



**HAL**  
open science

## Livrables Projet BiogazRIO - WP 3.2

Emilie Fernandez, Eric Latrille, Virginie Rossard, Diana Bernet-Garcia,  
Audrey Battimelli

► **To cite this version:**

Emilie Fernandez, Eric Latrille, Virginie Rossard, Diana Bernet-Garcia, Audrey Battimelli. Livrables Projet BiogazRIO - WP 3.2. Inrae Occitanie Montpellier, 2 place Pierre Viala, 34060 Montpellier. 2022. hal-04673849

**HAL Id: hal-04673849**

**<https://hal.inrae.fr/hal-04673849v1>**

Submitted on 9 Sep 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 4.0 International License

# Projet Biogaz-RIO

## L3.2.1.01

# SYSTÈME D'INFORMATION DE GESTION DES DONNÉES SCIENTIFIQUES

**Rédaction** : Emilie Fernandez

Date : 01.12.2022

**Relecture** : Eric Latrille, Virginie Rossard, Diana Bernet-Garcia, Audrey Battimelli

1. Contexte.....
2. Méthode.....
3. Résultats et Discussion.....
4. Conclusions et recommandations pour le projet.....
5. Bibliographie.....
6. Annexes.....

## Abréviations

FAIR : Findable, Accessible, Interoperable, Reusable

LBE : Laboratoire de Biotechnologies de l'Environnement

Bio2E : Biotechnologie et Bio Raffinerie Environnementales

## 1. CONTEXTE

Le Projet Biogaz RIO(1), Recherche et Innovation en Occitanie, porté par la plateforme Bio2E(2) vise la consolidation des performances des installations de méthanisation pour accompagner la transition énergétique des territoires. Pour cela, les travaux de recherche ciblent la résolution de plusieurs verrous rencontrés lors du développement de projets de méthanisation, avec notamment « le développement de systèmes pour l'acquisition et le traitement de données, le pilotage et la supervision des installations pour optimiser la production de biogaz ».

Dans ce cadre, le principal objectif de la tâche 3.2 est de concevoir et installer un système d'information comprenant une chaîne d'acquisition des données issus des capteurs et des analyses des échantillons.

## 2. MÉTHODE

OpenSILEX(3) est un ensemble de modules logiciels interopérables (structuration, visualisation, analyse, reproductibilité) qui permet la réalisation de systèmes d'information scientifiques dédiés à l'exploitation des données à grande échelle

en agronomie et axé sur des ontologies conçues spécifiquement qui permettent d'implémenter les principes FAIR.

La suite logicielle OpenSILEX est développée par l'UMR MISTEA de l'INRAE sous licence libre. Il s'agit dans ce livrable de s'approprier les méthodes et les outils de cette suite et de construire et configurer un système d'information complet pour gérer l'ensemble des données issues des bio-procédés du LBE-INRAE.

De plus, des développements de nouveaux modules (cf L3.2.2.03), non existants, ont été réalisés afin de compléter la suite logicielle. Ces modules ont été développés sous licence libre.

Une ontologie spécifique nommée Environmental Biorefinery Ontology (cf L3.2.1.02) a également été conçue afin de décrire les différentes approches étudiées dans le laboratoire.

### 3. RÉSULTATS ET DISCUSSION

Un système d'information spécifique aux besoins du LBE a été développé, il est nommé EnviBIS : Environmental Biorefinery Information System.

EnviBIS a été déployé sur deux serveurs :

1. Un serveur de test <sup>[1]</sup> interne afin de pouvoir contrôler la robustesse du logiciel avant la mise en œuvre finale.
2. Un serveur en production <sup>[2]</sup> hébergé dans une machine virtuelle sur le DataCenter INRAE.

Les codes informatiques ont été déposés sous licence libre via un outil de gestion de version, sur le Gitlab <sup>[3]</sup> de l'Institut, et archivés via une infrastructure de la science ouverte, Software Heritage <sup>[4]</sup>.

Plusieurs documentations ont été écrites pour permettre l'installation en mode développement <sup>[5]</sup> ou en mode production <sup>[6]</sup> ainsi qu'une documentation pour les utilisateurs <sup>[7]</sup> de l'outil.

### 4. CONCLUSIONS ET RECOMMANDATIONS POUR LE PROJET

### 5. BIBLIOGRAPHIE

(1) Projet Biogaz RIO (2020-2022), convention n°24001371, soutenu par la Région Occitanie et le FEDER

(2) Bio2E, INRAE, 2018. Plateforme Biotechnologie et Bioraffinerie Environnementales,

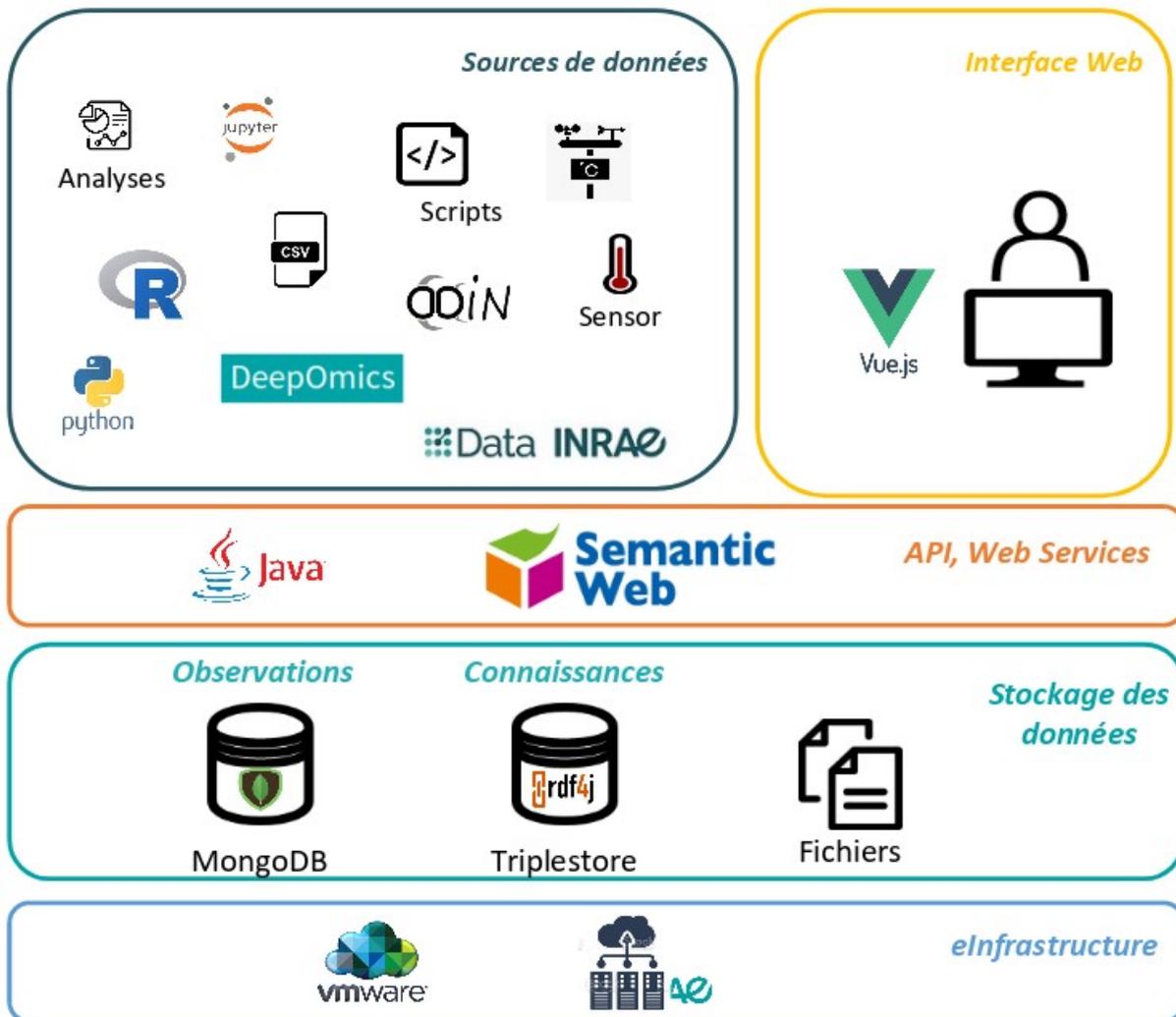
<https://doi.org/10.15454/1.557234103446854E12>

(3) : OpenSILEX : <http://www.opensilex.org/>

### 6. ANNEXES

[1] : Dépôt des codes informatiques : [Forgemia](#)

[2] : Architecture



# Projet Biogaz-RIO

## L3.2.1.02

# MODULES DE TRAITEMENT SPÉCIFIQUE DES DONNÉES

**Rédaction** : Emilie Fernandez

Date : 01.12.2022

**Relecture** : Eric Latrille, Virginie Rossard, Diana Bernet-Garcia, Audrey Battimelli

1. Contexte.....
2. Méthode.....
3. Résultats et Discussion.....
4. Conclusions et recommandations pour le projet.....
5. Bibliographie.....
6. Annexes.....

## Abréviations

LBE : Laboratoire de Biotechnologies de l'Environnement

Bio2E : Biotechnologie et Bio Raffinerie Environnementales

EBO : Environmental Biorefinery Ontology

INRIA : Institut National de Recherche en sciences et technologies du numérique

## 1. CONTEXTE

Le Projet Biogaz RIO(1), Recherche et Innovation en Occitanie, porté par la plateforme Bio2E(2) vise la consolidation des performances des installations de méthanisation pour accompagner la transition énergétique des territoires. Pour cela, les travaux de recherche ciblent la résolution de plusieurs verrous rencontrés lors du développement de projets de méthanisation, avec notamment « le développement de systèmes pour l'acquisition et le traitement de données, le pilotage et la supervision des installations pour optimiser la production de biogaz ».

Dans ce cadre, les principaux objectifs de la tâche 3.2 sont de concevoir et d'installer un système d'information comprenant une chaîne d'acquisition des données issus des capteurs et des analyses des échantillons.

## 2. MÉTHODE

La spécificité du projet est de concevoir un modèle de données, utilisé par le système d'information EnviBIS (cf L3.2.1.01), sous forme d'une ontologie du domaine de la bioraffinerie environnementale. Il s'agit dans ce livrable de

développer cette ontologie en utilisant des vocabulaires contrôlés afin qu'elle devienne une référence.

Ce système d'information permet l'acquisition de données hétérogènes issues de capteurs, d'observations ou d'appareils analytiques. Il s'agit également d'assurer la récupération de ses données pour un suivi en ligne des bio-procédés.

### 3. RÉSULTATS ET DISCUSSION

#### A. ONTOLOGIE

Une ontologie spécifique à la bioraffinerie environnementale nommée Environmental Biorefinery Ontology(3) a été conçue afin de décrire les différentes opérations étudiées au LBE selon les étapes suivantes :

##### 1. Capturer les connaissances :

Il s'agit de l'identification des concepts et des relations du domaine à modéliser, ainsi que les termes associés ont permis de réaliser un thésaurus des composants ainsi que des listes de types de classes pour les capteurs, installations et étapes de procédés.

##### 2. Coder l'ontologie :

Cela est réalisé en modélisant les connaissances définies à l'étape précédente dans un langage de formalisme d'ontologies [1].

##### 3. Intégrer des ontologies existantes :

Le système d'information EnviBIS est piloté par un ensemble d'ontologies. L'ontologie centrale d'OpenSILEX(4) s'appuie sur des ontologies de hauts niveaux et a été étendue avec des ontologies de domaine [2].

#### B. ODIN+

L'intégration des données issues de capteurs se fait via une autre brique logicielle, de la nouvelle version du logiciel d'acquisition de données en ligne ODIN+(5) développé par l'INRIA, qui a été testée et installée.

#### C. SCRIPTS

L'insertion des données en ligne a nécessité l'écriture de scripts utilisant les services web du système d'information EnviBIS. L'hétérogénéité des sources a conduit à trois modes de récupération différents :

1. Récupération de fichiers plantml
2. Récupération de fichiers CSV
3. Récupération via des requêtes SQL dans une base de donnée Postgres

Les codes informatiques ont été déposés via un outil de gestion de version, sur le Gitlab [3] de l'Institut.

### 4. CONCLUSIONS ET RECOMMANDATIONS POUR LE PROJET

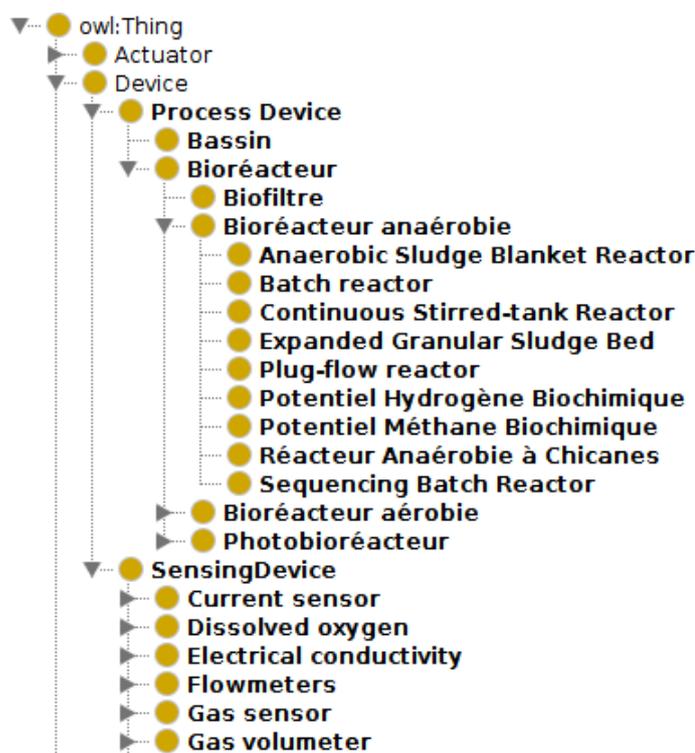
Des alignements d'ontologies, notamment entre EBO et PO2(6) facilitera, à court terme, les échanges d'information entre les deux systèmes d'informations et permettra à plus long terme à la communauté scientifique du domaine de : publier, partager et échanger des données à l'aide des outils du web sémantique. L'utilisation d'ontologies de référence permettra l'exploitation des données de la bioraffinerie environnementale produites par l'unité INRAE LBE et plus généralement de données expérimentales de procédés de transformation produites sur l'ensemble du département de recherche TRANSFORM d'INRAE.

## 5. BIBLIOGRAPHIE

- (1) Projet Biogaz RIO (2020-2022), convention n°24001371, soutenu par la Région Occitanie et le FEDER
- (2) Bio2E, INRAE, 2018. Plateforme Biotechnologie et Bioraffinerie Environnementales, <https://doi.org/10.15454/1.557234103446854E12>
- (3) EBO: [Agroportal](#)
- (4) OpenSILEX: [Github](#)
- (5) Odin : [Inria](#)
- (6) J. Dibie, S. Dervaux, L. Ibanescu, J. Raad (2018) . "Process and Observation Ontology", [10.15454/1.4702114192525708E12](https://doi.org/10.15454/1.4702114192525708E12), Portail Data INRAE, V4

## 6. ANNEXES

[1] : Aperçu EBO via l'outil Protégé



[2] : Récapitulatif des ontologies présentes dans EnviBIS

TYPE	SOUS TYPE	SIGLE	NOM	SPECIFICITÉ
Ontologies de référence	Ontologie de haut niveau	BFO	Basic Formal Ontology	Haut niveau d'abstraction et de généralité
	Ontologies standards	Time	Time Ontology	Description des propriétés temporelles
		DC	Dublin Core	Description des documents
		SOSA/SSN	Semantic Sensor Network Ontology	Description des capteurs et observations
		FOAF	Friend Of A Friend	Description des personnes et leurs relations
		OA	Open Annotation	Annoter, créer des associations entre des informations distinctes
		PROV-O	PROV-Ontology	Description de la provenance d'une donnée
		NCBI	National Center for Biotechnology Information	Classification et nomenclature organisées pour tous les organismes
Ontologies d'application	Ontologie centrale	OESO	The Object Ontology	Fonctionnement Opensilex
	Ontologies de domaine	EBO	Environmental Biorefinery Ontology	Description des procédés de bioraffinerie environnementale
		PO2	Process and Observation	Description de l'itinéraires d'un procédé de transformation

[3] : Dépôt des codes informatiques : [Forgemia](#)

# Projet Biogaz-RIO

## L3.2.2.03

# NOUVEAU MODULE DE LA SUITE LOGICIELLE SILEX

**Rédaction** : Emilie Fernandez

Date : 01.12.2022

**Relecture** : Eric Latrille, Virginie Rossard, Diana Bernet-Garcia, Audrey Battimelli

1. Contexte.....
2. Méthode.....
3. Résultats et Discussion.....
4. Conclusions et recommandations pour le projet.....
5. Bibliographie.....
6. Annexes.....

## Abréviations

### 1. CONTEXTE

Le Projet Biogaz RIO(1), Recherche et Innovation en Occitanie, porté par la plateforme Bio2E(2) vise la consolidation des performances des installations de méthanisation pour accompagner la transition énergétique des territoires. Pour cela, les travaux de recherche ciblent la résolution de plusieurs verrous rencontrés lors du développement de projets de méthanisation, avec notamment « le développement de systèmes pour l'acquisition et le traitement de données, le pilotage et la supervision des installations pour optimiser la production de biogaz ».

Dans ce cadre, le principal objectif de la tâche 3.2 est de concevoir et installer un système d'information comprenant une chaîne d'acquisition des données issus des capteurs et des analyses des échantillons.

### 2. MÉTHODE

Le démarrage des développements de la nouvelle version de la suite logicielle OpenSILEX(3) a eu lieu en mars 2020, aboutissant à une version bêta test en avril 2021 et une version release candidate fin 2022.

Il s'agit dans ce livrable de réaliser des développements de nouveaux modules, non existants, afin de compléter cette suite logicielle. Ces modules ont été développés sous licence libre.

Pour cela, l'intégration à la cellule Opensilex était nécessaire. L'organisation de l'équipe se fait selon la méthode agile scrum [1]. Les développements sont suivis

par des outils de gestion de projet tel qu'un tableau de bord Trello et un logiciel de version, le Gitlab [3] de l'Institut. Ces outils ont permis de suivre l'évolution, la maintenance corrective et les mises à jour des différents modules. La rédaction de la documentation et celle des scripts de sauvegardes des données ont été réalisées en parallèle.

### 3. RÉSULTATS ET DISCUSSION

Les modules réalisés sont découpés selon deux objectifs :

- Une première partie a été développée pour être réutilisée par l'ensemble des utilisateurs de la suite logicielle OpenSILEX, quelle que soit la spécialité des données intégrées.
- Une deuxième partie a été développée pour être utilisée par des utilisateurs ayant des données liées à des procédés de transformation.

#### A. OPENSILEX

Un module de gestion de document, s'appuyant sur une ontologie de référence DublinCore(4), qui permet la description des documents à l'aide d'un vocabulaire standardisé et leur stockage.

Une API et des web services ont été développés pour permettre l'interopérabilité avec d'autres systèmes. Les tests unitaires ont été écrits pour garantir une robustesse des codes. Une interface utilisateur a été développée dans le respect de l'ergonomie, adaptée à tout type d'écran et en respectant une homogénéité avec l'ensemble de l'application.

Un module de gestion des dispositifs de mesure s'appuyant sur des ontologies de référence Sensor, Observation, Sample, and Actuator Ontology(5). Une interface utilisateur a été développée dans le respect de l'ergonomie, adaptée à tout type d'écran et en respectant une homogénéité avec l'ensemble de l'application.

#### B. ENVIBIS

Un module de gestion des procédés de transformation, en utilisant certains concepts de l'ontologie Process and Observation(6) (cf L3.2.1.02) permet d'offrir une représentation des enchaînements d'opérations au sein d'un procédé.

Un module de gestion des étapes permet de détailler les composants d'entrées et de sorties d'une opération unitaire ainsi que ses modalités au sein d'un procédé.

Une API et des web services ont été développés pour ces deux modules afin de permettre l'interopérabilité avec d'autres systèmes. Les tests unitaires ont été écrits pour garantir une robustesse des codes. Une interface utilisateur a été développée dans le respect de l'ergonomie, adaptées à tout type d'écran et en respectant une homogénéité avec l'ensemble de l'application.

### 4. CONCLUSIONS ET RECOMMANDATIONS POUR LE PROJET

## 5. BIBLIOGRAPHIE

(1) Projet Biogaz RIO (2020-2022), convention n°24001371, soutenu par la Région Occitanie et le FEDER

(2) Bio2E, INRAE, 2018. Plateforme Biotechnologie et Bioraffinerie Environnementales,

<https://doi.org/10.15454/1.557234103446854E12>

(3) [OpenSILEX](#)

(4) [Dublin Core](#)(5) [SOSA \(Semantic Sensor Network Ontology\)](#)

(6) J. Dibie, S. Dervaux, L. Ibanescu, J. Raad (2018) . "Process and Observation Ontology", <[10.15454/1.470211](#)>, Portail Data INRAE, V4

## 6. ANNEXES

Réunion	Qui ?	Fréquence ?	Quoi ?
Mêlée	MISTEA 11 personnes : Chef de projet Développeurs Testeurs	Tous les matins	2 minutes par personne : présenter ce qui a été fait la veille, prévoir ce qui va être fait ce jour, identifier les problèmes
Point d'avancement		Une fois / semaine	Démonstration, discussion, retards possibles
Retrospective		Une fois / sprint*	Discussion sur le sprint* passé : météo, point positif, point négatif, amélioration
Planification		Une fois / sprint*	Définition des objectifs à réaliser sur le prochain sprint*
Point d'avancement	LBE	Une fois / semaine	Discussion, test
Réunion de travail		Une fois / mois	Réunion avec le service informatique du LBE : tâches à faire/accomplies
Point d'avancement		Une fois / trimestre	Point d'avancement du projet <a href="#">BiogazRio</a>

\* période d'un mois environ

Figure 1: [1] : Méthode scrum

[2] : [Forgemia](#)