



# Workflows in the framework of the HELIO and (Europlanet) VOs

**Xavier Bonnin & Baptiste Cecconi**

*LESIA, Observatoire de Paris*



## e-infrastructure FP7 project (No. 238969)



- June 2009 → November 2012
- VO dedicated to solar physics and heliophysics
- Service Oriented Architecture
- Next step → HELIOSpace (?)
- Main data providers are implied :  
**BASS2000, CDPP, MEDOC, VSO**

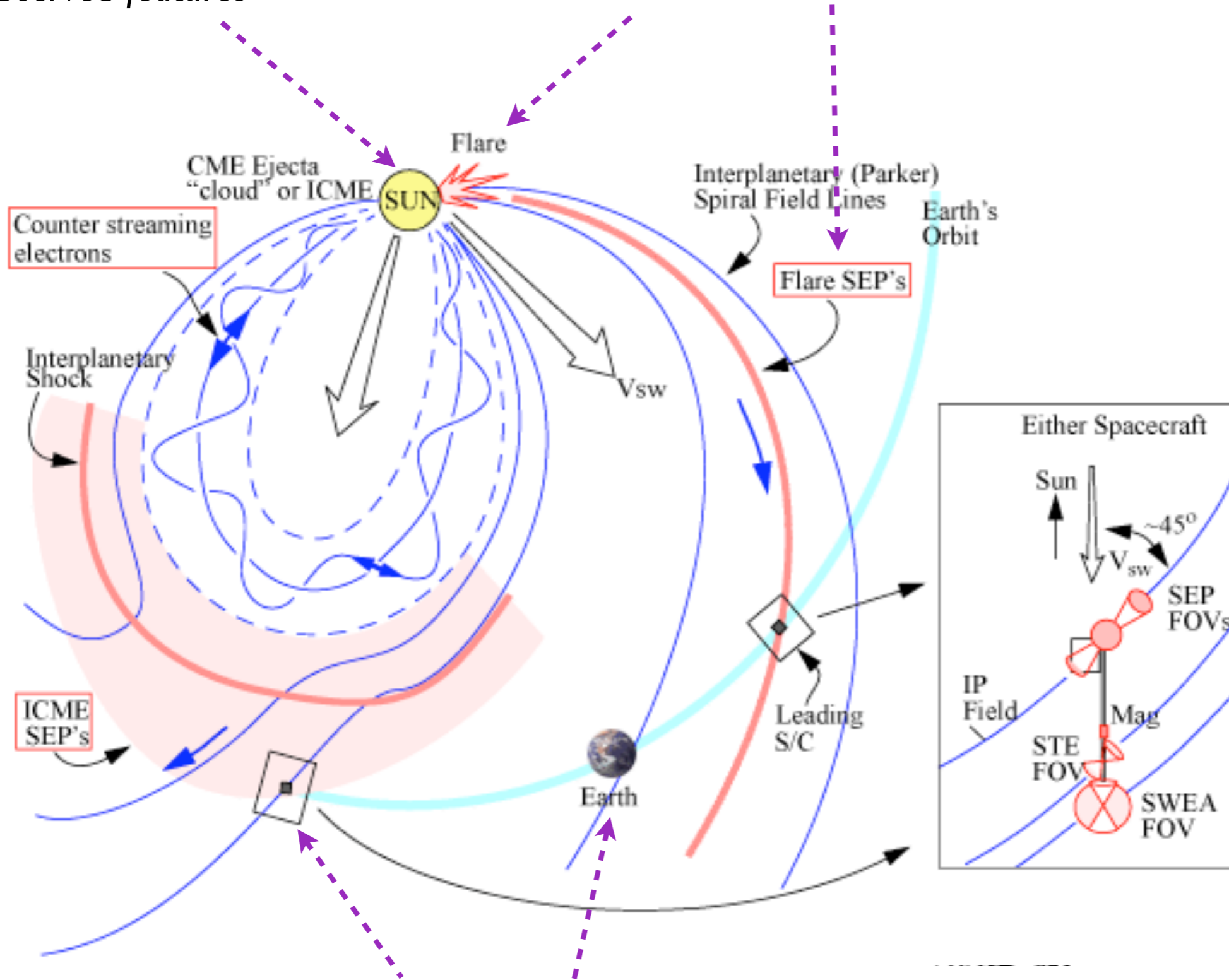
<http://www.helio-vo.eu/>

### 13 members

MSSL/UCL (UK)  
FHNW (Suisse)  
**Observatoire de Paris (France)**  
**CESR/UPS (France)**  
STFC/RAL (UK)  
**IAS/UPS (France)**  
INAF Trieste (Italie)  
**Univ. Manchester (UK)**  
Trinity College Dublin (Irlande)  
NASA/GSFC/HDMC (USA)  
RPI (USA)  
LMTAC (USA)  
ESA (Europe/Espagne)

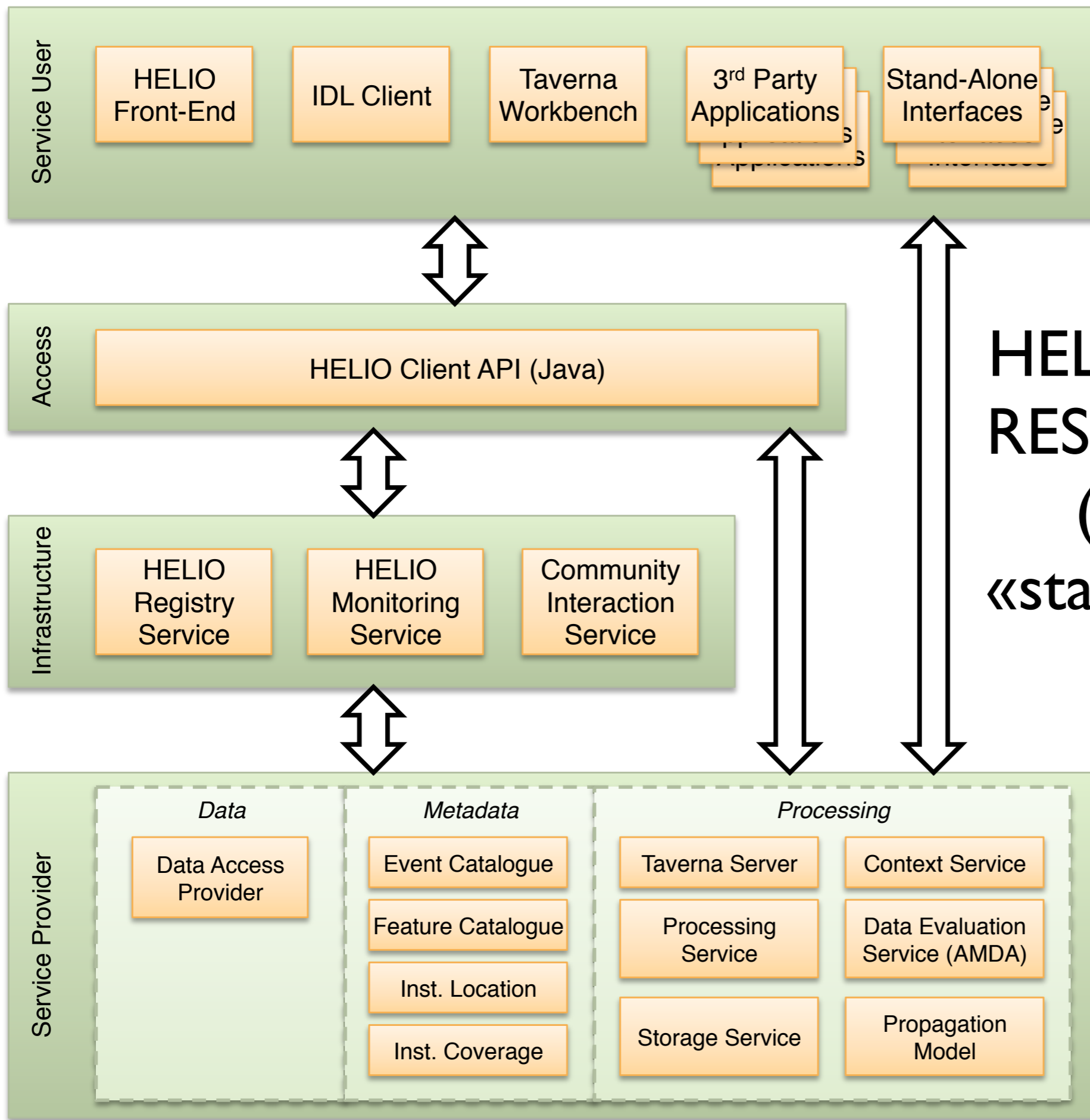
Information about observed features

Information about observed events

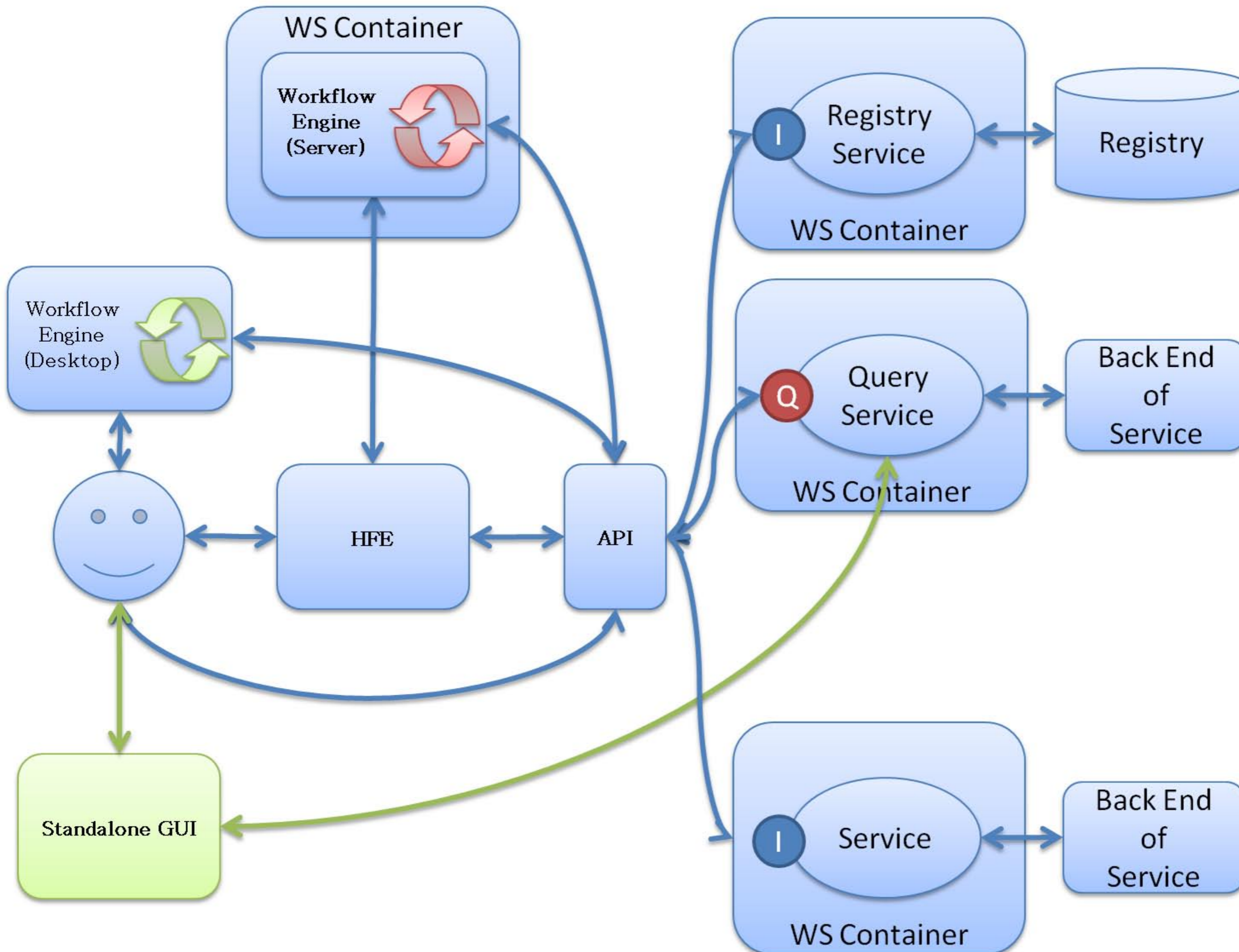


Information about observations available

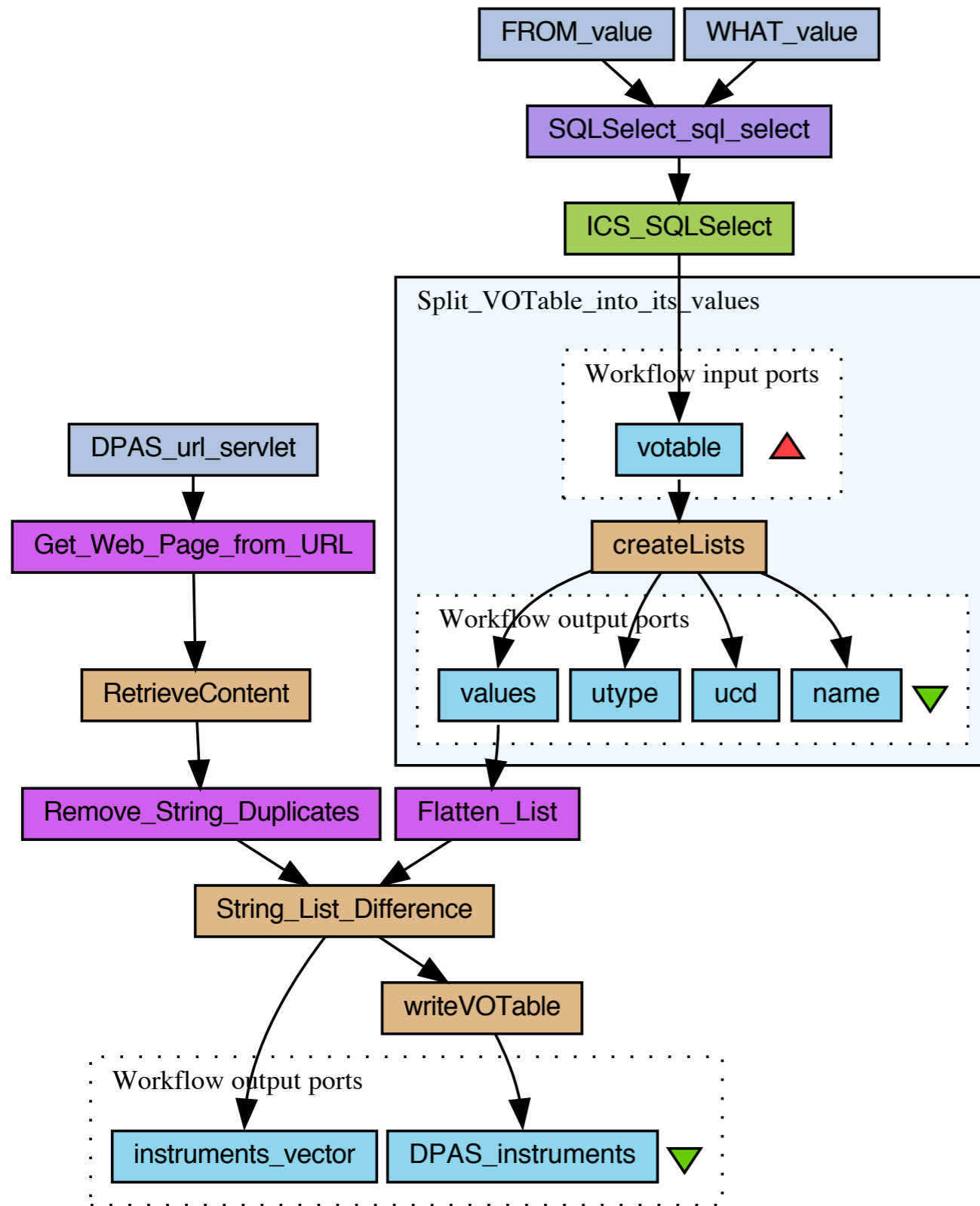
**Idea :**  
Associate data/meta-data according to physics  
↓  
Use propagation models to associate data



HELIO services has a REST/SOAP interface (HQL) allowing «standard» queries in the system



# Workflow use case (1/3)



This workflow checks the consistency between instrument IDs which are used in the DPAS and are defined in the ICS

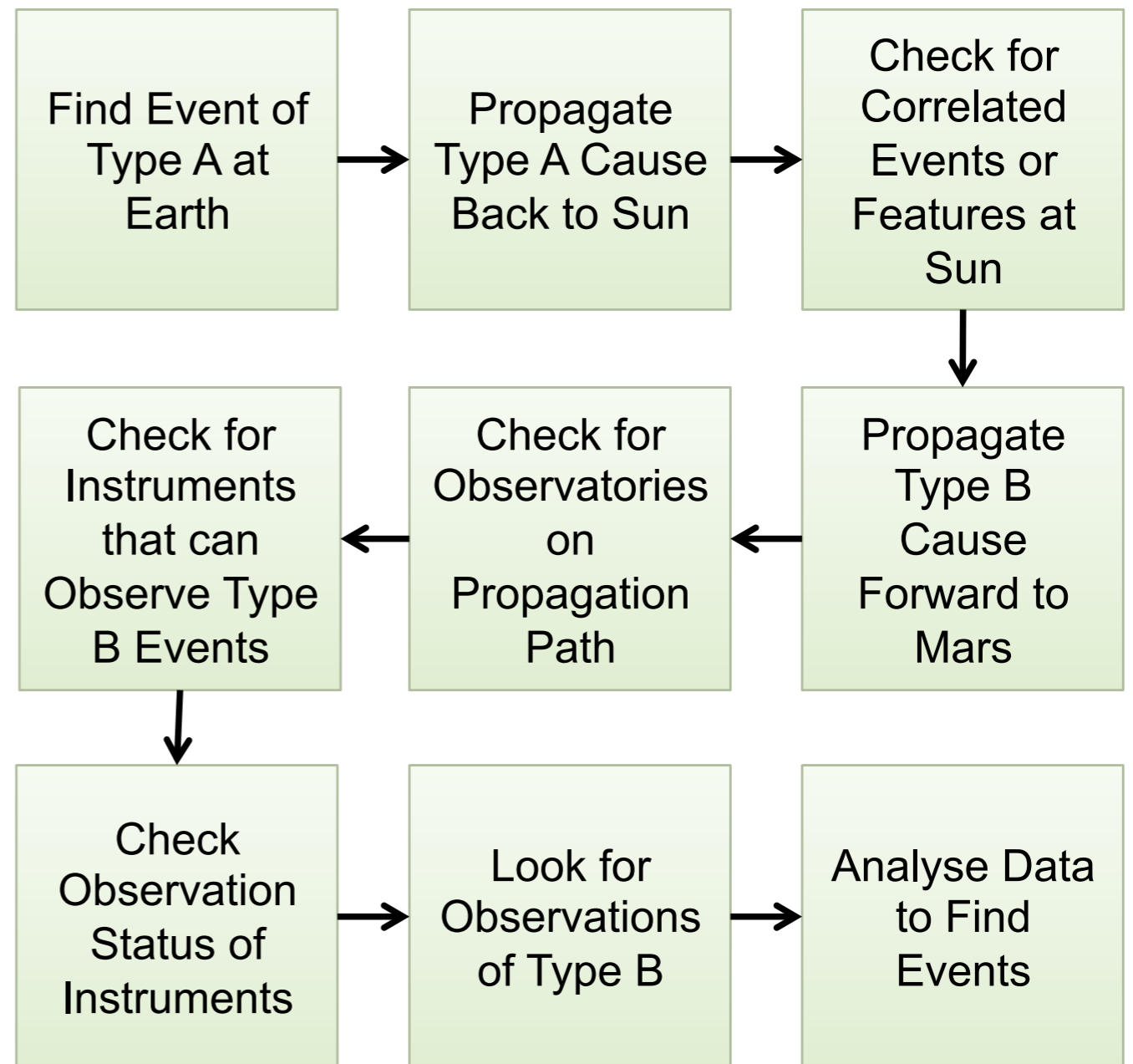
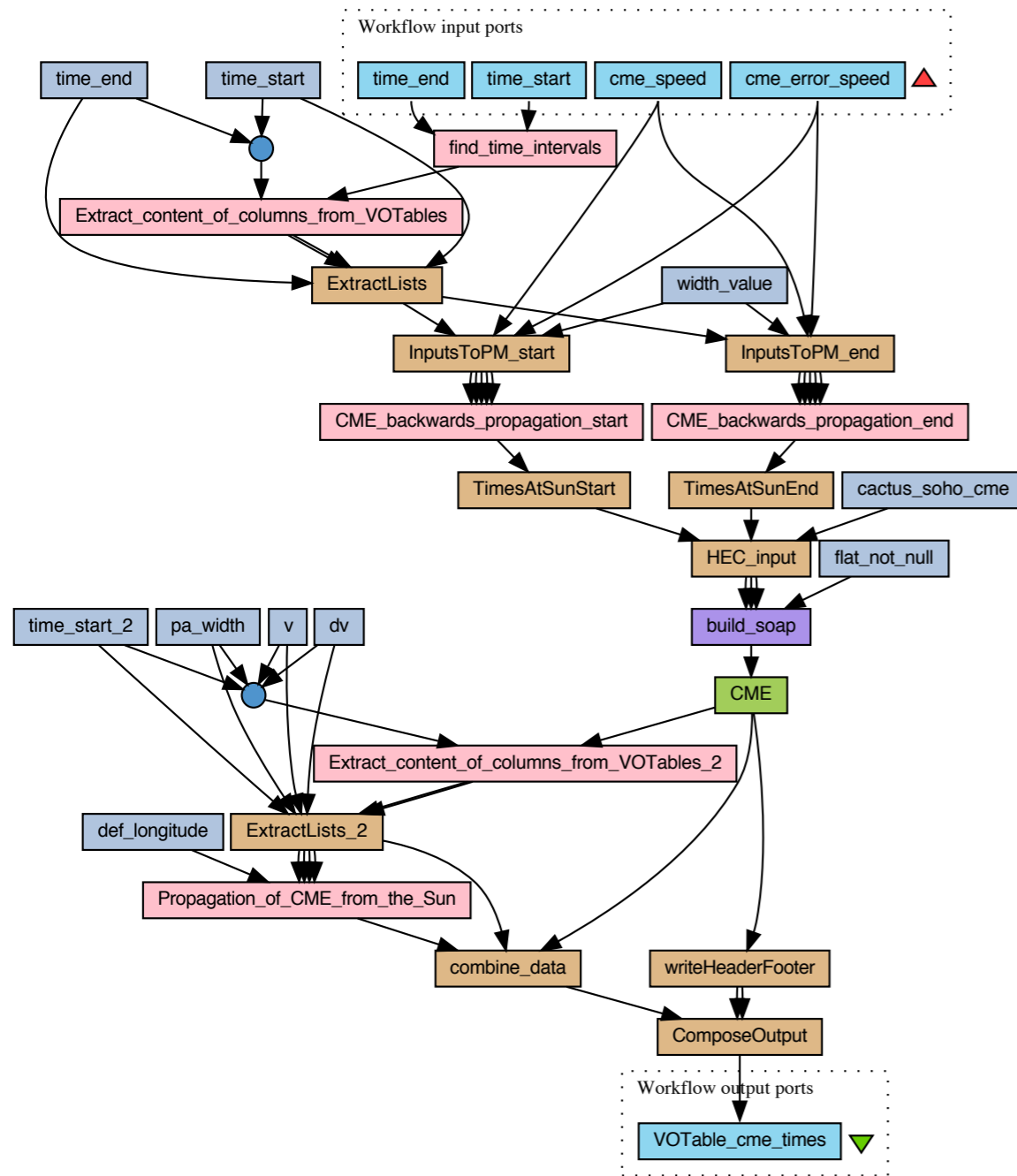
(from Anja et al., submitted)

DPAS = Data Provider Access Service  
ICS = Instrument Coverage Service

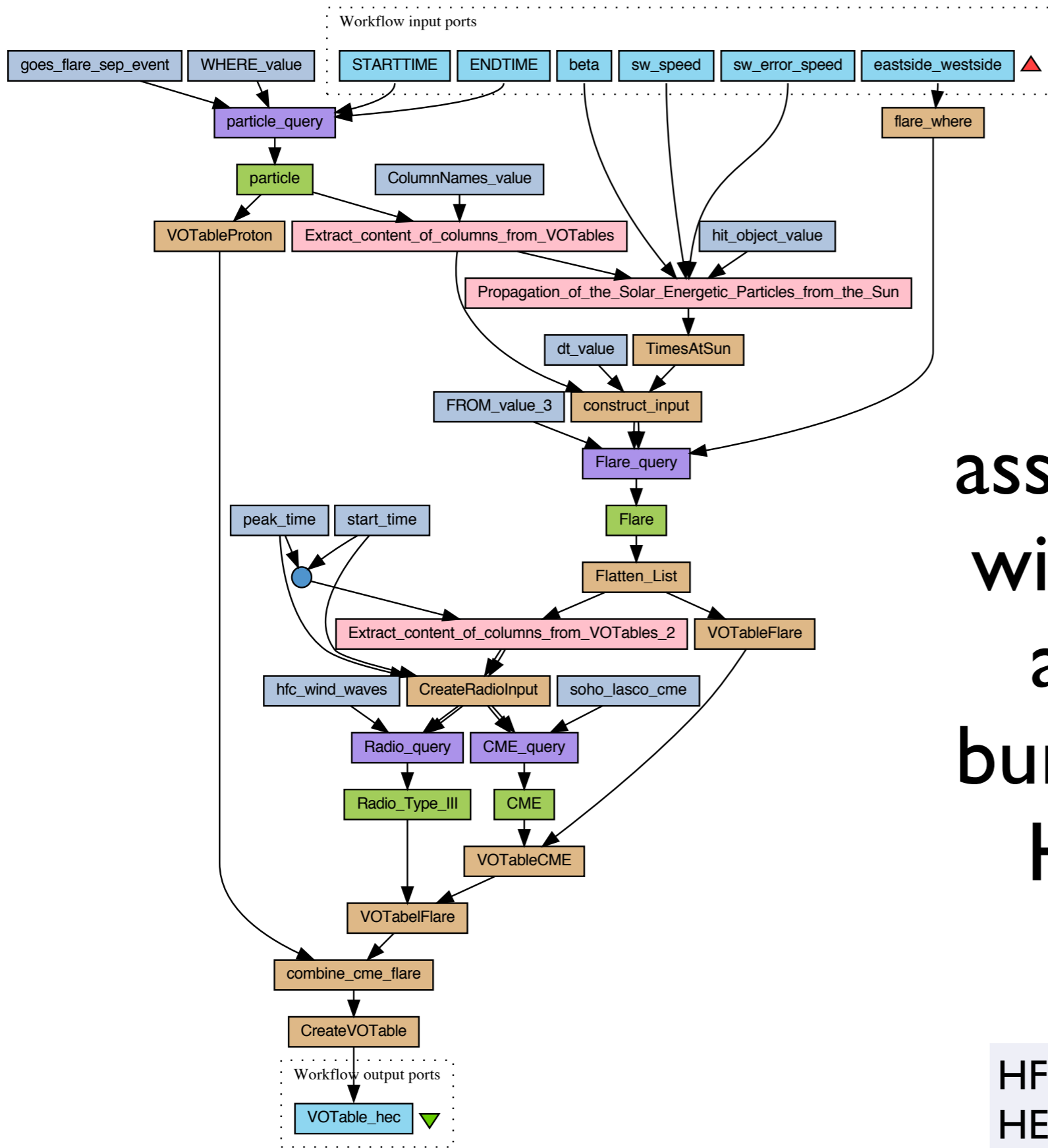
# Workflow use case (2/3)

## Propagation of CMEs at Earth and Mars

(from Bentley et al., 2012)



# Workflow use case (3/3)



This workflow associating SEP events with solar flare, CME and Type III Radio burst events using the HFC and the HFE

(from Anja et al., submitted)

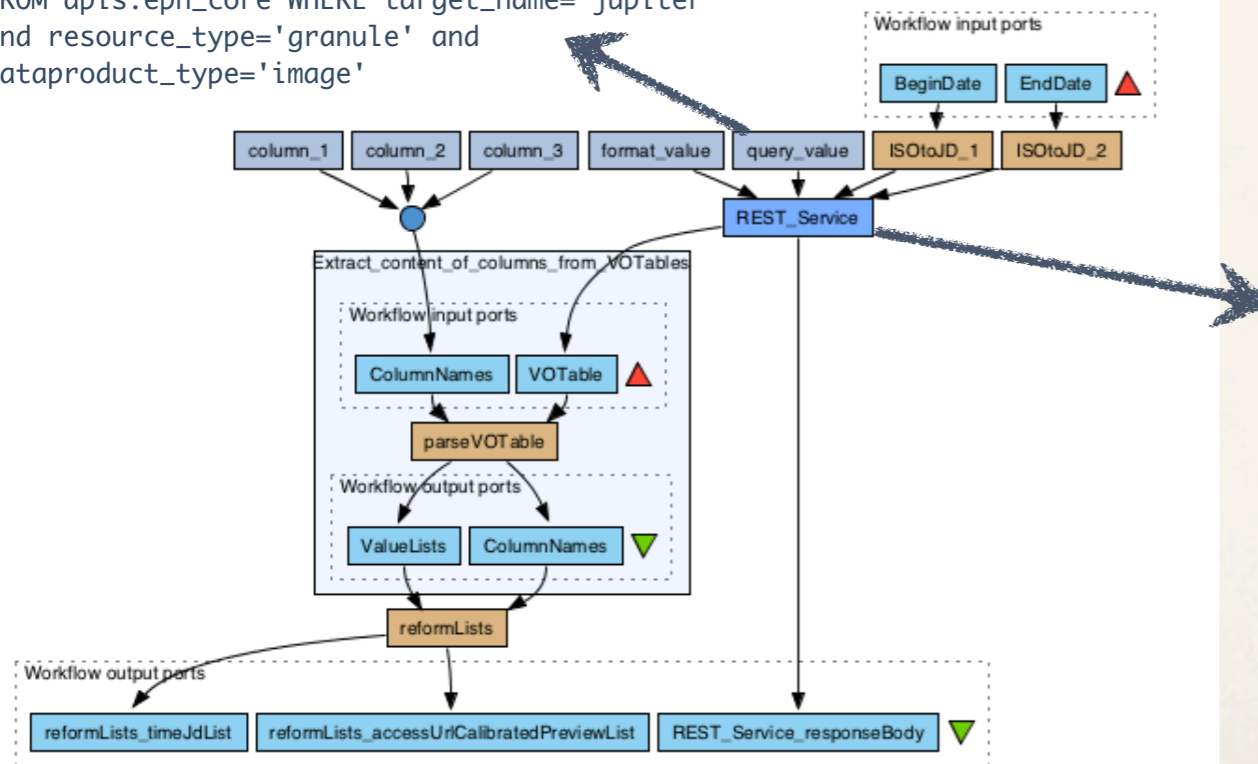
HFC = Heliophysics Feature Catalogue  
HEC = Heliophysics Event Catalogue



# Interroger un service TAP (Table Access Protocol)

- \* Les services TAP (e.g., ceux qui utilisent le framework DaCHS) incluent une interface REST. On peut donc très facilement construire un workflow utilisant un service TAP.
- \* En particulier, dans le cadre d'Euoplanet IDIS, on a développé une déclinaison de TAP adaptée aux sciences planétaires: EPN-TAP.

```
SELECT t_min,t_max,access_url_processed_preview  
FROM apis.epn_core WHERE target_name='jupiter'  
and resource_type='granule' and  
dataprodect_type='image'
```



Workflow explorer Details Validation report

### REST Service Details

HTTP Method: GET

URL Template: `http://voparis-tap.obspm.fr/__system__/adql/query/form?__nevow_form__=genForm&query={query}%20and%20t_max%3E[jdmin]%20and%20t_min%3C[jdmax]&_TIMEOUT=5&_FORMAT={format}&_VERB=H&_TDENC=on&submit=Go`

'Accept' header: application/xml

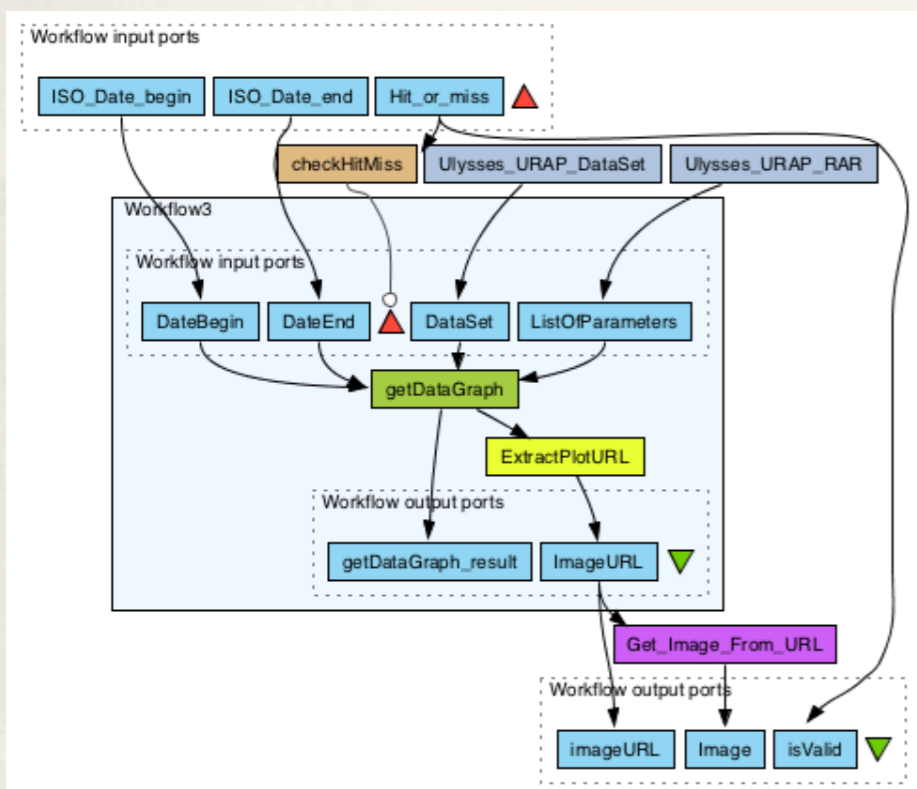
Configure

- + List handling
- + Predicted behavior
- + Advanced
- + Annotations

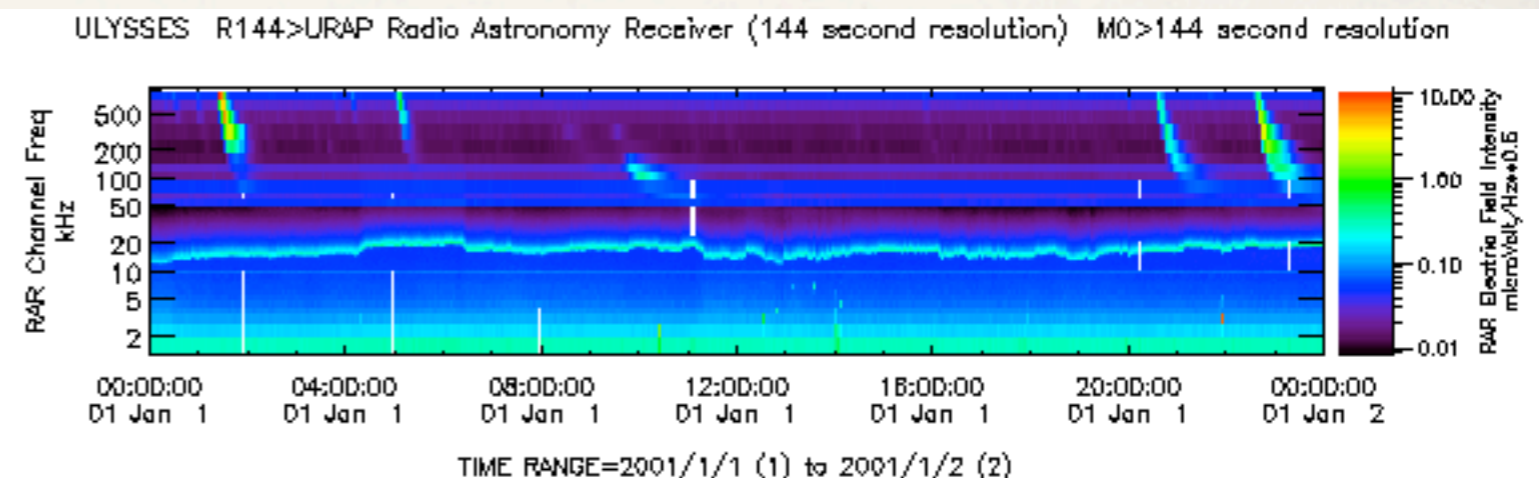
# Interroger des WebServices pour obtenir des données

- ❖ La plupart des centres d'archivages (en physique spatiale, en tout cas) ont des webservices permettant de récupérer des images résumée des données, ou bien les données elles-mêmes.
- ❖ Exemple: images calculées lors d'une requête sur le webservice du CDAWeb (NASA) pour un interval temporel et un instrument d'une mission spatiale.

## Workflow:



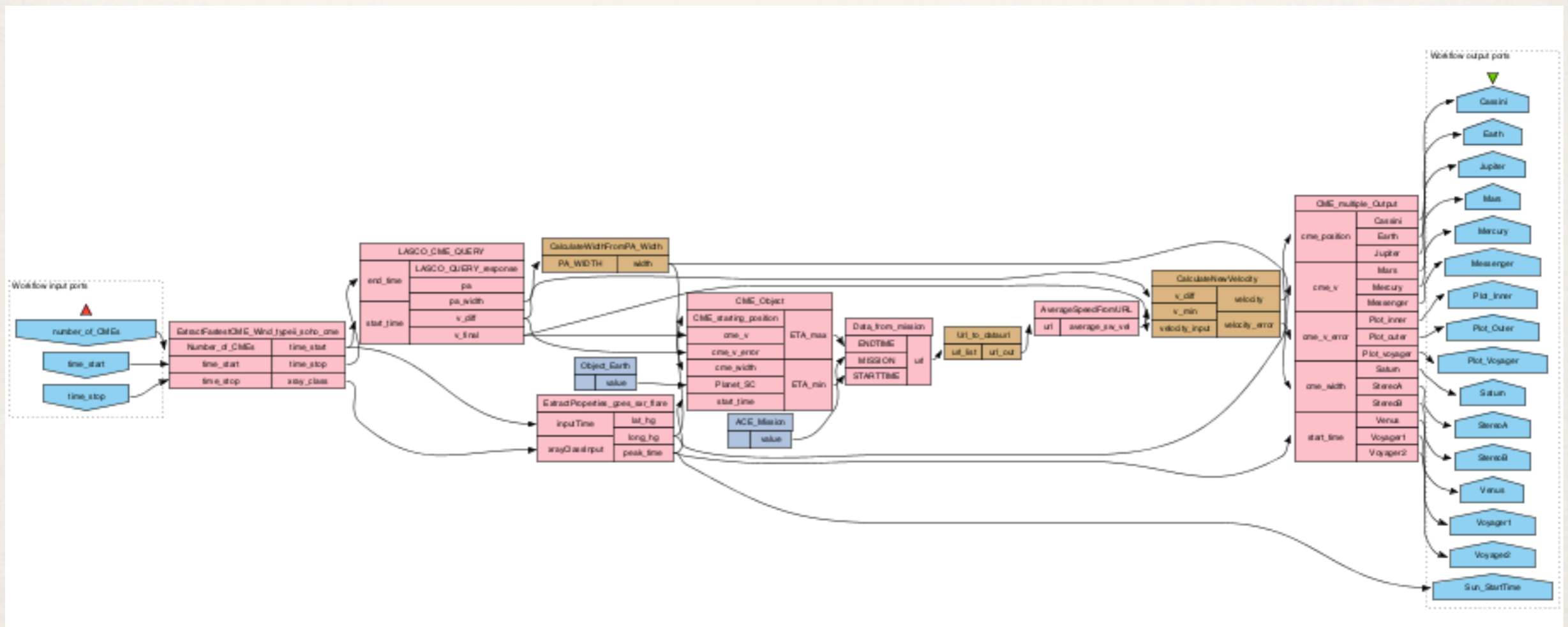
## Réponse du webservice:



Please acknowledge data provider, R MacDawall and CDAWeb when using these data.  
Generated by CDAWeb on Wed Nov 28 12:47:39 2012

# Un Workflow développé dans le cadre d'HELIO

- ❖ **But:** construire un workflow qui repère les 100 CME (éjections solaires) les plus intenses, calcule les temps d'arrivée à chaque planète ou sonde spatiale, vérifie quels instruments sont disponibles au moment de l'arrivée du choc aux différents endroits puis récupère les données (images résumées, et vraies données).



# Limites (actuelles ?) des workflows

---

- \* Taverna est une «plaie» (*désolé, mais il fallait que ça sorte :-)*  
Le principe est excellent, mais dès qu'on essaye de faire des tâches un peu complexe et des workflows un peu long, c'est inutilisable.
- \* Edition des scripts en Beanshell uniquement.
- \* Gestion des VOTables pas native, donc on est souvent obligé d'en extraire l'info pour la passer d'un service à l'autre.
- \* En général les webservices sont assez mal documentés, donc pas facilement utilisables (voire pas du tout!)  
*Ce n'est pas de la faute de Taverna, pour le coup...*

# Avantage des workflows

---

- ❖ Les services OV font des tâches simples et bien identifiées.
- ❖ Pour utiliser l'OV, les utilisateurs doivent :
  - soit jongler entre les différents services «à la main»
  - soit utiliser les outils existants, mais avec des tâches prédéfinies et génériques
  - soit créer un workflow pour faire les traitements complexes et pas génériques dont il a besoin pour aller plus loin.
- ❖ C'est clairement le «chaînon manquant» entre les services proposés dans le cadre de l'OV !
- ❖ myExperiment permet de partager les workflows qu'on a produit : base de connaissance et de partage !

HELIO group on  **myexperiment**

<http://www.myexperiment.org/groups/101.html>

Web page of the HELIO project

<http://www.helio-vo.eu/>

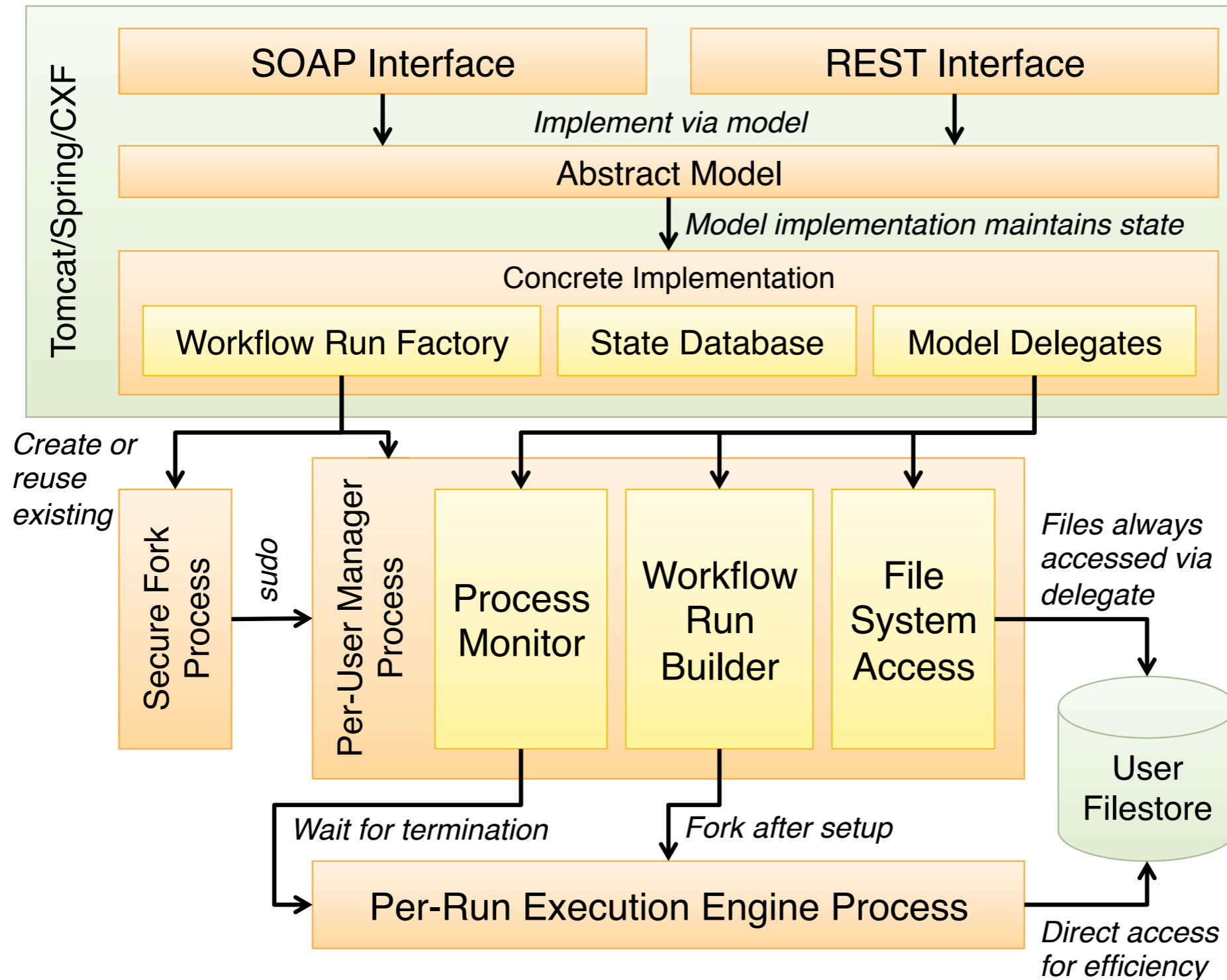
Web page of the Europlanet project

<http://www.europlanet-eu.org/>

Paper

Anja et al., «Workflows for Heliophysics»,  
*Journal of Grid Computing*, Submitted

# Serveur Taverna (3/3)



## HELIO Taverna server:

Provides remote execution of Taverna workflows

# Workflow's Legend

*green:*

SOAP operator

*purple:*

XML splitter — decomposes complex SOAP types into their components

*brown:*

local beanshell[22] scripts — user written scripts to provide custom functionality

*violet:*

local operator — predefined functions within Taverna

*pink:*

nested workflows — workflows re-used inside another workflow

*blue:*

string constant — a string or text which does not change

Input and output ports are of a different *blue* colour and separately labelled.