

Pratique responsable en statistique et science des données

Individuellement

- S'assurer d'avoir les **connaissances et compétences nécessaires** pour nos tâches ; se **former en continu** sur les meilleures pratiques.
- Se baser sur les **codes de déontologie** pertinents, notamment ceux de la SSC et de l'ASA.

1. Enjeux éthique : Conception de l'étude

Pertinence Pourquoi ce projet est-il mené ? Poursuit-il un objectif légitime et clairement défini ?

Équité et inclusion Qui est inclus ou exclu ? Les critères d'inclusion risquent-ils d'introduire des biais ?

Transparence Les éléments clés du projet sont-ils clairement documentés ? Y a-t-il des conflits d'intérêts ou des apparences de conflits d'intérêts ?

Intégrité A-t-on obtenu les approbations éthiques nécessaires pour le projet ?

2. Enjeux éthique : Collecte des données

Consentement Les participants ont-ils donné un consentement libre, éclairé et réversible ?

Proportionnalité Les données recueillies sont-elles limitées à ce qui est nécessaire aux objectifs du projet ?

Équité Certains groupes sont-ils sous-représentés ou exclus de la collecte, notamment à cause des exigences de participation (langue, durée, moyens techniques) ?

Confidentialité Les données sont-elles protégées dès la collecte et leur traitement respecte-t-il les lois sur la protection des renseignements personnels (ex. Loi 25) ?

3. Enjeux éthique : Gestion et stockage

Sécurité Les accès, transferts et sauvegardes sont-ils protégés par des mécanismes de sécurité adéquats ?

Qualité Les données sont-elles complètes, cohérentes et suffisamment documentées pour être réutilisées ?

Traçabilité Peut-on retracer l'origine des données et leurs mises à jour depuis leur collecte ? Peut-on identifier le propriétaire légitime (*data ownership*) ?

Conformité La gestion des données (conservation, suppression, protection) est-elle conforme aux obligations légales et institutionnelles ?

4. Enjeux éthique : Analyse et modélisation

Rigueur Les méthodes utilisées sont-elles clairement justifiées ? Les résultats obtenus sont-ils reproductibles ? Demeurent-ils robustes face à des hypothèses ou méthodes d'analyse alternatives ?

Transparence Le code, les méthodes et les sources externes sont-ils clairement documentés et attribués, dans le respect de la propriété intellectuelle ?

Équité Les modèles introduisent-ils des biais ? Risquent-ils d'entraîner de la discrimination ? D'amplifier les inégalités déjà présentes ?

Confidentialité Les méthodes d'analyse limitent-elles les risques de divulgation ou de réidentification ?

Explicabilité Les résultats ou prédictions peuvent-ils être expliqués de façon claire et compréhensible ? A-t-on évalué si un modèle plus simple ou plus interprétable serait suffisant ?

Indépendance Quelqu'un cherche-t-il à orienter les analyses pour servir un intérêt particulier ?

Responsabilité Les risques d'erreur d'utilisation ou de mésusage du modèle ont-ils été identifiés et évalués ? Le modèle pourrait-il générer des boucles de renforcement, en influençant les comportements ou données ultérieures d'une façon qui renforce ses propres biais ?

5. Enjeux éthique : Communication des résultats

Exactitude et intégrité Les résultats sont-ils présentés sans exagération ni simplification trompeuse ? Les incertitudes, marges d'erreur et limites méthodologiques sont-elles clairement communiquées ?

Clarté Les tableaux, graphiques et indicateurs peuvent-ils être mal interprétés ? Les explications sont-elles compréhensibles et adaptées au public visé ?

Confidentialité Les données sensibles sont-elles présentées de manière agrégée ou anonymisée ?

Responsabilité Les résultats risquent-ils d'être interprétés hors contexte ou d'influencer indûment des décisions ?

Traçabilité : Les sources, codes et conditions de production des résultats sont-ils décrits de façon transparente ?

6. Enjeux éthique : Diffusion et utilisation

Conformité Le partage des données, codes et/ou résultats respecte-t-il les accords de confidentialité et les licences d'usage ? Les conditions de réutilisation sont-elles explicites ?

Confidentialité Les données diffusées peuvent-elles être croisées avec d'autres sources de façon risquée ?

Durabilité Les formats et plateformes utilisés favorisent-ils une conservation pérenne ?

Intégrité La documentation est-elle suffisante pour permettre une réutilisation utile ? Les limites d'usages sont-elles clairement documentées ?

Responsabilité Un mécanisme de suivi permet-il de détecter les effets inattendus ou disproportionnés de l'utilisation des données, modèles ou résultats ? Les données pourraient-elles être combinées à d'autres sources pour surveiller ou cibler des personnes ?

A. Impacts sociaux à considérer

- Le projet risque-t-il de renforcer des stéréotypes, des préjugés ou des inégalités structurelles ?
- Des mécanismes existent-ils pour signaler, corriger ou compenser les effets sociaux négatifs ?
- Le projet peut-il miner la confiance du public envers les institutions ou la science ?
- Les coûts sont-ils soutenables à long terme ?
- Les bénéfices sont-ils partagés équitablement ?
- Le projet contribue-t-il au bien commun, à la justice sociale ou à l'inclusion ?

B. Impacts environnementaux à considérer

- La consommation énergétique est-elle proportionnée aux bénéfices attendus du projet ?
- Les choix méthodologiques privilégient-ils des options moins énergivores lorsqu'elles sont suffisantes ?
- L'utilisation des modèles ou des résultats pourrait-elle avoir des effets négatifs sur l'environnement ?
- Le projet contribue-t-il à des objectifs environnementaux plus larges ?