



HAL
open science

ODK, un outil open source pour la création de formulaires embarqués

Eric Quinton, Sylvie Ladet, Alain Benard, Florent Blaise

► To cite this version:

Eric Quinton, Sylvie Ladet, Alain Benard, Florent Blaise. ODK, un outil open source pour la création de formulaires embarqués. Printemps de la Donnée 2024, INRAE; Université Haute-Alsace; Université de Strasbourg; INSA; PNDB; AgroParisTech; Université de Lille; Sorbonne Université; Data Terra, Mar 2024, Paris, France. hal-04708638

HAL Id: hal-04708638

<https://hal.inrae.fr/hal-04708638v1>

Submitted on 25 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

ODK, un outil open source pour la création de formulaire embarqués



Contexte & constats à Inrae

- ❑ Développement des carnets de terrain électronique pour la collecte de données facilitée, sécurisée (Cf. qualité des données) par de multiples acteurs (agents Inrae; science participative/crowdsourcing...)
- ❑ Possibilité de s'adapter facilement et rapidement aux n protocoles des scientifiques
- ❑ Difficulté de maintien des solutions maison développées from scratch par manque de moyens humains/temps

ODK dans le CATI Gedeop, c'est une équipe projet ad-hoc (A. Bénard, F. Blaise, W. Heintz, S. Ladet, E. Quinton) réunis depuis 2022 avec comme objectifs :

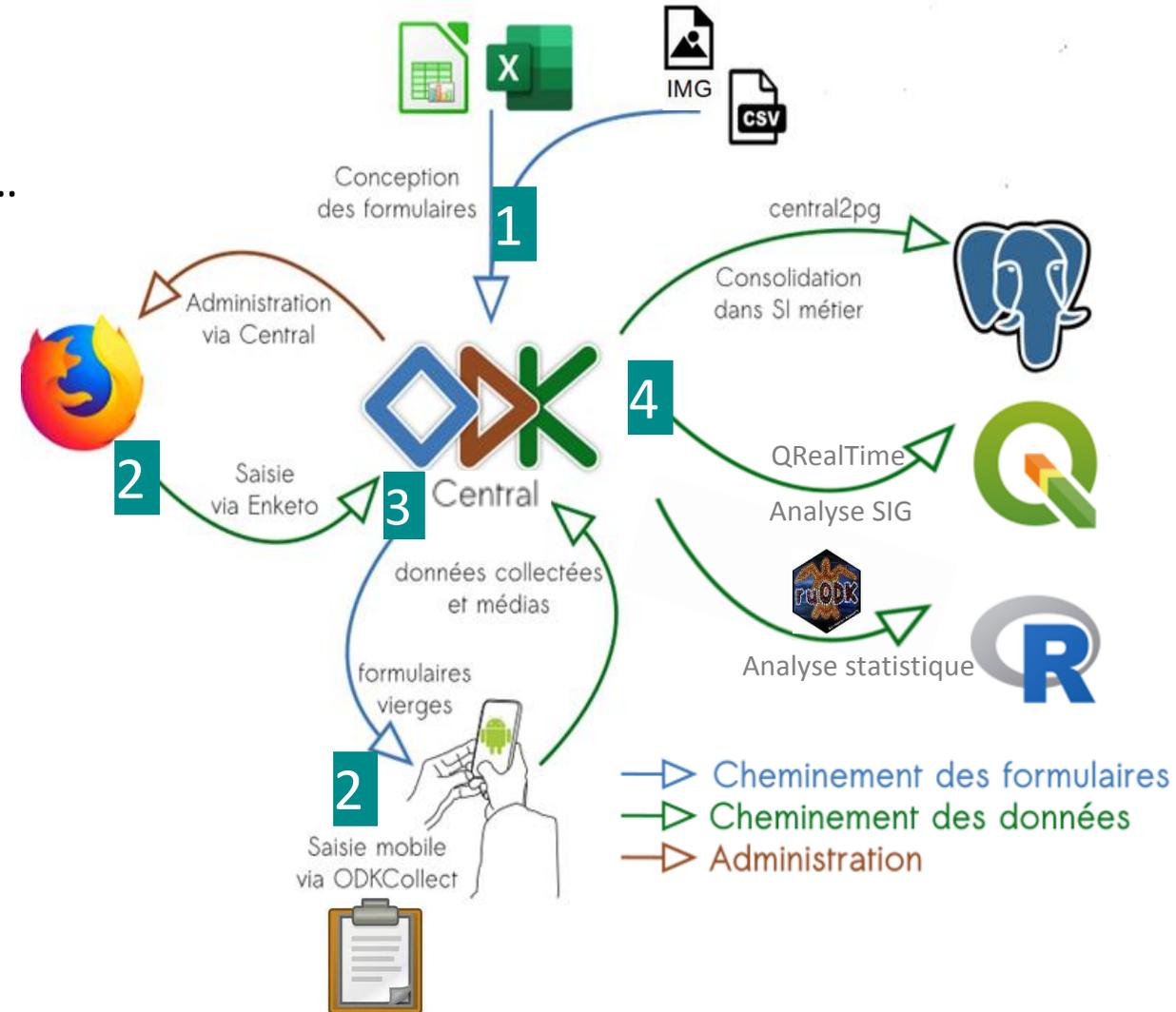
- **montée en compétences au sein du CATI Gedeop** en 3 temps : avril 2022: formation de l'équipe projet (2 jours par M. Bossaert CEN, financé par Gedeop), puis janvier 2023 atelier ODK à l'AG Gedeop puis mars 2024 : au-delà de Gedeop à Inrae
- **mise en place d'un serveur ODK Central**, hébergé dans l'infrastructure technique d'INRAE
- construction de **divers formulaires ODK** (naturaliste, inventaire, enquête pratiques...).



Résumé de fonctionnement d'ODK

□ Ensemble d'outils Open Source pour :

1. **Construire un formulaire, sous la forme d'un tableau, avec des questions à réponses de diverse nature:** nombre, texte, date, donnée géographique point/ligne/polygone, image, audio, vidéo, QRcode... cf liste des types <https://docs.getodk.org/form-question-types/> avec saisie contrôlée (bornes de valeurs, listes déroulantes simple / en cascade avec des référentiels internes ou externes ...) et champs calculés
2. **Collecter avec un appareil nomade (ou web) ces données avec l'application mobile ODK Collect** même en conditions contraintes (déconnecté)
3. **Centraliser ces données, temporairement, dans ODK Central, serveur unique avant transfert**
4. afin de pouvoir ensuite **les exploiter** dans des SGBD (PostgreSQL/PostGIS) et/ou des logiciels d'analyse (QGIS, R,...) via extensions, scripts API...



Gouvernance d'ODK et documentation

- ❑ **Projet initié en 2008 et porté par la société [GetODK](#)**, qui développe ODK [composée de 14 personnes](#) qui s'appuie sur un comité technique TAB (Technical Advisory Board) impliquant des utilisateurs de toute la planète et de différents domaines. [Ce TAB se réunit tous les 15 jours de manière très transparente.](#)
 - En 2020 ODK a été utilisé dans tous les pays du monde excepté au Groenland et en Corée du Nord !
 - ODK Collect, la partie mobile a été téléchargé plus d'un million de fois sur le playstore de google.
cf. <https://getodk.org/#success-stories>
- ❑ **ODK Collect et ODK Central s'appuient sur des composants ouverts et interopérables**, produits / maintenus par GetODK, qui sont devenus des briques de base de l'industrie de la collecte de données mobile :
 - XLSForm - Spécification de formulaires basée sur Excel
 - pyxform - Librairie Python qui convertit les XLSForms en ODK Xforms
 - ODK XForms - Spécification de formulaires basée sur la norme Xforms du W3C
 - OpenRosa - APIs pour communication client/serveur des outils ODK
 - Enketo Express - formulaires web compatible avec ODK
- ❑ **Un forum vivant <https://forum.getodk.org>** en langue anglaise
 - développé sous Discourse (comme le forum Dipso!) avec partage de fichiers tableurs (debug)
 - 13000 membres inscrits
 - qui facilite la co-construction des outils



Présentation de formulaire en quelques captures 1/2

Type de la question
(tous prédéfinis)

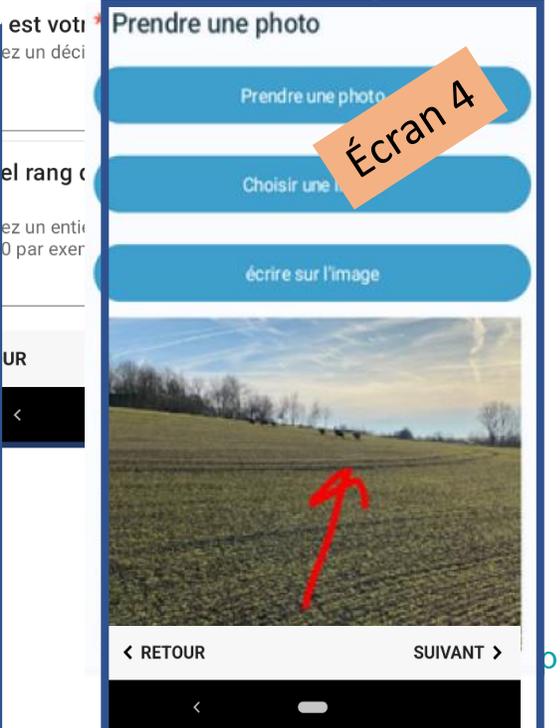
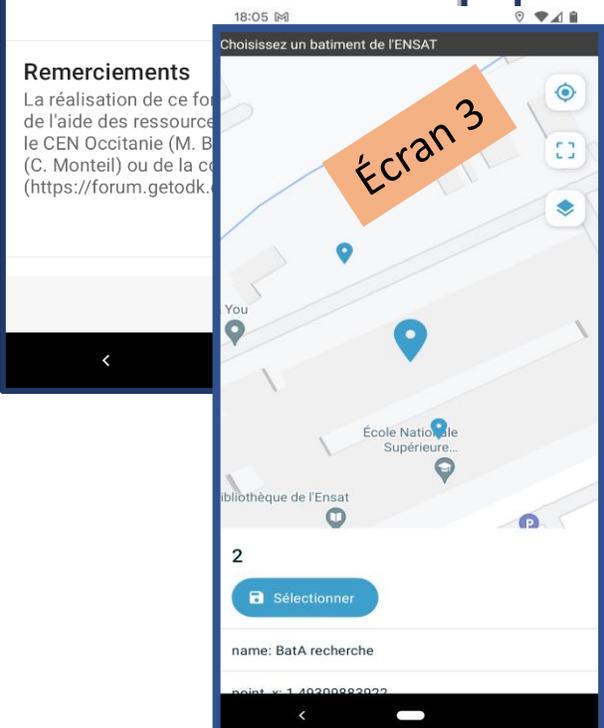
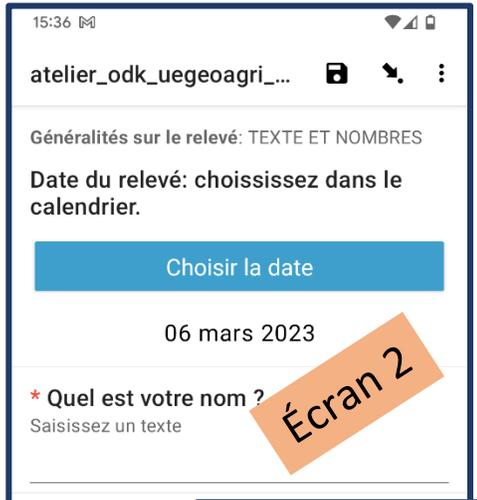
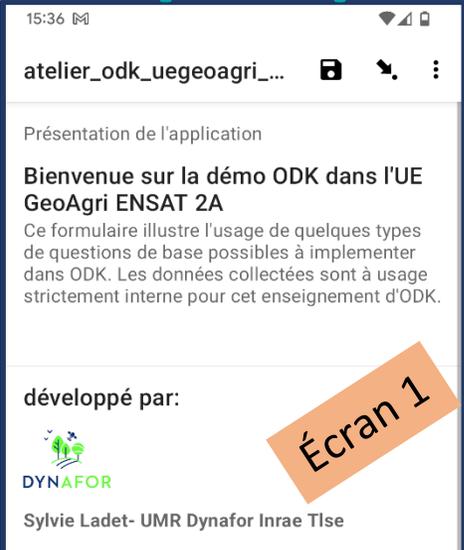
Nom interne
choisi

Titre affiché
(mis en gras dans
le formulaire)

type	name	label
begin_group	presentation	Présentation de l'application
username	username	username
	note_bienvenue	Bienvenue sur la démo ODK dans l'UE GeoAgri ENSAT 2A
note		
note	devlp	développé par:
note	remerciement	Remerciements
note	contact	Contact
end_group		
begin_group	releve	**Généralités sur le relevé** : TEXTE ET NOMBRES
date	releve_date	Date du relevé: choisissez dans le calendrier.
text	nom	Quel est votre nom ?
decimal	taille	Quelle est votre taille ?
integer	travees_nb	A quel rang dans la salle êtes-vous assis?
end_group		

begin_group	position	**Où êtes-vous situé?**: CARTOS: question liée à position
geopoint	lat_lon	Position actuelle
calculate	latitude	
calculate	longitude	
select_one_from_file centre_bati_ensat_WGS84.geojson	centre_salle	Choisissez un bâtiment de l'ENSAT
geopoint	geo	Positionnez la salle sur la carte
end_group		

begin_group	photo	**Et pour finir...**: MEDIAS
image	photo	prendre et annoter une photo
end_group		



Présentation de formulaire en quelques captures 2/2

☐ **Plein d'autres colonnes facultatives** dans le tableau pour définir :

- **les valeurs par défaut** : date d'aujourd'hui (colonne default)
- **la saisie contrôlée** : bornes de valeurs de la taille (colonne constraint)
- **Les champs calculés**: le nom du collecteur

hint	constraint	constraint_message	calculation	required	appearance	default
Saisie d'un texte, d'un nombre entier ou décimal (avec contrainte)					field-list	
Saisissez un texte			\${username}	yes		today() \${username}
Saisissez un décimal en m contraint entre 2 bornes	`\${taille} >= 0.5 and \${taille} <= 2.5`	Entre 0.5 m et 2.5 m SVP !				round(1.50 + 0.4 * random(),2)
Saisissez un entier contraint entre 2 bornes (taper 30 par exemple)	`. > 0 and . <= 25`	Entre 1 et 25 ans SVP !		yes		10

Quelques utilisations Inrae/Gedeop [nombre de soumissions reçues ODK Central]

- Suivi floristique, protocole Baïonnette, avec référentiel TaxRef, en mode déconnecté avec sélection des transects sur carte et fonction GoTo sur fond de carte Mbtiles ScanIGN (Dynafor, 2022) [58]
- Enquête web via Enketo sur études Inrae dans réserve MAB (Dynafor, 2023) [11]
- Collecte d'enregistrements sonores géolocalisée à Florence en Italie avec fond de carte OSM/GoogleHybride (Dynafor, 2023) [111]
- Travaux pratiques avec étudiants ENSAT (2ièmeA ; master 2) avec centralisation sur GoogleDrive (Dynafor, 2023) [16]
- Saisie d'informations relatives aux arbres d'un arboretum avec sélection sur fond carto (Silva, 2023) [55]
- Description de stations de prélèvements ADNe en rivière : saisie des paramètres lus sur les sondes, caractérisation du prélèvement (EABX, 2023,2024) [22]
- Saisie du résultat de tris de poissons en laboratoire avec mode web Enketo (EABX, projet DILEME, 2022,2023) [800]



INRAE

Open Data Kit (ODK)

CAFE NUMERIQUE DIPSO 25/03/2024 / Groupe de travail GEDEOP ODK

Administration du serveur ODK Central mutualisé

☐ Service proposé par Cati Gedeop:

<https://odk.gedeop.inrae.fr> (34 projets, 46 mbres au 18/03/24)

- Gestion des droits via divers rôles: administrateurs(équipe ODK Gedeop)/gestionnaires/simple utilisateurs

<https://docs.getodk.org/central-users/#central-users-web-roles>

- Les gestionnaires de projets :

- gère les formulaires (versionnés) regroupés par projets projet=1 unité INRAE
- les utilisateurs mobiles et leurs accès aux formulaires
- les soumissions et leur validation. Les soumissions sont consultables en ligne ou téléchargeable (archive zip)
- les utilisateurs web et leurs droits
- la publication web des formulaires avec Enketo

☐ Architecture et maintenance INRAE

- Machine virtuelle hébergée dans le datacenter de Toulouse (4 vCPU, RAM: 8 Go et 75 Go de disque)
- Machine sauvegardée
- Backup de la base de données vers le CINES :
 - PgBackRest : sauvegarde incrémentale horaire en journée ouvrable
 - Taille actuelle de la base de données : 2,5Go
 - Rétenion des sauvegardes pendant un mois
- Demande de création de compte et de projet via tickets Ariane: cf diapos suivantes
- Rédaction en cours des règles d'usage
<https://forum.dipso.inrae.fr/t/odk-gedeop/2027>

The screenshot shows the ODK Central web interface. At the top, there is a navigation bar with links for 'ODK Central', 'Projets', 'Utilisateurs', and 'Système'. Below this, a header reads 'Bienvenue sur Central'. The main content area features four statistics cards: '34 Projets', '46 Utilisateurs', 'Documents', and 'Forum'. Below these cards, there is a section for 'Projets' with a 'Nouveau' button and a 'Trier' dropdown menu. The first project listed is 'Silva' with a 'test_trace' submission from 'Arboretum etiquetage 2023'.

Statistique	Valeur
Projets	34
Utilisateurs	46
Documents	0
Forum	0

Central est organisé en projets, chacun contenant ses propres formulaires et données associées.

Les utilisateurs peuvent être assignés à des projets pour les gérer, ou pour collecter ou réviser les données soumises.

Un guide de démarrage et la documentation utilisateur sont disponibles sur le site internet d'ODK.

Vous pouvez toujours obtenir le forum, où vous pouvez accéder aux précédentes.

Projets Nouveau Trier Dernière soumission

Silva

test_trace
Arboretum etiquetage 2023

9 0 0 -20 j 10
55 0 0 -6 m. 55

Ticket Ariane n°1: Demander un compte/espace

https://ariane.inrae.fr/block?id=ariane_sc_cat_item&sys_id=1e08275987db251096b2a7580cbb3566&sysparm_category=2ca7d5920f3b7500db98936792050e22

- Demande sur Ariane dans Tous les catalogues => Informatique de proximité => Amélioration continue => ODK. (non automatisé, traitement manuel de la demande)
- Entrer le bénéficiaire d'un espace ou d'un compte, possibilité de demander l'espace (1 espace = 1 unité) et le compte sur le même formulaire.
- Possibilité d'ajouter un commentaire sur la demande.

The screenshot shows the Ariane web interface for the 'Open Data Kit (ODK)' form. The page header includes the Ariane logo and navigation links. The main content area is titled 'Open Data Kit (ODK)' and contains a description of the service, a 'Bénéficiaire de la demande' field with 'Florent BLAISE' entered, and a 'Description succincte de la demande' field. There are also buttons for 'Ajouter à vos brouillons', 'Soumettre', and 'Poser une question'. The footer includes the INRAE logo and contact information.



INRAE

Open Data Kit (ODK)

CAFE NUMERIQUE DIPS0 25/03/2024 / Groupe de travail GEDEOP ODK

Copyright - ©INRAE - CGU - Cookies - Mentions légales - Contact - À propos

INRAE

Assistance virtuelle

Ticket Ariane n°2: Support

https://ariane.inrae.fr/block?id=ariane_sc_cat_item&sys_id=906528a10a0a0bf000a85bc20dfb6541&sys_parm_service=fd4305b81b57a110ffd09821b24bcb3f

- ❑ Possibilité de déclarer un incident ou de poser des questions sur Ariane.
 - ❑ Nous ne fournissons pas de support pour coder les formulaires.
- ❑ Fiche de présentation du service dans la base de connaissances

The screenshot shows the 'Je déclare un incident ou je pose une question' form. It includes a search bar at the top, a 'Soumettre' button, and several input fields: 'Utilisateur concerné' (Florent BLAISE), 'Domaine' (Applications métier), 'Description courte', and 'Personne(s) à informer pour le suivi du ticket'. There are also checkboxes for 'Options' and a 'Champs obligatoires' warning.

The screenshot shows a Knowledge Base article titled 'Utiliser la plate-forme ODK fournie par GEDEOP - qu'est-ce qu'un projet, quels sont les droits?'. It includes a table of contents, a 'Présentation' section, and a 'Comment utiliser la plate-forme ODK Central, gérée par GEDEOP?' section. The article is created by Eric QUINTON and has 33 views.

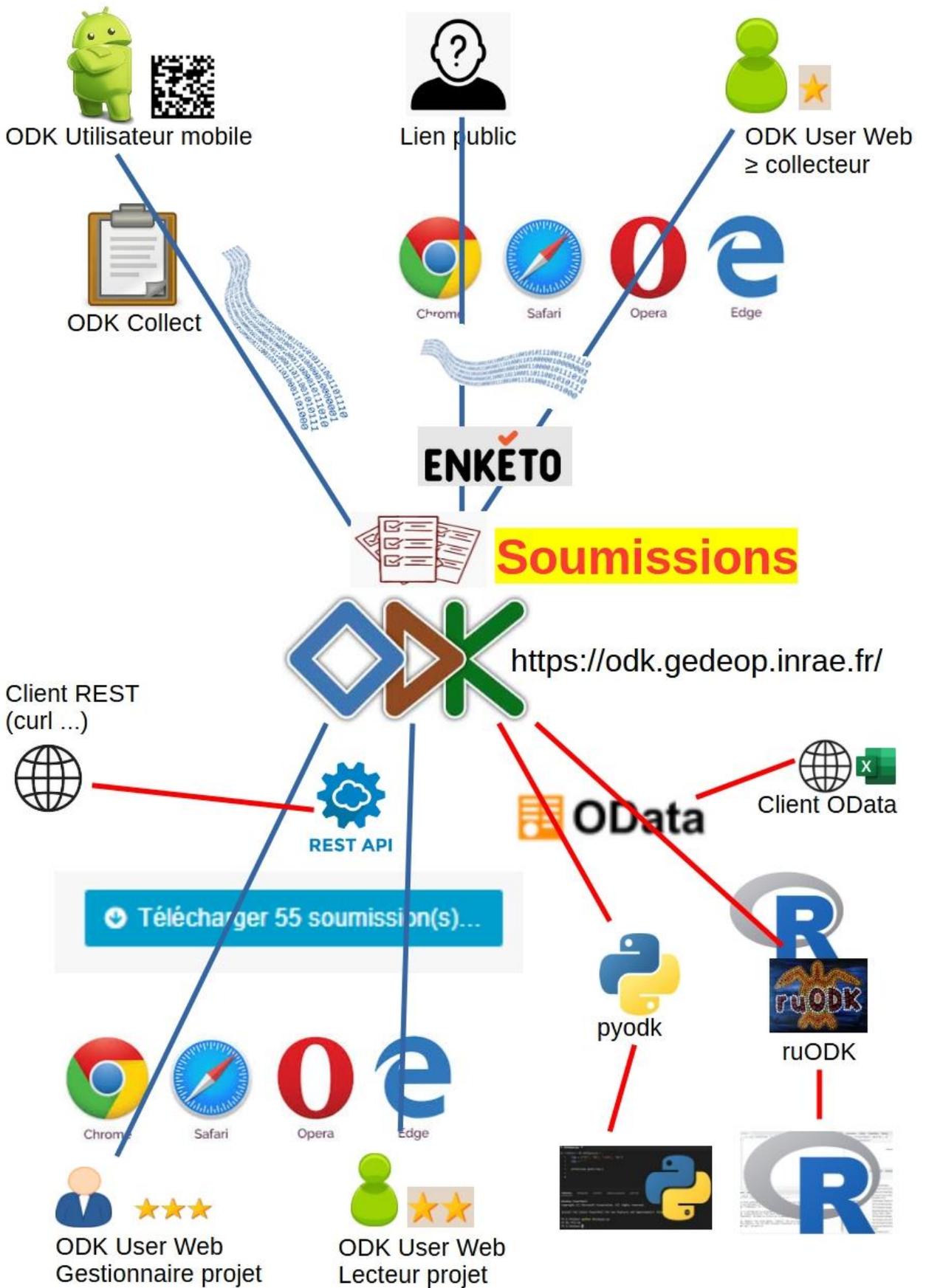


Open Data Kit (ODK)

CAFE NUMERIQUE DIPS0 25/03/2024 / Groupe de travail GEDEOP ODK



Exploitation des données



Télécharger les soumissions d'un formulaire

<https://docs.getodk.org/central-submissions/#downloading-submissions-as-csvs>

Via l'interface ODK Central

- Statut gestionnaire ou lecteur de projet
- Sélection du projet / formulaire / onglet 'Soumissions'
- Bouton 'Télécharger N soumissions' :
 - 1 ou plusieurs fichiers csv selon les boucles de répétitions
 - Des fichiers 'médias' (photos, sons ...)

Télécharger les soumissions ✕

Options d'exportation

- Séparer le contenu des questions à choix multiples en plusieurs colonnes.
- Supprimer les noms de groupe
- Inclut les champs qui ne sont pas dans le formulaire publié
Utilisez cette option si vous avez besoin de champs référencés dans les versions précédentes du formulaire

 Principale table de données (sans "repeats")

 .csv

 Toutes les tables de données

 .zip

 Toutes données et fichiers joints

 .zip

Via pyodk (<https://www.youtube.com/watch?v=slaquyz5bP8>)

- Installer et importer 'import pyodk'
- Instancier la classe Client()  connexion
- Utiliser les méthodes de l'objet submissions
 - monclient.submissions.get_table(...)
 - Gérer les blocs de répétition par jointure

Télécharger les soumissions d'un formulaire

<https://docs.getodk.org/central-submissions/#downloading-submissions-as-csvs>

Via ruODK (<https://docs.ropensci.org/ruODK/reference/ruODK-package.html>)

- Installer et déclarer 'library ('ruODK')
- Utiliser la fonction `ru_setup()` ? connexion
- Lister les tables à récupérer :
`odata_service_get()`
- Récupérer les tables :
`odata_submission_get(table = ...)`
- Gérer les blocs de répétition par jointure

Bonus

<https://appgeodb.nancy.inrae.fr/donnees/documentations/ruodk/>

- Un fichier pdf d'une demi-page qui décrit ce qu'il faut faire / savoir
- Un tuto video (11')
- Les sources

- Fonction qui récupère / fusionne
 - 1 seul csv
 - Multi csv
 - Json
- Perspectives :
 - Export SQLite
 - Porter le travail sur pyodk

PHPOdkRead

<https://github.com/inrae/phpodkread>

Principe :

Le programme permet de créer un fichier JSON contenant l'ensemble des données présentes dans le fichier ZIP fourni par ODK

Structure hiérarchique (entrée *CHILDREN* créée à chaque niveau)

Positionne correctement les « boucles »

Il est possible d'alimenter directement une base de données PostgreSQL (pas testé avec MySQL)

Une classe à écrire pour réaliser les opérations :

```
<?php
```

```
include_once "lib/db.class.php";
```

```
class Dileme extends Db implements Database {
```

```
(...)
```

Lancement du code en ligne de commande

```
php odkread.php
```

Gère le déplacement des fichiers zip traités

Ressources

- <https://docs.getodk.org/> – Guide d'utilisation de la **suite ODK** (anglais)
- <https://docs.getodk.org/collect-intro/> – Guide Utilisation **ODK Collect**
- <https://docs.getodk.org/form-question-types/> – Inventaire des **questions (widgets)** utilisables dans les formulaires
- https://static.geotribu.fr/articles/2021/2021-06-08_odk_postgis_1/ – Série de 3 articles présentant **Open Data Kit en français**
- <https://xlsform.org/en/> – Guide complet d'utilisation (en anglais)
- <https://www.youtube.com/watch?v=22l0xHxJ3vo> – xlsform template
- <https://xlsform.org/en/ref-table/> – Table de tous les items utilisables dans les 3 feuilles **survey**, **choices** et **settings** d'un classeur XLSForm
- <https://biodiversityforms.org/> – Des formulaires relatifs à la biodiversité

Conclusion & Perspectives

- ❑ Opensource / Mode Déconnecté / Cartographique
- ❑ Communauté +++
- ❑ Facilité de mise en œuvre de la solution
- ❑ Grande souplesse (workflows simples ou complexes)
- ❑ Conception aisée de formulaires de saisie
→ Par des collègues « thématiciens »
- ❑ Intégration facilitée au SI préexistant
- ❑ Généricité
→ +/-

- ❑ Formation Inrae (tous CATIs), en cours de discussion avec Dipso/FTLV Inrae
- ❑ Elargir l'équipe projet
- ❑ Roadmap GetODK riche avec plus de fonctionnalités sur entités, widgets cartographiques
→ A la faveur d'un financement important de la Fondation Gates

ODK receives \$3.7M Gates Foundation grant to improve global health data

ODK is used to collect the data that powers the fight against poverty, inequity, and disease around the world. In global health alone, the platform's impact has been notable. ODK was used to:

- Eradicate wild polio from Africa
- Prove the efficacy of the RTS,S malaria vaccine
- Measure the global burden of disease

This impact is only possible because of the incredible partners, big and small, ODK has had over the years. To everyone who has been an ODK champion, thank you. Together, we've changed the world!

This grant from the Gates Foundation enables us to continue making the world better through data. In particular, it will accelerate ongoing work extending ODK to better support global health workflows with geospatial and longitudinal data.

We have an incredible opportunity to further improve the health of the billions of people who live in the places where ODK is used. It's an opportunity we're excited to deliver on with you.



INRAE

Open Data Kit (ODK)

CAFE NUMERIQUE DIPSO 25/03/2024 / Groupe de travail GEDEOP ODK

[gedeop-odk-admin\[at\]inrae.fr](mailto:gedeop-odk-admin[at]inrae.fr)

Merci de votre attention et place aux questions

Sources:

- Apéro technique OSU OREME du 30/01/24 avec intervention de Mathieu Bossaert (CEN Occitanie) devant CEFÉ et UMR G-eau INRAE Pierre Balzergue et Bruno Bonté : [lien à la présentation](#)

