

**ANNALES 2023**

**SELECTION PROFESSIONNELLE**

***CORPS DES INGENIEURS 2<sup>e</sup> GRADE-ITM***  
***(Domaine d'activité de la météorologie)***

**DU STATUT PARTICULIER DES**  
**PERSONNELS TECHNIQUES DE LA**  
**NOUVELLE-CALÉDONIE**

**EXAMEN PROFESSIONNEL POUR LE RECRUTEMENT  
D'INGÉNIEURS DES TRAVAUX DE LA MÉTÉOROLOGIE  
SESSION 2023**

\*\*\*\*\*

**ÉPREUVE ÉCRITE OBLIGATOIRE D'ADMISSIBILITÉ  
NOTE DE PROBLÉMATIQUE**

Durée : 3 heures

Coefficient : 1

Outre l'appréciation du fond, l'évaluation de la copie tient compte de la présentation, de l'orthographe, de la ponctuation, de la correction de la langue et du respect des consignes.

Par ailleurs, le jury appréciera une écriture lisible.

L'utilisation de toute documentation (hors dossier joint) sur support papier ou électronique n'est pas autorisée.

Il est rappelé que cette épreuve vise à apprécier les qualités rédactionnelles du candidat, sa capacité de raisonnement, sa capacité à mobiliser ses connaissances mais aussi à comprendre les textes fournis et à en extraire les informations pertinentes, sa capacité à cerner les enjeux et à les replacer dans une cohérence générale, sa capacité à structurer une analyse, éventuellement à exprimer des propositions et à les argumenter.

**Sujet : 2 pages**

**Dossier support : 30 pages**

A partir de vos connaissances et des documents figurant dans le dossier joint, il vous est demandé d'établir une note de problématique d'une longueur indicative recommandée de 4 pages.

**MISE EN SITUATION :**

La digitalisation croissante de notre société s'appuie sur l'exploitation de plus en plus massive de données, notamment grâce à des données de plus en plus nombreuses et accessibles, et grâce à l'évolution des technologies (ex : big data, intelligence artificielle).

Dans ce contexte, des politiques publiques sont mises en place pour organiser l'ouverture de certaines données, et ainsi favoriser de nouveaux usages. Des évolutions importantes ont ainsi récemment eu lieu dans la politique de partage de données, dans des domaines qui concernent directement les activités de Météo-France.

Vous êtes ITM travaillant à la DSM, et votre Directeur est sollicité pour intervenir lors d'un colloque dans une table-ronde consacrée à l'ouverture des données et aux nouveaux usages que cela ouvre.

Il vous est ainsi demandé de rédiger une note de problématique, destinée à votre Directeur, exposant les enjeux de l'ouverture des données et de la massification des échanges de données, et leurs conséquences pour Météo-France et pour son environnement concurrentiel ; vous y indiquerez notamment les éléments de langage qu'il vous semble les plus importants à faire valoir publiquement lors du colloque.

- Dans un premier temps, vous présenterez les **principaux enjeux des politiques publiques** en matière d'ouverture des données (leurs grands objectifs, les tendances ou les principales évolutions récentes notamment au niveau national ou européen).
- Dans un second temps, vous présenterez les **principales implications** de cette ouverture et de cette massification des données, **concernant la mise à disposition de données par Météo-France**, notamment sur le plan budgétaire et sur le plan opérationnel.
- Enfin, vous présenterez quelques **implications dans le domaine des services liés à la météo** (ex : paysage concurrentiel, nouvelles offres de service autour de la donnée, utilisations innovantes des données, partenariats...), en soulignant les **enjeux et opportunités que vous identifiez pour Météo-France**.

#### Liste des documents du dossier :

**DOCUMENT 1** : Extrait du Contrat d'Objectifs et de Performance (COP) de Météo-France pour la période 2022-2026 : objectif #14. [1 page]

**DOCUMENT 2** : Circulaire du Premier Ministre du 27 avril 2021 concernant la politique publique de la donnée, des algorithmes et des codes sources. [3 pages]

**DOCUMENT 3** : Synthèse de la directive (UE) 2019/1024 concernant les données ouvertes et la réutilisation des informations du secteur public. [3 pages]

**DOCUMENT 4** : FAQ de la Commission Européenne sur les jeux de données à haute valeur (HVD), 2023. [2 pages]

**DOCUMENT 5** : Règlement d'exécution du 21/12/2022 de la Commission européenne établissant une liste d'ensembles de données de forte valeur spécifiques et les modalités de leur publication et de leur réutilisation (EXTRAIT concernant les HVD météorologiques) [2 pages]

**DOCUMENT 6** : Article sur la politique unifiée de données de l'OMM (2021, extraits). [5 pages]

**DOCUMENT 7** : Extraits d'une note DG interne (été 2021) sur les impacts des évolutions prévues. [4 pages]

**DOCUMENT 8** : Extrait du rapport d'information sur Météo-France, fait au nom de la commission des finances du Sénat, par M. Vincent CAPO-CANELLAS, sénateur et « rapporteur » pour la commission. Septembre 2021 : § « le développement de l'open data s'accélère ». [3 pages]

**DOCUMENT 9** : Extraits d'une présentation interne sur les évolutions envisagées dans notre offre de données publiques (2022). [1 page]

**DOCUMENT 10** : Généralités sur les API (extraits issus d'internet). [1 page]

**DOCUMENT 11** : Extraits de la présentation du Plan d'Actions Commerciales 2023 de Météo-France (janvier 2023). [1 page]

**DOCUMENT 12** : Actualité EUMETSAT (juillet 2021) concernant un programme pilote innovant visant à acheter, pour la toute première fois, des données à un opérateur commercial de flotte de satellites à des fins de prévision météorologique. [1 page]

**DOCUMENT 13** : Actualité intramet : « Une mission dédiée à l'intelligence artificielle vient d'être lancée au sein de l'établissement » (mars 2021). [2 pages]

**DOCUMENT 14** : Actualité intramet : « Lancement de l'incubateur de start-up 'MétéoFab' de Météo-France » (juillet 2021). [1 page]

## DOCUMENT 1 :

### Objectif #14

#### Servir la croissance et l'innovation par l'accès aux données publiques de l'établissement susceptibles de créer de la valeur, notamment via la poursuite de leur mise en ligne et le développement d'API

L'action de Météo-France s'inscrit, depuis près de dix ans, dans la politique volontariste de l'État d'ouverture des données publiques. Ainsi, depuis 2012, Météo-France maintient un portail de données publiques (<https://donneespubliques.meteofrance.fr>) permettant d'accéder à plus de 30 jeux de données. Mi-2021, ce sont d'ores et déjà de l'ordre de 0,6 To de données qui sont mis en ligne chaque jour sur le site de données publiques de l'établissement et près de 1,5 To de données sont téléchargés quotidiennement. Au total, le portail de données publiques de l'établissement permet d'accéder à près de 15 To de données publiques. Les volumes de données publiques, accessibles sur demande, en ligne ou via API, ont vocation à augmenter très fortement.

Météo-France ne percevra plus de redevances de réutilisation de ses données publiques, dans un contexte de forte augmentation du volume de ces données, conformément à la circulaire du Premier ministre N° 6264/SG du 27 avril 2021. L'établissement poursuivra sa stratégie d'amélioration des conditions techniques d'accès à ces données, via des développements de l'accès direct en ligne et d'API à un rythme compatible avec les moyens humains et financiers dont il dispose pour assurer cette mission et en fonction des enjeux liés à la réutilisation des différentes catégories de données. En particulier, Météo-France mettra en ligne les jeux de données qualifiés de jeux de données à forte valeur dans les conditions prévues par la 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. Une action spécifique sera menée auprès de ses partenaires institutionnels, et notamment en région, afin de mieux faire connaître le portail de données publiques de l'établissement et

favoriser la réutilisation de ses données auprès de cette catégorie d'utilisateurs.

Météo-France cherchera également à faciliter le croisement de ses données avec d'autres catégories de données météorologiques et climatiques issues de réseaux de partenaires, et en participant, y compris par la recherche des financements dédiés nécessaires, à de possibles initiatives de plates-formes de mutualisation de données, notamment environnementales permettant de croiser données météorologiques et autres catégories de données (débits, ressources en eau, état du sol, qualité de l'air, données sanitaires, épidémiologiques..).

Enfin, Météo-France poursuivra sa politique d'ouverture de ses codes de recherche, excepté ceux faisant l'objet de droits de propriété intellectuelle détenus par des tiers.

#### INDICATEUR

- **Indicateur 14.1** : volume de données publiques mises en ligne sur le portail des données publiques de Météo-France.

# DOCUMENT 2 :

*Le Premier Ministre*

Paris, le 27 avril 2021

N° 6264/SG

à

Mesdames et Messieurs les ministres,  
Mesdames et Messieurs les ministres délégués,  
Mesdames et Messieurs les secrétaires d'État,  
Mesdames et Messieurs les préfets de région

## **Objet : Politique publique de la donnée, des algorithmes et des codes sources**

Au cours des dernières années, la France s'est dotée d'un cadre juridique à l'avant-garde en matière d'accès aux documents administratifs et de réutilisation des informations publiques<sup>1</sup>, en instaurant l'ouverture libre, gratuite et par défaut de toutes les données dont la publication représente un intérêt économique, social, sanitaire ou environnemental, dans le respect de la protection des données à caractère personnel et des secrets protégés par la loi.

La crise sanitaire sans précédent que nous connaissons vient démontrer que l'ouverture et l'utilisation des données publiques, relatives notamment au suivi de l'épidémie, sont indispensables à l'efficacité de l'action publique, au suivi de l'action gouvernementale et au débat démocratique.

Lors de la convention managériale de l'État du 8 avril 2021, le Président de la République a appelé de ses vœux la mise en place d'une politique publique de la donnée ambitieuse qui constitue un enjeu de souveraineté, mais également de transformation de l'action publique.

L'exploitation des données des administrations permet ainsi une meilleure évaluation des politiques publiques et la simplification des relations entre les usagers et le service public. À cet égard, le baromètre des résultats de l'action publique, publié en janvier et qui sera régulièrement enrichi et actualisé, donne à voir aux Français les résultats des politiques prioritaires et constitue un outil de pilotage de l'action de l'État au niveau territorial. Enfin, la diffusion, le partage et la valorisation des données et des algorithmes soutiennent l'innovation, la recherche, la création de valeur et le développement de nouveaux usages, comme en matière d'intelligence artificielle.

Or, si notre pays figure en tête des classements internationaux, le rapport du député Éric Bothorel remis en décembre 2020 a souligné la nécessité de redoubler nos efforts, dans un contexte de transformation numérique croissante de nos sociétés.

Dans ce contexte, comme j'en ai pris l'engagement à l'occasion du comité interministériel de la transformation publique du 5 février 2021, je souhaite que le Gouvernement porte une ambition renouvelée en matière d'exploitation, d'ouverture et de circulation des données, des algorithmes et des codes sources publics au profit des usagers, des chercheurs, des innovateurs et de l'ensemble de nos concitoyens.

.../...

---

<sup>1</sup> Notamment le code des relations entre le public et l'administration, tel que modifié par la loi n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public et par la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

I. **La politique de la donnée doit constituer une priorité stratégique de l'État dans ses relations avec tous ses partenaires, notamment les collectivités territoriales et les acteurs privés**

Cette ambition nécessite que vous vous y impliquiez personnellement.

Les administrations placées sous votre autorité devront rechercher en permanence la meilleure circulation de la donnée, des algorithmes et des codes, dans des formats ouverts et exploitables par les tiers. Pour ce faire, elles doivent mobiliser les dispositifs techniques disponibles comme l'accès sécurisé aux données, la pseudonymisation, l'anonymisation ou l'agrégation.

Cette ambition renouvelée implique, en outre, un renforcement de l'ouverture des codes sources et des algorithmes publics, ainsi que de l'usage du logiciel libre et ouvert, et l'extinction, à horizon 2023, des redevances perçues pour la réutilisation des données notamment sur le fondement de l'article L. 324-1 du code des relations entre le public et l'administration.

Il vous appartient également de mettre en place l'organisation la plus adaptée pour favoriser les synergies entre les services chargés du numérique dans votre ministère, notamment les services statistiques. À cette fin, vous nommerez d'ici le 15 mai un administrateur ministériel des données, chargé d'élaborer la stratégie de votre ministère dans ce domaine, de coordonner les parties prenantes et d'être le point de contact des utilisateurs des données et des applications numériques relevant de votre périmètre. Son positionnement hiérarchique devra lui permettre d'assurer ses missions en lien étroit avec les services placés sous votre autorité.

Le directeur interministériel du numérique, en sa qualité d'administrateur général des données, algorithmes et codes sources, en assurera la coordination, en s'appuyant sur le département Etalab. La direction interministérielle du numérique (DINUM) peut à cet égard être saisie, par toute personne morale de droit public, de toute question portant sur la circulation des données. Je vous encourage en particulier à la saisir pour accélérer le partage de données entre administrations, qui souffre encore trop souvent de freins et de lenteurs, alors qu'il est un gage indispensable de l'efficacité de l'action publique.

L'ouverture et la circulation des données, déjà pratiquées dans les territoires, est un enjeu majeur de simplification et de transparence de l'action publique locale, qui doit être promu et accompagné. Je souhaite ainsi qu'un référent « données, algorithmes et codes sources » soit désigné auprès de chaque préfet de région pour accompagner les services déconcentrés de l'État dans la mise en œuvre de ces orientations.

De même, la ministre de la cohésion des territoires et des relations avec les collectivités territoriales sera chargée d'animer la coopération entre l'État et les collectivités territoriales en matière de circulation des données, en s'appuyant notamment sur l'Agence nationale de la cohésion des territoires.

S'agissant enfin des relations avec les acteurs privés détenant des données dites d'« intérêt général » à forte valeur ajoutée pour la puissance publique, une mission de préfiguration de la fonction de médiateur de la donnée d'intérêt général sera lancée et rendra ses travaux avant le 1<sup>er</sup> décembre 2021.

En lien avec la ministre de l'enseignement supérieur, de la recherche et de l'innovation, la ministre de la transformation et de la fonction publiques établira, d'ici le 1<sup>er</sup> juillet 2021, une charte d'engagement de l'État pour accélérer le traitement des demandes des chercheurs souhaitant accéder à des données ayant vocation à être mises à disposition du public et en particulier réutiliser des algorithmes et des codes sources.

.../...

Enfin, je souhaite que vous animiez les écosystèmes de réutilisateurs des données et codes sources de votre ministère et des éventuels contributeurs, afin d'identifier les jeux de données à forte valeur ajoutée, de définir d'éventuels standards et d'alimenter le partage. Le forum du partenariat pour un gouvernement ouvert constituera l'une des enceintes à privilégier pour ces échanges.

## **II. L'année 2021 doit poser les fondements d'une politique ambitieuse de la donnée, des algorithmes et des codes sources dans chacun de vos ministères**

Afin de garantir la mise en œuvre de ces orientations, une stratégie interministérielle pour la filière numérique incluant les compétences liées aux données et aux algorithmes me sera présentée par la ministre de la transformation et de la fonction publiques d'ici le 15 mai 2021.

Vous élaborerez une feuille de route présentant la stratégie de votre ministère, qui devra m'être transmise, ainsi qu'à la ministre de la transformation et de la fonction publiques, d'ici le 15 juillet 2021 et sera rendue publique d'ici le 15 septembre 2021. Vous vous attacherez dans cette feuille de route à développer les compétences relatives aux données, aux algorithmes et aux codes sources au sein de vos administrations, notamment pour les cadres dirigeants, supérieurs et intermédiaires.

Dans les feuilles de route des services placés sous votre autorité, vous veillerez à intégrer systématiquement des objectifs relatifs au pilotage, à l'ouverture, à la circulation et au partage des données, des algorithmes et des codes sources. Ces objectifs devront constituer des critères d'évaluation de l'action de vos directeurs d'administration centrale et des chefs de services déconcentrés. Je vous demande également d'inclure des clauses analogues dans les contrats conclus avec les opérateurs et établissements publics placés sous la tutelle de votre ministère et d'intégrer cette dimension dans les relations contractuelles avec les acteurs privés chargés, sous votre autorité, d'une mission de service public (entreprises, professions réglementées, associations).

De plus, vous vous assurerez que les données ouvertes de votre ministère sont référencées sur les portails interministériels [data.gouv.fr](https://data.gouv.fr) et [api.gouv.fr](https://api.gouv.fr). La DINUM assure la gestion et la montée en gamme de ces portails et définira des standards de qualité et d'interopérabilité concernant les données qui y sont exposées. Un label des services producteurs de données sera créé pour soutenir les initiatives des administrations.

Une mission dédiée à l'animation et la promotion interministérielles en matière de logiciel libre et de communs numériques sera mise en place au sein de la DINUM et créera le portail interministériel [code.gouv.fr](https://code.gouv.fr). Vos administrations pourront s'appuyer sur cette mission, ainsi que sur l'Agence nationale de la sécurité des systèmes d'information pour les enjeux relatifs à la cybersécurité, pour les accompagner dans l'ouverture de leurs algorithmes et de leurs codes sources.

\*

Afin d'accélérer la mise à niveau numérique de l'administration, soutenue par les moyens de France Relance, je demande à la ministre de la transformation et de la fonction publiques de vous apporter son concours dans la mise en œuvre de cette politique de la donnée et d'assurer le suivi de ces orientations. Je ferai un point régulier sur son avancement en comité interministériel de la transformation publique.



Jean CASTEX

# DOCUMENT 3 :

## SYNTHÈSE de la directive (UE) 2019/1024

concernant les données ouvertes et la réutilisation des informations du secteur public

### QUEL EST L'OBJET DE CETTE DIRECTIVE?

- Elle établit le cadre juridique pour la réutilisation\* des informations du secteur public, notamment les informations dans les domaines géographique, cadastral, statistique ou juridique détenues par les organismes du secteur public\* ou par les entreprises publiques\*, ainsi que des données issues de la recherche financée par le secteur public.
- Elle vise à stimuler le potentiel socio-économique des informations du secteur public, par exemple en les rendant plus accessibles aux jeunes pousses et aux petites et moyennes entreprises, en accroissant la fourniture de données dynamiques\* et d'ensembles de données ayant une incidence économique particulièrement importante et en promouvant la concurrence et la transparence sur le marché de l'information.
- Elle fait partie d'un ensemble de mesures destinées à renforcer l'économie des données de l'Union européenne (UE), y compris le développement de l'intelligence artificielle.
- Aussi appelée «directive données ouvertes», elle refond la directive 2003/98/CE en l'abrogeant à compter du 17 juillet 2021.

### POINTS CLÉS

La refonte de la directive se fonde sur le principe général selon lequel les données publiques et obtenues grâce à un financement public devraient être réutilisables à des fins commerciales ou non commerciales.

#### **Données ouvertes**

La directive encourage l'utilisation des données ouvertes (données présentées dans des formats ouverts\* qui peuvent être librement utilisées et partagées par tous, quelle qu'en soit la finalité). Les organismes du secteur public et les entreprises publiques doivent mettre à disposition leurs documents dans tout format ou tout langage préexistant et, s'il y a lieu, sous forme électronique dans des formats ouverts, lisibles par machine, accessibles, traçables et réutilisables, en les accompagnant de leurs métadonnées.

#### **Dispositions pratiques permettant la réutilisation**

- Les organismes du secteur public doivent, s'il y a lieu sous forme électronique, traiter les demandes de réutilisation de documents, en les mettant à disposition dans un délai raisonnable.
- Dans le même temps, ils doivent prendre les dispositions nécessaires afin de faciliter la recherche en ligne et la localisation des documents qu'ils détiennent.
- Les pays de l'UE doivent également faciliter une réutilisation efficace des documents, en particulier en fournissant des informations sur les droits énoncés dans la directive et en offrant assistance et conseils.

#### **Données dynamiques et en temps réel**

Les données dynamiques doivent être mises à disposition pour une réutilisation immédiate aussitôt qu'elles ont été recueillies par l'intermédiaire d'une interface de programme d'application (API) et, le cas échéant, sous la forme d'un téléchargement de masse.

## **Données issues de la recherche**

- Les pays de l'UE doivent adopter des politiques et des mesures visant à mettre à disposition les données issues de la recherche financée par le secteur public de manière ouverte, conformément au principe d'«ouverture par défaut», et soutenir la diffusion de données de la recherche qui soient traçables, accessibles, interopérables et réutilisables (principe «FAIR»).
- Il y a lieu de tenir compte des préoccupations liées aux droits de propriété intellectuelle, à la protection des données à caractère personnel et à la confidentialité, à la sécurité et aux intérêts commerciaux légitimes dans le respect du principe «aussi ouvert que possible, mais aussi fermé que nécessaire».
- Les données issues de la recherche financée par le secteur public peuvent être réutilisées à des fins commerciales et non commerciales dans les cas où elles sont déjà mises à la disposition du public par l'intermédiaire d'une archive ouverte institutionnelle ou thématique.

## **Commerce équitable et non-discrimination**

- La réutilisation des documents est ouverte à tous sur le marché et toute condition de réutilisation applicable doit être non discriminatoire.
- En règle générale, les accords conclus entre les organismes du secteur public ou les entreprises publiques détenant les documents et des tiers ne peuvent attribuer de droits d'exclusivité.
- Dans les cas très strictement définis où la directive autorise la conclusion de tels accords, leur validité est soumise à un réexamen régulier et des conditions de transparence particulières s'appliquent.

## **Ensembles de données de forte valeur**

La Commission européenne a la possibilité d'adopter une liste d'ensembles de données de forte valeur qui doivent être mis à disposition dans des formats lisibles par machine et gratuitement par l'intermédiaire d'API. Ces ensembles de données seront sélectionnés dans six catégories thématiques énoncées à l'annexe I:

- géospatiales;
- observation de la terre et environnement;
- météorologiques;
- statistiques;
- entreprises et propriété d'entreprises; et
- mobilité.

De nouvelles thématiques peuvent être ajoutées par la Commission, par l'intermédiaire d'un acte délégué.

## **Redevances**

- La réutilisation de documents devrait être gratuite. Néanmoins, le recouvrement des coûts marginaux occasionnés pour reproduire et diffuser les documents, anonymiser les données personnelles et protéger les informations confidentielles à caractère commercial peut être autorisé.
- De façon exceptionnelle, les institutions culturelles, les entreprises publiques et certains organismes du secteur public qui remplissent les conditions énoncées à l'article 6, paragraphe 2, peuvent facturer la mise à disposition des informations, au-delà du principe général par défaut, afin de recouvrir les coûts éligibles.

## **Exceptions**

La directive ne s'applique pas aux:

- documents dont les droits de propriété intellectuelle sont détenus par des tiers;
- documents dont l'accès est exclu ou restreint en raison d'un régime national particulier ou lorsqu'ils relèvent de la protection des infrastructures critiques;
- documents dont la fourniture ne relève pas de la mission de service public dévolue à un organisme du secteur public ou dont la production ne relève pas de la fourniture de services d'intérêt général par une entreprise publique;
- documents liés aux activités d'une entreprise publique directement exposée à la concurrence et par conséquent non soumise aux règles relatives aux marchés publics en vertu de l'article 34 de la directive 2014/25/UE;
- autres documents mentionnés à l'article 1<sup>er</sup>, paragraphe 2, de la directive.

## **DEPUIS QUAND CETTE DIRECTIVE S'APPLIQUE-T-ELLE?**

Elle est en vigueur depuis le 16 juillet 2019. Elle refond et remplace la directive 2003/98/CE (et ses modifications ultérieures) à compter du 17 juillet 2021. Elle doit être intégrée dans le droit national des pays de l'UE au plus tard à la même date.

## **TERMES CLÉS**

*Réutilisation: l'utilisation par des personnes ou des entités juridiques de documents détenus par des organismes du secteur public ou des entreprises publiques.*

*Organisme du secteur public: l'État, les autorités régionales ou locales, les organismes de droit public ou les associations formées par ces mêmes autorités ou organismes de droit public.*

*Entreprise publique: toute entreprise sur laquelle les organismes du secteur public exercent une influence dominante du fait de la propriété de l'entreprise, de la participation financière qu'ils y détiennent ou des règles qui la régissent; les entreprises publiques incluent celles qui agissent au titre d'opérateurs de transport ferroviaire ou routier de voyageurs, de transporteurs aériens et d'armateurs de l'UE remplissant des obligations de service public.*

*Données dynamiques: des documents se présentant sous forme numérique, faisant l'objet d'actualisations fréquentes ou en temps réel à cause de leur volatilité ou de leur obsolescence rapide; typiquement les données émanant de capteurs.*

*Format ouvert: un format de fichier indépendant des plates-formes utilisées et mis à la disposition du public sans restriction empêchant la réutilisation des documents.*

Source : <https://eur-lex.europa.eu/legal-content/FR/LSU/?uri=CELEX%3A32019L1024>

Éléments publiés par la Commission Européenne,  
après l'adoption en 2019 de la directive 2019/1024 « concernant les données ouvertes et la réutilisation des informations du secteur public »,  
et après l'adoption du règlement d'exécution 2023/138 de la Commission du 21 décembre 2022 « établissant une liste d'ensembles de données de forte valeur spécifiques et les modalités de leur publication et de leur réutilisation »

### Que sont les jeux de données à haute valeur (High-Value Datasets, HVD), et pourquoi sont-ils importants ?

- Certaines données produites par le secteur public sont particulièrement intéressantes pour les créateurs de services et d'applications à valeur ajoutée. Par exemple, la réutilisation de jeux de données tels que la mobilité ou la géolocalisation des bâtiments peut ouvrir des opportunités commerciales pour les secteurs de la logistique ou du transport, ainsi qu'améliorer l'efficacité de la prestation de services publics, par exemple en comprenant les flux de circulation pour rendre les transports plus efficaces.
- Les ensembles de données tels que les données d'observation météorologique, les données radar, les données sur la qualité de l'air et la contamination des sols et les données sur le niveau de bruit peuvent également soutenir la recherche et l'innovation numérique ainsi que l'élaboration de politiques mieux informées, notamment dans la lutte contre le changement climatique et ses impacts ; cela a un impact positif sur la qualité de vie.
- La recherche et l'expérience ont montré que, pour que les informations du secteur public aient un impact positif sur l'économie, les données publiques devraient être disponibles par défaut le plus largement possible. Cependant, la persistance d'obstacles techniques, juridiques et financiers conduit à une situation dans laquelle les informations du secteur public en Europe ne sont pas suffisamment utilisées, malgré le fait que la politique et la législation de l'UE soient en place depuis près de 20 ans et que des investissements importants aient été réalisés au niveau national. La directive sur les données ouvertes renforce les règles existantes sur les formats, permettant divers scénarios de réutilisation, y compris les données dynamiques fournies en temps réel. La directive s'attaque aux pratiques exclusives qui restreignent la disponibilité des données du secteur public au-delà des accords d'exclusivité explicites. Les nouvelles règles laissent des limites plus strictes à la tarification, éliminant les coûts prohibitifs comme obstacle à la réutilisation.
- Pour lever ces obstacles restants, la directive sur les données ouvertes a obligé la Commission européenne à adopter un règlement d'application spécifiant des HVD concrètes. Les organismes du secteur public devront mettre ces HVD à disposition gratuitement, dans un format lisible par machine, via des interfaces de programmation d'applications (API) et, le cas échéant, sous forme de téléchargement en masse.
- La directive sur les données ouvertes définit les HVD comme "des documents détenus par un organisme du secteur public, dont la réutilisation est associée à des avantages importants pour la société, l'environnement et l'économie". Les HVD seront réutilisables à toutes fins (comme c'est le cas pour toutes les données ouvertes).

### De quels types de données sont les HVD ?

- La directive sur les données ouvertes définit six catégories de HVD : géospatiales, observation de la Terre et environnement, météorologiques, statistiques, entreprises et propriété des entreprises, et mobilité.
- En principe, la liste des jeux de données à haute valeur ajoutée ne fait pas référence aux données personnelles. Toutefois, si certains jeux de données sont considérés comme des données à caractère personnel dans les États membres ou si les États membres choisissent d'étendre la liste aux données à caractère personnel, la conformité au RGPD doit être assurée.
- La directive "Open Data" permet à la Commission d'étendre cette gamme thématique initiale par le biais d'un acte délégué si nécessaire pour refléter les évolutions technologiques et du marché.

### Comment ont été identifiés les jeux de données ?

- Afin de choisir les ensembles de données spécifiques dans chacune des six catégories, la Commission a consulté un large éventail de parties prenantes représentant à la fois les détenteurs et les utilisateurs de données.

- Les entretiens avec les parties prenantes ont permis de dresser une première liste de souhaits concernant les ensembles de données considérés comme ayant une grande valeur (d'un point de vue économique, sociétal et de réutilisation).
- En parallèle, l'étude d'appui à l'analyse d'impact a recensé toute la législation pertinente au niveau européen afin d'identifier les champs de données déjà couverts par les règles de l'UE. L'étude a ensuite évalué dans quelle mesure chaque ensemble de données pouvait effectivement être considéré comme ayant une "grande valeur", conformément aux critères d'évaluation prévus par la directive sur les données ouvertes.

### **Comment la mise en œuvre de la réglementation sur les HVD va-t-elle affecter les européens, et quels sont les cas d'usage possibles ?**

- Les HVD réduiront considérablement les obstacles à l'entrée sur le marché européen des données et augmenteront le volume des ensembles de données réutilisés. Cela stimulera à son tour la création de nouveaux services numériques et l'amélioration des services ou des processus commerciaux existants.
- Une meilleure disponibilité de divers ensembles de données statistiques (marché du travail, démographie, production industrielle) permettra de prévoir plus facilement l'impact d'éventuelles mesures politiques. En outre, l'ouverture des données sur les entreprises augmentera la transparence du marché, ce qui permettra une meilleure répartition des investissements privés ou des aides publiques.
- Une plus grande disponibilité des informations sur les entreprises présente des avantages sociaux évidents dans des domaines tels que la lutte contre la criminalité (y compris la criminalité financière), un engagement public accru et la transparence des transactions économiques.
- Des initiatives récentes visant à endiguer la pandémie de COVID-19 ont également montré l'importance de la géolocalisation des hôpitaux, des événements et des populations.

### **Quels sont les impacts sur les PME et la compétitivité ?**

- Les HVD ont été conçus au profit des PME et des jeunes entreprises, qui restent souvent exclues du marché car elles ne disposent pas des ressources humaines et financières suffisantes pour acquérir et améliorer la qualité des données du secteur public.
- L'augmentation de l'offre de données stimulera l'esprit d'entreprise et entraînera la création de nouvelles sociétés. En outre, les HVD peuvent devenir un outil important pour les jeunes entreprises afin de valider leurs analyses de rentabilité et d'attirer les investisseurs.
- Les données ouvertes peuvent également renforcer les PME, pour lesquelles les solutions logicielles utilisant les données du secteur public peuvent devenir leur principal produit ou service.
- Enfin, la mise à disposition gratuite des données contribuera à rééquilibrer la position des PME par rapport aux grandes entreprises technologiques, pour lesquelles le prix de l'acquisition des données ne constitue pas un obstacle important.

### **La fourniture de HVD entraînera-t-elle une charge importante pour les organismes publics ?**

- Les autorités publiques devront uniquement mettre à disposition les données existantes. Aucune obligation ne leur sera faite de commencer à produire de nouvelles données.
- Le secteur public supportera les coûts liés à la nécessaire mise à niveau technique de la publication des données (via les API) et devra cesser de facturer des frais pour la réutilisation des HVD.
- Cet impact devrait être atténué par les avantages tirés d'une meilleure disponibilité et d'une plus grande facilité d'utilisation des données au sein du secteur public et par la réduction de la charge administrative liée au traitement des demandes de réutilisation, aux négociations relatives aux licences et au traitement des redevances. La Commission soutient également les organismes du secteur public par des actions concrètes dans le cadre du programme Digital Europe.
- La liste des HVD a été établie en collaboration avec les États membres et a été préparée après une analyse d'impact approfondie. En tant que telle, elle ne contient que des données dont la réutilisation générera des avantages qui compenseront largement les coûts liés à leur mise à disposition gratuite.

# DOCUMENT 5 :

## RÈGLEMENT D'EXÉCUTION (UE) 2023/138 DE LA COMMISSION

du 21 décembre 2022

établissant une liste d'ensembles de données de forte valeur spécifiques  
et les modalités de leur publication et de leur réutilisation

(EXTRAIT de l'annexe du règlement)

### 3. MÉTÉOROLOGIQUE

#### 3.1. Ensembles de données pris en compte

La catégorie thématique «météorologique» inclut les ensembles de données relatifs à des données d'observation mesurées par les stations météorologiques, des observations validées (données climatiques), des alertes météorologiques, des données radar et des modèles de données de prévision météorologique numérique (PMN) présentant la granularité et les attributs clés figurant dans le tableau ci-après.

Ensembles de données	Données d'observation mesurées par les stations météorologiques	Données climatiques: observations validées	Alertes météorologiques	Données radar	Modèles de données PMN
Granularité	Par station météorologique, résolution temporelle complète	Par station météorologique, résolution temporelle complète	Alertes, au moins 48 heures à l'avance	Par station radar dans l'État membre et composite national	Au moins 48 heures à l'avance par paliers d'une heure, national, à 2,5 km/meilleur maillage disponible
Attributs clés	Toutes les variables d'observation mesurées	Toutes les variables d'observation mesurées et validées; moyenne journalière par variable		Réflexivité, rétrodiffusion, polarisation Précipitation, vent et écotops	Déterministes et/ou ensembles si disponibles, pour les paramètres et niveaux d'intérêt météorologique

#### 3.2. Modalités de publication et de réutilisation

a) Les ensembles de données sont mis à disposition en vue de leur réutilisation:

- aux conditions de la licence Creative Commons BY 4.0 ou de toute licence ouverte équivalente ou moins restrictive,
- sous l'un des formats figurant dans le tableau ci-après ou un autre format ouvert et lisible par machine, reconnu au niveau de l'Union ou au niveau international,
- par l'intermédiaire d'API et sous la forme d'un téléchargement en masse, sauf pour l'ensemble de données «modèles de données PMN», lequel n'est disponible qu'en recourant à des API,
- conformément à la fréquence et l'actualité de mise à jour indiquées dans le tableau ci-après.

b) Les métadonnées décrivant l'ensemble de données sont complètes et mises à disposition sur le web sous

un format ouvert et lisible par machine largement utilisé.

- c) Les ensembles de données sont décrits dans un document en ligne complet et public, lequel précise au moins la structure et la sémantique des données.

Ensembles de données	Données d'observation mesurées par les stations météorologiques	Données climatiques: observations validées	Alertes météorologiques	Données radar	Modèles de données PMN
Format	BUFR, NetCDF, ASCII, CSV, JSON	NetCDF, JSON, CSV	XML (Cap ou RSS/Atom), JSON	HDF5, BUFR	GRIB (ou NetCDF)
Fréquence et actualité de mise à jour	Toutes les 5 à 10 minutes en temps réel pour les stations automatisées, données horaires non validées pour toutes les stations, pendant les dernières 24 heures	Données validées chaque jour toutes les heures (et avec une meilleure résolution temporelle) et données d'observation moyennes quotidiennes; toutes les données historiques numérisées	En fonction des publications ou toutes les heures	En temps quasi réel à intervalles de 5 minutes (ou à l'intervalle le plus court disponible)	Toutes les 6 heures, ou à une meilleure résolution temporelle, sur les dernières 24 heures

Source : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32023R0138&from=EN>

# DOCUMENT 6 :

*Traduit et extrait d'un article publié dans le bulletin n°70 (2) – 2021 de l'OMM,  
consacré à la politique unifiée de données de l'OMM*

## **Politique de données de l'OMM pour le 21<sup>e</sup> siècle**

[...]

### **La politique de données de l'OMM et les moteurs actuels du changement**

L'un des principaux rôles de l'OMM est de faciliter et de coordonner les échanges internationaux de données nécessaires pour la délivrance de services. Le rôle de sa politique de données est d'articuler les principes de cet échange et les pratiques qui soutiennent sa mise en œuvre : quels types de données seront échangées, par qui et avec qui seront-elles échangées, à quelles fins, et dans quelles conditions ?

Alors que la nécessité d'un échange international de données est clair et presque universellement comprise par les Membres de l'OMM, il n'est pas nécessairement facile de formuler une politique qui soit à la fois utilisable et acceptable pour tous les membres. Dans le but commun des gouvernements nationaux de fournir des services météorologiques et climatiques à leurs citoyens, il existe un large éventail de différentes mises en œuvre nationales, avec différentes attributions de responsabilités aux divers acteurs. Il n'y a, par exemple, pas de compréhension commune et universelle du rôle du public par rapport au secteur privé. Il n'y a pas d'accord universel sur quels services doivent être gratuitement fournis par les institutions gouvernementales, versus le rôle des services payants, qui peuvent être fournis soit par des entités privées ou par des institutions gouvernementales agissant comme des entités privées. Une politique de données efficace doit fournir suffisamment de clarté pour permettre aux Membres de générer des produits et de fournir des services à leur échelle. En même temps, elle doit rester suffisamment large et non prescriptive pour tenir compte de ces différentes politiques nationales et des différentes mises en œuvre nationales de la chaîne de valeur météorologique.

Au fil des ans, la politique de données de l'OMM a dû évoluer pour s'adapter aux nouvelles exigences, aux nouveaux domaines d'application, aux nouvelles technologies, et à l'évolution des réalités politiques et économiques. La politique des données continue d'évoluer au XXI<sup>e</sup> siècle, avec une mise à jour majeure soumise à la délibération de la Session extraordinaire du Congrès de l'OMM en octobre 2021. Il y a plusieurs grands moteurs derrière ce nouveau développement, et certains des plus importants seront listés dans les prochains paragraphes.

Tout d'abord, les progrès phénoménaux et le succès de la surveillance et de la prévision météorologique et climatique ont conduit à une **croissance explosive de la demande en informations météorologiques, climatiques et relatives au Système Terre**, de la part de tous les secteurs de la société. Cela a conduit à une reconnaissance accrue de la valeur économique de tous les types de données du système Terre, ce qui à son tour a conduit à un groupe d'acteurs d'une diversité de plus en plus grande impliqués dans la génération ou dans l'utilisation de ces données. De plus, malgré notre capacité technologique en constante progression, il peut être avancé que notre vulnérabilité aux aléas météo augmente dans de nombreux domaines. Plus de gens que jamais auparavant dans l'histoire vivent dans des zones à haut risque telles que les zones côtières basses et exposées et les plaines inondables. Les mégapoles ont leur propre vulnérabilité aux conditions météorologiques à fort impact liées à l'importance du maintien des infrastructures critiques opérant dans toutes les situations et les difficultés liées aux évacuations potentielles. De plus, le changement climatique en cours est déjà en train de modifier la fréquence des phénomènes météorologiques à fort impact. Une surveillance accrue et des capacités de prévision améliorées sont nécessaires non seulement pour gérer l'impact des événements météorologiques actuels, mais aussi pour aider la société à comprendre et à s'adapter au temps auquel il faudra nous attendre dans le futur.

Deuxièmement, la demande toujours croissante de données du Système Terre et les progrès constants de la technologie, couvrant les observations, les télécommunications et l'analyse de données, ont conduit à une **énorme augmentation du volume de données disponibles sur le système Terre**. Les développements en technologie de télédétection au sol et dans l'espace et dans la vitesse de traitement et la taille de la mémoire des ordinateurs utilisés pour la modélisation météorologique nécessitent l'adoption de **nouvelles approches**

**pour la distribution des données et l'accès aux données.** Des solutions plus anciennes, telles que le Système mondial de télécommunications (GTS), ne peut plus supporter correctement l'échange de données et doit être remplacé par des technologies basées sur internet et le cloud. Le système d'information de l'OMM (WIS) fournit les normes requises pour les formats de données et les métadonnées.

Troisièmement, ces dernières années, **de nombreux pays (en particulier les pays les plus développés) se sont efforcés de rendre toutes les données financées par des fonds publics généralement disponibles en tant que « open data ».** Les fournisseurs de données, tels que les SMHN, sont confrontés à une demande croissante d'accès ouvert à toutes les données qu'ils produisent. L'Union européenne a été le fer de lance de ces démarches avec sa directive « sur les données ouvertes et sur la réutilisation des informations du secteur public » (directive 2019/1024), que ses États membres sont tenus de mettre en œuvre dans leur législation nationale. La justification socio-économique de cette directive est énoncée clairement dans son préambule (par. 9 et 16) : « *Les informations du secteur public constituent une source extraordinaire de données qui peuvent contribuer à améliorer le marché intérieur et à développer de nouvelles applications pour les consommateurs et les personnes morales. L'utilisation intelligente de données, y compris leur traitement par des applications utilisant l'intelligence artificielle, peut avoir un effet de transformation sur tous les secteurs de l'économie* ».

Quatrièmement, il y a eu une **expansion constante de pratiques d'observation et de modélisation météorologiques dans les domaines d'application adjacents** tels que la surveillance de l'environnement, et la recherche d'une extension à la fois de la précision et la plage de prévision de divers phénomènes météorologiques, climatiques et hydrologiques. Cela montre clairement qu'une **approche intégrée du système Terre** est nécessaire. Pour réussir, cette approche doit englober la conception de réseaux d'observation, la production et l'échange d'observations, la modélisation d'un système terrestre intégré, et l'échange des données modèles en résultant.

Après une analyse approfondie de ces moteurs et de leurs implications, l'OMM a décidé que l'adoption d'une déclaration de politique globale unique, qui énonce l'étendue de l'échange de données nécessaire au XXI<sup>e</sup> siècle, serait clairement et sans ambiguïté la meilleure façon de répondre et d'aider le monde de la météorologie pour continuer à progresser.

## **Contexte sociétal et évolutions externes**

Comme indiqué plus haut, ces dernières années, un groupe diversifié de parties prenantes ont rejoint les SMHN comme participants actifs ou potentiels à l'échange de données. Il s'agit notamment d'agences gouvernementales autres que les SMHN, des organisations à but non lucratif et diverses entités du secteur privé et du monde universitaire. Cette diversité est hautement souhaitable et, si elle est gérée correctement, aidera à atteindre l'ampleur des échanges de données nécessaire pour la mise en œuvre de l'approche du système Terre que l'on vient de décrire. Afin d'exploiter la valeur qu'une participation aussi large et diversifiée fournira, l'élaboration et la mise en œuvre de la politique de données de l'OMM doit pouvoir accueillir cette diversité.

En reconnaissance de ces opportunités, la 18<sup>e</sup> session du Congrès météorologique mondial a adopté la Déclaration de Genève de 2019 : « *bâtir une communauté pour agir dans les domaines du temps, du climat et de l'eau* », et a convenu de convoquer un dialogue consultatif de haut niveau et ouvert sur le partenariat et l'innovation pour la prochaine génération d'intelligence météo et climatique.

La nouvelle politique de données unifiées de l'OMM a été développé avec cette Déclaration de Genève en toile de fond. Des représentants des secteurs public, privé et universitaire ont participé activement à sa rédaction. La politique vise à fournir des opportunités gagnant-gagnant en facilitant une large participation à l'échange libre et sans restriction de données météorologiques et de données connexes du système Terre. Grâce à une formulation adaptée de la politique de données et grâce à son utilisation de définitions précises, elle apporte de la clarté à toutes les parties concernant les attentes placées en elles et les avantages qu'elles peuvent en retirer. Une large consultation, impliquant tous les participants à l'échange de données, restera un élément clé des révisions régulières de la politique et des pratiques, avec l'intention de s'assurer que la politique reste à jour et réponde à l'évolution du contexte.

Les principales initiatives stratégiques prises dans le monde, à la fois dans les domaines privé et gouvernemental, illustrent déjà certaines des tendances mentionnées ci-dessus. Un fil conducteur est l'implication de consortiums multi-acteurs où toutes les parties bénéficient d'investissements substantiels dans des infrastructures collaboratives. Quelques exemples de telles initiatives, qui ont soit aidé à éclairer l'élaboration de la nouvelle politique de données de l'OMM et/ou qui bénéficieront de sa mise en œuvre par les Membres de l'OMM, sont mis en évidence ci-dessous.

**Copernicus** – Via son programme Copernicus, la Commission européenne au nom de ses États membres investit dans la consolidation des plates-formes de modélisation et d'intégration de données, pour mieux exploiter et maximiser l'utilisation des capacités existantes et pour construire sur sa masse critique d'expertise. La mission du programme Copernicus est d'exploiter les observations de la Terre et les prévisions numériques pour générer de la valeur ajoutée couvrant tous les composants de l'environnement au profit des décideurs, chercheurs, utilisateurs commerciaux et privés ainsi que la communauté scientifique mondiale. Copernicus couvre six principaux domaines de service, dont l'un est axé sur le changement climatique et ses impacts. Le programme est mis en œuvre en partenariat avec les États membres de l'Union européenne, l'Agence spatiale européenne (ESA) et EUMETSAT, le CEPMMT et diverses autres agences de l'UE ainsi que Mercator Océan.

**Microsoft Earth** – En décembre 2017, Microsoft a annoncé un investissement de 50 millions de dollars dans leur « IA pour programme Terre » pour développer les technologies d'intelligence artificielle pour mieux comprendre et faire face aux problèmes environnementaux auxquels la planète est confrontée. Comme ils l'ont formulé : « Fondamentalement, l'IA peut accélérer notre capacité à observer les systèmes environnementaux et la façon dont ils évoluent à l'échelle mondiale, convertir les données en informations utiles et appliquer ces informations pour prendre des mesures concrètes pour mieux gérer nos ressources naturelles. » L'IA pour la Terre vise à offrir un meilleur accès aux outils informatiques et d'IA. Il abaisse les barrières à l'entrée pour les petites organisations qui ne peuvent pas se permettre l'investissement dans une infrastructure informatique et crée une communauté de contributeurs qui fera progresser les solutions d'IA innovantes et évolutives.

En parallèle, Microsoft et d'autres fournisseurs de cloud connectent les bases de données gouvernementales à leurs infrastructures cloud propriétaires pour augmenter l'offre de données qu'ils mettent à disposition de leurs clients.

**IBM Environmental Intelligence Suite : Geospatial Analytics** – le composant Geospatial Analytics de la suite d'*Environmental Intelligence Suite* d'IBM est une plateforme spécifiquement conçue pour les données géospatiales-temporelles massives (cartes, satellite, météo, drone, Internet des objets), services de requête et d'analyse à partir d'une base de données d'informations géospatiales-temporelles continuellement mise à jour et qui comprend les informations météorologiques et climatiques par satellite. Elle effectue diverses transformations pour rendre à la fois des ensembles de données historiques et en temps réel plus faciles à utiliser, puis permet une découverte rapide des données. En même temps, Geospatial Analytics est une plateforme de calcul pour exécuter des modèles physiques et statistiques sur les ensembles de données organisées. Il vise à soutenir le développement d'applications incluant l'IA et d'autres types d'analyses de données.

**Politique de données unifiées de l'OMM sur l'échange international de données du système Terre et ses impacts attendus** – Comparée aux trois résolutions existantes de politique de données de l'OMM qu'il est prévu de remplacer, la nouvelle politique de données unifiées de l'OMM fournira une approche plus complète, plus flexible et plus facilement applicable des échanges de données.

La politique actuelle de l'OMM en matière de données (septembre 2021) est énoncée dans trois résolutions distinctes du Congrès, chacune couvrant un domaine spécifique : Résolution 40 (Cg-XII, météo), Résolution 25 (Cg-XIII, eau) et Résolution 60 (Cg-17 climat). A contrario, la résolution de la politique de données unifiées de l'OMM couvre sept disciplines et domaines – englobant ensemble toutes les données du système Terre de l'OMM – dans une seule résolution de politique. De plus, la nouvelle politique s'étend – partant d'abord uniquement de l'échange entre SMHN – pour englober les Membres dans leur ensemble, et ainsi toutes les entités participant à l'échange, y compris le secteur privé et le milieu universitaire.

Semblable<sup>1</sup> à la résolution 40, la nouvelle résolution distingue les « données de base », pour lesquelles l'échange de données est considéré comme obligatoire, et les « données recommandées », pour lesquelles l'échange est fortement recommandé. Cependant, contrairement à la résolution 40, la nouvelle résolution

---

1 Le terme « core data » remplace « essential data » utilisé dans la résolution 40, tout comme les « recommended data » remplacent le terme « additional data » de la résolution 40.

prend une approche modulaire pour spécifier ce qui relèvera précisément de ces deux catégories. Des détails spécifiques sur ce qui est considéré comme « données de base » et « données recommandées » sont – ou seront dans certains cas – fournis dans les règlements techniques de l’OMM. De nouveaux développements peuvent donc être pris en compte par des amendements de ces règlements techniques, plutôt que d’exiger une mise à jour de la résolution de la politique elle-même. Cela rend la mise en œuvre beaucoup plus facile à faire vivre et évoluer que ce n’est actuellement le cas.

En termes d’impact attendu, la mise à jour proposée de la politique de données stimulera et renforcera l’échange international d’observations de toutes parts du globe. Augmenter le nombre d’observations qui sont partagées au niveau international pour être exploitées dans les modèles globaux et régionaux de prévision numérique du temps contribueront à améliorer considérablement la qualité des résultats obtenus. Cette amélioration se fera sentir partout sur le globe, mais sera particulièrement prononcée dans les zones où la couverture actuelle des données d’observation est faible, ce qui est le cas dans de nombreux pays en développement. En retour, la politique de données aidera à fournir un accès gratuit et sans restriction à une gamme beaucoup plus large de produits de données de modèle du système Terre pour tous les membres, ce qui les aidera à améliorer et à élargir l’éventail des services qu’ils peuvent fournir.

Un accès à des données supplémentaires, au-delà du traditionnel domaine des observations météorologiques, sera critique au développement continu et à la mise en œuvre des systèmes de modélisation couplés utilisés dans l’approche intégrée du système Terre. Au cours des prochaines décennies, ces systèmes, couvrant des échelles de temps allant des prévisions météo de court terme jusqu’aux projections climatiques de long terme, devra résoudre des éléments à des résolutions de plus en plus élevées et devra incorporer des caractéristiques détaillées de la surface du sol, répondant au niveau de détail des emplacements géographiques intéressant les utilisateurs. Le système Terre, tel que défini par les activités de l’OMM, est susceptible de continuer à s’étendre au-delà de sa portée actuelle. On peut ainsi voir arriver des systèmes bio-géo-chimiques couplés en appui à la production primaire, la prédiction de floraison, l’apport de carbone, la gestion des pêches, etc. Les paragraphes suivants énumèrent des exemples de domaines d’application émergents et qui sont construits sur différents niveaux d’approches intégrées du système Terre. Il s’agit de domaines qui bénéficieront de la politique de données unifiées et qui influenceront son évolution future via des revues régulières.

**Marine : système de sécurité de navigation maritime à l’état de l’art :** plusieurs pays travaillent à obtenir de meilleures informations de navigation pour les marins. Ces pays sont engagés à développer une opérationnalisation robuste en utilisant des solutions de modèles océaniques dynamiques pour fournir des solutions numériques de navigation et hydrographiques pour les marins et les navigateurs. De telles associations nationales d’efforts entre les services de météorologie, d’océanographie et d’hydrographie visent à fournir des solutions opérationnelles grâce à des capacités de modélisation sans couture. Cette approche permettra des opérations sécurisées et optimales pour les secteurs de la navigation commerciale et du transport maritime en favorisant une approche écosystémique de la gestion des activités humaines.

**Sécurité et Urgence :** Dans le domaine de la caractérisation des événements nucléaires, il est nécessaire d’améliorer l’échange de données en temps réel, les capacités de calcul, le *machine learning* et l’intelligence artificielle liés à l’analyse d’ensemble des données météorologiques ainsi que d’autres types de données. Cela nécessitera d’autres améliorations dans la précision des modèles de prévisions météorologiques et de transport atmosphériques, et dans la caractérisation des incertitudes associées. Des travaux supplémentaires sont également nécessaires pour développer et intégrer davantage de modélisations de la chimie atmosphérique dans la modélisation du transport atmosphérique. Ces mises à niveau amélioreront la capacité d’identification des événements nucléaires, la reconstruction du terme-source et la modélisation prédictive des résultats sanitaires liés aux rejets radiologiques ou d’autres rejets d’origine humaine.

**Changement climatique : apporter des innovations scientifiques pour adresser les conséquences sur la santé de la pollution de l’air et du changement climatique.** Des recherches ont montré que plusieurs sources de données peuvent être incorporées dans les modèles de qualité de l’air pour améliorer leurs détails et leur précision. L’explosion des méthodes de saisie et d’analyse des données peut accroître la gamme des données d’entrées pour le modèle. En même temps, le *deep learning* et d’autres méthodologies ont le potentiel d’améliorer la compréhension des relations sous-jacentes et d’améliorer les capacités de prévision, à la fois pour les prévisions quotidiennes et les projections à long terme. Une capacité de modélisation élargie sera utilisée pour renforcer la science qui sous-tend la prise de décision réglementaire et pour suivre les bénéfices sanitaires d’instruments fondés sur le marché, tels que la tarification du carbone. Des outils seront développés pour soutenir les autorités de santé publiques locales pour anticiper et faire face aux défis environnementaux.

**Développement des terres et des ressources : soutenir le développement durable des terres et des ressources avec des données prêtes pour l'analyse.** Les marchés dépendent de la durabilité du développement des terres et des ressources. Il y a un besoin continu de surveiller et d'évaluer les changements dynamiques sur le paysage terrestre local, régional et mondial. L'approche écosystémique de la gestion de l'environnement se concentre sur le maintien de la capacité d'un système entier pour produire des biens et services écologiques. Cela commence par le suivi et la gestion, par exemple, des ressources en eau, de la qualité de l'air et de l'eau, et des ressources génétiques, qui maintiennent l'économie mondiale, la sécurité, la santé et le bien-être. De solides infrastructures de données et de calcul sont nécessaires pour adopter une approche écosystémique pour permettre des progrès dans l'évaluation de l'état et des tendances des changements terrestres. Cette capacité émergente facilitera l'analyse des impacts cumulatifs qui fournira des services spécifiques ainsi qu'une aide à la décision pour les réglementations environnementales et pour le développement des politiques environnementales.

**Agriculture : informations météorologiques et climatiques pour surveiller la sécheresse et gérer le risque agroclimatique.** Certains pays ont mis en place des programmes de surveillance de la sécheresse qui utilisent une diversité d'observations de la Terre et d'autres données pour fournir en temps utile des informations et des cartes sur des paramètres météo et climatiques particulièrement pertinents pour leurs secteurs agricoles nationaux. L'agriculture a besoin de divers services météorologiques et hydrométéorologiques, adaptés selon la culture spécifique, sa phase de croissance ou le type et l'état actuel du sol. Le labour, l'irrigation, les semences, la récolte ou l'application efficace d'engrais et de pesticides avec la contrainte de la protection des eaux souterraines... dépendent beaucoup de la météo. Les cartes de l'humidité du sol sont un exemple de produit créé à l'aide de données géophysiques. De tels produits permettent aux agriculteurs de voir où les conditions sont plus humides ou plus sèches que la normale, permettant une résilience face à un climat changeant.

**Santé : Sciences de l'observation de la Terre et maladies infectieuses émergentes et réémergentes.** Les maladies infectieuses apparaissent et réapparaissent sous l'influence de facteurs clés tels que l'environnement, le climat, la démographie, et les changements socio-économiques et de comportements humains. Elles sont un défi pour la santé publique aux niveaux local et mondial. En élargissant l'échange d'observations environnementales et en comprenant comment ces facteurs affectent l'occurrence des maladies, les autorités peuvent prédire quand, comment et où la maladie apparaîtra ainsi qu'identifier les populations à risque celles les plus vulnérables. Le changement climatique devrait exacerber les risques de maladies à transmission vectorielle en permettant la propagation d'animaux-hôtes, d'agents pathogènes, de vecteurs, l'établissement de vecteurs exotiques et des maladies qu'ils transmettent (dengue, Zika, chikungunya, fièvre jaune), au-delà de leurs domaines historiques. Le changement climatique devrait également augmenter la réémergence (épidémies) de maladies infectieuses déjà endémiques dans des pays des régions sous-continentales.

---

Les améliorations continues des services de suivi et de prévision du temps, du climat et du système Terre, dont le monde a été témoin au cours des 70 dernières années, sont liées à une meilleure science, de meilleures technologies, et à l'échange en temps réel de sources plus variées d'observations. Nous vivons à une époque de technologies de plus en plus sophistiquées et le rythme de l'innovation s'accélère. Nous sommes inondés d'observations de la Terre. Les médias sociaux offrent un accès à des informations contextuelles et à des mécanismes de diffusion sans précédent. Les plates-formes de calcul haute-performance nous permettent d'aborder des problèmes auparavant insolubles.

Ce n'est qu'une question de temps avant que la fusion des données sur la météo, l'eau et le climat, les technologies du big data, et leur exploitation par les entreprises ne se généralisent. Cela changera la façon dont les gens et les entreprises perçoivent les données météo et hydro - et expérimentent leurs effets démultiplicateurs sur l'amélioration des décisions commerciales sensibles aux conditions météorologiques. Ces évolutions vont vraisemblablement amener de nombreux Membres de l'OMM à réévaluer leurs politiques de données et leurs stratégies de partenariat au niveau national, et elles auront un impact profond sur la communauté météorologique mondiale. La nouvelle politique de données unifiée de l'OMM vise à aider les Membres à s'adapter à ces changements et à continuer à fournir les meilleurs services possibles dans toutes les disciplines et domaines de l'OMM.

Source : <https://public.wmo.int/en/resources/bulletin/wmo-data-policy-21st-century>

# DOCUMENT 7 :

## **EXTRAITS d'une note interne de la DG (juillet 2021)**

**concernant l'impact anticipé des changements réglementaires annoncés au niveau national, dans un contexte de forte augmentation du volume de données publiques produites par l'établissement.**

### **1 Contexte**

L'action de Météo-France s'inscrit, depuis près de dix ans, dans la politique volontariste de l'Etat d'ouverture des données publiques. Ainsi, 70 jeux de données sont aujourd'hui déclarés sur le Service Public de la Donnée (data.gouv.fr) ; les premières données ayant été mises en ligne dès 2012. Depuis 2012, Météo-France maintient également un portail de données publiques (<https://donneespubliques.meteofrance.fr>) permettant d'accéder à 30 jeux de données, dont la majeure partie sont téléchargeables en ligne par paquets et/ou par API<sup>1</sup>. Un service de hotline est également proposé sur ce portail pour répondre aux demandes et aux questions des usagers sur les données publiques. Les autres jeux de données publiques de l'établissement sont accessibles en réponse à des demandes ponctuelles. Des frais de mise à disposition sont alors facturés selon le barème en ligne sur le portail de données publiques.

Les moyens humains consacrés actuellement par l'établissement pour ce qui concerne l'évolution et le maintien en conditions opérationnelles du portail de données publiques de Météo-France et la constitution des jeux de données demandés par les utilisateurs et qui ne sont pas en ligne, sont de l'ordre de la dizaine d'ETP. En matière d'infrastructure, l'établissement consacre de l'ordre de 100 k€/an à l'hébergement, la mise à disposition et la diffusion des données du portail de données publiques (bande passante, serveurs).

**Météo-France ne dispose d'aucun financement spécifique pour assurer cette mission de mise à disposition de ses données publiques.**

Le contexte réglementaire conduira Météo-France à devoir, à horizon 2023, mettre en ligne gratuitement de nouvelles catégories de données jusqu'ici soumises à redevance de réutilisation avec une perte de revenus immédiate. Il est de plus prévu que les volumes des catégories de données publiques déjà accessibles (gratuitement ou avec redevance) augmentent fortement. Cela va nécessiter de réaliser des développements conséquents en matière d'accès à ces données sur le portail de données publiques de l'établissement, pour gérer de plus grands volumes de données, et, le cas échéant, en faciliter l'accès, notamment via des API. Cela va également impacter les recettes commerciales de l'établissement.

Plus précisément, le passage à la gratuité de l'ensemble des données pouvant actuellement faire l'objet de redevances de réutilisation à horizon 2023 tel qu'annoncé dans la circulaire du premier ministre N°6264/SG du 27 avril 2021, conduira à une perte directe de ressources financières de l'ordre de 1 400 k€/an.

La 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 est également susceptible de conduire à imposer la mise en ligne de données d'ores et déjà non soumises à redevance de réutilisation ce qui ajouterait de l'ordre de 200 k€/an de perte de ressources budgétaires (facturation des coûts de mise à disposition des données basée sur le défraiement des coûts liés à l'extraction et l'envoi des données souhaitées).

<sup>1</sup> Dans cette note, le terme API (interface de programmation d'application) sera utilisé pour désigner des API ciblées (de type WS Inspire) permettant à l'utilisateur d'accéder uniquement aux données dont il a besoin sans avoir à télécharger l'intégralité d'un paquet de données.

Cette note détaille les conséquences de ces changements pour l'établissement et les besoins correspondants, en matière de modernisation de l'infrastructure technique de mise en ligne de ses données publiques, de développements liés aux données qui seront considérées comme HVD, et enfin d'évolution de l'offre commerciale pour s'adapter à ce nouveau contexte.

**Sans ressources spécifiques (investissement, fonctionnement, ETP supplémentaires), Météo-France n'aura pas les moyens humains & budgétaires de supporter les coûts engendrés par ces changements, et en particulier d'offrir une qualité de service répondant aux attentes des utilisateurs.**

## **2 Estimation des coûts d'investissement et de fonctionnement nécessaires à la modernisation de l'infrastructure technique de mise en ligne des données publiques de l'établissement.**

Les évolutions réglementaires et la forte augmentation du volume de données produites par l'établissement rendent nécessaire la modernisation de l'infrastructure technique d'hébergement/production/mise en ligne/diffusion des données publiques de l'établissement. Ainsi, l'infrastructure technique actuelle est dimensionnée et permet la mise en ligne quotidienne de 0,7 To de données. En 2023, 2 To de données publiques devront être mises en ligne quotidiennement sur le portail de l'établissement. En 2026, ce sont de l'ordre de 7 To de données publiques qui pourraient devoir être mises en ligne quotidiennement, soit plus de 10 fois plus de données qu'aujourd'hui. Le dépôt de 0,7 To occasionne actuellement des téléchargements quotidiens à hauteur de 1,4 To, en saturant la bande passante mise à disposition.

Sans modernisation de cette infrastructure technique, le niveau de service rendu, en termes de temps de transfert par exemple, aux utilisateurs risque de se dégrader fortement par rapport à aujourd'hui, alors que se produisent déjà des situations de saturation de la bande passante réseau. Le passage de cette dernière de 0,4 à 1 Gbit/s prévu en 2022 devrait permettre de revenir à une très bonne qualité de service, y compris pour les usagers des modèles de prévision.

**Dans le contexte d'une mise en ligne quotidienne de données publiques d'un volume de 7 To (prévision 2026), une estimation des coûts d'investissement et de fonctionnement a été réalisée pour trois scénarios afin d'arbitrer la modernisation de cette infrastructure technique :**

- **Scénario A : Gel de l'infrastructure au niveau de 2022**
- **Scénario B : Modernisation de l'infrastructure technique dimensionnée de manière à assurer un niveau de service équivalent au niveau de service actuel.**
- **Scénario C : Modernisation de l'infrastructure technique dimensionnée de manière à assurer un niveau de service amélioré.**

Dans le scénario A, la bande passante reste au niveau de 2022 (1 Gbit/s), mais le reste de l'infrastructure n'est pas modifié. Dans la mesure où les dépôts quotidiens seront 10 fois plus volumineux qu'actuellement, un nombre important de tentatives de téléchargement risque d'échouer avec une dégradation progressive au fur et à mesure de l'augmentation du volume des dépôts quotidiens de données publiques.

Dans le scénario B, la bande passante nécessaire pour accompagner un dépôt quotidien de 7 To est portée à 4 Gbit/s avec des investissements en matériel réseau spécifiques. Le nombre de serveurs est également augmenté, et un silo de production spécifique pour la diffusion de données est mis en place, avec 40 serveurs et 90 To de stockage. Le niveau de bande passante est 10 fois supérieur au niveau actuel, en stricte proportion de l'augmentation du volume de données (de 0,7 To à 7 To). Le niveau de service, en termes de temps nécessaire pour télécharger les modèles, restera au niveau actuel avec peu d'échecs de téléchargement pour autant que les téléchargements de prévisions ensemblistes restent modérés. C'est sensiblement le niveau de service actuel, mais sur un ensemble de données élargi.

Dans le scénario C, la bande passante internet dédiée passe à 10 Gb/s, le silo spécifique abrite 60 serveurs, et le stockage rapide dédié aux caches passe à 350 To. Ceci permet de conserver sur l'intégralité de la période 2022-2026 l'amélioration du niveau de service qui sera apportée en 2022 par l'amélioration du débit internet, mais sur l'ensemble du spectre des services : données complètes et historiques, y compris pour les ensembles, mécanismes de cache renforcés pour limiter les risques d'interruption de téléchargement.

Le tableau suivant résume les coûts globaux de ces scénarios :

	Investissement initial	Fonctionnement sur 5 ans		RH informatiques	
		Actuel	Prévu	actuelles	prévues
scénario A	0 k€	500 k€	500 k€	3	3
scénario B	896 k€	500 k€	1781 k€	3	7
scénario C	1402 k€	500 k€	4008 k€	3	9

Les coûts d'investissement et de fonctionnement mentionnés dans ce tableau portent sur l'ensemble de l'infrastructure technique d'hébergement/production/mise en ligne/diffusion des données. Les coûts de fonctionnement des scénarios B et C sont très majoritairement liés aux coûts d'achat de bande passante (B : 45% / C : 50%) et de dispositifs de protection du portail des données publiques en cas d'attaques par déni de service (B : 34% / C : 37%).

A noter qu'aux coûts d'investissement et de fonctionnement directement liés à la modernisation de l'infrastructure technique, s'ajoutent également des ressources humaines supplémentaires d'administration et de pupitrage (colonne RH informatiques du tableau ci-dessus). Le scénario B nécessite ainsi de mobiliser 4 ETP supplémentaire par rapport aux ressources actuelles (6 ETP pour le scénario C).

### **3 Estimation des coûts d'investissement et de fonctionnement nécessaires au développement et au maintien d'une offre d'accès en ligne et via API pour les données qui seront considérées comme HVD.**

Les données qui seront considérées comme HVD (High Value Datasets – Ensembles de données de forte valeur) en application de la directive (UE) 2019/1024 devront être accessibles en ligne via des API. A ce stade, la liste définitive de ces données HVD n'est pas connue mais le cabinet Deloitte a réalisé une étude (été 2020) pour le compte de la Commission Européenne en proposant deux scénarios qui permettent de cerner ce que pourraient être ces données.

Les coûts induits pour l'établissement sont principalement des coûts d'investissement liés à la mise en ligne et au développement d'API. En se basant sur les deux scénarios de l'étude Deloitte, ces coûts sont estimés entre 5 et 6 hommes.année. Une sous-traitance est envisageable pour un montant estimé entre 700 k€ (scénario de « faible intensité d'intervention ») et 840 k€ (scénario de « forte intensité d'intervention »). Le pilotage interne et l'accompagnement de cette sous-traitance est alors estimé entre 1 à 2 hommes.année.

Enfin, la maintenance évolutive associée à la mise en ligne et au maintien de ces API est estimée à 30 % par an du coût d'investissement initial, soit de l'ordre de 2 ETP, sous forme de ressources d'ingénierie interne.

Météo-France entend généraliser le principe d'une facturation des frais de mise à disposition pour compenser une partie des coûts de mise à disposition et diffusion de ses données publiques, y compris pour les données mises en ligne sur son portail de données publiques mais, sauf à facturer des montants très élevés, cela ne pourra pas couvrir l'ensemble des coûts.

#### **4 Pertes commerciales et perspectives de compensation grâce à une évolution rapide et la modernisation de l'offre de services de l'établissement.**

Aux 1 600 k€/an de pertes immédiates engendrées par la suppression des redevances de réutilisation des données prévue par la circulaire du premier ministre N°6264/SG du 27 avril 2021 et l'application de la directive (UE) 2019/1024 du parlement européen et du conseil du 20 juin 2019, s'ajouteront des pertes associées de recettes commerciales liées au passage à la gratuité de ces données en raison des modifications induites sur le marché des services météorologiques payants (baisse de la valeur globale du marché, induite par la baisse des coûts d'acquisition des données, accroissement de la pression concurrentielle, nouveaux entrants..), même si ces pertes de recettes commerciales sont difficiles à estimer.

Afin d'anticiper les conséquences de ces possibles pertes commerciales, Météo-France devra investir fortement dans la modernisation de son offre de services. Deux axes en particulier vont nécessiter des investissements significatifs :

- Le développement d'une nouvelle offre de diffusion des produits de l'établissement basée sur la mise en place d'une boutique d'API.
  - Le développement de nouveaux produits basés sur des techniques d'Intelligence Artificielle (IA),
- Ces deux axes requièrent un effort conséquent de R&D et de développements pour l'établissement.

Le développement d'une nouvelle offre de diffusion des produits de l'établissement basée sur la mise en place d'une boutique d'API vise un objectif double. Permettre d'une part une plus grande souplesse dans l'ajout d'API adaptées aux besoins des clients, qui éprouvent des difficultés à s'adapter aux API standards prévues par la directive Inspire et sa déclinaison nationale. D'autre part, mettre en place des mécanismes de paiement et de facturation induisant des tarifs dégressifs à l'usage, des périodes d'essai avec limitations qualitatives, et des paiements en ligne. Le premier sujet est de nature architecturale aussi bien qu'organisationnelle, et demande à développer une capacité technico-commerciale forte sur les questions d'échanges de données. C'est un véritable enjeu économique pour les clients de Météo France, qui sont parfois relativement démunis pour transformer leurs besoins internes en commandes précises et efficaces.

Dans un contexte de volumes de données toujours plus importants (arrivée de nouveaux instruments satellitaires de télédétection, données issues des prévisions d'ensemble ou encore des simulations climatiques), le recours aux techniques d'IA est une opportunité pour Météo-France, non seulement en termes d'innovation pour ses clients mais aussi pour ses activités de recherche ou encore pour développer de nouveaux systèmes d'aide à la décision pour ses prévisionnistes. A cette fin, Météo-France a mis en place, dans le cadre d'un projet soutenu par le FTAP, un laboratoire dédié à l'exploration des techniques d'IA. Dans le cadre de ce projet, Météo-France dispose d'un financement jusqu'à fin 2021 qui a notamment permis de recruter 7 experts spécialistes en Machine Learning. Le bilan associé à la création de ce laboratoire est très positif et a permis à l'établissement de mettre en œuvre de nouveaux projets innovants d'utilisation des techniques d'IA.

Disposer de financements permettant de pérenniser les moyens humains de ce laboratoire, à même hauteur qu'actuellement (soit 7 ETP pendant 5 ans), permettrait à Météo-France de parvenir au déploiement de nouvelles offres de produits innovantes, compétitives et à forte valeur ajoutée, basées sur les techniques d'IA.

Un tel effort de R&D permettrait d'une part de poursuivre les projets en cours et valoriser pleinement l'effort de R&D entrepris dans le cadre du projet FTAP et d'autre part, d'explorer le potentiel de ces techniques sur de nouveaux domaines d'application : prévision immédiate des précipitations intenses, observation automatique de phénomènes météorologiques à partir de webcams (neige sur la route et brouillard notamment), estimation des précipitations par satellite, détection par deep learning d'objets météorologiques dans les sorties de modèles, génération automatique de bulletins et de textes (du type de ceux rédigés à destination des usagers de la mer), etc.

## 2. Le développement de l'open data s'accélère

**L'ouverture des données publiques**, notamment les données météorologiques, constitue un levier de croissance économique et d'innovation en permettant le développement de nouvelles activités et la création de nouveaux produits à forte valeur ajoutée. Elle est aussi une **opportunité pour Météo-France de valoriser son savoir-faire et d'accroître les bénéfices socio-économiques qu'il génère**. Ainsi, à titre d'exemple, l'opérateur a récemment mis en ligne un jeu de données d'archives météorologiques, baptisé « **Météonet** », sous un format adapté aux data scientists dans la perspective d'**attirer les activités de recherches de laboratoires d'intelligence artificielle (IA) vers la météorologie**. Néanmoins, au-delà du simple enjeu d'accessibilité, et parce que l'ouverture ne doit pas être considérée comme une fin en soi, la valorisation des données publiques de l'établissement suppose **des démarches d'accompagnement des utilisateurs** dans l'appropriation des données météorologiques. **Cet enjeu va aller croissant avec la dilatation du volume de données mises en ligne et reste très largement à explorer et à traiter**.

**La mise en ligne des données publiques de l'opérateur s'accélère. Sur la seule année 2021, le volume de ses données accessibles gratuitement sur son portail doit être multiplié par trois**. Le volume de données quotidiennes mises en lignes gratuitement doit ainsi passer de 600 à 2 000 giga-octets (Go). À horizon 2025, le volume devrait avoir plus que décuplé avec un objectif de 7 000 Go de données.

Aujourd'hui, l'intérêt de Météo-France serait de **parvenir à maîtriser la progression de la mise à disposition de ses données** pour que celle-ci ne vienne pas compromettre sa situation financière. **Cependant, l'opérateur est tenu par des normes européennes et nationales de mettre en ligne gratuitement de nouvelles données dont certaines étaient jusqu'ici soumises à des redevances de réutilisation**.

**Sur le fondement de l'article L.324-1 du code des relations entre le public et l'administration, la circulaire du premier ministre n° 6224/SG du 27 avril 2021 prévoit notamment l'extinction, d'ici 2023, des redevances de réutilisation. Le droit européen, en introduisant la notion de « jeux de données à forte valeur » dites HVD (High Value Dataset)<sup>1</sup>, qui s'applique notamment au domaine météorologique, va imposer, à Météo-France, dans un délai de deux ans, d'assurer la mise en ligne de nouveaux jeux de données accessible via des API (interface de programmation d'application).**

**Un autre sujet d'inquiétude pour Météo-France doublé d'un enjeu de souveraineté tient à la perspective d'ouverture des codes de calculs des modèles de PNT**. Les codes de calcul du SMN français sont réputés être parmi les tout meilleurs au monde et leur mise à disposition pourrait être

---

<sup>1</sup> À travers la directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019.

exploitée par de grands groupes privés ou des puissances étrangères disposant de capacités de calcul intensif supérieures à celle de Météo-France.

### **3. Un effet ciseau financier qui pourrait affecter le budget de Météo-France de 3 millions d'euros d'ici 2025**

**Le processus de développement et d'accélération de l'ouverture des données se traduit par un effet ciseau susceptible d'affecter la situation financière fragile de l'établissement. D'une part il induit une diminution des ressources** de l'opérateur à travers l'extinction des redevances de réutilisation **et d'autre part il suppose un accroissement des dépenses** nécessaires à la mise en ligne des données dans de bonnes conditions d'accès.

Au regard des perspectives de développement du volume de données mises en ligne, **les coûts supplémentaires induits sont estimés par l'opérateur entre 0,6 et 1,4 millions d'euros annuels** selon le niveau de qualité d'accès retenu<sup>1</sup>. **Les baisses de recettes** résultant de l'extinction des redevances de réutilisation sont quant à elles estimées à **environ 1,6 millions d'euros** par l'établissement. **Le cumul** des pertes de recettes et des hausses de dépenses pourrait ainsi atteindre jusqu'à **3 millions d'euros à horizon 2025**.

Par ailleurs, l'essor de la gratuité et le développement de la concurrence qui pourrait résulter de l'accroissement du volume de données libres d'accès pourrait provoquer **des phénomènes indirects** conduisant à **une déflation sur le marché des services météorologiques**. Cette conséquence serait susceptible d'affecter encore davantage le budget de Météo-France à travers une **contraction de ses recettes commerciales**. L'estimation de ces pertes est très difficile à ce stade et doit être interprétée avec toute la prudence requise. Néanmoins, **selon Météo-France, ce phénomène indirect de l'ouverture des données pourrait, à terme, affecter encore davantage son budget que les conséquences financières directes**. Sur un volume de recettes commerciales de 29,5 millions d'euros en 2020, s'il venait à se matérialiser, l'effet serait alors loin d'être négligeable.

### **4. Une compensation de l'État est nécessaire et des pistes de mutualisation doivent être recherchées**

**Une compensation financière de l'État sera nécessaire pour que le coût des obligations nouvelles imposées à Météo-France en termes d'ouverture des données publiques ne crée pas un effet d'éviction au détriment des dépenses d'investissements ou de recherche de l'opérateur.**

---

<sup>1</sup> La mise en ligne via API, un standard de plus en plus généralisé car facilitant la récupération des données nécessite des développements importants et onéreux.

Le rapporteur a pris note que cette question sera traitée dans des arbitrages prochains portant sur les enjeux de l'open data qui ne sont pas propres à Météo-France. Le 26 mai 2021, devant la commission des finances du Sénat, le rapporteur a également accueilli favorablement les propos de la ministre de la transition écologique, Madame Barbara Pompili, qui a indiqué que les dépenses liées à la mise à disposition de ses données publiques par Météo-France pourraient être au moins pour partie couverts par l'État.

En parallèle, pour réduire et mutualiser les coûts liés à la mise en ligne de nouvelles données, des partenariats européens doivent être explorés dans le but de standardiser les formats d'échanges d'informations ou encore d'assurer l'interopérabilité des plateformes d'accès.

Avec le CEPMMT et EUMETSAT, Météo-France est d'ores et déjà engagé dans la phase pilote d'un projet de mutualisation des infrastructures de technologie de cloud qui seront nécessaires à la mise à disposition de très gros volumes de données.

**Recommandation n° 23 :** promouvoir et concrétiser les projets de coopérations européens destinés à mutualiser les infrastructures informatiques, définir des standards ou encore harmoniser les règles relatives aux transferts et à la mise en ligne des données publiques.

**Recommandation n° 25 :** Compenser au moins pour partie les surcoûts engendrés par l'ouverture des données publiques.

### *C. L'APPROPRIATION DES NOUVELLES OPPORTUNITÉS NUMÉRIQUES EST UN ENJEU CRUCIAL POUR MÉTÉO-FRANCE*

#### **1. Les opportunités liées à la révolution numérique doivent être saisies de façon anticipée par Météo-France**

Les potentiels de l'IA en matière de météorologie apparaissent très prometteurs et son appropriation par Météo-France constitue un enjeu majeur à court et moyen termes. Ce sera probablement l'une des conditions pour que l'opérateur français reste dans la course au niveau international.

Les principales opportunités liées à l'IA portent sur l'assimilation et le traitement de volumes de données toujours plus importants. Les apports de l'IA en termes d'aide à la décision en présence de telles masses

# DOCUMENT 9 :

## La nouvelle offre de données publiques

### Les grands principes :

- **Une offre de données publiques de base sans redevances de réutilisation ni Frais de Mise à Disposition:**
  - ciblées sur les données HVD
  - accessibles gratuitement par API (*hypothèse où c'est la mise à disposition des packages qui conduit aux coûts de mise à disposition les plus importants*)
- **Une offre complémentaire en ligne composée :**
  - de packages en ligne des données HVD avec perception de frais de mise à disposition sur le principe d'un coût unique par Go.
  - de données publiques à valeur ajoutée (i.e. données SAFRAN) avec perception de frais de mise à disposition sur le principe d'un coût unique par Go.
  - de services de fourniture de données publiques à valeur ajoutée tels que :
    - des API « évoluées »
    - le Service Premium pour les données de modèle
- **Une offre offline de fourniture à la demande avec FMD + Frais de mise en place :**
  - pour répondre aux demandes de données non disponibles en ligne

17/06/2022

COS Digital & Données

6/15

## La nouvelle offre de données publiques (à ajuster selon le périmètre final des HVD)

	Offre de base sans FMD	Offre complémentaire avec FMD ou <b>tarif spécifique</b>
Observations	API par station	Packages existants (pack RADOME...)
Données climatologiques	API par station de données observées et moyennes journalières	Packages de données observées et moyennes journalières et de données SAFRAN <i>(P1) API et API « évoluées » (à définir) des autres catégories de données climatologiques (données type SAFRAN, séries homogénéisées, ETP spatialisées, produits finalisés climatologiques)</i>
Vigilance	API dernière production de la Vigilance. <i>Packages (à définir)</i>	<i>Packages (à définir)</i>
Radars	API par radar & composite nationale (réflectivité, rétrodiffusion, polarisation, précipitations, vent et echotops)	<i>Packages (à définir)</i>
Modèles	API à pleine résolution spatiale par pas d'une heure de prévision pour les « meteorologically relevant parameters and levels » : <ul style="list-style-type: none"><li>- API existantes (modèles déterministes)</li><li>- nouvelles API pour les nouveaux modèles</li></ul>	Packages existants et packages des nouveaux modèles <i>(P1) Service Premium avec téléchargement dans un délai donné.</i> <i>(P2) En option : API évoluées (à définir)</i>
Données expertisées	<i>Données fournies à la demande : aucun utilisateur excepté la D2C</i>	

17/06/2022

COS Digital & Données

7/15  
Page 26/32

## Généralités sur les API (Interface de programmation d'application)

**Définition** : (source : <https://www.cnil.fr/fr/definition/interface-de-programmation-dapplication-api>)

Une API (*application programming interface* ou « interface de programmation d'application ») est une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service afin d'échanger des données et des fonctionnalités.

-----

**La monétisation des API, qu'est-ce que c'est ?** (EXTRAIT d'un article publié en 2019 par l'entreprise informatique « Red Hat », source <https://www.redhat.com/fr/topics/api/what-is-api-monetization>)

### ➤ Présentation

La monétisation des API désigne le processus par lequel les entreprises génèrent des revenus grâce à leurs API (interfaces de programmation d'application). Les API représentent un enjeu majeur de ce que beaucoup considèrent comme l'avenir du développement des entreprises dans l'économie numérique. Bien conçues, elles seront indispensables pour nouer et entretenir des relations commerciales. Elles vous assurent une présence globale sur le Web en permettant aux autres entreprises d'accéder à vos données et à vos ressources pour les intégrer à leurs applications et sites publics ou privés.

En principe, si vous avez mis en place un plan de gestion des API, cela signifie que vous disposez déjà d'un modèle économique solide, qui constitue un cadre de référence pour vos objectifs de monétisation. La monétisation des API ne consiste pas uniquement à déterminer comment générer des revenus via vos API, mais aussi comment vous assurer de leur fonctionnement et de leurs performances pour les utilisateurs.

Toutes les API ne sont pas créées pour répondre aux mêmes objectifs et le déploiement d'une API peut revêtir des formes très diverses. Voici quelques exemples de schémas courants en matière de monétisation des API.

### ➤ Gratuit

Pour donner accès à leur API, de nombreuses entreprises offrent un niveau gratuit qui permet à n'importe qui de s'inscrire, de commencer à l'utiliser et d'en découvrir la valeur. Les utilisateurs peuvent ainsi l'évaluer et s'assurer qu'elle répond à leurs besoins avant d'engager des dépenses. Si la *gratuité* est une bonne option pour de nombreuses stratégies de monétisation des API, elle fonctionne aussi très bien en association avec d'autres stratégies. Évitez de proposer uniquement une version gratuite de l'API, car si votre entreprise ne dispose d'aucune stratégie pour vendre des services aux utilisateurs plus exigeants, elle risque de se heurter à des difficultés.

### ➤ Paiement par le client

Après l'accès gratuit, une autre stratégie de monétisation consiste à définir un prix que le client paie pour utiliser les services ou les ressources offerts par votre API. Voici les trois principaux types de tarification appliqués aux API :

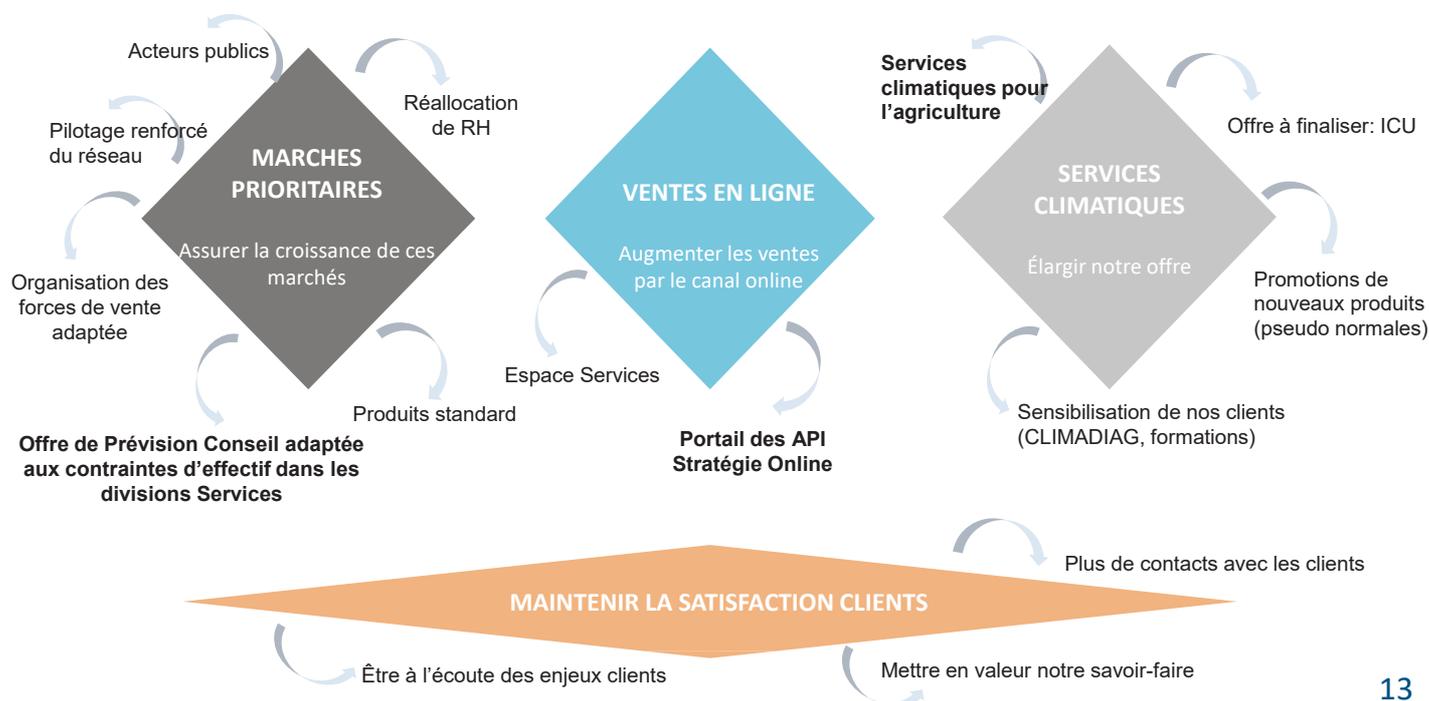
- **La tarification par niveau.** Certains fournisseurs d'API mettent en place plusieurs niveaux d'accès payant, par exemple Bronze, Or et Platine. Chaque niveau offre des services et des quotas spécifiques pour l'accès aux ressources de l'API ; le prix augmente en fonction du niveau.
- **Paiement à la consommation.** Une autre option se base sur la consommation d'un service : les utilisateurs de l'API paient pour ce qu'ils utilisent. Les fournisseurs établissent leurs factures sur la base de la quantité de bande passante, du stockage et d'autres frais fixes liés à l'utilisation de l'API, puis y intègrent leurs propres coûts ainsi qu'une marge de profit.
- **Paiement par unité.** Enfin, d'autres fournisseurs d'API considèrent chaque ressource de l'API comme une unité à laquelle ils attribuent un coût. Les clients de l'API paient pour le nombre d'unités qu'ils prévoient d'utiliser et ont la possibilité d'en acheter plus au besoin.

Certains fournisseurs d'API associent différentes combinaisons de tarification par niveaux, à la consommation et par unité afin de couvrir leurs coûts d'exploitation et de générer des revenus.

# DOCUMENT 11 :

Extraits de la présentation du Plan d'Actions Commerciales 2023 de Météo-France  
Source : <https://intramet.meteo.fr/d2c/pac-2023>

## Nos actions doivent s'inscrire dans la stratégie commerciale



13

## Focus API

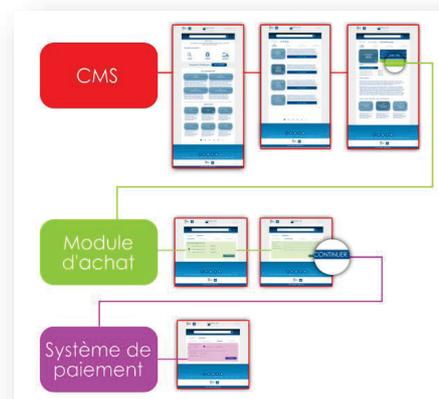
Le portail on line des API (données publiques et commerciales)

➔ Externalisation des développements de l'APIM (à partir de l'instance actuelle) + TMA/MCO -> AO lancé bientôt

- Portail / Site Internet unique pour l'Établissement permettant de commander et de payer en ligne des abonnements à des API de données publiques et commerciales.
- API de données publiques, vendues au volume afin de recouvrir les coûts de mise à disposition
- API commerciales, commercialisées par la vente à la requête
- Les API de Météo-France pourront donc être mises à disposition :
  - « en ligne » : Paiement par CB ou Prélèvement, facturation avec PEP Premium
  - Ou
  - « hors ligne » pour certaines API commerciales
- Mise en opérationnel du portail des API avec monétisation « en ligne », T4 2023

➔ En parallèle finalisation de la stratégie commerciale sur les API :

- Mise en place de la médiation avec la DSI
- Propositions de tarification
- Formation des RCR
- Stratégie des API Online



## DOCUMENT 12 :

### **EUMETSAT, l'agence européenne de satellites météorologiques, s'apprête à lancer un programme pilote innovant visant à acheter, pour la toute première fois, des données à un opérateur commercial de flotte de satellites à des fins de prévision météorologique.**

Publié le 21 juillet 2021

Le Conseil d'EUMETSAT a approuvé l'acquisition de données auprès de Spire Global Luxembourg S.a.r.l. pour un montant maximum de 9 millions d'euros répartis sur trois ans.

EUMETSAT réceptionnera les données, dites de radio-occultation, de Spire, les traitera et les diffusera pour qu'elles soient utilisées dans des modèles de prévision météorologique.

*« Les données de radio-occultation nous renseignent sur la température et la quantité d'humidité dans l'atmosphère. Elles sont précieuses pour améliorer la précision des prévisions météorologiques », a déclaré Phil Evans, Directeur général d'EUMETSAT.*

*EUMETSAT fournit déjà aux services météorologiques de ses États membres des données de radio-occultation provenant d'un instrument embarqué à bord de ses satellites Metop évoluant en orbite basse, mais des recherches ont montré que la multiplication de ce type de données augmentait la précision des modèles de prévision météorologique.*

*En décidant d'acheter ces données auprès d'un fournisseur commercial dans le cadre d'un programme pilote, EUMETSAT et ses États membres se donnent les moyens d'évaluer le rapport coûts/bénéfices des solutions proposées par le secteur du New Space. »*

Les données de radio-occultation du Système mondial de navigation par satellite (GNSS) offrent une lecture instantanée de l'état de l'atmosphère. Ce résultat est obtenu en mesurant la réfraction d'un signal émis par un satellite GPS lorsqu'il traverse l'atmosphère en direction d'un satellite en orbite basse. Ce sont la température et la quantité d'humidité dans l'atmosphère qui déterminent l'ampleur de cette réfraction.

Spire exploite une flotte qui compte plus de 100 nanosatellites en orbite basse.

En procédant au traitement et au contrôle des données du secteur privé, EUMETSAT garantira la cohérence avec d'autres sources de données de radio-occultation. L'Organisation n'aura besoin d'aucun autre matériel ou logiciel pour traiter, diffuser et archiver ces données supplémentaires.

Aux termes du contrat qu'a approuvé son Conseil, EUMETSAT recevra de Spire les données en temps quasi réel grâce à une licence mondiale l'autorisant à les partager immédiatement avec toute tierce partie. Cette approche, que soutiennent résolument les États membres d'EUMETSAT, est en phase avec les discussions en cours sur les politiques en matière de données de l'Organisation météorologique mondiale (OMM), qui préconisent d'élargir et de renforcer l'échange international libre et sans restriction des données du système Terre.

*« Sachant que de nombreux secteurs économiques dépendent de la précision des prévisions météorologiques, EUMETSAT est résolue à explorer tous les moyens efficaces pour évoluer et innover, afin de fournir aux Services météorologiques de nos États membres les données dont ils ont besoin pour protéger leurs populations », a déclaré M. Evans.*

*« Notre capacité à nous adapter aux besoins des États membres est particulièrement importante à une époque où le dérèglement climatique impacte nos systèmes météorologiques. »*

Source : <https://www.eumetsat.int/fr/une-premiere-eumetsat-va-acheter-des-donnees-meteorologiques-un-fournisseur-commercial>

# Une mission dédiée à l'intelligence artificielle vient d'être lancée au sein de l'établissement

Par [Xavier Freymuth](#) — Date de publication : 19/03/2021 08:05 — Dernière modification : 19/03/2021 08:05

Contributeurs : Sarah Bardis



*L'intelligence artificielle (IA) est un enjeu majeur pour Météo-France qui a déjà entrepris ces dernières années plusieurs initiatives. Pour poursuivre ces progrès, Alain Joly a été chargé d'une mission de développement et de promotion de l'utilisation des méthodes d'intelligence artificielle pour tout l'établissement. Il a répondu à nos questions sur le sujet. Entretien.*

Au cours des deux dernières décennies, la collecte, la circulation et l'analyse de grands nombres de données ont connu un essor formidable. La météorologie est basée sur le traitement d'un nombre de données sans cesse croissant dont on obtient des prévisions météo et des services climatiques toujours plus pertinents.

**Alain Joly, vous êtes chargé d'une mission sur l'intelligence artificielle. Quel est votre parcours à Météo-France ?**

Je travaille à Météo-France depuis les années 1980 sur des sujets qui touchent à la prévision du temps, numérique ou pas. Je suis passé par l'université de Reading, puis j'ai coordonné une campagne de mesures internationale à l'échelle de l'océan Atlantique Nord. Les tempêtes de décembre 1999 m'ont conduit à me pencher sur la question de la prévisibilité.

Ces dernières années, j'étais responsable du groupe de prévision numérique du CNRM, le GMAP (Groupe de modélisation et d'assimilation de données pour la prévision). Ce groupe est tout à la fois un service opérationnel qui fait évoluer les systèmes Arome et Arpege de Météo-France dans un cadre coopératif assez unique, et un service de recherches. La co-évolution d'un code informatique riche mais lourd et les transferts majeurs vers l'opérationnel dont nous nous occupons sont des sujets essentiels. Ce groupe est aussi le premier utilisateur du calculateur scientifique de Météo-France. À ce titre, j'ai été beaucoup impliqué dans la phase initiale de dimensionnement des moyens de calcul dont la mise en place s'achève en ce moment.

**En quoi consiste votre mission ?**

L'objectif est d'élaborer une « proposition de plan d'action global en vue de favoriser la diffusion, le partage et la formation à l'intelligence artificielle au sein de l'établissement et de nous faire progresser dans la valorisation des données existantes ou nouvelles au moyen de cette technique ».

**Qu'est-ce que l'intelligence artificielle et comment sert-elle la prévision du temps et l'étude du climat ?**

L'histoire des sciences prête au mathématicien britannique Alan Turing l'origine de l'intelligence artificielle. Turing a posé les bases logiques du processeur informatique avant d'en créer des formes électro-mécaniques avec lesquelles les Britanniques ont pu casser les codes de messages allemands pendant la Seconde Guerre mondiale (Voir le MétéoTips sur ce sujet [ici](#)). Et par ailleurs, il s'est intéressé à la reproduction du langage par automate ou encore à reproduire la prise de décision dans une partie de jeu d'échecs : deux des thèmes classiques de l'intelligence artificielle.

[Le rapport de la Commission dite Villani Donner un sens à l'intelligence artificielle de 2018](#) présente l'intelligence artificielle dans le contexte actuel comme un « *programme, fondé autour d'un objectif ambitieux : comprendre comment fonctionne la cognition humaine et la reproduire ; créer des processus cognitifs comparables à ceux de l'être humain.* »

« *Nous sommes dans un monde [...] de données. Ces données qui sont au cœur du fonctionnement des intelligences artificielles actuelles. [...] Ces technologies [...] déterminent notre capacité à organiser les connaissances, à leur donner un sens, à augmenter nos facultés de prise de décision et de contrôle des systèmes. Et notamment à tirer de la valeur des données. L'intelligence artificielle est donc une des clés du pouvoir de demain dans un monde numérique.* ». C'est pourquoi on parle aussi de « science des données » pour l'intelligence artificielle, et c'est peut-être la bonne dénomination.

Fondées sur l'observation, la météorologie et le climat sont des sciences de données. Météo-France collecte et archive des données d'observations depuis le 19<sup>e</sup> siècle. La quantité de données observées ne cesse de croître avec l'avènement de la télédétection. L'arrivée de multiples projections climatiques ou celle des ensembles en assimilation de données et en prévision sont aussi à l'origine de l'explosion des volumes. Enrichir les systèmes de prévision soit en les adossant à des observations toujours plus diversifiées, soit en leur permettant de représenter davantage de milieux (au-delà de l'atmosphère) et de processus est un autre facteur d'enrichissement des données.

Ces nouvelles générations d'instruments et la montée en puissance de la prévision d'ensemble constituent des sources de données volumineuses qu'il convient de valoriser pleinement au bénéfice de la qualité des prévisions et des services que Météo-France peut proposer.

**L'IA représente donc un enjeu stratégique majeur pour l'établissement. Où en est-on actuellement ?**

Une des richesses de Météo-France est son personnel doté d'une forte culture scientifique : la curiosité, l'intérêt pour de nouvelles idées ou de nouvelles approches est une dimension de cette culture. Ici et là dans l'établissement, des collègues sont devenus d'excellents connaisseurs de méthodes que l'on rattache à la science des données et s'en servent depuis longtemps.

Il y a donc, à Météo-France, un certain nombre de réalisations, de travaux de recherche publiés ou en cours, fondés sur des méthodes qui relèvent de l'intelligence artificielle. Nous consacrerons d'ailleurs prochainement une série d'articles sur l'intelligence artificielle et à ses applications à Météo-France.

La DirOP a ainsi introduit dans le post-traitement des sorties des modèles de prévision, des méthodes d'apprentissage qui apportent des améliorations significatives sur certains paramètres sensibles, en particulier la température. Elle s'investit dans le traitement des ensembles de prévision au bénéfice du projet 3P. L'établissement dispose aussi du LabIA, rattaché actuellement à la DSI.

**Au sein de Météo-France, quels sont les domaines qui pourront bénéficier le plus vite de ses évolutions technologiques ?**

En termes de réalisations, les plus avancées (au sens d'une utilisation opérationnelle) tournent autour des projets de transformation de la prévision 3P : alimentation de la base dite « Alpha », production de bulletins. Une application sur le point de déboucher entre dans le domaine de la prévision immédiate.

Et il faut retenir que, face à l'augmentation aussi bien des observations que des simulations, des méthodes de sciences des données peuvent apporter quelque chose à toutes les étapes de la chaîne de traitement, absolument toutes.

# Lancement de l'incubateur de start-up « MétéoFab » de Météo-France

Par [Emilie Gilberton](#) — Date de publication : 12/07/2021 09:00 — Dernière modification : 12/07/2021 11:41  
Contributeurs : Matthieu Philippot



*Depuis le début de l'année 2021, Météo-France assure la gestion de l'incubateur de start-up hébergé sur la Météopole, lancé en 2017 par le ministère. L'occasion de lui donner un nouveau nom « Météo Fab » et une nouvelle ambition au service de l'innovation. Explications.*

Depuis le 1<sup>er</sup> janvier 2021, l'établissement dispose de son propre incubateur de start-up. Renommé pour l'occasion Météo Fab, il se substitue à l'incubateur « GreenTech Verte » qui était hébergé depuis 2017 sur le site de la Météopole, dans le cadre d'un dispositif national mis en place et financé par le ministère de la Transition écologique.

### **Valorisation de l'expertise et des données de Météo-France**

En appuyant de nouveaux acteurs dans le champ des politiques publiques de l'établissement, Météo Fab se donne notamment pour objectif de favoriser l'utilisation des données produites par Météo-France afin de maximiser leur utilité pour la collectivité, tout en contribuant au développement de l'innovation dans l'établissement.

L'incubateur vise ainsi à héberger des projets innovants développant des produits et des services en lien avec les prévisions, les données et les techniques d'observation météorologiques et climatologiques. En partageant son expertise scientifique et technique, l'établissement vise ainsi à créer une relation mutuellement bénéfique.

Les start-up sont sélectionnées au fil de l'eau, en s'appuyant sur l'expertise des directions métiers pour évaluer les candidatures reçues.

En complément des entreprises Waves'n See, OgoXe et HD Rain déjà hébergées, deux nouvelles start-up rejoignent ainsi l'incubateur :

**Keyros** développe une application informatique (appelée « Projections ») visant à accompagner les gestionnaires de crises. Le caractère innovant du projet réside notamment dans l'utilisation du temps comme variable dans la représentation de la situation et la visualisation de l'impact des plans d'actions sur toute la durée de la crise.

**Lean Connected** connecte les territoires isolés aux réseaux Internet et téléphone grâce à des modules autonomes solaires sans empreinte carbone, notamment pour des habitations isolées non desservies par les réseaux publics et des sites professionnels (stations de contrôle, télémaintenance, refuges, stations météo, etc.)