

## **AIC TransformON (2021-2022) : Construire une ontologie sur les procédés alimentaires et non alimentaires pour permettre une meilleure interopérabilité des données et des outils développés pour l'intégration de connaissances au sein du Département TRANSFORM.**

Magalie Weber<sup>1</sup>, Liliana Ibanescu<sup>2</sup>, Patrice Buche<sup>3</sup>, Stéphane Dervaux<sup>2</sup>, Hervé Guillemain<sup>4</sup>, Caroline Pénicaud<sup>5</sup>, Elisabeth Guichard<sup>6</sup>, Sophie Aubin<sup>7</sup>.

<sup>1</sup> INRAE, UR BIA ; <sup>2</sup> UMR MIA-Paris ; <sup>3</sup> UMR IATE ; <sup>4</sup> Plateforme Plastic <sup>5</sup> UMR SayFood, <sup>6</sup> CSGA ; <sup>7</sup> DipSO

### **Contexte**

Le Département TRANSFORM (Sciences pour l'ingénierie des aliments, des produits biosourcés et des résidus de l'activité humaine) a été créé au 1er janvier 2020 au sein d'INRAE. Ses recherches s'articulent autour de deux grands axes structurants qui sont la promotion d'une alimentation saine et durable et la bioéconomie.

L'AIC « TransformON » porte une action collective visant à construire un référentiel termino-ontologique spécifique aux connaissances expertes produites par le Département TRANSFORM dans le cadre du Nexus Santé-Environnement-Agriculture-Alimentation.

Cette démarche s'inscrit dans le processus global de la science ouverte (Open Science/ Open Data), impliquant la mise en œuvre des principes FAIR (Findable, Accessible, Interoperable, Reusable) et les technologies du web sémantique.

### **Objectifs**

La finalité de cette AIC est de permettre une meilleure interopérabilité des données et des outils développés dans le département TRANSFORM, afin de pouvoir intégrer des données hétérogènes et améliorer les échanges entre bases de données ayant des structurations différentes, facilitant ainsi notre capacité à mobiliser les connaissances acquises sur la complexité des bioressources, les procédés/bioprocédés de transformation et les systèmes alimentaires.

Notre objectif est double :

- vérifier qu'il est possible de construire un vocabulaire contrôlé unique, unifié, bilingue (français/anglais), permettant de couvrir l'ensemble des domaines de connaissances générées par TRANSFORM sur les itinéraires de construction/déconstruction de la qualité des aliments et des bioproduits, en lien avec les bénéfices/risques pour la santé humaine et l'environnement
- proposer une méthodologie adaptée pour maintenir et faire évoluer ce référentiel de façon incrémentale au fur et à mesure de l'intégration de nouveaux projets de recherche

### **Méthodologie**

Ce travail se basera sur les développements réalisés précédemment pour l'intégration de données dans deux bases de données élaborées depuis une dizaine d'années dans le cadre de différents projets de recherche : les bases @Web et BaGaTel qui sont basées sur deux modèles d'ontologie différents : l'ontologie noyau PO<sup>2</sup> (Process and Observation Ontology) qui permet de représenter des itinéraires de transformation et les différents paramètres de contrôle associés à une opération unitaire ou à une étape de caractérisation à une échelle donnée, et le modèle @Web basé sur l'ontologie naRyQ (n-ary Relations between Quantitative experimental data), qui permet de représenter une expérience impliquant un objet étudié, des paramètres de contrôle et un résultat en mettant en relation des arguments de type Quantité ou Symbolique à travers le concept Relation.

La constitution de ce référentiel spécifique au Département TRANSFORM se fera en lien étroit avec le Thésaurus INRAE récemment élaboré par la DipSO et des alignements seront également recherchés avec des référentiels internationaux.

La méthodologie adoptée devra permettre d'organiser le vocabulaire de façon modulaire, avec une granularité qui soit adaptée aux projets développés au sein du Département tout en visant la complémentarité avec le Thésaurus INRAE. L'enjeu est de lever le verrou de l'organisation « en silo » découlant de projets souvent déconnectés entre eux en proposant un pipeline d'outils commun et une structuration compatible pour l'intégration de nouveaux termes et concepts.

### **Résultats attendus**

**Les objectifs opérationnels du projet sont d'élaborer une méthode et de sélectionner les outils implémentant cette méthode qui permettront de :**

- harmoniser** les 6 ontologies de domaine déjà développées en utilisant le modèle PO2 (dairy products, meatylab et biorefinery) et @Web (matter transfer, milk microfiltration et biorefinery),
- étendre** le vocabulaire sur le périmètre d'intérêt pour TRANSFORM à l'aide de différents cas d'étude (aliments et bioproduits),
- déterminer les référentiels nationaux et internationaux** de référence pour le Département TRANSFORM,
- aligner les vocabulaires** du périmètre TRANSFORM avec ces référentiels et
- tester la capacité de la méthode** à gérer de manière efficace la mise à jour ultérieure du référentiel TransformON.