

Café data



netCDF

19/03/2026 - LIENSs

NetCDF, c'est quoi ?

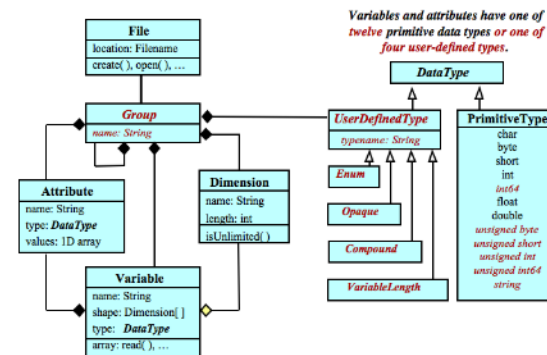


NetCDF (Network Common Data Form) créé en 1988 par [Unidata](https://unidata.ucar.edu)

un **format de fichier** pour les données multidimensionnelles (.nc)

+

un **modèle de représentation de données**

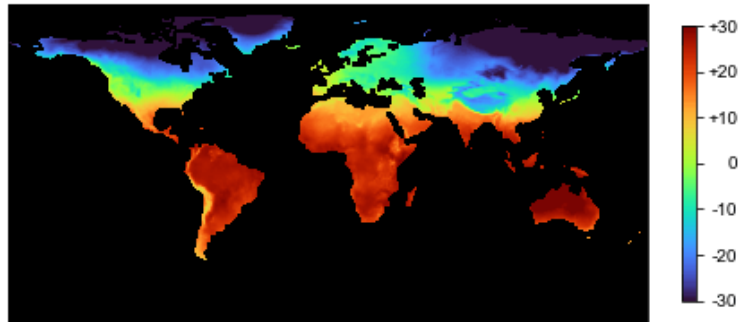


A file has a top-level unnamed group. Each group may contain one or more named subgroups, user-defined types, variables, dimensions, and attributes. Variables also have attributes. Variables may share dimensions, indicating a common grid. One or more dimensions may be of unlimited length.

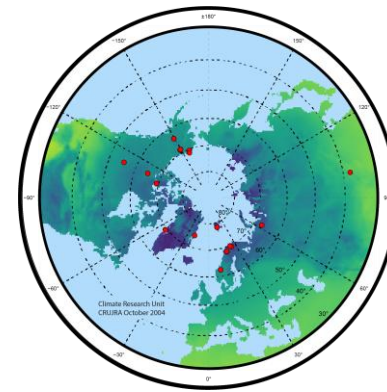
NetCDF, c'est quoi ?

- Développé pour la communauté des sciences de la Terre pour gérer des informations géospatiales représentant des phénomènes variant dans l'espace/le temps : des **données multidimensionnelles**.
- Très utilisé dans le domaine scientifique : climatologie, météorologie, océanographie, médecine, etc...

Average monthly temperature [°C] from 1901 to 2020
01-1901



Data source: CRU TS monthly high-resolution gridded multivariate climate dataset
Visualization: Johannes H. Uhl, University of Colorado Boulder (USA), 2022.



<https://warmingpermafrost.nau.edu>

- Vous connaissez d'autres formats de ce type?

Pourquoi ?



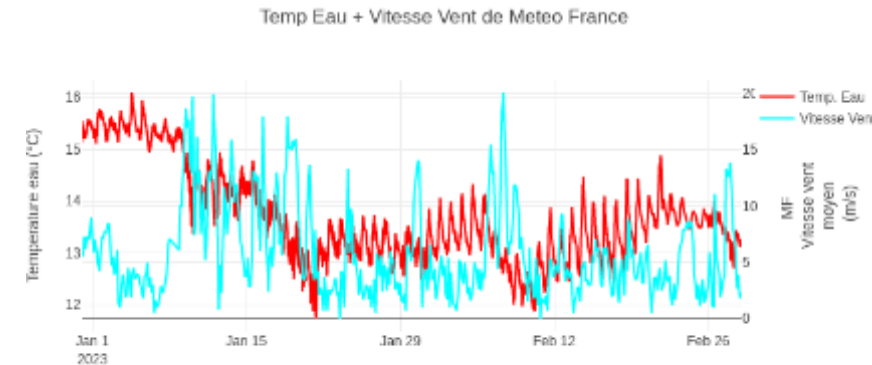
Format de fichier
interopérable
pour la
science ouverte

Pour répondre à plusieurs besoins :

- Stocker des données multidimensionnelles efficacement
 - (Spatio)temporelles*
 - Mais pas que !

- Suivre les principes FAIR → Accessible, Interopérable, Réutilisable

- ✓ Format ouvert
- ✓ Lisible par de nombreux outils (logiciels, programmation / code)
- ✓ Pérenne / maintenu (version 5.8 en mai 2025)
- ✓ Reconnu par des conventions internationales (ex : [Convention Climate and Forecast](#))



https://sist.pages.in2p3.fr/webinaire_netcdf_2024/introduction.html

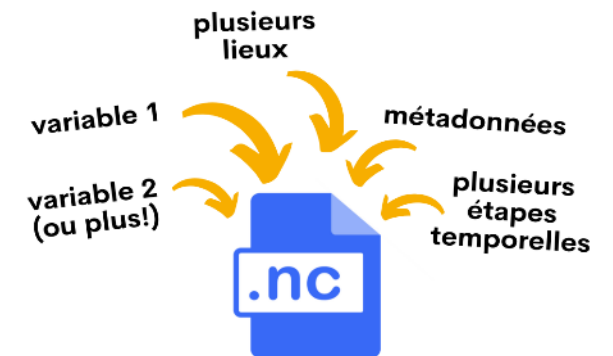
*Le netCDF accepte bien les données 1D (expl : séries temporelles)

Validé par la communauté scientifique

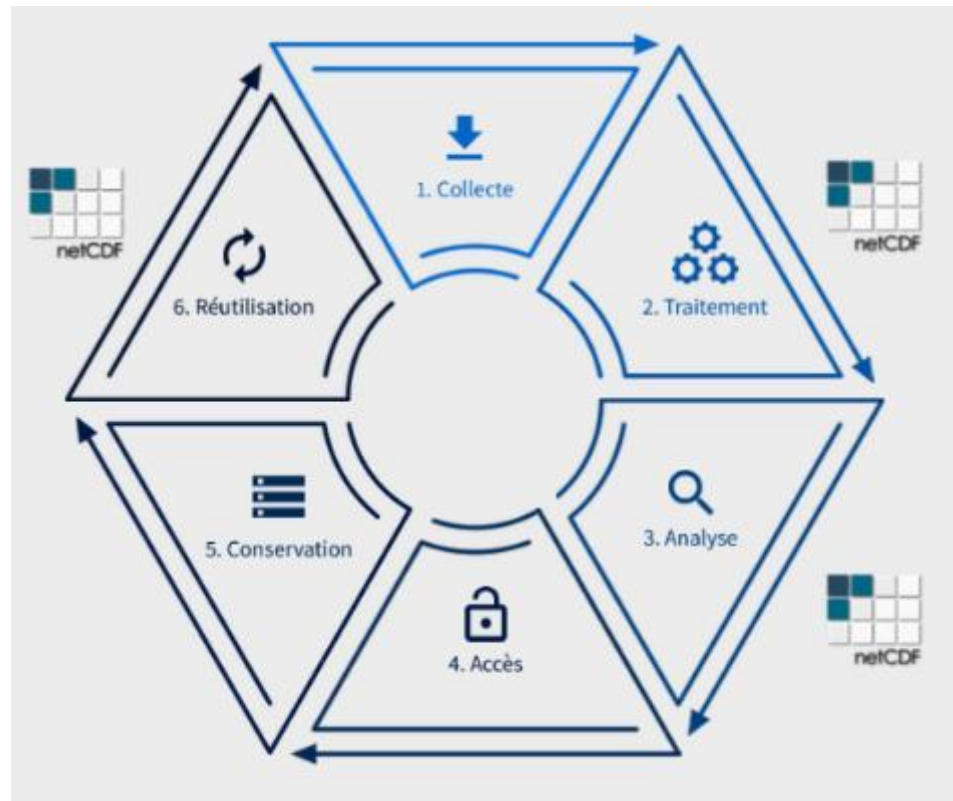


Un seul fichier **auto-descriptif** pour :

- Les données multi-dimensionnelles
- Les métadonnées (signification des variables et unités, méthode, matériel...)
- Optimiser le stockage :
 - Gros volumes
 - Peu de duplication de données
 - Efficace (binaire + compression)
 - Plus écologique / responsable



Le NetCDF dans le cycle de vie des données



<https://doranum.fr/>

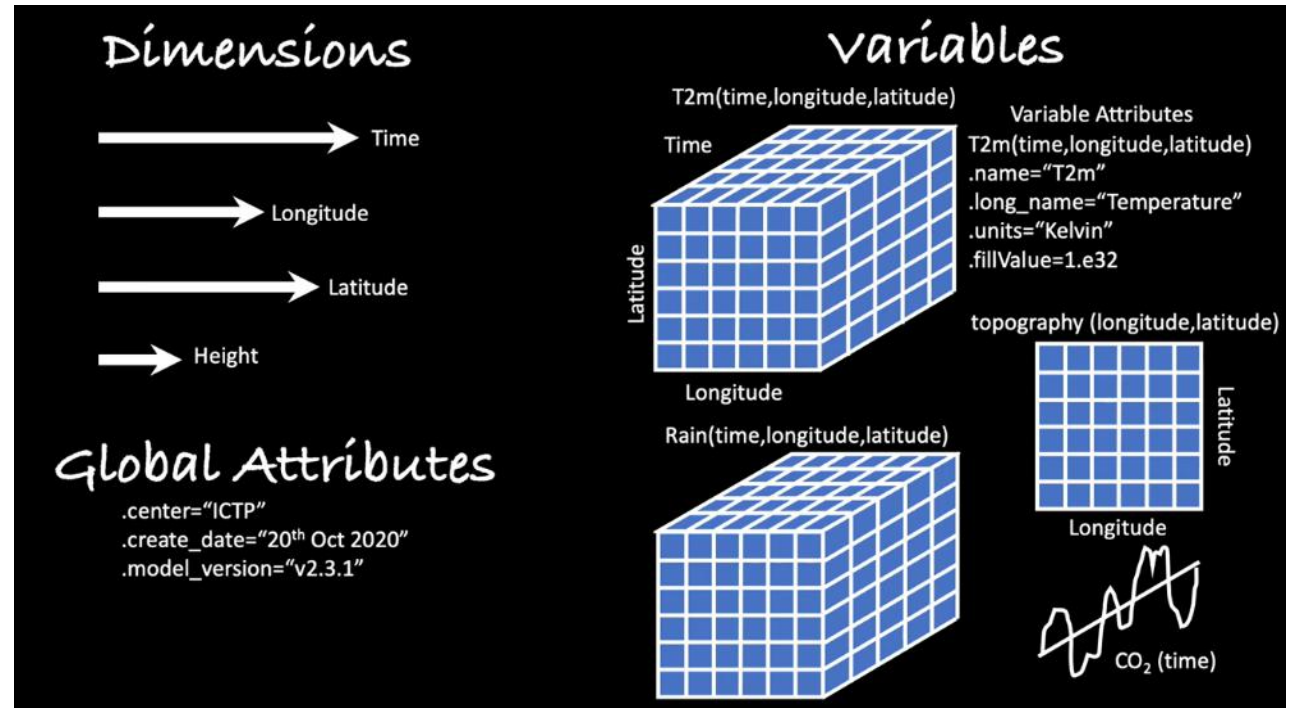
recommandé par le [pôle de données Odatis](#) et l'[IR DataTerra](#)

Le concept

Les données sont stockées dans des **variables** sous forme de tableaux à plusieurs **dimensions**.

A savoir :

- Chaque dimension est d'une taille fixe.
- Les variables peuvent avoir 1 à n dimensions



<https://www.youtube.com/watch?v=UvNBnjiTXa0>

Comprendre un fichier

Quatre termes à connaître :

- **Dimension** : nombre de données associées à un axe.
- **Variable** : Ensemble d'attributs + tableau de valeurs associé à plusieurs dimensions
- **Coordonnée** : Variable particulière associant des valeurs à une dimension.

Quatre sections

dimensions

variables

attributs globaux

données

```
netcdf cdl_example {
  dimensions:
    t = 8 ;
    x = 6 ;
    y = 6 ;
  variables:
    double time(t) ;
      time:long_name = "Chronological Julian UTC Datetime" ;
      time:standard_name = "time" ;
      time:units = "days since -4713-01-01T00:00:00Z" ;
      time:axis = "T" ;
      time:calendar = "julian" ;
    double latitude(y) ;
      latitude:long_name = "Latitude" ;
      latitude:standard_name = "latitude" ;
      latitude:units = "degree_north" ;
      latitude:axis = "Y" ;
    double longitude(x) ;
      longitude:long_name = "Longitude" ;
      longitude:standard_name = "longitude" ;
      longitude:units = "degree_east" ;
      longitude:axis = "X" ;
    short precipitation(t,x,y) ;
      precipitation:long_name = "Precipitation" ;
      precipitation:standard_name = "precipitation" ;
      precipitation:units = "mm" ;
      precipitation:coordinates = "longitude latitude" ;
  // global attributes:
    :Conventions = "CF-1.13" ;
    :featureType = "profile" ;
    :title = "Example of dataset." ;
  data:
    ...
}
```

Les conventions

NetCDF très flexible => mauvaises pratiques possibles

Conventions => limiter les mauvaises pratiques



[CF conventions](#)

Une convention ==

- + Structure de fichier (dimensions et coordonnées)
- + Attributs (métadonnées) obligatoires et optionnels
- + Vocabulaires contrôlés (nommage des variables)



[Webinaire SIST - NetCDF : les "conventions"](#)

Mais...

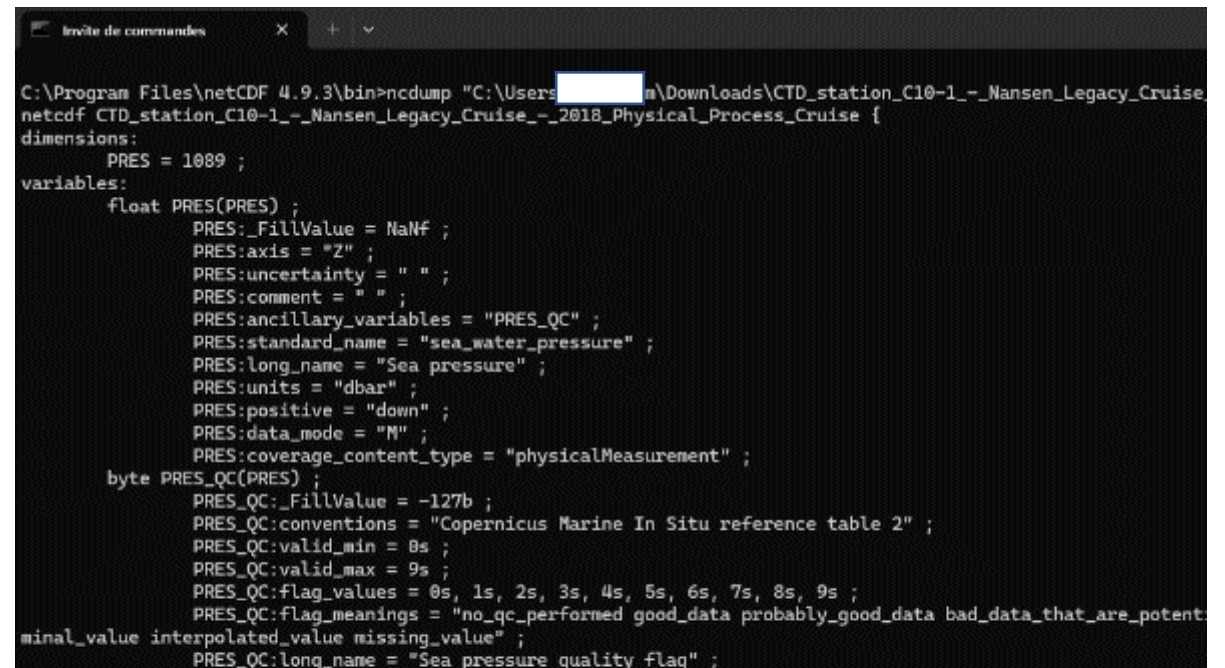
Ce **format binaire** n'est pas toujours simple à prendre en main pour visualiser, manipuler ou créer

- Commandes de base fournies par Unidata (en shell)

<https://downloads.unidata.ucar.edu/netcdf/>

Visualiser le contenu du fichier avec :

- `ncdump` : `ncdump -h` (en-tête), `ncdump -v` (variables), `ncdump -c` (coordonnées)



```
invite de commandes
C:\Program Files\netCDF 4.9.3\bin>ncdump "C:\Users\...m\Downloads\CTD_station_C10-1_-_Nansen_Legacy_Cruise.ncdf"
netcdf CTD_station_C10-1_-_Nansen_Legacy_Cruise_-_2018_Physical_Process_Cruise {
dimensions:
    PRES = 1089 ;
variables:
    float PRES(PRES) ;
        PRES:_FillValue = NaNf ;
        PRES:axis = "Z" ;
        PRES:uncertainty = " " ;
        PRES:comment = " " ;
        PRES:ancillary_variables = "PRES_QC" ;
        PRES:standard_name = "sea_water_pressure" ;
        PRES:long_name = "Sea pressure" ;
        PRES:units = "dbar" ;
        PRES:positive = "down" ;
        PRES:data_mode = "M" ;
        PRES:coverage_content_type = "physicalMeasurement" ;
    byte PRES_QC(PRES) ;
        PRES_QC:_FillValue = -127b ;
        PRES_QC:conventions = "Copernicus Marine In Situ reference table 2" ;
        PRES_QC:valid_min = 0s ;
        PRES_QC:valid_max = 9s ;
        PRES_QC:flag_values = 0s, 1s, 2s, 3s, 4s, 5s, 6s, 7s, 8s, 9s ;
        PRES_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potential_value interpolated_value missing_value" ;
        PRES_QC:long_name = "Sea pressure quality flag" ;
```

Logiciels graphiques

Outils de visualisation :

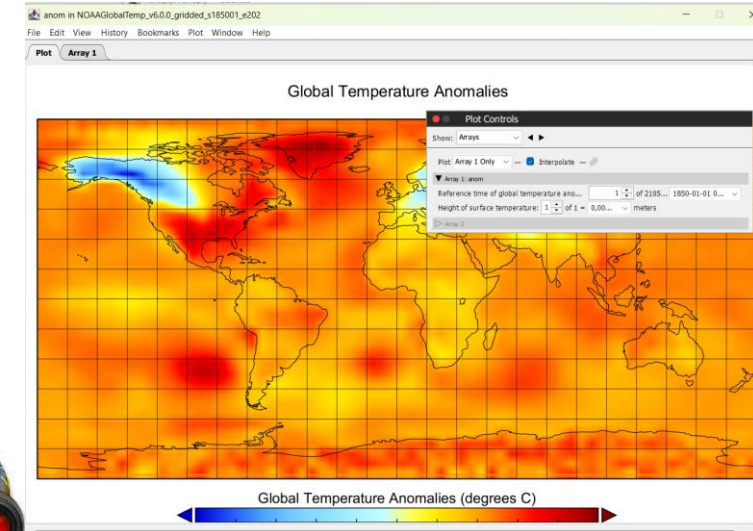
Panoply <https://www.giss.nasa.gov/tools/panoply/>

NCVlew https://cirrus.ucsd.edu/ncview/cdo_showdate_example.nc

Ferret <https://ferret.pmel.noaa.gov/Ferret/> (un peu vintage!)

netcdf-three : <https://github.com/umrlastig/netcdf-three>

Et aussi QGIS, ArcGIS...



Les langages informatiques

Création et édition : Plusieurs moyens :

- Le texte brut : langage CDL + ncgen
- Le programmation : Python, R, Matlab, Fortran, C++, C, etc...
- Outils sur mesure : GENIFAIR, nc_conv

* Bibliothèques : NetCDF4 et Xarray

Common Data Language (CDL)

Prérequis :

- Editeur de texte brut
- Connaissance de la syntaxe
- ncgen

Avantages :

- Simple et minimaliste

Inconvénients :

- Syntaxe lourde
- Peu d'éditeurs disponibles
- Métadonnées et données écrites à la main

```
netcdf cdl_example {
dimensions:
    t = 8 ;
    x = 6 ;
    y = 6 ;
variables:
    double time(t) ;
        time:long_name = "Chronological Julian UTC Datetime" ;
        time:standard_name = "time" ;
        time:units = "days since -4713-01-01T00:00:00Z" ;
        time:axis = "T" ;
        time:calendar = "julian" ;
    double latitude(y) ;
        latitude:long_name = "Latitude" ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degree_north" ;
        latitude:axis = "Y" ;
    double longitude(x) ;
        longitude:long_name = "Longitude" ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degree_east" ;
        longitude:axis = "X" ;
    short precipitation(t,x,y) ;
        precipitation:long_name = "Precipitation" ;
        precipitation:standard_name = "precipitation" ;
        precipitation:units = "mm" ;
        precipitation:coordinates = "longitude latitude" ;

// global attributes:
    :Conventions = "CF-1.13" ;
    :featureType = "profile" ;
    :title = "Example of dataset." ;

data:
    ...
}
```

La programmation

Prérequis :

- Expérience en programmation
- Familiarité avec le langage

Avantages :

- Versatile
- Conversion de jeux de données issus d'autres formats (CSV, etc.)
- Automatisation

Inconvénients :

- Expérience / Syntaxe du langage
- Temps de familiarisation avec la librairie
- Métadonnées écrites à la main

```
import netCDF4 as nc

ds = nc.Dataset('path/to/file.nc')

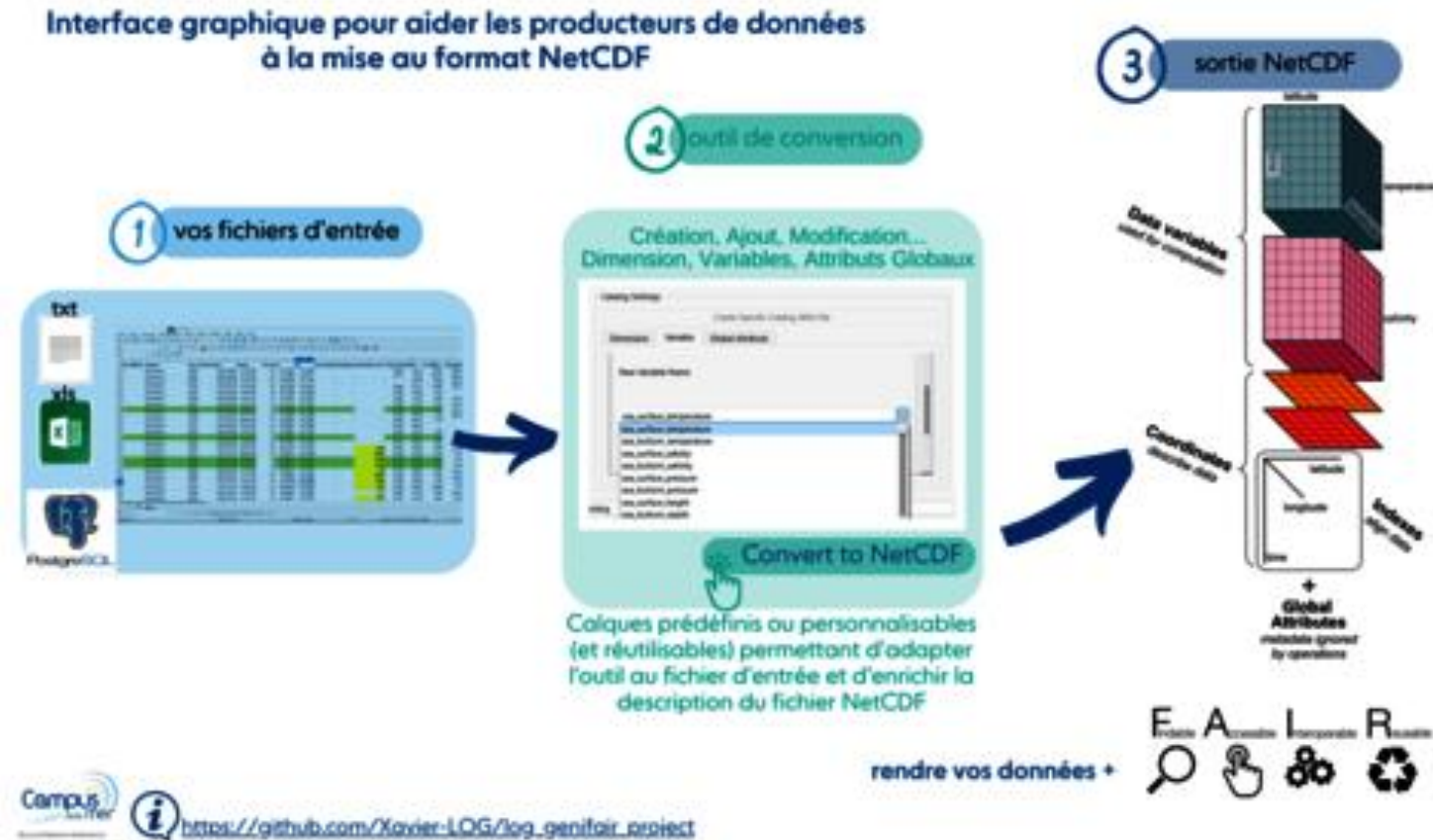
# Create dimensions
ds.createDimension('t',8)
ds.createDimension('x',6)
ds.createDimension('y',6)

# Create variables
time = ds.createVariable('time','f8',('t'))
time.units = 'days since -4713-01-01T00:00:00Z'
time.axis = 'T'
time[:] = [...]
...

lat = ds.createVariable('latitude','f8',('y'))
lat.units = 'degree_north'
lat.axis = 'Y'
lat[:] = [...]
...

# Create global attributes
ds.Conventions = 'CF-1.13'
ds.featureType = 'profile'
ds.title = 'Example of dataset.'
```

GENIFAIR : création de fichiers au format NetCDF



https://github.com/Xavier-LOG/log_genifair_project

nc_conv

Prérequis :

- Editeur de texte / code
- Installer le package Python
- (Expérience sur Python)

Avantages :

- Simple et minimaliste
- Fonctionnalités no-code
- Conversion de jeux de données issus d'autres formats (CSV, etc.)
- Automatisation
- Semi-automatisation des métadonnées

Inconvénients :

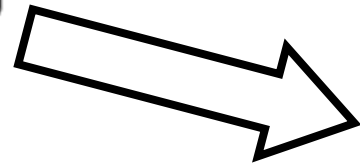
- Prise en main
- Remplissage des bases de métadonnées



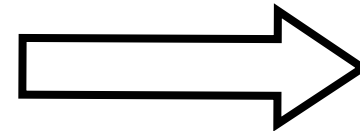
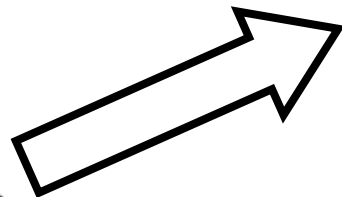
nc_conv



Metadata



Data



Références

- <https://docs.unidata.ucar.edu/netcdf-c/current/index.html>
- [https://sist.pages.in2p3.fr/webinaire netcdf 2024](https://sist.pages.in2p3.fr/webinaire_netcdf_2024)
- <https://sist.cnrs.fr/wp-content/uploads/2021/12/SIST19-Atelier-NetCDF-3-Generer-avec-python.pdf>
- <https://cfconventions.org/>
- <https://doranum.fr/tag/netcdf/>
- <https://youtu.be/UvNBnjiTXa0>

Merci à tous!

et à Quentin !

