

Annexe B - Observations météorologiques d'Environnement et Changement climatique Canada dans Wikimedia Commons

Rapport de la séance du 10 février 2021

Rapportrice : Camille Vézy pour Wikimedia Canada

Relecture : Ha-Loan Phan, Wikimedia Canada; Pierre Choffet pour Wikimedia Canada;
Miguel Tremblay, ECCC

Mercredi 10 février 2021, de 9 h à 11 h : Remue-méninges

Présences	1
Déroulement de la séance	2
Résumé des idées générées	3
Visualisations de données	3
Valorisation descriptive des données météo pour « augmenter » les informations relatives à des événements, des lieux ou des personnes	6
Valorisation prescriptive des données météo pour aider à la décision dans différents secteurs et d'autres pays	9
Création de pages Méta pour inviter à reproduire le projet	10
Conclusion	11

Présences

- Organiseurs:
 - Pierre Choffet et Ha-Loan Phan (pour Wikimedia Canada)
 - Miguel Tremblay (Environnement et Changement climatique Canada)
 - Partenaires de diffusion : Acfas (Johanne Lebel, Frédéric Macé) et IVADO (Guillaume Chicoisne)
- Participants :
 - Jean-Hugues Roy, professeur à l'École des médias, en journalisme des données, à l'UQAM
 - Nadia Tahiri, postdoctorante en apprentissage automatique, intelligence artificielle et santé publique, à l'Université de Montréal
 - Pascal Martinolli, bibliothécaire à l'Université de Montréal

- Sandrine Edouard, coordonnatrice nationale des données ouvertes du Service météorologique, à Environnement et Changement climatique Canada
- Simon Villeneuve, wikimédien et professeur de physique au Cégep de Chicoutimi
- Trevor Smith, spécialiste scénarios et services climatiques chez Ouranos
- Christian Jauvin, programmeur - Plateformes climatiques, données et opérations chez Ouranos
- Jean-Christophe Voyer, Stratégie et développement des affaires, cartographie et outils d'aide à la décision, chez Anagraph
- Aurélie Dussenne, analyste de données chez ImpactBee
- Personne anonyme, climatologue d'une organisation météorologique gouvernementale française
- Benoit Rochon, administrateur de Wikimedia Canada

Animation des discussions par Sébastien Paquet (Service Now) et Ha-Loan Phan (Wikimédia Canada)

Les discussions se déroulent en français.

Déroulement de la séance

L'objectif de la séance est de faire ressortir des idées originales de valorisation des données, d'outils à créer dans les projets de la Fondation Wikimedia, réalisables à court et à long terme.

Certaines des idées à court terme seront réalisées par Pierre Choffet d'ici le 31 mars 2021, date de fin du projet « Observations météorologiques d'Environnement et Changement climatique Canada dans Wikimedia Commons ».

La séance s'est déroulée en quatre temps :

- 1) **Introduction** (10 min) : L'équipe d'organisation et d'animation a brièvement introduit la séance en **présentant les parties prenantes**, en rappelant le **but** du projet et de ce remue-méninges en particulier et en annonçant le déroulement de la séance. Avant de diviser le groupe en deux sous-groupes de discussions, les **règles** du remue-méninges ont été annoncées : se concentrer sur la quantité d'idées, pas de critique, encourager les idées folles, combiner et améliorer les idées.

Pierre Choffet a également partagé quelques liens pour :

- Rappeler quelles données sont disponibles actuellement (les données [mensuelles](#) et d'[almanach](#) sur Wikimedia Commons). Pour un résumé du projet sur Wikidata : https://www.wikidata.org/wiki/Wikidata:WikiProject_Weather_observations

- Donner des exemples de ce qu'il serait possible de faire, comme automatiser le [tableau de la section "climat" sur la page Wikipédia de Montréal](#) ou encore réaliser des visualisations telles que des histogrammes, des diagrammes à barre, y ajouter de l'interactivité, etc. comme ce qui se fait sur la [page Wikipédia de la pandémie de Covid-19 en France](#). Un exemple un peu plus complexe, qui serait à plus long terme, serait de réaliser des diagrammes interactifs comme l'outil [Year selector](#) pour la ville de New York qui permet de sélectionner les données météo des années qui intéressent l'utilisateur.

2 et 3) **Discussions** (2 x 30 min) : Les participants et participantes ont été répartis en 2 sous-groupes de 6 personnes (4 hommes, 2 femmes dans chaque sous-groupe) pour réfléchir à des idées d'outils à créer à court et long terme.

Après un rapide **brise-glace** en sous-groupe où chaque participant a fait part de sa saison préférée, le groupe « Hiver », animé par Sébastien Paquet, a commencé par des **idées à court terme** tandis que le groupe « Printemps », animé par Ha-Loan Phan, a réfléchi à des **idées à long terme**.

Au bout de 30 minutes, les 2 groupes ont été réunis en une courte **plénière** pour partager les principales idées.

Chaque groupe est ensuite reparti pour 30 minutes de **discussion** en inversant les thèmes : générer des idées à long terme pour le groupe « Hiver » et à court terme pour le groupe « Printemps ».

4) **Conclusion** (15 min) : les deux groupes sont une dernière fois rassemblés en plénière et les participants ont été invités à faire un retour d'expérience sur la matinée.

Résumé des idées générées

Les deux groupes ont abordé des idées similaires : imaginer des façons de visualiser les données; « augmenter » les informations sur les événements, lieux et personnes par l'ajout des données météo correspondantes, en particulier via les photos sur Wikimedia Commons; utiliser les données météo pour aider à la décision dans différents secteurs et pays; et à un niveau plus méta, partager le processus du projet pour inviter à le reproduire.

La valorisation des données météo a ainsi été envisagée sous l'angle de la visualisation - à la fois par graphiques et par les photos sur Wikimedia Commons - dans une perspective descriptive (améliorer l'information) et prescriptive (aider à la décision). Le fort potentiel de sensibilisation du grand public aux changements climatiques par la visualisation des données météo a notamment été exploré.

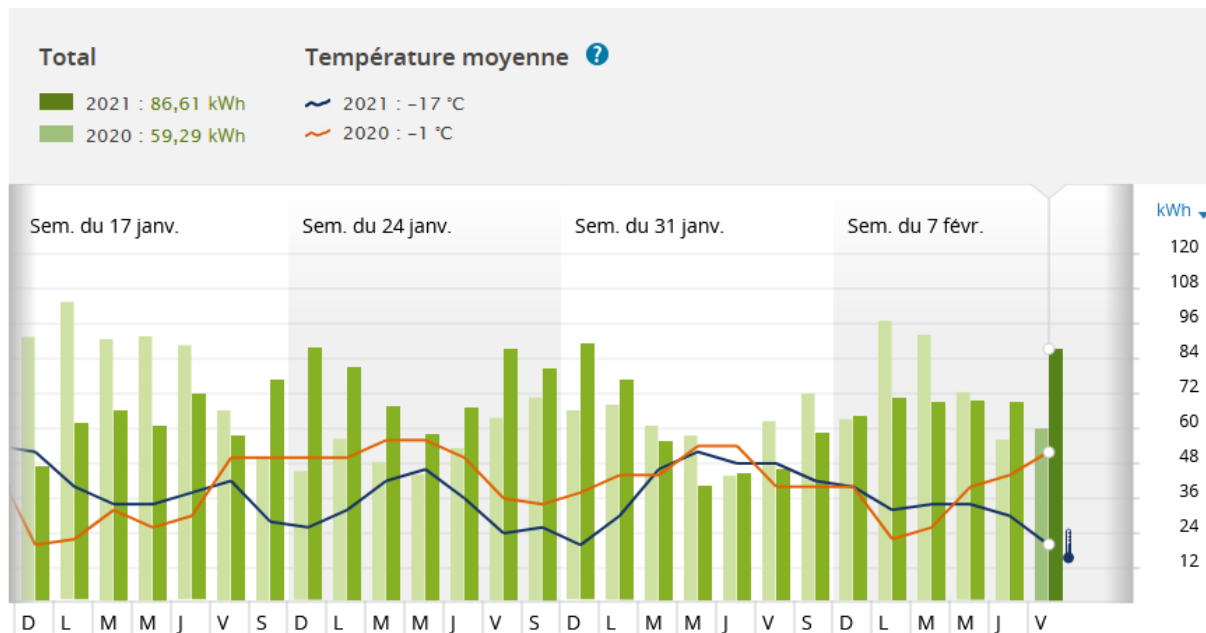
1) Visualisations de données

Des **graphes** à l'image de [ceux réalisés pour le climat de la ville de New York](#) pourraient être développés **pour toutes les localités proches des stations**. Un modèle pourrait permettre d'aller chercher les températures de façon automatisée, mais il reste encore à définir quel modèle intégrer.

D'autres types de graphes pourraient être développés **pour suivre l'évolution du climat, afin de sensibiliser aux changements climatiques** :

- **Des graphes sur une valeur mensuelle et/ou annuelle liée au climat** (température, précipitations, etc.) pourraient permettre d'avoir une visualisation de ces évolutions.
- Il serait possible de **calculer les nouvelles normales dans différentes régions**. Par exemple, en France, ce type de graphes existe et permet de voir une augmentation de 0,4 degrés dans certaines régions.
- Sur un **graphe interactif**, on pourrait permettre à un usager de **choisir une date de début et une date de fin pour voir l'évolution du climat sur cette période dans tel lieu**¹.
- Dans le même ordre d'idée, pour l'instant sur Wikipédia, on ne peut pas vraiment **comparer** l'année actuelle pour voir où on en est par rapport à la moyenne, à la température maximale, minimale, la médiane... Il serait intéressant d'avoir accès à une façon plus visuelle de comparer l'année et le mois actuels par rapport au même mois plusieurs années auparavant, un peu comme l'outil de comparaison utilisé par Hydro-Québec dans son interface client de comparaison d'énergie.

¹ Il faudrait toutefois préciser des détails méthodologiques, car certaines évolutions dans les données brutes peuvent être liées à des changements liés aux mesures (un changement d'outil de mesure par exemple).



- Des [graphes en rond](#) pourraient aussi permettre de visualiser l'évolution du climat.
- Ces graphes pourraient être **liés à une carte et aux données démographiques correspondant** à ce point sur la carte. On pourrait ainsi voir combien de personnes sont touchées dans telle zone par les changements de précipitations, de chaleur, etc.
 - En ajoutant la **géolocalisation**, quelqu'un qui se connecte sur Wikipédia pourrait voir, selon la zone où il ou elle est, un impact "personnalisé" des changements climatiques. Cela permettrait de se rendre compte de changements à un niveau très local, plutôt que de penser que c'est loin et que ça n'arrive qu'aux autres.
- Des visualisations de données pourraient également permettre de voir s'il y a des **corrélations entre des phénomènes météorologiques et des phénomènes de santé publique** :
 - À partir de cartes géographiques où sont situées les [zones bleues](#) (de longévité) et celles de forte mortalité, des corrélations pourraient être faites avec les données météo. On pourrait ainsi visualiser s'il existe des liens entre l'espérance de vie, le climat et les changements climatiques dans ces zones.
 - Des **graphes d'analyse de sentiment à partir de publications sur les réseaux sociaux** pourraient être liés à des données météo pour voir s'il y a des corrélations entre des sentiments et des phénomènes météorologiques. Cela permettrait par exemple de visualiser le phénomène de dépression hivernale, et éventuellement de prévoir des campagnes de prévention du suicide ou d'accompagnement en **santé mentale**.
 - En matière d'**épidémiologie**, les graphes visualisant l'évolution d'épidémies pourraient être liés aux données météo dans les zones concernées pour voir les corrélations entre circulation des virus et climat selon les périodes

historiques. Peut-être que des corrélations pourraient être vérifiées avec les indicateurs des changements climatiques.

Pour développer ces graphes, plusieurs pistes ont été suggérées :

- Comme les **outils de visualisation** sont pour l'instant limités **sur Wikipédia**, il serait intéressant d'**améliorer les engins existants**. Cela pourrait commencer par une **liste** d'améliorations à faire, classées selon un ordre de priorités et de fonctionnalités, pour développer un outil plus polyvalent.
- **Crowdsourcing** : des **partenariats** avec les différentes communautés en analyse de données existantes pourraient permettre de trouver des façons créatives de visualiser ces données avec impact. Par exemple, [Kaggle](#) organise des mini-compétitions pour créer de nouvelles sortes de graphes à partir de jeux de données. Des synergies intéressantes pourraient émerger de partenariats entre Wikimedia et ces communautés.
 - Pour maintenir ces communautés de sciences des données dans la boucle wikipédienne, on pourrait instaurer un système de suivi automatique de certains mots clés (comme « météo », « permafrost », ...) liés à des pages concernant le Canada sur Wikipédia. Ces communautés seraient ainsi invitées régulièrement (voire automatiquement?) à contribuer à des visualisations créatives de ces thèmes et jeux de données.

2) Valorisation *descriptive* des données météo pour « augmenter » les informations relatives à des événements, des lieux ou des personnes

Les deux groupes ont suggéré des idées à propos de l'ajout des données météo à des informations déjà datées et localisées, dans les fiches et photos déjà existantes dans Wikipédia. Les données météo pourraient être ajoutées soit dans les boîtes d'information de certaines fiches (villes, événements), soit dans les métadonnées des photos dans Wikimedia Commons. Les liens créés entre catégories d'événements, informations météo, et documents qui s'y rattachent permettraient en même temps d'améliorer l'information quant aux changements climatiques - ce qui pourrait contribuer à étoffer l'information sur la page Wikipédia du [changement climatique](#) (voire du [réchauffement climatique](#)).

Pour faire cela, plusieurs pistes sont suggérées :

- Comme il y a beaucoup de **pages liées à des lieux et des dates**, il serait possible de **cibler ces pages et indiquer le temps qu'il faisait**.
 - Par exemple, pour une journée précise d'une année précise : https://fr.wikipedia.org/wiki/27_juillet_1945 ou une journée précise sans année : https://fr.wikipedia.org/wiki/27_juillet. De façon ludique, cela

permettrait de retrouver le temps qu'il faisait à des dates comme le jour de sa naissance, le jour du mariage de ses grands-parents, etc. Cela contribuerait également à la **reconstitution ou la mise en contexte d'événements historiques**.

- Pour un **événement**, par exemple « crise du verglas », « JO de Montréal », « éclipse lunaire », on pourrait préciser les informations météo, selon les données de la station météo la plus proche de cet événement.
 - On pourrait donc cibler en priorité les **événements qui dépendent des conditions météorologiques** pour y ajouter les informations météo.
- On pourrait cibler des **personnages historiques** canadiens entre 1840 et 2018 (balises des données météo qui sont sur Wikimedia Commons à l'heure actuelle dans le cadre de ce projet) et sur chaque fiche bibliographique (dans les infoboîtes, par exemple), préciser le climat à cette époque pour telle personne. Par exemple :
 - « À l'époque de [La Bolduc](#), le climat ressemblait à ça ».
 - Illustrer le climat pendant l'écriture de la chanson « Mon pays, ce n'est pas un pays, c'est l'hiver » de [Gilles Vigneault](#) : pendant les 10 ans avant l'écriture de cette chanson, il est tombé tant de neige à Natashquan, 20 ans plus tard, il est tombé tant de neige, ce qui a pu inspirer l'auteur.
- Pour des exemples de **lieux**, les données météo pourraient être mises en lien avec les **ouvrages d'art** (ponts, tunnels, digues, quais...). Cela pourrait permettre de voir l'évolution du climat proche de ces ouvrages d'art et contribuer à faire évoluer les standards de construction en tenant compte des changements climatiques.

En faisant des liens entre **les événements climatiques et le quotidien des gens, des lieux précis, on peut ramener un côté individuel et local** aux changements climatiques, tout en restant informatif. Pour cela, les photos peuvent être une bonne ressource :

- Les informations météo pourraient être accrochées aux **photos** téléversées dans Wikimedia Commons, photos dont les métadonnées des lieux et des dates correspondent aux données météo disponibles.
 - On pourrait trouver la photo la plus proche de chaque point d'observation (station météo) pour cet événement ou lieu (intéressant pour des projets comme Wikivoyage). Et inversement, à partir de la photo, on pourrait trouver le point d'observation le plus proche dans le temps et dans l'espace. C'est intéressant pour les gens qui vont aller lire les pages en question, cela contribue à restituer des événements historiques, et cela permet en même temps de créer des bases de données d'entraînement pour que d'autres personnes aillent utiliser ces informations.
 - Par extrapolation, avec de l'IA, on pourrait avoir une suggestion « quand il y a telle photo, il fait tel temps » !

- De la même façon que des graphiques ont été proposés pour suivre l'évolution du climat, cette évolution pourrait être documentée par des **photos anciennes à comparer avec des photos actuelles d'un même endroit** (l'estuaire, la Baie James, par exemple)
 - **Des liens pourraient être établis entre les photos et les graphiques qui permettent de suivre l'évolution du climat** : par exemple, des photos anciennes du bas du fleuve, l'estuaire pris dans les glaces à telle période telle année, pourraient être en lien avec les graphiques correspondant et être comparées à des photos d'aujourd'hui, avec les valeurs des graphiques correspondantes.
 - D'une manière similaire, **des liens pourraient être établis avec la position dans l'espace** : sur la page d'une ville, en sélectionnant la température d'une date précise sur un graphique (exemple: 2 juin 2012), on pourrait voir – avec photos à l'appui – que cette **température** arrivait deux semaines plus tard pendant les dix années précédentes, ou que le **gel** durait plusieurs mois à cet endroit alors qu'il dure maintenant 3 jours en février. Les dates et photos d'apparition des **bourgeons** seraient également parlantes. Cela permettrait de voir l'évolution du climat en repositionnant le lecteur dans l'espace et jouant avec les échelles de temps (10, 20, 100 ans)².
 - Il serait alors intéressant d'**avoir accès à des photos de fonds muséaux** qui rendraient disponibles les images correspondantes aux périodes concernées par le jeu de données météo.

Cela nécessite de **catégoriser les données pour créer des climats-types et relier les photos qui correspondent** à cette période.

- On pourrait créer un climat-type (avec température, précipitations, vent par exemple) et faire des scores avec les métriques. Ces métriques permettraient d'identifier une fourchette de valeurs pour un ou des éléments du temps en particulier. On pourrait **faire des groupes de photos** en reliant celles dont les métadonnées correspondent à cette période-là. Un exemple de ce que cela permettrait : « À Chicoutimi pendant la période du déluge, les précipitations sont de tel type, voilà les photos dans Wikimedia Commons. Par la suite, quand on regarde une photo ou une analyse qui donne une métrique semblable, on pourrait voir les photos qui y sont associées. »
 - Pour faire cela, on pourrait **créer une catégorie dans Wikimedia Commons qui s'appellerait « Précipitations entre tel score et tel score »**.
 - Pour le grand public, cela apparaîtrait :
 - **dans les pages événements de Wikipédia, avec un lien « autres événements similaires »** qui renverrait au groupe de photos correspondantes.

² Par exemple, avec Google Street View, il est possible de remonter à travers le temps. Cette perspective historique est parlante : on peut par exemple voir l'évolution urbaine avec l'apparition d'immeubles dans des zones désertiques il y a 10 ans.

- **dans les pages des villes**, dans les sections sur le climat de la ville, couplé aux graphiques correspondants (cf idée de couplage photo-graphique ci-dessus).

D'un point de vue d'**accessibilité**, pour les [lecteurs d'écran](#) utilisés par les personnes qui ont des difficultés de lecture, il serait intéressant d'**ajouter aux photos des métadonnées qualitatives** en précisant par exemple qu'il s'agit d'une image d'extérieur, prise une journée d'été particulièrement humide et chaude, si cette température était normale ou non.

Dans une perspective **interculturelle**, dans des langues ou des cultures qui ont **différents mots ou différentes notions temporelles pour définir un phénomène climatique**, il serait possible d'**illustrer ces nuances par les données météo et des photos correspondantes**.

- Quand plusieurs mots sont utilisés pour définir un phénomène climatique (la neige par exemple qui se décrit par plusieurs mots au Japon et au Canada), ces mots pourraient être illustrés par les données météo et photos dans le **Wiktionnaire**.
- Quand les saisons ne sont pas déterminées de la même façon selon les cultures, ce serait intéressant de relier les descriptions trouvées dans des textes/photos/vidéos aux données météo correspondantes. Par exemple, [un ensemble de photographies](#) a été récemment téléversé dans Wikimedia Commons, à partir d'un fonds d'archives Atikamekw, il s'agit de photos prises dans les années 1970, réalisées parallèlement à la série *Les Six Saisons Attikamek* des cinéastes Pierre Diné et Pierre Hivon. Il serait intéressant de relier ces photos aux données météo correspondant à cette période et ainsi rendre compte des six saisons.
- À partir de cas similaires, des données pourraient être cherchées auprès d'autres pays et cultures pour montrer d'autres photos sur ces phénomènes et leurs évolutions.

La valorisation descriptive des données météo a ainsi un fort potentiel à la fois informatif mais également de sensibilisation aux manifestations des changements climatiques grâce à des retombées à la fois locales, individuelles, collectives et interculturelles via les comparaisons par photos et graphiques. Les gouvernements pourraient par la suite avoir recours à ces projets-là qui partent de citoyens pour prendre des décisions aux niveaux nationaux et internationaux.

3) Valorisation *prescriptive* des données météo pour aider à la décision dans différents secteurs et d'autres pays

Les deux groupes ont abordé plusieurs idées d'utilisation des données météo *passées* dans une perspective *prédictive* à différents niveaux (individuel, sectoriel, international).

À titre individuel, il serait possible de faire **ses propres prédictions météo à partir des données passées disponibles sur Wikimedia Commons**. Bien qu'il y ait une limite méthodologique aux prévisions à partir de ce jeu de données météo composé d'observations et non de prévisions³, il est possible de faire des **simulations** de prédiction à partir de données passées, sans données de prévisions en tant que telles : un réseau de neurones peut apprendre à partir des données sur une période précise et simuler une prédiction selon ce modèle⁴. Le résultat serait une **infographie réalisée à partir d'apprentissage machine**.

À l'inverse de la prédiction, **on pourrait utiliser ces données de façon rétrospective, pour évaluer les prévisions passées** d'organismes qui réalisent des prévisions météorologiques.

À titre pratique pour différents secteurs, les données météo sont utilisées à des fins d'**aide à la décision** spécifiques à chaque domaine. ECCC répond déjà régulièrement à des demandes d'organisations de secteurs variés (agriculture, criminologie, cabinets d'avocats, aviation, santé, assurance, marketing pour les vitrines de magasins, etc.) qui utilisent les données météo pour faire leurs propres prévisions. Par exemple, certains musées utilisent les données météo pour prévoir l'achalandage et ajuster leur personnel (voire leurs tarifs) en fonction de la météo ; le secteur du transport peut prévoir les conditions de circulation selon les prévisions météo ; les données météo sont utilisées pour de la planification en urbanisme et par les assurances pour prévoir certains impacts climatiques.

- Pour **générer d'autres pistes de corrélations entre les données météo et d'autres secteurs**, des **hackathons** ont déjà eu lieu pour utiliser les données d'ECCC, ouvrant la voie à de nombreuses pistes créatives dans différents domaines. D'autres hackathons de ce type pourraient avoir lieu. *(Cela rejoint l'idée de crowdsourcing pour des visualisations graphiques mentionnées plus haut; ici il s'agit de crowdsourcing d'idées d'applications des données météo).*

³ il s'agit de données d'observation climatiques, pas de prévision. Les modèles de prévisions se composent de données sur davantage de niveaux (dont l'atmosphère) que les données d'observation (qui ne sont qu'en surface).

⁴ Dans un même ordre d'idée, pour sensibiliser aux changements climatiques, un participant du groupe Printemps a mentionné le projet du MILA "[Visualiser le changement climatique](#)" qui n'a pas tant recours à des données météorologiques mais à une très grande quantité de photos d'inondations pour modéliser ce à quoi ressemblerait votre quartier avec la montée des eaux.

- Dans le contexte agricole, plusieurs institutions dont le MAPAQ ou Agriculture Canada font des modélisations à partir de données météo plus complexes que ce qui est disponible dans notre jeu de données (leurs données incluent par exemple les radiations solaires). **Si des modèles simples pouvaient être transférés dans l'écosystème Wikimedia** à partir de ce qui est déjà développé par ces institutions, **les personnes de pays moins développés pourraient utiliser ces modèles** par simples transferts de données de leur côté.

4) Création de pages Méta pour inviter à reproduire le projet

Sur la [page Wikidata du projet](#) et celle sur [MetaWiki](#), un tutoriel pas à pas préciserait toutes les démarches techniques et de gestion faites pour ce projet (comment faire, combien de temps ça prend, combien de personnes et dans quels rôles, quels sont les outils existants, quels formats, quelles activités, etc.).

Ce projet pourrait servir de fer de lance pour initier d'autres projets à partir d'autres organismes, dans d'autres secteurs, et à l'international. Il est donc important de partager le processus par lequel il a été mené, que tout le monde ait accès à la recette.

Plus particulièrement, la mise à disposition de données météo par des institutions météorologiques d'autres pays pourrait permettre de faire des comparaisons à l'échelle internationale et contribuer à rendre visibles les changements climatiques au niveau mondial. Cela repose sur la volonté des gouvernements à revoir certains modèles d'affaires axés sur les données météorologiques et la mouvance vers les gouvernements ouverts est en marche.

Conclusion

Cette séance a permis de révéler un fort potentiel des données météorologiques, à la fois à titre informatif mais aussi prescriptif dans de multiples secteurs. Les parcours différents des participants et participantes réunies autour de l'intérêt commun de la valorisation de grands jeux de données permettent de croiser les réflexions parfois divergentes pour faire grandir le projet.

Face au potentiel de ces données que l'on a tendance à imaginer publiques et qui ne sont pourtant pas si facilement accessibles, le mouvement des données ouvertes gagne à s'étendre aux données météo de tous les pays pour partager à la fois des modèles qui peuvent bénéficier en particulier au pays en développement, mais également permettre des comparaisons internationales pour sensibiliser mondialement aux changements climatiques.



Ces comparaisons auraient notamment un fort impact via des visualisations à la fois par graphiques et par les photos disponibles sur Wikimedia Commons.

Wikimédia Canada propose un premier pas dans cette sensibilisation : une campagne de communication pourrait être lancée à partir d'une exposition photos liées aux données météo d'Environnement et Changement climatique Canada, comme suite à ce projet. Cela pourrait être le point de départ d'un mouvement citoyen qui pourrait inspirer de nouvelles mesures par les gouvernements en matière de changement climatique et de données ouvertes.