



HAL
open science

Plan de Gestion de Données de la plateforme MassLor

Frédéric Aubriet, François Dupire, Jasmine Hertzog, Lionel Vernex-Loset

► **To cite this version:**

Frédéric Aubriet, François Dupire, Jasmine Hertzog, Lionel Vernex-Loset. Plan de Gestion de Données de la plateforme MassLor. Université de Lorraine; CNRS. 2025. hal-04895780

HAL Id: hal-04895780

<https://hal.univ-lorraine.fr/hal-04895780v1>

Submitted on 18 Jan 2025

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

DMP de la structure MassLor

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "AgroParisTech - modèle de PGD "Structure"" fourni par AgroParisTech - Institut des sciences et industries du vivant et de l'environnement.

Renseignements sur le plan

Titre du plan	DMP de la structure MassLor				
Version	Version initiale				
Domaines de recherche (selon classification de l'OCDE)	Chemical sciences				
Langue	fra				
Date de création	2023-06-14				
Date de dernière modification	2025-01-15				
Licence	<table><tr><td>Nom</td><td>Creative Commons Attribution 4.0 International</td></tr><tr><td>URL</td><td>http://spdx.org/licenses/CC-BY-4.0.json</td></tr></table>	Nom	Creative Commons Attribution 4.0 International	URL	http://spdx.org/licenses/CC-BY-4.0.json
Nom	Creative Commons Attribution 4.0 International				
URL	http://spdx.org/licenses/CC-BY-4.0.json				

Renseignements sur le projet

Titre du projet DMP de la structure MassLor

Acronyme MassLor

Résumé MassLor est une Infrastructure de Recherche en Spectrométrie de Masse de l'Université de Lorraine et du Pôle Scientifique Chimie et Physique Moléculaires (CPM). Elle a pour objectif de proposer son expertise et son savoir-faire à des partenaires académiques et industriels, lorrains et nationaux, dans le domaine de la Recherche & Développement, des prestations de service et de la formation.

Elle regroupe des moyens humains, des équipements et des compétences en spectrométrie de masse présents sur un site Messin et sur un site Nancéien de l'Université de Lorraine.

Le site MassLor-Metz est de plus intégré au **Réseau National Infranalytics** dédié à la caractérisation et à l'analyse chimique en Spectrométrie de Masse FT-ICR à très haut champ (FT-ICR - FR2054). **Les chercheurs peuvent ainsi avoir un accès privilégié à la spectrométrie de masse à très haute résolution.**

Les domaines d'applications de nos techniques sont nombreux et concernent par exemple la caractérisation de milieux complexes (solide ou liquide) comme les produits pétroliers, les polymères, les particules, les extraits végétaux, les matériaux, les nanomatériaux ou encore les produits de synthèse.

Partenaires

- Centre national de la recherche scientifique <https://ror.org/02feahw73>
- Université de Lorraine <https://ror.org/04vfs2w97>

Produits de recherche :

1. Données issues de nos "processus recherche" dans le cadre de projets identifiés (Jeu de données)
2. Rapports d'analyses issus de nos "processus prestation" dans le cadre de service d'appui à la recherche (Texte)

Contributeurs

Nom	Affiliation	Rôles
Aubriet Frederic - 0000-0003-2457-7974	Université de Lorraine	
Dupire Francois - 0000-0001-5038-723X	Université de Lorraine	<ul style="list-style-type: none">• Coordinateur de projet• Personne contact pour les données (Recherche, Prestation)• Responsable du plan
Hertzog Jasmine - 0000-0002-7946-1904	Université de Lorraine	
Vernex-Loset Lionel	Université de Lorraine	

DMP de la structure MassLor

1. INFORMATIONS SUR LA STRUCTURE

Acronyme

MassLor

Nom

MassLor

Nature

MassLor est une Infrastructure de Recherche en Spectrométrie de Masse de l'Université de Lorraine et du Pôle Scientifique Chimie et Physique Moléculaires (CPM).

Elle regroupe des moyens humains, des équipements et des compétences en spectrométrie de masse présents sur un site Messin et sur un site Nancéien de l'UL.

Le site MassLor-Metz est de plus intégré au Réseau National Infranalytics dédié à la caractérisation et à l'analyse chimique en Spectrométrie de Masse FT-ICR à très haut champ (FT-ICR - FR2054).

Résumé/Objet(s)

Elle a pour objectif de proposer son expertise et son savoir-faire en spectrométrie de masse et techniques couplées à des partenaires académiques et industriels, lorrains et nationaux, dans le domaine de la Recherche & Développement, des prestations de service et de la formation.

Champ(s) thématique(s)

Le champ instrumental est la spectrométrie de masse et les techniques couplées, telles que la chromatographie (liquide et gazeuse) et l'électrophorèse capillaire.

MassLor propose l'accès à la spectrométrie de masse FT-ICR à très haute résolution et à haut champs, équipée de nombreuses sources d'ionisation (ESI, APPI, APCI, DIP, DART, (MA)LDI).

Les analyses réalisées couvrent les domaines de la caractérisation de composés issus de la synthèse organique aux matrices organiques complexes comme les biocarburants, les extraits de plantes, la matière cosmique, les polymères naturels et de synthèse, et les échantillons environnementaux.

Mots-clés

Spectrométrie de masse

Chromatographie gazeuse
Chromatographie liquide
Electrophorèse capillaire
FTICR MS
Imagerie
Métabolomique
Pétroléomique
Biocarburant
Matrices organiques complexes
Sources d'ionisation

Responsable(s) de la structure

Prof. Frédéric Aubriet, Responsable scientifique
Dr Lionel Vernex-Lozet, Responsable technique

Responsable(s) de la gestion et de l'ouverture des données

François Dupire, Responsable qualité, ambassadeur de la Donnée

Partenaires impliqués

Unités : L2CM UMR 7053 (site de Nancy) et LCP-A2MC EA 4632 (site de Metz)
Tutelles : CNRS et Université de Lorraine
Fédération de recherche : Infranalytics FR 2054

2. INFORMATIONS SUR LES DONNEES - Origine(s) des données

Données produites par la structure

Les données produites par MassLor sont issues d'expériences ou d'analyses réalisées sur l'un de nos instruments. Ces dernières peuvent être :

- des spectres de masses
 - des données de mobilité ionique
 - des données issues de techniques séparatives (chromatographie, électrophorèse capillaire, ...) couplée à de la détection MS
 - de l'imagerie par spectrométrie de masse
 - des données MS après traitement informatisé
-

Données utilisées par la structure et provenant de sources extérieures

Il existe plusieurs types de données provenant de sources extérieures :

- des données déjà publiées par les pairs dans la littérature scientifique
- des données propres à chaque partenaire, acquises en amont du projet
- des métadonnées concernant les échantillons apportés par le partenaire / le client
- des données d'expériences issues d'autres méthodes de caractérisations issues d'autres plateformes (spectroscopie, RMN,)

Objectif(s)/Utilisation(s) de chaque jeu de données dans la structure

Les objectifs et l'utilisation dépendent du projet de recherche ou de la question du client dans le cadre d'une prestation.

- Dans le cadre de la recherche se référer au Plan de Gestion de Données (PGD) du projet
- Dans le cadre de la prestation se référer à la demande du client, tracée dans eLabFTW (cahier de laboratoire électronique)

Protocole(s) de création des jeux de données

Les procédures de préparation d'échantillon (mode opératoire) et d'acquisition des données brutes (paramètres opératoires, nom de la méthode, ...) seront conservés dans le cahier de laboratoire électronique (eLabFTW) de MassLor ou dans celui du client dans le cadre de prestation interne à l'Université de Lorraine.

Certains paramètres d'acquisition sont conservés dans le fichier échantillon (raw data), dépendant du constructeur. Ils seront stockés et sauvegardés selon ce qui est précisé respectivement dans la partie 5 et 6 du présent PGD.

Les paramètres de traitement des données sont conservés avec la même attention dans le cahier de laboratoire électronique eLabFTW et pour certaines d'entre elles dans les fichiers de process.

Outils spécifiques pour lire les données

Les raw data ne sont lisibles que par les logiciels propriétaires des constructeurs.

Les données traitées peuvent être de natures diverses

- soit rester au format constructeur, dans ce cas, elles ne resteront lisibles que par le logiciel correspondant
- soit exporter sous un format libre pour ensuite être travaillées dans des logiciels non-propriétaires (MZmine2, Origine,)
- soit sous forme de rapport dans un format pdf

3. INFORMATIONS SUR LES DONNEES - Qualité et documentation des données

En cas de besoin, consulter la fiche pratique sur les [bonnes pratiques de gestion des données](#) et celle sur les [métadonnées](#)

Données issues de nos "processus recherche" dans le cadre de projets identifiés

Responsable(s)

Dans le cadre d'un projet de recherche, le responsable des données est le chercheur pilotant le projet au sein de MassLor.

Référentiel(s) de données (taxonomiques, géographiques, etc...)

Les métadonnées dépendent du projet. Cette partie devra donc être précisée dans le Plan de Gestion de Données du dit projet.

Méthode(s) de contrôle de la qualité des données

Le contrôle de la qualité des données se décline en deux points :

- l'étalonnage & la surveillance des instruments sont précisés dans les manuels d'utilisation et protocole.
- une traçabilité de la méthode utilisée en se basant sur la convention de nommage ci dessous :

La convention de nommage des méthodes est celle recommandée par DoRaNum (https://doranum.fr/stockage-archivage/comment-nommer-fichiers_10_13143_wgqw-aa59/).

Dans le cadre d'un projet de recherche une méthode à vocation à évoluer rapidement. De plus MassLor est amené à réaliser différents tests en modifiant les méthodes sans pour autant les conserver tous.

Concernant les tests, les règles d'arborescence et de nommage des fichiers devront tracer les paramètres étudiés lors de mise au point.

Concernant le nommage des méthodes, la procédure mise en place pour la "prestation" est un bon guide. Un exemple de nom de méthode peut être : « *FD-synthese-LC-ESI-qTOF-pos-automsms-100-1000-v200708-1* ». Ceci indique que c'est François Dupire qui a mis au point une méthode destinée aux analyses issues de la synthèse organique en LC/ESI/qTOF en mode positif. La méthode fragmente en mode autoMSMS dans la gamme de masse 100-1000. Enfin c'est la première version de la méthode qui a été mise au point le 8 juillet 2020.

Une modification de la méthode induisant une modification du signal (en masse ou en intensité) induit un changement de version de la méthode.

Standard(s) de métadonnées

La convention de nommage des fichiers d'analyses est celle recommandée par DoRaNum.

Pour une meilleure connaissance du contenu d'un fichier de données d'analyses à travers le nom, nous la nommons en concaténant les informations contenues dans le Tableau 3. Ceci vaut autant pour les rawdata que d'éventuels fichiers issus de processus de données

Un exemple de nom de fichier peut être : « *LC-IM-1025-EM256* ». Ceci indique que l'expérience étudie le couplage LC avec la mobilité ionique sur un échantillon enregistré sur ElabFTW le 25 octobre dont le nom de l'expérience est EM256. C'est la première version de l'analyse sinon nous ajouterons Rx pour xième répétition.

Cette nomenclature peut être adaptée en fonction des instruments

Note01 : les dates sont sous la forme mmjj puisque les données sont rangées par années. C'est donc bien l'arborescence qui indique l'année, il n'est donc pas utile de reprendre cette information dans les noms de fichiers.

Note02 : Il est interdit de mettre des espaces entre les mots ni de caractères spéciaux (accents et cédille y compris). Pour séparer les mots, nous utilisons un tiret "-" ou un underscore "_".

information(s) sur le mode d'analyse :	Selon les besoins préciser le mode de couplage (LC ; GC ; DI ; EC), la source (ESI ; APCI ; ASAP), la polarité (pos ; neg) ou le mode de scan (DDA ; DIA ; MSe ; SCAN ; MRM) Il n'y a pas ici de règles bien strictes. Cela dépend principalement de la méthode utilisée et du paramètre que l'expérimentateur désire mettre en valeur au regard des objectifs décrits dans l'expérience sur Elab.
date d'enregistrement :	Date à laquelle l'expérience est enregistrée dans le cahier de laboratoire électronique ElabFTW sous la forme de mois/jour (mmjj)
Nom de l'expérience :	Nom de l'expérience indiquée par le client et enregistré dans ElabFTW. Il est conseillé d'utiliser des tailles de noms relativement restreintes, Si besoin il peut être utile de les tronquer si cela n'apporte pas de confusion. Conformément à la Note02, ne pas utiliser de caractères spéciaux mais uniquement des tirets.
Répétition :	Dans le cas d'une seule répétition de l'expérience, rien est indiqué. Si l'expérience doit être répétée, nous ajoutons -Rx à la suite du nom pour indiquer que nous utilisons la xième répétition de l'expérience.

Tableau 2: Détermination du nom d'un fichier de données d'analyses
les fichiers de données brut ou retraités sont ainsi nommés pour faciliter leur identification.

Thésaurus/ontologie

L'utilisation de TAG et mots clés (particulièrement dans ElabFTW) est une bonne manière de simplifier les recherches futures de nos données. Loterre rubrique chimie (<https://skosmos.loterre.fr/fr/>) est un bon guide pour déterminer les mots clés utiles à notre activité.

Dans tous les cas l'anglais est doit être utilisé.

Localisation

L'arborescence de travail dépend du fonctionnement du logiciel constructeur et doit donc être adapté tout en respectant la ligne directrice décrite ci dessous.

Dans le cadre d'une activité de recherche, il y a création d'un répertoire par projet, puis un sous-répertoire par expérimentateur, puis par objectif de l'expérience puis enfin par date. ceci s'adapte aux logiciels constructeurs.

Un exemple d'arborescence peut être : "REALYTIC/NT/opt-DADIP-APCIneg/240909" ceci indique que ces données concerne l'ANR realytic, qu'il ont été généré par Nathan dont l'objectif était l'optimisation de la méthode DADIP par APCI en mode neg. Cette série d'analyse date du 9 septembre 2024.

Note01 : les dates sont sous la forme année, mois, jour (aammjj)

Autres éléments de documentation des données

Question sans réponse.

Rapports d'analyses issus de nos "processus prestation" dans le cadre de service d'appui à la recherche

Responsable(s)

Dans le cadre de la prestation, le responsable des données est le personnel réalisant l'analyse.

Référentiel(s) de données (taxonomiques, géographiques, etc...)

Les métadonnées liées à l'échantillon analysé sont retranscrites dans le cahier de laboratoire électronique ElabFTW. Ils sont de deux catégories :

- les métadonnées fournies par le client sous sa responsabilité
 - les métadonnées liées à l'analyse sous la responsabilité de MassLor.
 - Nom de la méthode utilisée
 - Nom du fichier résultats utilisé pour le rapport d'analyse
 - Toute information pertinente (calibration par exemple) qui ne serait pas indiquée dans la méthode.
 - Toute adaptation par rapport à la méthode.
-

Méthode(s) de contrôle de la qualité des données

Le contrôle de la qualité des données se décline en deux points :

- l'étalonnage & la surveillance des instruments sont précisés dans les manuels d'utilisation et protocole.
- une traçabilité de la méthode utilisée en se basant sur la convention de nommage ci dessous :

La convention de nommage des méthodes est celle recommandée par DoRaNum (https://dorandum.fr/stockage-archivage/comment-nommer-fichiers_10_13143_wgqw-aa59/).

Pour une meilleur connaissance du contenu d'une méthode à travers le nom, nous la nommons en concaténant les informations contenues dans le Tableau 1.

Un exemple de nom de méthode peut être : « *FD-synthese-LC-ESI-qTOF-pos-automsms-100-1000-v200708-1* ». Ceci indique que c'est François Dupire qui a mis au point une méthode destinée aux analyses issues de la synthèse organique en LC/ESI/qTOF en mode positif. La méthode fragmente en mode autoMSMS dans la gamme de masse 100-1000. Enfin c'est la première version de la méthode qui a été mise au point le 8 juillet 2020.

Une modification de la méthode induisant une modification du signal (en masse ou en intensité) induit un changement de version de la méthode.

Note01 : les dates sont sous la forme années/mois/jours (aammjj)

Note02 : Il est interdit de mettre des espaces entre les mots ni de caractères spéciaux (accents et cédille y compris). Pour séparer les mots, nous utilisons un tiret "-" ou un underscore "_".

Initiales de la personne :	Ayant mise au point la méthode. Il doit s'agir de personnel permanent ou non de MassLor (incluant thésard ou postdoc). En cas de stagiaire utiliser les initiales de son encadrant. Ex : FD,
Acronyme du projet :	Projet de recherche ou de la prestation de service. Eventuellement la famille de molécule analysée ou des mots clés. L'acronyme doit être concis et aisément compris ex : Bioval
Technique séparative :	Indiquer ici la présence ou pas d'une méthode séparative préalable, la liste est donc exhaustive : Ø DI pour introduction directe (ou du moins pour absence de séparation on-line) Ø LC pour la chromatographie liquide Ø GC pour la chromatographie phase gaz Ø EC pour l'électrophorèse capillaire Ø TLC pour chromatographie sur couche mince
Type de source :	Indiquer juste la source d'ionisation Ex : MALDI ; ESI ; ...
Analyseur :	Indiquer le type d'analyseur de l'instrument sur lequel la méthode a été mise au point. Ex : qTOF ; FT ; TQ ; IT ; ...
Polarité :	Indiquer : pos ; neg ou posneg
Type de balayage :	Indiquer le type de balayage utilisé lors de la méthode, incluant la méthode MSMS s'il y a lieu. Ex scan, MRM, autoMSMS,
Gamme de masse :	Indiquer la gamme de masse (modes scans) ou rien dans les cas évident comme les modes SIM ou MRM
Version :	Comme tous les documents de MassLor la version se note de la manière suivante : VAAMMJJ-x. x étant une incrémentation du numéro de version Ex : v201124-1 pour une méthode mise au point le 24 novembre 2020 Ex : v201215-3 pour la deuxième mise à jour de la méthode en date du 15 décembre 2020.

Tableau 1: Détermination du nom d'une méthode

La méthode sera ainsi nommée dans l'ordinateur pilotant l'instrument. Le nom de la méthode étant généralement sauvegardé comme métadonnée avec les spectres. Ceci permet donc de tracer d'une part la méthode utilisée et d'autre part la version de la méthode utilisée.

Standard(s) de métadonnées

La convention de nommage des fichiers d'analyses est celle recommandée par DoRaNum (https://doranum.fr/stockage-archivage/comment-nommer-fichiers_10_13143_wgqw-aa59/).

Pour une meilleure connaissance du contenu d'un fichier de données d'analyses à travers le nom, nous la nommons en concaténant les informations contenues dans le Tableau 2. Ceci vaut autant pour les rawdata que d'éventuels fichiers issus de processus de données

Un exemple de nom de fichier peut être : « 231026-LCESIMSe-1025-EM256 ». Ceci indique que l'analyse a été effectuée le 26 octobre 2023 en couplage LC/ESI en mode MSe sur un échantillon enregistré sur ElabFTW le 25 octobre dont le nom est EM256. C'est la première version de l'analyse sinon nous ajouterons Rx pour xième répétition

Note01 : les dates sont sous la forme mmjj puisque dans le cas de prestation, les données sont rangées par années. C'est donc bien l'arborescence qui indique l'année, il n'est donc pas utile de reprendre cette information dans les noms de fichiers.

Note02 : Il est interdit de mettre des espaces entre les mots ni de caractères spéciaux (accents et cédille y compris). Pour séparer les mots, nous utilisons un tiret "-" ou un underscore "_".

Date de réalisation de l'analyse : sous la forme année/mois/jour (aammjj)

information(s) sur le mode d'analyse : Selon les besoins préciser le mode de couplage (LC ; GC ; DI ; EC), la source (ESI ; APCI ; ASAP), la polarité (pos ; neg) ou le mode de scan (DDA ; DIA ; MSe ; SCAN ; MRM)
Il n'y a pas ici de règles bien strictes. Cela dépend principalement de la méthode utilisée (car beaucoup d'informations sont contenues dans la méthode) et des éventuelles divergences par rapport à la méthode (polarité, source, mode de scan)

date d'enregistrement : Date à laquelle l'échantillon est enregistré dans le cahier de laboratoire électronique ElabFTW sous la forme de mois/jour (mmjj)

Nom de l'échantillon : Nom de l'échantillon indiqué par le client et enregistré dans ElabFTW.
Il est conseillé d'utiliser des tailles de noms relativement restreintes, Si besoin il peut être utile de les tronquer si cela n'apporte pas de confusion.
Conformément à la Note02, ne pas utiliser de caractères spéciaux mais uniquement des tirets

Répétition : Dans le cas d'une première version de l'analyse, rien n'est indiqué.
Si pour une raison quelconque nous devons répéter l'analyse (problèmes techniques ou problème de concentration, changement de polarité, de mode de scan...) nous ajoutons -R2 à la suite du nom pour indiquer que nous donnons les résultats à partir d'une la répétition N°2

Tableau 2: Détermination du nom d'un fichier de données d'analyses

La méthode sera ainsi nommée dans l'ordinateur pilotant l'instrument. Le nom de la méthode étant généralement sauvegardé comme métadonnée avec les spectres. Ceci permet donc de tracer d'une part la méthode utilisée et d'autre part la version

Thésaurus/ontologie

L'utilisation de TAG et mots clés (particulièrement dans ElabFTW) est une bonne manière de simplifier les recherches futures de nos données. Loterre rubrique chimie (<https://skosmos.loterre.fr/fr/>) est un bon guide pour déterminer les mots clés utiles à notre activité.

Dans tous les cas l'anglais est doit être utilisé.

En s'inspirant des paramètres de la méthode il convient d'indiquer (en fonction de la pertinence) dans les TAG :

- méthode d'introduction : LC, GC, DI, CE,
- méthode d'ionisation : ESI, APCI, ASAP, DIP, APPI, EI, CI,
- méthode de scan : MS, MSMS, DDA, DIA,
- analyseur, SQ, TQ, qToF, ...
- nom de l'instrument : Xevo,

Localisation

L'arborescence de travail dépend du fonctionnement du logiciel constructeur et doit donc être adapté tout en respectant la ligne directrice décrite ci-dessous.

Dans le cadre d'une prestation, il y a création d'un répertoire par année, puis un sous-répertoire par type d'échantillon, par

méthode ou par utilisateurs, selon les cas.

Un exemple de nom d'arborescence peut être : « 2024/synthese ». Ceci indique que les échantillons datent de 2024 et correspondent à une méthode pour les composés issus de la synthèse organique par exemple.

Un autre exemple de nom d'arborescence peut être : « 2024/Sabrina ». Ceci indique que les échantillons datent de 2024 et correspondent aux échantillons de Sabrina Touchet. Cette arborescence est plutôt adaptée dans le cas de la mise à disposition. Dans le cadre d'une analyse quantitative, un sous répertoire supplémentaire par série est créé. Par exemple « 2024/alcaloide/241003/ ». Ceci indique que les échantillons datent de 2024, correspondent à la quantification des alcaloïdes et proviennent de la série du 03 octobre 2023.

Autres éléments de documentation des données

Question sans réponse.

4. MESURE LEGALES, REGLEMENTAIRES ET ETHIQUES

AgroParistech met à disposition une fiche pratique sur le [cadre juridique](#) des données de recherche

Responsable(s) du respect du cadre légal

La responsabilité du respect du cadre légal est partagée entre le responsable scientifique (Frédéric Aubriet) et le responsable technique (Lionel Vernex-Loiset) de MassLor.

Le responsable qualité (François Dupire), en cohérence avec sa fiche mission, a un rôle de veille réglementaire, de conseils, d'accompagnement et de mise en oeuvre du présent Plan de Gestion de Données.

Titularité des droits sur les données, propriété intellectuelle

Dans le cadre de prestations nous distinguons

- Les données scientifiques obtenues pour les projets en mode « **simple mise à disposition** » sont la propriété exclusive du laboratoire de recherche et/ou de l'utilisateur qui mène l'expérimentation. Toute utilisation de ces données, au titre de la plateforme, ne pourra se faire qu'après accord explicite de l'utilisateur ou de son responsable.
- Pour les projets menés en mode « **prestation de service** », la plateforme se réserve le droit d'exploiter les données techniques dans le but d'amélioration continue de ses services et de la qualité de ses prestations.

Dans le cadre de projet de recherche, la titularité et les droits de propriété intellectuelle sur les résultats et les données sont régies par un accord de gré à gré entre les différents partenaires dans le cadre d'un projet de recherche et décrit dans le Plan de Gestion des Données dudit projet. Il concerne :

- **Les connaissances propres** des partenaires qui restent leurs propriétés respectives. Dans le cadre général, un partenaire ne reçoit aucun droit sur les connaissances propres des autres partenaires du fait de son implication dans un projet.
- **Les résultats propres** qui sont la propriété du partenaire qui les a générés seul sans la contribution des autres partenaires. Dans le cadre général, les éventuels brevets nouveaux et les autres titres de propriété intellectuelle sur lesdits résultats propres seront déposés à ses seuls frais, à son seul nom et à sa seule initiative.
- **Les résultats communs** qui appartiennent à l'ensemble des partenaires à proportions de leurs contributions. Dans le cadre général, les partenaires seront responsables de la répartition entre eux de la quote-part de copropriété qui leur revient conformément à la convention et/ou le PGD qui les lie et désigneront parmi eux un mandataire unique en application du décret N°2020-21 du 13 janvier 2020 relatif au mode de désignation et aux missions du mandataire prévu à l'article L.353-1 du code de la recherche.

Dans tous les cas, chaque partie s'engage à ne pas publier, ni divulguer, les informations scientifiques ou techniques appartenant à une autre partie, dont elle pourrait avoir connaissance à l'occasion de l'exécution du projet.

Partenariats internationaux

Question sans réponse.

Cadre éthique

Sans objet

Données personnelles

MassLor ne traite pas de données personnelles.

En charge au partenaire d'anonymiser les échantillons en lien avec des données personnelles.

Données issues de ressources génétiques

Sans objet

Données confidentielles

Sans objet

Données ouvertes

Dans le cadre d'une prestation :

La propriété intellectuelle étant au partenaire / client, **MassLor ne peut ouvrir l'accès aux données dont il n'est pas propriétaire.**

L'apport technologique ou technique de MassLor doit être reconnu et mentionné clairement dans tout document utilisant les résultats obtenus sur la plateforme : **"We thank xx from the mass spectrometry MassLor platform of Lorraine University"**.

Dans le cadre d'un projet de recherche:

Les résultats scientifiques obtenus et les données achevées, acquis au cours du projet, seront partagés avec la communauté scientifique autant que possible via l'espace institutionnel de l'Université de Lorraine dans Recherche Data Gouv <https://recherche.data.gouv.fr> (DOREL). **Le partage ou non des données et des résultats sera discuté entre les partenaires au cas par cas et inscrit dans le PGD du projet**, avec un souci d'un juste équilibre entre la protection et la valorisation du savoir-faire et des connaissances acquis au cours du projet. (aussi ouvert que possible, aussi fermé que nécessaire)

Les résultats générés seront valorisés au cours du projet sous formes d'articles scientifiques dans des revues à communauté de lecture et sous formes de communications écrites ou orales dans des congrès nationaux ou internationaux. **Le libre accès aux articles sera garanti soit par la publication dans des revues en libre accès, soit par le dépôts du post-print dans une archive ouverte et en libre accès tel que HAL** en respectant les 6 mois d'embargo prévus par la loi N° 2016-1321 du 7 octobre 2016 pour une République numérique.

MassLor en tant que plateforme doit être mentionné dans les remerciements lors de présentations, posters ou publications : **"We**

thank the mass spectrometry MassLor platform of Lorraine University".

5. FONCTIONNEMENT ET INFRASTRUCTURES - Stockage

Responsable(s)

La Direction Numérique de l'Université de Lorraine est responsable du maintien des infrastructures et des outils mis à disposition dans le cadre du stockage des données.

Les responsables de site vérifient le bon fonctionnement des outils de stockage des données et alertent la direction Numérique en cas de dysfonctionnement.

Méthode(s), supports

Concernant les données brutes issues de nos instruments : PETA

PETA est un outil permettant de **stocker des données à caractère confidentiel**. Les données sont uniquement accessibles par le logiciel Storage Made Easy (SME), soit par https ou soit via un client lourd. Outre **la sécurité très forte du datacenter d'une part et de l'authentification SSO/CAS d'autre part**, SME gère en interne sa gestion des droits.

Concernant les données de volumes plus faible : B'UL

Les données déposées sur la B'UL sont stockées sur des serveurs de l'Université de Lorraine qui sont gérés par la Direction du Numérique. Les transferts de documents utilisent des canaux de communication chiffrés. Le contenu du document ne peut pas être lu par une tierce personne au cours du transfert. Le système est sauvegardé en totalité toutes les nuits. En cas de problème le service peut être restauré dans l'état dans lequel il se trouvait la nuit précédant l'incident. En cas de suppression accidentelle d'un document, il peut être retrouvé dans la corbeille puis restauré. Les versions précédentes d'un document sont accessibles via l'interface web et il est possible, à tout moment, de revenir sur une version précédente.

L'espace B'UL est sécurisé au moyen de l'identifiant et du mot de passe UL. L'accès peut être ouvert à des personnels extérieurs à l'UL par le biais de la création de comptes invités numériques sécurisés.

Concernant les données expérimentales : ElabFTW

le cahier de laboratoire électronique "ElabFTW" est notre outils pour tracer les informations liées à l'échantillon et l'analyse. Nous y retrouvons l'ensemble des métadonnées afférentes à l'analyses. A l'instar des outils ci dessus, ElabFTW est stocké sur des serveurs universitaires et géré par la direction du numérique de l'UL.

Concernant les données extrêmement lourdes issues du FTICR :

En raison du gros volume de données générées par les analyses FT-ICR MS (plusieurs dizaines de téraoctets) et de leurs utilisations pour du retraitement, l'outil PETA n'est utilisé que pour la sauvegarde des données présent sur le PC d'acquisition du FT-ICR MS.

Afin d'éviter de saturer, l'espace mémoire du PC d'acquisition, les utilisateurs sont amenés à stocker leurs données sur un serveur dédié nommé Archivage. Le serveur est accessible sur le réseau interne de l'établissement aux seuls utilisateurs possédant un compte local.

Le serveur possède un système d'exploitation Ubuntu non virtualisé, avec une architecture RAID 5:1 appliquée aux disques durs, ce qui permet une tolérance de panne sur 2 disques en même temps.

L'entretien et la maintenance du serveur sont réalisés par l'équipe informatique de l'UFR SciFA.

Localisation géographique et institutionnelle

Quels que soient les outils numériques institutionnels (B'UL, PETA, ElabFTW,), ils se trouvent sur des serveurs de l'Université de

Lorraine et sont gérés directement pas la Direction Numérique.

Financement

L'espace de stockage sur PETA est gratuit pour MassLor (coût assuré par l'Université de Lorraine) dans la limite de 30 To par laboratoire. Des incréments sont possibles par tranche de 5 To moyennant une participation financière.

Le financement des autres outils mis à disposition (DOREL, B'UL, ElabFTW,) est assuré par l'Université de Lorraine.

Méthode de maintenance et de sécurisation

La maintenance des outils institutionnels est assurée par la Direction Numérique.

Les outils institutionnels sont sécurisés aux moyens de l'identifiant et de mot de passe UL (authentification SSO/CAS) . L'accès peut être ouvert à des personnels extérieurs à l'Université de Lorraine par le biais de la création de comptes invités numériques sécurisés.

En outre certains outils tel que eLabFTW (cahier de laboratoire électronique) sont accessibles depuis un réseau externe uniquement par un VPN.

6. FONCTIONNEMENT ET INFRASTRUCTURES - Sauvegardes

Responsable(s)

La sauvegarde des serveurs de stockage institutionnels sont exclusivement à la charge de la Direction Numérique.

Concernant le stockage des données FTICR sur le serveur :

L'entretien et la maintenance du serveur "Archivage" des données FT-ICR MS sont réalisés par l'équipe informatique de l'UFR SciFA.

Méthode(s), supports

Question sans réponse.

Localisation géographique et institutionnelle

Question sans réponse.

Financement

Question sans réponse.

Méthode de maintenance et de sécurisation

Question sans réponse.

7. FONCTIONNEMENT ET INFRASTRUCTURES - Partage et accès collaboratifs aux fichiers

Responsable(s)

Dans le cadre de la prestation : le responsable du partage des données est le personnel réalisant l'analyse.

Dans la cadre d'un projet de recherche : le responsable du partage des données est le chercheur pilotant le projet au sein de MassLor.

Infrastructure de travail et d'échange de fichiers

Divers outils sont proposés par l'Université de Lorraine :

- PETA pour le partage des données brutes
- B'UL ou filesender pour le partage de documents
- ElabFTW pour le partage d'un cahier de laboratoire électronique

Plus généralement, l'ensemble des outils du réseau institutionnel Renater.

Sécurité

Les outils institutionnels sont sécurisés au moyen de l'identifiant et de mot de passe UL (authentification SSO/CAS) . L'accès peut être ouvert à des personnels extérieurs à l'Université de Lorraine par le biais de la création de comptes invités numériques sécurisés.

Pour la B'UL et PETA un lien peut être créé par un utilisateur et envoyé à un partenaire avec ou sans mot de passe. Un mot de passe doit toujours être préféré. Ce dernier doit contenir des caractères spéciaux, des majuscules et des minuscules. En outre le mot de passe doit impérativement être envoyé aux partenaires par un autre moyen que le lien (SMS, autre mail).

Pour eLabFTW (cahier de laboratoire électronique), il est accessible depuis un réseau externe uniquement par un VPN.

Nommage des dossiers et des fichiers

Question sans réponse.

Versionnement des fichiers

Question sans réponse.

Les règles diffèrent-elles selon les formats de fichiers ? Si oui, détailler

Question sans réponse.

8. OUVERTURE DES DONNEES

AgroParisTech met à disposition des fiches pratiques sur les :

- [entrepôts de données](#)
- [data papers](#)
- [licences de diffusion](#)

Responsable(s)

Temporalité

Question sans réponse.

Méthodes d'ouverture des données

Question sans réponse.

Méthodes de publicité des données ouvertes

Question sans réponse.

Publication et archivage du PGD

Le présent PGD sera archivé sous HAL, et rendu public avec DMP OPIDor et HAL.

Chaque partenaires, lors de la rédaction d'un PGD recherche, pourra faire référence au PGD MassLor concernant la gestion des analyses de spectrométrie de masse réalisées sur la plateforme.

Réutilisation des données

L'ouverture et la réutilisation des données est sous la responsabilité du chercheur. Il reste donc à sa charge de prévoir dans son propre PGD recherche les conditions d'ouverture des données. MassLor en tant que plateforme n'as pas d'apposition de principe à **une ouverture, sous licence, des données dans un entrepôt institutionnel** (par exemple data.gouv.fr). Il convient évidemment de discuter avec la plateforme tous projets d'ouverture des données.

9. ARCHIVAGE

Responsable(s)

Périmètre

Question sans réponse.

Volume

Question sans réponse.

Infrastructure

Question sans réponse.

Financement

Question sans réponse.