



Refonte du thésaurus de la filière fruits et légumes : structurer, valider, diffuser - du descripteur au tag

Webinaire Horti'doc : « Un thésaurus a-t-il encore du sens dans un monde d'IA ? »
18 novembre 2025

Chloé HERVY, CTIFL
Brittany BOIVIN, ENSSIB

PRÉSENTATION DU CTIFL

Centre technique interprofessionnel des fruits et légumes

- Institut technique agricole dédié aux professionnels de la filière des fruits et légumes.
- Objectif : Impacter la filière des fruits et légumes par ses actions

RECHERCHE ET INNOVATION

FORMATIONS ET PRESTATIONS

OUTILS ET APPLICATIONS

PUBLICATIONS

ÉVÈNEMENTS

CERTIFICATION





01

LE PROJET

02

STRUCTURER ET NETTOYER

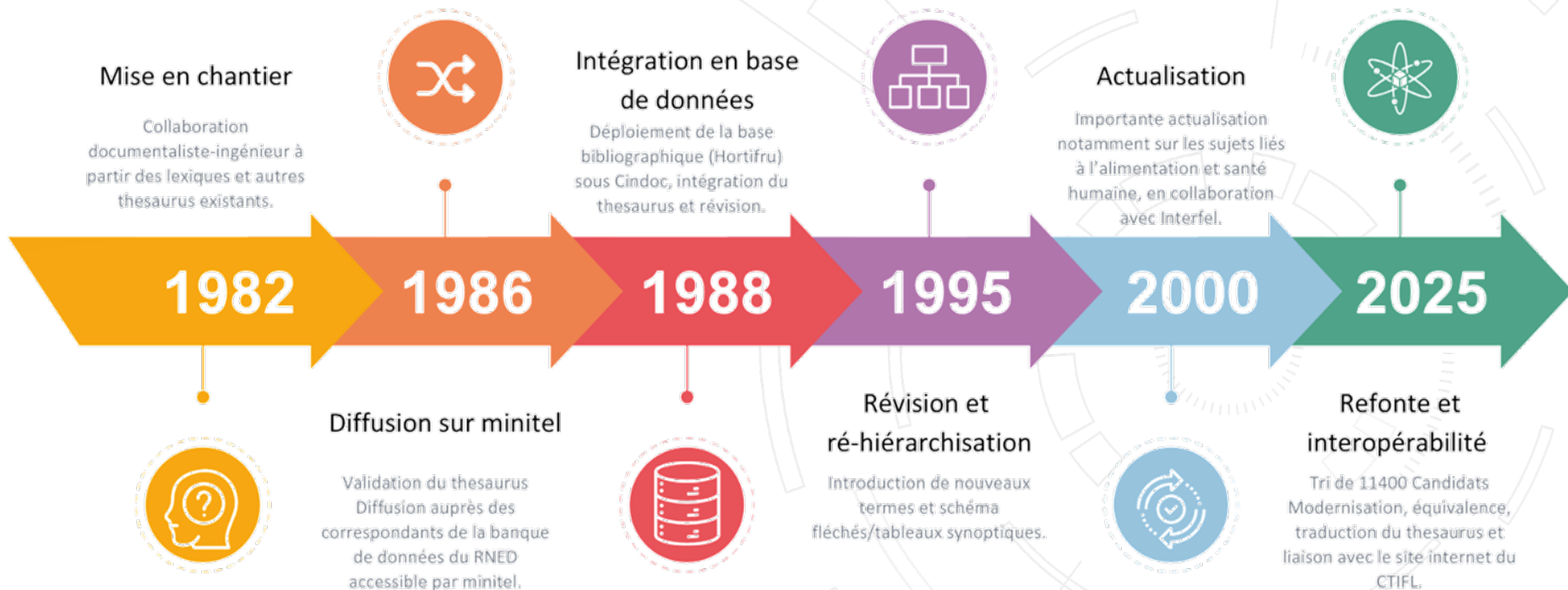
Candidats et obsolescence

03

RELIER

*Relations sémantiques,
doublons et tags*

HISTOIRE DU THESAURUS



OBJECTIFS

- Adapter le vocabulaire pour aux pratiques et problématiques professionnelles actuelles (durabilité, agroécologie, sociologie de l'alimentation)
- Faciliter la recherche d'information
- Rendre l'indexation plus pertinente et plus simple
- Préparer son interconnexion avec les outils numériques



Le thésaurus n'est pas un vestige du passé mais un socle de sens, de cohérence et d'interopérabilité.



01

LE PROJET

02

STRUCTURER ET NETTOYER

Candidats et obsolescence

03

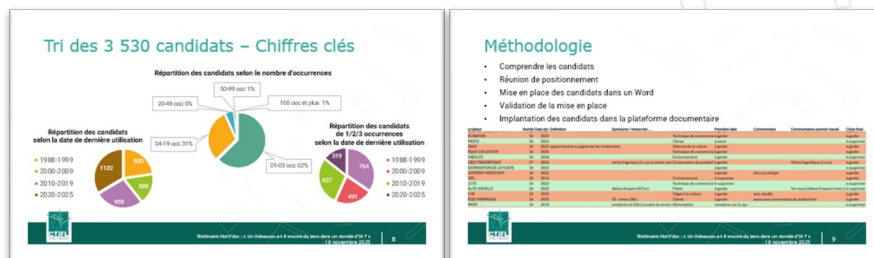
RELIER

*Relations sémantiques,
doublons et tags*

NETTOYAGE DES CANDIDATS

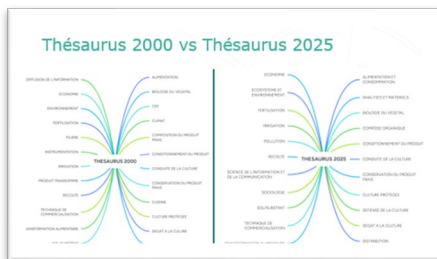
Tri des 3 530 candidats

- Suppression de 1 891 candidats à faible occurrence
- Analyse selon leur dernière date d'utilisation
- Suppression de 922 autres candidats
- Intégration de 717 nouveaux mots-clés, validés et hiérarchisés



HIÉRARCHISATION DES DESCRIPTEURS

- Analyse des branches existantes : 14 conservées, 4 renommées, 7 supprimées, 4 créés.
- Réflexion sur la profondeur de la hiérarchie (jusqu'à 7 niveaux)
- Création d'une hiérarchie plus lisible : du général au spécifique (exemple : environnement > climat > mesure climatique)
- Tentative d'analyse des branches avec ChatGPT



Analyse avec ChatGPT (pro) - Exemple

Voici la structure d'une branche du thésaurus sur la Conservation du produit frais. Qu'en penses-tu ?

...
Créer des sous-branches homogènes sans ajouter ni modifier les notions existantes.
Suggère une organisation différente et tu peux uniformiser la terminologie.
Explique chacun de tes choix en donnant la définition et la raison pour chaque terme.

...
"Ensemble des altérations physiologiques et chimiques survenant pendant la conservation." Le terme Putrescence a-t-il sa place ?
...





LE PROJET



STRUCTURER ET NETTOYER

Candidats et obsolescence



RELIER

*Relations sémantiques,
doublons et tags*

RATIONALISATION DES RELATIONS

- **TG : Terme générique**, correspond au terme parent.
- **TS : Terme spécifique**, correspond au terme enfant.
- **TA : Terme associé**, un terme existant dans le thésaurus et en lien avec le descripteur.
- **Synonyme ou Équivalent** : un terme synonyme, quasi-synonyme ou terme rejeté ("terme employé pour").
- Note d'application : La définition du descripteur.
- Traduction : Une traduction des termes en anglais

DOUBLON - Cas concret

117 termes présents dans plusieurs branches du thésaurus, voire plusieurs fois dans une même branche.

- Décision : ne pas multiplier les termes mais clarifier leurs usages contextuels.
- Tension entre rigueur documentaire et souplesse numérique.

• Dégât à la culture → Dégât non parasitaire

• Environnement → Climat

Ses enfants :

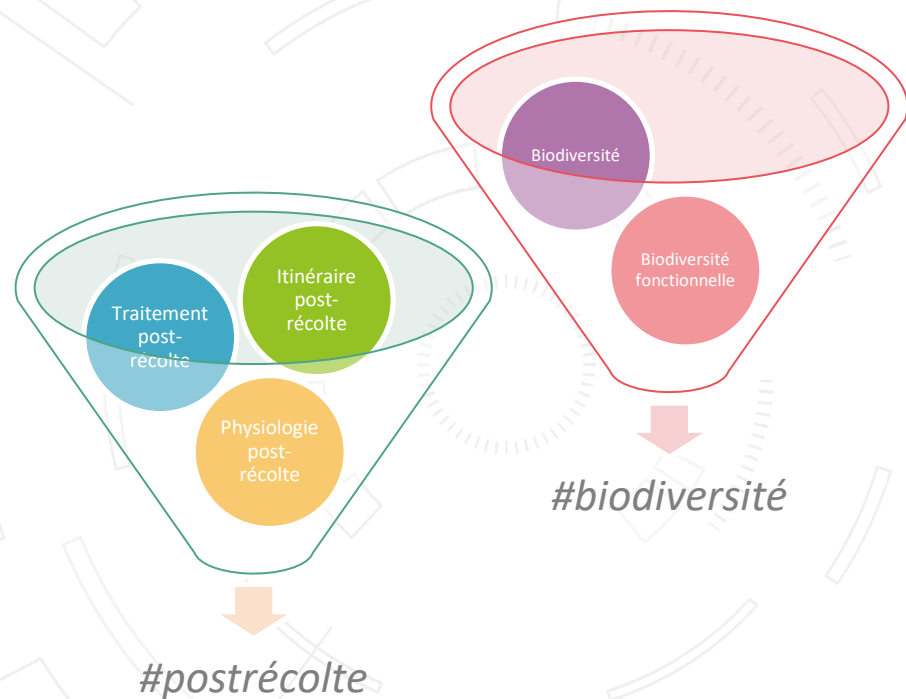
- Assurance dommage (49 docs)
- Stress climatique (9 docs)
- Stress thermique (17 docs)
- Coup de soleil (88 docs)
- Gel (479 docs)
 - Gel de printemps (153 docs)
 - Lutte contre le gel (540)
 - Asperion anti-gel (111 docs)
 - Bougie (15 docs)
 - Chauffeferette (18 docs)
 - Tour à vent (44 docs)
- Grêle (82 docs)
 - Lutte contre la grêle (158 docs)
 - Protection par filet (819 docs)
- Sécheresse (335 docs)
- Vent (136 docs)
 - Lutte contre le vent (41 docs)
 - Brise vent (83 docs)
- Climat chaud (175 docs)
- Climat froid (84 docs)
- Exigence climatique (381 docs)

DIFFUSION : DU DESCRIPTEUR AU TAG

Tag : passerelle entre le thésaurus et le web

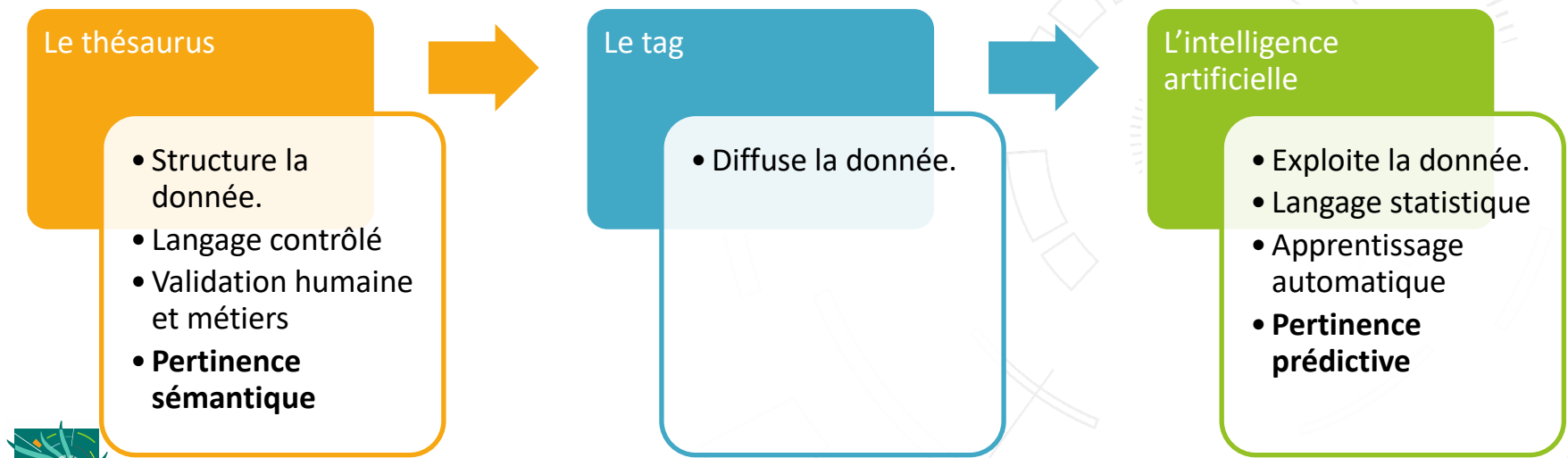
- tags issus du thesaurus
- permettent de relier les contenus entre eux
- non hiérarchisés

Le thésaurus devient la base lexicale normalisée qui alimente et structure les tags.



LIEN THÉSAURUS - INTELLIGENCE ARTIFICIELLE

Le thésaurus apporte la fiabilité lexicale qui manque aux algorithmes d'analyse automatique.



Merci pour votre attention !

Chloé HERVY
chloe.hervy@ctifl.fr

Découvrez le thesaurus
sur la plateforme documentaire
(accessible après authentification)

<https://www.ctifl.fr/plateforme-documentaire>





Réussir aujourd'hui, c'est imaginer demain.

CENTRE TECHNIQUE INTERPROFESSIONNEL
DES FRUITS ET LÉGUMES

Pour en savoir plus
www.ctifl.fr

Action financée par

