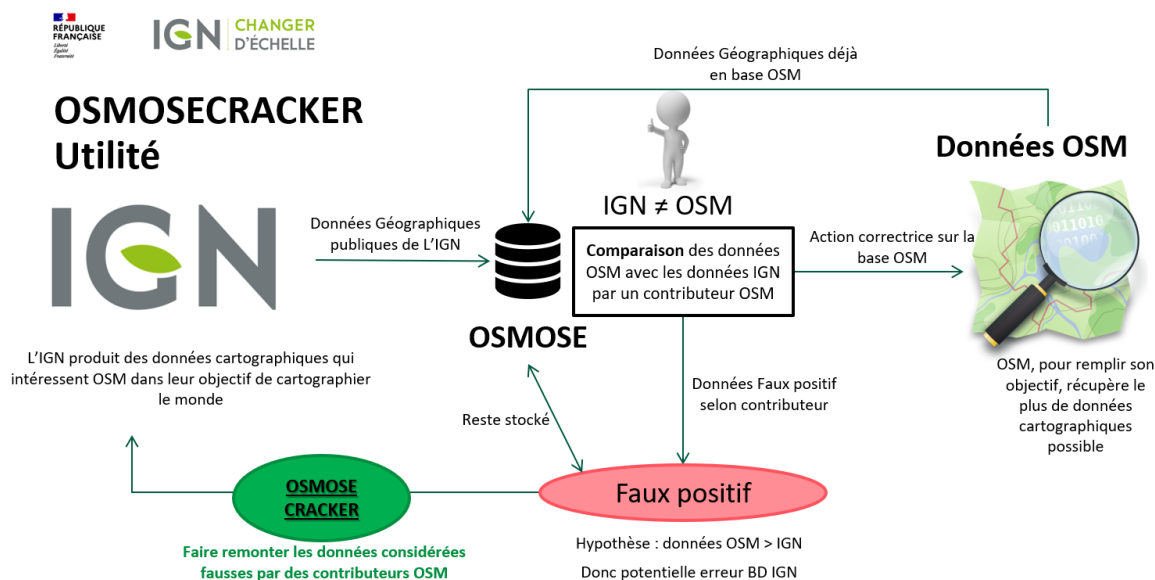


READ ME

OSMOSE compare des données open source dont celle de l'IGN à celle d'OSM, pour créer des signalements pour les contributeurs OSM indiquant une différence entre leurs données et des données externe.

L'objectif principal du programme OSMOSECRAKER est la création de signalements IGN en cas de retour d'informations jugées mauvaises par un contributeur OSM par rapport à une donnée OSM. En théorie, cela impliquerait que la donnée OSM soit supérieure à la donnée de l'IGN, ce qui suggérerait un problème dans notre base de données.



INSTALLATION

Il est avant tout primordial d'installer Python pour pouvoir lancer le programme. Il est également conseillé d'avoir une application capable de lire un fichier SQLite, telle que DBeaver. *Pour ouvrir le fichier "Analyse.fmw", vous aurez besoin de FME.*




- Python : [Téléchargement de Python](#)
- DBeaver : [Téléchargement de DBeaver](#)
- FME : [Téléchargement de FME](#)















Pour assurer le bon fonctionnement du programme, vous devrez installer un ensemble de bibliothèques Python certaines sont préinstallées. Voici la liste des bibliothèques Python à installer :

__future__, argparse, dataclass_csv, dataclasses, datetime, json, logging, math, os, pathlib, psycopg2, psycopg2.extras, sqlite3, sys, requests, requests.auth, typing, urllib3, urllib.parse, uuid

Pour obtenir des informations sur l'installation des bibliothèques Python, consultez la procédure d'installation de bibliothèques : [Procédure d'installations de bibliothèque .pdf](#)

Ensuite, il vous faudra récupérer tous les dossiers du programme et les placer dans un même répertoire. Les fichiers à récupérer se trouvent dans le dossier "Script". Pour les récupérer, suivez le chemin suivant :

Nom	Modifié le	Type	Taille
 Doc	30/10/2023 09:52	Dossier de fichiers	
 Logs	30/10/2023 09:52	Dossier de fichiers	
 Script	30/10/2023 09:52	Dossier de fichiers	

Nom	Modifié le	Type	Taille
 __pycache__	30/10/2023 09:52	Dossier de fichiers	
 Analyse.fmw	04/10/2023 16:18	FME Workspace	116 Ko
 Analyse.log	04/10/2023 16:15	Document texte	14 Ko
 osmosecracker.py	30/10/2023 10:26	Fichier source Python	27 Ko
 osmosecracker_config.py	30/10/2023 10:23	Fichier source Python	8 Ko
 osmosecracker_database.sqlite	14/10/2023 11:53	Fichier SQLITE	3 812 Ko
 osmosecracker_database_management.py	30/10/2023 10:22	Fichier source Python	89 Ko
 osmosecracker_espacecollaboratifign.py	30/10/2023 10:21	Fichier source Python	4 Ko
 osmosecracker_exceptions.py	05/10/2023 16:28	Fichier source Python	2 Ko
 osmosecracker_issue.py	30/10/2023 10:20	Fichier source Python	26 Ko
 osmosecracker_query_bduni.py	30/10/2023 10:19	Fichier source Python	22 Ko
 osmosecracker_query_osm.py	02/10/2023 14:16	Fichier source Python	1 Ko
 osmosecracker_query_osmose.py	30/10/2023 10:19	Fichier source Python	12 Ko
 osmosecracker_workflow.py	02/10/2023 14:16	Fichier source Python	5 Ko

Ou via Git :

https://gitlab.ign.fr/bduni_prod/autres-projets/osmosecracker/-/tree/main/src?ref_type=heads

The screenshot shows the GitHub interface for the 'osmosecracker' repository. At the top, there's a navigation bar with 'main' selected, the repository path 'osmosecracker / src', and buttons for 'Find file', 'Edit', and 'Code'. Below this, a message indicates a recent edit to 'osmosecracker_issue.py'. The main area displays a table of files and their commit history.

Name	Last commit
..	
Analyse.fmw	Initial commit
Analyse.log	Initial commit
osmosecracker.py	osmosecracker.py
osmosecracker_config.py	Nouvelles class erreur amenant a des modifications de
osmosecracker_database_management.py	Deuxième commit avec les nouvelles modifications V1.1
osmosecracker_espacecollaboratifign.py	Deuxième commit avec les nouvelles modifications V1.1
osmosecracker_exceptions.py	Initial commit
osmosecracker_issue.py	Edit osmosecracker_issue.py all the details_descriptions
osmosecracker_query_bduni.py	Deuxième commit avec les nouvelles modifications V1.1 10 months ago
osmosecracker_query_osm.py	Deuxième commit avec les nouvelles modifications V1.1 10 months ago
osmosecracker_query_osmose.py	Deuxième commit avec les nouvelles modifications V1.1 10 months ago
osmosecracker_workflow.py	Deuxième commit avec les nouvelles modifications V1.1 10 months ago

A dropdown menu is open from the 'Code' button, showing options to clone the repository via HTTPS, open it in an IDE (Visual Studio Code or IntelliJ IDEA), or download the source code in various formats (zip, tar.gz, tar.bz2, tar). The 'Download source code' section is highlighted, with 'zip' selected.

Télécharger le Zip d'osmosecracker_main puis extraire le zip

VARIABLES D'ENVIRONNEMENT REQUISES

Pour sécuriser les identifiants (Login) et les mots de passe (Password) de l'ESPACECO et de la BDUNI, nous avons pris des variables environnements contenus dans l'OS des ordinateurs normalement habilités à lancer OSMOSECRAKER.

Ces Variables d'environnement sont appelées dans le fichier « osmosecracker_config.py »

LANCEMENT

Vous pouvez normalement démarrer le programme avec la commande suivante :

```
« "C:/Program Files/Python310/python.exe"
  "J:\DEP_ECHANGE\DATA C2023_OsmoseCracker\Script\osmosecracker.py"
  -{VARIABLE UTILISATEUR} {VALEUR DONNÉE À LA VARIABLE UTILISATEUR} -{...} {...} ... »
```

À noter que les variables utilisateur ne sont pas obligatoires à remplir, elles sont toutes prédéfinies. Cependant, si vous souhaitez les modifier, voici les variables utilisateur que vous pouvez utiliser :

- Si la requête n'a pour but que de rafraîchir les données sur le signalement de l'espace CO, « -ecso {...} » [False, True]. Par défaut, « False » pour ne pas qu'actualiser les données mais créer des signalements.
- Le pays/la zone géographique « -c {...} » [lien vers la liste des pays, régions et départements](#). Par défaut, « france* » pour collecter des données sur tout le territoire français.
- Le code INSEE des départements pour lesquels les signalements Osmose doivent être téléchargés « -fdep {...} » Tableau de string.
- Le code INSEE des régions pour lesquels les signalements Osmose doivent être téléchargés « -freg {...} » Tableau de string.
- La plage temporelle sur laquelle nous voulons récupérer des informations OSMOSE :
Pour le commencement « -sd {YYYY-MM-DD} ». Par défaut, le début du mois.
Pour la date de fin « -ed {YYYY-MM-DD} ». Par défaut, la date d'aujourd'hui.
- Le statut des données OSMOSE collectées « -s {...} » ["false", "done", "open"]. Par défaut, « "false" », car les données nous intéressant sont des faux positifs.
- L'utilisation des objets développeur ou non « -udi {...} » ["false", "true", "all"]. Par défaut, « "false" ».
- La source d'où proviennent les données, un paramètre **à ne pas utiliser** « -so {...} » pour plus d'informations [Wiki API OSMOSE](#) et [discussion avec OSMOSE issue #2363](#). Par défaut, ["*"] c'est une liste. ~~((Source [14708] À ne pas prendre pour une généralisation de la provenance des données IGN, les informations sur les sources d'OSMOSE sont très floues))~~
- Les objets que OSMOSECRAKER recherche, un paramètre à ne pas toucher sans modification de la source et des classes d'erreurs ((en développement)) « -it {...} ». Par défaut [7170] c'est une liste.
- La verbosité, c'est-à-dire le niveau de journalisation pour connaître les informations que la console vous renvoie « -ve {...} » ["DEBUG", "INFO", "WARN"]. Par défaut, la valeur est "INFO" pour ne pas avoir une console surchargée d'informations, mais seulement les informations sur le déroulement du programme.
- Si des signalements de l'espace CO sont émis ou non et de quel type « -ts {...} » ["skip", "test", "submit", "repost"]. Par défaut, « "skip" » pour ne pas émettre de signalement sans le faire exprès.
- Si le programme rédige un fichier CSV des données collectées ou non « -w {...} » [False, True]. Par défaut, False pour ne pas écrire automatiquement des CSV à chaque lancement du programme.

 <small>INSTITUT NATIONAL DE L'INFORMATION GÉOGRAPHIQUE ET FORESTIÈRE</small>	<h1>Documentation Utilisateur OSMOSECRAKER</h1>	DATAc Édition : 1 Date : 06/10/2023 Page: 5/15
---	---	--

Par défaut, le programme fait appel à ces 4 class erreur du même item (l'item 7170) :

Title	Item	Class
7170 (Objet de type Route)		
possibly missing highway in the area (BD Topo IGN)	7170	1
missing ref=* or misaligned road compared to BDTopo IGN	7170	3
Misaligned road compared to BDTopo IGN or bad highway=* type	7170	4
Road not integrated	7170	13
lanes=* missing on highway with more than 2 lanes (BD Topo)	7170	20

Pour information, il est complexe d'obtenir des informations très claires sur la provenance des combinaisons Item/class. Cependant, nous avons eu des informations sur la manière utilisée par OSMOSE pour diriger ses 'analyses' (les différents outils qui comparent les données OSM à une autre source de données) afin de renseigner une combinaison Item/class. Chaque classe d'erreur a **une** source et **une** erreur associée. Une personne d'OSMOSE nous a informés sur un certain nombre d'analyses qui utilisent les données IGN :

- [analyser merge cemetery FR.py](#)
- [analyser merge geodesie.py](#)
- [analyser merge geodesie support FR.py](#)
- [analyser merge man made FR.py](#)
- [analyser merge natural FR.py](#)
- [analyser merge poi FR.py](#)
- [analyser merge reservoir FR.py](#)
- [analyser merge restriction motorway FR.py](#)
- [analyser merge road FR.py](#)

- [NEW 24.10.21 analyser_merge_road_FR.py](#)
- [analyser_merge_water_FR.py](#)

Pour obtenir plus d'informations sur ces analyses, il faut aller voir [le GitHub d'OSMOSE](#) (ou lien cliquable direct vers les analyseurs). Pour trouver facilement des analyseurs liés à l'IGN, créez-vous un compte GitHub et cherchez le mot-clé « SourceIGN », qui est théoriquement présent dans tous les analyseurs en lien avec l'IGN.

PRINCIPAUX FICHIERS

Base de la structure du programme

`_pycache_` -> est un répertoire créé par l'interpréteur Python lorsqu'il importe un module.

`osmosecracker.py` -> est le squelette du programme il est le main, il contient le code capable d'effectuer des tests de connexion aux différents serveurs qui seront interrogés pour savoir si une erreur est constatée dans OSMOSE ou dans la BDuni. Il effectue également l'appel aux différentes parties du programme nécessaires au bon fonctionnement de ce dernier. Il fait l'intermédiaire entre les appels au serveur et la classe qui stocke en mémoire les informations clés des appels au serveur ; il fait aussi par la suite l'intermédiaire entre la classe et la base de données SQLite. Il fait enfin l'écriture en CSV des données collectées, si cette dernière est demandée.

`osmosecracker_config.py` -> contient tous les paramètres développés, parmi ces paramètres se trouvent les définitions en français des objets, des classes d'erreurs prévues initialement et les variables environnements.

`osmosecracker_exceptions.py` -> permet d'écrire des erreurs types pour simplifier la compréhension de ces dernières.

`osmosecracker_workflow.py` -> permet de connaître le temps écoulé entre chaque étape tout au long du programme de l'exécution du programme pour stocker ces informations. Ainsi, on peut savoir quand et comment notre programme a été utilisé.

Récolte des données

Externe à l'IGN

`osmosecracker_query_osm.py` -> permet de faire les appels à l'API d'OSM pour avoir des données complémentaires sur l'objet OSMOSE (en développement).

`osmosecracker_query_osmose.py` -> permet de faire les appels à l'API d'OSMOSE pour extraire les objets OSMOSE et avoir des informations complémentaires sur eux grâce à l'UUID collecté. Nous collectons également des informations sur l'Item et la Classe d'erreur.

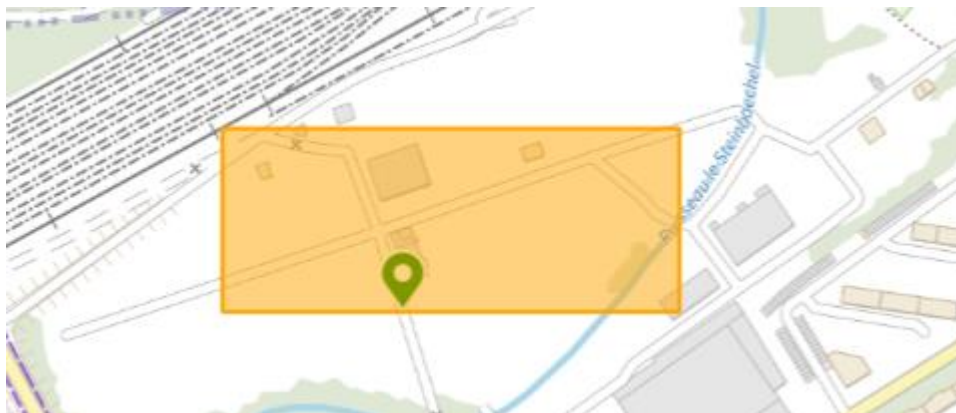
Interne à l'IGN

`osmosecracker_espacecollaboratifign.py` -> émet des signalements sur l'espace collaboratif IGN si de nouveaux signalements sont créés par OSMOSECRAKER. Il permet également d'interroger la base de données de l'espace CO pour actualiser le statut des signalements dans notre base de données SQLite.

`osmosecracker_query_bduni.py` -> permet de collecter des données de la BDUni pour avoir des informations sur l'objet du type renseigner à la coordonnée stockée.

Les informations collectées sont sur un point. Nous collectons des données telles que : le collecteur lié à la zone géographique, la reprojection des coordonnées, la commune, des informations sur l'objet et si le point se situe dans une ZICAD (pour ne pas émettre de signalement si c'est le cas).

Création de cluster pour réduire le nombre de signalements lorsque plusieurs signalements remplissent simultanément les conditions suivantes : proximité géographique (moins d'un kilomètre), appartenance à une même classe d'erreur, et proviennent de la même exécution du programme (récupérés au même moment). Ce regroupement vise à limiter les doublons dans les signalements.



bounding box = rectangle englobant les différents signalements d'un même cluster

Gestion du stockage en mémoire des données

`osmosecracker_issue.py` -> crée la classe Python qui stocke les données en mémoire en attendant leur persistance en base SQLite et l'écriture de la syntaxe du signalement de l'espace CO.

Gestion de la persistance et du stockage en base SQLite

`osmosecracker_database.sqlite` -> est le fichier qui stocke la base de données SQLite d'OSMOSECRAKER contenant les données OSMOSE, BDUni et autres dans la table `osmoseissue` et les informations liées au lancement du programme dans la table `workflowexecution`.

osmosecracker_database_management.py -> permet de stocker les données dans la base de données SQLite. Il permet également d'actualiser les informations sur le statut des signalements de l'espace CO et aussi de savoir où des données peuvent manquer pour pouvoir les ajouter automatiquement.

ANALYSE DES DONNÉES

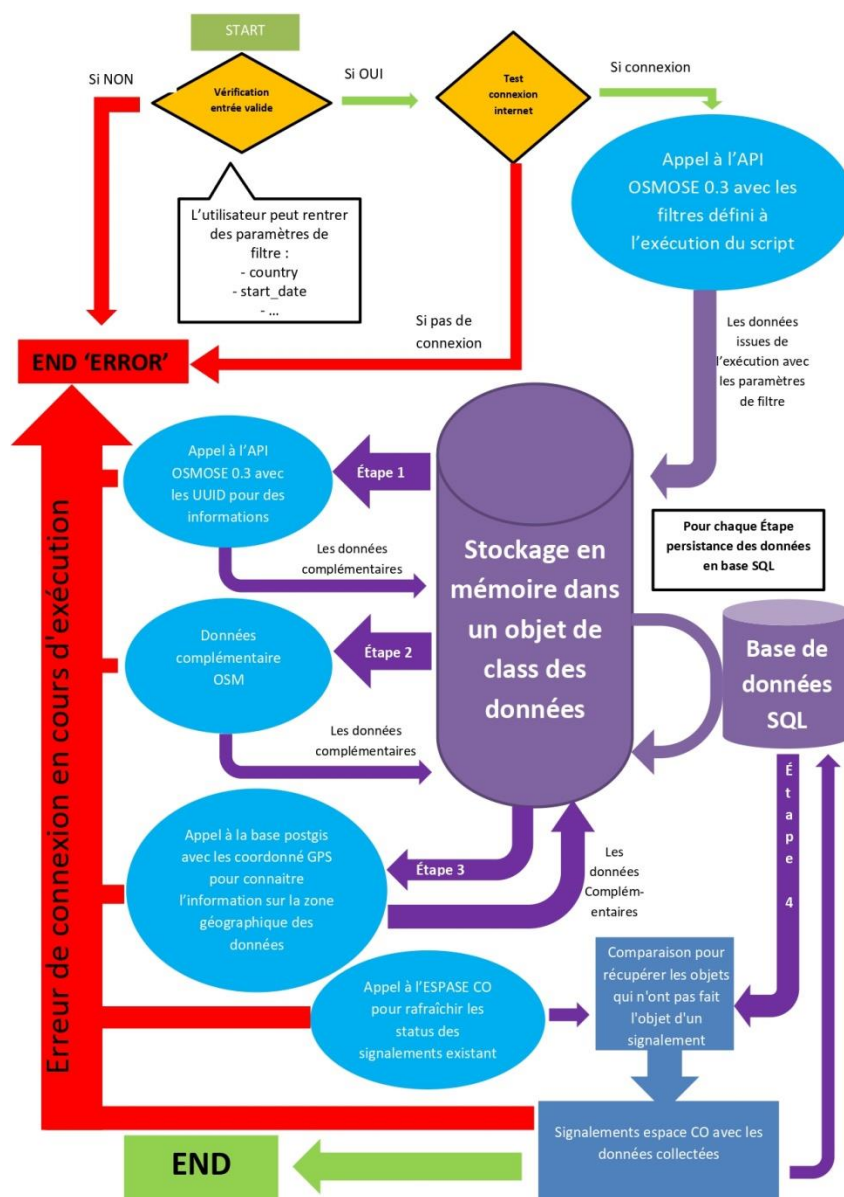
Pour connaître si les signalements ont été utiles ou non, il faut étudier les données «espaceco_signalement_status» les différents statuts sont le suivant :

Status	title
<u>submit</u>	<u>Reçu dans nos services</u>
pending0	En demande de qualification
pending	En cours de traitement
pending1	En attente de saisie
pending2	En attente de validation
<u>valid</u>	<u>Pris en compte</u>
<u>valid0</u>	<u>Déjà pris en compte</u>
<u>reject</u>	<u>Rejeté (hors spéc.)</u>
<u>reject0</u>	<u>Rejeté (hors de propos)</u>
<u>test</u>	<u>En mode test</u>
<u>None</u>	<u>En raison de la création d'un cluster incluant le signalement, le statut n'a pas été créé par le programme.</u>

(« Valid » étant donc le meilleur statut !)

ANNEXE

SCHEMA GENERAL DE OSMOSECRAKER



Donnée retourner

Nom de la colonne	Obligatoire	Type	Information	Provenance
core_id	x	VARC HAR	ID \n Identifiant unique de l'objet OSMOSE, plus d'informations disponibles sur " https://wiki.openstreetmap.org/wiki/ID ". Nous l'utilisons donc comme clé primaire pour OSMOSECRAKER.	OSMOSE general DATA
core_status	x	VARC HAR	Statut de signalement OSMOSE [false, done, open] \n Les statuts permettent de savoir l'état du signalement OSMOSE : "open" (pas encore traité) et "done" (changement effectué sur la base OSM) ne nous concernent pas. Nous recherchons des fausses positives, celles-ci ont un statut "false"	OSMOSE general DATA
core_lat	x	FLOA T	Latitude de l'objet en EPSG:4326 (WGS84)	OSMOSE general DATA
core_lon	x	FLOA T	Longitude de l'objet en EPSG:4326 (WGS84)	OSMOSE general DATA
core_item_id	x	VARC HAR	Numéro de l'item OSMOSE \n Tous les items OSMOSE ont un identifiant. Vous pouvez vous référer au lien suivant pour plus d'informations à leur sujet : " https://wiki.openstreetmap.org/wiki/Osmose/issues "	OSMOSE general DATA
core_item_name_auto	x	TEXT	Nom EN de l'item	OSMOSE general DATA
core_item_name_fr	x	TEXT	Nom FR de l'item \n Le nom français des items qui nous intéressent n'est pas toujours référencé. Nous avons donc lié chaque item ID à un nom français pour la future création de signalement	CREAT

core_source	x	INT	Numéro de la source \n La source est la combinaison de l'analyse utilisée pour la comparaison de la donnée OSM et de la donnée externe, ainsi que la zone géographique de la donnée. Cependant, cette dernière est très peu documentée	OSMOSE general DATA
core_class_id	x	INT	Numéro de l'erreur liée à l'item \n Chaque item a un certain nombre d'erreurs qui lui sont liées, et chacune d'elles a un numéro. Il convient cependant de noter que deux classes peuvent avoir le même numéro sans avoir de lien entre elles, tant que l'item auquel elles sont liées est différent	OSMOSE general DATA
core_class_name_auto	x	TEXT	Nom EN de l'erreur	OSMOSE general DATA
core_class_name_fr	x	TEXT	Le nom en français des classes d'erreur qui nous intéressent n'étant pas toujours référencé, nous avons associé chaque identifiant de classe (lié à un identifiant d'item) à un nom français en vue de la future création de signalements	CREAT
core_subtitle	x	TEXT	Informations complémentaires sur l'erreur \n Ce texte fournit des informations complémentaires sur le signalement OSMOSE. Cependant, la qualité des subtitles est variable	OSMOSE general DATA
core_country	x	TEXT	Zone de provenance (france*) \n La zone de provenance des données est la zone géographique sur laquelle nous avons effectué un filtre lors de l'appel au serveur OSMOSE, ici la france*. Cela permet de sélectionner les données de toute la France	OSMOSE general DATA
core_level	x	INT	Niveau de gravité de l'erreur entre 1 et 3	OSMOSE general DATA

core_update_timestamp	x	TEXT	Date du lancement du programme OSMOSE pour la création des signalements OSM	OSMOSE general DATA
core_usernames	x	TEXT	Nom de l'utilisateur qui a rapporté l'erreur (NULL) \n Jamais renseigné, probablement en raison du fait que les signalements sont générés de manière automatique	OSMOSE general DATA
core_osm_ids_elems	x	TEXT	Info OSMOSE sur nodes, way et relation	OSMOSE general DATA
core_osm_ids_nodes	x	TEXT	Noeuds OSM de l'objet	OSMOSE general DATA
core_osm_ids_ways	x	TEXT	Passages OSM de l'objet	OSMOSE general DATA
core_osm_ids_relations	x	TEXT	Relations OSM de l'objet	OSMOSE general DATA
details_descriptionstr		TEXT	Résumé (créé) au format Markdown des données pour le signalement \n À partir de toutes les données récoltées, nous créons un rapport Markdown à l'intention des lecteurs des signalements que OSMOSECRAKER crée. Ce rapport filtre les données utiles et crée un signalement dans un format fixe pour permettre une meilleure compréhension de celui-ci	OSMOSE uuid DATA
details_minlat		FLOAT	Zoom maximal prédéfini pour l'application Osmose sur l'issue \n latitude minimale de la Bbox	OSMOSE uuid DATA
details_maxlat		FLOAT	Zoom maximal prédéfini pour l'application Osmose sur l'issue \n latitude maximale de la Bbox	OSMOSE uuid DATA
details_minlon		FLOAT	Zoom maximal prédéfini pour l'application Osmose sur l'issue \n longitude minimale de la Bbox	OSMOSE uuid DATA

details_maxlon		FLOAT	Zoom maximal prédéfini pour l'application Osmose sur l'issue \n longitude maximale de la Bbox	OSMOSE uid DATA
details_b_date_datetime		TEXT	Date du signalement OSMOSE \n Date à laquelle OSMOSE a comparé les données IGN aux données OSM et a créé un signalement s'il y avait une incohérence entre les deux	OSMOSE uid DATA
details_osm_json_nodes		TEXT	Noeuds OSM de l'objet, détails	OSMOSE uid DATA
details_osm_json_ways		TEXT	Passages OSM de l'objet, détails	OSMOSE uid DATA
details_osm_json_relations		TEXT	Relations OSM de l'objet, détails	OSMOSE uid DATA
details_new_elems		TEXT	Nouveau élément lié à l'objet	OSMOSE uid DATA
osm_objects		TEXT	Objet OSM lié à l'objet OSMOSE [En développement]	DEVELOPEMENT
bduni_zone_collecte_collecteur		TEXT	Collecteur lié à la zone géographique de la BDuni où l'objet BDuni lié à l'objet OSMOSE a été collecté	BDUINI
espaceco_theme		TEXT	Objet BDuni lié à l'objet OSMOSE	CREAT
bduni_commune_code_insee		VARCHAR	ID INSEE BDuni de la commune	BDUINI
bduni_commune_nom_officiel		VARCHAR	Nom BDUni de la commune	BDUINI
bduni_canton_code_insee		VARCHAR	ID INSEE BDuni du canton	BDUINI
bduni_arrondissement_code_insee		VARCHAR	ID INSEE BDuni de l'arrondissement \n Plus d'information *****	BDUINI
bduni_arrondissement_nom_officiel		TEXT	Nom BDUni de l'arrondissement	BDUINI
bduni_collectivite_terr_code_insee		VARCHAR	ID INSEE BDuni de la collectivité territoriale \n Plus d'information *****	BDUINI
bduni_collectivite_terr_nom_officiel		TEXT	Nom BDUni de la collectivité territoriale	BDUINI

bduni_departement_code_insee		VARC HAR	ID INSEE BDuni du departement \n Plus d'information *****	BDUINI
bduni_departement_nom_officiel		TEXT	Nom BDuni du département	BDUINI
bduni_region_code_insee		VARC HAR	ID INSEE BDuni de la region \n Plus d'information *****	BDUINI
bduni_region_nom_officiel		TEXT	Nom BDUni de la region	BDUINI
bduni_territoire_nom		TEXT	Type de Territoire	BDUINI
bduni_territoire_srid		INT	Code EPSG du système de projection légal du territoire bduni_territoire_nom	BDUINI
bduni_x		FLOA T	Coordonnée X, système de projection légal du territoire bduni_territoire_nom	BDUINI
bduni_y		FLOA T	Coordonnée Y, système de projection légal du territoire bduni_territoire_nom	BDUINI
bduni_object_cleabs		TEXT	ID objet Bduni	BDUINI
bduni_objet_attribut_1		TEXT	SELECT BDuni atribut1	PARAMET RE
bduni_objet_attribut_2		TEXT	SELECT BDuni atribut2	PARAMET RE
bduni_objet_attribut_3		TEXT	SELECT BDuni atribut3	PARAMET RE
bduni_objet_attribut_4		TEXT	SELECT BDuni atribut4	PARAMET RE
bduni_objet_attribut_5		TEXT	SELECT BDuni atribut5	PARAMET RE
espaceco_signalement_id		INT	ID Espaceco du signalement de l'objet. Si négatif informe qu'il fait partie du même cluster que les signalements avec le même id	ESPACE CO
espaceco_signalement_status		TEXT	Statut Espaceco du signalement de l'objet	ESPACE CO
espaceco_signalement_status_r efresh_timestamp		TEXT	Date du dernier rafraîchissement de statut Espaceco du signalement de l'objet	ESPACE CO
bduni_zicad		BOOL EAN	Dans une ZICAD ou non \n Vérification que l'objet ne soit pas	BDUINI

			dans une ZICAD pour ne pas créer un signalement inutile	
core_classe_bduni	*	TEXT	Class de l'objet en BDUi	PARAMETRE
bduni_objet_date_modification		TEXT	La date de la dernière modification de l'objet en Bdunu n'est pas actualisée à chaque lancement de OSMOSECRAKCR. Cette information n'est pas mise à jour en raison du fait que les informations contenues dans le suivi d'OSMOSECRAKCR lié à la Bduni n'ont pas vocation à être mises à jour. Pour plus d'informations, veuillez suivre ce lien *****	BDUINI