

Atelier vocabulaires SIST22

Groupe 1 : vocabulaires existants

Objectif : expliciter le panorama existant et le confronter avec ce que les gens en pensent / connaissent / utilisent + éventuellement compléter.

Inventaire <https://docs.google.com/spreadsheets/d/13nYZkdkX-Eu6Z1u7sOtW-CojXUpoGaUHL96Ahx53I6k/edit?usp=sharing>

- **Que pensez-vous de ce panorama ? est-il complet par rapport à vos besoins ?**
panorama très bien et nécessaire... à faire connaître davantage pour qu'on s'y réfère de manière automatique et coutumière, en se posant moins de questions (ML)

Quid des variables géophysiques ? gravimétrie, inclinométrie, etc ?

Variables en écologie : voir le thésaurus ANAEE ? Sur quoi est-il basé ? (ontologie OBOE je crois, + alignements vers Gemet, Agrovoc).

- **quels sont ceux que vous utilisez, comment et pourquoi ?**
j'utilise NERC BODC, en oceano, parce que je les connais dans ma discipline... la question est de rendre l'utilisation de ceux qu'on n'a pas l'habitude d'utiliser, de manière plus régulière et familière... bref les faire connaître aux Ingé et Scientifiques

- **connaissez-vous en d'autres ?**
Thésaurus de l'INRAE : <https://vocabulaires-ouverts.inrae.fr/a-propos-du-thesaurus-inrae/>

- >>> <https://consultation.vocabulaires-ouverts.inrae.fr/thesaurus-inrae/fr/>

Un référentiel identique à Taxref Espèces existe pour les habitats : HABREF

- >>> <https://inpn.mnhn.fr/programme/referentiel-habitats>
Peut-être un thésaurus développé dans Copernicus ?

- **Divers :**
Comment intégrer de nouveaux vocabulaires ? Data Terra pourrait-il proposer la mise en place de thésaurus spécifiques pour les labo/osu. Cela permettrait à Data Terra de regarder si certains termes sont de plus en plus utilisés ?

Construction/mise en place d'un registre de thésaurus standardisés pour utiliser par exemple dans Geonetwork (comme cela est fait pour INSPIRE et Gemet) : Le logiciel Re3gistry est un outil de publication de registres. Les registres permettent de gérer des listes (dans différentes langues). Le registre INSPIRE fournit un point d'accès central à un certain nombre de registres INSPIRE gérés de manière centralisée. : <https://inspire.ec.europa.eu/registry>

En parallèle se rapprocher du forum des développeurs de geonetwork pour leur demander d'intégrer les liens vers d'autres thésaurus connus et standardisés (ex celui du NERC en oceano)

Groupe 2 : Comment faire en pratique (stack technique, transformer son vocabulaire en SKOS, ..)

Quand on produit de la donnée, la première étape consiste à traduire les noms de variables issues des capteurs en concepts déjà bien définis, comme décrits dans les thésaurus existant. Mais explorer les thésaurus de référence est fastidieux. Utiliser un outil comme onaGUI pourrait aider à pointer rapidement vers les vocabulaires correspondants.

L'idée serait donc de produire un thésaurus local (par exemple avec SKOSplay) qui n'a pas vocation à être publié, à seule fin d'avoir un "dictionnaire" de nos données. Puis on le "mouline" dans onaGUI avec le thésaurus de référence qu'on vise.

Une fois validé manuellement, un export en .csv ou .xml permettra de récupérer les infos obtenues et les intégrer dans nos bases de données.

Groupe 3 : Comment structurer la communauté autour de ça

- Quels besoins ?
 - mutualisation / échanges / recommandations / formations / outils
- Quel cercle ? SIST ? + large ? interpoles ?
- Quels besoins de formalisation de la communauté ?
- Quels objectifs aurait cette communauté ?

Danger de l'ouverture des données exploitées massivement par des machines

Besoin de construire une communauté ?

Articulation entre la communauté Ingénieur Data et communauté de chercheurs pour formaliser la connaissance

--> avancées à différents niveaux des communautés

SHS ne sont pas à ce stade de standardisation, difficultés

- -
- question autour du danger de l'ouverture des données via la formalisation de vocabulaires

- - différencier fair et ouverture
 -
- problème de RH: effort vs besoins et priorités

- -
- existant mais dans d'autres langues ?

- -
- comment capturer le lexique des chercheurs ? ou comment articuler ce travail de formalisation avec les chercheurs ?

- -

--> difficulté de mobiliser les chercheurs