

# Science Ouverte et Interoperabilité en Héliophysique



<https://ihdea.net>

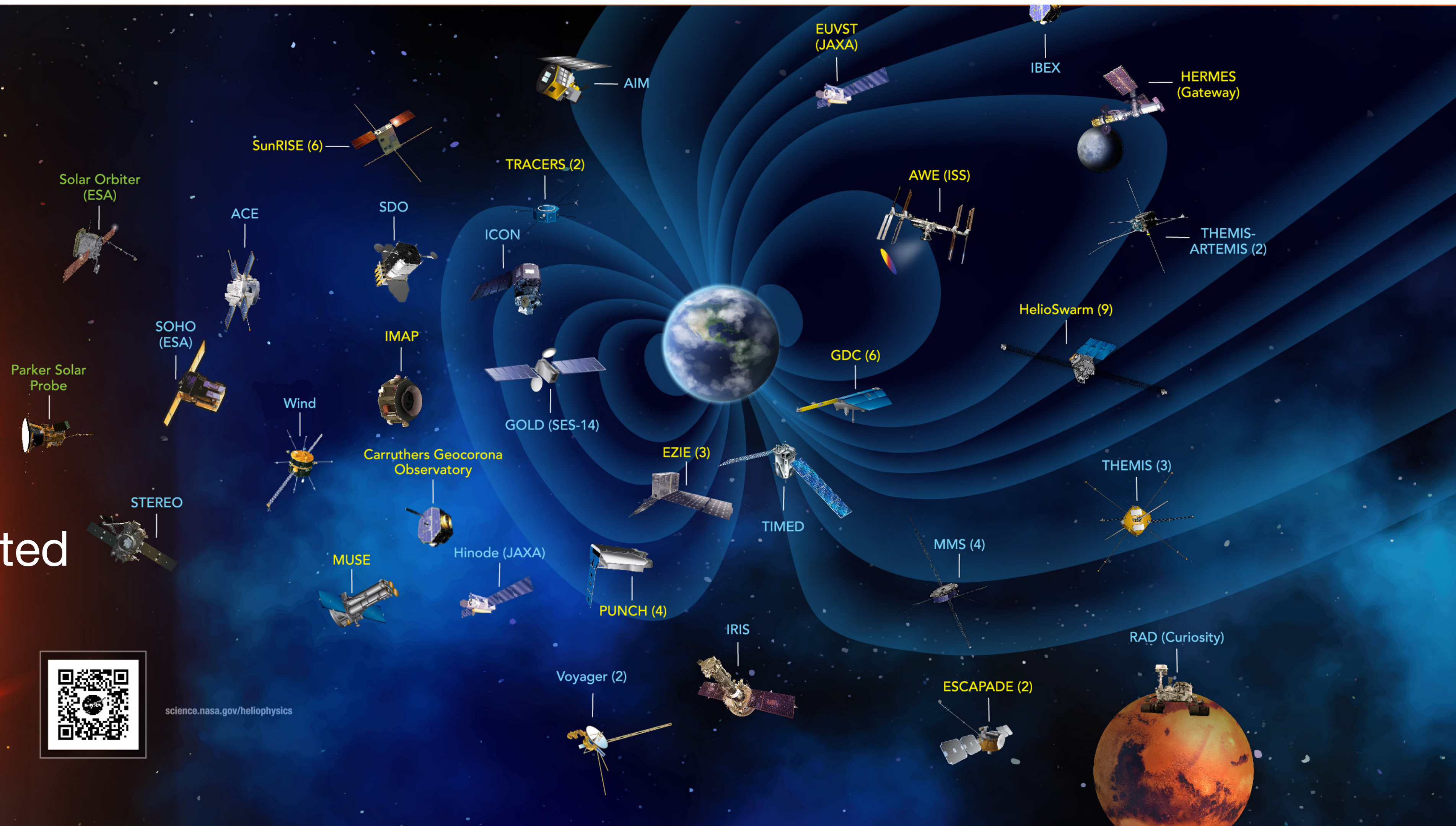
*International Heliophysics Data Environment Alliance*

# Héliophysique & physique spatiale

multi-longueur d'onde, multi-point, multi-messenger

missions NASA

- kHz to X-ray
- ground + space
- single + distributed
- in-situ + remote



[science.nasa.gov/heliophysics](https://science.nasa.gov/heliophysics)

# La partage n'a jamais été optionnel

## Science ouverte depuis 1998

- **Objets étudiés nécessitent le partage des données, à partir de plusieurs sources. Standardisation de l'accessibilité pour la réutilisation et la découverte.**
- CDF+ISTP (1998): standard pour l'échange de données
- SPASE : un modèle de données commun + un registre (2005)
- premier version du Virtual Solar Observatory (2005)
- Base de données d'événements solaires HEK (2007)
- FITS+WCS pour la physique solaire (2011)
- Extension de SPASE pour les simulations numériques (projet FP7 IMPEX : 2014)
- HAPI : accès aux séries temporelles (2021)  
Heliophysics API: <https://github.com/hapi-server>

# L'héliophysique aujourd'hui sur le chemin du FAIR

- Principes FAIR (Findable, Accessible, Interoperable, Re-usable)
- **Findable:**
  - VSO (<https://sdac.virtualsolar.org/>),
  - HDP (<https://heliophysicsdata.gsfc.nasa.gov/>)
  - VESPA (<https://vespa.obspm.fr/>)
  - AMDA (<http://amda.cdpp.eu/>)
- **Accessible:**
  - toutes les données sont publiques, avec des formats standards
  - modules Python pour accéder aux données (PyHC)
- **Interoperable**
  - HAPI = accès interoperable pour les séries temporelles
  - format de données (CDF, FITS)
- **Reusable:**
  - Registre SPASE contient des DOI pour la citation des données
  - modules Python pour travailler/processer les données (PyHC)

The screenshot shows the Heliophysics Data Portal interface. At the top, it features the NASA logo and the text 'GODDARD SPACE FLIGHT CENTER Space Physics Data Facility'. Below this is a navigation bar with tabs for 'Help', 'Geo Orbits', 'Helio Orbits', 'SPASE Registry', 'ADS Abstracts', and 'Feedback'. The main content area is divided into two columns. The left column contains search filters: 'Text Restriction', 'Time Span Restriction' (with fields for 'from:' and 'to:'), and 'Element Restriction' (with sub-filters for Resource type, Measurement type, Observatory Group, Observatory, Instrument, Observed region, Spectral range, Cadence, Repository Name, Access rights, and Format). The right column displays 'Current Product Restrictions' (none are currently set) and a list of 10 data products. Each product entry includes a number, a title, a DOI link, and a list of access links (e.g., 'Canada ABOVE', 'Geographic and magnetic coordinate stations', 'Polar-Wind-Geotail 'gif-walk' site', 'FTP from SPDF', 'CDWeb', 'CDWeb HAPI Server').

#	Products (& SPASE descriptions)	Information and Access Links
1	ABOVE Level-2 Amplitude and Phase VLF Transmitter	Canada ABOVE Geographic and magnetic coordinate stations
2	ABOVE VLF Level-1 time series uncalibrated CDF	Canada ABOVE Geographic and magnetic coordinate stations
3	ACE 27-day Survey Plots	Polar-Wind-Geotail 'gif-walk' site Get Images/Plots
4	ACE Cosmic Ray Isotope Spectrometer (CRIS) 1-Hour Level 2 Data <a href="https://doi.org/10.48322/g72t-0814">https://doi.org/10.48322/g72t-0814</a>	FTP from SPDF (not with most browsers) HTTPS from SPDF CDWeb CDWeb HAPI Server ACE Science Center FTP area ACE Science Center ACE Cosmic Ray Isotope Spectrometer (CRIS) Level 2 data Home Page Get Data/Plots
5	ACE Cosmic Ray Isotope Spectrometer (CRIS) Daily Level 2 Data <a href="https://doi.org/10.48322/wbmv-4z80">https://doi.org/10.48322/wbmv-4z80</a>	FTP from SPDF (not with most browsers) HTTPS from SPDF CDWeb CDWeb HAPI Server ACE Science Center ftp area ACE Science Center ACE Cosmic Ray Isotope Spectrometer (CRIS) Level 2 data Home Page Get Data/Plots
6	ACE Daily Survey Plots	Polar-Wind-Geotail 'gif-walk' site Get Images/Plots
7	ACE Definitive Attitude, Hourly Values, Direction Cosines in RTN, GSE, and J2000 GCI Coordinates <a href="https://doi.org/10.48322/f2a8-nm36">https://doi.org/10.48322/f2a8-nm36</a>	FTP from SPDF (not with most browsers) HTTPS from SPDF CDWeb CDWeb HAPI Server ACE Attitude and Orbit Data Website Satellite Situation Center, SSCWeb, System and Services Get Data/Plots
8	ACE Definitive Orbit, Daily Values, in GSE and J2000 GCI Coordinates <a href="https://doi.org/10.48322/mh1w-x115">https://doi.org/10.48322/mh1w-x115</a>	FTP from SPDF (not with most browsers) HTTPS from SPDF CDWeb CDWeb HAPI Server ACE Attitude and Orbit Data Website Satellite Situation Center, SSCWeb, System and Services Get Data/Plots
9	ACE Electron Proton Alpha Monitor (EPAM) 1-Hour Level 2 Data <a href="https://doi.org/10.48322/pn1t-zx03">https://doi.org/10.48322/pn1t-zx03</a>	FTP from SPDF (not with most browsers) HTTPS from SPDF CDWeb CDWeb HAPI Server ACE Science Center ftp area ACE Science Center ACE Electron Proton Alpha Monitor (EPAM) Level 2 data Home Page Rules of use, and caveats Get Data/Plots
10	ACE Electron Proton Alpha Monitor (EPAM) 12-sec Count Rate Data <a href="https://doi.org/10.48322/qaw3-9z28">https://doi.org/10.48322/qaw3-9z28</a>	FTECS site EPAM Overview at ASC Explanation of print-groups and CSV file nomenclature. FTP from SPDF (not with most browsers)

### Search for Solar Physics Data Products:

If you're new to the VSO, see [How To Search](#), the [FAQ](#) or click the icons for online help.

Please select which values you wish to use to search for data products:

- Time**  
Search by time interval.  
[Derive time intervals from event catalogs](#)
- Observable**  
Search based on physical observables
- Instrument / Source / Provider**  
Search based on instruments or data archives
  - Compact listing
  - Instrument / Source (not provider dependent)
  - Instrument Only (not source or provider dependent)
- Spectral Range**  
Search based on a spectral range
- Nicknames**  
Search based on common terms used to describe data products  
**Note:** Nicknames generate an intersection with other search terms, so searching for a nickname observable (or other parameter) when a nickname defines other physical observables will result in fewer results.  
 Show Nickname Definitions



VSO  
PyHC  
AMDA  
VESPA  
Autoplot  
...

### Projects

To add a project to this page, please refer yourself to the [project addition instructions](#).

### Core packages

These packages each offer a wide range of functionality in their area, and conform to the PyHC community [standards](#).

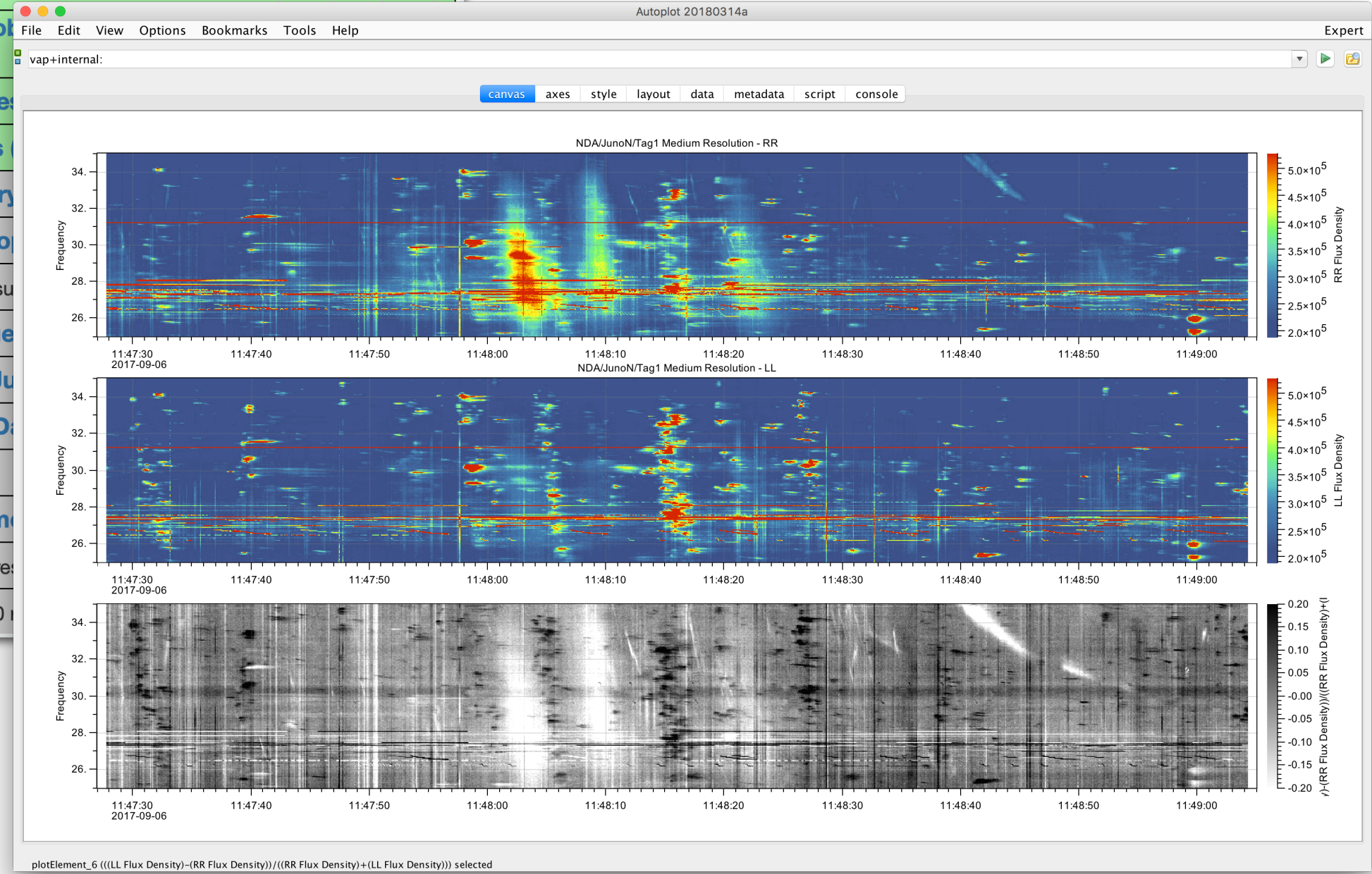
Name	Description	Code	Docs	Site	Contact	Community	Documentation	Testing	Software Maturity	Python 3
Python for Solar Physics					Stuart Mumford	Good	Good	Good	Good	Good
Space science library for Python	Includes file I/O, time and coordinate conversions, common analysis techniques.				Steve Morley	Good	Good	Partially met	Good	Good
Tools for loading, analysis and plotting of data from various heliophysics missions and ground magnetometers					Nick Hatzigeorgiu, Eric Grimes	Good	Good	Good	Good	Good

### Searching against current VSO instances

The screenshot shows the VSO software interface. On the left, the 'Workspace Explorer' displays a tree view of resources and operations. The 'Plot Manager' shows a list of plots for 'Cassini-Solar-Wind', including 'Panel (#0, Time Plot)', 'Panel (#1, Time Plot)', and 'Panel (#2, Time Plot)'. The main window displays three stacked plots: a time series plot of 'Cassini MAG', a spectrogram of 'Cassini Ring Frequency', and a time series plot of 'Cassini Ring Altitude'. The bottom status bar shows 'Afficher un menu', 'Workspace Explorer', 'Plot Manager', and 'Cassini-Solar-Wind'.

The screenshot shows the VESPA search interface. The 'Refine your search' section includes filters for 'Main Parameters' (Target Name: Sun), 'Target Class', 'Dataproduct Type', 'Instrument Host Name', 'Instrument Name', and 'Processing level'. The 'Data Services' list includes:
 

- AMDA - Planetary and heliophysics plasma data at CDPP/AMDA 743128 results
- bass2000 - Bass2000 solar survey archive 219295 results
- CLIMSO - CLIMSO coronagraphs at Pic du Midi de Bigorre 1098499 results
- eit\_syn - Synchronous synoptic maps of the solar corona from EIT/SoHO 18482 results
- Gaia-DEM - Thermal structure maps of the solar corona from SDO 783668 results
- HFC1AR - Heliophysics Feature Catalog active regions 1211449 results
- HFC1T3 - Heliophysics Feature Catalog type 3 radio bursts 90845 results
- NDA Obs. Database - Nancay Decameter Array of 9811 results
- spectro\_planets - Spectra of planets and satellites
- VizieR\_planets - VizieR Solar System Catalogues
- abs\_cs - Data for numerical modeling of planetary
- APIS - Auroral Planetary Imaging and Spectroscopy
- BaseCom - The Nancay Cometary Database 0 results
- BDIP - IAU database of historical planetary images
- cassini\_jupiter - Cassini RPWS/HFR Calibrated Jupiter
- cpstasm - CLUSTER STAFF-SA Spectral Matrix Database
- CRISM\_speclib - CRISM spectral library 0 results
- DynAstVO - Asteroid orbital database and ephemerides
- ExoPlanet - Extrasolar Planets Encyclopaedia 0 results
- Exotopo - Simulated Topography of Exoplanets 0 results

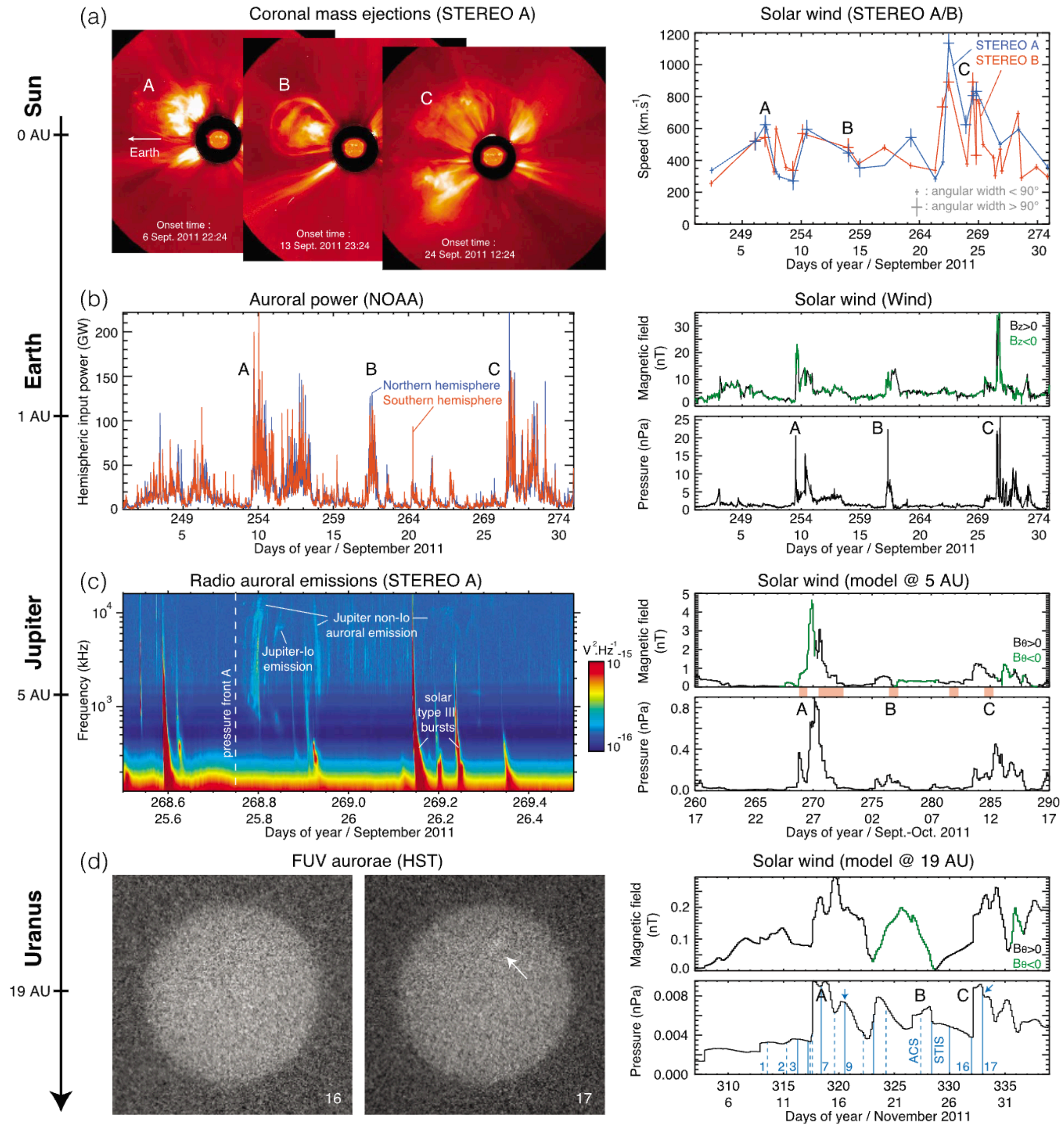


# Cas d'usage: *événement interplanétaire*

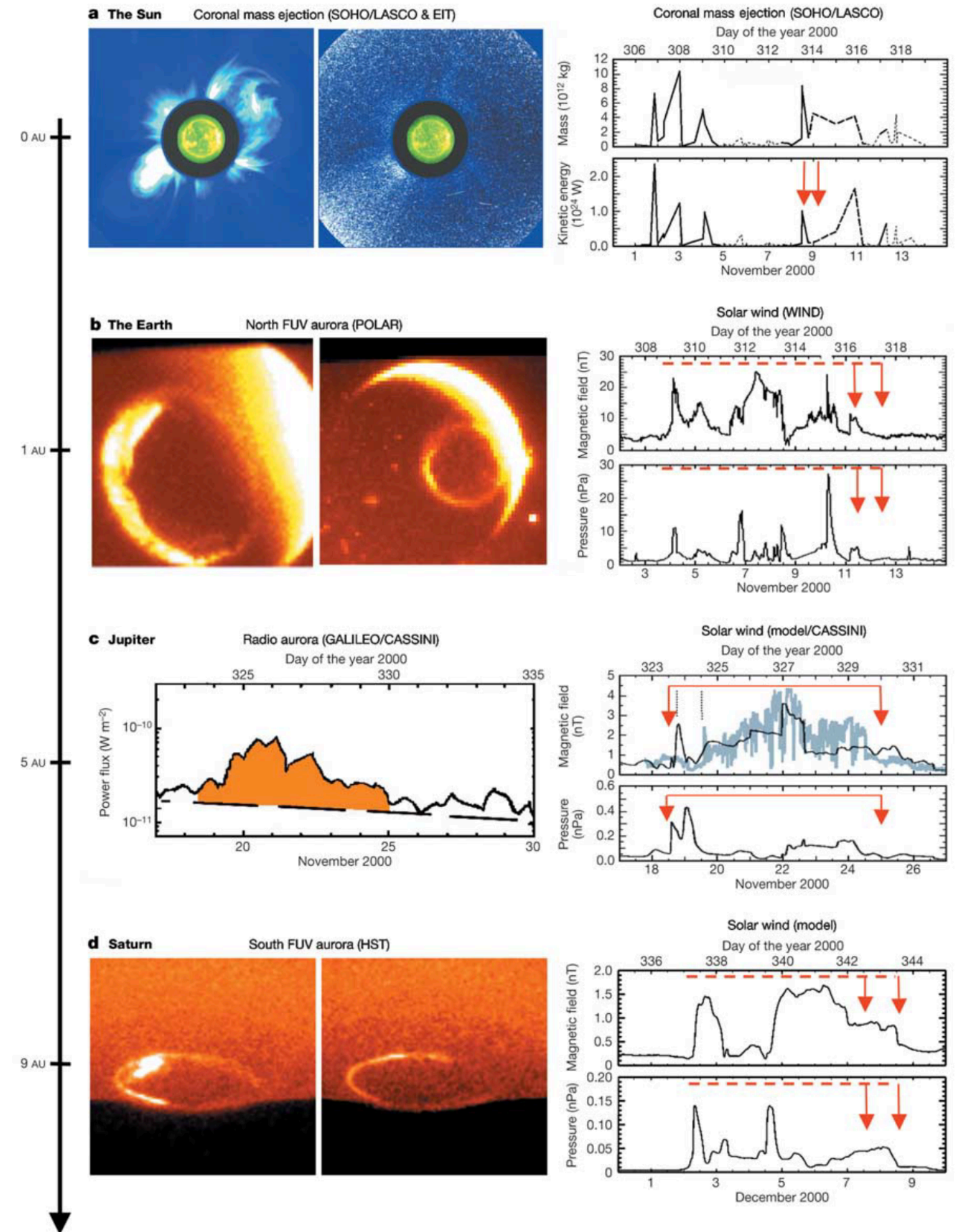
## Suivre un événement à travers le système solaire

- **Cas d'usage : Suivre l'impact d'un événement solaire à chaque planète.**
  - analyse des signatures observées => retrouver les processus initiaux
  - propager l'événement => vérifier si une signature est observée
- **Besoins:**
  - découverte de données (par exemple: données spatiales in-situ, données HST, données radio sol)
  - affichage et processing des données
  - découverte de modèles => conditions de vent solaire aux planètes (outils de propagation)
  - paramètre d'entrée des modèles => découverte de données
  - capacité de lancer un code
  - récupérer les résultats sur la même plateforme que les données initiales
- **Plusieurs études publiées** (faites à la main)
  - Cecconi et al. FrASS 9, (2022) doi:10.3389/fspas.2022.800279
  - Witasse et al. JGR 122, 7865–7890 (2017) doi:10.1002/2017ja023884
  - Lamy et al. GRL 39, (2012). doi:10.1029/2012gl051312
  - Prangé et al. Nature 432, 78–81 (2004). doi:10.1038/nature02986

- Lamy et al. (2012)



- Prangé et al. (2004)



# Accès aux données

## données ouvertes

- Données héliophysiques sont en accès ouvert
  - images, séries temporelles, événements...
- Outils and portails pour la découverte de données
  - plusieurs solutions (VSO, AMDA, CDAWeb, VESPA...)
  - besoin d'une meilleure interopérabilité sur les interfaces/protocoles de découverte
- Logiciels en open source (Python, IDL) pour le traitement des données
  - PyHC 🙌



# Simulation numériques

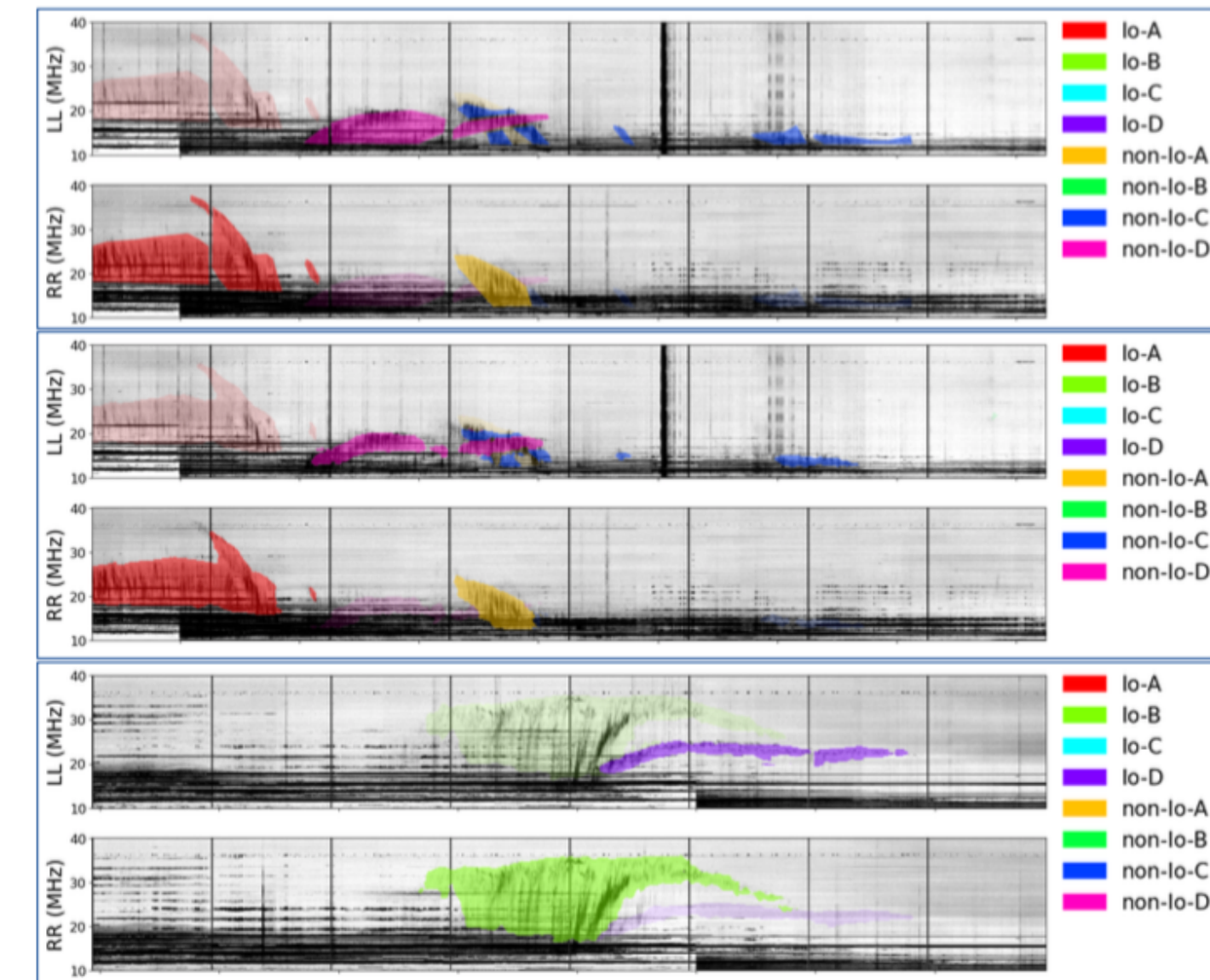
## modéliser et exécuter

- Simulations: modélisation, interprétation ou prédiction
- Codes numériques: des codes locaux aux codes communautaires
- Ouvrir les codes: pour la réutilisation (calcul à la demande) et pour la contribution (open source).
- Calcul à la demande: questions du passage à l'échelle, de la maintenance, des interfaces, du stockage, des autorisations...
- Actuellement:
  - SPASE peut décrire des modèles et des résultats de simulation (contribution IMPEx)
  - CCMC et IMPEx ont des interfaces de calcul à la demande, mais pas standardisé
  - IVOA propose une interface de calcul à la demande (UWS)
- Le *cloud* peut-il aider pour le passage à l'échelle et les autorisations ?

# Intelligence artificielle

## détection d'événement et analyse de texte

- Principale application de l'IA en héliophysique est la **détection/classification d'événements/signatures dans les collections de données** (exemple: émissions radio, signature plasma in situ)  
=> nécessaire car observations continues sur des années/décennies
- Besoin d'accéder à des données d'entraînement (que faire pour les événements rares); et des données d'entrées (locales ou en flux) de qualité (ou contrôlées)
- Stockage des résultats pour la réutilisation: y a-t-il des standards communautaires ?
- Métadonnées de provenance pour la reproductibilité:
  - problématique générique pour l'IA
  - le partage du code n'est pas suffisant; le partage des coefficients du réseau non plus.
  - question de confiance dans les résultats
- Autre application (action ESA, NASA, CDPP) autour de **l'analyse textuelle des articles** pour extraire les instruments utilisés, les intervalles associés, etc.



# L'héliophysique demain

## vers le cloud et au delà

- le Big data nécessite de rapprocher le code des données.
  - besoin d'interopérabilité et portabilité du code (plutôt que de l'optimisation?).
  - gestion de workflow, reproductibilité
  - infrastructure commune pour les logiciels et les déploiements
- NASA <http://heliocloud.org/>  
ESA <https://datalabs.esa.int/>  
et d'autres initiatives dans d'autres domaines (EOSC, SKA...)
- Cloud de clouds:  
workflow à travers les clouds / science platforms

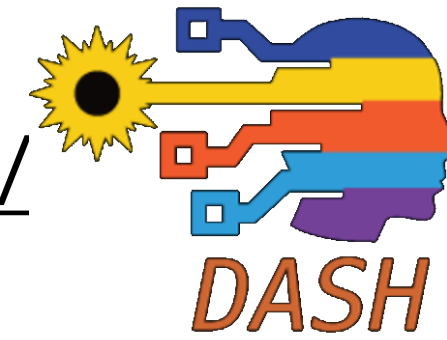
# l'héliophysique et l'IVOA

## IHDEA (International Heliophysics Data Environment Alliance)



<https://ihdea.net/>

- Oct 2023: DASH+IHDEA meeting: <https://dash.heliophysics.net/> (similaire à ADASS+IVOA)
- Pendant ces réunions: des liens plus étroits avec l'IVOA ont été évoqués.



Sujets identifiés:

### - Découverte de Ressources

Comment réutiliser ou s'inspirer du registre IVOA pour améliorer le registre SPASE ?

Utiliser EPNTAP comme interface de découverte ?

### - Sémantique

S'appuyer sur les vocabulaires du IVOA Semantics WG (reference frames)

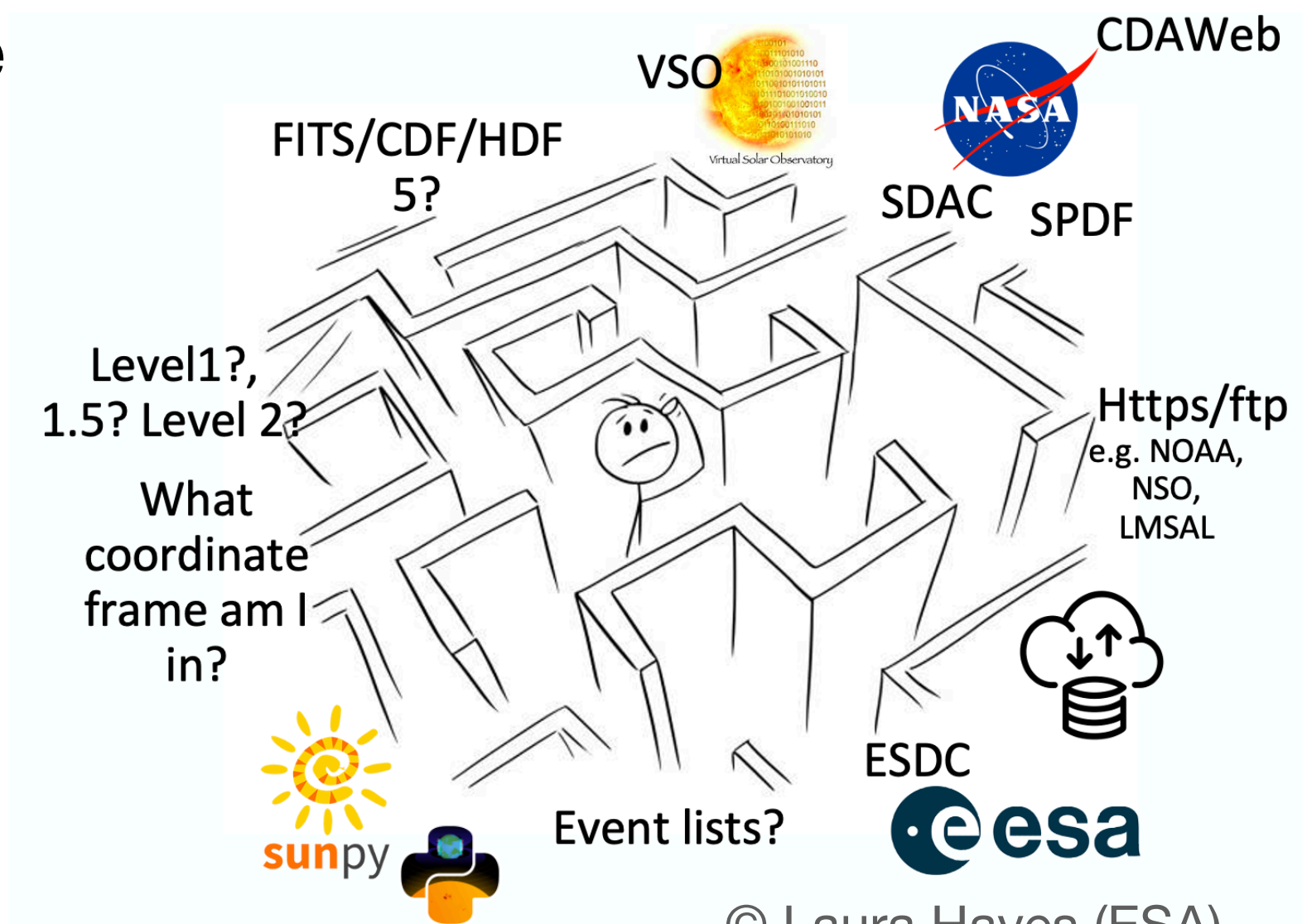
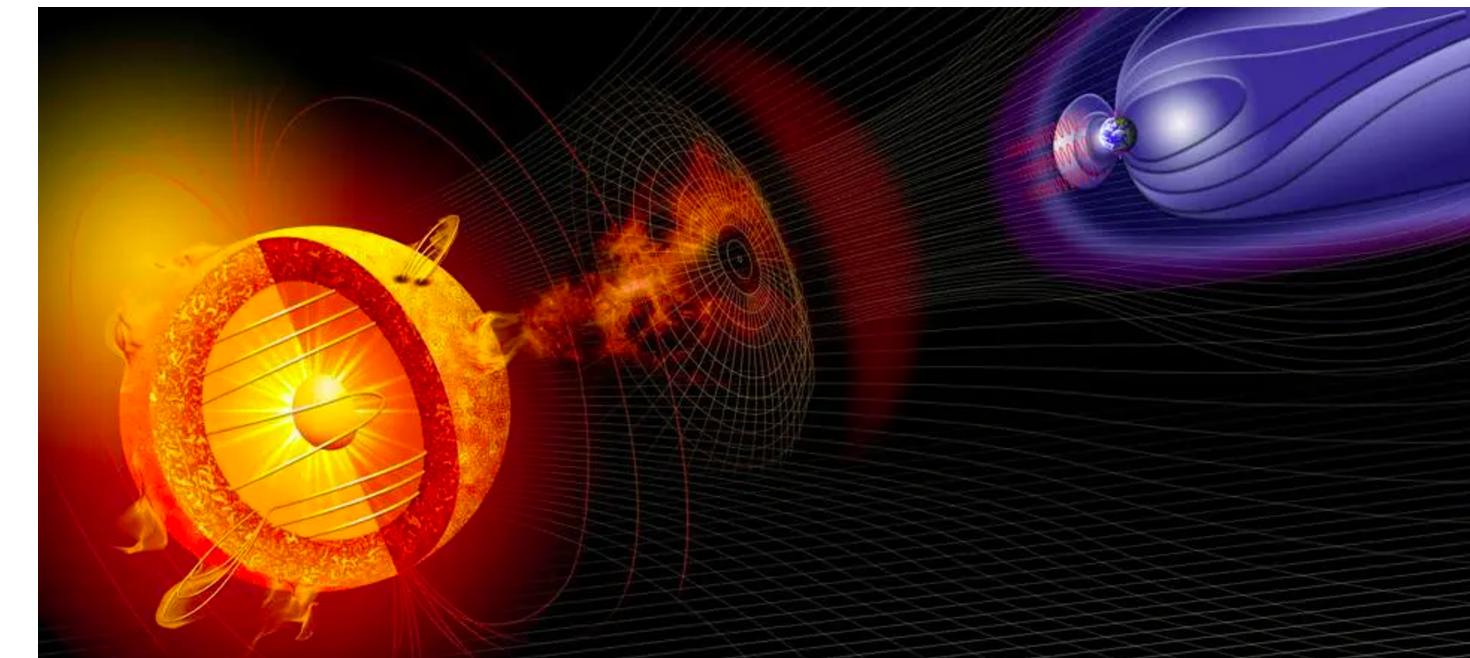
### - Domaine temporel

Améliorer les VOEvent pour l'héliophysique (RFC en préparation)

Intégrer HAPI dans l'IVOA (discussion TDIG + déjà intégré dans TOPCAT)

### - Accès

pousser pour l'adoption de TAP



# Résumé

## héliophysique et science ouverte

- Science ouverte par construction depuis le début
- Principaux moteurs: accès et réutilisation des données
- Standardisation des interfaces: améliorer ou réutiliser
- Modélisation: besoin d'interfaces standardisées (description, calcul, accès)
- IA: besoin de données de qualité; sorties aux standards communautaires et défi de la reproductibilité
- Prochaines étapes:
  - Le *Cloud* avec beaucoup de défis techniques
  - Rapprochement avec l'IVOA