



HAL
open science

Plan de gestion de données du projet MEMENTO

David Benaben, Nicolas Creusot, Joan Artigas, Isabelle Batisson, Chloé
Bonnineau, Soizic Morin

► **To cite this version:**

David Benaben, Nicolas Creusot, Joan Artigas, Isabelle Batisson, Chloé Bonnineau, et al.. Plan de gestion de données du projet MEMENTO. INRAE; UCA. 2024. hal-04707847

HAL Id: hal-04707847

<https://hal.inrae.fr/hal-04707847v1>

Submitted on 24 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

DMP du projet MEMENTO

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe : modèle structuré" fourni par Science Europe.

Renseignements sur le plan

Titre du plan	DMP du projet MEMENTO				
Livrable	D5.1				
Version	Version initiale				
Domaines de recherche (selon classification de l'OCDE)	1.5 Sciences de la terre et de l'environnement				
Langue	français				
Date de création	2024-07-01				
Date de dernière modification	2024-07-02				
Identifiant					
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	Approche métabolomique pour décrypter l'acquisition de tolérance induite par les (bio)fongicides dans les biofilms d'eau douce
Acronyme	MEMENTO
Résumé	<p>Face à une préoccupation sociétale et économique grandissante sur les risques associés aux pesticides de synthèse pour l'Homme et l'environnement, les bio-pesticides sont de plus en plus utilisés. Cependant, des études récentes montrent qu'ils peuvent avoir un effet néfaste sur des organismes non-ciblés. Cette situation est d'autant plus préoccupante que le besoin de remplacer les fongicides de synthèse retrouvés de manière ubiquitaire dans l'environnement par des composés naturels peut conduire à des situations absurdes telle que l'utilisation du cuivre dans les vignobles «bio» alors que son écotoxicité est parfaitement connue. Ainsi, il existe un besoin urgent d'évaluer l'impact potentiel des fongicides «bio» sur les écosystèmes naturels. Dans ce contexte, le projet MEMENTO vise à caractériser et comparer la réponse et la tolérance des écosystèmes aquatiques aux fongicides «bio» et à ceux de synthèse. Pour répondre à ce défi scientifique, MEMENTO propose un changement de paradigme en écotoxicologie à travers la mise en œuvre d'une approche métabolomique non ciblée basée sur la spectrométrie de masse haute résolution sur les biofilms aquatiques permettant de fournir une vision globale de la réponse moléculaire à l'échelle des communautés microbiennes. En effet, ces communautés sont un modèle très pertinent au regard de leur diversité taxonomique (tous les règnes y sont présents) et de leur rôle clef dans le fonctionnement des écosystèmes, ces communautés contribuent ainsi à</p>

des services écosystémiques majeurs. De même, la métabolomique non ciblée est une méthode de choix à travers sa capacité à mesurer simultanément l'exposition et la réponse associée tout en caractérisant les voies de signalisation impliquées dans ces réponses, ainsi qu'en permettant l'étude des interactions chimiques microbiennes. Aussi, dans le cadre conceptuel de l'adverse outcome pathway, cette approche innovante sera combinée à la caractérisation de descripteurs physiologiques/fonctionnels/structurels et de la tolérance sur les biofilms (i.e. périphyton et litière de feuille) et leurs composantes (mono/co-culture) afin de faire le lien entre l'exposition, la réponse moléculaire/fonctionnelle/structurelle et l'acquisition de tolérance à l'échelle des communautés microbiennes (i.e. tolerance outcome pathway). Ainsi, l'approche expérimentale de MEMENTO est organisée selon un gradient croissant de complexité en termes d'expérimentation (de l'étude d'un composé seul à l'étude des facteurs confondants), de modèles (de la population à la communauté) et de descripteurs biologiques (de la réponse moléculaire à la réponse structurelle et fonctionnelle). Ainsi, ces investigations consisteront à décrypter les mécanismes de réponse et d'acquisition de la tolérance à des fongicides «bio» et de synthèse seuls et en mélange et d'évaluer l'influence de facteurs environnementaux (nutriments et mélange complexe) sur l'acquisition de tolérance. Dans son ensemble, MEMENTO apportera des connaissances sur: (i) l'identité des voies métaboliques et les métabolites impliqués dans la réponse et l'acquisition de tolérance des communautés microbiennes aquatiques aux fongicides «bio» (e.g. cuivre) vs de synthèse (e.g. azole); (ii) le lien entre l'exposition à ces fongicides et les réponses moléculaires, physiologiques, fonctionnelles et structurelles des communautés aquatiques et certaines de leurs composantes (e.g. champignons, diatomées); (iii) l'influence des facteurs environnementaux dans la réponse et l'acquisition de tolérance des communautés microbiennes aquatiques aux fongicides «bio» vs de synthèse; (iv) les théories de la co-tolérance en lien avec ces composés et les facteurs environnementaux. Ce projet permettra ainsi de mieux comprendre les mécanismes d'acquisition de tolérance des communautés microbiennes à ces composés et leur conséquence pour les écosystèmes ainsi qu'à identifier des marqueurs moléculaires sensibles, spécifiques et précoces en lien avec ces réponses.

Sources de financement

- Agence Nationale de la Recherche : ANR-23-CE34-0003

Date de début

2023

Partenaires

- LABORATOIRE MICROORGANISMES : GÉNOME ET ENVIRONNEMENT

Produits de recherche :

1. Documents de gestion du projet (Texte)
2. Données expérimentales (Jeu de données)
3. Protocoles (Texte)
4. Données compilés et traités (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Artigas Joan - https://orcid.org/0000-0002-7764-7860	Laboratoire Microorganismes : Génome et Environnement - https://rnsr.adc.education.fr/structure/200012122F	<ul style="list-style-type: none"> • Personne contact (expérimentations)
Batisson Isabelle - https://orcid.org/0000-0003-2827-7768	Laboratoire Microorganismes : Génome et Environnement - https://rnsr.adc.education.fr/structure/200012122F	<ul style="list-style-type: none"> • Personne contact (protocoles)
Benaben David - https://orcid.org/0000-0001-9930-9363	BFP Biologie du fruit et pathologie - https://rnsr.adc.education.fr/structure/201119465P	<ul style="list-style-type: none"> • Responsable du plan
Bonnineau Chloé - https://orcid.org/0000-0001-6341-003X	Ecosystèmes aquatiques et changements globaux - https://rnsr.adc.education.fr/structure/201421785Y	<ul style="list-style-type: none"> • Personne contact (expérimentations)
Creusot Nicolas - https://orcid.org/0000-0002-4240-1529	Ecosystèmes aquatiques et changements globaux - https://rnsr.adc.education.fr/structure/201421785Y	<ul style="list-style-type: none"> • Coordinateur du projet • Personne contact (documents, données traitées)
Morin Soizic - https://orcid.org/0000-0003-0360-9383	Ecosystèmes aquatiques et changements globaux - https://rnsr.adc.education.fr/structure/201421785Y	<ul style="list-style-type: none"> • Personne contact (protocoles)

DMP du projet MEMENTO

Description des données et collecte ou réutilisation de données existantes

Documents de gestion du projet

Description générale du produit de recherche

Nom	Documents de gestion du projet
Description	Documents de gestion du projet: <ul style="list-style-type: none">• réponse à l'appel à projet• études<ul style="list-style-type: none">◦ design expérimental de la campagne d'échantillonnage (Mi3.1)◦ design expérimental pour l'isolement (Mi4.1)• délivrables<ul style="list-style-type: none">◦ plan de gestion de données (T5.1)• compte-rendu de réunion, notamment :<ul style="list-style-type: none">◦ Mi1.1. Minutes of the kick-off meeting◦ Mi6.1a,b. Minutes on the mid-term and final workshop• rapports, notamment :<ul style="list-style-type: none">◦ D1.1.a-d Annual/final reports for ANR• présentations• brouillons• photos (illustrations du projet)• publications et communications : publications scientifiques, communication en congrès (abstract, présentation), site web MEMENTO, support pour les journées scientifiques publiques, messages sur les réseaux sociaux (tout au long du projet, tel que X/Twitter)
Type	Texte
Workpackage	WP1, WP2, WP3, WP4, WP5, WP6
Mots clés (texte libre)	
Langue	
Date de publication	
Contient des données personnelles ?	Non
Contient des données sensibles ?	Non
Prend en compte des aspects éthiques ?	Non

Est-ce que des données existantes seront réutilisées ?

Comment seront produites/collectées les nouvelles données ?

Nom de la méthode	Edition collaborative et stockage
Description	Edition collaborative de documents « OnlyOffice » et stockage de document produits localement (éditeur de texte, tableur, etc.).
Nature des données	

Données expérimentales

Description générale du produit de recherche

Nom	Données expérimentales
Description	<p>Données expérimentales :</p> <ul style="list-style-type: none">• métabolomique non ciblée (UPLC-HRMS)• lipidomique ciblée (GC-MS/MS : LC-MS/MS) sur des classes spécifiques (phospholipides, glycolipides, acides gras)• activité photosynthétique et teneur en chlorophylle (PhytoPAM)• pourcentage de tératologies, si nécessaire (observations au microscope)• activités enzymatiques (spectrofluorimètre)• respiration microbienne (tests à base de résazurine)• production primaire (mesure de l'O₂)• diversité microbienne (metabarcoding basé sur le séquenceur de nouvelle génération au LMGE)• mesures des conditions d'expositions: T°C, pH, luminosité, dosage des nutriments, dosage des fongicides• biomasse/biovolume de chaque composant du biofilm initial (wp4) <p>Ces données expérimentales peuvent également être temporelles (mesures PICT).</p> <p>Données compilées et traitées :</p> <ul style="list-style-type: none">• liste des descripteurs moléculaires (métabolites) et des voies de tolérance aux (bio)fongicides (WP2)• liste des candidats biomarqueurs faiblement influencés par le nutriment ou le cocktail de contaminants (WP3)
Type	Jeu de données
Workpackage	WP2, WP3, WP4
Mots clés (texte libre)	
Date de publication	
Contient des données personnelles ?	Non
Contient des données sensibles ?	Non
Prend en compte des aspects éthiques ?	Non

Est-ce que des données existantes seront réutilisées ?

Justification	<p>MEMENTO pourra bénéficier des résultats des projets en cours MICROBIOMIQ et TAPIOCA, liées à l'exposition et à l'écotoxicité des pesticides et de leurs produits de transformation sur les biofilms.</p> <p>Les teneurs en nutriments seront sélectionnées sur la base de leurs valeurs sur le terrain issu du projet PSGAR MAIA.</p>
----------------------	--

Comment seront produites/collectées les nouvelles données ?

Nom de la méthode	Expérimentations et mesures
Description	Données expérimentales issues de diverses expérimentations (PICT, mesures des conditions d'expositions:, etc) et divers instruments (UPLC-HRMS, GC-MS/MS, LC-MS/MS, spectrofluorimètre, ...) Traitement des données via les outils d'analyses tel que DRomics, W4M, MS DIAL, MzMine3, GNPS, MS-FINDER, MetFrag, Sirius6, Cytoscape, MetExplore, R, QIIME2, Past.
Nature des données	Données expérimentales
Equipements, plateaux techniques utilisés	<ul style="list-style-type: none"> Bordeaux Metabolome : https://doi.org/10.15454/1.5572412770331912E12

Protocoles	
Description générale du produit de recherche	
Nom	Protocoles
Description	Protocoles expérimentaux produits pour le projet MEMENTO <ul style="list-style-type: none"> D4.1. Standard Operating Procedure (SOP) for isolation Protocole d'isolement d'au moins une espèce de champignon et de bactérie pour la litière de feuilles, et de diatomée et cyanobactérie pour le périphyton. D4.2. Standard Operating Procedure (SOP) for building co-culture
Type	Texte
Workpackage	WP4
Date de publication	
Contient des données personnelles ?	Non
Contient des données sensibles ?	Non
Prend en compte des aspects éthiques ?	Non
Est-ce que des données existantes seront réutilisées ?	
Comment seront produites/collectées les nouvelles données ?	
Nom de la méthode	

Données compilés et traités	
Description générale du produit de recherche	

Nom	Données compilés et traités
Description	<p>Analyser les données collectées dans les WP2 et WP3.</p> <p>Données compilées et traitées :</p> <ul style="list-style-type: none"> • Annotation des métabolites et des voies associées (T5.2) • caractérisation des descripteurs physiologiques, fonctionnels et structurels (T5.3) • méta-analyse des empreintes métaboliques (T5.4) • caractérisation des relations entre le métabolome et les autres descripteurs (T5.5) • caractérisation de l'influence des facteurs environnementaux sur les empreintes métabolomiques et les marqueurs d'intérêt identifiés (T5.6)
Type	Jeu de données
Workpackage	WP5
Mots clés (texte libre)	
Date de publication	
Contient des données personnelles ?	Non
Contient des données sensibles ?	Non
Prend en compte des aspects éthiques ?	Non

Est-ce que des données existantes seront réutilisées ?

Justification	De nombreuses banques en ligne seront utilisées (MassBank, MetaCyc, AOP Wiki, BioCyc , PubChem, MoNA, PeakForest, COCONUT, KEGG, KNAPSAcK) pour identifier les métabolites, voie métabolique, annoter, etc.
Données réutilisées	<ul style="list-style-type: none"> • MassBank : https://massbank.eu/MassBank/ • MetaCyc : https://metacyc.org/ • AOP : https://aopwiki.org/ • BioCyc : https://biocyc.org/ • PubChem : https://pubchem.ncbi.nlm.nih.gov/ • MoNA : https://mona.fiehnlab.ucdavis.edu/ • PeakForest : https://peakforest.org/ • COCONUT : https://coconut.naturalproducts.net/ • KEGG : http://www.genome.jp/kegg/ • KNAPSAcK : http://www.knapsackfamily.com/KNAPSAcK/

Comment seront produites/collectées les nouvelles données ?

Nom de la méthode	Analyse des données
Description	<p>Les données métabolomiques seront traitées par les solutions adaptées:</p> <ul style="list-style-type: none"> • W4M • MS-DIAL • MzMine • GNPS • MetaboAnalyst • R (PCA, HCA et (O)PLS-DA, voire ASCA, AMOPLS, analyses multi-blocs, méthodes basées sur l'apprentissage automatique : GLM, PLS-R) • DRomics • MixOmics/DIABLO et mmec (QIIME2) (analyse des réseaux de cooccurrence)

Documentation et qualité des données

Documents de gestion du projet

Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Description

Les dossiers sont organisés par WP comme suit:

- Biblio
- Photos
- WP1 - Project management
- WP2 - Tolerance acquisition
- WP3 - Confounding factors
- WP4. Isolation
- WP5. Data management & Analysis
- WP6. Dissemination & Communication

Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Données expérimentales

Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Description

Les données seront décrits à l'aide d'un fichier descriptif (fichier structuré ou simple fichier texte type "README.txt") ou via un onglet dédié dans les tableaux de données.

Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Données compilés et traités

Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Exigences légales et éthiques, code de conduite

Documents de gestion du projet

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 ?

Données expérimentales

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 ?

Description

Les données produites seront placés sous licence libre conformément aux recommandations (<https://anr.fr/fr/lanr/engagements/la-science-ouverte/>, <https://www.data.gouv.fr/fr/pages/legal/licences/>, <https://science-ouverte.inrae.fr/fr/les-donnees-et-le-numerique-scientifiques/partager-publier-des-donnees-et-des-codes/comment-choisir>).

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

Protocoles

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 ?

Description

Les données produites seront placés sous licence libre conformément aux recommandations (<https://anr.fr/fr/lanr/engagements/la-science-ouverte/>, <https://www.data.gouv.fr/fr/pages/legal/licences/>, <https://science-ouverte.inrae.fr/fr/les-donnees-et-le-numerique-scientifiques/partager-publier-des-donnees-et-des-codes/comment-choisir>).

Données compilés et traités

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 ?

Description

Les données produites seront placés sous licence libre conformément aux recommandations (<https://anr.fr/fr/lanr/engagements/la-science-ouverte/>, <https://www.data.gouv.fr/fr/pages/legal/licences/>, <https://science-ouverte.inrae.fr/fr/les-donnees-et-le-numerique-scientifiques/partager-publier-des-donnees-et-des-codes/comment-choisir>).

Traitement et analyse des données

Documents de gestion du projet

Comment et avec quels moyens seront traitées les données ?

Données expérimentales

Comment et avec quels moyens seront traitées les données ?

Références associées

- DRomics : <https://lbbe.univ-lyon1.fr/fr/dromics>
- Workflow4Metabolomics (W4M) : <https://workflow4metabolomics.org/>
- MS-DIAL : <https://systemsomicslab.github.io/compms/msdial/main.html>
- MetaboAnalyst : <http://www.metaboanalyst.ca>
- MetFrag : <http://msbi.ipb-halle.de/MetFragBeta/>
- Sirius : <http://bio.informatik.uni-jena.de/sirius>
- PAST : <https://www.nhm.uio.no/english/research/resources/past/>
- R : <http://www.r-project.org/>
- MixOmics : <http://mixomics.org/>
- MS-Finder : <http://prime.psc.riken.jp/compms/msfinder/main.html>
- MZmine 3 : <http://mzmine.github.io/download.html>
- GNPS : <https://gnps.ucsd.edu/>
- Cytoscape : <https://cytoscape.org/>
- MetExplore : <https://metexplore.toulouse.inrae.fr/>
- QIIME 2 : <https://qiime2.org/>

Protocoles

Comment et avec quels moyens seront traitées les données ?

Données compilés et traités

Comment et avec quels moyens seront traitées les données ?

Description

Les données seront traitées sur les postes de travail (MS-DIAL, MzMine, GNPS, R, DRomics, MixOmics/DIABLO, mmecc, PAST) ou via des plateformes en ligne (W4M, MetaboAnalyst)

Références associées

- Workflow4Metabolomics (W4M) : <https://workflow4metabolomics.org/>
- MS-DIAL : MS-DIAL
- MetaboAnalyst : <http://www.metaboanalyst.ca>
- R : <http://www.r-project.org/>
- DRomics : <https://lbbe.univ-lyon1.fr/fr/dromics>
- MixOmics : <http://mixomics.org/>
- MS-Finder : <http://prime.psc.riken.jp/compms/msfinder/main.html>
- MetFrag : <http://msbi.ipb-halle.de/MetFragBeta/>
- PAST : <https://www.nhm.uio.no/english/research/resources/past/>
- MZmine 3 : <http://mzmine.github.io/download.html>
- GNPS : <https://gnps.ucsd.edu/>

Stockage et sauvegarde des données pendant le processus de recherche

Documents de gestion du projet

Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage	Dépôt des documents ou création sur le NextCloud INRAE https://nextcloud.inrae.fr/ Espace dédié: MEMENTO
Volume estimé des données	100
Unité	Go
Equipements, plateaux techniques	<ul style="list-style-type: none">• NextCloud INRAE : https://nextcloud.inrae.fr/
Politique de sauvegarde	<ul style="list-style-type: none">• NextCloud INRAE Réplication: cela ne constitue pas une sauvegarde mais s'inscrit dans une logique de plan de continuité d'activité (PCA/PRA) permettant d'accéder aux données si le site principal (Toulouse) est injoignable (sous réserve d'un délai de bascule). NextCloud INRAE Rétention de 45 jours: permet de conserver les versions antérieures d'un élément sur une période donnée (s'apparente à une forme de sauvegarde). Sauvegarde: les données sont copiés sur un site distant, une fois par jours, sur la solution AgroDataRing.
Mesures prises pour la sécurité des données	La sécurité de l'infrastructure est assuré par la DSI INRAE. L'accès est limité aux utilisateurs authentifiés.

Données expérimentales

Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage

Les données brutes, mesures et résultats d'observations seront stockés sur le site de production (solution de stockage interne au site : NAS EABX, poste de travail).

Les données brutes de metabarcoding traitées par un prestataire ne sont pas récupérées (pas utile pour le projet et n'étant pas réutilisable), seules les données (fastq) curées sont gardées et seront publiées (NCBI)

Les données traitées seront partagées sur le NextCloud INRAE.

Elles pourront également être envoyés sur les dépôts institutionnels (tel que recherche.data.gouv.fr) en préparation de la publication des données.

Données:

- *.csv*, tableau de données, biomasse/biovolume
- données brutes séquenceur
- *.fastq* (données de séquençage traitées)
- liste des descripteurs moléculaires (métabolites) et voies de tolérance aux (bio)fongicides
- *.mzML* (lipidomique)

Volume estimé des données 15

Unité To

Equipements, plateaux techniques

- NextCloud INRAE : <https://nextcloud.inrae.fr/>
- EABX NAS (bxdata) :
- Collec-Science : <https://www.collec-science.org/>
- UCAdrive : <https://drive.uca.fr/>
- STORM (STOckage Recherche Mésocentre) : <https://mesocentre.uca.fr/ressources>
- Postes de travail personnel :

Politique de sauvegarde

- MEMENTO s'appuie sur des infrastructures supportées et sauvegardées par les DSI correspondantes (INRAE, UCA) ou par l'équipe informatique locale (tel que le NAS bxdata). Toutes les données sont copiées sur ces infrastructures sécurisés (NextCloud INRAE, NAS bxdata, STORM, UCAdrive) et déposés dans les entrepôts thématique tel que le NCBI pour les données omiques ou le CEBA (Cloud Environnemental au bénéfice de l'Auvergne).

Mesures prises pour la sécurité des données

MEMENTO s'appuie sur des infrastructures sécurisés par les DSI correspondantes (INRAE, UCA) ou par l'équipe informatique locale (tel que le NAS EABX). Les accès sont limités au personnel identifiés et authentifiés.

Protocoles

Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage

Les mesures et protocoles seront stockés sur le site de production (solution de stockage interne au laboratoire) puis partagé sur le NextCloud INRAE.

Volume estimé des données 100

Unité Mo

Equipements, plateaux techniques

- NextCloud INRAE : <https://nextcloud.inrae.fr/>

Politique de sauvegarde

- NextCloud INRAE Réplication: cela ne constitue pas une sauvegarde mais s'inscrit dans une logique de plan de continuité d'activité (PCA/PRA) permettant d'accéder aux données si le site principal (Toulouse) est injoignable (sous réserve d'un délai de bascule).
NextCloud INRAE Rétention de 45 jours: permet de conserver les versions antérieures d'un élément sur une période donnée (s'apparente à une forme de sauvegarde).
Sauvegarde: les données sont copiées sur un site distant, une fois par jours, sur la solution AgroDataRing.

Mesures prises pour la sécurité des données

La sécurité de l'infrastructure est assurée par la DSI INRAE. L'accès est limité aux utilisateurs authentifiés.

Données compilés et traités

Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage	Les résultats seront stockés sur le site d'analyse (postes de traitements, postes personnels, cluster de calcul) puis partagé sur le NextCloud INRAE.
Volume estimé des données	100
Unité	Mo
Equipements, plateaux techniques	<ul style="list-style-type: none">• NextCloud INRAE : https://nextcloud.inrae.fr/
Politique de sauvegarde	<ul style="list-style-type: none">• NextCloud INRAE Réplication: cela ne constitue pas une sauvegarde mais s'inscrit dans une logique de plan de continuité d'activité (PCA/PRA) permettant d'accéder aux données si le site principal (Toulouse) est injoignable (sous réserve d'un délai de bascule). NextCloud INRAE Rétention de 45 jours: permet de conserver les versions antérieures d'un élément sur une période donnée (s'apparente à une forme de sauvegarde). Sauvegarde: les données sont copiés sur un site distant, une fois par jours, sur la solution AgroDataRing.
Mesures prises pour la sécurité des données	La sécurité de l'infrastructure est assuré par la DSI INRAE. L'accès est limité aux utilisateurs authentifiés.

Partage des données et conservation à long terme

Documents de gestion du projet

Comment les données seront-elles partagées ?

Modalités de partage	Les livrables seront transmis aux parties prenantes ou publiés dans des dépôts.
Potentiel de réutilisation	

Comment les données seront-elles conservées à long terme ?

Justification	
Unité	
Date de début	
Date de fin	
Dispositions finales	

Données expérimentales

Comment les données seront-elles partagées ?

Modalités de partage

Les données seront rendues publiques dans les entrepôts thématiques (tel que le NCBI) ou institutionnels (tel que le CEBA) après une période d'embargo en accord avec le processus de valorisation des résultats du projet MEMENTO.

Les données expérimentales, les données traitées, les protocoles et les articles scientifiques seront disponible au moment de la publication scientifique afférente et au plus tard à la fin du projet.

L'exploitation sera favorisée par un partage adéquat des données (e.g. metabolic pathways, metabolites, AOP) dans les entrepôts de données (tel que RechercheDataGouv, MetaboLights) et les banques en lignes (tel que MassBank, MetaCyc, AOP Wiki).

En accord avec les principes F.A.I.R. (Findable, Accessible, Interoperable, Reusable), MEMENTO s'attachera à

- utiliser des formats de données ouverts (mzML en lieu et place des .raw),
- utiliser des logiciels au plus proches des principes F.A.I.R. pour les logiciels en recherche (« FAIR4RS Principles », « Evaluating LC-HRMS metabolomics data processing software using FAIR principles for research software »)
- mettre à disposition les codes sources développés sous une licence libre
- utiliser des identifiant pérennes (PID) dès que possible

Potentiel de réutilisation

Entrepôt/Catalogue de données

- CEBA Cloud Environnemental : [https://cat.opidor.fr/index.php/CEBA Cloud Environnemental](https://cat.opidor.fr/index.php/CEBA%20Cloud%20Environnemental) ()
- Entrepôt Recherche Data Gouv : [https://cat.opidor.fr/index.php/Entrepôt Recherche Data Gouv](https://cat.opidor.fr/index.php/Entrep%C3%B4t%20Recherche%20Data%20Gouv) ()
- NCBI : <https://submit.ncbi.nlm.nih.gov/> ()
- MetaCyc : <https://metacyc.org/> ()
- AOP : <https://aopwiki.org/> ()
- MetaboLights : <https://www.ebi.ac.uk/metabolights/> ()
- MassBank : <https://massbank.eu/MassBank/> ()

Comment les données seront-elles conservées à long terme ?

Justification

La conservation à long terme est assurée par les entrepôts thématiques (tel que le NCBI) ou institutionnels (tel que le CEBA).

Volume estimé des données 15

Unité To

Date de début

Date de fin

Dispositions finales

Données compilés et traités

Comment les données seront-elles partagées ?

Modalités de partage

Les résultats et processus de traitement seront partagés au moment de la publication.

En accord avec les principes F.A.I.R. (Findable, Accessible, Interoperable, Reusable), MEMENTO s'attachera à

- utiliser des formats de données ouverts,
- utiliser des logiciels au plus proches des principes F.A.I.R. pour les logiciels en recherche (« FAIR4RS Principles », « Evaluating LC-HRMS metabolomics data processing software using FAIR principles for research software »)
- mettre à disposition les codes sources développés sous une licence libre
- utiliser des identifiants pérennes (PID) dès que possible

Potentiel de réutilisation