



LIVRET DU CITOYEN

CHARTRE ETHIQUE DES USAGES DES DONNÉES



devient

ekitia

Data et IA en pleine confiance

OCCITANIE DATA

LA CHARTE

PRINCIPE 1 *SCIENCE DES DONNÉES ET SOCIÉTÉ*

1.1. BIENFAISANCE

Les projets contribuent à la réalisation des 17 objectifs de développement durable définis par les Nations Unies.

1.2. INNOVATION SOUTENABLE

Les projets sont réalisés dans des conditions respectueuses de l'environnement.

1.3. SOLIDARITÉ ET DIVERSITÉ

Les projets respectent la diversité et le pluralisme de la société : ils n'auront pas pour effet de créer ou de renforcer des inégalités économiques ou sociales, ni entre les citoyens, ni entre les petits et les grands acteurs du marché.

PRINCIPE 2 *SCIENCE DES DONNÉES ET INDIVIDU*

2.1. AUTONOMIE INDIVIDUELLE

Les personnes dont les données sont utilisées pour réaliser un projet sont toujours dûment informées de cette utilisation, de leurs droits sur ces données et de la finalité du projet.

2.2. PROTECTION DES DONNÉES PERSONNELLES ET DE LA VIE PRIVÉE

Afin que la vie privée des personnes soit protégée de manière optimale, elle est prise en considération dès les étapes de recherche et de conception des projets. Cela inclut l'application, lorsque cela est pertinent de techniques de pseudonymisation ou d'anonymisation des données personnelles.

PRINCIPE 3 QUALITÉ DES DONNÉES ET SÉCURITÉ DES SYSTÈMES D'INFORMATION

3.1. QUALITÉ DES DONNÉES

Les données utilisées pour réaliser des projets sont exactes, dépourvues de biais, interopérables et sont pertinentes par rapport à l'utilisation envisagée.

3.2. SÉCURITÉ DES CENTRES DE DONNÉES

Les centres de données choisis pour héberger les données, et notamment les données les plus sensibles, respectent les meilleures normes de sécurité.

3.3. ROBUSTESSE DES ALGORITHMES

Les résultats sortants d'une analyse des données sont fiables à un haut degré de certitude. Dans tous les cas, la marge d'erreur affectant le résultat est indiquée aux utilisateurs finaux.

PRINCIPE 4 TRANSPARENCE

4.1. INFORMATION CLAIRE ET ACCESSIBLE

Les utilisateurs finaux ont la possibilité de se renseigner facilement

(par exemple, à l'aide d'une notice d'information) à propos des caractéristiques de la solution développée : la finalité, les avantages et les risques d'utilisation, les données concernées, leur lieu de stockage et la durée de leur conservation.

4.2. EXPLICABILITÉ DES ALGORITHMES

La logique des méthodes utilisées pour analyser des données est documentée afin de pouvoir expliquer le résultat obtenu de manière intelligible aux utilisateurs finaux ; cela implique par exemple de mettre en évidence les variables déterminantes des calculs algorithmiques. Il peut toutefois arriver que la logique de la méthode d'analyse soit trop complexe pour être expliquée de manière intelligible. Dans ce cas, les porteurs de projet détaillent le choix stratégique et les avantages que représentent cette méthode.

4.3. AUDITABILITÉ DES PROJETS

Chaque élément essentiel à la réalisation d'un projet est bien documenté en vue de fournir des informations à toute personne intéressée, ou en vue à servir de base à un contrôle. Sont par exemple concernés les données utilisées, les systèmes algorithmiques utilisés pour les analyser et les résultats des évaluations des risques.

PRINCIPE 5 *GOVERNANCE DES DONNÉES DANS UN CADRE DE CONFIANCE*

5.1. DÉCLINAISON DE LA CHARTE EN RÈGLE

La Charte est détaillée en des critères plus précis permettant sa mise en œuvre concrète.

5.2. APPRENTISSAGE COLLECTIF

Chaque année, les Signataires de la Charte font remonter à Occitanie Data les difficultés qu'ils rencontrent concernant la mise en œuvre de certains principes. Cela permet d'identifier des problématiques communes et d'ouvrir des possibilités de collaboration pour les résoudre.

5.3. ÉVALUATION DES RISQUES

Dans la mesure du possible, les porteurs de projet évaluent les risques individuels, sociétaux et environnementaux que leur projet est susceptible de générer. À partir des résultats de ces évaluations, ils mettent en balance les bénéfices (y compris sociaux et environnementaux) et les risques et ne lancent le projet que lorsque les avantages sont plus nombreux que les risques.

5.4. INCLUSION DES CITOYENS

Lorsque cela est opportun, notamment en présence de projets visant à mettre en œuvre des politiques publiques, les porteurs de projet offrent aux individus (en leur qualité de citoyen ou d'utilisateur) la possibilité d'exprimer leur point de vue et leurs retours d'expérience. Ces retours sont pris en considération pour définir les axes d'évolution du projet.

5.5. PRINCIPE D'INTÉGRITÉ

Toute personne intervenant dans un projet respecte les règles déontologiques auxquelles elle est soumise et agit dans un esprit d'intégrité intellectuelle et de coopération.

PRINCIPE **6** *RÉCIPROCITÉ*

6.1. RECONNAISSANCE

Lorsqu'un projet implique la collaboration de plusieurs acteurs, cette collaboration est explicitement reconnue et rendue publique.

6.2. RÉPARTITION ÉQUITABLE DE LA CRÉATION DE VALEUR

La valeur créée grâce à un projet est équitablement répartie entre tous les acteurs ayant contribué à sa réalisation.

PRINCIPE **7** *ÉTHIQUE ET URGENCE*

En temps de crise sanitaire, environnementale ou sécuritaire, la Charte constitue un outil de référence, et notamment les principes liés au respect de la vie privée et à la transparence. Afin de répondre efficacement à des problématiques urgentes, les autres principes pourront être aménagés.

PRINCIPE **8** *RÉEXAMEN RÉGULIER DE LA CHARTE*

La Charte est un instrument dynamique : elle sera réévaluée régulièrement en vue d'être adaptée en continu au progrès technologique et à l'évolution du cadre légal applicable aux activités numériques.

LEXIQUE TECHNOLOGIQUE

Donnée (numérique)

Toute information pouvant être enregistrée en format numérique.

Données à caractère personnel

Il s'agit de données dont le traitement est encadré par la loi pour garantir aux personnes une protection adéquate de leur vie privée (que ces données reflètent). Précisément, il s'agit de toute information permettant d'identifier une personne physique directement (nom, prénom, photo, empreinte digitale ...) ou indirectement (n° de sécurité sociale, adresse IP ...).

Données sensibles

Il s'agit d'une catégorie de données personnelles faisant l'objet d'une protection spécifique : leur traitement est interdit par principe. Précisément, il s'agit de toute information relative à l'origine raciale ou ethnique, aux opinions politiques, aux convictions philosophiques ou religieuses, à l'adhésion syndicale, à la génétique, à la biométrie, à la santé, à l'orientation ou à la vie sexuelle d'une personne.

Données non personnelles

On en distingue deux sortes, d'une part les données qui sont par nature non rattachables à une personne physique et, d'autre part, les données personnelles qui ont fait l'objet d'une anonymisation. Ces données peuvent librement circuler en Union Européenne (sous réserve que cette circulation ne porte pas atteinte à l'intérêt public d'un État membre)².

²Règlement (UE) 2018/1807 du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne.



Lots de données

Il s'agit d'un ensemble de données formant un « tout » cohérent. Les lots de données sont souvent présentés sous une forme de tableau (les colonnes contiennent les variables (ex. âge) et les lignes les valeurs de ces variables (ex. 18 ans)).

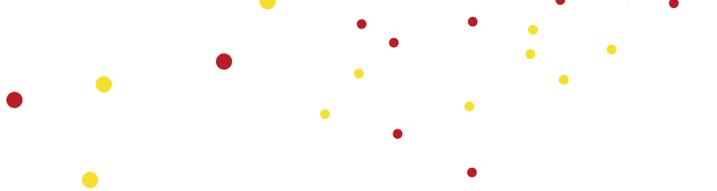
Cycle de vie d'une donnée

- **Collecte de données** : Tout processus dont l'objectif est de rassembler des données.
- **Stockage de données** : Les données sont stockées dans des centres de données permettant de sécuriser leur accès. La plupart de ces centres sont gérés par des prestataires externes.
- **Traitement de données** : Toute opération permettant d'extraire de l'information à partir des données collectées, quel que soit le procédé utilisé. Il peut notamment s'agir d'analyser des données. On distingue généralement trois types d'analyses :
 - Analyse descriptive : Elle consiste simplement à décrire le contenu d'un ensemble de données.
 - Analyse prédictive : Elle consiste à prévoir, à l'aide d'une mesure de probabilité, ce qui pourrait survenir dans le futur. Dans le domaine du Big Data, les systèmes d'apprentissage automatique (cf. ci-dessous) sont très utilisés pour faire des prédictions de ce type.
 - Analyse prescriptive : Elle reprend le principe de l'analyse prédictive en y ajoutant la notion d'action. C'est-à-dire que l'analyse va permettre d'identifier les cibles sur lesquelles agir.

• **Archivage de données** : L'archivage consiste à exporter les données inutilisées du lieu où elles sont stockées vers un autre support sécurisé afin de libérer de l'espace de stockage.

• **Partage de données** : Le partage de données consiste à fournir à des tiers intéressés un accès à ses propres données.





Big Data

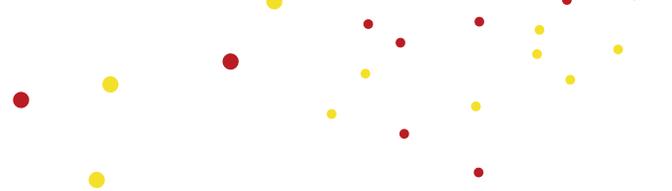
Littéralement ces termes renvoient à l'expression de « données massives ». En pratique, le Big Data correspond à la production, par différentes sources, de grands volumes de données. Si ces grands volumes de données permettent de faire des analyses à une échelle nettement supérieure à celle que permettent des lots de données plus petits, la contrepartie est que leur traitement nécessite d'importants moyens techniques (ce qui provoque le retard, voire l'exclusion, des petites structures sur le marché du Big Data).

Algorithme

Un algorithme est une suite d'instructions en nombre fini permettant d'obtenir un résultat souhaité (l'exemple le plus basique étant celui de la recette de cuisine). Afin de pouvoir être exécutée par un logiciel, cette suite d'instructions est traduite en langage informatique (le langage de programmation). Lorsque les instructions ainsi programmées visent à reproduire une fonction cognitive humaine, on dit du logiciel qu'il est un « système d'intelligence artificielle ». On en distingue principalement deux catégories :

Systèmes experts (déterministes)

Ils résolvent des problèmes en s'appuyant sur une base de connaissances relatives à un domaine spécifique. Ces connaissances correspondent en fait à toutes les informations relatives à un sujet donné, et sur lesquelles, traditionnellement, un expert humain se fonde pour établir son expertise. Cet ensemble de règles forme le modèle sur la base duquel, à la manière d'un expert du domaine traité, le système peut raisonner.



Systèmes d'apprentissage automatique (probabilistes)

Ils résolvent des problèmes de manière plus ou moins autonome, en s'appuyant sur un modèle qu'ils auto-génèrent en analysant des données existantes. En d'autres termes, il s'agit de systèmes qui « apprennent » un modèle de raisonnement à partir de données existantes : une phase d'apprentissage est nécessaire, afin que le système affine et optimise son modèle de raisonnement, au fur et à mesure qu'il est confronté à de nouveaux cas.

Intelligence Artificielle (IA)

Elle sera ici appréhendée en tant que composante des systèmes algorithmiques permettant d'analyser les données.

Science des données

Il s'agit d'une évolution de la démarche statistique classique, qui consistait à se poser une question, à formuler une hypothèse, à planifier une expérience, à recueillir des données et à estimer les paramètres d'un modèle pour *in fine* prendre une décision. Avec la science des données les hypothèses sont directement formulées à partir des données, et non plus théoriquement, ce qui constitue un changement de paradigme majeur.

Économie des données

Les données numériques que nous émettons tous lors de nos activités quotidiennes créent de nouvelles opportunités économiques liées à leur exploitation. Toutefois, dans le cadre d'Occitanie Data les usages ne se limitent pas à l'économie au sens strict : ils peuvent aussi concerner des innovations sociales, par exemple.

EN SAVOIR PLUS

Sur le développement durable

Consulter les 17 Objectifs de développement durable de l'Organisation des Nations Unies

Sur la régulation éthique de l'intelligence artificielle

Consulter la Déclaration de Montréal pour une Intelligence Artificielle responsable (et pour approfondir, les Lignes directrices du Groupe d'experts de haut-niveau sur l'Intelligence Artificielle désigné par la Commission européenne)

Sur la protection des données personnelles

Consulter le site internet de la Commission Nationale d'Informatique et des Libertés (<https://www.cnil.fr>)

Sur les impacts sociétaux et environnementaux des nouvelles technologies

Consulter le site internet du Think Tank de l'Institut du Numérique Responsable : <https://institutnr.org/>

Sur la méthode du design fiction

Julian Bleeker, A short essay on design, science, fact and fiction, mars 2009



À l'issue de cet atelier, vous aurez toujours la possibilité de nous faire parvenir vos réflexions éthiques concernant les usages des données via un formulaire à remplir en ligne :

<https://occitaniedata.fr/consultez-la-1ere-version-de-la-charte-ethique-du-big-data-elaboree-par-occitanie-data/>



ekitia

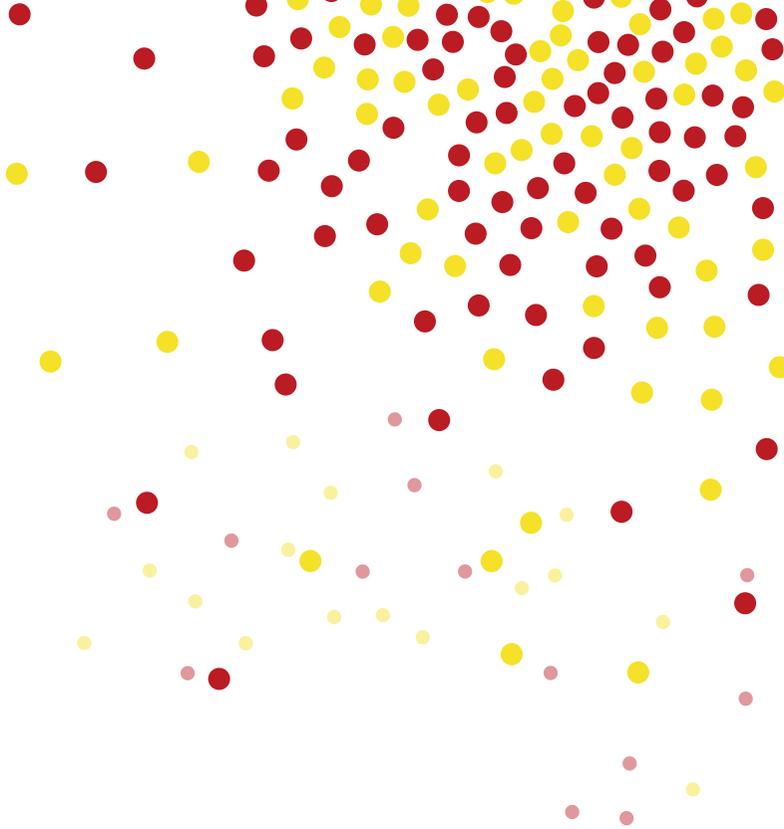
Data et IA en pleine confiance

Occitanie Data est une association de préfiguration d'un pôle d'économie de la donnée, en cours de transformation en un GIP national. Structurée en Occitanie mais rassemblant des acteurs au-delà, ses membres sont des entreprises privées et publiques, des collectivités, des acteurs académiques, des clusters et pôles de compétitivité. Occitanie Data propose un **cadre de confiance, éthique et souverain**, destiné à permettre aux acteurs de partager et de croiser leurs données tout en respectant les intérêts des individus et des propriétaires des données.

Pour construire ce cadre, une première version de la **Charte éthique des usages des données** a été publiée en avril 2020. Il ne s'agit toutefois pas d'une version définitive : elle sera enrichie par le fruit d'ateliers de co-construction que nous avons initié, en collaboration avec la **Chaire UNESCO «Ethique, Sciences et Société»**, au cours de l'été 2020.

En effet, après avoir réalisé 6 ateliers avec des experts métier sectoriels (mobilité, santé, énergie, environnement, agriculture, emploi et formation) et pris en considération leurs retours, nous ambitionnons d'organiser des ateliers de co-construction citoyenne, afin de confronter les principes de la Charte à des regards neutres.

L'atelier auquel vous participez aujourd'hui représente donc pour nous une réelle opportunité pour recueillir des premiers retours citoyens sur le contenu de la Charte.



Pour toute information complémentaire ou pour nous
contacter, retrouvez-nous sur notre site internet :

occitaniedata.fr

  @OccitanieData