

*Création d'un formulaire ODK  
(Open Data Kit)  
conforme au standard  
« Occurrences de Taxon » v2.0*

CEN Occitanie et CEN Nouvelle-Aquitaine



**SINP**  
Système d'Information  
de l'iNventaire du Patrimoine naturel



Rapport final



## Rappel des objectifs

La mise en conformité des données naturalistes au standard « Occurrences de Taxon » v2.0 est chronophage et nécessite des compétences techniques de gestion de données. Cela représente un frein à la mutualisation des données et donc à leur intégration dans le SINP.

Les observations sont de plus en plus collectées sur le terrain via les smartphones des naturalistes et diverses applications « propriétaires » et non génériques.

ODK est une suite de logiciels libres qui permet la collecte de données sur le terrain, dans des environnements contraints (faible couverture réseau, matériel standard). De nombreux acteurs du SINP et Parcs Nationaux l'utilisent pour leurs besoins de saisie mobile de données opportunistes ou pour la mise en œuvre de protocoles particuliers. Les CEN d'Occitanie et de Nouvelle-Aquitaine utilisent la solution depuis 2015 et ont développé de nombreux formulaires pour différents publics. L'adoption de cet outil par les naturalistes des deux structures est unanime : gain de temps (pas de double saisie, formulaire optimisé) et gain de qualité de la donnée (référentiels à jour et saisie dirigée limitent les erreurs).

XLSForm est la norme, dérivée d'XForms, qui permet de décrire, sans connaissance informatique particulière, les formulaires affichés par ODK Collect (et dérivés, Enketo, Survey 123, etc.).

Proposer un formulaire naturaliste pour ODK Collect, conforme au standard du SINP, selon la norme XLSForm, permettrait de réduire très fortement le temps consacré à la mise en conformité des données. Le format XLSForm facilitera aussi son appropriation par des structures souhaitant développer leurs propres formulaires.

Les données sont consolidées au sein d'un serveur « ODK Central » que nous avons mis en place dans le cadre de ce projet. Il peut servir de démonstrateur aux structures porteuses du SINP (pôles thématiques, plateformes régionales, INPN).

Ce serveur est interrogeable via une API REST « ODATA ».

# Déroulement

Plusieurs étapes ont été définies pour réaliser les objectifs.

## Réalisation d'une version du formulaire conforme au standard en vigueur au lancement du projet

Mise en place d'un dépôt Github pour le développement du formulaire et la documentation :

-> <https://github.com/Projet-ODK-SINP/>

## Structuration des questions et de la logique du formulaire

Les documents relatifs au standard ont été téléchargés sur le site de l'INPN et déposés ici :

-> <https://github.com/Projet-ODK-SINP/ODKform/tree/main/sources/standard>

Les attributs ont été regroupés selon leur concept ou extension, et leur caractère, obligatoire, facultatif ou conditionnel, a été renseigné.

L'ensemble de ces attributs définit un corpus de questions à implémenter dans le formulaire en respectant la logique du standard.

-> [https://github.com/Projet-ODK-SINP/ODKform/blob/main/sources/standard/occtaxODK\\_brut\\_initial\\_non\\_fonctionnel.xlsx](https://github.com/Projet-ODK-SINP/ODKform/blob/main/sources/standard/occtaxODK_brut_initial_non_fonctionnel.xlsx)

## Préparation des listes de choix en conformité avec le standard

Les nomenclatures ont quant à elles été transposées en CSV.

-> [https://github.com/Projet-ODK-SINP/ODKform/blob/main/sources/standard/nomenclatures\\_sinp.csv](https://github.com/Projet-ODK-SINP/ODKform/blob/main/sources/standard/nomenclatures_sinp.csv)

Ces deux « éléments » constituent la matière brute de notre formulaire :

- les questions posées et leur organisation
- les réponses proposées dans les questions fermées

Le classeur Excel [occtaxODK\\_brut\\_initial\\_non\\_fonctionnel.xlsx](#) constitue cet embryon de formulaire.

## Mise en forme du référentiel taxonomique

Le formulaire exploite TAXREF v16 (publié par PatriNat en décembre 2022), en distinguant les noms valides des synonymes. Cela est mis en œuvre en intégrant des balises HTML aux libellés des taxons.

Les requêtes SQL permettant de générer les fichiers csv (Faune, Flore, Fonge) exploités par ODK Collect sont mises à disposition ici :

-> <https://github.com/Projet-ODK-SINP/ODKform/tree/main/sources/taxref>

## Installation du serveur ODK Central

Une instance d'ODK Central a été déployée pour le projet à l'adresse suivante :

-> <https://aap-odk-sinp.cen-nouvelle-aquitaine.dev>

C'est la version v2023.3.0 qui est actuellement en service. L'instance est hébergée chez OVH.

## Création de comptes utilisateurs

- Utilisateur mobile (du téléphone sur le terrain)

Un utilisateur mobile nommé Naturaliste a été créé. Le QRCode correspondant, à « flasher » avec ODK Collect est celui-ci :



- Utilisateur web (qui pourra se connecter au serveur et récupérer les données)

identifiant : [lecteur@aap-odk-sinp.dev](mailto:lecteur@aap-odk-sinp.dev)

mot de passe : ve1pMXTnHIRwu9Ee

## Mise en production du formulaire réalisé

Une première version du formulaire est disponible sur ce serveur.

Il apparaîtra dans les formulaires disponibles à l'utilisateur qui aura scanné le QRCode ci-dessus.

## Documentation du fichier XLSForm

Chaque élément (question) du fichier xlsform est documenté avec la description de l'attribut issue du standard (colonnes *concept\_principal*, *attribut* et *description* dans la feuille *survey* du xlsform) .

-> <https://github.com/Projet-ODK-SINP/ODKform/blob/main/OccTaxODK.xlsx>

## Rédaction d'une « documentation utilisateur » du formulaire sous format écrit et/ou vidéo

Un manuel d'utilisation, illustré de nombreuses captures d'écran est présenté ici :

-> [https://github.com/Projet-ODK-SINP/ODKform/blob/main/manuel\\_utilisation.md](https://github.com/Projet-ODK-SINP/ODKform/blob/main/manuel_utilisation.md)

## Récupération des données

Le moissonnage du serveur est possible via l'[API ODATA](#).

Nous avons documenté l'utilisation des fonctions pl-pyODK et l'utilisation de FME :

- Dans PostgreSQL en utilisant les fonctions [pl-pyODK](#) :  
-> [https://github.com/Projet-ODK-SINP/ODKform/blob/main/recuperation\\_des\\_donnees\\_dans\\_postgresql.md](https://github.com/Projet-ODK-SINP/ODKform/blob/main/recuperation_des_donnees_dans_postgresql.md)
- Avec FME en utilisant un script JavaScript qui interroge l'API d'ODK Central :  
-> [https://github.com/Projet-ODK-SINP/ODKform/blob/main/recuperation\\_des\\_donnees\\_api\\_et\\_etl.md](https://github.com/Projet-ODK-SINP/ODKform/blob/main/recuperation_des_donnees_api_et_etl.md)

D'autres solutions d'exploitation des données existent :

- utilisation de la bibliothèque [ruODK](#) pour R
- pyODK : <https://github.com/getodk/pyodk> (maintenu par GetODK)
- les connexions aux solutions PowerBI et Excel de Microsoft

Les données peuvent aussi être téléchargées au format CSV depuis la page « Soumissions du serveur » :

-> <https://docs.getodk.org/central-submissions/#accessing-submissions>