

# Qui êtes-vous, quelle est votre communauté ?

- Joël SUDRE

- UAR Data Terra (Infrastructure de Recherche Data Terra)
- Ingénieur de Recherche CNRS, Océanographe
- Coordinateur technique Projet Equipex+ France 2030 Gaia Data



- Communauté Système Terre, Biodiversité et Environnement (5 pôles de données):

- Atmosphère (pôle AERIS)
  - Océan (Pôle ODATIS – Chargée de mission In situ)
  - Terre Solide (Pôle FormaTerre)
  - Surfaces continentales (Pôle Theia)
  - Biodiversité (Pôle PNDB)
- +
- DINAMIS (Dispositif Transversal : Accès aux données satellites THRS)
- +
- CLIMERI-France (Simulation de référence)



# Quelles données ? Type, formats, spécificités, volumétrie... Pratiques ou usages spécifiques à votre communauté?

- **Type de données:** Très hétérogènes (In situ, satellitaire, synthétique, ...)

Ex: Séries temporelles, images, vidéo, audio, matrice N dimensions, génomiques, ...

- **Formats:** De très nombreux types de format (brut, riche - propriétaire, ouvert - texte, natif, image, carte de points, vectoriel 2D,3D,ND, hybride, compression, stockage, échange...)

**Une liste non exhaustive:** 7Z, Archive Darwin Core, Audio Band Core, ASCII, AVI, BMP, BUFR, BZ2, CRD, CSV, DIMAP, EOL, EOS, EXIF, GEOJSON, GEOTIFF, GFC, GIF, GML, GNSS, GPS, GRD, GRIB, GZIP, HDF 4, HDF 5, HGT, HRIT, IBF, IMF, IAGA2, JSON, JPEG (JPG, jpeg2000,...), KML, Medaltas, MAT, MP3, MP4, NASA-AMES, NETCDF 3, NETCDF 4, NGS, ODS, ODV, PDF, PLY, PNG, R (Rdata et RDS), RAR, RINEX, SEED, SAFE, SHAPE, SINEX, TAR, TIFF, TSV, US-MGRS, VLBI, VRML, WAV, XLS, XLSX,XML, Zarr, ZIP

- **Volumétrie: Big data** Des centaines de Po qui ne cessent d'augmenter quotidiennement
  - Jeu de donnée: de quelques octets à plusieurs Téra
  - 150 Po de donnée distribuée par les pôles (estimation en 2021)
  - Arrivée de nouveaux capteurs à acquisition haute fréquence, drones, Satellite haute résolution, ...
- **Utilisateurs:** Communauté scientifique, acteurs publics et privés (Usages divers)