Appendix B - Weather Observations of Environment and Climate Change Canada in Wikimedia Commons

Report of the meeting of February 10, 2021

Reporterr: Camille Vézy for Wikimedia Canada
Proofreading: Ha-Loan Phan, Wikimedia Canada; Pierre Choffet for Wikimedia Canada; Miguel Tremblay, ECCC. Translated by: Jean-Philippe Béland, Wikimedia Canada

Wednesday, February 10, 2021, 9:00 a.m. to 11:00 a.m.: Brainstorming

Attendance

- Organizers:
  - Pierre Choffet and Ha-Loan Phan (for Wikimedia Canada)
  - Miguel Tremblay (Environment and Climate Change Canada)
  - Dissemination partners: Acfas (Johanne Lebel, Frédéric Macé) and IVADO (Guillaume Chicoisne)
- Participants:
  - Jean-Hugues Roy, professor at the School of Media, in data journalism, at UQAM
  - Nadia Tahiri, postdoctoral fellow in machine learning, artificial intelligence and public health, at the Université de Montréal
  - Pascal Martinolli, librarian at the Université de Montréal

- Sandrine Edouard, National Open Data Coordinator, Meteorological Service, Environment and Climate Change Canada
- Simon Villeneuve, wikimedian and physics professor at Cégep de Chicoutimi
- Trevor Smith, Climate Scenarios and Services Specialist at Ouranos
- Christian Jauvin, Programmer - Climate, Data and Operations Platforms at Ouranos
- Jean-Christophe Voyer, Strategy and business development, mapping and decision support tools, at Anagraph
- Aurélie Dussenne, data analyst at ImpactBee
- Anonymous person, climatologist of a French governmental meteorological organization
- Benoit Rochon, Wikimedia Canada administrator

Moderated by Sébastien Paquet (Service Now) and Ha-Loan Phan (Wikimedia Canada)

Discussions are held in French.

Flow of the session

The **objective** of the session is to bring out original ideas of data valorization, tools to be created in the Wikimedia Foundation projects, feasible in the short and long terms.

Some of the short-term ideas will be realized by Pierre Choffet by March 31, 2021, the end date of the Weather Observations of Environment and Climate Change Canada in Wikimedia Commons project.

The session consisted of four parts:

1) **Introduction** (10 min): The organizing and facilitating team briefly introduced the session by **introducing the stakeholders**, recalling the **purpose** of the project and of this particular brainstorming session, and announcing the flow of the session.
   Before dividing the group into two subgroups for discussions, the **rules** of brainstorming were announced: focus on the quantity of ideas, no criticism, encourage crazy ideas, combine and improve ideas.

   Pierre Choffet also shared some links for :
   - Recall what data is currently available (monthly and almanac data on Wikimedia Commons). For a summary of the project on Wikidata: https://www.wikidata.org/wiki/Wikidata:WikiProject_Weather_observations
   - Give examples of what could be done, such as automating the table in the "climate" section on the Montreal Wikipedia page, or making visualizations such as histograms, bar graphs, adding interactivity, etc., like what is done on the Covid-19

pandemic in france Wikipedia page. A more complex example, which would be longer term, would be to make interactive diagrams like the Year selector tool for New York City which allows users to select the weather data of the years that interest them.

2 and 3) **Discussions** (2 x 30 min): Participants were divided into 2 subgroups of 6 people (4 men, 2 women in each subgroup) to brainstorm ideas for tools to create in the short and long terms.
After a quick **icebreaker** in sub-groups where each participant shared their favourite season, the "Winter" group, facilitated by Sébastien Paquet, started with **short-term ideas** while the "Spring" group, facilitated by Ha-Loan Phan, brainstormed **long-term ideas**.
After 30 minutes, the 2 groups were brought together in a short **plenary** to share key ideas.
Each group then went back for 30 minutes of **discussions**, reversing the themes: generating long-term ideas for the "Winter" group and short-term ideas for the "Spring" group.

4) **Conclusion** (15 min): the two groups were brought together in plenary one last time and the participants were invited to give feedback on the session.

Summary of ideas generated

Both groups discussed similar ideas: imagining ways to visualize the data; "augmenting" information about events, places, and people with corresponding weather data, especially via photos on Wikimedia Commons; using weather data to help make decisions in different sectors and countries; and at a more meta level, sharing the project process to invite replication.

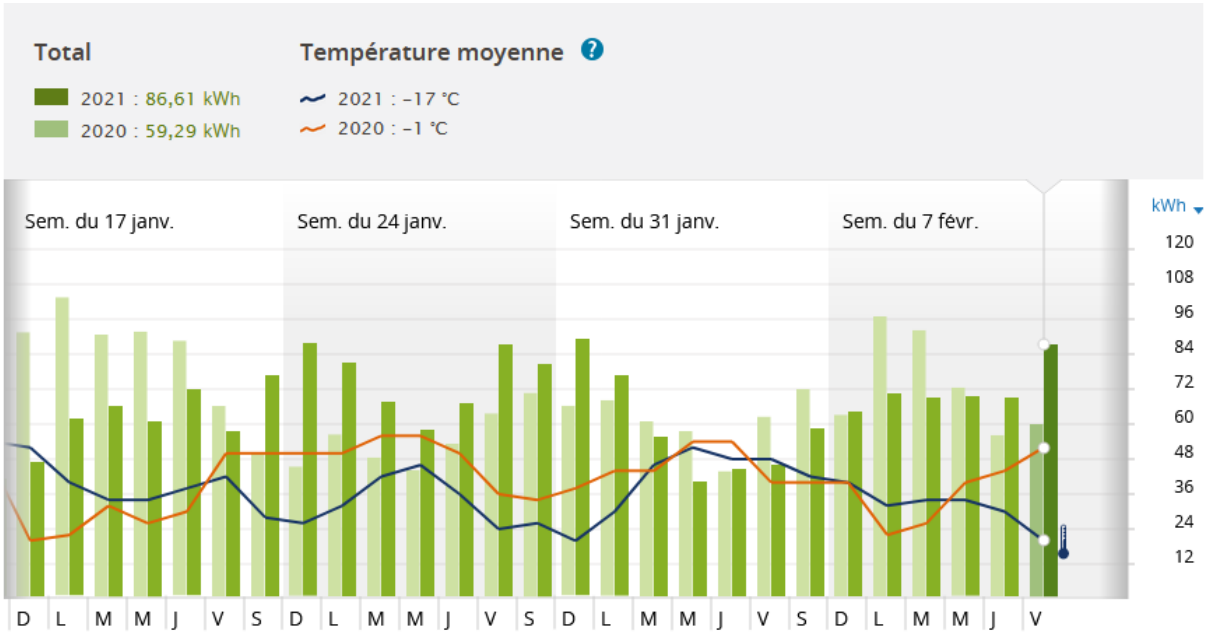
The valorization of weather data was thus considered from the point of view of visualization - both by graphics and by photos on Wikimedia Commons - in a descriptive (improving information) and prescriptive (helping in decision-making) perspective. The strong potential of public awareness on climate change through weather data visualization was explored.

1) Data visualizations

Graphs similar to those made for the climate of New York City could be developed **for all the localities near the stations**. A template could be used to retrieve the temperatures in an automated way, but it remains to be defined which template to integrate.

Other types of graphs could be developed **to track climate change to raise awareness of climate change**:

- **Graphs on a monthly and/or annual values linked to the climate** (temperature, precipitation, etc.) could allow a visualization of these evolutions.
- It would be possible to **calculate the new normals in different regions**. For example, in France, this type of graph exists and shows an increase of 0.4 degrees in some regions.
- On an **interactive graph**, we could allow a user to **choose a start date and an end date to see the evolution of the climate over this period in a given location[1].**
- In the same vein, for the moment on Wikipedia, we can't really **compare** the current year to see where we are in relation to the average, the maximum, the minimum temperature, the median... It would be interesting to have access to a more visual way to compare the current year and month to the same month several years ago, a bit like the comparison tool used by Hydro-Quebec in its energy comparison customer interface.



- Round graphs could also be used to visualize climate change.
- These graphs could be **linked to a map and the demographics corresponding to** that point on the map. This would show how many people are affected in a given area by changes in precipitation, heat, etc.
    - By adding **geolocation**, someone logging on to Wikipedia could see, depending on the area he or she is in, a "personalized" impact of climate change. This would make it possible to see changes at a very local level, rather than thinking that it is far away and that it is only happening to others.

[1] However, methodological details should be specified, as some changes in the raw data may be related to measurement changes (e.g., a change in measurement tool).

- Data visualizations could also be used to see if there are **correlations between weather and public health phenomena**:
    - Using geographical maps where **blue areas** (of longevity) and high mortality are located, correlations could be made with weather data. It would be possible to visualize if there are links between life expectancy, climate and climate change in these areas.
    - **Sentiment analysis graphs from social network publications** could be linked to weather data to see if there are correlations between feelings and weather phenomena. This would allow, for example, to visualize the phenomenon of winter depression, and possibly to plan suicide prevention or **mental health** support campaigns.
    - In **epidemiology,** graphs visualizing the evolution of epidemics could be linked to weather data in the concerned areas to see the correlations between virus circulation and climate according to historical periods. Perhaps correlations could be verified with indicators of climate change.

To develop these graphs, several avenues have been suggested:

- Since **visualization tools** are currently limited **on Wikipedia**, it would be interesting to **improve the existing ones**. This could start with a prioritized **list** of improvements and features to develop a more versatile tool.
- **Crowdsourcing**: **partnerships** with the various existing data analysis communities could help find creative ways to visualize this data with impact. For example, Kaggle organizes mini-competitions to create new kinds of graphs from data sets. Interesting synergies could emerge from partnerships between Wikimedia and these communities.
    - To keep these data science communities in the Wikipedia loop, we could set up a system of automatic tracking of certain keywords (such as "weather", "permafrost", ...) linked to pages about Canada on Wikipedia. These communities would then be regularly (or even automatically?) invited to contribute to creative visualizations of these themes and datasets.

2) *Descriptive* use of weather data to "augment" information about events, places or people

Both groups suggested ideas about adding weather data to already dated and localized information in existing Wikipedia entries and photos. Weather data could be added either in the information boxes of some records (cities, events) or in the metadata of photos in Wikimedia Commons. The links created between event categories, weather information, and related documents would at the same time improve the information about climate

change - which could contribute to the information on the climate change (or even global warming) Wikipedia page.

To do this, several paths are suggested:

- As there are many **pages linked to places and dates**, it would be possible to **target these pages and indicate the weather**.
    - For example, for a specific day in a specific year: https://fr.wikipedia.org/wiki/27_juillet_1945 or a specific day without a year: https://fr.wikipedia.org/wiki/27_juillet. In a playful way, this would make it possible to find the weather on dates such as the day of one's birth, the day of one's grandparents' wedding, etc. It would also help to **reconstruct or put into context historical events**.
    - For an **event**, for example "ice storm", "Montreal Olympics", "lunar eclipse", we could specify the weather information, according to the data of the weather station closest to this event.
        - Therefore, we could target **weather-dependent events** as a priority for adding weather information.
    - We could target Canadian **historical figures** between 1840 and 2018 (weather data tags that are on Wikimedia Commons right now as part of this project) and on each bibliographic record (in infoboxes, for example), specify the climate at that time for that person. For example:
        - "At the time of La Bolduc, the climate was like this.
        - Illustrate the climate during the writing of the song "Mon pays, ce n'est pas un pays, c'est l'hiver" by Gilles Vigneault: during the 10 years before the writing of this song, that much snow fell in Natashquan, 20 years later, that much snow fell, which could have inspired the author.
- For examples of **locations**, weather data could be linked to **engineering structures** (bridges, tunnels, dykes, docks...). This could allow to see the evolution of the climate close to these structures and contribute to the evolution of the construction standards by taking into account the climate changes.

By making connections between **climate events and the daily lives of people and places, we can bring an individual and local aspect** to climate change, while remaining informative. For this, photos can be a good resource:

- Weather information could be attached to **photos** uploaded to Wikimedia Commons, with location and date metadata matching the available weather data.
    - One could find the closest photo to each observation point (weather station) for this event or place (interesting for projects like Wikivoyage). And conversely, from the photo, we could find the closest observation point in time and space. This is interesting for the people who are going to read the pages in question, it helps to restore historical events, and at the same time it allows to create training databases for other people to use this information.

        - By extrapolation, with AI, we could have a suggestion "when there is such and such a picture, the weather is such and such"!
- In the same way that graphs have been proposed to follow the evolution of the climate, this evolution could be documented by **old photos to be compared with current photos of the same place** (the estuary, James Bay, for example)
    - **Links could be established between the photos and the graphs that allow us to follow the evolution of the climate**: for example, old photos of the lower river, the estuary frozen at such and such a period of time and such and such a year, could be linked to the corresponding graphs and compared to photos of today, with the corresponding graph values.
    - In a similar way, **links could be established with the position in space**: on the page of a city, by selecting the temperature of a precise date on a graph (example: June 2, 2012), one could see - with supporting photos - that this **temperature** arrived two weeks later during the ten previous years, or that the **frost** lasted several months at this place whereas it lasts now 3 days in February. The dates and photos of the appearance of the **buds** would also speak for themselves. This would make it possible to see the evolution of the climate by repositioning the reader in space and playing with time scales (10, 20, 100 years)[2].
    - It would then be interesting to **have access to pictures of museum backgrounds** that would make available the images corresponding to the periods concerned by the weather data set.

This requires **categorizing the data to create climate types and linking photos that correspond to** that time period.

- One could create a typical climate (with temperature, precipitation, wind for example) and make scores with the metrics. These metrics would identify a range of values for one or more elements of the weather in particular. We could **make groups of photos** by linking those whose metadata correspond to that period. An example of what this would allow: "In Chicoutimi during the flood period, the precipitation is of such and such a type, here are the photos in Wikimedia Commons. Then, when we look at a photo or an analysis that gives a similar metric, we could see the photos associated with it. »
    - To do this, we could **create a category in Wikimedia Commons called "Precipitation between this score and that score".**
    - To the general public, this would appear:
        - **in the Wikipedia events pages, with a link to "other similar events"** that would lead to the corresponding group of photos.
        - **in the city pages**, in the sections on the city's climate, coupled with the corresponding graphs (see idea of photo-graphic coupling above).

[2] For example, with Google Street View, it is possible to go back through time. This historical perspective is telling: we can for example see the urban evolution with the appearance of buildings in desert areas 10 years ago.

From an **accessibility** point of view, for **screen readers** used by people with reading difficulties, it would be interesting to **add qualitative metadata to the photos,** specifying, for example, that it is an outdoor image, taken on a particularly humid and hot summer day, and whether this temperature was normal or not.

From a **cross-cultural** perspective, in languages or cultures that have **different words or different temporal concepts to define a climatic phenomenon**, it would be possible to **illustrate these nuances with weather data and corresponding photos**.

- When several words are used to define a climatic phenomenon (e.g. snow which is described by several words in Japan and Canada), these words could be illustrated by weather data and photos in the **Wiktionary**.
- When seasons are not determined in the same way in different cultures, it would be interesting to link descriptions found in texts/photos/videos to corresponding weather data. For example, a set of photographs was recently uploaded to Wikimedia Commons, from an Atikamekw archive, taken in the 1970s, in parallel with the series *The Six Seasons of Attikamek* by filmmakers Pierre Dinel and Pierre Hivon. It would be interesting to link these photos to the weather data corresponding to this period and thus give an account of the six seasons.
- From similar cases, data could be sought from other countries and cultures to show other pictures of these phenomena and their evolution.

The descriptive valorization of weather data has a strong potential to inform but also to raise awareness of climate change manifestations thanks to local, individual, collective and intercultural impacts through photo and graphic comparisons. Governments could then use these projects, which start from citizens to make decisions at national and international levels.

3) *Prescriptive* valorization of weather data to support decision making in different sectors and other countries

Both groups discussed several ideas for using *past* weather data in a *predictive perspective* at different levels (individual, sectoral, international).

As an individual, it would be possible to make **your own weather predictions from past data available on Wikimedia Commons**. Although there is a methodological limit to forecasting from this weather dataset composed of observations and not forecasts[3], it is possible to make prediction **simulations** from past data, without forecast data as such: a neural network can learn from the data over a specific period and simulate a prediction according to this model[4]. The result would be an **infographic made from machine learning**.

In contrast to prediction, **this data could be used retrospectively, to evaluate past forecasts** of organizations that make weather forecasts.

As a practical matter for different sectors, weather data is used for domain-specific **decision support** purposes. ECCC already regularly responds to requests from organizations in various sectors (agriculture, criminology, law firms, aviation, health, insurance, storefront marketing, etc.) that use weather data to make their own forecasts. For example, some museums use weather data to forecast attendance and adjust their staff (and even their prices) according to the weather; the transportation sector can forecast traffic conditions according to weather forecasts; weather data is used for urban planning and by insurance companies to forecast certain climatic impacts.

- To **generate other avenues for correlations between weather data and other sectors**, **hackathons** have already been held to use ECCC data, opening up many creative avenues in different areas. *(This is similar to the idea of crowdsourcing for graphic visualizations mentioned above; this is crowdsourcing ideas for applications of weather data*.)
- In the agricultural context, several institutions such as MAPAQ or Agriculture Canada are modeling with more complex weather data than what is available in our dataset (their data includes for example solar radiation). **If simple models could be transferred into the Wikimedia ecosystem** from what is already developed by these institutions, **people from less developed countries could use these models** by simple data transfers on their side.

4) Creation of Meta pages to invite to reproduce the project

On the Wikidata page of the project as well as the one on MetaWiki, a step-by-step tutorial would specify all the technical and management steps done for this project (how to do it, how much time it takes, how many people and in which roles, what tools exist, what formats, what activities, etc.).

---

[3] it is about climate observation data, not forecasting. The forecast models are composed of data on more levels (including the atmosphere) than the observational data (which are only at the surface).

[4] In a similar vein, to raise awareness about climate change, a Spring participant mentioned MILA's "Visualizing Climate Change" project that uses not so much weather data but a huge amount of flood photos to model what your neighborhood would look like with rising water.

This project could serve as a spearhead to initiate other projects from other organizations, in other sectors, and internationally. It is therefore important to share the process by which it was conducted, so that everyone has access to the recipe.

In particular, the availability of weather data from meteorological institutions in other countries could enable international comparisons and help make climate change visible at the global level. This relies on the willingness of governments to rethink some of their business models around weather data and the move towards open government is underway.

Conclusion

This session revealed the strong potential of meteorological data, both for information and for prescriptive purposes in multiple sectors. The different backgrounds of the participants gathered around the common interest of the valorization of large datasets allow to cross the sometimes divergent thoughts to make the project grow.

Faced with the potential of this data that we tend to imagine as public and yet are not so easily accessible, the open data movement is gaining to extend to the weather data of all countries to share both models that can benefit in particular to developing countries, but also allow international comparisons to raise awareness globally on climate change. These comparisons would have a strong impact via visualizations both by graphs and by photos available on Wikimedia Commons.

Wikimedia Canada proposes a first step in this awareness: a communication campaign could be launched from a photo exhibition linked to Environment and Climate Change Canada's weather data, as a follow-up to this project. This could be the starting point of a citizen's movement that could inspire further action by governments on climate change and open data.