



Le directeur, administrateur général des données,
des algorithmes et des codes sources,

Ref. : Note/NBH/Etalab

Paris, le 23/09/2021

NOTE

à

Mesdames et Messieurs les administrateurs ministériels des données,
des algorithmes et des codes sources

Objet : Cadre interministériel d'administration de la donnée.

Annexe : Détail de l'offre de services de la DINUM

En écho à la circulaire n°6264/SG du 27 avril 2021 relative à la politique publique de la donnée, des algorithmes et des codes sources, ce document précise le cadre interministériel dans lequel les administrations doivent s'inscrire. Il décrit également l'offre de service mise à leur disposition afin de concevoir et de mettre en œuvre à leur échelle, ainsi qu'à celle de leurs opérateurs, des stratégies ambitieuses en matière de données, d'algorithmes et de codes sources.

1. Vision et enjeux de l'action publique fondée sur la donnée

La donnée est l'actif stratégique de la révolution numérique, et l'administration doit s'en saisir pleinement, dans tous les champs de l'action publique.

La mise en place d'une politique publique de la donnée ambitieuse constitue un enjeu de transformation de l'action publique, de transparence et d'efficacité, mais aussi de création de valeur économique et de souveraineté.

Une action publique fondée sur la donnée est une action publique :

- **Plus transparente**
 - Les citoyens ont accès à des données analysables sur l'action de l'Etat et des organisations chargées d'une mission de service public,
 - Les citoyens sont informés des données utilisées par l'administration dans le cadre des démarches administratives et peuvent s'assurer de leur exactitude,
 - L'administration fait un usage responsable et explicable des données, et est en mesure de rendre compte des règles qui sous-tendent des décisions collectives et individuelles
- **Qui produit un meilleur service public**
 - Les services sont mieux adaptés aux besoins des usagers,
 - L'expérience utilisateur des démarches administratives est simplifiée et améliorée (dématérialisation, pré-remplissage, « dites-le-nous une fois », France Connect..),
 - Des tiers peuvent proposer des services à valeur ajoutée qui complètent l'action publique à partir des données mises à disposition en open data
- **Plus efficace**
 - Les ressources sont allouées plus efficacement, des processus métier sont optimisés, des tâches administratives sont allégées ou supprimées,
 - Les données peuvent être utilisées pour optimiser les ciblage et la lutte contre la fraude, et réduire la pression administrative sur les usagers qui se conforment à la réglementation,
 - La conception, le pilotage et l'évaluation des politiques publiques sont améliorés,
 - La connaissance est plus facilement partagée et se diffuse en dehors des silos,
 - La prise de décision collective et individuelle des agents publics est renforcée par un appui sur les faits et leur mise en regard des problématiques rencontrées
- **Plus aidante**
 - Le recours à la donnée doit permettre d'évaluer les situations de non-recours des citoyens et des entreprises, et de favoriser la proactivité de l'administration (information, notification, voire attribution automatique des droits),
- **Créatrice de valeur économique**
 - Les entreprises, les chercheurs, les associations... peuvent exploiter les données publiques pour améliorer leur activité ou pour innover, créant de la valeur,
 - La donnée publique, répartie sur tout le territoire, constitue une véritable infrastructure numérique pérenne et souveraine, socle de services publics et privés.

2. Objectifs pour les administrations

« **La donnée** » comprend l'ensemble des informations collectées et produites par les administrations ou par leurs opérateurs, pour leur usage propre ou pour les échanges avec les parties prenantes des politiques publiques qu'elles opèrent, qui peuvent faire l'objet d'un traitement informatique aisé. Ce concept inclut les algorithmes et les codes sources produits par l'administration.

Afin de concrétiser les ambitions précitées, chaque ministère devra en conséquence poursuivre résolument les objectifs suivants :

- 1) Assurer la **simplification des démarches administratives** dont il a la charge, en appliquant sur toutes ses procédures le principe du « Dites-le-nous une fois », consistant à utiliser les données déjà disponibles au sein de l'Etat plutôt que les redemander aux usagers, et contribuer lui-même à mettre ses données à disposition des autres administrations pour qu'elles-mêmes puissent appliquer ce même principe ;
- 2) **Mettre à disposition de manière spontanée, et en recherchant un niveau de qualité, d'interopérabilité, de disponibilité adaptés aux usages identifiés, les jeux de données, les algorithmes et les codes sources** qu'il produit, et intégrer le plus en amont possible cette finalité dans la conception des systèmes d'information et des services ;
- 3) **API-fier les bases de données** et garantir un niveau de **qualité de service** quant à la disponibilité de l'accès aux APIs ;
- 4) Participer à la mutualisation et à la montée en qualité des données essentielles, en contribuant à étendre les **données de référence** et à faciliter leur réutilisation par les autres administrations et la société civile,
- 5) Exploiter le plus possible les données grâce aux **datasciences et à l'intelligence artificielle** afin d'améliorer l'action et la décision publique ;
- 6) Contribuer à la mise en œuvre de **hubs de données sectoriels ou intersectoriels**, permettant notamment de faciliter l'accès à des données couvertes par des secrets légaux aux parties prenantes disposant des droits d'accès, et d'éclairer la prise de décision des autorités gouvernementales et administratives ;
- 7) Favoriser l'essor de la **recherche** en France par le développement de la connaissance partagée, en répondant sans délais à toute demande de chercheurs du secteur public souhaitant accéder à des jeux de données publiques.

3. Cadre d'action et offre de service de la DINUM

1) Sur le plan juridique

L'ouverture, le partage et le traitement des données sont encadrés par différentes dispositions législatives et réglementaires conciliant la diffusion, la réutilisation et l'exploitation des données avec la protection des données à caractère personnel et des données protégées par la loi. Les administrations sont chargées **d'identifier et, le cas échéant, de lever, en sollicitant si nécessaire l'expertise de la DINUM, les verrous juridiques et sécuritaires** qui entravent la circulation maximale et l'exploitation de la donnée dès lors que celles-ci ne porteraient pas atteinte aux secrets protégés par la loi.

2) Sur le plan de l'ouverture, du partage et de la consolidation des données

Les administrations doivent avoir connaissance du patrimoine de données produites ou collectées au sein de leur structure. Leur **inventaire** facilite l'identification des données qui ont vocation à être ouvertes, partagées ou exploitées dans le cadre de projets numériques. Les administrations pourront solliciter la DINUM afin d'être accompagnées sur le choix de **l'outil de catalogage**.

Afin de mutualiser les efforts et faciliter la réutilisation des données, les administrations doivent s'appuyer sur **trois plateformes de référence** et sur les ressources associées mises à disposition par la DINUM :

- La plateforme interministérielle **Data.gouv.fr** est le dispositif de référence pour recenser, en un point unique, l'ensemble des données ouvertes par l'Etat et ses partenaires, et favoriser leur réutilisation. Des déclinaisons thématiques ou sectorielles existent également (transports, adresse.data.gouv.fr...), ainsi que des outils facilitant la publication.
- La plateforme **Api.gouv.fr** recense l'ensemble des interfaces applicatives (API) permettant l'interconnexion automatisée entre deux systèmes d'information. La gestion des habilitations et des droits à en connaître est gérée par un dispositif unique, appelé **Datapass**, qui devra être utilisé par l'ensemble des producteurs de données pour faciliter l'accès aux données auprès des administrations habilitées.
- Le site **Code.gouv.fr** vise à référencer l'ensemble des codes sources produits et publiés par l'administration.

Elles sont invitées également à contribuer aux travaux visant à améliorer la **qualité et l'interopérabilité des données**, notamment grâce à des schémas de données (schema.data.gouv.fr). Ceux-ci garantissent de plus une meilleure homogénéité des données lorsqu'elles sont produites par plusieurs entités et en permettent l'agrégation.

Les administrations en charge de la production de données et qui exposent des API doivent **garantir une disponibilité de 98%** afin de permettre à l'ensemble des réutilisateurs de bénéficier d'un niveau de qualité acceptable.

La création de **hubs de données sectoriels, et à terme intersectoriels**, est encouragée dans la mesure où les initiatives de partage de la donnée dans un cadre de confiance sont bénéfiques

pour la recherche, l'innovation ou encore le pilotage des politiques publiques. La DINUM mettra en place un accompagnement interministériel en la matière et proposera une doctrine d'engagement.

3) Sur le plan de l'exploitation des données

Les administrations **identifient les leviers par lesquels l'exploitation des données peut les aider à mieux mener leurs politiques publiques**, que ce soit par une optimisation des processus grâce à la donnée et notamment aux algorithmes d'IA, une meilleure coordination des acteurs grâce aux données, une meilleure évaluation de la mise en œuvre et de l'impact des politiques publiques ou encore pour faciliter la relation à l'utilisateur grâce à la donnée. Elles peuvent solliciter un accompagnement de la DINUM (Lab IA d'Etalab notamment).

4) Sur le plan des codes sources

Les codes sources correspondent à la version humainement lisible d'un programme informatique. Dans l'administration publique, ils sont soit développés par des agents publics, auquel cas les droits appartiennent en général à l'administration, soit développés par des tiers, auquel cas c'est à l'administration acheteuse de décider des droits patrimoniaux qu'elle souhaite récupérer auprès du prestataire. Les codes sources dont l'administration détient les droits sont des documents administratifs communicables au public, dans le **cadre légal fixé par le Code des relations entre le public et l'administration (CRPA)**.

Les administrations dressent un état des lieux de la production et de la publication des codes sources dans leur périmètre et elles organisent le référencement des codes sources communicables sur **code.gouv.fr**. Elles identifient des **référénts pour accompagner leurs services et opérateurs sur l'ouverture des codes sources** et des **référénts aidant à renforcer le recours à des codes sources libres** dans l'administration, notamment ceux publiés par d'autres administrations.

Elles mettent en place une politique **d'ouverture par défaut** pour les nouveaux développements. L'ouverture rétroactive des codes sources fermés est abordée en priorisant selon deux critères : les enjeux en termes de transparence de l'action publique et l'existence, au sein de ces codes sources, de parties génériques réutilisables. **Les administrations identifient et publient la liste des codes sources correspondants.**

5) Sur le plan des algorithmes et des modèles

Les administrations dressent l'**inventaire** des traitements algorithmiques prenant des décisions administratives individuelles non couverts par l'article L311-5 du Code des relations entre le public et l'administration. Elles s'appuient notamment sur le [guide « Expliquer les algorithmes »](#). Les administrations listent les algorithmes et modèles d'apprentissage automatique qui pourraient avoir une **valeur d'usage** pour d'autres administrations.

6) Sur le plan de la coopération entre producteurs et réutilisateurs des données, d'API et des codes sources

Les administrations sont chargées d'identifier en lien avec leurs écosystèmes, les jeux de données, les API et les codes sources dont l'ouverture présenterait une **valeur ajoutée élevée**, ainsi que ceux déjà ouverts, mais dont la piètre qualité empêche une réutilisation aisée.

Chacun des producteurs de données, d'API et de codes sources est **responsable de l'animation de son écosystème** de contributeurs et d'utilisateurs : Il doit s'assurer de la vitalité de cet écosystème et veiller à ce que l'effort puisse être le plus largement réparti entre ses membres de l'écosystème, afin d'assurer la viabilité à long terme des communautés.

7) Sur le plan des règles de développement applicatif

Les équipes de développement informatique de l'Etat, ainsi que leurs prestataires, doivent privilégier les **approches de développement** permettant :

- de recourir, chaque fois que possible, à une approche de développement en microservices, qui permet de segmenter les applicatifs par grandes fonctions autonomes et de faciliter leur intégration et la réversibilité ;
- d'exposer les données qu'ils manipulent, en masse ou unitairement par API (qui seront référencées sur api.gouv.fr), pour favoriser leur extraction et l'interopérabilité des applications ;
- de coder et documenter l'application de façon à ouvrir le code par défaut dès le début du développement, sauf contrainte de propriété intellectuelle avérée et à l'exception des paramètres et secrets sur lesquels repose la sécurité applicative ;
- de publier les données et logiciels sous les licences légales, et de contribuer activement aux communautés open source en cas de recours aux logiciels libres.

8) Sur le plan de la gouvernance des projets

Les administrations devront chercher à construire des alliances avec d'autres administrations (Etat et collectivités territoriales) afin de mutualiser les efforts financiers et humains, et de réutiliser des solutions ou des codes sources déjà existants dans ces administrations.

Par ailleurs la gouvernance interne des ministères en matière de conduite des projets devra veiller à la bonne application des règles définies dans ce cadre interministériel.

La DINUM intégrera ces vérifications au **contrôle d'exécution et aux missions de conseil aux grands projets numériques de l'Etat**, tels que défini aux articles 3 et 4 du décret n° 2019-1088 du 25 octobre 2019.

9) Sur le plan de la sécurité et de la souveraineté

Les questions de sécurité et de souveraineté sont à considérer **dès le lancement** de tout projet afin de définir les niveaux de sécurité et de souveraineté adaptés aux données dont il est question.

Les administrations s'appuient notamment sur la stratégie cloud mise en place par l'Etat et doivent mettre en place une gouvernance en matière de sécurité des systèmes d'information permettant de lever les freins internes qui ne seraient pas dûment justifiés. Elles pourront notamment à cette fin, bénéficier du concours de l'ANSSI.

10) Sur le plan des ressources humaines expertes en matière de donnée

Les ministères sont chargés de construire une **stratégie de recrutement, fidélisation et formation** dans le domaine du numérique, leur permettant de développer leur expertise et favoriser la mise en œuvre des orientations fixées en matière de donnée.

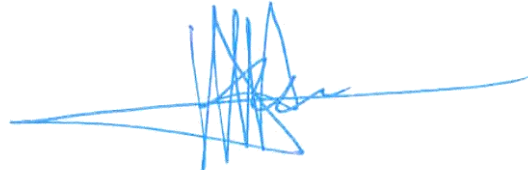
Ils peuvent s'appuyer utilement sur les actions conduites conjointement par la DINUM et la DGAFP visant à promouvoir les métiers du numérique au sein de l'Etat, et à développer l'attractivité de l'Etat afin de pouvoir recruter et préserver des expertises de haut niveau, indispensable à l'autonomie de l'Etat en matière de numérique. Ils pourront également s'appuyer sur l'expertise de l'INSEE, qui anime une filière professionnelle dédiée à la statistique et à la donnée.

Ils sont également incités à recourir, autant que possible, aux initiatives animées par la DINUM visant à recruter au sein de l'Etat des profils rares et experts en matière de données et de code informatique - entrepreneurs d'intérêt général, lab IA, brigade d'intervention numérique... - qui concourent à cette attractivité.

4. Gouvernance interministérielle

- **L'Administrateur général des données, des algorithmes et des codes sources** préside tous les semestres un Comité Interministériel de l'Administration de la Donnée (CIAD), rassemblant les Administrateurs ministériels des données, et autant que de besoin, les référents placés auprès des Préfets de région.
- Chaque Administrateur ministériel des données (AMD) est chargé de constituer et d'animer **un réseau de correspondants** au sein des directions de son périmètre ministériel.
 - o Chaque AMD est notamment responsable du référencement et de la tenue à jour des organisations et données relatives à son périmètre ministériel sur la plateforme data.gouv.fr, api.gouv.fr et code.gouv.fr
 - o Si l'AMD ne porte pas lui-même les sujets de code sources et d'algorithmes, il désigne un point de contact chargé de la mise en œuvre et de la promotion des actions ministérielles en la matière.
- Un **tableau de suivi des ouvertures annoncées et effectives** des jeux de données, API et codes sources a été publié¹ et sera tenu à jour sur une base trimestrielle.

¹ <https://www.data.gouv.fr/fr/datasets/tableau-de-suivi-des-ouvertures-de-donnees-codes-sources-et-api-publics-avril-2021/>



Nadi BOU HANNA

Directeur interministériel du numérique
Administrateur général des données, des algorithmes
et des codes sources

Copie à :

- Mesdames et messieurs les secrétaires généraux des ministères

Annexe : Détail de l'offre de services de la DINUM

La Direction interministérielle du numérique dispose en son sein du département Etalab, expert des problématiques associées, et anime de nombreuses initiatives contribuant directement aux finalités décrites, dont le programme TECH.GOUV (missions LABEL, DATA, TALENTS, PILOT, BETA notamment).

A/ Conseil et accompagnement

- Les services de la DINUM, et en particulier le département Etalab, se tiennent à la disposition des administrations pour les **accompagner** dans leurs initiatives relatives aux données, algorithmes et codes sources.
- Elles peuvent directement se référer aux **guides thématiques** produits par Etalab et disponibles sur <https://guides.etalab.gouv.fr> :
 - o [Préparer les données à l'ouverture et la circulation](#)
 - o [Identifier les données à ouvrir](#)
 - o [Publier les données sur data.gouv.fr](#)
 - o [Publier des réutilisations sur data.gouv.fr](#)
 - o [Expliquer les algorithmes publics](#)
 - o [Ouvrir les codes sources](#)
 - o [Créer des schémas de données](#)
 - o [Pseudonymiser des documents grâce à l'IA](#)
- Guichet « Dites-le nous une fois » et accompagnement : <https://www.numerique.gouv.fr/services/guichet-dites-le-nous-une-fois/>
- Formation : en offre en cours de développement par la DINUM

B/ Plateformes et outils

- data.gouv.fr et verticales thématiques : transport.data.gouv.fr, cadastre.data.gouv.fr, entreprise.data.gouv.fr, adresse.data.gouv.fr...
- Outils d'aide à la production et publication de données de qualité : <http://schema.data.gouv.fr> et <https://publier.etalab.studio/>
- [API.gouv.fr](https://api.gouv.fr) et solution DataPass <https://datapass.api.gouv.fr/>, « hubs » d'API thématiques : entreprise.api.gouv.fr et particulier.[api.gouv.fr](https://particulier.api.gouv.fr)
- code.gouv.fr (précédemment code.etalab.gouv.fr), et sill.etalab.gouv.fr

C/ Facilités relatives à l'incubation et au développement de projets, soutien RH et financier

- Accompagnement des **projets de l'investigation à l'autonomisation** : programmes Beta (beta.gouv.fr), Entrepreneurs d'Intérêt Général (EIG)
- Aide à la **mobilisation de compétences** : Programme EIG, Brigade d'Intervention numérique
- **Appui financier** :
 - o Programme EIG : <https://entrepreneur-interet-general.etalab.gouv.fr/>
 - o Fonds FAST : <https://beta.gouv.fr/approche/fast>
 - o Guichets ITN5 et ITN7 du Plan de Relance : <https://france-relance.transformation.gouv.fr/>