

---

# "DMP du projet ANR EGOUT - Observations Géochimiques des Trajectoires Urbaines"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

## Renseignements sur le plan

<b>Titre du plan</b>	"DMP du projet ANR EGOUT - Observations Géochimiques des Trajectoires Urbaines"
<b>Version</b>	Version initiale
<b>Domaines de recherche (selon classification de l'OCDE)</b>	Earth and related environmental sciences, Social and economic geography, Health sciences, Sociology, Earth and related environmental sciences
<b>Langue</b>	fra
<b>Date de création</b>	2022-04-06
<b>Date de dernière modification</b>	2023-01-23
<b>Plans de gestion en lien avec le projet</b>	<ul style="list-style-type: none"><li>• Site web du projet : <a href="http://egout.cnrs.fr">egout.cnrs.fr</a></li></ul>

## Renseignements sur le projet

<b>Titre du projet</b>	Observations Géochimiques des Trajectoires Urbaines
<b>Acronyme</b>	EGOUT
<b>Résumé</b>	Alimentation, santé, bien-être, niveau de stress... Comment obtenir des renseignements pertinents sur les caractéristiques et l'évolution des systèmes urbains ? Par les sondages, les recensements et... les égouts ! Inspirés par les derniers développements en information chimique extraite des eaux usées, appuyés par nos succès récents, nous développerons un observatoire objectif, réactif et intégré du quotidien des parisiens par le suivi d'indicateurs comparables ou supplétifs des métriques en usage en Sciences Humaines et Sociales. Les cartographies géochimiques seront confrontées à celles des inégalités socio-économiques. Les trajectoires socio-économiques seront retracées via l'évolution des indicateurs géochimiques. Des actions de recherche interventionnelle impliquant une co-construction et une interprétation partagée des résultats promouvront l'implication des citoyens, des pouvoirs publics et des opérateurs.

<b>Sources de financement</b>	<ul style="list-style-type: none"><li>• Agence Nationale de la Recherche : ANR-21-CE03-0005-01</li></ul>
-------------------------------	--

<b>Date de début</b>	2022-01-01
----------------------	------------

<b>Date de fin</b>	2025-12-01
--------------------	------------

<b>Partenaires</b>	<ul style="list-style-type: none"><li>• Laboratoire des sciences du climat et de l'environnement (200611689J)</li><li>• Laboratoire Dynamiques Sociales et Recomposition des Espaces (199712591D)</li><li>• Laboratoire Eau Environnement et Systèmes Urbains (200920634U)</li><li>• Milieux Environnementaux, Transferts et Interactions dans les hydrosystèmes et les Sols (199712664H)</li></ul>
--------------------	---

## Produits de recherche :

1. Default research output

## Contributeurs

Nom	Affiliation	Rôles
Jacob Jérémy		<ul style="list-style-type: none"><li>• Coordinateur du projet</li><li>• Personne contact pour les données</li></ul>
Rican Stéphane - 0000-0002-7918-1339	LADYSS	<ul style="list-style-type: none"><li>• Responsable du plan de gestion de données</li></ul>

### Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source.  
L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

# "DMP du projet ANR EGOUT - Observations Géochimiques des Trajectoires Urbaines"

---

## 1. Description des données et collecte ou réutilisation de données existantes

### 1a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

Le projet EGOUT considère trois catégories de données (1) les données personnelles liées à l'organisation du projet, (2) la création de nouvelles données et (3) l'exploitation de données existantes

Chaque catégorie est distinguée en "Rubrique" puis en "Sous-rubrique ":

- données personnelles (données de contact),
- données bio-physico-chimiques (géochimie, biologie, physique...),
- données socio-économiques (santé, alimentation, pratiques, niveau de vie, niveau d'éducation, type de foyer, âge...),
- données de SIG.

Les données peuvent être sous statut "Propriétaire" ou "Ouvverte".

Les données socio-économiques peuvent être à l'échelle individuelle ou être agrégées (type maille IRIS).

Les données bio-physico-chimiques produites sur des échantillons par des tiers (Ville de Paris ou SIAAP, par exemple), sont documentées par ces tiers. Leur transmission au Consortium et leur diffusion est régie par un Material Transfer Agreement.

Les données bio-physico-chimiques acquises sur des échantillons (eaux, sédiments, carottes de sédiments) par les partenaires du projet font l'objet d'une documentation par le Consortium. Cette documentation permet le traçage de la donnée sur la base de l'identifiant unique de l'échantillon (IGSN, si possible), des fiches de campagne et de prélèvement qui explicitent les conditions de leur obtention (date, lieu, conditions...). Les méthodes de nommage seront convenues entre partenaires. La documentation comprend également toute la chaîne opératoire de la construction de la donnée en prenant en compte les traitements subis par les échantillons, les conditions opératoires de mesure, les procédures de validation de la donnée.

Les données socio-économiques sont soit recueillies auprès de producteurs externes au projet (INSEE, APUR, ARS...), soit produites dans le cadre du projet à partir d'enquêtes et d'entretiens avec les habitants.

Les données de SIG utilisées dans le cadre du projet sont soit celles accessibles de manière publique – Open Data sous Licence ODbL - Open Database License (Paris Data, APUR...), soit celles fournies par la Ville de Paris dans le cadre d'un Material Transfer Agreement qui définit les conditions de transmission, d'analyse et de diffusion de ces données.

Pour une meilleure navigation dans les tableaux listant les données, un numéro d'ordre leur est attribué dans la suite du DMP.

---

### 1b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

Les types, formats et volumes de données sont détaillés dans le tableau ci-dessous.

#### **Données personnelles liées à l'organisation du projet**

Numéro d'ordre	Rubrique	Sous-rubrique	Statut	Individuel/agrégé	Type	Description des données	Type de collecte	Format des données	Type de données	Volume
1	Données personnelles		Données de contact	Propriétaire	Individuel	Nom, prénom, fonction, établissement, adresse mail, cercle de communication, présence sur les réseaux sociaux	Création	Numérique	Tableur	ko
2	Liste de diffusion egout-copil@services.cnrs.fr		Liste de diffusion de mail	?						
3	Liste de diffusion egout_acteurs@services.cnrs.fr		?							
	Liste de diffusion egout-volontairesparis@services.cnrs.fr									
4	Liste de diffusion egout_comitemiroir@services.cnrs.fr		?							
5	Utilisateurs site web http://egout.cnrs.fr/		Nom, prénom, adresse mail	BDD						
6	Liste de diffusion egout_citoyens@services.cnrs.fr		Liste de diffusion de mail	?						

#### Création de nouvelles données

Numéro d'ordre	Rubrique	Sous-rubrique	Statut	Individuel/agrégé	Type	Indicateur	Description des données	Type de collecte	Format des données	Type de données	Volume
7	Données biophysico-chimiques	Échantillons		Fiches de prélèvements		Fiche de campagne (=suite logique de prélèvements)	Création	Numérique	Fichier .doc	ko	
8	Fiche de prélèvement de sédiments										
9	Fiche de prélèvement de carottes de sédiments										
10	Fiche de prélèvement d'eaux										
11	Carottes	Fiche de description de carottes									
12	Fiche d'échantillonnage de carottes sédimentaires										
13	Géochimie	Géochimie organique isotopique	d13C et d15N sur matières totales	Cahier de laboratoire	Papier	Cahier de laboratoire	-				
14	Runs bruts IRMS et EA	Automatique	Numérique	Tableur .xlsx	ko						
15	Rapport par run mesures d13C, C, N	Création									
16	Rapport pour série d'échantillons										
17	Analyses 14C	14C sur carbone organique total (dissout ou particulière)	Cahier de laboratoire, méta-données et datations	Base de Données	Mo						

18	Mesures 14C validées	Tableur .xlsx	ko			
19	Age des matières	Tableur .xlsx				
20	Géochimie organique moléculaire	Identification et quantification des biomarqueurs	Fiche de suivi des préparations pour biomarqueurs moléculaires	Tableur .xlsx	100 Ko	
21	Cahier de laboratoire - suivi de la GC-MS	Papier	Cahier de laboratoire			
22	Fichiers résultat de la GC-MS	Automatique	Numérique	Fichiers .raw	30 Mo	
23	Inventaire des analyses pour biomarqueurs moléculaires	Création	Tableur .xlsx	100 Ko		
24	Inventaire des biomarqueurs dans les échantillons	100 Ko				
25	Concentrations en biomarqueurs dans les échantillons	200 Ko				
26	Détermination des rapports isotopiques sur molécules individuelles	Cahier de laboratoire - suivi de la GC-irMS	Papier	Cahier de laboratoire		
27	Fichiers résultat brut de la GC-irMS pour les échantillons	Automatique	Numérique	Fichiers isodat	30 Mo	
28	Fichier intermédiaire de calcul des rapports isotopiques déterminés sur molécules individuelles dans les échantillons	Création	Tableur .xlsx	200 Ko		
29	Évolution des paramètres calculés à partir des concentrations ou des rapports isotopiques déterminés sur biomarqueurs moléculaires dans les échantillons	Fichier .pdf	20 Mo			
30	Éléments trace métalliques	Détermination des teneurs en éléments trace métalliques				
31						
32						
33						
34	Radionucléides	Activité de 137Cs, 210Pb, 7Be en Bq/kg	Cahier de laboratoire pour la préparation des échantillons	Création	Papier	Cahier de laboratoire

35	Cahier de laboratoire pour le suivi des mesures								
36	Fichiers résultats comportant les spectres (nbre de coups par énergie)	Numérique	.spe	10 Ko					
37	Fichier de traitement des résultats	Tableur	.xlsx	10 Ko					
38	Activités de 137Cs, 210Pb, 7Be en Bq/kg validées et corrigées de la décroissance radioactive à la date d'échantillonnage	200 Ko							
39	Médicaments et drogues	Identification et quantification médicaments et drogues	Cahier de laboratoire de suivi de l'HPLC-QQQ	Papier	Cahier de laboratoire	-			
40	Fichier de résultats bruts de l'HPLC-QQQ	Automatique	Numérique	Fichiers .bin	4 Mo				
41	Fichier de calcul des concentration en biomarqueurs	Création	Tableur .xlsx	10 ko					
42	Analyses par HR-MS	Empreintes moléculaires d'échantillons environnementaux	Cahier de suivi de l'appareil	15 ko					
43	Fiches protocole	.docx	10 Ko						
44	Cahiers de laboratoire des manipulateurs	Papier	Cahier de laboratoire	-					
45	Fichier de suivi des échantillons	Numérique	Tableur .xlsx	10 Ko					
46	Fichiers bruts	.uep	3 Go						
47	Fichiers convertis	.mz5	400 Mo						
48	Fichier final par type de matrice	Tableur .xlsx	1 Mo						
58	Propriétaire	Agrégré	Données géochimiques de curage des bassins de dessablement	Métaux, HAP...	Cartes de distribution des HAP et métaux	SIG	Mo		
64	Données de débit dans le réseau	Débits/volumes mesurés par capteurs...	Mesures réglementaires réalisées dans le cadre des opérations régulières de la Ville de Paris	SIG					
61	Données SIG	SIG	Données extraites de la BDD TIGRE7 de la Ville de Paris	Réseau d'assainissement, localisation et propriétés des bassins de dessablement...	Cartes des bassins d'apport des bassins de décantation	SIG			
49	Données socio-économiques	Pratiques de consommation et rejet dans les eaux usées	-	Individuel	Entretiens individuels	Pratiques de consommation et rejets dans les eaux usées	Carnet de prise de notes pendant les entretiens	Papier	Carnet de notes

50	Captation audio des entretiens	Numérique	Fichier audio	10 Mo	
51	Agrégré	Transcription en encodage des entretiens	Fichier.pdf / .docx	10 ko	
52	Rapport de synthèse des entretiens	Fichier.pdf / Tableur.xlsx	1 Mo		
53	Individuel	Questionnaire en ligne	Questionnaire individuel anonymisé recensant les pratiques et usages	Tableur .xlsx	10 ko
54	Agrégré	Rapport de synthèse sur les résultats du questionnaire	Fichier.pdf / .docx	10 ko	

**Données existantes**

Numéro d'ordre	Rubrique	Sous-rubrique	Statut	Individuel/agrégé	Type	Indicateur	Description des données	Type de collecte	Format des données	Type de données	Volume
56	Données bio-physico-chimiques	Géochimie	Propriétaire	Agrégé	Données géochimiques de curage des bassins de dessablement	Métaux, HAP...	Tableaux de données comportant les mesures réglementaires réalisées lors des curages des bassins de dessablement	Existant	Numérique	Tableur	Ko
57	Données intégrées dans un SIG	Création	SIG	Mo							
59	Données SIG	SIG	Données extraites de la BDD TIGRE7 de la Ville de Paris	Réseau d'assainissement, localisation et propriétés des bassins de dessablement...	Export de la BDD TIGRE7 sous forme Json, xml, shp ou autre	Existant					
60	Données intégrées dans un SIG	Création									
62	Données bio-physico-chimiques	Physique	Données de débit dans le réseau	Débits/volumes mesurés par capteurs...	Mesures réglementaires réalisées dans le cadre des opérations régulières de la Ville de Paris	Existant	Tableur			Ko	
63	Création	SIG	Mo								
65	Données socio-économiques		Ouverte	Données de l'APUR	Données socio-économiques, de santé, de consommation, d'urbanisme....	Morphologie urbaine					
66				Données de l'INSEE	Caractéristiques socio-démographiques						
67		Ouverte	Agrégé	Caractéristiques socio-démographiques							
68		Données de Paris Data									
69	Santé			Données de l'ARS	Offre et consommation de soins						
70	Ouverte	Agrégé	Offre et consommation de soins								

## 2. Documentation et qualité des données

### 2a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?

Les données bio-physico-chimiques seront accompagnées de la documentation permettant de tracer le processus d'acquisition depuis la collecte des échantillons jusqu'à la mesure et sa validation. Pour les données provenant de tiers (Ville de Paris ou SIAAP), il sera systématiquement fait mention du producteur originel de la donnée. Si possible, les conditions d'acquisition de ces données fournies par des tiers seront indiquées. De même, les données socio-économiques seront accompagnées d'une documentation explicitant le mode de production, qu'il s'agisse de données existantes provenant de tiers ou de données créées dans le cadre du projet.



## Données personnelles liées à l'organisation du projet

Sans objet

### Création de nouvelles données

Numéro d'ordre	Documentation/qualité des données
7	Contrôle par l'équipe de prélèvement
8	
9	
10	
11	
12	
13	Procédures internes
14	Fichiers bruts
15	Fichiers bruts compilés
16	Le code échantillon associé à la fiche prélèvement, au cahier de laboratoire, aux runs bruts et rapports par mesure permettent le suivi qualité de la mesure.
17	Procédures internes
18	
19	Le code échantillon associé à la fiche prélèvement, au cahier de laboratoire, aux métadonnées dans la BDD 14C et les validations permettent le suivi qualité de la mesure.
20	Procédures internes
21	
22	Fichiers bruts
23	Validation par opérateur
24	Le code échantillon associé à la fiche prélèvement, à la fiche et au cahier de suivi et aux fichiers .raw permettent le suivi qualité de la mesure.
25	
26	Procédures internes
27	Fichiers bruts
28	Validation par opérateur
29	Le code échantillon associé à la fiche prélèvement, à la fiche et au cahier de suivi, au fichier de résultats bruts et aux calculs des rapports isotopiques (avec standards) permettent le suivi qualité de la mesure.
30	
31	
32	
33	
34	Procédures internes
35	
36	Fichiers bruts
37	Validation par opérateur
38	Le code échantillon associé à la fiche prélèvement, à la fiche et aux cahiers de suivi (préparation et mesure), au fichier de résultats bruts et aux calculs des activités (avec standards/bruit de fond) permettent le suivi qualité de la mesure.
39	Procédures internes
40	Fichiers bruts
41	Validation par opérateur
42	Suivi qualité et maintenance de l'appareil
43	Fiches explicitant les protocoles
44	Recense les échantillons et les traitements subis selon les fiches protocole
45	Recense les échantillons avec des numéros d'ordre interne, les traitements subis (selon les fiches protocole) et les analyses réalisées.
46	Résultat brut issu de l'appareil
47	Fichier converti, utilisable par les logiciels de traitement
48	Filtrage de l'ensemble des signaux alignés recensant l'ensemble des constituants (temps de rétention, m/z, intensité) d'une série d'échantillons

58	Lecture critique par C. Asselin dans le cadre de sa thèse. Retour à la Ville de Paris.
64	Lecture critique par Ingénieur SI. Retour à la Ville de Paris.
61	Lecture critique par C. Asselin dans le cadre de sa thèse. Retour à la Ville de Paris.
49	Données brutes acquises durant l'entretien
50	Données brutes acquises durant l'entretien
51	Qualification des données selon les protocoles en vigueur au LADYSS
52	Synthèse des données selon expertise et traitements en vigueur dans la communauté et en particulier au LADYSS
53	Données brutes acquises via le questionnaire
54	Synthèse des données selon expertise et traitements en vigueur dans la communauté et en particulier au LADYSS

#### Données existantes

Numéro d'ordre	Documentation/qualité des données
56	Lecture critique par C. Asselin dans le cadre de sa thèse. Retour à la Ville de Paris.
57	
59	
60	
62	Lecture critique par Ingénieur SI. Retour à la Ville de Paris.
63	
65	
66	
67	
68	
69	
70	

---

#### 2b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

Les mesures de contrôle de la qualité de la données seront mise en œuvre selon les pratiques de chaque partenaire, selon le tableau indiqué en 2.a.

---

### 3. Stockage et sauvegarde pendant le processus de recherche

#### 3a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

##### Données personnelles liées à l'organisation du projet

Numéro d'ordre	Sauvegarde existante	Sauvegarde prévue
1	Espace collaboratif CoRe	Espace collaboratif CoRe
2	Services CNRS	CNRS
3		
4		
5	ODS CNRS	
6	Services CNRS	

##### Création de nouvelles données

Numéro d'ordre	Sauvegarde existante	Sauvegarde prévue
7	Disque dur externe	Espace collaboratif CoRe
8		
9		
10		

11		
12		
13	LSCE	
14	Disque dur externe et serveur LSCE	Espace collaboratif CoRe
15	Disque dur externe et serveur LSCE	
16	Disque dur externe	Système d'information du projet
17	Bdd14C - Serveur LSCE	
18	Bdd14C - Serveur LSCE	Espace collaboratif CoRe
19	Bdd14C - Serveur LSCE	Système d'information du projet
20	Bdd14C - Serveur LSCE	Espace collaboratif CoRe
21	LSCE	
22	Disque dur externe	Espace collaboratif CoRe
23	Disque dur externe	
24	Disque dur externe	Système d'information du projet
25	Disque dur externe	
26	LSCE	
27	Disque dur externe	Espace collaboratif CoRe
28	Disque dur externe	
29	Disque dur externe	Système d'information du projet
30		
31		
32		
33		
34	LSCE	LSCE
35	LSCE	
36	Ordinateur de pilotage et Serveur LSCE	Ordinateur de pilotage et Serveur LSCE
37	Serveur LSCE	Serveur LSCE
38	Serveur LSCE	Serveur LSCE et Système d'information du projet
39	METIS	METIS
40	Serveur METIS	Serveur METIS
41	Disque dur externe	Espace collaboratif CoRe / Système d'information du projet
42	OSU-EFLUVE	OSU-EFLUVE
43	Espace collaboratif Leesu	Espace collaboratif CoRe
44	LEESU	LEESU
45	Espace collaboratif Leesu	Espace collaboratif Leesu
46	Ordinateur de pilotage de l'appareil et Serveur de l'OSU-EFLUVE	Ordinateur de pilotage de l'appareil et Serveur de l'OSU EFFLUVE
47	Ordinateur de l'utilisateur et Serveur du LEESU	Ordinateur de l'utilisateur et Serveur du LEESU
48	Ordinateur de l'utilisateur et Serveur du LEESU	Ordinateur de l'utilisateur, Serveur du LEESU et Espace CoRe
58	Ordinateur de Camille Asselin / Disque dur externe	Système d'information du projet
64	-	
61	Ordinateur de Camille Asselin / Disque dur externe	
49	Coffre fort dédié au LADYSS.	Destruction après transcription et codage des données.
50	Disque dur externe dans coffre fort dédié au LADYSS.	Suppression après transcription et codage des données.
51	Disque dur externe / Ordinateur professionnel	Espace collaboratif CoRe
52	Disque dur externe / Ordinateur professionnel	
53	Disque dur externe / Ordinateur professionnel	
54	Disque dur externe / Ordinateur professionnel	

**Données existantes**

Numéro d'ordre	Sauvegarde existante	Sauvegarde prévue
57	Ordinateur de Camille Asselin / Disque dur externe	Système d'information du projet
56	Ordinateur de Camille Asselin / Disque dur externe / Serveurs et fichiers de la Ville de Paris	
59		
60		
62	Serveurs et fichiers de la Ville de Paris	
63	-	
65	Publique	
66	Serveurs de l'INSEE (privé)	
67	Serveurs de l'INSEE (publique)	
68	Serveurs de la Ville de Paris (publique)	
69	Serveurs de l'ARS (privé)	
70	Serveurs de l'ARS (publique)	

### 3b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

La sécurité des données personnelles sera garantie par les partenaires selon leurs pratiques habituelles.

La définition des données sensibles (hors données à caractère personnel) sera établie avec l'ensemble des partenaires et le Comité Miroir.

## 4. Exigences légales et éthiques, codes de conduite

### 4a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

Les données à caractère personnel gérées dans le cadre du projet EGOUT sont de plusieurs types.

#### Données personnelles liées à l'organisation du projet

- Les données à caractère personnel (nom, prénom, adresse mail, structure de rattachement) des participants au projet (Consortium, Comité Miroir, Ville de Paris, SIAAP, ASTS et partenaires s'agrégeant au projet au fur et à mesure) sont inventoriées dans un tableur stocké sur l'Espace CoRE CNRS du projet.

- Les données de contact des citoyens impliqués dans le projet ont été collectées dans le cadre d'une enquête réalisée par la Ville de Paris à l'été 2022. Le fichier tableur de réponse au questionnaire contient le nom, prénom, numéro de téléphone et adresse mail des citoyens. Leur adresse mail est utilisée pour établir la liste de diffusion [egout-volontairesparis@services.cnrs.fr](mailto:egout-volontairesparis@services.cnrs.fr). Le tableur est stocké dans l'espace CoRE CNRS du projet EGOUT.

- Les adresses mails des participants au projet sont agrégées dans des listes de diffusion gérées par les services du CNRS.

Numéro d'ordre	Respect de la législation sur les données personnelles	Ethique et déontologie
1	Compatibilité RGPD	Données non diffusables
2		
3		
3bis		
4		
5		
6		

#### Création de nouvelles données

Pour les besoins du projet, les participants seront amenés à collecter des informations à caractère personnel dans le cadre d'enquêtes et d'entretiens réalisées par le LADYSS ou les partenaires du projet (Ville de Paris et ASTS). Le traitement de ces données sera réalisé selon les standards en vigueur chez chaque partenaire. Pour le Consortium, le LADYSS assurera le respect des dispositions légales sur les données à caractère personnel et sur la sécurité des données avec le soutien de la Déléguée à la Protection des Données du LADYSS. Les données Bio-Physico-Chimiques peuvent également revêtir un caractère personnel. Si tel est le cas (populations potentiellement stigmatisées), nous nous appuierons sur les règles déontologiques en vigueur.

Numéro d'ordre	Respect de la législation sur les données personnelles	Titularité / propriété	Ethique et déontologie
7	Sans objet	Usage interne - LSCE	Données non diffusées

8		
9		
10		
11		
12		
13		
14		
15		
16	LSCE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
17	Usage interne - LSCE	Données non diffusées
18		
19	LSCE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
20	Usage interne - LSCE	Données non diffusées
21		
22		
23		
24		
25	LSCE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
26	Usage interne - LSCE	Données non diffusées
27		
28		
29	LSCE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
30		
31		
32		
33		
34	Usage interne - LSCE	Données non diffusées
35		
36		
37		
38	LSCE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
39	METIS	Données non diffusées
40		
41		Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
42	LEESU	Données non diffusées
43		
44		
45		
46		
47		Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
48		
58	Données agrégées respectant l'anonymat	Ville de Paris / LSCE

64	Sans objet	Ville de Paris	
61	Données agrégées respectant l'anonymat	Ville de Paris / LSCE	
49	Compatibilité RGPD : le nom des personnes répondant aux entretiens sont archivés dans des conditions de sécurité respectant la RGPD. Les données sont ensuite anonymisées lors des traitements et n'apparaissent pas lors de la diffusion des résultats.	LADYSS	Données non diffusables
50			
51			
52	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir		
53	Données non diffusables		
54	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir		

#### Données existantes

Pour les besoins du projet, nous solliciterons les structures tiers (INSEE, APUR, ARS, Ville de Paris...) pour accéder à des données revêtant un caractère personnel, par exemple des résultats de recensements à l'échelle des foyers. Ces données feront l'objet d'un traitement défini dans le cadre de la convention d'accès aux données entre le Consortium et le fournisseur de données. Si nécessaire, nous solliciterons les données initiales non anonymisées auprès des producteurs (BDD du recensement avec géolocalisation, par exemple). Dans ce cas, nous ferons appel au Centre d'accès sécurisé aux données (CASD) et suivrons la procédure d'habilitation après avoir présenté le projet du Comité du Secret.

Numéro d'ordre	Respect de la législation sur les données personnelles	Titularité / propriété	Ethique et déontologie
56	Sans objet	Ville de Paris	Données non diffusées
57			
59			
60			
62			
63			
65	Données agrégées respectant l'anonymat	APUR	Données publiques
66	Compatibilité RGPD	INSEE	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
67	Données agrégées respectant l'anonymat	Données publiques	
68	Ville de Paris		
69	Compatibilité RGPD	ARS	Soumis aux règles de diffusion gérées par la convention et après validation par le Comité Miroir
70	Données agrégées respectant l'anonymat	Données publiques	

#### 4b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

Chaque producteur de données conservera la propriété des données qu'il produit. La publication ou le partage des données produites ne pourront être réalisés qu'avec l'accord du producteur de la donnée, qu'il soit interne au Consortium ou externe.

#### 4c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

Les règles de fonctionnement du projet, notamment celles gérant les questions de publications scientifiques respecteront les préconisations décrites dans le "Guide pour une pratique de la recherche intégrée et responsable" (Comité d'Ethique du CNRS, septembre 2016). Le Comité Miroir donnera au responsable du projet un avis sur la diffusion des données ou publications.

Les règles d'anonymisation des résultats de suivi des traceurs dans les eaux usées suivront les standards en vigueur en Wastewater-Based Epidemiology (Lancaster et al., 2019 - [10.1016/j.drugpo.2018.10.011](https://doi.org/10.1016/j.drugpo.2018.10.011) et Prichard et al., 2014 - [10.1016/j.scitotenv.2013.11.039](https://doi.org/10.1016/j.scitotenv.2013.11.039)).

La décision de diffusion des résultats et publications relèvera au final du porteur du projet, en accord avec l'Administrateur de la Donnée et les membres du COPIL. Le responsable du projet peut décider du niveau d'anonymisation des données publiées mais est responsable d'éventuels problèmes survenant après cette publication.

---

## 5. Partage des données et conservation à long terme

### 5a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

Seules les données interprétées et validées par les producteurs et/ou les membres du Consortium pourront être partagées. Les données brutes seront conservées par les parties. Les membres du Consortium se réservent le droit d'appliquer un embargo pour publication des données avant leur diffusion dans des publications scientifiques.

Les données pourront être partagées sous diverses formes au gré des habitudes de chaque partenaires/chaque discipline : plateformes de partage (Pangea, NOAA, Data Terra, Gaia Data, Paris Data), data paper, diffusion sur portail web.

Indépendamment du cadre de partage des données (portail du projet, publication scientifique, data paper, portail de données national ou international...), ce partage devra être réalisé après avoir obtenu un avis du Comité Miroir.

#### Création de nouvelles données

Numéro d'ordre	Partage des données et conservation à long terme
7	Non partagé
8	
9	
10	
11	
12	
13	
14	
15	
16	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
17	Non partagé
18	
19	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
20	Non partagé
21	
22	
23	
24	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
25	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
26	Non partagé
27	
28	
29	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
30	
31	
32	
33	
34	Non partagé
35	
36	
37	
38	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
39	Non partagé
40	
41	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)

42	Non partagé
43	
44	
45	
46	
47	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
48	
58	Système d'information du projet / Publications scientifiques / Paris Data
64	
61	
49	Non partagé
50	
51	
52	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)
53	Non partagé
54	Publication scientifique + Plateformes locales (Paris Data), nationales (DataTerra) et internationales (Pangea, Zenodo...)

#### Données existantes

Numéro d'ordre	Partage des données et conservation à long terme
57	Données non partagées / Système d'information du projet
59	
60	
63	
65	Publique
67	
68	
70	
56	Système d'information du projet / Data Paper ?
62	
66	Système d'information du projet / Publications scientifiques
69	

---

#### 5b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?

Les données brutes, les informations sur les campagnes de prélèvement, échantillons, traitements, analyses, seront gérées par les producteurs de données selon le tableau fourni en 5.a. L'archivage à long terme des données validées sera assuré sur des entrepôts de données comme décrit au point 5a.

---

#### 5c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?

Le Système d'Information du projet EGOUT, hébergé par des serveurs du LSCE, sera accessible via des webservices. Il permettra d'accéder aux données produites dans le cadre du projet. Il servira également de portail vers des données d'intérêt pour le projet hébergées ou accessibles sur d'autres serveurs/portails, par exemple les données des partenaires du projet : SIAAP et Ville de Paris, ou données publiques (APUR, INSEE, ARS...). Les données brutes, les informations sur les campagnes de prélèvement, échantillons, traitements, analyses, seront gérées par les producteurs de données selon le tableau donné en 5c. Elles seront disponibles sur requête auprès des producteurs de la donnée.

---

#### 5d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?



Les jeux de données validés et déposés sur les entrepôts de données seront identifiés selon les protocoles de ces entrepôts, au gré des habitudes des producteurs de données (cf. 5.a)

## 6. Responsabilités et ressources en matière de gestion des données

6a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?

Stéphane Rican (LADYSS) est Administrateur de la Donnée du projet EGOUT. Il veille à la gestion des données, à leur qualification, à leur stockage, à leur sauvegarde, à leur archivage sur le long terme, à leur partage et au respect des normes éthiques. Avec l'Ingénieur Système d'Information recruté sur le projet, il s'assure de la mise en œuvre du PGD et des éventuelles mises à jour. En lien avec le COPIL du projet, Stéphane Rican veille à la coordination entre les producteurs/responsables de chaque donnée décrites dans le tableau ci-dessous, organisé par Work Package.

### Données personnelles liées à l'organisation du projet

Numéro d'ordre	Nom du responsable	Administrateur de la donnée	WP
1	Jérémy Jacob	Stéphane Rican	WP1
2	WP2		
3	WP3		
4	WP4		
5	WP5		
6	Catherine Carré	WP3	

### Création de nouvelles données

Numéro d'ordre	Nom du responsable	Administrateur de la donnée	WP
7	Jérémy Jacob	Stéphane Rican	WP 2, 3, 4
8	WP 2		
9	WP 4		
10	WP 3, 4		
11	WP 4		
12	WP 4		
13	Caroline Gauthier	WP 2, 3, 4	
14	WP 2, 3, 4		
15	WP 2, 3, 4		
16	Christine Hatté	WP 2, 3, 4	
17	WP 2, 3, 4		
18	WP 2, 3, 4		
19	WP 2, 3, 4		
20	Jérémy Jacob	WP 2, 3, 4	
21	WP 2, 3, 4		
22	WP 2, 3, 4		
23	WP 2, 3, 4		
24	WP 2, 3, 4		
25	WP 2, 3, 4		
26	WP 2, 3, 4		
27	WP 2, 3, 4		
28	WP 2, 3, 4		
29	WP 2, 3, 4		
30	Sophie Ayrault	WP 2,4	
31	WP 2,4		
32	WP 2,4		
33	WP 2,4		
34	Olivier Evrard	WP 2, 4	

35	WP 2, 4	
36	WP 2, 4	
37	WP 2, 4	
38	WP 2, 4	
39	Fabrice Alliot	WP 2, 3, 4
40	WP 2, 3, 4	
41	Thomas Thiebault	WP 2, 3, 4
42	Julien Le Roux	WP 2, 3, 4
43	WP 2, 3, 4	
44	WP 2, 3, 4	
45	WP 2, 3, 4	
46	WP 2, 3, 4	
47	WP 2, 3, 4	
48	WP 2, 3, 4	
58	Camille Asselin	WP 2
64	Jérémy Jacob	WP 4
61	Camille Asselin	WP 2, 3, 4
49	Catherine Carré	WP 2, 3
50	WP 2, 3	
51	WP 2, 3	
52	WP 2, 3	
53	WP 2, 3	
54	WP 2, 3	

**Données existantes**

Numéro d'ordre	Nom du responsable	Administrateur de la donnée	WP
56	Camille Asselin	Stéphane Rican	WP2
57	WP2		
59	WP 2, 3, 4		
60	WP 2, 3, 4		
62	Jérémy Jacob	WP 4	
63	WP 4		
65	Publique	WP 2, 3, 4	
66	Stéphane Rican	WP 2, 3, 4	
67	Publique	WP 2, 3, 4	
68	WP 2, 3, 4		
69	WP 2, 3, 4		
70	WP 2, 3, 4		

**6b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?**

Chaque producteur de donnée dédiera une partie de son temps à la curation des données qu'il produit, selon ses pratiques actuelles. Un Ingénieur Système d'Information recruté sur le projet sera chargé prioritairement de la coordination de la curation des données, sous la responsabilité de l'Administrateur de la Donnée (S. Rican). Les frais de stockage seront pris en charge par les partenaires sur des crédits du projet (si nouveaux besoins de stockage) ou sur crédits propres (stockages existants). Le LSCE veillera à la maintenance des serveurs supportant le Système d'Information du projet.