

Enseigner le cycle de vie des données à la Faculté de médecine



Pourquoi enseigner le cycle de vie des données à la Faculté de médecine ?

Les données – plus communément considérés comme le matériel de recherche – constituent un élément central de la pratique médicale et de la recherche en médecine fondamentale et clinique. La bonne gestion des données de recherche (GDR) et l'ouverture des données de recherche (Open Research Data - ORD) sont essentielles pour garantir la qualité des résultats, assurer la fiabilité et la reproductibilité des travaux scientifiques, préserver la confiance des patient-es et des partenaires et, par conséquent, la qualité des soins. Les professionnel-les du secteur se doivent d'apprécier la qualité des données, d'identifier les biais susceptibles de les affecter, d'en assurer la préservation et de veiller à leur partage et réutilisation appropriée, tout en maîtrisant les cadres normatifs, légaux et éthiques applicables aux différents contextes nationaux et institutionnels. Enseigner la gestion des données, c'est préparer les étudiant-es à une future pratique professionnelle responsable, éthique et répondant aux réglementations en vigueur.

Quels sont les objectifs d'apprentissage visés ?

Les objectifs d'apprentissage doivent être adaptés au niveau des apprenant-es et alignés au format d'enseignement et à l'évaluation des apprentissages. Voici, pour chaque élément du cycle de vie des données, des exemples d'objectifs d'apprentissage à décliner selon les besoins.

Planifier	<ul style="list-style-type: none"> Comprendre l'importance de la planification rigoureuse pour assurer une gestion efficace des données Concevoir une recherche en intégrant un plan de gestion des données clair et adapté Prévoir et formaliser les modalités de consentement Élaborer des protocoles et des modèles standardisés pour la collecte et le traitement des données Identifier les services et ressources disponibles à l'UNIGE et à la Faculté de médecine pour soutenir les chercheurs/euses dans leurs démarches tout au long de la recherche Désigner une (des) personne(s) responsable(s) pour les éventuelles décisions à prendre dans le projet en matière de GDR et d'ORD
-----------	---

Générer	<ul style="list-style-type: none"> Maîtriser les méthodes de collecte de données pertinentes selon le contexte de recherche
Collecter	<ul style="list-style-type: none"> Identifier et évaluer les sources de données existantes susceptibles d'être utilisées ou croisées dans le projet mené Enregistrer, avec les données, les informations contextuelles (métadonnées) indispensables à leur interprétation et à leur réutilisation, en respectant les standards Acquérir et intégrer des données provenant de tiers en respectant les conditions d'utilisation et de confidentialité
Backup	<ul style="list-style-type: none"> Mettre en place une stratégie de sauvegarde (sauvegardes automatiques, copies multiples, stockage sur différents supports ou emplacements) pour prévenir toute perte ou altération des données pendant leur traitement et leur analyse Vérifier périodiquement l'intégrité et la restauration possible des fichiers sauvegardés Documenter la politique de sauvegarde afin d'en assurer la cohérence et la pérennité au sein de l'équipe ou du projet
Traiter	<ul style="list-style-type: none"> Appréhender les outils de traitement de données pour saisir, numériser, transcrire et traduire les données et garantir leur accessibilité et leur intégrité Appliquer des procédures de vérification, validation, nettoyage et anonymisation, pseudonymisation ou chiffrement des données Dériver de nouvelles données à partir des données brutes (transformations, calculs, catégorisation) Documenter précisément les données et leurs transformations pour assurer la traçabilité de leur traitement
Analyser	<ul style="list-style-type: none"> Analyser et interpréter les données de manière rigoureuse pour en extraire des conclusions fiables Produire des résultats de recherche exploitables et citer correctement les sources de données utilisées Documenter les procédés d'analyse pour la bonne traçabilité
Préserver	<ul style="list-style-type: none"> Connaître les principes et les solutions de stockage à long terme des données (format, migration, contrôle de l'intégrité) Appliquer des mesures de sécurité physiques et numériques adaptées à la sensibilité des données (droits d'accès) Appliquer les principes FAIR (Findable, Accessible, Interoperable, Reusable) pour optimiser la visibilité, l'accessibilité, l'interopérabilité et la réutilisation des données Documenter les stratégies de préservation afin d'assurer la continuité en cas de transfert ou de changement d'équipe
Publier Partager	<ul style="list-style-type: none"> Identifier les niveaux et modalités d'accès possibles (ouvert, restreint, sous conditions) Connaître les obligations légales et éthiques liées à la diffusion des données et au droit d'auteur Rédiger des conditions d'utilisation claires et conformes aux licences choisies ou au consentement obtenu des personnes participantes à la recherche Préparer un jeu de données et ses métadonnées pour un dépôt dans une plateforme ou un dépôt institutionnel
Réutiliser	<ul style="list-style-type: none"> Identifier des jeux de données existants répondant à une nouvelle question de recherche Évaluer la pertinence et la qualité des données pour un nouvel usage scientifique ou clinique Interpréter correctement les métadonnées et la documentation associée pour comprendre le contexte de production Fournir des métadonnées et la documentation nécessaire à la bonne réutilisation des données Respecter les droits d'auteur, licences et accords de partage lors de la réutilisation

	<ul style="list-style-type: none"> • Réfléchir aux modalités d'utilisation des jeux des données produits (licences, copyright, etc.) • Citer et référencer correctement les données sources afin de garantir la traçabilité scientifique • Indiquer clairement comment les jeux de données doivent être cités /mentionnés • Adapter et transformer les données réutilisées de manière transparente, en documentant toute modification
--	---

Pour décliner ces objectifs d'apprentissage et les adapter au niveau de votre public, plusieurs ressources sont disponibles :

- Le [référentiel de compétences de « Digital skills »](#)
- Le site [Research Data](#)
- Le service de Data Stewards de la Faculté de Médecine
- Les [Core Facilities de la Faculté de Médecine](#)
- Le Committee board du projet Data

Comment l'enseigner de manière stimulante et évaluer les apprentissages ?

Les formats d'enseignement actifs favorisent une meilleure intégration des connaissances. L'évaluation des apprentissages doit être cohérente avec les objectifs fixés et le type d'enseignement proposé, afin que les apprenant-es puissent s'exercer aux compétences qui seront évaluées. L'évaluation peut être formative (sans notation) ou sommative (avec notation), et doit autant que possible inclure un retour personnalisé (feedback) pour soutenir le processus d'apprentissage.

Voici quelques exemples de modalités d'enseignement et d'évaluation alignées avec des objectifs d'apprentissage :

Objectif	Format	Evaluation
Connaître les principes FAIR (Findable, Accessible, Interoperable, Reusable) permettant d'optimiser la visibilité, l'accessibilité et la réutilisation des données	Cours <i>ex cathedra</i>	Quizz en ligne ou sur table sur les principes FAIR.
Concevoir une recherche en intégrant un plan de gestion des données clair et adapté	Mini-projet en groupe : construire un mini plan de gestion des données pour un projet imposé ou un réel projet.	Présentation orale avec Q&A et feedback des pairs.
Enregistrer, avec les données, les informations contextuelles (métadonnées) indispensables	Cas pratique : élaborer une fiche de métadonnées ou un fichier README.	Soumission de la fiche de métadonnées.
Préparer un jeu de données et ses métadonnées pour un dépôt dans une plateforme ou un dépôt institutionnel	Cas pratique : créer un guide utilisateur pour un dépôt de données, déposer des données dans un bac à sable institutionnel.	Évaluation par les pairs : critique du guide ou du dépôt (données et métadonnées).
Documenter les stratégies de préservation afin d'assurer la continuité en cas de transfert ou de changement d'équipe	Cas pratique : remplir en groupe un diagramme de flux des données de recherche (DFDR) sur la base d'un scénario de recherche.	Etablir un DFDR sur la base d'un autre scenario de recherche et argumenter les choix (travail individuel).

Appliquer des procédures de vérification, validation, nettoyage et anonymisation des données	Exercice technique de nettoyage et anonymisation d'un jeu de données et documentation de la procédure.	Vérifier par l'enseignant-e de la qualité du travail de curation (nettoyage et anonymisation), et notamment qu'il n'y ait pas de recouplements possibles.
Documenter précisément les données et leurs transformations pour assurer leur traçabilité	Lecture critique d'un article scientifique sous l'angle de la transparence des données.	Intégration d'une section « gestion des données » dans un travail écrit accompagnant un projet de recherche et évalué par une grille critériée.
Identifier et évaluer les sources de données existantes susceptibles d'être utilisées	Atelier d'évaluation d'un jeu de données fictif avec erreurs injectées.	Evaluation de la capacité à détecter et justifier la qualité et la pertinence des données lors d'un exercice équivalent à celui effectué en classe.
Connaître les obligations légales et éthiques liées à la diffusion des données et au droit d'auteur	Discussion éthique sur un cas réel de fuite de données médicales et sur les moyens de l'éviter. Simulation (jeu de rôle) d'un comité d'éthique évaluant un protocole.	Rédaction réflexive sur les dilemmes rencontrés.
Prévoir et formaliser les modalités de consentement	Analyser plusieurs exemples de formulaires de consentement et rédiger un modèle adapté à une étude fictive, en identifiant les éléments essentiels à la conformité éthique et légale.	À partir d'un scénario de recherche, rédiger les conditions d'utilisation d'un jeu de données en veillant à respecter les engagements pris dans le formulaire de consentement et les contraintes légales.

Et après ?

La gestion des données de recherche et de santé est un domaine en constante évolution, tant sur le plan technologique que réglementaire. Pour garantir la pérennité des compétences acquises, il est essentiel de promouvoir une formation continue intégrée aux pratiques professionnelles. Les dispositifs de mentorat – qu'ils prennent la forme d'un accompagnement par des pairs expérimentés, de communautés de pratique ou de formations tutorées – renforcent l'apprentissage et soutiennent la mise en œuvre concrète des bonnes pratiques. Cette approche favorise une culture partagée de la qualité, de la transparence et de la responsabilité dans la gestion des données au sein de la Faculté.

Pour aller plus loin :

Trouvez des ressources pour vos formations ainsi que les formations déjà existantes sur le site institutionnel ResearchData : <https://www.unige.ch/researchdata/en/home/>