

# La dalle météo



Team TLS - Hackathon Météo France - 08 et 09 avril 2024

# Problème - Agent du service public

## “Prévisions T° de AROME sur Toulouse”



- Champs de surface

Paquets
<b>SP1</b> : P(mer), U(10m), V(10m), DD(10m), FF(10m), FF_RAF(10m), U_RAF(10m), V_RAF(10m), T(2m), HU(2m), NEBUL, PRECIP, NEIGE, FLSOLAIRE_D, GRAUPEL
<b>SP2</b> : ALTITUDE, P(sol), T(sol), NEBBAS, NEBHAU, NEBMOY, NEBCON, CAPE_INS, H_COULIM, EAU, TMIN(2m), TMAX(2m), TD(2m), Q(2m)
<b>SP3</b> : COLONNE_VAPO, FLEVAP, FLLAT, FLSEN, FLTHERM_D, FLSOLAIRE, FLTHERM, FLRASOL_CC, RATHIE_CC, USTR, VSTR

### Champs isobares

Paquets
<b>IP1</b> : T, HU, U, V, Z sur 24 niveaux (100 à 1000 hPa)
<b>IP2</b> : CLD_WATER, CLD_RAIN, CLD_SNOW, CIWC, CLD_FRACT sur 24 niveaux (100 à 1000 hPa)
<b>IP3</b> : TD, O, DD, FF, VV, VV2, TP sur 24 niveaux (100 à

 **arome\_\_001\_\_HP1\_\_01H\_\_2024-04-08T12:00:00Z.grib2**

Mis à jour aujourd'hui — grib2 (68.2Mo) — 66 téléchargements

 **arome\_\_001\_\_HP1\_\_02H\_\_2024-04-08T12:00:00Z.grib2**

Mis à jour aujourd'hui — grib2 (67.9Mo) — 14 téléchargements

 **arome\_\_001\_\_HP1\_\_03H\_\_2024-04-08T12:00:00Z.grib2**

Mis à jour aujourd'hui — grib2 (68.4Mo) — 15 téléchargements

**arome\_\_001\_\_HP1\_\_04H\_\_2024-04-08T12:00:00Z.grib2**

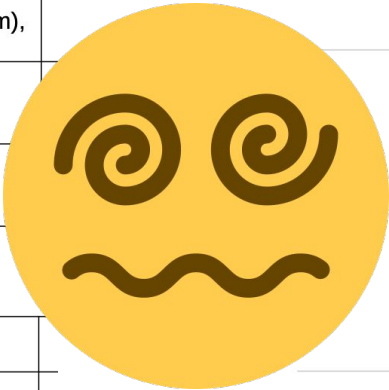
Mis à jour aujourd'hui — grib2 (68.8Mo) — 8 téléchargements

**arome\_\_001\_\_HP1\_\_05H\_\_2024-04-08T12:00:00Z.grib2**

Mis à jour aujourd'hui — grib2 (68.7Mo) — 10 téléchargements

 **arome\_\_001\_\_HP1\_\_06H\_\_2024-04-08T12:00:00Z.grib2**

Mis à jour aujourd'hui — grib2 (69.0Mo) — 11 téléchargements



# 3 innovations

- Format de données “.parquet”

- compresser les données
- indexer les données
- limiter le coût de stockage



- Duckdb

- requêter des fichiers “.parquet” **distants**
- en langage SQL
- depuis une page web (pas de serveur back-end)



- Le petit plus ( optionnel ) : LLM

- transformer la demande naturelle utilisateur en requête SQL





# Problème résolu : “Prévisions de AROME sur Toulouse”

## 1 FICHIER PRINCIPAL

 **arome-001-sp1-00h-2024-04-09t06-00-00z-heightaboveground-2.parquet**

Mis à jour aujourd'hui — parquet (16.5Mo) — 0 téléchargements

## 1 FICHIER PRINCIPAL

 **q-75-previous-1950-2022-rr-t-vent.parquet**

Mis à jour aujourd'hui — parquet (2.5Mo) — 0 téléchargements

### Métadonnées

#### URL

<https://static.data.gouv.fr/resour...>  Créée le 9 avril 2024

#### URL stable

<https://www.dat...>



Donne moi la requête pour les données de Toulouse, avec un pas quotidien, pour 2021, sur la température maximale







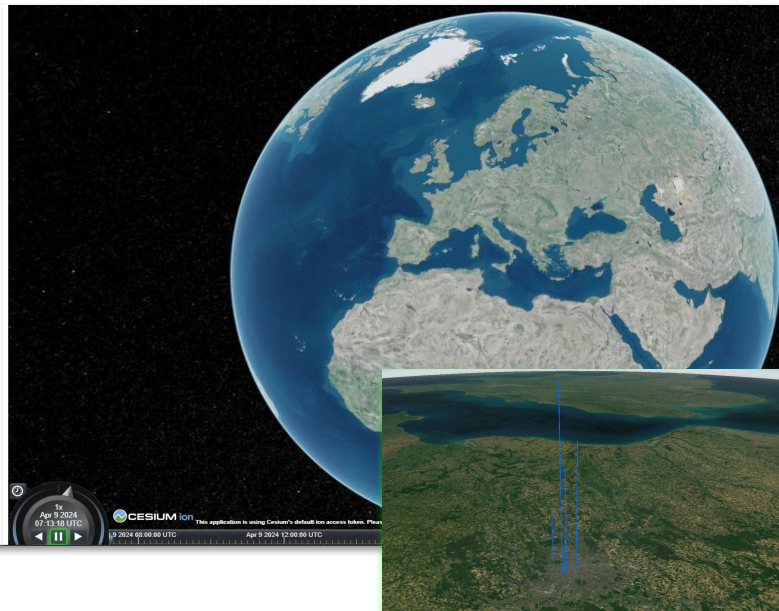
ChatGPT

#### Requête SQL

```
sql
SELECT
  TX/10 as TX_CELSIUS
FROM
  read_parquet('https://object.files.data.gouv.fr/meteofrance/data/synchro_ftp/BASE
WHERE
  date >= '2021-01-01'
  AND date <= '2021-12-31'
```



 Dashboard  
 Data Selection  
  
Export Data  
 To CSV format  
 To JSON format



# Équipe et Perspectives

IDEATION  
-  
09/04/2024



Sébastien Bourgeois  
Ingénieur Data



Maxime Pawlak  
Ingénieur Logiciel



Thomas Sigward  
Ingénieur Spatial



CONCRETISATION

-  
Formalisation d'un  
partenariat via  
Fondation d'entreprise



?

# Annexes

# Cas d'utilisation:



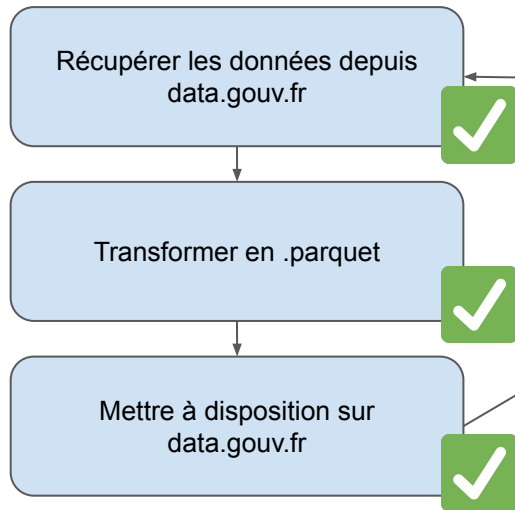
Le fichier JSON exporté des données vents permettra aux **ballons manoeuvrant** de ThalesAlenia Space (StratoBus) et Héméria (Balman) d'optimiser leur trajectoire en vol



Les **assurances paramétriques** nécessitent les données météo pour évaluer les risques ( inondations, perte culture agricole ) et concevoir des produits d'assurance novateurs en fonction d'un indice climatique (température, pluviométrie, taux d'humidité)

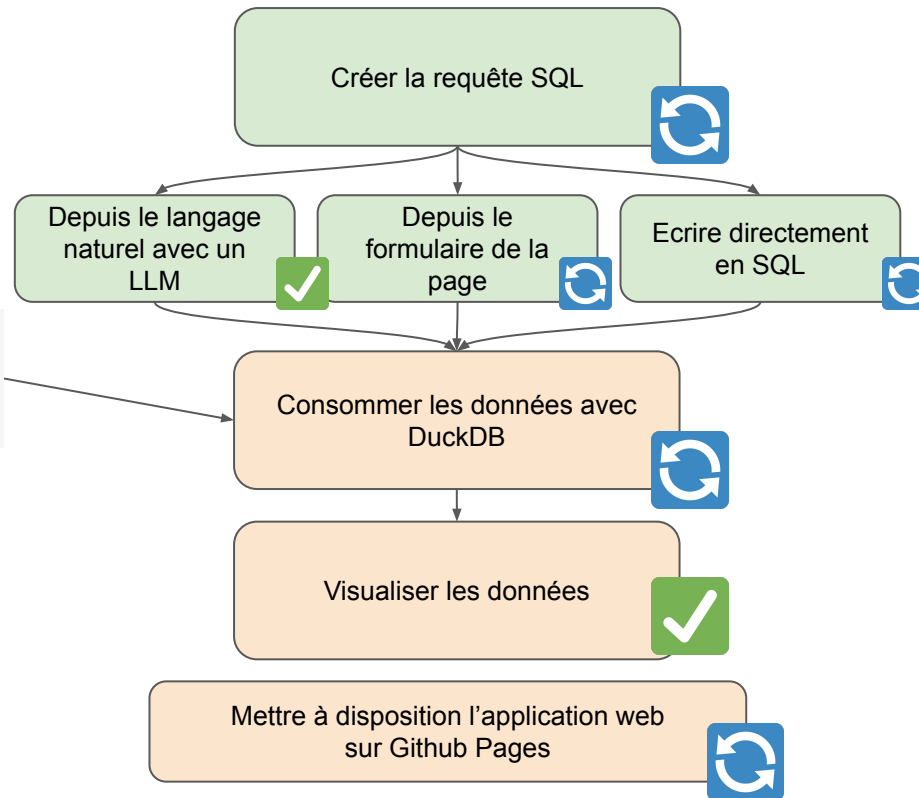
# Faisabilité technique

Transformer et rendre disponible les données ouvertes



data.gouv.fr

Faciliter l'accès aux données avec l'application web





# Solution

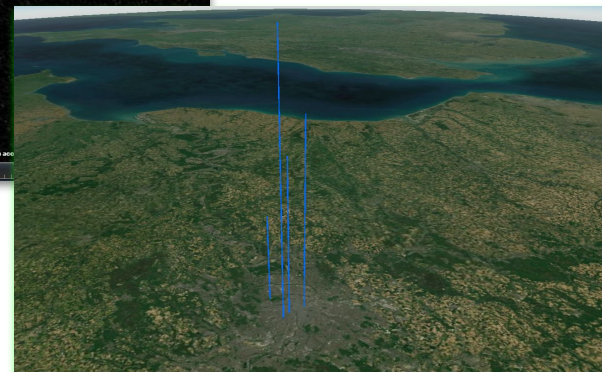
Une interface épurée et ergonomique pour récupérer et visualiser les données météo



L'innovation réside dans la facilitation de la donnée aussi bien à grande échelle que sur une zone précise.

Les éléments de présentation de données peuvent être toutes formes géométriques plaçable à toute altitude, afin de représenter:

- les datas météo
- les risques naturels à anticiper ( feux de forêt, inondations, îlot de chaleur ) .



Exemple d'affichage de pluviométrie sur le bassin parisien le 24/03/2024 (récupération données station météo)





# Données .parquet uploadées sur data.gouv.fr

- <https://www.data.gouv.fr/fr/datasets/donnees-climatologiques-quotidiennes-de-base-de-toutes-les-stations-du-departement-75-pour-2023-et-2024/>
- <https://www.data.gouv.fr/fr/datasets/paquets-sp1-arome-resolution-0-01deg-prevision-du-9-avril-2024-a-6h00/>

1 FICHIER PRINCIPAL

 **q-75-previous-1950-2022-rr-t-vent.parquet**  
Mis à jour aujourd'hui — parquet (2.5Mo) — 0 téléchargements



Métadonnées

URL	Créée le	Taille
<a href="https://static.data.gouv.fr/resour...">https://static.data.gouv.fr/resour...</a> 	9 avril 2024	2.5Mo
URL stable	Modifiée le	
<a href="https://www.data.gouv.fr/fr/data...">https://www.data.gouv.fr/fr/data...</a> 	9 avril 2024	
ID	sha1	
6dd7d069-af37-4b88-85d5-df5f... 	a79b3a1e15b1be5054a7b3ddd...	
Type MIME	application/octet-stream	

1 FICHIER PRINCIPAL

 **arome-001-sp1-00h-2024-04-09t06-00-00z-heightaboveground-2.parquet**  
Mis à jour aujourd'hui — parquet (16.5Mo) — 0 téléchargements

Métadonnées

URL	Créée le	Taille
<a href="https://static.data.gouv.fr/resour...">https://static.data.gouv.fr/resour...</a> 	9 avril 2024	16.5Mo
URL stable	Modifiée le	
<a href="https://www.data.gouv.fr/fr/data...">https://www.data.gouv.fr/fr/data...</a> 	9 avril 2024	
ID	sha1	
abc8b9bc-a590-4cae-ada0-f2ba... 	fffe4f6862307f691ff02c88b2e26... 	
Type MIME	application/octet-stream	

# .csv.gz -> .parquet

## 1 FICHIER PRINCIPAL

### q-75-previous-1950-2022-rr-t-vent.parquet

Mis à jour aujourd'hui — parquet (2.5Mo) — 0 téléchargements

#### Métadonnées

URL	Créée le	Taille
<a href="https://static.data.gouv.fr/resour...">https://static.data.gouv.fr/resour...</a>	9 avril 2024	2.5Mo
URL stable	Modifiée le	
<a href="https://www.data.gouv.fr/fr/data...">https://www.data.gouv.fr/fr/data...</a>	9 avril 2024	
ID	sha1	
6dd7d069-af37-4b88-85d5-df5f...	a79b3a1e15b1be5054a7b3ddddd...	
Type MIME		
application/octet-stream		

```
JS duckdb_meteo.js U X  README.md
paramo-misc > nodejs > JS duckdb_meteo.js > [ⓧ] request
1  const duckdb = require("duckdb");
2  const db = new duckdb.Database(":memory:"); // or a file name for a persistent DB
3
4
5  const request = () => {
6    const parquet_files = [
7      "https://static.data.gouv.fr/resources/donnees-climatologiques-quotidiennes-\
8      de-base-de-toutes-les-stations-du-departement-75-pour-2023-et-2024/20240409-081431/\
9      q-75-latest-2023-2024-rr-t-vent.parquet"
10   ].join(",");
11
12   const request =
13     `SELECT * FROM read_parquet(['${parquet_files}']) LIMIT 1000`;
14
15   console.log(request);
16
17   return new Promise((resolve, reject) => {
18     db.all(request, function (err, res) {
19       console.log(res);
20       if (err) {
```

PROBLÈMES SORTIE CONSOLE DE DÉBOGAGE TERMINAL PORTS

```
},
{
  NUM_POSTE: 75106001n,
  NOM_USUEL: 'LUXEMBOURG',
  LAT: 48.844667,
  LON: 2.333833,
  ALTI: 46n,
  AAAAMJJJ: 20230410n,
  RR: 5.4,
  QRR: 1,
  TN: 10.5,
  QTN: 1,
  HTN: 439,
  QHTN: 9,
  TX: 17.4,
  QTX: 1,
  HTX: 1357,
  QHTX: 9,
  TM: 13.1,
  QTM: 1,
  TNTXM: 14,
  QTNXM: 1,
  TAMPLI: 6.9,
  QTAMPLI: 1,
  TNSOL: null,
  QTNOL: null,
  TNS0: null,
```

# .grib2 -> .parquet

## 1 FICHIER PRINCIPAL

 **arome-001-sp1-00h-2024-04-09t06-00-00z-heightaboveground-2.parquet**  
Mis à jour aujourd'hui — parquet (16.5Mo) — 0 téléchargements

### Métadonnées

URL	Créée le	Taille
<a href="https://static.data.gouv.fr/resour...">https://static.data.gouv.fr/resour...</a> 	9 avril 2024	16.5Mo
URL stable	Modifiée le	
<a href="https://www.data.gouv.fr/fr/data...">https://www.data.gouv.fr/fr/data...</a> 	9 avril 2024	
ID	sha1	
abc8b9bc-a590-4cae-ada0-f2ba... 	fffe4f6862307f691ff02c88b2e26... 	
Type MIME		
application/octet-stream		

JS duckdb\_meteo.js U X README.md

paramo-misc > n Fermer (%W) kdb\_meteo.js > request

```
3 // q-75-latest-2023-2024-rr-t-vent.parquet"
4 const duckdb = require("duckdb");
5 const db = new duckdb.Database(":memory:"); // or a file name for a persistent DB
6
7
8 const request = () => {
9   const parquet_files = ["https://static.data.gouv.fr/resources/paquets-sp1-arome-resolution-0-01deg-prevision\
10 -du-9-avril-2024-a-6h00/20240409-105147/arome-001-sp1-00h-2024-04-09t06-00-00z-heightaboveground-2.parquet"
11 ].join(",");
12
13   const request =
14     `SELECT * FROM read_parquet(['${parquet_files}']) LIMIT 1000`;
15
16   console.log(request);
17
18   return new Promise((resolve, reject) => {
19     db.all(request, function (err, res) {
20       console.log(res);
21       if (err) {
22         console.warn(err);
```

PROBLÈMES

SORTIE

CONSOLE DE DÉBOGAGE

TERMINAL

PORTS

```
latitude: 55.4,
longitude: -11.03000000000002
},
{
  time: 2024-04-09T06:00:00.000Z,
  step: 0n,
  heightAboveGround: 2,
  valid_time: 2024-04-09T06:00:00.000Z,
  t2m: null,
  r2: null,
  latitude: 55.4,
  longitude: -11.020000000000021
},
{
  time: 2024-04-09T06:00:00.000Z,
  step: 0n,
  heightAboveGround: 2,
  valid_time: 2024-04-09T06:00:00.000Z,
  t2m: null,
  r2: null,
  latitude: 55.4,
  longitude: -11.010000000000021
},
... 900 more items
]
```

# Lien projet

- <https://www.data.gouv.fr/fr/reuses/projet-la-dalle-meteo/>

data.gouv.fr


Recherche

Données Réutilisations Organisations Commencer sur data.gouv.fr Actualités Nous contacter

Accueil > Réutilisations > Projet La Dalle Météo

Ajouter aux favoris Voir la réutilisation Modifier

## Projet La Dalle Météo



Maxime Pawlak

<b>Thématique</b>	<b>Type</b>
Outils open data	Application

**Mots-clés**

duckdb hackathon-meteo ilm parquet

**Date de création**  
9 avril 2024

**Dernière mise à jour**  
9 avril 2024

# Repo Github

- <https://github.com/MaximePawlakFr/hackathon-meteo-team-tls>

## Présentation

Equipe (en distanciel) ACCÈS ET COMPREHENSION DES DONNEES

## Membres

- Sébastien Bourgeois (Data engineer)
- Maxime Pawlak (Développeur)
- Thomas Sigward (Project management / Développeur front-end)

## Rendus

Les rendus sont disponibles dans le dossier [rendus](#)

- [Intermédiaire - Jour 1](#)
- [Rendu](#)

## Applications

Le projet est composée de plusieurs applications :

- [Application data](#) qui gère la transformation des données
- [Application front](#) qui est l'interface web
- [Module LLM](#) qui est les éléments pour faire fonctionner le LLM dans l'application front

# Problème

Assurance paramétrique

je veux récupérer les données de pluie d'une station dans le Gers depuis 1900, à la journée

- QUOT\_departement\_32\_periode\_1872-1949\_RR-T-Vent.csv.gz
- QUOT\_departement\_32\_periode\_1950-2022\_RR-T-Vent.csv.gz
- QUOT\_departement\_32\_periode\_1950-2022\_RR-T-Vent.csv.gz

Je veux plus précis :

- HOR\_departement\_32\_periode\_1872-1949\_RR-T-Vent.csv.gz
- HOR\_departement\_32\_periode\_1950-2022\_RR-T-Vent.csv.gz
- HOR\_departement\_32\_periode\_1950-2022\_RR-T-Vent.csv.gz

L'API de météo France limite à un an de donnée, ce n'est pas plus simple.

# Problème - Agent du service public

"Récupérer les données de température minimale des stations de Boulogne-Billancourt et de Paris entre le 20 décembre 2022 et le 10 janvier 2023".

- QUOT\_departement\_92\_periode\_1950-2022\_RR-T-Vent.csv.gz
- QUOT\_departement\_92\_periode\_2023-2024\_RR-T-Vent.csv.gz
- QUOT\_departement\_75\_periode\_1950-2022\_RR-T-Vent.csv.gz
- QUOT\_departement\_75\_periode\_2023-2024\_RR-T-Vent.csv.gz



Données téléchargées inutiles :

- 🌧️ la pluie et 🌬️ le vent
- 🗺️ les stations hors de la zone
- 📅 hors de la période (99,9995%)

