**Thésaurus INRAE : un outil central pour l’interopérabilité sémantique des systèmes d’information INRAE**

**Le thésaurus INRAE est le référentiel thématique ouvert et partagé couvrant les domaines de recherche d’INRAE** avec plus de 15 000 concepts, tous identifiés de manière unique et pérenne par un URI. Chaque concept est représenté par des termes préférentiels et alternatifs en français, et 2/3 d’entre eux sont traduits en anglais.

Le thésaurus a été constitué dans le cadre de la création d’INRAE à partir de deux vocabulaires existants (le thésaurus Irstea et le référentiel mots-clés Inra). Il est composé de 12 domaines : 10 thématiques couvrant le champ des recherches de l’institut (Environnement ; Agriculture et agronomie ; Transformation des bioressources,  etc.) et 2 domaines transversaux (Approche scientifique ; Organismes vivants). Ces domaines sont eux-mêmes constitués de microthésaurus qui permettent d’envisager une exploitation modulaire du contenu. Il est aussi possible de créer des collections qui regroupent certains concepts de manière alternative pour des besoins particuliers liés à des applications, par exemple.

Il a été conçu et est administré de manière à permettre une utilisation large au sein d’INRAE, dans l’esprit des données ouvertes, selon les principes FAIR :

* représentation selon le standard SKOS-XL et exposé en SKOS (W3C) ;
* des identifiants uniques et pérennes sous forme d’URI déréférencés ;
* une interface de navigation et d’interrogation ;
* une exposition par web services assurée par la DSI pour les applications internes ;
* des alignements vers d’autres thésaurus et ontologies (à ce jour,  GEMET et Agrovoc) ;
* une licence ouverte (Etalab)

Le thésaurus INRAE est maintenu par un comité éditorial animé par la DipSO. Sa mission est d’assurer l’accessibilité, la mise à jour et la maintenance du thésaurus avec l’aide des experts scientifiques. Il reçoit et traite les demandes des utilisateurs et initie des chantiers d’enrichissement à l’aide de l’outil Vocbench 3. Par exemple, il a déjà prévu d’ajouter des définitions textuelles et des alignements vers d’autres thésaurus et ontologies développés ou utilisés par les SI INRAE. Le comité accompagne l’intégration du thésaurus dans des systèmes d’information et forme les utilisateurs à son utilisation.

Le **thésaurus INRAE a été conçu afin d’être un outil central pour l’interopérabilité sémantique des systèmes d’information de l’institut**. Ses concepts peuvent être utilisés directement pour **décrire ou annoter de manière univoque et cohérente**  des documents (indexation dans HAL-INRAE), des jeux de données, des pages web, des descriptions d’activités, etc.

Nous aborderons les questions liées à l’**utilisation du thésaurus INRAE dans le cadre d’activités liées à l’ingénierie des connaissances, à la mise en œuvre d’applications et de systèmes d’informations sémantiques et aux « linked data »** : réutilisation d’URIs, alignements entre ressources sémantiques, intégration dans des SI pour sémantiser des champs de métadonnées, créer des « data dictionaries », entre autres applications possibles.

Sonia Bravo, Magalie Weber et Sophie Aubin pour le comité éditorial.