

# Projet de guide de bonnes pratiques en matière de traitement des données géolocalisées

Élaboré de manière collaborative par les participant(e)s au Groupe de travail « Éthique de l'usage de la donnée et de l'information géolocalisée », rattaché à la Commission Règles et Qualité du Conseil National de l'Information Géolocalisée (CNIG)

## TABLE DES MATIERES

Préambule .....	3
Objectif et méthode d'élaboration du livrable .....	3
Traiter les données avec éthique n'est pas une démarche nouvelle .....	3
<i>Définition de l'éthique</i> .....	3
<i>La démarche éthique</i> .....	4
<i>Historique de l'éthique appliquée</i> .....	4
Lien entre l'éthique et le droit .....	6
<i>Le cadre juridique applicable aux données géolocalisées</i> .....	7
Chapitre 1 – Introduction : pourquoi un tel GT ?.....	9
Le mandat du GT.....	9
Différence entre donnée et information géolocalisée .....	9
Chapitre 2 – Panel de cas d'usage étudiés pendant le GT .....	11
1 <sup>ère</sup> thématique : L'ouverture au grand public d'informations géolocalisées .....	11
2 <sup>ème</sup> thématique : L'utilisation de données et d'informations géolocalisées .....	15
3 <sup>ème</sup> thématique : La mise en commun de données géolocalisées .....	17
Chapitre 3 – Des bonnes pratiques en matière d'éthique de l'usage de la donnée/information géolocalisée .....	20
Introduction.....	20
Bonnes pratiques organisées selon le cycle de vie d'une donnée.....	20
<i>Bonnes pratiques en matière de collecte/production</i> : .....	20
<i>Bonnes pratiques pour le choix des infrastructures d'hébergement et d'analyse des données</i> .....	21
<i>Bonnes pratiques en matière de mise à disposition à des tiers</i> .....	21
<i>Bonnes pratiques en matière de réutilisation</i> .....	22
<i>Bonnes pratiques transversales</i> .....	22
ANNEXE : Points communs à différentes chartes d'éthique.....	24
REMERCIEMENTS AUX PARTICIPANT.E.S AU GROUPE DE TRAVAIL .....	25

# Préambule

## Objectif et méthode d'élaboration du livrable

Ce livrable propose les principales bonnes pratiques de traitement de données géolocalisées, qui ont été identifiées dans le cadre du Groupe de Travail du CNIG dédié à l'éthique des usages des données et des informations géolocalisées (GT). L'objectif est de promouvoir au travers de l'application de ces bonnes pratiques, une approche éthique, dès la conception et tout au long du cycle de vie des projets impliquant des données géolocalisées. Il s'adresse ainsi à toute la diversité d'acteurs qui interviennent dans le cycle de vie des données, c'est-à-dire autant à ceux qui la produisent, qu'à ceux qui en assurent la gestion et à ceux qui les utilisent.

Il s'agit d'un document élaboré de manière collaborative par les participant(e)s au GT. Instance consultative placée auprès du ministre en charge du développement durable, le CNIG rassemble une grande variété d'acteurs qui composent l'écosystème de la géo-donnée en France : ministères, établissements publics, collectivités territoriales, entreprises privées, associations professionnelles, organisations syndicales, association de citoyens, qui peuvent se rencontrer, décider et coproduire, notamment dans le cadre de commissions et de groupes de travail tel que celui à l'origine du présent livrable.

Ce document a vocation à évoluer, de sorte que toute suggestion future visant à enrichir son contenu provenant de la part de l'un des acteurs concernés, soit la bienvenue.

## Traiter les données avec éthique n'est pas une démarche nouvelle

### *Définition de l'éthique*

Selon Axel Kahn, « l'éthique est une boussole »<sup>1</sup> qui permet d'identifier les questions à se poser et de décider du meilleur chemin à suivre.

Plus précisément, elle est une discipline qui vise à réfléchir sur les comportements humainement acceptables par la société à un moment donné de son histoire. « *L'éthique possède un caractère plus dynamique et plus réflexif que la morale [...] Elle est une démarche. Elle a pour but de réinterroger les principes moraux et les règles déontologiques et juridiques, en particulier quand ils ne permettent pas des réponses conformes aux souhaits des personnes concernées ou que la situation qui se présente n'y trouve pas de solution. L'éthique procède ainsi d'une réflexion active, collective, interactive sur les valeurs humaines et les tensions entre des volontés et d'autres logiques. Elle étudie les valeurs, les tensions et les critères de choix et les différents scénarii possibles [...]. À un niveau collectif, elle participe aux questionnements sur les choix de société et contribue à l'évolution des normes et règles qui nous régissent. C'est donc un processus dynamique d'interrogations et de questionnements dont la visée est une aide à la décision.* »<sup>2</sup>

<sup>1</sup> KAHN Axel, *L'éthique est une boussole - En chemin vers une vie bonne*, Ed. Marabout, 2022.

<sup>2</sup> AMELINE Sylvie, LEVANNIER Muriel, Chapitre 11 - La démarche éthique, In : *DEAS Tout-en-un*. Ed. Vuibert AS/AP, p.524-527.

## La démarche éthique

En ce qui concerne la méthode permettant de parvenir aux décisions les plus justes, il existe aujourd'hui deux courants majeurs qui sont : l'éthique déontologique (fondée sur des principes considérés comme intrinsèquement moraux, à respecter en toutes circonstances) et l'éthique conséquentialiste (fondée sur la situation la plus souhaitable, ce qui peut justifier de hiérarchiser des principes au cas par cas).

En France, le Comité National Pilote d'Éthique pour le Numérique<sup>3</sup> semble se positionner plutôt en faveur de cette seconde approche, estimant que « *même si chacun des principes est désirable, toute situation concrète peut faire émerger des tensions entre eux [par exemple, entre innovation soutenable et explicabilité, entre bienfaisance et protection de la vie privée, ...].* » Il précise que « *la prise de décision au cas par cas est alors nécessaire en tenant compte des contextes de conception et d'usage, dans le respect du principe de proportionnalité et des droits fondamentaux [et que dans] chaque cas concret, la délibération doit s'appuyer autant sur les finalités recherchées que sur les contraintes techniques, ainsi que sur la prise en compte des intérêts des utilisateurs à court ou à long termes.* »<sup>4</sup> Cette définition sera donc prise comme référence dans le présent livrable.

Comme le présente la Chaire Hoover d'éthique économique et sociale de l'Université de Louvain<sup>5</sup>, l'éthique peut se décliner en au moins trois champs distincts :

- La *méta-éthique*, qui désigne la réflexion philosophique sur l'origine et la portée des jugements moraux.
- L'*éthique normative*, qui désigne la réflexion sur le contenu de nos obligations morales, qu'elles soient individuelles, collectives ou sociales.
- L'*éthique appliquée*, « qui est un prolongement de l'éthique normative dans l'étude de cas très concrets, [...]. Il s'agit ici moins de justifier une perspective normative que de l'appliquer afin de résoudre un dilemme ou de recommander un certain type d'action. »

Il est à noter que les cas concrets soumis à la démarche éthique peuvent être issus de n'importe quelle discipline, telle que la politique, l'économie, l'organisation du travail, les sciences du vivant, l'éducation, ou encore la communication. Dans le présent livrable, les cas qui ont fait l'objet de discussions étaient par exemple, des projets impliquant des traitements de données géolocalisées.

## Historique de l'éthique appliquée

---

<sup>3</sup> Placé sous l'égide du Comité Consultatif National d'Éthique (CCNE), le Comité National Pilote d'Éthique du Numérique (CNPEN) a été mis en place en décembre 2019 à la demande du Premier ministre. Il a pour rôle de rendre des avis sur les saisines qui lui sont confiées, d'éclairer le débat public sur les enjeux d'éthique du numérique et de rédiger des propositions relatives à la pérennisation de la réflexion nationale sur l'éthique du numérique. Le rôle de ce comité est donc d'éclairer les décisions tant individuelles que collectives en veillant à la sensibilisation et à l'information de la société civile, des institutions publiques ou privées et du gouvernement.

<sup>4</sup> Avis n°3 du CNPEN, « Agents conversationnels : enjeux d'éthique », adopté le 15 septembre 2021.

<sup>5</sup> <https://www.enseignerlethique.be/concepts-de-base/#ethique-et-morale>

« L'éthique appliquée est née dans le contexte du changement entre le moderne et le contemporain, où apparaissent des produits de la science, de la technologie et de la médecine, qui posent de nouveaux problèmes pour le traitement desquels il faut une éthique qui ne soit pas seulement principielle et théorique, mais une éthique capable d'« appliquer » sur ces objets des valeurs, valeurs humaines, principalement la générosité et la convivialité, valeurs de la vie et du vivant, valeurs environnementales de la conservation et de la diversité, valeurs épistémiques garantissant les formes d'honnêteté qui accompagnent la démarche scientifique. L'éthique appliquée s'allie aux pratiques et se différencie aussi en fonction de ses objets (technologie, environnement, affaires, etc.). »<sup>6</sup>

Au cours du siècle dernier, la démarche éthique s'est beaucoup développée dans le contexte de l'après Seconde Guerre Mondiale afin de questionner l'avancement et l'application des technosciences biomédicales. Cela a conduit à l'émergence des 4 grands principes de bioéthique<sup>7</sup> :

- Le principe d'autonomie, reconnaissant la capacité d'une personne à faire des choix personnels.
- Le principe de bienfaisance, visant à ce que toute action ait pour finalité le bien.
- Le principe de non-malfaisance, son corollaire, visant à limiter au maximum le dommage potentiel d'une action.
- Le principe de justice, promouvant équité et impartialité dans le traitement des personnes.

Concernant les nouvelles technologies, des valeurs éthiques universelles visant à aider les acteurs de la société à faire des choix et orienter les développements scientifiques et technologiques de manière respectueuse par rapport à l'humanité et à l'environnement, ont été définies par l'UNESCO dès 2005. Les principes sont similaires à ceux mentionnés ci-dessus, mais interprétés dans ce nouveau contexte. Concernant les traitements de données, les textes définissant les principes éthiques auxquels se référer en la matière se sont multipliés depuis 2018, essentiellement au travers de l'identification de principes relatifs à la conception, au développement et à l'utilisation de systèmes d'IA. Sans prétendre à l'exhaustivité, ci-dessous se trouvent quelques-unes de ces initiatives :

Au niveau international :

- [Déclaration de Toronto « Protéger le droit à l'égalité et à la non-discrimination dans les systèmes reposant sur l'apprentissage automatique »](#) (2018),
- [Déclaration de Montréal pour une IA responsable](#) (2018),
- [Recommandation 0449 de l'OCDE sur l'IA](#) (2019, mise à jour en 2024),
- [Recommandation sur l'éthique de l'IA de l'UNESCO](#) (2020).

Au niveau européen :

- [Déclaration sur l'IA, la robotique et les systèmes autonomes du Groupe Européen d'Éthique](#) (2018),
- [Lignes directrices pour une IA digne de confiance rédigées par un groupe d'experts de haut-niveau nommé par la Commission européenne](#) (2019),

---

<sup>6</sup> SCHMID Anne-Françoise, Les ambiguïtés de l'éthique appliquée. *Revue française d'éthique appliquée*, 2016/1 n° 1, p.92-106.

<sup>7</sup> La bioéthique peut se définir comme un « ensemble de recherches, de discours et de pratiques, généralement pluridisciplinaires, ayant pour objet de clarifier ou de résoudre des questions à portée éthique suscitées par l'avancement et l'application des technosciences biomédicales » (Gilbert Hottois).

- [Rapport de l'Agence de l'UE pour les droits fondamentaux « Bien préparer l'avenir. L'IA et les droits fondamentaux »](#) (2020),
- [Convention-cadre du Conseil de l'Europe sur l'IA](#) (2024).

Au niveau national (France) :

- [Charte éthique de la donnée 3DOK](#) (2010),
- [Charte éthique et big data](#) (2013),
- [Charte éthique des usages des données d'Ekitia](#) (2019, mise à jour en 2023).

Au niveau local :

- [Charte métropolitaine de la donnée de Nantes](#) (2019),
- [Charte de la donnée de Rueil-Malmaison](#) (2019),
- [Charte métropolitaine de la donnée d'Aix-Marseille-Provence](#) (2022),
- [Charte éthique de la donnée de Brest Métropole](#) (2024).

### *Essai de définition de l'expression « donnée éthique »*

Une donnée éthique pourrait être définie comme une donnée dont le cycle de vie (collecte, pré-traitement – notamment la mise en conformité de données brutes à des règles de validation -, stockage, analyse - au sens d'interprétation - partage, réutilisation, archivage, suppression) et les conditions d'utilisation ont été pensés et définis en cohérence avec des principes éthiques pertinents au regard du contexte de sa production et de son usage, préalablement définis.

Comme présenté ci-dessus, appliquer une démarche éthique à un cas concret, à un projet, peut nécessiter de mettre en balance les principes pertinents pour identifier les exigences qui seront proportionnées aux enjeux spécifiques au contexte. Il s'agit donc d'une démarche à répéter de manière systématique pour chaque nouveau projet. Une telle démarche pouvant être traduite par une méthodologie ou des processus de gouvernance auxquels une organisation pourra se référer, mais la réponse pourra être différente à chaque fois.

## **Lien entre l'éthique et le droit**

Certains des principes consacrés dans les textes mentionnés ci-dessus trouvent écho dans des textes juridiques (par exemple en France la protection de la vie privée, la transparence des algorithmes publics ...). Les deux disciplines se nourrissent mutuellement, mais si l'objectif du droit est de déterminer les règles dans lesquelles une société accepte de vivre ensemble, l'éthique permet quant à elle la remise en question régulière de certaines de ces règles et vise à rechercher la décision la plus juste, au cas par cas.

A ceci, l'on peut ajouter - comme le précise la Chaire Hoover d'éthique économique et sociale de l'Université de Louvain - que « l'éthique [...] se distingue du droit positif en ceci que les normes morales ne sont pas nécessairement incarnées dans des lois et que des lois peuvent très bien être immorales » et, en parallèle, que « le droit positif, quant à lui, évolue en partie en fonction de l'évolution des normes éthiques ou morales au sein de la société, en particulier quand elles sont portées par des combats ou revendications politiques. » Il est également à noter que « tout ce qui paraît souhaitable d'un point de vue éthique ou moral n'a pas

nécessairement vocation à prendre forme de loi. »<sup>8</sup> Enfin, si le non-respect d'une règle de droit entraîne des conséquences juridiques pouvant aller jusqu'à la sanction, ce n'est pas le cas du non-respect d'une règle éthique.

La réflexion menée afin d'élaborer ce livrable se veut en ce sens complémentaire aux normes juridiques existantes applicables aux données géolocalisées et devrait permettre d'interroger certaines d'entre elles.

### *Le cadre juridique applicable aux données géolocalisées*

Au niveau international, si des standards techniques relatifs à la manière de produire ou de diffuser certaines données existent, aucune norme juridique multilatérale de nature à encadrer les pratiques des multiples parties prenantes qui traitent des données n'a encore été adoptée. Sur ce point, force est par ailleurs de constater que certains acteurs économiques dominants sur ce marché des données cherchent à s'affranchir de toute contrainte juridique concernant ce sujet.

Au niveau européen, la situation est différente : la [Convention européenne des droits de l'Homme](#) ainsi que la [Charte des droits fondamentaux de l'Union Européenne](#), deux textes définissant les droits fondamentaux dont bénéficient les citoyens européens, leur confèrent une portée de nature à encadrer les pratiques à la fois des États comme de tout autre acteur en la matière. On y trouve notamment les principes de respect de la vie privée, de protection des données personnelles, de protection de l'environnement, de non-discrimination et de droit à un recours effectif. Tous ces principes ont par ailleurs été précisés et complétés par de nombreux autres textes à valeur juridique (directives et règlements adoptés par les institutions européennes et jurisprudence de la Cour Européenne des Droits de l'Homme et de la Cour de Justice de l'Union Européenne). Pour l'essentiel, ils sont porteurs des obligations suivantes :

- Ouverture au grand public des données produites et financées par les acteurs publics<sup>9</sup>.
- Mise à disposition des données environnementales produites et financées par les acteurs publics de manière interopérable, avec celles mises à disposition dans les autres États membres de l'UE<sup>10</sup>.
- Accès du grand public aux informations environnementales produites et financées par les acteurs publics<sup>11</sup>.
- Protection des données personnelles<sup>12</sup>.

---

<sup>8</sup> Pour aller plus loin : <https://www.enseignerlethique.be/concepts-de-base/#ethique-et-droit>

<sup>9</sup> Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public (refonte) et Règlement d'exécution (UE) 2023/138 de la Commission du 21 décembre 2022 établissant une liste d'ensemble de données de forte valeur spécifiques et les modalités de leur publication et de leur réutilisation.

<sup>10</sup> Directive 2007/2/CE du Parlement européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE).

<sup>11</sup> Directive 2003/4/CE du Parlement européen et du Conseil du 28 janvier 2003 concernant l'accès du public à l'information en matière d'environnement.

<sup>12</sup> Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

- Encadrement des modalités de transfert de données non personnelles « stratégiques »<sup>13</sup>.
- Accès des acteurs publics à des données détenues par des acteurs privés, pour résoudre des situations de crises et pour répondre à des besoins exceptionnels<sup>14</sup>.
- Accès pour les utilisateurs, aux données produites par les objets connectés<sup>15</sup>.
- Mise en place de processus de gouvernance des données utilisées pour entraîner des modèles d'IA<sup>16</sup>.
- Transparence concernant les données utilisées pour entraîner des modèles d'IA de fondation<sup>17</sup>.
- Respect des droits de propriété intellectuelle existants sur les données<sup>18</sup>.

Toutes ces obligations sont applicables en France, la loi française étant en outre parfois plus exigeante, par exemple en matière de droit d'accès aux documents administratifs<sup>19</sup>, en matière d'ouverture des données produites par les acteurs publics<sup>20</sup> et de transparence des algorithmes développés et utilisés par ces derniers<sup>21</sup>.

---

<sup>13</sup> Règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données.

<sup>14</sup> Règlement (UE) 2023/2854 du Parlement européen et du Conseil du 13 décembre 2023 concernant des règles harmonisées portant sur l'équité de l'accès aux données et de l'utilisation des données.

<sup>15</sup> *Id.*

<sup>16</sup> Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle.

<sup>17</sup> *Id.*

<sup>18</sup> *Id.*

<sup>19</sup> Voir les articles L300-1 et suivants du Code des Relations entre le Public et l'Administration.

<sup>20</sup> Voir notamment l'article L312-1-1 du Code des Relations entre le Public et l'Administration, l'article L1115-1 du Code des transports et l'article 25 de la loi n°2019-1428 du 24 décembre 2019 d'orientation des mobilités. Pour les modalités de diffusion exigées, se référer au décret du 27 avril 2017, celles-ci ne devant néanmoins pas aller à l'encontre des règles juridiques applicables en matière de protection des données à caractère personnel, du secret des délibérations du Gouvernement, du secret de la défense nationale ou de la sûreté de l'Etat, du secret médical et du secret des affaires.

<sup>21</sup> Voir les articles L311 et L312 du Code des Relations entre le Public et l'Administration, qui impliquent de publier en ligne les règles définissant les principaux traitements, de faire figurer en ligne et sur les documents une mention précisant les finalités du traitement et de fournir à la demande de l'intéressé des informations concernant le degré et le mode de contribution du traitement algorithmique à la prise de décision, les données traitées et leurs sources, les paramètres de traitement et leur pondération et les opérations effectuées par le traitement.

# Chapitre 1 – Introduction : pourquoi un tel GT ?

## Le mandat du GT

La constitution de ce Groupe de Travail en fin d'année 2023 s'est inscrite dans la volonté du CNIG de fournir aux acteurs de la donnée et de l'information géolocalisée en France un espace de réflexion et de partage de bonnes pratiques concernant les enjeux éthiques liés à leurs activités. Deux objectifs identifiés étaient notamment :

- De sensibiliser un maximum de parties prenantes à ces questions,
- D'établir des recommandations et cadre de référence : ceci étant la raison d'être du présent livrable.

Lors de la réunion visant à définir le mandat de ce Groupe de travail, il était ressorti des discussions que les enjeux éthiques liés aux données deviennent de plus en plus importants, notamment du fait de la quantité croissante de celles-ci (*massive data*) et de la multiplication de leurs réutilisations. Les participant(e)s avaient également mis en avant le fait que les acteurs de la donnée géolocalisée avaient parfois des intérêts différents/divergents et qu'il était important d'être à l'écoute de tous les points de vue. Les participant(e)s avaient par ailleurs manifesté un intérêt pour une synthèse des principes généraux applicables à tous cas d'usage. Depuis son lancement jusqu'à la fin d'année 2024, chacune des séances a été consacrée à la présentation, par les participant(e)s, de cas d'usage réels de traitement de données géolocalisées ayant donné lieu à des débats collectifs autour des enjeux éthiques identifiés.

Le Groupe de Travail est ouvert à tous, acteurs structurants dans le champ de la donnée géographique, acteurs ministériels, organismes publics d'expertise, collectivités territoriales, acteurs de la recherche, entreprises, tous producteurs ou utilisateurs de données géolocalisées, qu'ils soient français ou étrangers.

## Différence entre donnée et information géolocalisée

Une donnée est un élément initial, une observation, une description, une localisation, une mesure, ..., quelque chose qui existe et n'est pas re-traité. C'est une matière première. Elle peut être structurée ou non structurée. Elle peut être produite ou acquise.

Une information est une interprétation, une analyse d'une ou plusieurs données. Par exemple une donnée de température peut être interprétée comme étant chaude ou froide, suivant le contexte de la personne qui l'interprète.

La connaissance pour les données, signifie que celles-ci font partie d'un patrimoine connu et qu'elles sont cataloguées, avec leurs métadonnées<sup>22</sup>.

---

<sup>22</sup> Une métadonnée (metadata) est une information descriptive relative à une donnée (producteur, date de mise à jour, qualité, etc...).



## Chapitre 2 – Panel de cas d’usage étudiés pendant le GT

Lors des réunions du GT éthique menées de décembre 2023 à décembre 2024, les participant(e)s ont été invités à présenter des cas d’usage de données géolocalisées dont ils avaient connaissance, afin de partager et débattre des enjeux éthiques pouvant y être liés. Dans le présent livrable, nous avons fait le choix de présenter ces cas d’usage par thématique. Les points essentiels des discussions entre les participants concernant les enjeux éthiques liés à ces cas d’usage sont repris ci-dessous.

### 1<sup>ère</sup> thématique : L’ouverture au grand public d’informations géolocalisées

#### 1° L’ouverture d’informations géolocalisées existantes détenues par les acteurs publics

##### *Cas d’usage : L’ouverture de données non actualisées par la Ville de Montréal*

*En 2010, la Ville de Montréal a dû appliquer une politique d’ouverture des données qu’elle détenait. Dans ce contexte, une carte créée dans les années 70 indiquant les hauteurs des bâtiments de la ville a été ouverte au public par un autre Service que celui de la géomatique, sans vérification préalable du fait que les données étaient encore à jour ou non. Or, celles-ci ont été réutilisées pour produire une maquette 3D des bâtiments, qui s’est avérée erronée. En interne, s’est posée la question de qui pouvait / devait être habilité à ouvrir les données détenues par la Ville. En effet dans ce cas précis, le producteur des données initiales, à savoir la Division de la géomatique, n’avait pas été consulté, alors-même qu’en tant que producteur sa responsabilité juridique aurait pu être engagée du fait de la mauvaise qualité des données. Depuis, un protocole a été mis en place afin de mieux garantir la qualité et l’intégrité des données qui sont ouvertes : lorsque la Ville souhaite ouvrir une base de données, les producteurs sont contactés pour validation et remplissent un formulaire précisant les caractéristiques des données, celles-ci étant vérifiées par le Chef du Service concerné, avant d’être renseignées dans le « portail de données ouvertes ».*

##### **Enjeux éthiques discutés avec les participant(e)s du GT**

1° *La communication relative au niveau de qualité de données rendues accessibles au public est essentielle : Cela permet de fournir une information sur les usages appropriés des données au regard de leurs caractéristiques.*

2° *Les ré-utilisateurs de données ouvertes devraient être davantage incités à formuler des retours aux producteurs concernant les usages qu’ils en font et les défauts de qualité qu’ils ont rencontrés. Aujourd’hui, au Canada comme en France, cela n’est pas obligatoire et, malgré des statistiques démontrant que certaines données ouvertes sont très consultées, très peu de retours sont formulés aux producteurs. Or ils en seraient preneurs car cela permettrait d’améliorer en continu la documentation des usages pour lesquels elles pourraient être adaptées. Sur ce point, le GT souligne un enjeu : certains ré-utilisateurs seraient davantage à la recherche de données confortant leurs théories et leur permettant de réaliser leurs projets plutôt qu’à la recherche de données d’une qualité suffisante par rapport à l’usage qu’ils souhaitent en faire. Néanmoins, quelques bonnes pratiques ont été identifiées par le GT :*

- *Au minimum, la mise en place de procédures simples permettant aux ré-utilisateurs de données*

- ouvertes de formuler des retours aux producteurs concernant le niveau de qualité des données par rapport à l'usage qui en a été fait ;*
- *La mise en place d'une Charte éthique expliquant l'utilité de tels retours pour les producteurs de données ouvertes, et pour les futurs réutilisateurs (une telle Charte a vocation à être expérimentée dans le cadre du catalogue de données Green Data for Health mis en œuvre par Ecolab – le laboratoire d'innovation du Ministère de la Transition Ecologique ; cette Charte a été co-construite avec les utilisateurs du site, dans l'objectif de proposer un document "bonus"/"gagnant-gagnant" basé sur la conception d'un système d'incitations équilibré pour la promotion du partage des données et des réutilisations qui en sont issues) ;*
  - *La mise en place d'incitations ou de « récompenses » à destination des ré-utilisateurs qui formuleraient des retours aux producteurs*

## 2° La production d'informations géolocalisées ayant vocation à être ouvertes au public

### **Cas d'usage 1 : La cartographie d'un phénomène social par la Ville de Montréal**

*Une unité de la Ville de Montréal, indépendante du service de géomatique, avait développé une carte permettant de visualiser un indice d'équité des milieux de vie. Juste avant publication, l'unité a contacté la Division de la géomatique pour les informer de ce projet. Malgré la constatation de défauts par les géomaticiens, qui aurait supposé une refonte du projet (ex. selon l'option choisie de visualisation de données identiques, les résultats étaient différents), le caractère tardif de cette sollicitation et l'absence d'obligation d'obtenir leur aval pour publier la carte n'ont pas empêché la publication de celle-ci.*

#### **Enjeux éthiques discutés avec les participant(e)s du GT**

*1° Toute carte devrait être accompagnée d'un rapport, facilement accessible, détaillant la méthodologie appliquée pour la produire : Les producteurs d'une carte devraient en effet proposer, de manière très visible, un rapport exposant les choix méthodologiques effectués pour produire la carte ainsi que le contexte dans lequel elle a été produite. En effet, ces choix ont un impact direct sur le sens des données visualisées par les utilisateurs. Il est ici important de bien distinguer les enjeux liés à la qualité des données initiales de ceux liés à la représentation de ces données : en l'occurrence, malgré des données initiales de qualité, ici ce sont les choix méthodologiques faits pour les représenter qui ont affecté la fiabilité de l'indice. Autrement dit, des mêmes données peuvent faire dire des choses différentes. Les participants au GT ont souligné que cette problématique se retrouvait aujourd'hui dans les représentations de données faites dans les jumeaux numériques. Ils ont par ailleurs précisé le fait que la documentation accessible concernant la méthodologie appliquée devrait également détailler les algorithmes utilisés pour analyser les données, notamment en vue de garantir leur explicabilité.*

*2° Si les producteurs ont bien une part de responsabilité concernant la qualité des données qu'ils produisent, les utilisateurs devraient également se responsabiliser quant à la vérification de cette qualité, notamment si l'utilisation qu'ils souhaitent faire de ces données a un impact sur la population : Les participants au GT soulignent, comme dans le cas d'usage précédent, le fait qu'aujourd'hui les utilisateurs de ces outils, qu'il s'agisse de cartes ou de jumeaux numériques, se comportent comme de simples consommateurs ne se posant pas de question sur la qualité, la fiabilité et l'exactitude des données. Cet enjeu par ailleurs est accentué lorsque que l'outil est proposé par un acteur public car une confiance plus grande leur est accordée. Or, lorsque de tels outils sont utilisés en vue d'appuyer des décisions politiques, cela devient encore plus problématique.*

*3° Toute représentation cartographique devrait être validée par un ou plusieurs experts en géomatique avant sa publication : Tout en préservant une certaine autonomie des producteurs*

de cartes, une bonne pratique serait de faire valider la méthodologie appliquée par un géomaticien. En effet, ces derniers appartiennent à un ordre professionnel et doivent à ce titre se conformer à des règles déontologiques visant à leur assurer une expertise en traitement de données géolocalisées.

4° Une traçabilité des personnes intervenues dans une représentation cartographique devrait être assurée : Si l'identité des personnes ayant participé à la production des cartes était mentionnée, cela les responsabiliserait davantage. Les participants au GT estiment que si auprès du public la publicité du nom du service ayant produit une carte semble suffisante, une traçabilité nominative interne est importante, au moins sur le rapport méthodologique.

### **Cas d'usage 2 : La Base de Données Topo de l'IGN**

La BD Topo est une base de données librement accessible et réutilisable en licence ouverte, ce qui – pour l'IGN – génère une connaissance limitée sur les usages qui en sont faits réellement. En effet, pour télécharger et utiliser ces données, les utilisateurs n'ont plus besoin depuis 2021 de s'enregistrer par la création d'un compte sur le site Geoservices. Face à ce constat, l'IGN essaye de constituer depuis quelques mois une base de contacts des utilisateurs, par l'intermédiaire de plusieurs canaux : inscriptions aux lettres d'infos, aux événements organisés par l'IGN (webinaire, journée d'échanges) ou par les travaux engagés au fil de temps avec ses partenaires.

#### **Enjeux éthiques discutés avec les participant(e)s du GT**

1° À travers ces sondages l'IGN tente d'identifier des usages « acceptables » qui pourraient être mis en avant et ainsi servir à valoriser la base de données et ses évolutions. Se pose néanmoins la question de ce qui peut être considéré comme « acceptable ». Au-delà de sa mission de décrire topographiquement le territoire, la BD TOPO s'oriente de plus en plus à répondre à des usages experts en lien avec de politiques publiques.

2° L'IGN souligne par ailleurs le fait que le principe de liberté de réutilisation entraîne l'impossibilité d'identifier de manière exhaustive les usages « inacceptables » de la base de données, c'est-à-dire qui ne respecteraient pas les conditions d'utilisation définies par l'IGN, à travers la publication des métadonnées sur la donnée ou de CGU de l'offre générale IGN.

3° A été discuté un enjeu concernant l'évolution des finalités d'utilisation des données relatives au tracé des cours d'eau. En effet, si au départ ces données étaient exclusivement descriptives, l'évolution du cadre réglementaire a nécessité de les rendre opposables aux tiers et a ouvert la possibilité de modifications de ces données par des acteurs autres que l'IGN (ex. DDT). Afin de limiter le risque de description alternative du territoire par ces contributeurs externes, l'IGN envisage de mettre en place un dispositif de labellisation des contributeurs, en y ajoutant des contrôles dédiés et de consolider les spécifications de contenu et de saisie, afin de ne pas laisser place à l'interprétation. L'IGN apportant le support géométrique pour apposer la donnée réglementaire, le complétage attributaire relève de la responsabilité des contributeurs.

### **3° Le contenu des informations géolocalisées communiquées au grand public : l'exemple de la communication autour des risques naturels**

#### **Cas d'usage 1 : La prévision à court terme des risques d'incendie en Méditerranée par l'entreprise Murmuration et le National Observatory of Athens**

Objectif : Améliorer la planification opérationnelle, la préparation et la capacité d'intervention en cas d'urgence, des pompiers et d'autres décideurs (par exemple, les autorités de protection civile, les ministères de l'environnement, etc.), en leur fournissant un système de prévision des

*risques d'incendie robuste et fondé sur des données probantes, avec une résolution spatiale et une précision supérieures à celles de l'état de l'art.*

*Pour répondre aux besoins des utilisateurs finaux, trois produits de base sont proposés : Produit 1 - Prévisions quotidiennes du danger d'incendie ; Produit 2 - Graphiques d'explicabilité ; Produit 3 - Évaluation quotidienne du risque d'incendie*

*Données : Conditions météorologiques (températures, précipitations, vent), observations du satellite MODIS (capteur multispectral dans le visible et l'infrarouge, qui permet notamment de mesurer la santé de la végétation), données d'humidité dans le sol produites par le JRC (« Joint Research Center », laboratoire de recherche de la Commission Européenne), données statiques environnementales (topographie, couverture des sols, densité de population, réseau routier, surfaces en eaux (lacs, rivières...)).*

*Méthodologie : Combiner l'utilisation de modèles physiques et d'un modèle de « Deep Learning » basé sur des données historiques de feux de forêt en Grèce, l'objectif étant d'identifier les facteurs influant le déclenchement d'un feu, pour ensuite produire un modèle de prédiction qui fournit un indice quotidien de risque de feu.*

### ***Enjeux éthiques discutés avec les participant(e)s du GT***

*1° Les producteurs des modèles devraient fournir aux utilisateurs finaux une information loyale concernant les limites des données utilisées pour les alimenter, ainsi que sur leurs conditions optimales d'utilisation (par exemple concernant le biais lié à l'alimentation du modèle, uniquement avec des données historiques).*

*2° Les conditions d'utilisation et de réutilisation des données sortantes du modèles devraient être bien encadrées et prédéfinies (limitées à un objectif strict de prévention des incendies : il serait en effet discutable qu'elles puissent être réutilisées par des promoteurs immobiliers ou des assureurs, par exemple)*

*3° Il existe une incertitude concernant la communication à la population d'une information liée à une probabilité d'incendie. Les acteurs publics sont soumis à une obligation de communication des informations environnementales qu'ils détiennent, ainsi que d'ouverture de certaines données publiques. Leur responsabilité peut donc être mise en cause en cas de manquement à ces devoirs. Toutefois éthiquement, une question se pose concernant la communication d'informations peu fiables / peu probables. Il ne serait en effet peut-être pas bienvenu qu'une prédiction concernant un très faible risque d'incendie entraîne un mouvement de panique au sein de la population. Il pourrait par ailleurs être reproché aux autorités de ne pas être intervenues, alors qu'elles détenaient une information utilisable, même si celle-ci relèverait d'une faible probabilité de survenance d'incendie. Sur cette question de la responsabilité, un participant au GT rappelle une situation similaire s'étant produite dans le domaine sismique : après une série de séismes survenus en Italie, des experts avaient été consultés pour avis sur la probabilité de répliques et, sur la base des données qu'ils détenaient, ils se sont positionnés en faveur d'un faible risque de répliques. Les répliques se sont en réalité avérées importantes et certains experts ont été condamnés à de la prison. L'idée à retenir de cette comparaison est qu'il faudrait éviter que, par craintes de retombées négatives, des chercheurs ou des autorités procèdent à une rétention de certaines informations.*

### ***Cas d'usage 2 : La cartographie de risques d'inondation par la Ville de Montréal***

*Suite à des inondations ayant affecté la Ville de Montréal en 2017 et 2019, la Division de la géomatique*

*a produit des cartes visant à actualiser l'identification des zones inondables, avec beaucoup de documentation contextuelle et méthodologique.*

### ***Enjeux éthiques discutés avec les participant(e)s du GT***

1° Toute carte devrait être accompagnée d'un rapport, facilement accessible, détaillant la méthodologie appliquée pour la produire : Les producteurs d'une carte devraient en effet proposer, de manière très visible, un rapport exposant les choix méthodologiques effectués pour produire la carte ainsi que le contexte dans lequel elle a été produite. En effet, ces choix ont un impact direct sur le sens des données visualisées par les utilisateurs. Il est ici important de bien distinguer les enjeux liés à la qualité des données initiales de ceux liés à la représentation de ces données : en l'occurrence, malgré des données initiales de qualité, ici ce sont les choix méthodologiques faits pour les représenter qui ont affecté la fiabilité de l'indice. Autrement dit, des mêmes données peuvent faire dire des choses différentes. Les participants au GT ont souligné que cette problématique se retrouvait aujourd'hui dans les représentations de données faites dans les jumeaux numériques. Ils ont par ailleurs précisé le fait que la documentation accessible concernant la méthodologie appliquée devrait également détailler les algorithmes utilisés pour analyser les données, notamment en vue de garantir leur explicabilité.

2° Si les producteurs ont bien une part de responsabilité concernant la qualité des données qu'ils produisent, les utilisateurs devraient également se responsabiliser quant à la vérification de cette qualité, notamment si l'utilisation qu'ils souhaitent faire de ces données a un impact sur la population : Les participants au GT soulignent, comme dans le cas d'usage précédent, le fait qu'aujourd'hui les utilisateurs de ces outils, qu'il s'agisse de cartes ou de jumeaux numériques, se comportent comme de simples consommateurs ne se posant pas de question sur la qualité, la fiabilité et l'exactitude des données. Cet enjeu par ailleurs est accentué lorsque que l'outil est proposé par un acteur public car une confiance plus grande leur est accordée. Or, lorsque de tels outils sont utilisés en vue d'appuyer des décisions politiques, cela devient encore plus problématique.

3° Toute représentation cartographique devrait être validée par un ou plusieurs experts en géomatique avant sa publication : Tout en préservant une certaine autonomie des producteurs de cartes, une bonne pratique serait de faire valider la méthodologie appliquée par un géomaticien. En effet, ces derniers appartiennent à un ordre professionnel et doivent à ce titre se conformer à des règles déontologiques visant à leur assurer une expertise en traitement de données géolocalisées.

4° Une traçabilité des personnes intervenues dans une représentation cartographique devrait être assurée : Si l'identité des personnes ayant participé à la production des cartes était mentionnée, cela les responsabiliserait davantage. Les participants au GT estiment que si auprès du public la publicité du nom du service ayant produit une carte semble suffisante, une traçabilité nominative interne est importante, au moins sur le rapport méthodologique.

## 2<sup>ème</sup> thématique : L'utilisation de données et d'informations géolocalisées

### 1° A des fins environnementales

#### ***Cas d'usage 1 : La création d'un tableau de bord environnemental pour l'aide à la prise de décision dans le domaine du tourisme par l'entreprise Murmuration***

Contexte : « Comme toute activité économique et sociale, le tourisme a besoin de données et de recherches pour permettre et soutenir son développement réussi et durable. Pour contrôler la durabilité, il est nécessaire de disposer de données sur l'utilisation des ressources, telles que la terre, l'énergie et l'eau, les émissions et le traitement des déchets, la dégradation, la (sur)fréquentation, la congestion et les nuisances. L'évaluation des impacts sociaux, économiques et environnementaux permet de concevoir et de mettre en œuvre des politiques et

des stratégies visant à maximiser la contribution positive du tourisme et à minimiser ses effets négatifs. »

Production d'indicateurs environnementaux à partir de données d'observation de la Terre par satellite, dans plusieurs domaines : pollution de l'air, pollution de l'eau (zone côtières, lacs), santé de la végétation, artificialisation des sols, îlots de chaleur urbaine, changement climatique  
Croisement avec des données de fréquentation touristique : service « FluxVision » de « Orange Business Service » (statistiques anonymisées issues du réseau de téléphonie mobile, donne le nombre de résidents et de touristes présent sur une zone, par tranche de 2H dans la journée ou par nuitée) et les séries de données de fréquentation et d'indicateur environnementaux sont injectés dans des modèles d'IA pour identifier des tendances, des corrélations et isoler l'impact spécifique de l'activité touristique.

Intégration et présentation de ces indicateurs dans un tableau de bord interactif : visualisation cartographique, visualisation temporelle (courbes), mise à jour temps réel, historique dans le passé et projections dans le futur pour certains indicateurs (changement climatique), échelle locale au territoire ciblé (Commune, Communauté de Communes, Département, Région) et possiblement aussi plus large (nationale), pour pouvoir se comparer à des territoires voisins.

**Enjeux éthiques pré-identifiés** : Qualité des données représentant les flux de mobilité touristique, fiabilité des indicateurs produits pour mesurer l'impact environnemental du tourisme à l'échelle territoriale, formation des utilisateurs finaux concernant l'utilisation de cet outil, dans le cadre d'un processus décisionnel, contribution du projet au développement d'un tourisme durable.

**Enjeux éthiques discutés avec les participant(e)s du GT**

1° Les producteurs de données/modèles devraient fournir aux utilisateurs de ceux-ci une information loyale concernant leur qualité, afin d'éviter toute mauvaise interprétation. Toute la chaîne de valorisation de la donnée devrait donc être attentive à la documentation de cette qualité et notamment renseigner les limites des données/modèles produits, ainsi que les utilisations non pertinentes. Concernant spécifiquement les modèles, les producteurs devraient définir et être transparent sur le seuil de précision jugé acceptable pour appuyer une décision.

2° Les utilisateurs (ici les institutionnels du tourisme) devraient eux-mêmes informer les destinataires finaux des incertitudes affectant les données sur lesquelles ils fondent leurs décisions.

3° La formation de ces utilisateurs à la bonne compréhension des outils mis à leur disposition serait très pertinente. Pour dispenser une telle formation, le fournisseur de l'outil semble être l'acteur le mieux placé. Au-delà de cette formation continue dans le milieu professionnel, une pédagogie dans le milieu scolaire/supérieur concernant les enjeux liés à la fiabilité de l'information apparaît complémentaire.

4° Il serait intéressant que chaque acteur de la chaîne de valorisation des données nomme un garant de l'usage éthique des données, à qui toute personne intéressée pourrait s'adresser pour se renseigner sur la qualité des données/modèles, les conditions optimales d'utilisation d'un outil,

### **Cas d'usage 2 : La création d'un outil permettant le suivi des migrations induites par le climat en Afrique par l'entreprise Murmuration et l'Université de Valence**

Objectifs : Modéliser, anticiper et comprendre les flux migratoires induits par le climat (vagues de chaleur, sécheresses et inondations) en Afrique, à partir de données fiables et dévoiler les dynamiques existantes entre toutes les variables considérées, grâce à un modèle d'inférence causale, afin de comprendre, quantifier et prévoir les effets des flux migratoires à partir du contexte socio-économique et de variables environnementales extraites de données d'observation de la Terre.

Utilisateurs ciblés : agences nationales et inter-gouvernementales telles que l'iDMC (« Internal Displacement Monitoring Centre ») et le WFP (« World Food Program »)

Données : Conditions météorologiques (températures, précipitations, vent), observations du satellite MODIS : capteur multispectral dans le visible et l'infrarouge, utilisé pour mesurer la

santé de la végétation et l'artificialisation des sols, données socio-économiques (prix du marché local, conflits violents, pandémies, achat/vente de bétail...).

### **Enjeux éthiques discutés avec les participant(e)s du GT**

1° Les producteurs des modèles devraient être attentifs à la qualité des données socio-économiques utilisées et la documenter (par exemple en accompagnant leurs publications scientifiques de data papers).

2° Les producteurs des modèles devraient informer les utilisateurs finaux des hypothèses à l'origine de liens de causalité supposés entre deux variables.

3° Les conditions d'utilisation des modèles devraient être prédéfinies, notamment afin d'éviter toute récupération à des fins politiques, voire démagogiques, des données sortantes. Il serait pour cela important de définir une typologie de personnes habilitées à les utiliser et d'encadrer à quelles fins elles peuvent être utilisées.

4° Afin de faciliter la réutilisation des résultats de leurs recherches, les chercheurs peuvent par exemple s'appuyer sur les principes FAIR (Findable/Facile à trouver, Accessible/Accessible, Interoperable/Interopérable, Reusable/Réutilisable). Il conviendrait toutefois de préciser les points de vigilance à prendre en considération concernant l'application de ces modèles à d'autres zones géographiques.

5° Les utilisateurs finaux des données sortantes des modèles devraient être transparents (à l'égard des destinataires de l'information fournie sur la base de ces données) sur le niveau d'incertitude affectant celles-ci.

## **2° A des fins de contrôle social**

### **Cas d'usage : Le ciblage des contrôles des bénéficiaires de prestations sociales par la CNAF**

#### **Enjeux éthiques discutés avec les participant(e)s du GT**

La question de la finalité de contrôle de la CNAF n'est pas remise en cause. Ce sont les moyens utilisés qui peuvent l'être, à travers l'usage d'algorithmes potentiellement biaisés.

1° La nécessité de transparence des algorithmes développés par des acteurs publics dont l'utilisation affecte les droits des personnes : Les participants sont unanimes sur le fait que cette transparence est essentielle, afin de permettre l'identification des critères ayant le plus de poids dans un calcul algorithmique et ainsi de potentiels biais discriminatoires. En l'espèce, afin de développer/tester l'outil, 7 000 dossiers liés à des personnes ayant déjà fait l'objet d'un contrôle avaient été tirés au sort, afin d'identifier les caractéristiques communes aux foyers contrôlés comme fraudeurs. Toutefois, un premier enjeu réside dans le fait que certaines de ces caractéristiques sont des critères pouvant être discriminatoires. Cet enjeu est renforcé par le fait que, bien que les dossiers frauduleux aient des caractéristiques communes et que celles-ci puissent être corrélées à un acte de fraude, la caractérisation d'un lien de causalité réel entre de telles caractéristiques et le caractère frauduleux des dossiers n'a pas été recherchée.

Les participants se sont par ailleurs posé la question du niveau de transparence adéquat : une transparence générale par principe serait-elle plus adaptée qu'une transparence uniquement à l'égard des allocataires ciblés par l'algorithme pour faire l'objet d'un contrôle ?

2° La nécessité de proportionnalité entre la finalité de l'outil basé sur le traitement de données personnelles et les moyens mis en œuvre pour y parvenir : En l'espèce, peut-être y aurait-il des moyens moins sensibles en termes de protection des données personnelles et des droits fondamentaux pour permettre à la CNAF de cibler ses contrôles.

## **3<sup>ème</sup> thématique : La mise en commun de données géolocalisées**

## **Cas d'usage : La création d'un commun numérique ouvert de photos terrain**

*En 2022, la Fabrique des géocommuns de l'IGN, en collaboration avec OSM (Open Street Map) France et l'incubateur beta.gouv.fr de la DINUM (Direction du Numérique de l'État), lance Panoramax : le géocommun pour partager ses photos de terrain. Cette base de vues immersives collaborative et ouverte, a pour but de mettre en commun les initiatives existantes en France (et aussi à l'étranger) : une seule photo, initialement prise pour une raison particulière, peut ainsi servir différents utilisateurs/trices et donc générer beaucoup d'usages différents. Cette mise en commun peut aider à réduire les déplacements sur le terrain et à accéder à des données récentes, notamment sur des zones qui n'avaient pas encore été couvertes par ailleurs. Cette solution opensource et souveraine permet une gouvernance partagée par une communauté nationale et internationale, la mise en place de règles d'utilisation claires, mais aussi l'accessibilité, l'autonomie et la pérennisation de ce type de données. Elle repose sur un système d'hébergement décentralisé et une API qui garantit l'anonymisation des données : n'importe qui peut ouvrir une instance Panoramax à l'image de celles de l'IGN et d'OSM France et rejoindre le catalogue fédéré. Les contributeurs et contributrices ont alors le choix de l'instance à laquelle envoyer les photos et une seule organisation n'a pas à héberger les photos du monde entier.*

### **Enjeux éthiques discutés avec les participant(e)s du GT**

*1° Enjeu d'anonymisation par le floutage des visages et des plaques d'immatriculation : Chacune des instances qui héberge des photos est responsable de leur traitement au sens du RGPD, et donc responsable de leur anonymisation. La start-up d'Etat joue ici le rôle de sous-traitant dans la mesure où c'est elle qui développe les outils et le modèle d'anonymisation, qui est amélioré en continu, notamment afin de renforcer sa performance et minimiser les défauts d'anonymisation.*

*2° Enjeu lié aux zones protégées (pour des raisons sécuritaires ou de confidentialité) : en France, il existe des zones interdites à la captation aérienne des données (ZICAD), qui sont donc exclues de la base de données de Panoramax. Le projet étant d'envergure internationale, il existe un enjeu quant à la connaissance de telles réglementations dans les autres pays du monde. Comme pour l'anonymisation, l'administrateur d'une instance en est a priori responsable. Se pose également la question de l'intérêt / de la possibilité de privatiser certaines séquences au sein des instances d'hébergement afin d'en permettre l'accès et la réutilisation de manière restreinte.*

*3° Enjeu de l'impact environnemental du projet : Le projet impliquant des déplacements motorisés pour faire les photos ainsi que l'hébergement de nombreuses photos, son impact environnemental est pourrait devenir de plus en plus important. Concernant les déplacements, il pourrait être intéressant de développer des indicateurs permettant mettre en balance les déplacements motorisés économisés grâce à Panoramax avec les déplacements motorisés générés pour contribuer à Panoramax. Concernant le stockage des photos, les CGU de Panoramax pourraient par exemple définir un niveau de qualité adéquat et un format standardisé à respecter (cf. GT CNIG standards d'image de captation terrestre).*

*4° Enjeu lié à l'intégrité des données versées par les contributeurs : Aujourd'hui rien ne garantit que les photos versées dans les instances d'hébergement soient des photos réelles et intègres, au sens de non modifiées, altérées, ou encore générées par des outils à base d'IA. Il existe simplement une possibilité de signalement de photos non réelles, altérées, etc. mais des travaux sont en cours pour automatiser la modération.*

*5° Enjeu d'inclusion d'une diversité de contributeurs individuels : Aujourd'hui la majorité des contributeurs individuels sont issus de la communauté OSM. Plusieurs méthodes peuvent être imaginées pour inclure une communauté plus large de citoyens : la gamification du projet (challenges de contribution, hackathon, ...), des subventions données aux collectivités territoriales afin qu'elles puissent équiper leurs agents de GOPRO, l'animation des communautés SIG dans les collectivités ou encore, de manière plus générale, l'acculturation et la sensibilisation*

*des citoyens au projet et à ses objectifs. L'application mobile jouera un rôle important aussi dans la démocratisation et l'accessibilité à la contribution.*

*6° Enjeu de gouvernance multi-acteurs : Un modèle de gouvernance permettant la représentation des différentes catégories de contributeurs / d'utilisateurs est à imaginer pour permettre la pérennisation du projet suite à la fin de son incubation (la création d'une association semble être la piste privilégiée).*

*7° Enjeu d'intégration de la base de données Panoramax dans des projets parallèles mis en œuvre à l'IGN (ex. Cartes.gouv, Cartes IGN).*

## Chapitre 3 – Des bonnes pratiques en matière d'éthique de l'usage de la donnée/information géolocalisée

### Introduction

Les débats concernant les enjeux éthiques liés aux cas d'usage présentés en Chapitre 2 ont débouchés sur l'identification de bonnes pratiques générales ayant pour objectif d'aider les différentes parties prenantes à la production, la gestion et l'usage de données géolocalisées à rechercher, au cas par cas, la décision la plus juste au regard de principes de référence. Ces bonnes pratiques sont complémentaires au respect des règles juridiques applicables aux acteurs concernés.

### Bonnes pratiques organisées selon le cycle de vie d'une donnée

#### *Bonnes pratiques en matière de collecte/production :*

##### *Bonnes pratiques générales*

- Définir (= documenter) les conditions de production des données :
  - Données d'observation : identifier le producteur (organisation publique comme privée) pour savoir s'il s'agit d'un professionnel ou bien d'un citoyen (anonyme dans ce cas).
  - Données de mesure : identifier le matériel utilisé, sa calibration, le protocole de mesure mis en œuvre.
  - Renforcer les critères RSE relatifs au mode de production et de collecte des données dans tout cahier des charges en cas visant à contracter avec des acteurs tiers
- Questionner la nécessité de la géolocalisation :
  - Prise en compte du caractère sensible ou stratégique de la géolocalisation (ex. existence de zones interdites à la captation aérienne).
  - Respect de la vie privée des personnes concernées par la géolocalisation, ceci amenant à privilégier l'anonymisation des données.
- Questionner le niveau de précision nécessaire à la géolocalisation :
  - Mise en balance du niveau de précision nécessaire par rapport aux usages qui en seront faits.
  - Mise en balance du niveau de précision nécessaire avec un principe de sobriété numérique (éviter la « sur-qualité »). Exemple avec [Panoramax](#) : résolution des images, métadonnées, ... en plus de la géolocalisation.
- Documenter le processus de mise en qualité des données (règles de gestion, valeurs limites...)
- Documenter le respect de ces règles et des informations qui en découlent (limites, hypothèses à l'origine de liens de causalité supposés entre des variables, ...).

#### *Cas particulier des représentations cartographiques*

- Assurer en interne une traçabilité des personnes responsables de la production des données.

#### *Cas particulier de la co-production de données, par exemple pour constituer un commun numérique*

- Mettre en place des mesures visant à garantir l'intégrité des données produites par chacun/e des contributeurs/trices.
- Mettre en place une gouvernance représentative de la diversité des contributeurs/trices.
- Favoriser la contribution d'un maximum de parties prenantes, dans le cas d'un commun ouvert.
- Diffuser les bonnes pratiques à tous les co-producteurs, notamment en terme de sécurité, de collecte sur le terrain : formations, guides, tutoriels, consignes, ...

#### *Bonnes pratiques pour le choix des infrastructures d'hébergement et d'analyse des données*

- Maîtriser la production et les usages des données nécessaires à la mise en œuvre des politiques publiques :
  - Questionner la réversibilité des données à l'égard des technologies utilisées.
  - Questionner les règles juridiques applicables au fournisseur de cloud utilisé.
  - Privilégier des infrastructures souveraines permettant de tracer les accès aux données.
- Stocker les données dans un standard, qui peut-être interne, mais qui doit garantir la pérennité de leur conservation et leur intégrité. Le respect des principes FAIR<sup>23</sup> est souhaité.

#### *Bonnes pratiques en matière de mise à disposition à des tiers*

##### *Bonnes pratiques générales*

- Définir le plus en amont possible les conditions d'utilisation des données/informations/outils (personnes ayant le droit d'y accéder, finalités d'utilisation, ...) – ceci étant particulièrement important concernant les données sensibles, par ex. sur les risques naturels.
- Être transparent sur le niveau de qualité des données (ceci permettant d'en favoriser des réutilisations pertinentes), par exemple en communiquant les informations suivantes (métadonnées) :
  - Limites identifiées dans la qualité des données.
  - Hypothèses à l'origine de liens de causalité supposés entre des variables d'analyse.
  - Possibilité de s'en servir pour appuyer une décision.
  - Utilisations non pertinentes.
  - Source.

---

<sup>23</sup> Définition : [https://fr.wikipedia.org/wiki/Fair\\_data](https://fr.wikipedia.org/wiki/Fair_data)

- Date de production/collecte ou de dernière mise à jour des données<sup>24</sup>.
- Durée de validité.
- Respecter les principes FAIR.
- Respecter des standards validés pour favoriser l'interopérabilité et les réutilisations.
- Proposer des formations et de la documentation pour la bonne compréhension des données/informations/outils mis à disposition.
- Mettre en place des procédures ou outils simples permettant aux ré-utilisateurs de données ouvertes de formuler des retours aux producteurs concernant le niveau de qualité des données.
- Interroger les ré-utilisateurs sur leur besoin d'amélioration de leur expérience utilisateur.

#### *Cas particulier des représentations cartographiques*

- Assortir toute cartographie d'un rapport détaillé de la méthodologie appliquée pour les produire.
- Faire valider celle-ci par un ou plusieurs experts en géomatiques avant publication.

#### *Bonnes pratiques en matière de réutilisation*

- Engager un dialogue avec le/la producteur/trice de la donnée.
- Questionner le contexte dans lequel la donnée a été collectée :
  - Identifier pour quels cas d'usage la donnée pourrait être réutilisée.
- Questionner le niveau de précision nécessaire de la géolocalisation :
  - Prise en compte des usages qui en seront faits.
- Être transparent à l'égard des destinataires finaux d'une décision, prise avec l'appui de données géolocalisées, du niveau d'incertitude pouvant affecter ces données.

#### *Bonnes pratiques transversales*

##### *Bonnes pratiques générales*

- Nommer un(e) garant(e) de l'usage éthique des données dans chaque organisation.
- Participer à la gouvernance territoriale des données.
- Participer aux instances de normalisation des données.
- Prendre en compte l'accessibilité à plusieurs niveaux, notamment technique et économique, pour toujours plus d'inclusion.

##### *Cas concernés par l'intelligence artificielle*

- Avoir des données d'apprentissage représentatives, pour éviter biais, discriminations, oublis/erreurs par les modèles. Représenter tous les territoires, toutes les personnes pour assurer l'inclusion (et le respect de la loi qui souvent va de pair). Exemple avec [Panoramax](#) : le modèle IA d'anonymisation a dû être amélioré avec l'internationalisation du projet pour la reconnaissance de visages de toute ethnicité et des plaques

---

<sup>24</sup> Concernant les données entrant dans le champ d'application de la Directive INSPIRE, les standards à respecter en la matière y sont précisés. Voir notamment : [https://www.geoinformations.developpement-durable.gouv.fr/document-de-recommandations-pour-des-metadonnees-a2563.html?id\\_rub=296](https://www.geoinformations.developpement-durable.gouv.fr/document-de-recommandations-pour-des-metadonnees-a2563.html?id_rub=296)

d'immatriculation étrangères, bien que dans le cas présent il s'agit surtout d'un sujet juridique (ceci dit, la réglementation peut varier d'un pays à l'autre).

- Développer des modèles en suivant une démarche frugale, c'est-à-dire « en prenant en compte les mécanismes d'optimisation connus de l'état de l'art pour en réduire les besoins en ressources matérielles et énergétiques et les émissions associées, tout en conservant une performance suffisante »<sup>25</sup>.
- Mutualiser les modèles.

---

<sup>25</sup> AFNOR SPEC 2314, Référentiel général pour l'IA frugale – Mesurer et réduire l'impact environnemental de l'IA, juin 2024.

## ANNEXE : Points communs à différentes chartes d'éthique

Le travail visant à identifier les principes faisant consensus dans la plupart des initiatives va être réalisé par la coanimation du GT et devrait être disponible au second semestre 2025.

## REMERCIEMENTS AUX PARTICIPANT.E.S AU GROUPE DE TRAVAIL

Sylvain Riss, Advances  
Arnaud Garnier, BRGM  
Christian Giraud, CGD Conseil  
Ferdinand Causse, CNIG  
Benoît Morando, CNIG  
Pierre Laulier, CNIG  
Mathilde Wind, CNIG  
Arnauld Gallais, CEREMA  
Guillaume Chrétien, CEREMA  
Richard Mongeau, ÉTS Montréal  
Pierre-Alban Bonin, FNCCR  
Amélie Crépin, IGN  
Marie Lambois, IGN  
Thierry Lallemand, IGN  
Serge Hombert, Métropole Européenne de Lille  
Fabien Castel, Murmuration  
Grégory Delobelle, Région Centre Val de Loire  
Bruno Thenail, Région Normandie  
Agnès Ruffat, Région Occitanie  
Cécile Tamoudi, Rennes Métropole  
François Robida, Robida Consultant  
Pierre Pages, Toulouse Métropole  
Marie Parent, Ville de Montréal  
Simon Gignac, Ville de Montréal