



**HAL**  
open science

# Utiliser des données méta-omiques pour piloter les procédés de méthanisation : le projet DeepOmics Ariane Bize

Ariane Bize

## ► To cite this version:

Ariane Bize. Utiliser des données méta-omiques pour piloter les procédés de méthanisation : le projet DeepOmics Ariane Bize. JRI Biogaz méthanisation 2020 (journées recherche innovation), CTBM (ATEE), Sep 2020, Toulouse, France. hal-04362402

**HAL Id: hal-04362402**

**<https://hal.inrae.fr/hal-04362402>**

Submitted on 22 Dec 2023

**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.



# Utiliser des données méta-omiques pour piloter les procédés de méthanisation : le projet DeepOmics



Ariane Bize  
[ariane.bize@inrae.fr](mailto:ariane.bize@inrae.fr)

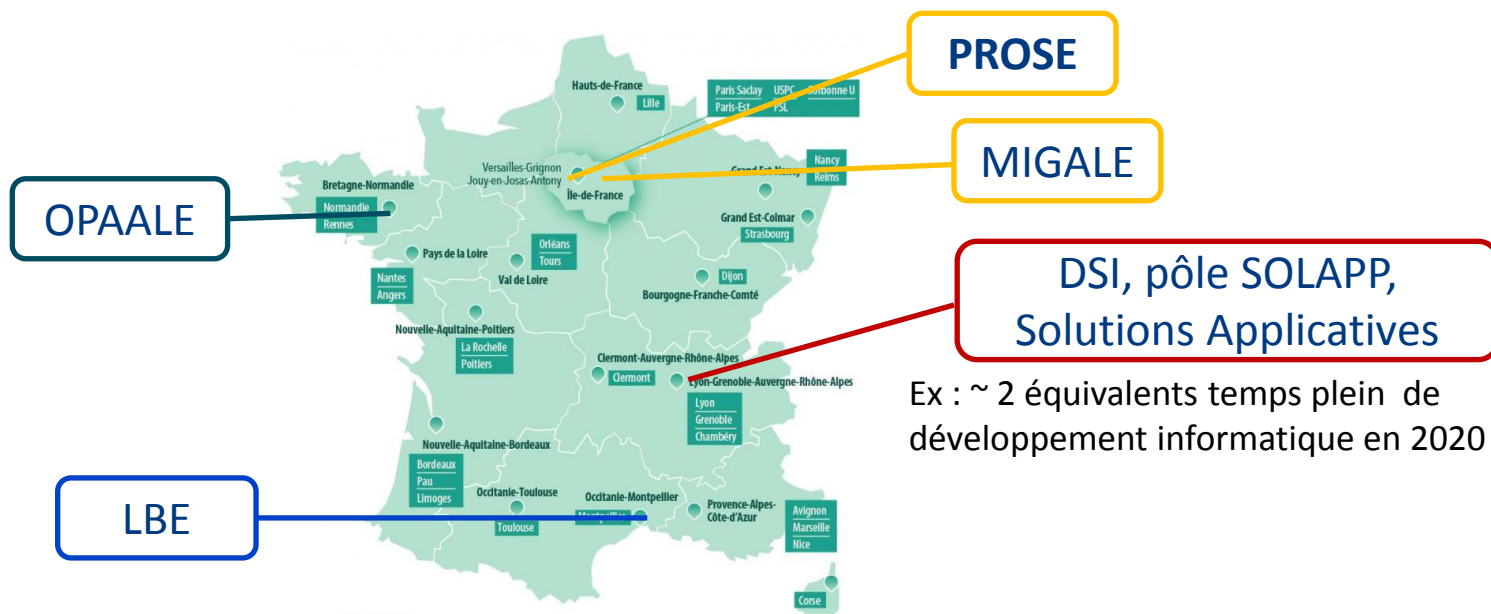


*en partenariat avec*



# DeepOmics: un projet collaboratif INRAE

- **Projet INRAE** PROSE Antony, DSI Lyon, LBE Narbonne, OPAALE Rennes



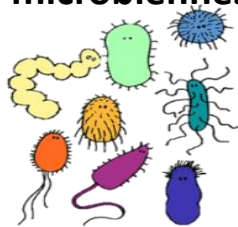
Ex : ~ 2 équivalents temps plein de développement informatique en 2020

- Perspectives : poursuivre le développement informatique et opérationnel de DeepOmics
- Etendre les collaborations à des partenaires extérieurs

# Lier les performances des procédés au fonctionnement des communautés microbiennes

...

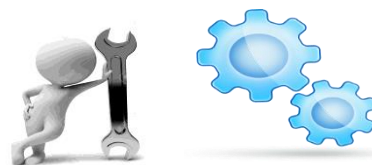
Structure, diversité et fonctions métaboliques des communautés microbiennes



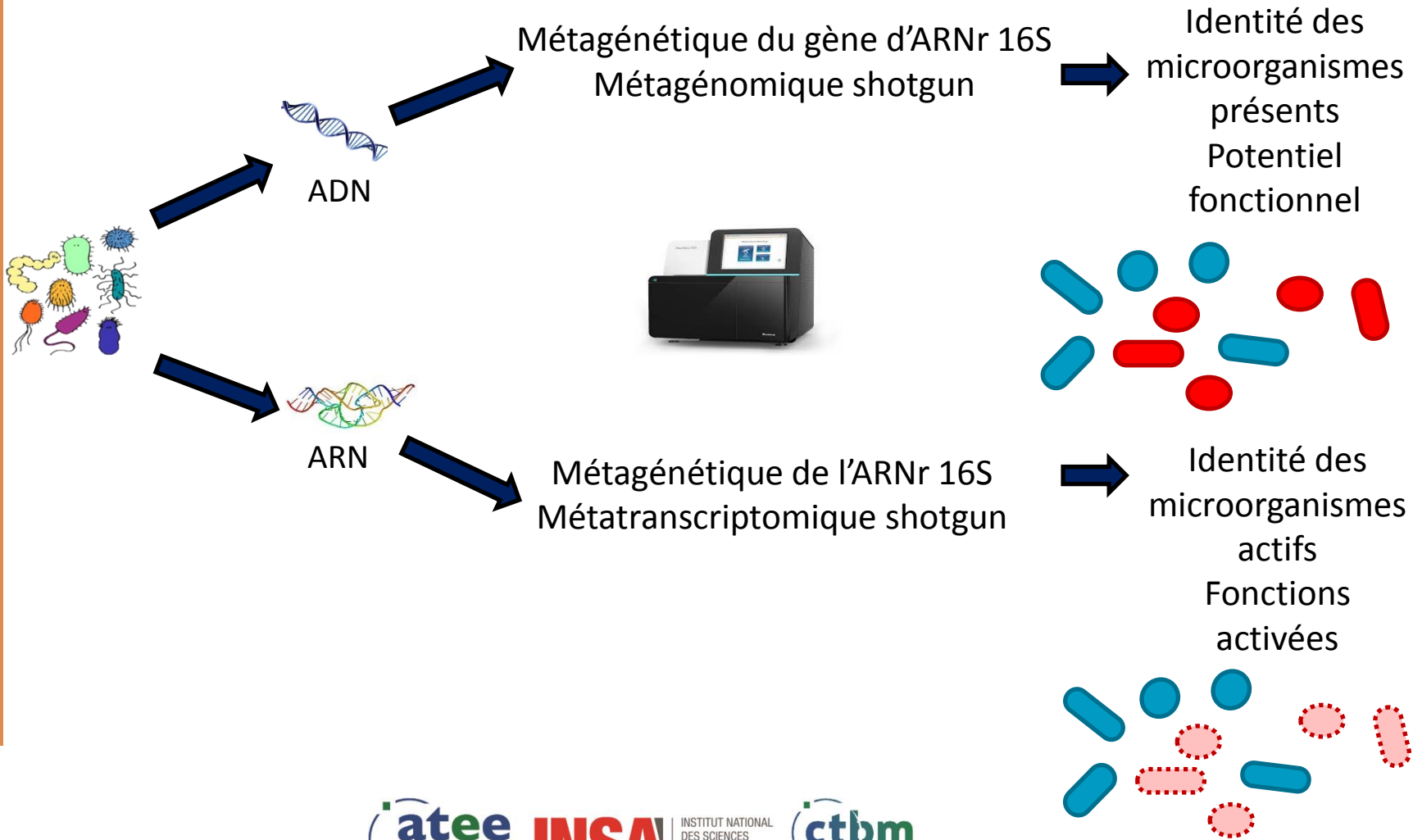
Performances et fonctions du procédé



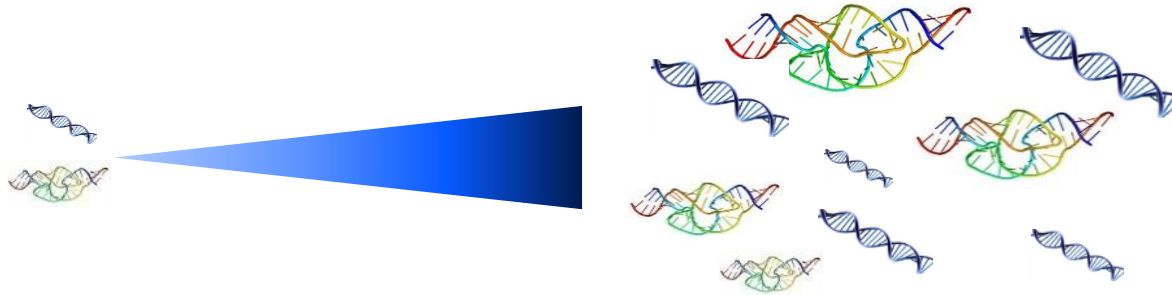
Paramètre opérationnels  
Design du procédé



# ... à l'aide de méthodes « méta-omiques »



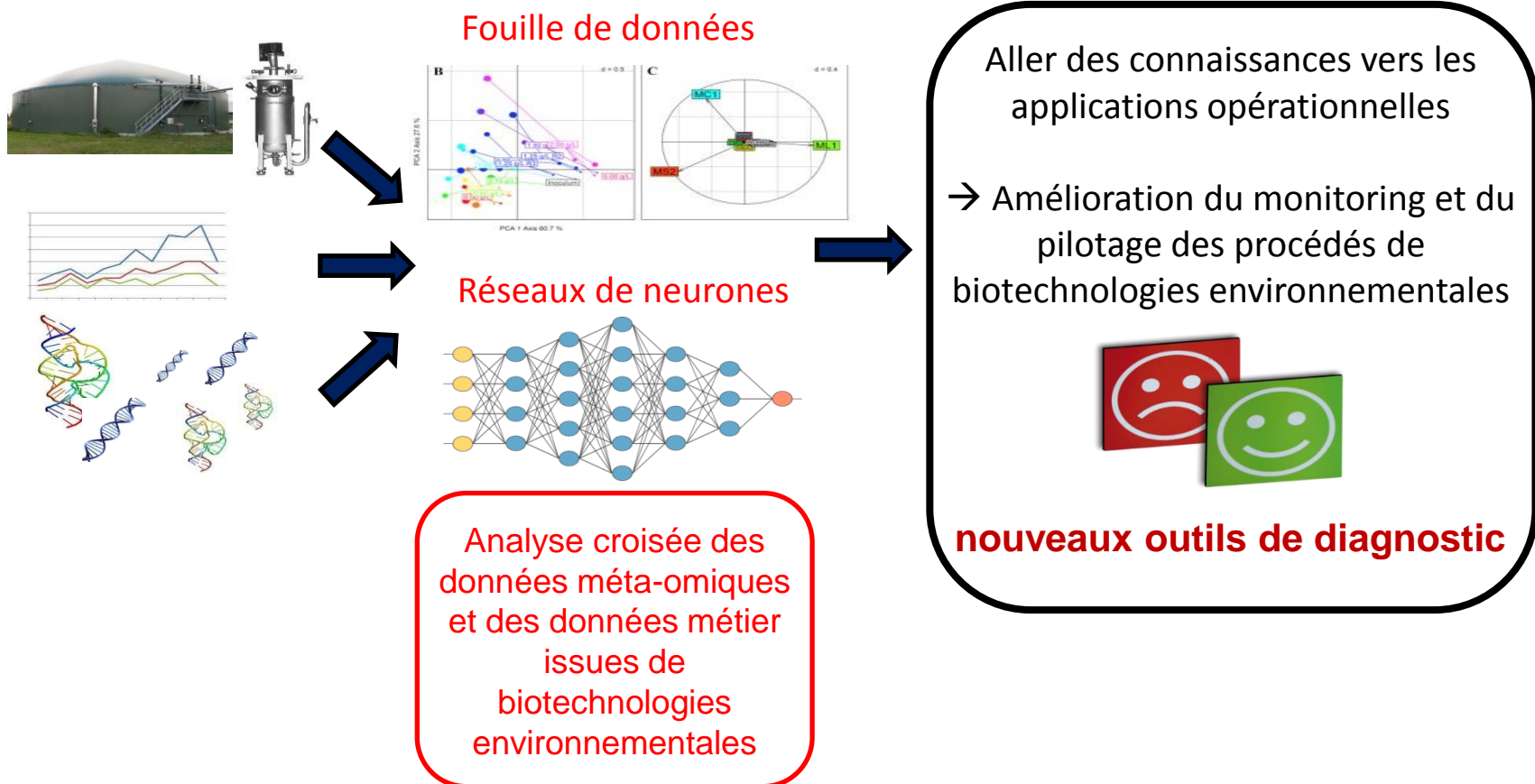
# Les données méta-omiques s'accumulent rapidement ...



mais les jeux de données sont souvent analysés de manière isolée, à l'échelle de chaque projet



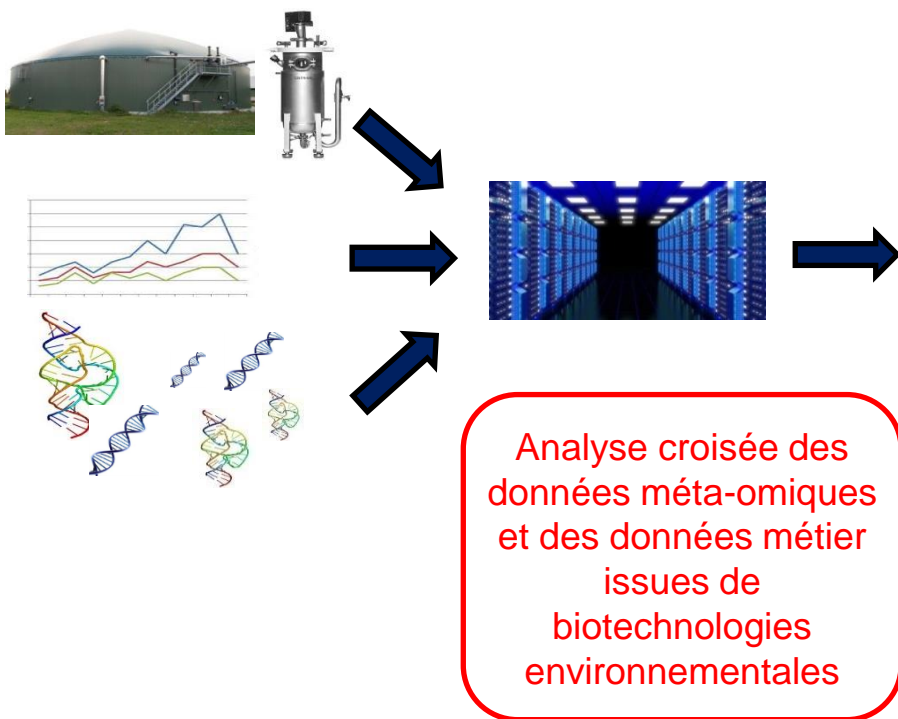
# Pourrions-nous capitaliser sur ce type de données à plus large échelle ?





# Le projet DeepOmics

- “Digital Environmental Engineering Platform for OMICS data”
- Développement d’un entrepôt dédié aux données méta-omiques de procédés de méthanisation et d’autres biotechnologies environnementales

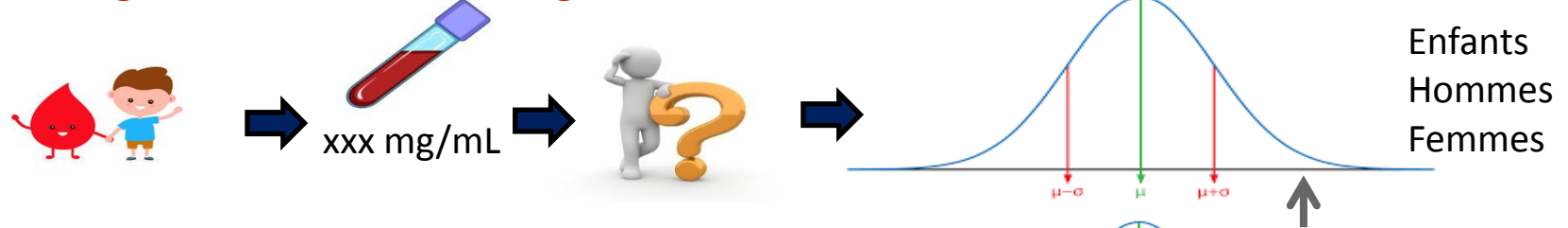


- Entreposage rationalisé des données méta-omiques
- Et des métadonnées métier associées, structurées
- Requêtes au sein des données et méta-données
- Autres fonctionnalités conviviales pour favoriser la préparation de jeux de données pour des méta-analyses (fouille de données, ...)

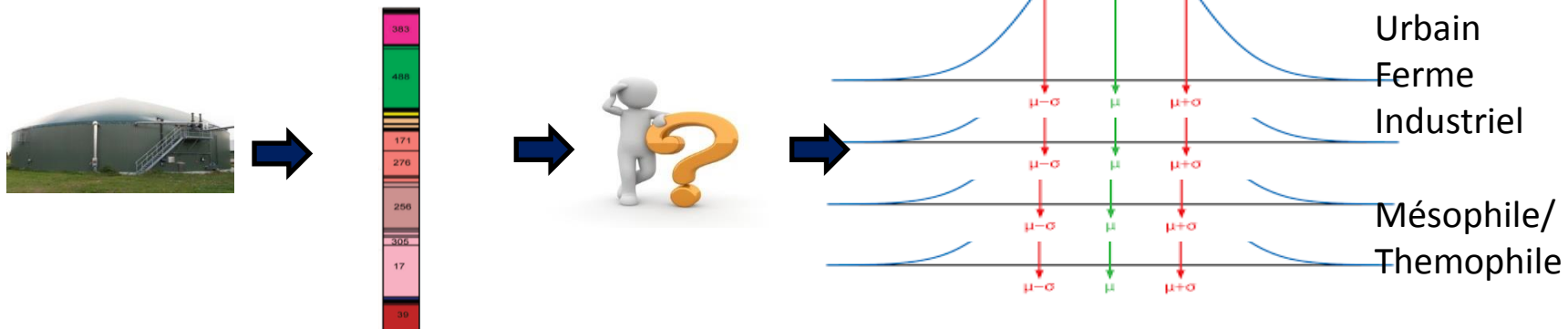


# Le projet DeepOmics

## Diagnostic médical, test sanguin

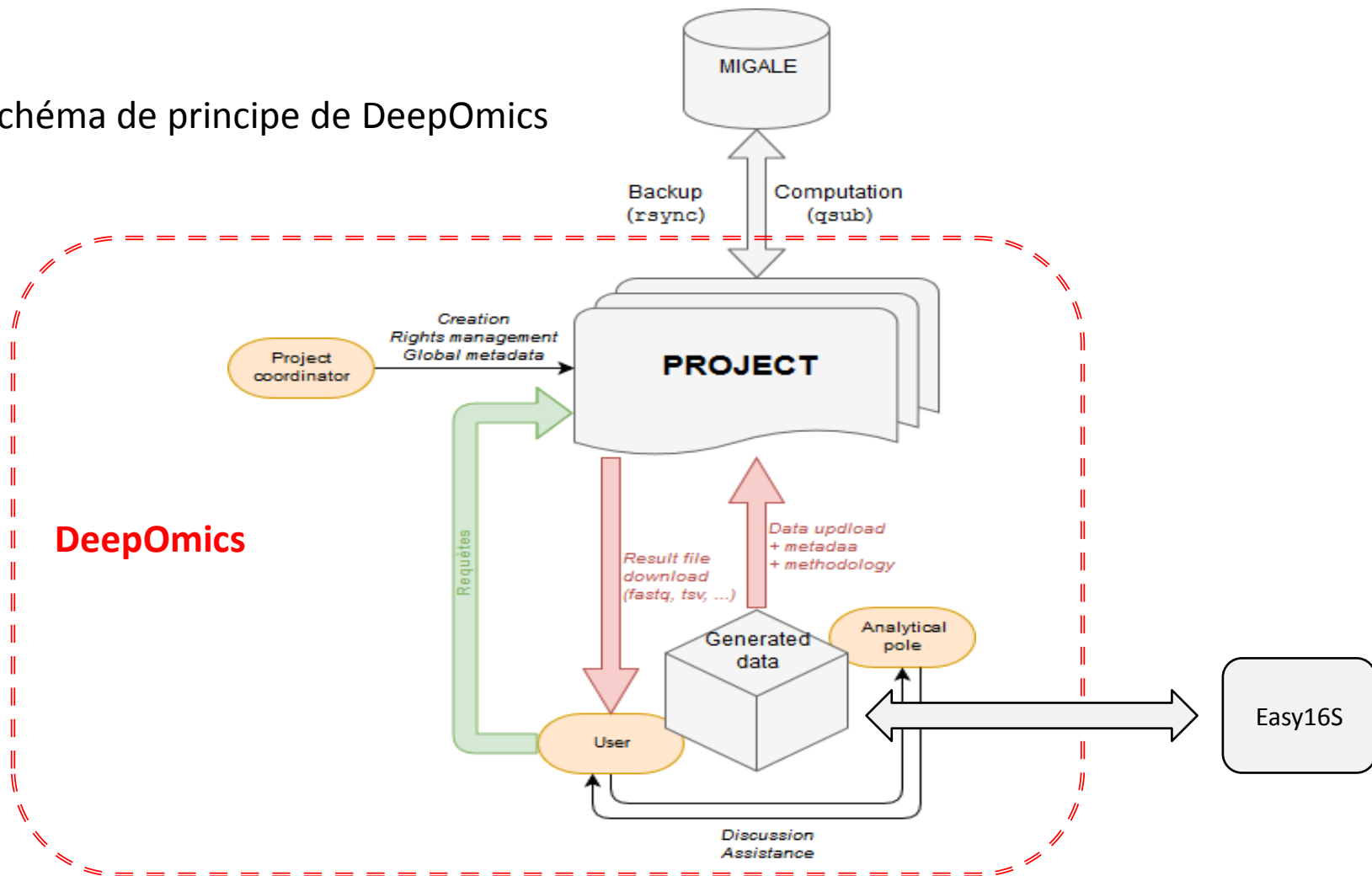


## Diagnostic de la méthanisation, analyses « 16S »



# Le projet DeepOmics

Schéma de principe de DeepOmics



# Le projet DeepOmics

## Objectifs

- Fournir un outil pour entreposer de manière organisée les données métiers et méta-omiques issues de procédés de méthanisation
- Favoriser les démarches de fouille de données
- Prendre en compte des données de laboratoire (pilotes) ou de terrain

<https://deepomics.test.irstea.fr/home>

# Fonctionnalités clés de DeepOmics

- **Entreposage rationalisé** de données méta-omiques et de **métadonnées** issues d'installations industrielles et de réacteurs pilotes, à l'échelle de projets de recherche individuels

The screenshot displays the DeepOmics web application interface. At the top, there is a teal header with the INRAE logo and the text 'DeepOmics'. On the right side of the header, the user profile 'ariane.bize' is visible. Below the header is a navigation bar with 'Home', 'Projects', and 'Meta-analysis' options. The main content area shows the 'Project "Cellulose Nelly"' with a 'View details' link. There are three tabs: 'Lab-scale process data' (selected), 'Industrial and field process data', and 'Meta-omics analysis'. On the left, a sidebar contains 'Laboratory' with sub-options for 'Experimental series' and 'Project's input management'. The main area is titled 'List of experimental series' and includes a search bar and a '+ New series' button. Below this is a table with the following data:

	Experimental series	Contact person	Start date	Operators
>	<a href="#">LignocellulosicSubstrates</a> Lignocellulosic substrates (complex substrates)	Ariane BIZE D'AVEZAC DE CASTERA	2011-01-15	Ariane BIZE D'AVEZAC DE CASTERA
>	<a href="#">CellulosicSubstrates</a> Cellulosic substrates (simple substrates)	Marguerite BAZIN	2012-08-22	Martin BAILLY

# Fonctionnalités clés de DeepOmics

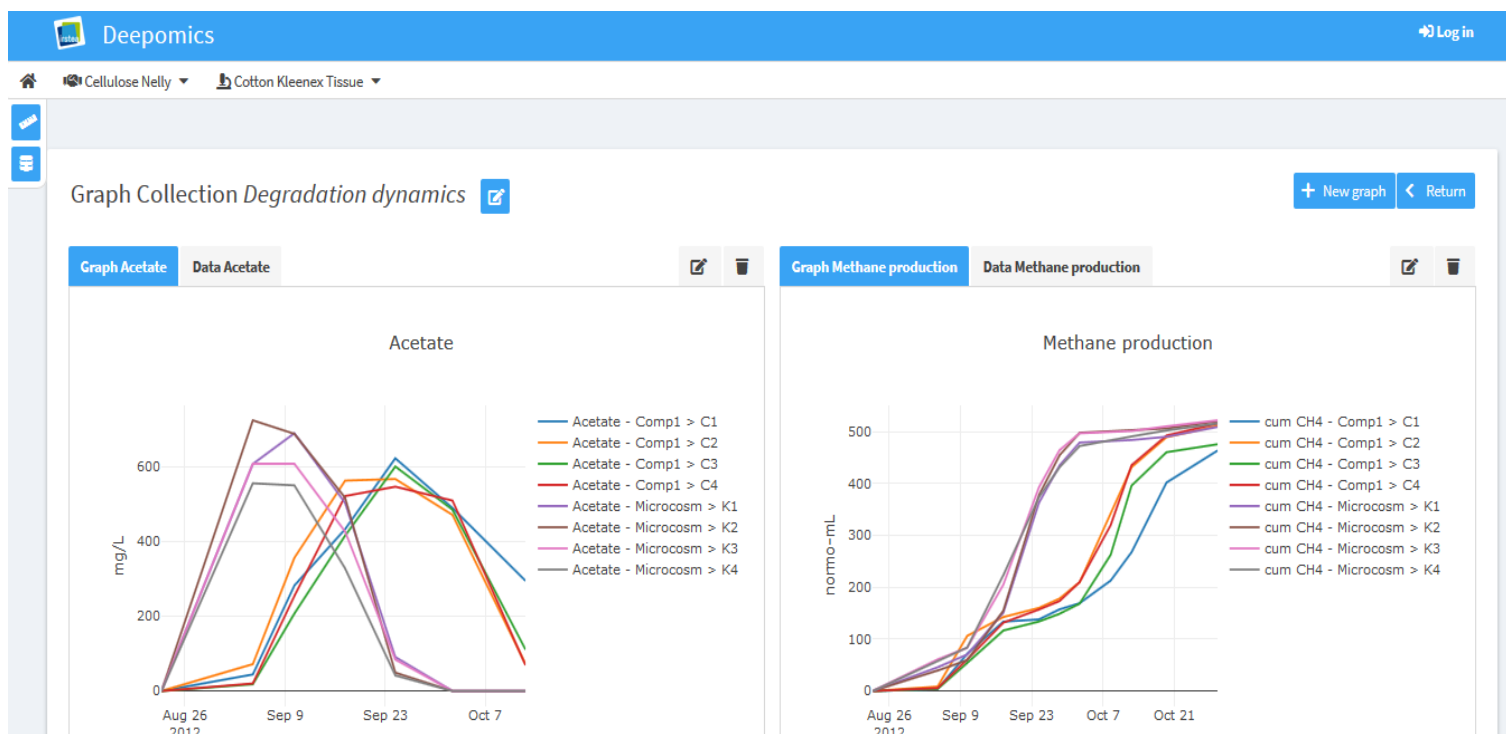
- **Import et export de données** dans des formats standards (biom, fastq, fasta, csv, ... selon la nature des données)

The screenshot shows the DeepOmics web interface. At the top, there is a blue header with the 'Deepomics' logo and a 'Log in' button. Below the header, there are navigation elements including a home icon, a dropdown menu for 'Cellulose Nelly', and another dropdown for 'Cotton Kleenex Tissue'. The main content area is titled 'Data Management' and contains two buttons: '1) Generate csv file' and '2) Import data'. Below these buttons, the text 'Number of Data :' is followed by a table.

Monitored Measure Type	Replicate Location															
	T1	T2	K1	K2	K3	K4	C1	C2	C3	C4	P1	P2	P3	P4		
Acetate concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data		
Butanol concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data		
Butyrate concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data		
Dissolved Inorganic Carbon (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data		

# Fonctionnalités clé de DeepOmics

- **Visualisation** de graphiques en courbes définis par l'utilisateur, basés sur les métadonnées (mesures physico-chimiques), au sein d'un tableau de bord



# Outil complémentaire : Easy16S

- **Analyses statistiques** et visualisation de données de métabarcoding 16S et métagénomiques grâce à l'utilisation de l'application conviviale **Easy16S** (<http://genome.jouy.inra.fr/shiny/easy16S/>)





# Fonctionnalités clé de DeepOmics

- Réaliser des requêtes au sein de données et de méta-données, de réacteurs de laboratoires et d'installations industrielles
- Utiliser la partie publique de l'entrepôt et ses propres données privées
- **Exemple d'analyse**

DeepOmics

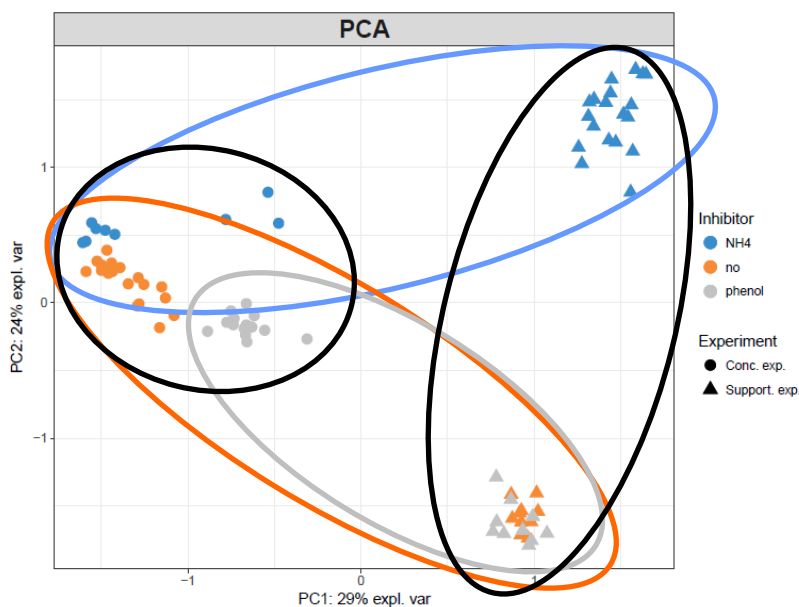
- *Requête dans DeepOmics pour trouver toutes les données de séquençage 16S et toutes les analyses physico-chimiques d'échantillons issus de séries d'expériences pour lesquelles du phénol ou de l'ammonium ont été introduits dans certains des réacteurs*
- *Première exploration statistique avec Easy16S*
- *Export des résultats dans des formats standards*
- *Analyses statistiques spécifiques avec les outils préférés de l'utilisateur*

# Exemple d'application possible

2 séries d'expériences  
 144 échantillons analysés par métabarcoding 16S

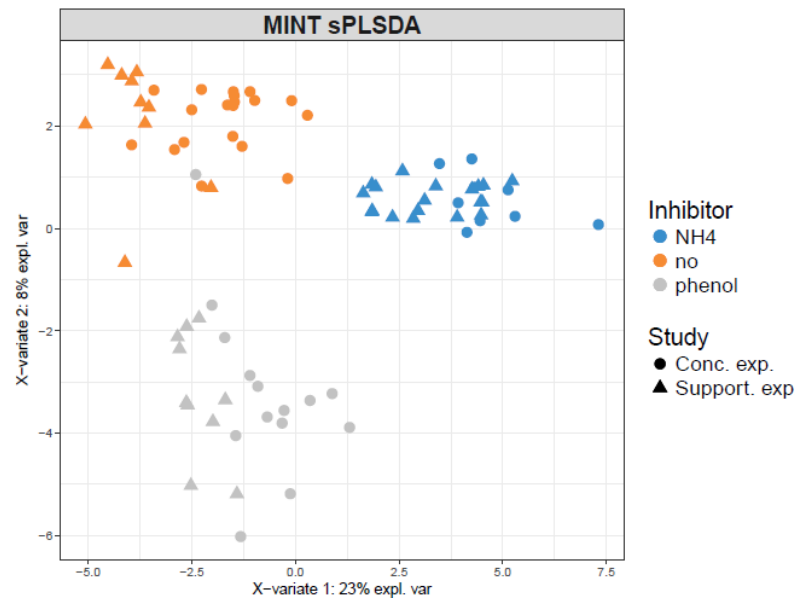


Ordination des échantillons par ACP



Important effet "étude"  
 Pas de possibilité de généraliser ?

Multivariate INTeegrative method, MINT



Correction de l'effet "étude"  
 Les résultats peuvent être généralisés



# Conclusions

- **DeepOmics**
  - Un entrepôt de données dédié au **stockage et à l'analyse croisée** de données méta-omiques et de procédés, sur la méthanisation
  - Un environnement pour préparer des approches de méta-analyse et de fouille de données
  - Un outil pour aider les professionnels à produire des données FAIR\*
- **Notre approche est très complémentaire de la base de données de connaissance "MiDAS", du groupe danois de Per Nielsen**

McIlroy, Simon Jon, et al. "MiDAS: the field guide to the microbes of activated sludge." *Database* (2015).

\* Findable, Accessible, Interoperable, and Reusable

# Merci pour votre attention !

## Remerciements aux différentes personnes impliquées



Théodore Bouchez  
Cédric Midoux  
Yannick Fayolle

Elie Le Quémener  
Virginie Rossard  
Eric Latrille



**DSI**  
**Irstea**  
**puis**  
**INRAE**

Nicolas Raidelet  
Guillaume Perréal  
Aurélie Gramusset  
Manon Predhumeau  
Christophe Biasini  
Harold Boissenin  
Walter Roa Serrano

Patrick Dabert  
Valentin Loux

