

CoreTrustSeal (CTS) et IDOC

Contexte

Objectifs

Retour d'expérience

IDOC Missions et engagements

- En préparation :
Euclid, Plato,
Juice, Bepi-Colombo,
Jovial, Exomars..
- En cours :
JWST, Mars-express,
SoHo, Stereo, SDO,
Solar Orbiter
- Mise à disposition : Rosetta,
CoRot, Planck, Herschel,
Trace, Coronas, Picard, Iras



19 ANOs portés par IDOC au sein de l'OSU Paris Saclay

Pourquoi CoreTrustSeal, le discours officiel

- La confiance est au cœur de la Science Ouverte
- Les entrepôts sont une source majeure de données
- Il leur faut avoir la confiance
 - de leurs communautés scientifiques
 - Les personnes qui produisent et déposent les données
 - Les utilisateurs des données
 - de leurs autorités et de leurs financeurs

Les aspects positifs d'une approche vers CoreTrustSeal

- Critères établis par des personnes compétentes et applicables quel que soit le cadre disciplinaire
- Au préalable, auto-évaluation selon les critères, qui permet de vérifier l'organisation et les process et d'identifier des améliorations possibles
- Evaluation externe par des personnes compétentes
- Le dépôt dans un centre de données certifié est un point important dans les Plans de Gestion des Données (PGD)
- Stratégie nationale : Plan National pour la Science Ouverte



Pourquoi CoreTrustSeal, le vécu

- L'approche des 16 questions force à repenser les pratiques et les procédures
- La couverture des différents aspects est complétée au sein de son environnement
- Cela permet d'explicitier et de formaliser des habitudes "évidentes"
- Cela rassure (je n'ai pas de "trou" quelque part)
- Cela permet de disposer d'un corpus documentaire cohérent qui sert ensuite en toutes occasions (Réponses à appels d'offre, bilans, écriture de DMPs,...) et qui est rentable.
- Cela procure de la respectabilité

Ce à quoi il faut se préparer

- Ecrire la majorité des documents en anglais
- Digérer le vocabulaire (OAIS, RDA,...) pour bien utiliser les termes appropriés (curation, provision,..)
- Participer à RDA-France
- Faire des résumés synthétiques de ce qui semble évident
- Exister dans les lieux nationaux ou internationaux de la donnée
 - Re3data,
 - cat opidor,
 - référencement plateformes CNRS
- DOIs sur les données présentées à la labellisation
- Déterminer le périmètre de la labellisation et l'ancrer dans les documents

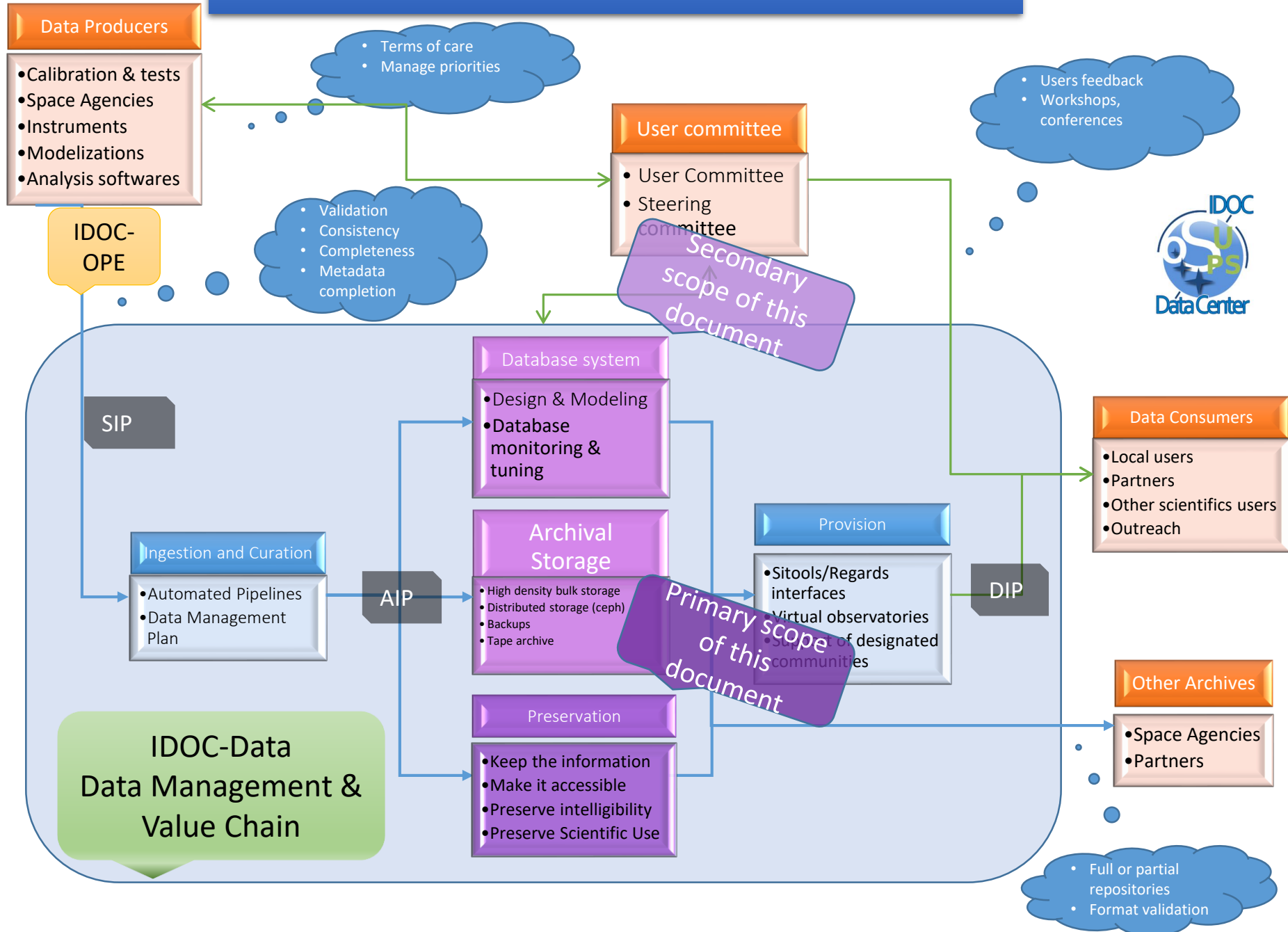
Ce à quoi il faut faire attention

- Bien identifier la portée des questions CTS (lire des réponses existantes dans un domaine/structure proche)
- Questions organisationnelles R01 à R06
- Questions sur la gestion de la donnée R07 à R14
- Questions techniques R15 à R16
- Résister à la tentation de répondre en avance (surtout sur les points "techniques", il y en a finalement que deux !)
 - Par exemple la question R03 donne envie de répondre trop tôt à R07, R09, R15 and R16
- Possible par contre ensuite de renvoyer aux réponses précédentes

Les pièges communs

- Réponses trop courtes : il faut que l'ensemble du sujet soit couvert
- Réponses trop longues : Utiliser des liens vers vos sites officiels pour schémas, documents détaillés,...
- Corollaire : les sites web, portails, doivent prendre en compte la certification et être rendus cohérents dans le contenu et la structure.
- Corollaire : quand on répond, on dessine une organisation, on réalise des schémas, on écrit des documents, on modifie les pages web, puis on écrit la réponse

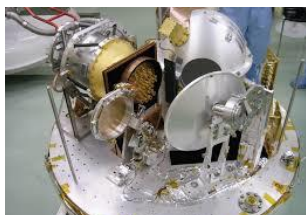
IDOC-DATA and the OAIS model



Les autres pièges communs

- Oublier les aspects humains (compétences, disponibilité, présence à long terme,..)
- Oublier l'environnement des données (codes, portails d'accès, communautés utilisatrices, lieux d'échanges...)
- Oublier de prévoir le plan "vous n'existez plus du tout"

Quels types de projets ?



Avant le lancement

0) Conception, développement contrôle commande et tests de l'instrument

Agences spatiales

1) Instrument Operations

Segments sols

2) Pipelines

Gestion des données

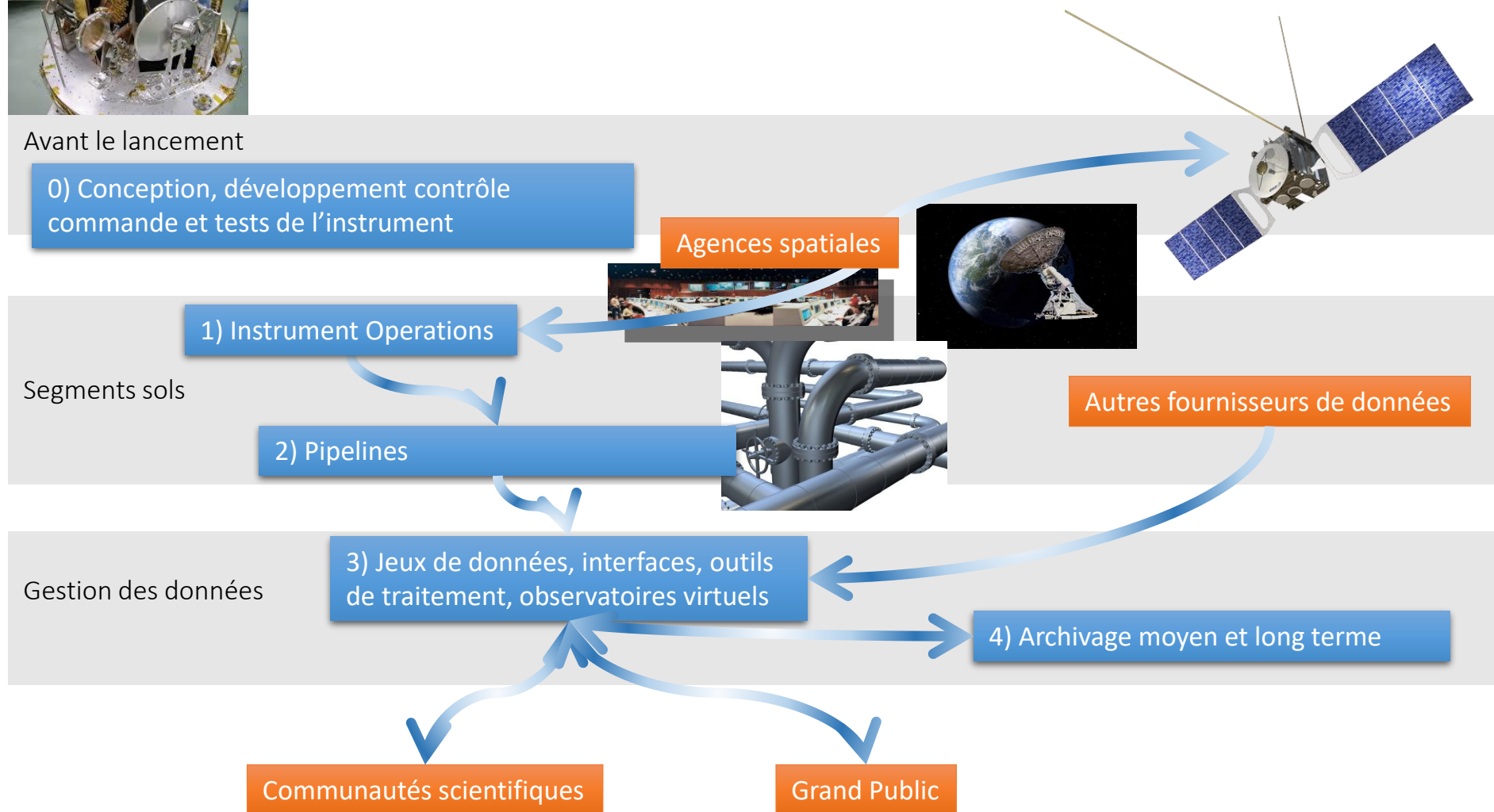
3) Jeux de données, interfaces, outils de traitement, observatoires virtuels

Autres fournisseurs de données

4) Archivage moyen et long terme

Communautés scientifiques

Grand Public



IDOC Steering committee

Technical leader

Scientific leader

IDOC-INSTR

IDOC-OPE

IDOC-DATA

Informations systems,
computing & networks

Training & Communication

Technical fields

Future missions design,
conception, realization & tests

Missions Center

Ground Segments

Data ingestion and curation

Data Provision

Data analysis & modelization

Virtual observatories

Data Preservation

Scientific Themes

MEDOC

AMIS

COSMO

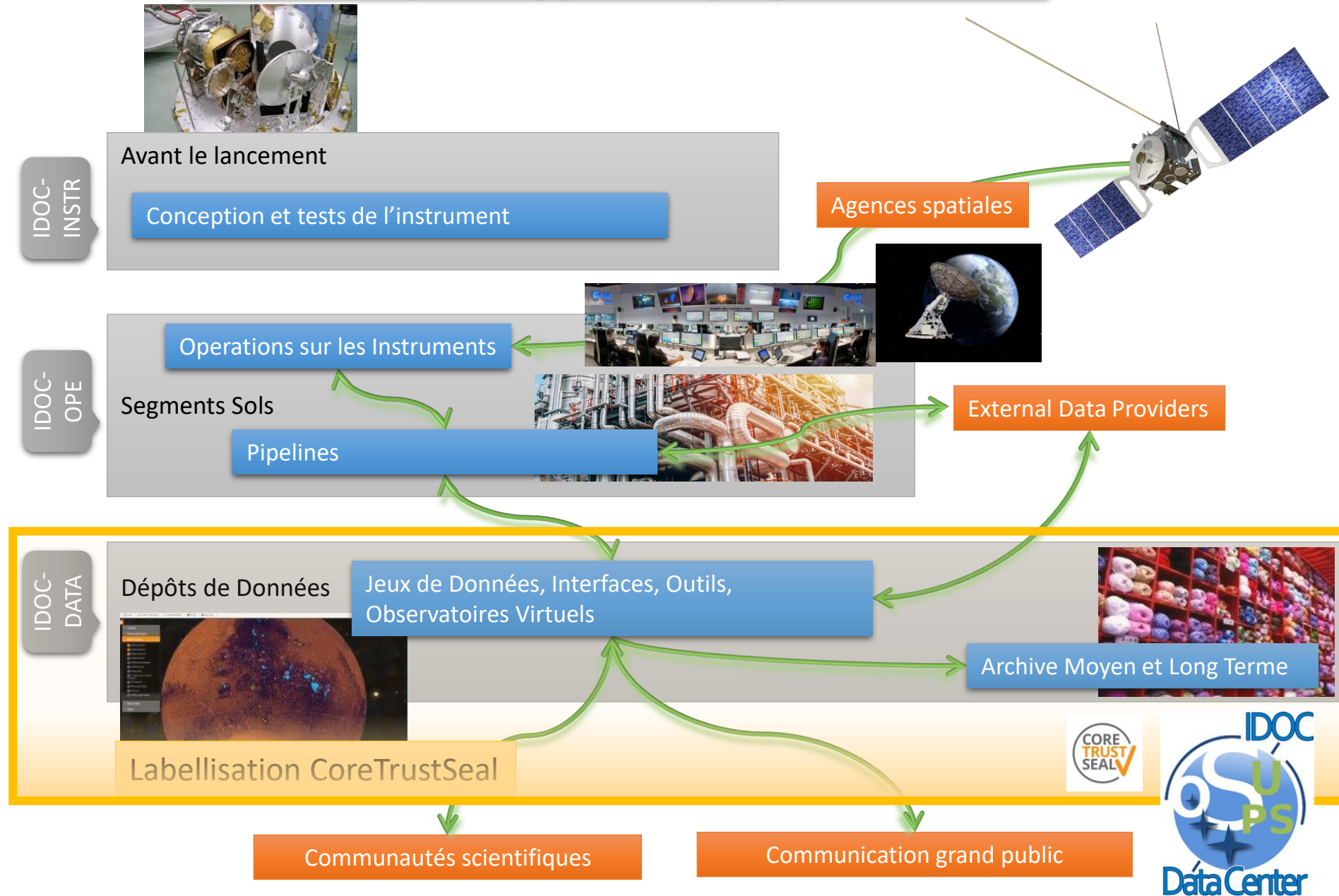
DS2

PSUP

GEOPS

AIM

Quels types de projets



Problématique CNRS

Pas de Mission
explicite (pas de
contrat) pour une
plateforme

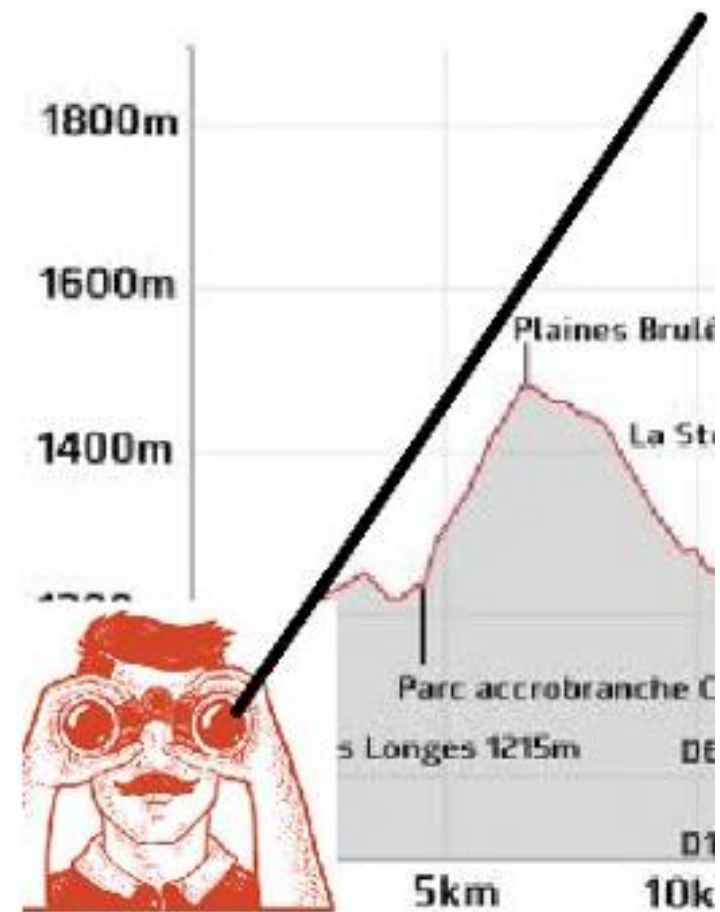
Pas d'engagement
formel sur le long
terme

- 1. Mission/Scope
 - R1. The repository has an **explicit mission to provide access to and preserve data** in its domain.
 - I detected the following stand-alone sentence on the about page of IDOC: "Dataset archive deals with the long term preservation of a dataset hosted at IDOC". But a **clear LTP mission statement is hard to find** for someone who wants to check.
- 3. Continuity of access
 - R3. The repository has a **continuity plan to ensure ongoing access** to and preservation of its holdings
 - Although the described context provides reassurance about the sustainability of IDOC holdings, I understand that **there is no formal written agreement in place with the university and/or CNRS** given any guarantees that they will take over the responsibility in case IDOC no longer exists.

Ces points
durs,
comment les
lever ou faire
avec

- Mission : le site web doit expliciter cette mission
 - Renoncer à expliquer le mille-feuille administratif français -> CNRS/univ./Départements/OSUs/UARs/Labos
- Continuité d'accès : pas grand-chose à faire : IDOC a renoncé à "**Fully implemented**" et est en "**Implementation phase**"
- Lettres de soutien Université et CNRS
 - Commentaire rapporteur "Very good to see the formal letters of support also on the website!"
- J'essaie de convaincre toute personne du CNRS que je rencontre de la nécessité de formaliser au niveau global des réponses communes à ces deux questions (sinon pourquoi pousser à la labellisation)
- Le CNRS doit aider aussi pour répondre à "vous n'existez plus du tout"

Où en êtes
vous ?



GRAND TRAIL DU SANCY

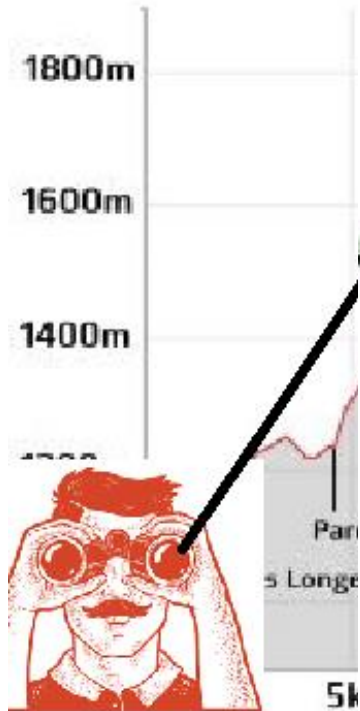
Où en êtes
vous
vraiment ?



GRAND TRAIL DU SANCY 60 km - 3350 mD+ (R) : Ravitaillement (E) : Point d'Eau

Toutes les actions doivent servir à avancer dans le processus

Too busy to improve?
Need more slack in your system!



GRAND TRAIL



Henrik Kniberg

