



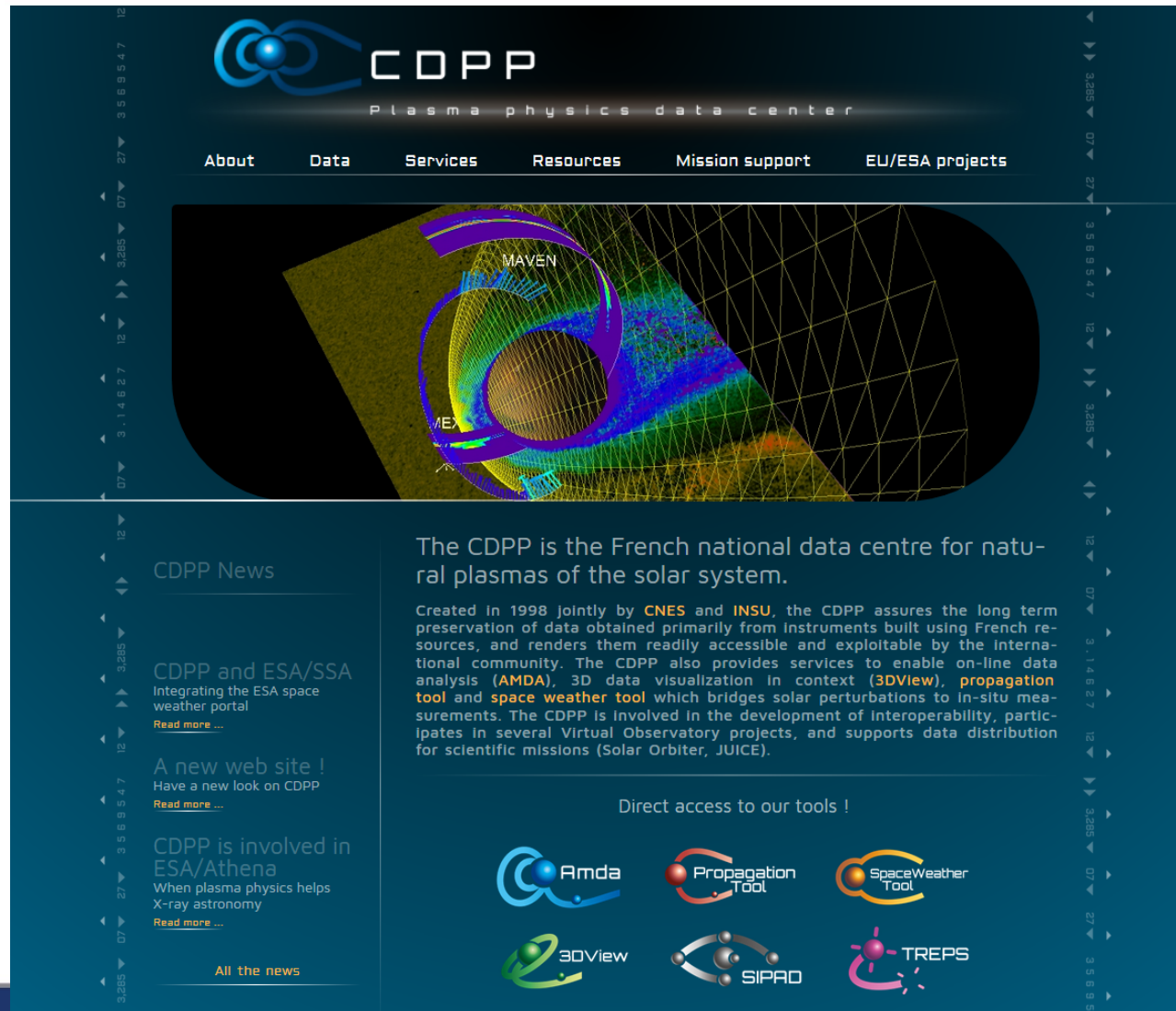
Les 20 ans de PIN

Des données éternelles ?

20 ans d'expérience du centre de données CDPP au CNES

Danièle Boucon (daniele.boucon@cnes.fr), Nicolas Dufourg, Dominique Heulet





The screenshot shows the CDDP website interface. At the top, the CDDP logo is displayed with the text "Plasma physics data center". Below the logo is a navigation menu with links: About, Data, Services, Resources, Mission support, and EU/ESA projects. The main content area features a large 3D visualization of a plasma structure, labeled "MAVEN". Below this, there is a section titled "CDPP News" with three news items: "CDPP and ESA/SSA Integrating the ESA space weather portal", "A new web site ! Have a new look on CDDP", and "CDPP is involved in ESA/Athena When plasma physics helps X-ray astronomy". Each news item has a "Read more ..." link. To the right of the news section, there is a paragraph describing the CDDP: "The CDDP is the French national data centre for natural plasmas of the solar system. Created In 1998 jointly by CNES and INSU, the CDDP assures the long term preservation of data obtained primarily from instruments built using French resources, and renders them readily accessible and exploitable by the international community. The CDDP also provides services to enable on-line data analysis (AMDA), 3D data visualization in context (3DView), propagation tool and space weather tool which bridges solar perturbations to in-situ measurements. The CDDP is involved in the development of interoperability, participates in several Virtual Observatory projects, and supports data distribution for scientific missions (Solar Orbiter, JUICE)." Below this paragraph is a section titled "Direct access to our tools !" which displays six logos: Rmda, Propagation Tool, SpaceWeather Tool, 3DView, SIPAD, and TREPS.



Les évolutions du CDPP

Les évolutions contractuelles, avec de nouveaux partenaires



D'une convention inter-organismes CNES – INSU (CESR) à une convention étendue à l'Université Paul Sabatier Toulouse III et Observatoire de Paris





Années 70-80
"Bibliothèque"



Années 90
STAF
Robot Storagetek



Années 2000
STAF
Robot Storagetek
(NG)

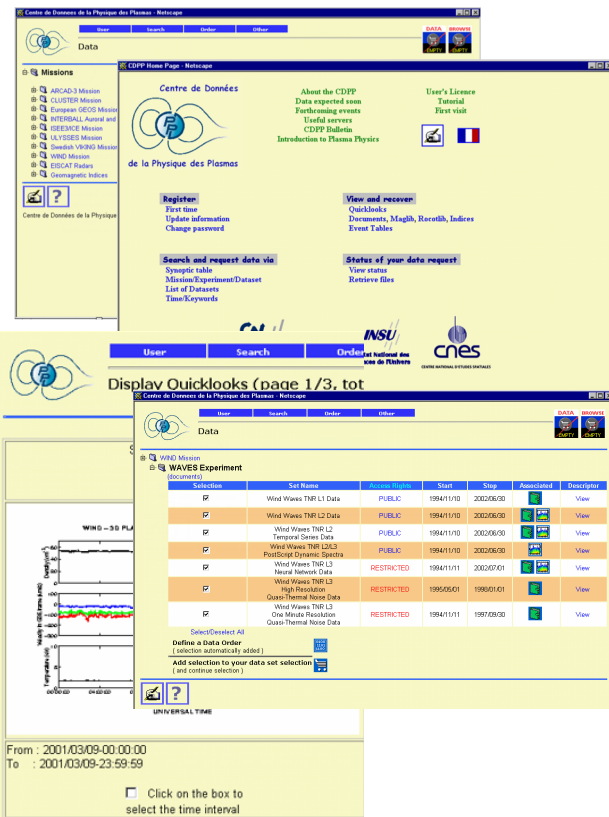


Année 2023
Data Lake



Les évolutions du COPP

Les évolutions technologiques côté serveur archive



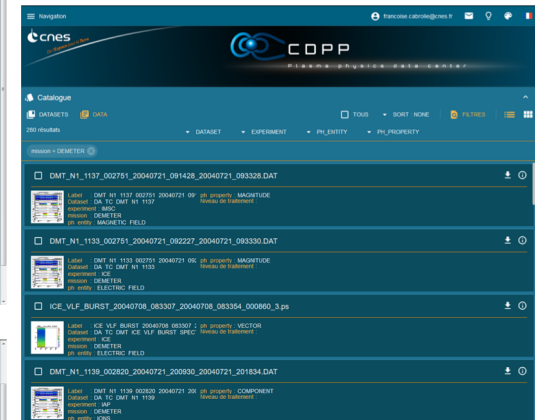
SIPAD (old generation)
Solaris/Métaphase
~1998



SIPAD-NG
Solaris/Oracle
2006



SIPAD-NG
Linux/Oracle
2015



(proto)



REGARDS
Linux/PostgreSQL
+ElasticSearch
2022





Les évolutions du CDPP

Les évolutions de formats côté archive



"Historiquement" des **formats natifs, différents en fonction des missions**

- Non standards
- Nécessité de proposer des lecteurs pour ces données
- Données non directement utilisables par les outils (IDL, matlab ...)

Une initiative de standardisation a abouti à la définition d'un format commun "le format CDPP"

Des standards émergeant dans la thématique, toutes les missions récentes sont par défaut proposées à ces standards

- CDF
- ISTP/IACG

L'avantage de ces **standards**, outre le fait qu'ils sont **largement utilisés**, est qu'ils sont **outillés**

- L'archive a adopté ce standard et propose aux fournisseurs de l'adopter ou réalise la conversion



Les évolutions du CDPP

Quelques chiffres côté archive



Volume de l'archive 169 Toctets

22 missions

Données :

1142 jeux de données

4 6 400 941 objets de données

Quicklooks :

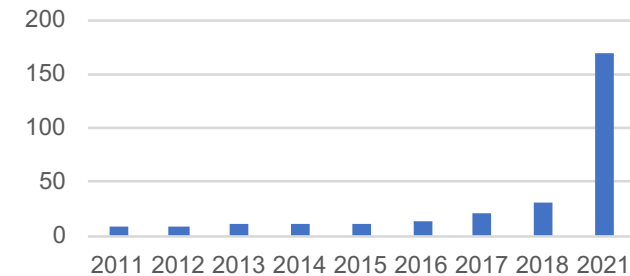
117 jeux de quicklooks

800 004 objets quicklooks

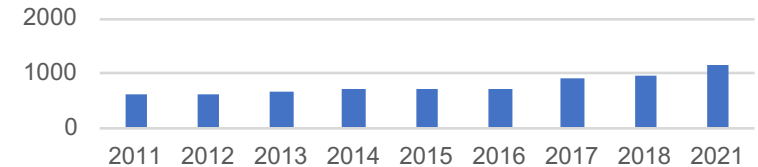
Documents :

548 documents

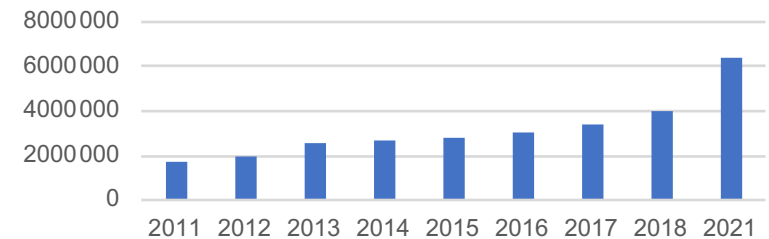
Volume data (To)



Nombre de jeux de données



Objets de données

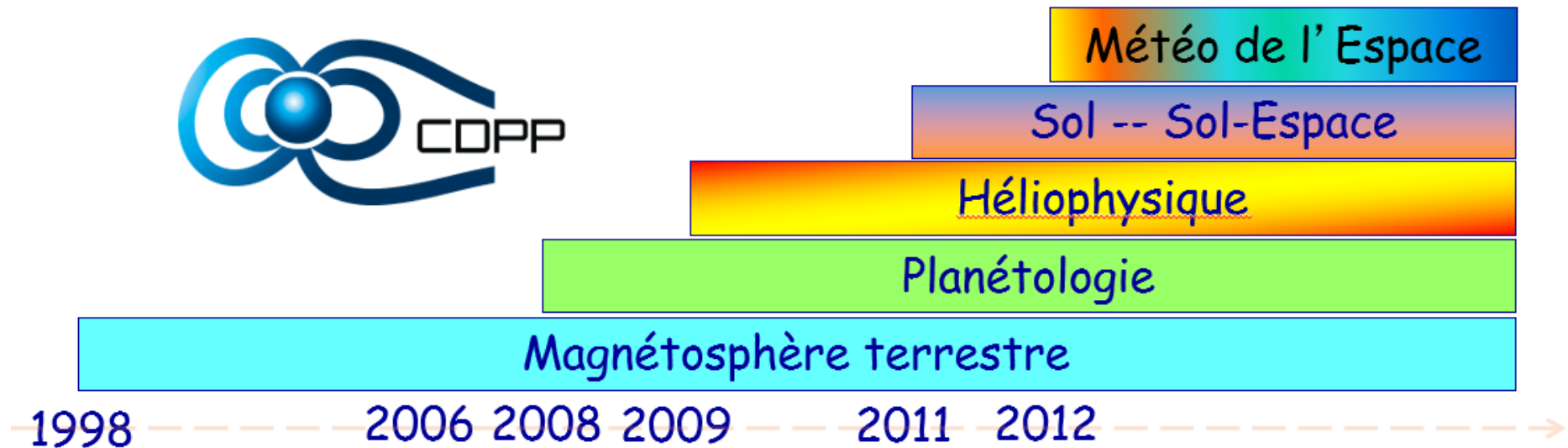






Les évolutions du CDPP

Les évolutions thématiques



outils





Les évolutions du CDPP

Les évolutions de périmètre



De la mise en place d'une archive pérenne accompagnée de SVA techniques vers des activités de valorisation des données et des outils d'analyse scientifique

"Les données ne sont rien sans les outils"



Mais aussi des efforts importants de revalorisation des données anciennes en parallèle des missions en cours

"Les outils ne sont rien sans données"



Les évolutions du CDPP

Les évolutions de périmètre



Nombreux **partenaires** et nombreuses **collaborations** (synergie avec le pôle français de physique solaire MEDOC)

Implication dans les missions (support à la production/validation des données, diffusion, pérennisation, caractérisation des environnements, etc.)

Participation aux **projets européens**

Participation aux groupes de **normalisation**

- Le CDPP est fournisseur de services à la communauté et aux missions spatiales
- La pérennité de son organisation et de son financement permet de fournir ces services sur le long terme



Et demain ? De nouveaux challenges



Des données « FAIR »

- Favoriser la "découverte" des données
- Ouvrir l'archive aux outils
- Élargir la thématique, corrélér des ensembles de données

En assurant la pérennisation, avec des volumes de données croissants, en étant à l'écoute de la communauté des utilisateurs

S'adapter à un monde en perpétuelle évolution ? Maintenir et faire évoluer une archive (budget) ? Standards accès simple, rapide, valorisé ? Acteurs européens et internationaux ?

- Veille
- Rationaliser exploitation, interfaces
- Standards (dont format), protocoles d'accès
- Echanges acteurs européens et internationaux, rôle des agences, accès public aux données

