
DMP du projet "FUTUROC"

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

Titre du plan	DMP du projet "FUTUROC"
Livrable	
Version	Version initiale
Domaines de recherche (selon classification de l'OCDE)	5.1 Psychologie et sciences cognitives
Langue	fra
Date de création	2024-06-25
Date de dernière modification	2024-06-25
Identifiant	

Renseignements sur le projet

Titre du projet	DPM FUTUROC
Acronyme	FUTUROC
Sources de financement	<ul style="list-style-type: none">Agence Nationale de la Recherche : ANR-23-CE28-0001-01
Date de début	
Date de fin	

Produits de recherche :

- Produit de recherche 1

Contributeurs

Nom	Affiliation	Rôles
Ernst Alexandra		<ul style="list-style-type: none">Coordinateur du projetPersonne contactResponsable du plan

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 "FUTUROC"

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 FUTUROC est un projet en psychologie cognitive avec une approche développementale. Ce projet vise à tester le rôle de la métacognition dans la Pensée Future Épisodique (PFE). Pour ce faire, nous allons étudier l'impact de variations développementales dans la disponibilité des "ingrédients cognitifs" formant les croyances métacognitives à travers les âges de la vie.

Les nouvelles données seront collectées dans le cadre de 3 études comportementales, après l'obtention des consentements libres, éclairés et informés des participants. Les données seront collectées chez 4 groupes de participants (enfants (8-11 ans), adolescents (13-16 ans), jeunes adultes (20-29 ans) et adultes âgés (plus de 65 ans)).

Aucune donnée préexistante ne sera réutilisée dans le cadre de cette étude.

Les outils de mesures suivants seront utilisés pour récolter les nouvelles données :

Étude 1

Questionnaires en ligne : Version 5.2.9 (*LimeSurvey Community Edition*)

Questionnaire papier : Réponses aux questionnaires au format papier.

Autres données physiques : Formulaire de consentement

Étude 2

Fiches d'entretien et jugements métacognitifs: Tableur et texte (Microsoft Pack Office 2019 Word et Excel)

Questionnaires FoFTS : Version 5.2.9 (*LimeSurvey Community Edition*) / Format papier

Enregistrement audio des entretiens : Dictaphone OLYMPUS VN-541PC, Audacity

Retranscription des entretiens : Python, Microsoft Pack Office 2019 Excel, Microsoft Pack Office 2019 Word

Données d'eye tracking (oculométrie) : PupilLabs (<https://github.com/pupil-labs/pupil>)

Étude 3

Fiches d'entretien et jugements métacognitifs: Tableur et texte (Microsoft Pack Office 2019 Word et Excel)

Enregistrement audio des entretiens : Dictaphone OLYMPUS VN-541PC, Audacity

Retranscription des entretiens : Python, Microsoft Pack Office 2019 Excel, Microsoft Pack Office 2019 Word

Données d'eye tracking (oculométrie) :

Pupil Core (présentation du matériel : <https://youtu.be/kjjPL7qLy7s>)

PupilLabs : (<https://github.com/pupil-labs/pupil>)

Traitement des données :

Logiciels et langage prévu : R, Python, Jamovi, Microsoft Pack Office 2019 Excel, Audacity

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

Des disques durs d'1To seront mis à disposition pour le stockage des données.

Les volumes indiqués sont prévisionnels et seront mis à jour dans les prochaines versions du PGD.

→ Formulaire de consentements : Format papier ou numérique

◆ Volumes :

● Etude 1 : 80 participants x 4 groupes

● Etude 2 : 45 participants x 4 groupes

● Etude 3 : 45 participants x 4 groupes

→ Fiches d'entretien et jugements métacognitifs : Format numérique, .csv

◆ Volumes :

● Etude 2 : 45 participants x 4 groupes

● Etude 3 : 45 participants x 4 groupes

- Enregistrement audio des entretiens : .wav
- ◆ Volumes :
- Etude 2 : 45 participants x 4 groupes
- Etude 3 : 45 participants x 4 groupes

A noter que les enregistrements audio des entretiens ne seront pas conservés après transcription. Seule la version transcrite des entretiens sera archivée.

→ Retranscription des entretiens : .csv ; .txt ; .doc

- ◆ Volumes :
- Etude 2 : 45 participants x 4 groupes
- Etude 3 : 45 participants x 4 groupes

→ Données d'eye tracking (oculométrie) : .npy ; .csv

- ◆ Volumes :
- Etude 2 : 45 participants x 4 groupes
- Etude 3 : 45 participants x 4 groupes

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 ?

Nous allons utiliser les standards de métadonnées BIDS pour faciliter la recherche et l'identification des données. La structure et l'organisation des données, les noms des fichiers et les métadonnées seront conformes aux recommandations BIDS (Brain Imaging Data Structure; [Gorgolewski et al., 2016](#)).

→ Les BIDS pour les données comportementales sont spécifiés ici : <https://bids-specification.readthedocs.io/en/stable/04-modality-specific-files/07-behavioral-experiments.html>

→ L'extension BIDS pour les données d'Eye-Tracking : <https://docs.google.com/document/d/1eggzTCzSHG3AEKhtnEDbckd-2avXN6I94X8aUPEBVsw/edit?usp=sharing>.

Nous allons mettre en place un système de dénominations de fichier dans le but de standardiser le processus de traitement des données et minimiser les risques d'erreurs, et en faciliter l'accès et l'utilisation.

Afin de faciliter le contrôle de la qualité des données par les différents collaborateurs, ainsi que leur partage, les données seront documentées par un fichier "code book" (i.e., un dictionnaire de variables) regroupant la signification des différents codes et acronymes utilisés, noms des variables, ainsi que la cotation des items des questionnaires.

Les scripts de traitement et d'analyse de données seront accompagnés d'un fichier "README" contenant la documentation nécessaire à la reproduction des analyses et à leur potentielle réutilisation (information sur la méthodologie utilisée pour collecter les données, sur les procédures et méthodes d'analyse utilisées, sur la définition des variables, des unités de mesure, etc.)

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

Une procédure standardisée de calibration de l'oculométrie sera mise en place afin de s'assurer de la qualité des données d'eye tracking.

Un questionnaire post-expérimental permettra de recueillir des données supplémentaires sur les stratégies mises en place par les participants pendant la tâche dans le but de vérifier la bonne réalisation de celle-ci.

Un fichier tableur (.xls) sera mis en place pour permettre de faire l'inventaire des données récoltées à chaque session, le statut de stockage, nombres et emplacements de sauvegarde, statut du traitement et statut des données brutes.

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 ?

Les données seront stockées via le service ShareDocs de la TGIR Huma-Num (<https://www.huma-num.fr/les-services-par-etapes/#collecte>) sur un serveur sécurisé, conforme aux normes et recommandations RGPD, permettant une sauvegarde automatique des fichiers. Alexandra Ernst sera gestionnaire du compte projet sur ShareDocs et les membres suivants de l'équipe investigatrice auront accès au partage de données : Wassila El Mardi, Claire Vallat-Azouvi, Marianne Habib et Geoffrey Tissier.

Deux sauvegardes supplémentaires seront effectuées sur deux disques durs externes cryptés.

Les disques durs externes ainsi que les données physiques (formulaire de consentements, questionnaires et fiches d'entretiens, etc) seront stockés dans un casier fermé à clé dans un bureau au sein de l'unité de recherche DysCo (seule la doctorante Wassila EL MARDI et la Coordinatrice scientifique du projet, Dr. Alexandra ERNST, auront accès à cette clé de casier).

À la fin de l'étude, les formulaires de consentement seront archivés dans une enveloppe scellée placée dans un coffre verrouillé situé dans l'unité de recherche DysCo, à laquelle seule la responsable scientifique du projet aura accès.

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

Les données seront stockées et partagées via le service ShareDocs de la TGIR Huma-Num (<https://www.huma-num.fr/les-services-par-etapes/#collecte>). Ce système constitue un système de sauvegarde pérenne et sécurisé des données. Alexandra Ernst sera gestionnaire du compte projet sur ShareDocs et les membres suivants de l'équipe investigatrice auront un droit d'accès aux données via ce service : Wassila El Mardi, Claire Vallat-Azouvi, Marianne Habib et Geoffrey Tissier.

Concernant la politique institutionnelle des données, une inscription au registre des traitements de l'université Paris 8 sera réalisée.

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 ?

Aucune donnée ne sera recueillie auprès des participants avant l'obtention d'un consentement libre, éclairé et informé. Les données recueillies dans le cadre de cette étude - tout au long de la procédure, du traitement ainsi qu'après la réalisation des études - seront conservées et traitées statistiquement dans le respect de l'anonymat des participants selon le Règlement général sur la protection des données (RGPD).

Dans les trois études, nous utiliserons le procédé de pseudonymisation suivant :

Les données seront pseudonymisées, c'est-à-dire que chaque participant se verra attribuer un code d'anonymat unique pour l'ensemble de ses données. Le code de pseudonymisation sera construit de la façon suivante :

- La combinaison du code de l'étude (i.e., ETD+[Numéro de l'étude]; e.g., ETD1 = étude 1) avec deux lettres d'identification du groupe d'âge (EN = enfant, AD = adolescent, AJ = adultes jeunes et AA adultes âgés) sera combinée à un numéro de participant (e.g., ETD1EN11, ETD3AD20).

Le code sera conservé par le participant sur l'exemplaire de la notice d'information qui lui est destinée. Ainsi, chaque participant pourra exercer son droit de retrait en contactant les responsables scientifiques de l'étude et en leur communiquant le code associé à ses données.

Pour l'étude 3, le participant recevra son code lors de la première et seconde session de test et les données collectées pendant ces deux jours porteront le nom de code du participant. Ce code servira à mettre en lien les différentes données pseudonymisées des différentes sessions de test. Un fichier de correspondance indiquant le code du participant et le nom du participant sera conservé le temps de l'étude dans un fichier crypté dont seule la doctorante (expérimentatrice) et la responsable scientifique de l'étude auront le mot de passe. À l'issue du projet, cette table de correspondance sera détruite. Cela permettra de pouvoir recontacter les participants pour la seconde session de passation au temps prévu (1 mois plus tard) et le fichier sera aussi utilisé en cas d'oubli de son code de pseudonymisation par le participant.

Les enregistrements audios ne seront réalisés qu'avec l'accord du participant qui sera informé que celui-ci ne sera accessible qu'aux personnes impliquées dans la recherche. Dans le but d'identifier l'enregistrement audio pour sa transcription ultérieure, le code de pseudonymisation du participant sera prononcé au début de l'enregistrement et apparaîtra dans le nom du fichier d'enregistrement, dans le format suivant :

[CodeParticipant]_[TypeDeFichier]_[NumeroDeSession] => (e.g., ETD3AA15_REC_S1)

Ces enregistrements audio seront supprimés après transcription et seule la version écrite et pseudonymisée sera conservée et archivée, en faisant uniquement mention du code pseudonymisation du participant dans le nom du fichier.

dans le format suivant :

[CodeParticipant]_[TypeDeFichier] => (e.g., ETD3AA15_TRANSCRIPT)

Les noms de personnes citées dans les récits audio ne seront pas transcrits de manière complète, mais remplacés par une initiale. Toute autre information pouvant avoir un caractère identifiant (e.g., lieux, adresses) sera remplacée dans la transcription par une initiale ou une version simplifiée de l'information.

Chaque participant sera libre de pouvoir réclamer, supprimer ou récupérer personnellement ses données à tout moment auprès de l'investigateur principal, en précisant le code de pseudonymisation présent sur le document d'information qui lui a été remis (ce code n'est noté que sur la version remise au participant). en cas de perte du code par le participant, l'exécution de la demande de rétractation du participant ne pourra être effectuée que si la demande de rétractation est envoyée avant la destruction du fichier de correspondance entre le code de pseudonymisation et le nom du participant.

Enfin, les données personnelles de contact de chaque participant, recueillies lors du recrutement pour l'organisation du planning des passations expérimentales (e.g., adresse mail) seront également détruites dès la fin de leur participation.

Seules les informations du genre, le nombre d'années d'études et l'âge des participants seront conservés en vue de l'analyse des données.

Les données personnelles ne seront consultables que par la responsable scientifique de l'étude et par la doctorante qui effectuera le traitement des données.

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 ?

Alexandra Ernst sera gestionnaire du compte projet sur ShareDocs et attribuera un droit d'accès aux données aux différents membres suivants de l'équipe investigatrice : Wassila El Mardi, Claire Vallat-Azouvi, Marianne Habib et Geoffrey Tissier. L'équipe investigatrice aura une propriété partagée des données.

L'ensemble des études seront pré-enregistrées sur Open Science Framework (OSF) et les données anonymisées brutes seront rendues disponibles sur OSF à l'issue du projet, suivant ainsi le cadre de la Loi pour une République Numérique (Loi Lemaire), qui prévoit que les données issues d'une activité de recherche sont librement réutilisables si l'activité de recherche est financée au moins pour moitié par des fonds publics.

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

La mise en œuvre des trois études a été conditionnée par l'obtention au préalable d'un accord éthique auprès du CEEI/IRB de l'INSERM.

Les investigateurs s'engagent à respecter et à suivre le protocole de l'étude. Ils s'engagent également à respecter les textes réglementaires et législatifs en vigueur, relatifs à la recherche et plus précisément aux recherches non interventionnelles. Ce travail de recherche sera mené dans le respect de la réglementation française en vigueur, notamment la loi informatique et liberté modifiée (loi 78-17 du 6 janvier 1978), la Déclaration d'Helsinki (JAMA ; Octobre 19, 2013. Doi : 10.1001/jama.2013.281053). Les investigateurs-coordonnateurs s'engagent à ce que la recherche soit menée conformément à ces dispositions éthiques et réglementaires.

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 ?

Les publications associées aux études seront mises en ligne en version pre-print sur le répertoire HAL une fois acceptée.

Les délais de publication prévus sont les suivants :

- Étude 1 : année 1
- Étude 2 : année 3
- Étude 3 : année 3

L'ensemble des études seront pré-enregistrées sur Open Science Framework (OSF) et les données anonymisées brutes (sous la forme de tableaux de données Excel) seront rendues disponibles sur OSF à l'issue du projet, c'est-à-dire après la publication des résultats dans des revues internationales à comité de lecture. Cette procédure renvoie à la Loi pour une République Numérique (Loi Lemaire), qui prévoit que les données issues d'une activité de recherche sont librement réutilisables si l'activité de recherche est financée au moins pour moitié par des fonds publics.

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) ?

Le promoteur (Université Paris 8) et la responsable scientifique (Alexandra Ernst) conservent les documents et données relatifs à la recherche qui leur sont spécifiques pendant une période de deux à dix ans, selon les données, après la fin de la recherche ou son arrêt anticipé.

Ces documents et données comprennent l'ensemble des données personnelles collectées durant la recherche décrite au présent protocole. Une fois la durée de conservation écoulée, le service des archives du promoteur peut procéder à un passage des documents et données en archives définitives, après une opération de tri. Ces documents et données comprennent l'ensemble des données personnelles collectées durant la recherche décrite au présent protocole. Les consentements seront conservés pour un délai de 10 ans après valorisation de la recherche. Les retranscriptions des entretiens, les questionnaires et les tableaux de données (qui correspondent à des données anonymisées) seront conservés sans limite de durée, permettant une éventuelle réutilisation dans le cadre de projets de recherche (e.g., méta-analyse).

Les données issues des études 1, 2 et 3 seront utilisées à des fins de publications et de communications scientifiques par les membres de l'équipe investigatrice.

L'utilisation du service via le service ShareDocs de la TGIR Huma-Num pour le stockage des données assure leur sécurité ainsi que leur pérennité au-delà de la durée du financement du projet.

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

Liste des logiciels nécessaires pour accéder et utiliser les données :

- Microsoft Pack Office 2019 Word et Microsoft Pack Office 2019 Excel. / Libre Office
- R (pour les scripts d'analyse des données)

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

Un DOI sera attribué à chaque ensemble de données des études décrites grâce à la plateforme OSF, ce qui permettra l'attribution d'un identifiant pérenne aux données. Les données pourront être réutilisées à des fins de publications ou de communications scientifiques (e.g., méta-analyse).

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) ?

Les rôles des différents membres de l'équipe investigatrice sont décrits ci-dessous. Le partage des données collectées au sein de l'équipe investigatrice se fera via ShareDocs, comme décrit précédemment (cf. 3. Stockage et sauvegarde pendant le processus de recherche).

Le PGD sera publié à 6 mois après le démarrage de l'étude, puis mis à jour à 27 mois et 36 mois pour guider le partage et la réutilisation des données. Ces mises à jour seront sous la responsabilité de la responsable scientifique du projet.

- **Alexandra Ernst : MCF**, Gestionnaire des données; Responsable scientifique du projet et responsable de la mise en oeuvre du PGD ; co-directrice de thèse de Wassila El Mardi.
→ Qualité et utilisation des données (analyses statistiques), archivage et partage des données.
- **Wassila El Mardi : Doctorante**
→ Saisie des données, production des métadonnées, stockage et sauvegarde, archivage et partage des données.
- **Claire Vallat-Azouvi : PR**, Collaboratrice scientifique et directrice de thèse de Wassila El Mardi
→ Qualité et utilisation des données (analyses statistiques), archivage et partage des données.
- **Marianne Habib : MCF**, Collaboratrice scientifique
→ Qualité et utilisation des données (analyses statistiques)
- **Geoffrey Tissier : Ingénieur d'études**, collaborateur scientifique
→ Qualité et utilisation des données (analyses statistiques)

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) ?

Wassila El Mardi dispose d'un contrat doctoral (CDD de 36 mois à temps plein, financé par l'ANR N° ANR-23-CE28-0001-01) couvrant la durée complète des trois études. 100% de son temps de travail est donc dédié à ces trois études, incluant donc le temps nécessaire à la préparation des données, leur partage et préservation.

Par ailleurs, les temps recherche dédiés à ce projet pour les autres membres de l'équipe investigatrice, conformément au projet ANR, sont les suivants :

- Alexandra Ernst : 80% Research Time
- Claire Vallat-Azouvi : 25% Research Time
- Marianne Habib : 25% Research Time
- Geoffrey Tissier : 10% Research Time

Les coûts matériels de stockage (achat de disques durs externes) seront couverts par le financement ANR (N° ANR-23-CE28-0001-01). Le stockage sur ShareDocs, ainsi que la diffusion ultérieure des données à l'issue du projet sur OSF, ne requièrent aucun financement.

Concernant la diffusion des données et des résultats de ces recherches, les frais de publication et de communications scientifiques seront également pris en charge par le financement ANR N° ANR-23-CE28-0001-01.

Nous avons sollicité l'avis de M. Goran Sekulovski (Chargé de mission Données de la recherche et Science Ouverte - Université Paris 8) sur notre plan de gestion des données, nous l'avons en particulier consulté concernant la question du stockage et du partage des données. M. Sekulovski nous a mis en contact avec le service HumaNum pour faire une demande de création d'un compte projet sur l'outil ShareDocs. Une présentation de l'outil ShareDocs par M. Sekulovski est prévue pour permettre de faciliter la prise en main de cet outil par les membres de ce projet.