,1], Gildas Le Corguillé[c], Jean-François Martin[b], Mélanie Pétéra[d], Pierrick Roger-Mele[e], Alexis Delabrière[e], Sophie Goulitquer[f], Misharl Monsoor[c], Christophe Duperier[d], Cécile Canlet[b], Rémi Servien[b], Patrick Tardivel[b], Christophe Caron[g], Franck Giacomoni[d,*,2], Etienne A. Thévenot[e,*,2]

[a] *LUNAM Université, Oniris, Laboratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA), Nantes, F-44307, France*
[b] *Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, MetaboHUB, Toulouse, France*
[c] *UPMC, CNRS, FR2424, ABiMS, Station Biologique, 29680, Roscoff, France*
[d] *INRA, UMR 1019, PFEM, MetaboHUB, 63122, Saint Genes Champanelle, France*
[e] *CEA, LIST, Laboratory for Data Analysis and Systems' Intelligence, MetaboHUB, F-91191 Gif-sur-Yvette, France*
[f] *INSERM-UBO UMR1078-ECLA, IBSAM, Faculty of Medicine, University of Brest, 29200 Brest, France*
[g] *INRA, Ingenium, Toulouse, France*

ARTICLE INFO

Keywords:
Metabolomics
Data analysis
E-infrastructure
Workflow
Galaxy
Repository

ABSTRACT

Metabolomics is a key approach in modern functional genomics and systems biology. Due to the complexity of metabolomics data, the variety of experimental designs, and the multiplicity of bioinformatics tools, providing experimenters with a simple and efficient resource to conduct comprehensive and rigorous analysis of their data is of utmost importance. In 2014, we launched the Workflow4Metabolomics (W4M; http://workflow4metabolomics.org) online infrastructure for metabolomics built on the Galaxy environment, which offers user-friendly features to build and run data analysis workflows including preprocessing, statistical analysis, and annotation steps. Here we present the new W4M 3.0 release, which contains twice as many tools as the first version, and provides two features which are, to our knowledge, unique among online resources. First, data from the four major metabolomics technologies (i.e., LC–MS, FIA–MS, GC–MS, and NMR) can be analyzed on a single platform. By using three studies in human physiology, alga evolution, and animal toxicology, we demonstrate how the 40 available tools can be easily combined to address biological issues. Second, the full analysis (including the workflow, the parameter values, the input data and output results) can be referenced with a permanent digital object identifier (DOI). Publication of data analyses is of major importance for robust and reproducible science. Furthermore, the publicly shared workflows are of high-value for e-learning and training. The Workflow4Metabolomics 3.0 e-infrastructure thus not only offers a unique online environment for analysis of data from the main metabolomics technologies, but it is also the first reference repository for metabolomics workflows.