



HAL
open science

Plan de gestion de données de l'unité expérimentale EASM (Unité Expérimentale Système d'Élevage Avicoles Alternatifs)

Karine Germain, Anissa Dieudonné, Karine Météau

► To cite this version:

Karine Germain, Anissa Dieudonné, Karine Météau. Plan de gestion de données de l'unité expérimentale EASM (Unité Expérimentale Système d'Élevage Avicoles Alternatifs). INRAE. 2024. hal-04706090

HAL Id: hal-04706090

<https://hal.inrae.fr/hal-04706090v1>

Submitted on 23 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

PLAN DE GESTION DES DONNEES (PGD) DE L'UNITE EXPERIMENTALE SYSTEMES D'ELEVAGE AVICOLES ALTERNATIFS (UE EASM)

INFORMATIONS GENERALES

RENSEIGNEMENTS SUR LE PLAN

Titre du plan	Plan de gestion des données (PGD) de l'unité expérimentale systèmes d'élevage avicoles alternatifs (UE EASM)
Version	Version 1
Domaines de recherche (selon classification de l'OCDE)	Agriculture, forestry, and fisheries
Langue	Français
Date de création	2024-07-30
Date de dernière modification	2024-07-30
Plans de gestion en lien avec le projet ou l'activité de l'entité	PGD-SICPA (Systèmes d'Informations et Calcul pour le Phénotypage Animal) : 10.57745/ACTK8H

DETAILS DE LA STRUCTURE

Nom de l'entité	UE 1206 EASM (Unité Expérimentale Systèmes d'Elevage Avicoles Alternatifs)
Acronyme	EASM
Identifiant	10.15454/1.5572418326133655E12
Type d'identifiant	DOI
Description	L'unité EASM est une Unité Expérimentale du département PHASE. Les recherches menées à l'UE EASM visent à concevoir des systèmes de production avicole innovants et durables (élevage biologique ou pratiques en rupture), à conduire des expérimentations dans le domaine de la nutrition avicole et à définir des conditions d'élevage préservant la santé et le bien-être des animaux pour répondre au mieux aux préoccupations de la filière. D'autre part, les laboratoires de l'unité réalisent le suivi sanitaire des lignées aviaires du département SA et des tests d'évaluation sensorielle des viandes.
Date de création	2003-01-01
Tutelle(s) de l'entité	Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement https://ror.org/003vg9w96

CONTRIBUTEURS

Nom	Affiliation	Rôles
Dieudonné Anissa 0000-0003-3477-949X	INRAE	https://ror.org/003vg9w96 Personne contact
Germain Karine 0009-0005-6638-9404	INRAE	https://ror.org/003vg9w96 Responsable de l'entité
Meteau Karine 0009-0005-7819-8285	INRAE	https://ror.org/003vg9w96

PLAN DE GESTION DES DONNEES

INFORMATIONS SUR LA STRUCTURE

❖ *NOM DE LA STRUCTURE*

UE 1206 EASM (Unité Expérimentale Systèmes d'Elevage Avicoles Alternatifs)

❖ *TYPE DE STRUCTURE*

Unité de recherche, Unité ou Installation Expérimentale, Infrastructures Scientifiques Collectives (ISC)

❖ *IDENTIFIANT DE LA STRUCTURE PRÉCISER LE FOURNISSEUR DE L'IDENTIFIANT (ISNI, VIAF, FUNDREF, DATA CITE...).*

DOI : 10.15454/1.5572418326133655E12 (DataCite)

❖ *RESPONSABILITES DANS LA STRUCTURE*

Nom, Prénom	Courriel	Rôle
Germain, Karine	karine.germain@inrae.fr	Directrice d'unité
Dieudonné, Anissa	anissa.dieudonne@inrae.fr	Référente données opérationnelles

❖ *ETABLISSEMENT(S) TUTELLE(S)*

INRAE - Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

❖ *DEPARTEMENT DE RATTACHEMENT INRAE (OU ANCIENS DEPARTEMENTS INRA)*

PHASE : Physiologie animale et élevages

❖ *FINANCEUR(S) (PERMETTANT L'ACQUISITION DES JEUX DE DONNEES – HORS PROJET)*

INRAE (département, CNUE, prévention)

INFORMATIONS SUR LE PLAN DE GESTION

❖ *DOI (VERSION PUBLIEE DU PLAN DE GESTION)*

A prévoir

❖ *HISTORIQUE*

DES

VERSIONS

Date	N° de version	Statu	Auteur	Affiliation de l'auteur	Validé par	Validé le
05/09/2024	V1	En rédaction	Anissa Dieudonné	EASM -Unité Expérimentale Système d'Elevage Avicoles Alternatifs - INRAE	Karine Germain	06/09/2024

PRESENTATION GENERALE DES DONNEES

❖ *MODE D'OBTENTION DES DONNEES*

- Données générées par la structure : Les données sont collectées directement dans les élevages et laboratoires de l'unité EASM. Ces données peuvent être acquises avec ou sans instruments, avec ou sans intervention humaine. La majorité des données collectées sont retranscrites sur des fichiers informatisés.
- Données produites par un tiers : analyses sous-traitées (prestataire de service ou partenaire INRAE) et données ensuite envoyées à l'unité EASM.

❖ *ORIGINE*

- Analyse
- Enquête
- Expérimentation
- Observation

❖ *TYPE DE DONNEES*

- Audiovisuel
- Dataset
- Image
- Sound
- Text

❖ *NATURE DES DONNEES*

L'unité distingue les données :

- Expérimentales (produites par l'unité ou par un tiers)
 - Les données d'élevage issues d'observations ou d'expérimentations : Données de bien-être animal, Données d'ambiance, Données de performances zootechniques, Mesures de faune, Mesures de couvert végétal, Mesures comportementales, Mesures post-mortem
Type de données : Dataset, Image, Audiovisuel, Sound, Text
 - Les données de laboratoires issues d'analyses : Valeur des aliments, Quantification des rejets animaux, Composition des tissus, Données de Parasitologie, Données de Sérologie, Données de Biologie Moléculaire, Données de Bactériologie, Données de Qualité organoleptique
Type de données : Dataset, Image
 - Les données de traitements statistiques
Type de données : Dataset
- Données de fonctionnement : Quantification des rejets collectés dans le cadre du SME, Suivi des consommations des fluides, Données équarrissage
Type de données : Dataset

❖ *FORMAT DES DONNEES*

- Format papier : Cahier de traçabilité, fiche d'élevage
- Tableur : xls,xlsx, csv

- Texte : .doc, .docx, .pdf, .txt
- Video : mpg, mp4, mkv
- Image : .jpeg, .tif, .jpg, .png
- Son : .wav, .mp3
- Données d'analyses sensorielles : .frp
- Données statistiques : .rtf, mpx, rmd, .py

❖ *PERIMETRE THEMATIQUE DES DONNEES*

- Animal Breeding and Animal Products
- Animal Health and Pathology
- Biodiversity and Ecology
- Climate
- Farming Systems and Practices
- Human Nutrition and food security
- Insects and Entomology
- Microorganisms
- Plant Breeding and Plant Products
- Rural and Agricultural Sociology
- Soils and soil sciences

DROITS DE PROPRIETE INTELLECTUELLE

❖ *QUI DETIENDRA LES DROITS SUR LES DONNEES ET LES AUTRES INFORMATIONS CREEES ?*

L'unité est en accord avec les principes de l'open data et respecte la [charte des Infrastructures Scientifiques Collectives](#) (ISC) d'INRAE.

Les données, hors prestation avec un partenaire privé, sont la propriété d'INRAE. Pour les autres, les droits de propriété sont précisés dans la convention.

SENSIBILITE DES DONNEES

❖ *IDENTIFICATION DU NIVEAU DE SENSIBILITE DES JEUX DE DONNEES*

- Public
- Diffusion limitée
- Confidentiel
- Diffusion restreinte

❖ *QUELLES SONT LES MESURES PRISES ET LES NORMES AUXQUELLES IL EST NECESSAIRE DE SE CONFORMER POUR GARANTIR LA SECURITE DES DONNEES SENSIBLES ?*

Les données soumises au RGPD font l'objet d'une déclaration à la CNIL.

Les participants aux tests organoleptiques signent un consentement éclairé.

Les données de l'unité sont stockées sur des data centers INRAE.

❖ *S'IL Y A DES DONNEES A CARACTERE PERSONNEL, QUELLES SONT LES MESURES ENVISAGEES POUR LES PROTEGER AU COURS DU PROJET OU DANS LE CADRE D'UNE REUTILISATION ?*

Pour protéger les données à caractère personnel d'analyse sensorielle, celles-ci sont anonymisées ou pseudo anonymisées du traitement des données à la diffusion.

PARTAGE DES DONNEES

❖ *Y A T'IL UNE OBLIGATION DE PARTAGE (OU A L'INVERSE UNE INTERDICTION OU UNE RESTRICTION) ?*

L'UE EASM est soumise aux règles institutionnelles (formalisées par la [charte](#) pour le libre accès aux données et aux publications).

Seules les données acquises dans le cadre d'un partenariat privé peuvent nécessiter une restriction particulière.

Selon les données produites, la communication peut être obligatoire, interdite par principe ou sous conditions : <https://datapartage.inrae.fr/Partager-Publier>

❖ *QUELLES SONT LES REUTILISATIONS POTENTIELLES DE CES DONNEES ?*

- Réutilisation pour de nouvelles études ou études complémentaires
- Agrégation avec de nouvelles données
- Publications scientifiques
- Vulgarisation scientifique
- Travaux de Recherche et Développement
- Enseignement

❖ *LA LECTURE DES DONNEES NECESSITE-T-ELLE LE RECOURS A UN LOGICIEL OU UN OUTIL SPECIFIQUE ? SI OUI, LEQUEL ?*

Une fois les données brutes extraites, celle-ci sont accessibles via la plupart des logiciels de traitement de texte (DOC, TXT) ou tableur (XLS, CSV) ou lecteurs multimédia (JPEG, PNG, TIFF, MP4, MP3, WAV).

Les données RFID sont lues avec le logiciel EADS.

❖ *COMMENT LES DONNEES SERONT-ELLES PARTAGEES ?*

Échange de fichiers avec les partenaires des projets par dépôt sur outils collaboratifs INRAE (Sharepoint, Onedrive, Nextcloud) ou non INRAE (Filesender)

Avec qui ?

- Partenaire(s) identifié(s)

❖ *SOUS QUELLE LICENCE ?*

Aucun jeu de données étant déposé, actuellement aucune licence n'a été choisie.

ORGANISATION ET DOCUMENTATION DES DONNEES

❖ *QUELS METHODES ET OUTILS SONT UTILISES POUR ACQUERIR ET TRAITER LES DONNEES, DEPUIS LEUR ACQUISITION JUSQU'A LEUR MISE A DISPOSITION, LEUR ARCHIVAGE OU LEUR DESTRUCTION ?*

Les méthodes et outils de collecte dépendent du type de données.

Les méthodes de collecte de données sont décrites dans les protocoles expérimentaux ou dans les modes opératoires. Les protocoles expérimentaux sont à disposition de tous les agents de l'unité sur le réseau EASM hébergé sur les datacenters INRAE. Les modes opératoires sont stockés via l'application AQ-tools sur le serveur du Cirad et accessibles à tous les agents de l'unité. Les procédures P-DC-11 Gestion Doc Activités, P-DC-20 Gestion Echantillons, P-DC-80 Gestion Cahier Traçabilité encadrent l'expérimentation et la production des données (nommage, classement...).

La collecte des données ainsi que la saisie est assurée par le personnel de l'unité.

La nature et la quantité des données expérimentales collectées sont directement dépendantes du projet expérimental.

Les données expérimentales :

- Les données d'élevage :
 - Données de bien-être animal sont issues d'observations et sont recueillies sur tablette (outil EBENE) ou des fiches papier et reportées sur l'ordinateur (XLS).
 - Données d'élevage sont recueillies soit sur des fiches papier reportées sur l'ordinateur (XLS), soit via des balances connectées dont les fichiers sont exportés sous format XLS. Les données de températures et d'hygrométrie recueillies automatiquement par les sondes sont exportées durant les périodes d'expérimentations sous format XLS. Les pesées automatiques de consommation d'aliment sont exportées sous format XLS.
 - Mesures de faune, sont issues d'observations directes recueillies sur des fiches papier reportées sur l'ordinateur (XLS), ou enregistrées en vidéo au format .mp4.
 - Mesures de couvert végétal sont obtenues soit par traitement d'image via le logiciel Image J, soit à partir d'un herbomètre, soit par des relevés sur papier reportés ensuite sur tableur.
 - Mesures comportementales sont recueillies sur des fiches papier reportées sur l'ordinateur (XLS), ou par des vidéos au format mp4, ou grâce à des balises RFID actives produisant des données au format CSV.
 - Mesures post-mortem sont recueillies sur des fiches papier reportées sur l'ordinateur (XLS).
 - Données météorologiques au format XLS, données d'ambiances (température, hygrométrie) exportées au format XLS.
- Les données de laboratoire :
 - Les données issues d'analyses sont renseignées sur les cahiers de laboratoire avant d'être retranscrites dans des formulaires standardisés (DOC) ou des tableaux (XLS).

- Les données issues d'appareils spécifiques sont collectées avec les logiciels dédiés et référencés dans le tableau ci-dessous.
- Données brutes des tests organoleptiques sont recueillies sur questionnaire papier puis scanné grâce au logiciel fizz (FRP). Les données sont ensuite exportées sous excel (XLS).

Les données de fonctionnement :

Ces données sont collectées sur papier ou via le réseau centre INRAE Nouvelle Aquitaine-Poitiers, puis reportées sur tableur (XLS).

Tableau récapitulatif des logiciels informatiques utilisés :

Activité	Logiciel	Format
Analyse sensorielle	Fizz papier	FRP
Station Météo	ZL6	XLS
Calibration des pipettes	Proline soft	
Balances connectées du laboratoire	LabX direct balance	XLS
Balances connectées de l'élevage	AnimalScale	CSV
Pesons (aliment)	Catman	XLS
Acquisition images de gels d'agarose	UVdek	TIF
SPIR	Opus	001
CPG	GC ChemStation	DAT
Lecteur de plaques	AllSheng	XLS
Mesure de température et hygrométrie	ThermoClient et ThermoServer	
Mesure BEA	ebene	
Mesure comportement par RFID	EADS v2.0.2 (eads gestion lecteurs rfid) PCW v1.0.0.25 (creation de dossiers des tags par jours) ELA Read Manager v2.0.1 (programmation tags ancien modeles) ELA Read Write v4.5.8 (programmation tags rfid) ELA Terminal v3.0.8(beta) (programmation lecteurs rfid)	
Vidéos microscopes	Motic	
Etiquetage	Brady	
Nanodrop	IMPLEN nanophotometer	

Traitement et curation :

Les données d'élevage et de laboratoire saisies sous Excel font l'objet de contrôles par les opérateurs.

Les données de Qualité organoleptique sont contrôlées par les opérateurs avant d'être exportées vers Excel (XLS).

Tableau récapitulatif des logiciels informatiques utilisés :

Activité	Logiciel	Format
Analyse sensorielle traitement	Fizz traitement	RTF
Traitement statistique	Minitab	MPX
Traitement statistique	R	RMD

Traitement de données	Python	PY
Traitement d'images	Image J	
Analyse gels agarose	Quantity One	

Mise à disposition :

Actuellement les données sont disponibles aux acteurs des projets concernés et aux agents de l'unité.

Stockage et archivage :

- Les versions papier (fiches pesées, consommation d'aliments, baguage, comportement, de prélèvements, de mesures ou d'analyses, fiches résultats...) des laboratoires et des élevages sont regroupées par année et par protocole ou poste d'activité.

Chaque équipe archive ses enregistrements à proximité des utilisateurs dans ses propres locaux pour permettre une facilité d'accès. Les contenants recevant les archives sont identifiés.

- Les versions numériques sont stockées sur les datacenters INRAE selon les règles décrites dans la procédure P-DC- 11 Gestion Doc Activités, les sauvegardes sont gérées selon les procédures des datacenters.

Elimination :

Les données archivées au format papier et les versions numériques sont sauvegardées sans limite de temps.

Les coordonnées personnelles soumises à déclaration auprès de la CNIL sont détruites au bout de 10 ans après l'arrêt de participation au jury d'analyse sensorielle.

❖ *QUELLES METADONNEES SERONT UTILISEES POUR ACCOMPAGNER LE JEU DE DONNEES ? QUELS SERONT LES STANDARDS, VOCABULAIRES, TAXONOMIES... UTILISES POUR DECRIRE ET REPRESENTER LES DONNEES ET ELEMENTS DE METADONNEES ? COMMENT LES METADONNEES SERONT-ELLES PRODUITES ET MISES A JOUR ?*

Le référencement des métadonnées est en cours pour préparer le dépôt des données et des métadonnées associées.

❖ *UNE DOCUMENTATION COMPLEMENTAIRE AUX METADONNEES EST-ELLE NECESSAIRE POUR DECRIRE LES DONNEES ET ASSURER LEUR REUTILISABILITE SUR LE LONG TERME ?*

L'ensemble des modes opératoires et procédures régissant les activités de l'unité est disponible dans la base qualité AQTools afin de décrire et de préciser les modes d'obtention des données.

❖ *COMMENT LES FICHIERS DE DONNEES SONT-ILS GERES ET ORGANISES : CONTROLE DES VERSIONS, CONVENTIONS DE NOMMAGE DES FICHIERS, ORGANISATION DES FICHIERS*

Les données expérimentales sont classées dans des dossiers informatiques sous le réseau puis stockées sur un datacenter. Les règles de classement utilisées sont différentes en fonction de l'équipe :

- Elevage : Les données sont classées par année puis par nom de protocole,
- Laboratoire d'analyse sensorielle : Les données sont classées par espèces puis année et nom de protocole,
- Laboratoire de biologie : Les données sont classées par type d'activité puis année et nom de protocole,

- Laboratoire de chimie : Les données sont classées par année puis par nom de protocole.
La mise en place d'un serveur local pour un stockage temporaire et en secours est en cours.

❖ *QUEL EST LE PROCESSUS DE CONTROLE QUALITE DES DONNEES ?*

Nous réalisons des actions à différents niveaux pour fiabiliser les données de l'unité :

- Rédaction des modes opératoires,
- Validation des méthodes,
- Contrôle métrologique des outils de mesure (balance, pipettes, thermomètres),
- Participation au circuit Bipéa (<https://www.bipea.org/fr/>) et circuit interlabos (centre Nouvelle Aquitaine-Poitiers),
- Ajout de témoins internes de contrôle pour les analyses de laboratoires.

STOCKAGE ET SECURITE DES DONNEES

❖ *LES SYSTEMES D'INFORMATION DE LA STRUCTURE ONT-ILS FAIT L'OBJET D'UNE ANALYSE DE RISQUES OU D'UNE HOMOLOGATION ?*

Non

❖ *QUELS TYPES DE SUPPORTS PHYSIQUES SONT UTILISES POUR STOCKER LES DONNEES ?*

- Terminal mobile de saisie (tablette...)
- Capteur (sonde t°...)
- Ordinateur (portable, fixe)
- Papier
- Espace de stockage sur serveur : données température
- Un datacenter,
- Une offre de service type cloud : cas données météo

Les données sont stockées au final sur le réseau INRAE dans un datacenter et sur un serveur local configuré par le CATI SICPA.

❖ *QUELLES SONT LES MESURES DE SECURITE MISES EN PLACE LORS DES ETAPES DE TRANSFERT DES DONNEES ?*

Pour les transferts de données d'un appareil vers les postes de travail personnel : utilisation de clés USB dédiées (non cryptées).

Pour le transfert de données en réseau (mail INRAE, Onedrive, filesender, réseau intranet en interne) : aucune mesure de sécurité.

❖ *QUELLE EST LA VOLUMETRIE ACTUELLE ET PREVISIONNELLE ?*

Nous disposons actuellement d'un espace de stockage de 2.84 terra sur le datacenter INRAE dont 2.45 sont utilisés, et prochainement sur le serveur local de 20 terra.

❖ *L'ENTITE HEBERGEANT PHYSIQUEMENT LES DONNEES A-T-ELLE UNE POLITIQUE DE SECURITE DE L'INFORMATION ET A-T-ELLE UN PLAN D'ASSURANCE SECURITE ?*

Pour l'hébergement des données numériques, c'est la PSSI INRAE qui s'applique (ns2023-31).

Nous travaillons directement et enregistrons sur les serveurs (data center CEA - Commissariat à l'Energie Atomique et aux Energies Alternatives - et INRAE), qui garantissent des sauvegardes régulières sur 2 sites en cas d'incident technique.

La sécurité des données est aussi incluse dans la charte des infrastructures de recherche INRAE.

Pour les données physiques : l'accès aux données papier n'est accessible qu'avec la possession d'un badge personnel.

❖ *SECURITE - CONFIDENTIALITE : LES DONNEES FONT-ELLES L'OBJET D'ECHANGE OU DE PARTAGE AVEC DE TIERS ACTEURS ET SELON QUELLES MODALITES ? COMMENT SONT DETERMINES LES DROITS D'ACCES AUX DONNEES AVANT LEUR PUBLICATION ?*

Dans l'unité, les agents échangent les données via le datacenter nécessitant une connexion personnelle (login et mot de passe).

Hors unité, le partage des données est défini dans le contrat de recherche.

❖ *SECURITE - INTEGRITE - TRACABILITE : QUELLES SONT LES MESURES DE PROTECTION MISES EN ŒUVRE POUR SUIVRE LA PRODUCTION ET L'ANALYSE DES DONNEES ?*

- Pour suivre la production et l'analyse des données, l'unité possède :
- Des cahiers de laboratoire,
- Des protocoles de recherche,
- Des registres de consignation (ex cahier Bien Être Animal),
- Une authentification personnelle sur les outils (ex balance)

❖ *LES AGENTS DE LA STRUCTURE ONT-ILS BENEFICIE D'UNE SENSIBILISATION AUX BONNES PRATIQUES D'HYGIENE NUMERIQUE ?*

- Non

Ce point sera abordé en conseil de service.

ARCHIVAGE ET CONSERVATION DES DONNEES

❖ *QUELLES SONT LES DONNEES A CONSERVER SUR LE MOYEN OU LE LONG TERME ET QUELLES SONT LES DONNEES A DETRUIRE ?*

L'ensemble des données (brutes et traitées) est conservé sur le long terme avec les métadonnées nécessaires permettant leur réutilisation.

Les données vidéos sont détruites selon le temps défini par le protocole.

❖ *SUR QUELLE PLATEFORME D'ARCHIVAGE PERENNE SERONT ARCHIVEES LES DONNEES A CONSERVER SUR LE LONG TERME ? SINON, QUELLES PROCEDURES SERONT MISES EN PLACE POUR LA CONSERVATION A LONG TERME ?*

Actuellement, les données sont stockées sur le data center INRAE.

❖ *QUELLE EST LA DUREE DE CONSERVATION DES DONNEES ?*

Actuellement les données sont conservées sans limite de durée sauf les données vidéos qui sont détruites selon le temps défini par le protocole.

❖ *QUELLES GARANTIES DE FINANCEMENTS COUVRIRONT LES COUTS ASSOCIES A LA CONSERVATION A LONG TERME ?*

Les coûts associés au stockage sur le datacenter seront financés sur les fonds propres de l'unité.