



MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE

*Liberté
Égalité
Fraternité*

COMMISSARIAT
GÉNÉRAL AU
DÉVELOPPEMENT
DURABLE

Suites de la mission Bothorel

Commission Données, CNIG, 23 mars 2021



Contexte de la mission

- Confiée par le PM à **Eric Bothorel**, assisté de **Renaud Vedel** (coord. national pour l'IA) et **Stéphanie Combes** (directrice Health Data Hub)
- Appui de **l'IGF (et coordination), l'IGAS, l'IGESR, CGE, IGA, CGDD/Ecolab** + Nicolas Amar (stratégie nationale IA) et Lorien Benda(HDH) + Maxime Donadile, l'assistant du député
- Direction du groupe « des inspecteurs » : Florence Gomez, IGF
- Soutien d'Etalab (Simon chignard et Bastien Guerry) + aide de Romain Tales, Mathilde Huong et Perica Sucevic,
- Près de **200 auditions** dont 5 en contradictoire, des dizaines de questionnaires retournés.
- 2 points presse
- Site Internet présentant le rapport intermédiaire et soumission de 10 cas d'usage de données privées à commentaires.
- Contexte du covid prégnant : concentration assez forte de la mission sur les données de santé

Les tendances des auditions : données publiques et codes sources

Les constats englobant tous les ministères (MTE plutôt bon):

- Hétérogénéité importante des administrations
- équilibre non trouvé entre ouverture et sécurisation des données (données personnelles, secrets, sécurité SI) : fermetures abusives de données. Autorité du DPO trop prononcée par rapport à l'action de l'AMD.
Démarches trop longues = perte de chance pour l'innovation ou risque qu'elle se fasse sur données étrangères.
=> nécessité de contrebalancer vers l'ouverture
=> Le RGPD offre des marges qui ne sont pas assez exploitées
=> assainir le processus d'autorisation
- En parallèle, les opérateurs se rapprochent des bonnes pratiques de la recherche : comité pour une science ouverte, principe européen « **aussi ouvert que possible, aussi fermé que nécessaire** ».
- Flux de données entre Etat et territoires déficient dans les deux sens

Les tendances des auditions

Les opportunités/améliorations:

- La crise du COVID a démontré l'utilité des données publiques et des niveaux d'acculturation très hétérogènes des administrations => nécessité d'un **portage politique fort**
- Comme **alternative** entre ouverture et fermeture des données : le **partage et l'accès** ne sont pas assez expérimentés, en particulier entre administrations
- Besoin d'experts et de cadres connaisseurs à tous niveaux : **ne pas attendre le remplacement de génération** / importance de la formation **à tout niveau**

Les tendances des auditions

- **Focus sur les infrastructures (Hubs, Espaces communs de données) :**
 - Importance de la standardisation et interopérabilité des données
 - Pas d'uniformisation des hubs car très dépendant des écosystèmes
 - Nécessité de mettre en place l'interopérabilité inter-Hub pour lutter contre les silos
- **Sur réponse aux besoins de l'IA :**
 - Pour la mission, l'IA a été l'occasion d'un focus sur la **qualité des données** :
 - La qualité n'est pas assez renseignée dans les métadonnées
 - Trop peu de producteurs se soucient des réutilisations, conduisant à l'insuffisance de la qualité des données et métadonnées
 - Concernant **l'IA entraînée sur des données sensibles** et qui ne permet pas par la suite de reconstituer les données d'apprentissage :
 - Besoin de données sensibles **en masse** et notamment de données publiques
 - Besoin de les conserver dans le temps (non prévu par RGPD)
 - Idée : D'autres pays mettent en place des **bacs à sable** pour tester des dérogations au RGPD. C'est aussi une possibilité ouverte par la loi Energie-Climat (en cours de déploiement par le CRE).

Les tendances des audits

- **Sur données d'intérêt général :**

- Absence de remontée des données produites durant les délégations de service public malgré la loi Lemaire: opportunité d'amélioration
- Séparer le partage des données privées pour les politiques publiques (B2G) et le partage des données privées entre privés (B2B)
- Impossibilité de mettre en place un cadre juridique strict mais besoin de cohérence entre pratiques

- **Au niveau gouvernance : Besoin d'un portage politique et administratif plus fort, Renforcement du réseau AGD – AMD**, cheville ouvrière de la politique publique des données et codes sources

Les recommandations

- Recommandations transversales :
 - But : **associer la société civile**. Consultation Internet sur des cas d'usage durant la mission => grande méfiance du public à l'égard de l'Etat, y compris quand la donnée est gérée dans un cadre sécurisé.
 - Associer la société civile à **l'identification de jeux de données à ouvrir**.
 - Évaluation de **l'impact économique, social et scientifique** de l'ouverture et du partage : chiffres européens : 184 Mds€ en 2019, 334 en 2025 à affiner

Les recommandations

- Portage de la politique :
 - Au niveau du 1^{er} ministre
 - Inscription à l'ordre du jour des **CITP**
 - Ediction d'une **circulaire** établissant les principes
- Mise en œuvre de l'ouverture des données et codes sources
 - Nommer un AGDAC (AC pour « algorithmes et codes sources) auprès du DINUM et piloter la stratégie nationale d'ouverture en s'appuyant sur les AMDAC
 - Réfléchir le **pilotage interministériel** (indicateurs de performance, insertion dans les études d'impact)
 - Créer un **OSPO** à la DINUM (Open source program office)
 - Participer plus activement aux **communs numériques**

Les recommandations

- Mise en œuvre de l'ouverture des données et codes sources (suite)
 - Recommandation n° 9 : **Élargir et renforcer la fonction d'administrateur ministériel des données**, des algorithmes et des codes sources (AMDAC) :
 - en redéfinissant leurs missions dans une fiche de poste type
 - en dotant les AMDAC d'une **lettre de mission signée par les ministres** concernés après consultation des **directions générales** et de la DINUM
 - en s'assurant que l'AMDAC a des **moyens d'intervention suffisants**
 - en systématisant des formations conjointes entre AMDAC et délégués à la protection des données
 - **Mission d'accompagnement des collectivités à l'ANCT**
 - Prise en compte de l'opendata dans l'évaluation et le financement de la **recherche**

Les recommandations

- Droit et régulation
 - Pouvoir de sanction à la CADA « en cas de non-respect des dispositions du CRPA relatives à la communication et à la publication des données et documents »
 - Renforcer la CNIL en etp pour renforcer son rôle de conseil et accompagnement
 - Introduire dans les collèges CNIL et CADA 2 personnalités qualifiées en sécurité des SI et nouveaux usages de la donnée
 - Associer plus l'ANSSI : CADA, CNIL, AGDAC pourraient saisir ou solliciter l'ANSSI : but de « dédramatisation », audit de bibliothèques de logiciels libres
 - Vérifier que la loi garantit l'ouverture des données de services publics mis en œuvre par des acteurs privés

Les recommandations

- Acculturation et politique RH
 - NB : Constat qu'on ne peut pas attendre le remplacement de génération
 - **Formation** plus ambitieuse de la fonction publique, de la formation initiale aux séminaires de cadres dirigeants
 - Poursuivre les travaux relatifs à la **gpec** : valorisation de parcours, **pérennisation en CDI** des contractuels avec compétences absentes des corps existant
 - Diversification des parcours des administrateurs et attachés de l'INSEE
 - Accroître l'attractivité de l'Etat, développer le programme EIG
 - Offre de **formation pour les élus**

Les recommandations

- Qualité de la donnée
 - Créer un **label de service producteur** de la donnée
 - Définir et mettre en œuvre une **politique interministérielle d'interopérabilité et de qualité** de la donnée
 - Encourager les **écosystèmes** à définir des **principes de gouvernance de la qualité** : référent, **communautés** de réutilisation avec participation active des producteurs de la donnée

Les recommandations

- Infrastructure, partage et accès sécurisé
 - **Orienter les investissements du plan de relance** vers les projets favorables à la circulation de la donnée
 - **Hubs** sectoriels ou intersectoriels : assurer leur **interopérabilité** (danger de re-silotage)
 - **Bac à sable expérimental** permettant à la CNIL de déroger aux textes existants pour les besoins de l'IA (NB : idée que l'IA apprenante sur des données personnelles n'en restitue aucune une fois déployée, nécessité de conserver les données personnelles d'apprentissage dans le temps)
 - Prise en charge des **demandes de chercheurs** dans les administrations

Les recommandations

- Données d'intérêt général
 - **Approche incitative** et concertée plutôt que coercitive
 - **Sécuriser** par un guide de conformité de la CNIL **le cadre juridique du partage** volontaire de DIG, clarifier l'application du droit d'accès et de réutilisation des données du secteur privé reçues par les administrations
 - Encourager les initiatives de **portabilité citoyenne des données** au service de l'intérêt général

Les recommandations

- Données d'intérêt général B2G
 - Clarifier **le régime juridique de la réquisition** pour motif impérieux d'intérêt général et d'urgence
 - Confier au réseau AGDAC et AMDAC **une mission de facilitation et de médiation** en lien avec la DGE
 - Garantir l'efficacité des 2 articles de la loi Lemaire :
 - Remonté des données détenues par les délégataires et concessionnaires (=> Clausier-type à destination des structures territoriales)
 - Article 19 : utilisation de données privées à des fins statistiques
- B2B
 - Développer le **B2B au service de l'intérêt général** (comités de filières, au sein des projets PIA, soutien d'initiatives associatives et privée)

L'après mission

- MTFP désigné pour suites mission (avec Cedric O), supervisé par Adnène Trojette (cab PM)
- Eric Bothorel invité par le MTFP à suivre les suites.
- 7/12 : Rim informelle : présentation par le député et Florence Gomez

*Cf Synthèse du rapport : « **Que se passera-t-il après ce rapport? En l'état actuel des choses, la mission craint que ses recommandations ne soient pas portées et suivies.** Elle formule donc plusieurs recommandations pour donner les moyens à cette politique d'être transformée en actes. »*

- 14/12 : Rim recommandations

***Recommandation n° 4** : Assurer un portage politique au niveau du Premier ministre des enjeux de la donnée et des codes source. Inscrire à l'ordre du jour des comités interministériels présidés par le Premier ministre le suivi et la mise en œuvre de cette politique. Édicter une circulaire établissant les principes (gouvernance, missions et responsables dans les administrations, interopérabilité, qualité, guides juridiques)*

- => Inscription d'un point au CITP du 5/2/21
- D'autres RIM préparatoires pour lesquelles un tableau de bord d'ouverture de certaines données ministérielles est validé par le cab PM

L'après mission

- Les 12 engagements du 5ème CITP. Engagement 6 sur la donnée mais données et API présentes dans les engagements 5, 7 et 9:

Engagement # 5

Le baromètre des résultats de l'action publique sera mis à jour tous les trimestres et sera étendu à 40 réformes d'ici juillet 2021. Dès avril, le baromètre sera enrichi de 11 réformes prioritaires.

Engagement # 6

Une politique publique de la donnée ambitieuse sera mise en œuvre par tous les ministères pour mieux exploiter, ouvrir, partager et valoriser les données publiques, au bénéfice de la transparence et de l'efficacité de l'action publique.

Engagement # 7

10 démarches et 100 formulaires administratifs seront simplifiés significativement avec un impact mesurable en janvier 2022.

Engagement # 8

Un accès téléphonique sans surfacturation sera garanti pour tous les services publics. Tous les sites Internet publics afficheront un numéro de téléphone pour pouvoir être contactés par téléphone. Les réseaux s'engagent à converger vers un taux de décroché de 85 %.

Engagement # 9

Le partage d'informations entre administrations sera accéléré pour simplifier les démarches en ligne et éviter de demander de nombreuses fois la même chose aux Français.

L'après mission

- Engagement 6 du CITP: Une politique publique de la donnée ambitieuse mise en œuvre par tous les ministères
 - clairement positionnée à la suite du rapport Bothorel;
 - Annonce d'ouvertures de nouvelles bases de données;
 - Chaque ministère sera chargée de décliner une feuille de route ambitieuse en matière de valorisation des données s'appuyant sur une gouvernance renouvelée plus efficace;
 - Charte d'engagement sur le traitement des demandes de données des chercheurs;
 - Renforcement de la politique de ressources humaines dans le domaine de la donnée;
 - Mission de médiateur de la DIG

(extrait du dossier de presse du CITP)

Par ailleurs, de nouvelles bases de données, interfaces de partage de données et codes sources publics de qualité seront ouverts d'ici fin 2021. Parmi cette liste figurent notamment :

- les données de la carte scolaire des collèges ;
- les données sur les écoles, collèges et lycées (effectifs, langues vivantes, formations, spécialités, etc.) ;
- les données librement accessibles du répertoire opérationnel de ressources qui recense l'ensemble de l'offre sanitaire et du médico-social ;
- les données sur les auto-écoles ayant reçu un agrément préfectoral ;
- les données sur l'agriculture biologique à la maille communale ;
- les données des permis de construire des personnes physiques ;
- les données d'information routière temps réel fournies par les acteurs publics ;
- la liste des centres de formation d'apprentis avec indicateurs de performance ;
- le code source de FranceConnect ;
- le code source d'Affelnet Lycée ;
- la base de données des fichiers fonciers MAJIC pour les personnes morales ;
- le code source du calcul du prélèvement à la source ;
- les données de mise en œuvre du plan de relance ;
- la limite terre-mer haute résolution.

A venir..

- Circulaire PM concernant les recommandations générales
- Chaque ministère devra produire une feuille de route ambitieuse (données, codes sources et algorithmes)
- Création d'une mission de médiateur de la donnée d'intérêt général