



Observatoire de Physique du Globe de Clermont-Ferrand

Données, observatoire virtuel et interopérabilité

Emmanuel Delage

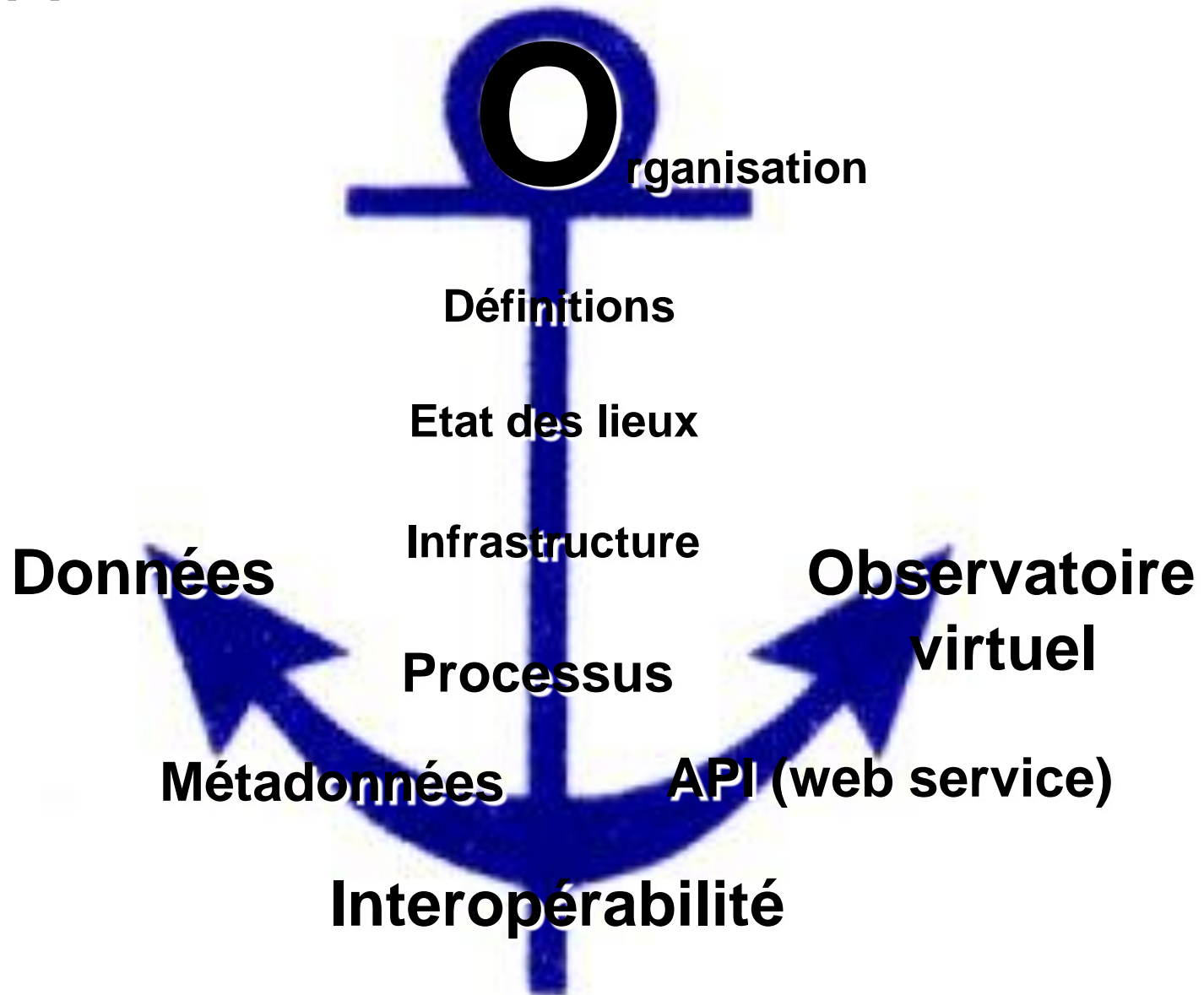
Responsable technique du centre de données

Service de Développement Informatique



29 septembre 2016

PLAN



Organisation du centre de données

La connaissance de « qui fait quoi » est bénéfique

=> Mise en place d'un **Comité de pilotage**

=> Mettre en œuvre de **bonnes pratiques** adoptées par tous:

- Adapter les solutions simples
- Fournir un environnement de développement commun
- Définir des méthodes et outils pour la mise en œuvre des processus
- Uniformiser les bases de données

=> **Centralisation** ↗ traçabilité ↗ sécurité ↗ qualité => **pérennisation**

- Directeur, **Responsable scientifique** : Patrick Bachèlery
- **Responsable technique** : Emmanuel Delage
- Responsable SDI, **ASR et infrastructure** : Philippe Cacault
- Développeur web et bases de données (**Remplacement d'un départ janvier 2016 prévu en septembre 2017**)

Non permanents :

- Ingénieur développement web et BD (**peut-être avant le remplacement**)
- Romain Martin, **Ingénieur stagiaire** interopérabilité, observatoire virtuel (**départ fin septembre 2016**)
- Yannick Guéhenneux, **PostDoc**. Interopérabilité EPOS (**départ fin mars 2017**)
- José Nicolas, **PostDoc**. Interopérabilité ACTRIS (20%) (**départ fin mars 2017**)

Définitions

« un **Observatoire virtuel** (OV) est ainsi un cadre organisationnel incontournable des bases de données et des services, impliquant l'usage de standards de description des données et de protocoles d'accès pour permettre de relever ce défi de la gestion de grandes masses de données, de leur analyse et de leur diffusion. » **Obspm.fr**

« Cette intégration dans un cadre standardisé permet le dialogue entre systèmes c'est-à-dire **l'interopérabilité**. Il devient alors possible de croiser des données d'origines différentes, de façon transparente pour l'utilisateur, et d'utiliser des outils génériques permettant d'effectuer les mêmes opérations sur des jeux de données d'origines variées. Cette activité a ouvert ainsi de nouveaux horizons à la recherche astronomique. » **Obspm**

L'interopérabilité web des données d'observation implique :

- 1) Métadonnées : description des données spatio-temporelles et lien vers les données
- 2) API (Service web) : communication et échange de données entre applications
- 3) Observatoire virtuel : Plateforme de gestion des données scientifiques

Etat des lieux (fin 2015)



Terre solide



Observatoire de Physique du Globe de Clermont-Ferrand



Atmosphère

- **Pôle de télédétection volcanologique SNO**
6 Services d'observation:
Piloté par l'IPGP et l'OPGC

- **SISMO**

- **Instrument national :
Presse multi-enclumes**

- **Hydrogéologie**

- **Site instrumenté COPDD**
6 services d'observation:
 - 1 SNO piloté à l'OPGC
 - 2 SNO
 - 3 SO

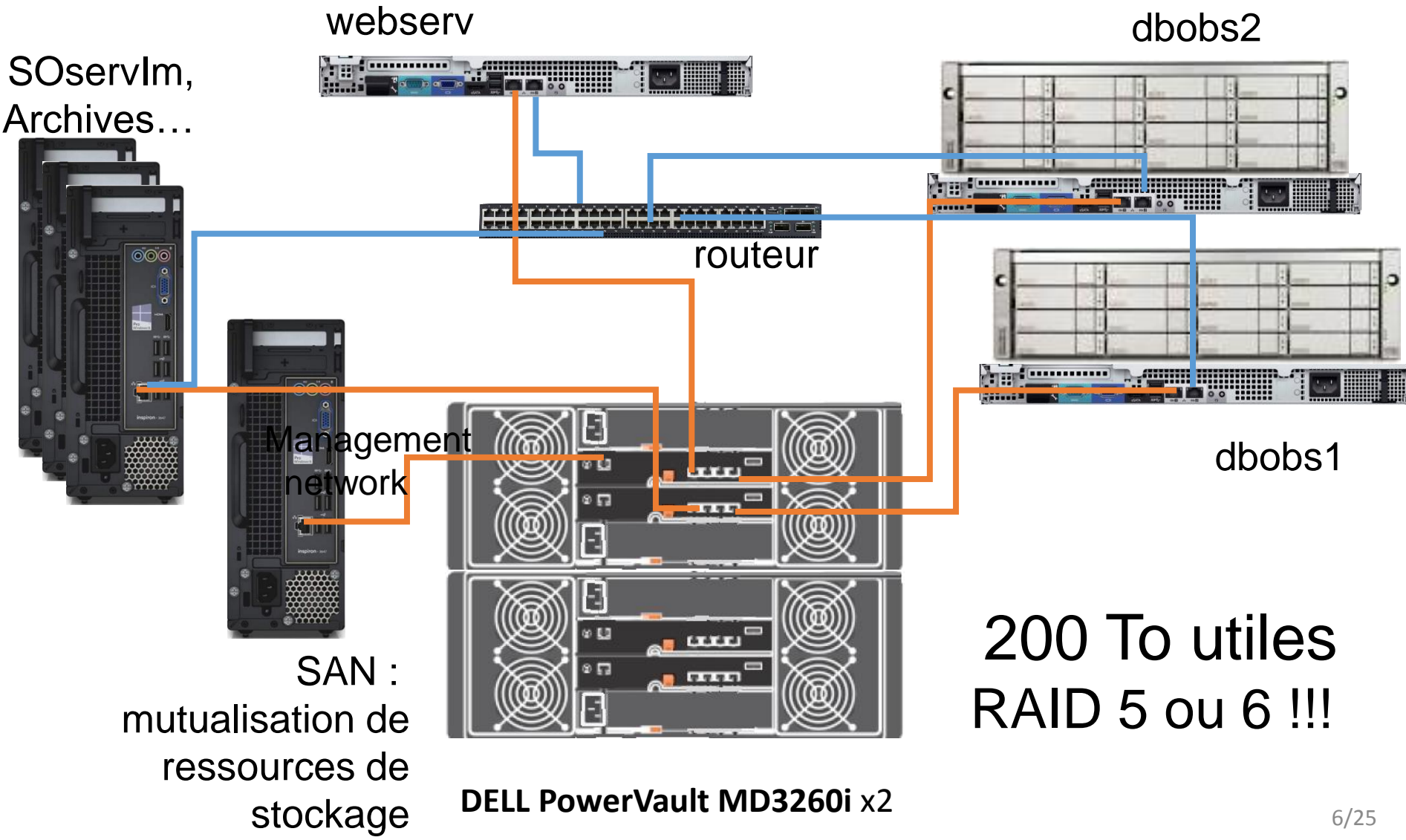
- **Instrument national :
Plateforme de Mesure
Aéroportée**

15 SO services d'observation

- 1) Bases de données (SQL, Texte, images)
- 2) Métadonnées (pas d'implémentation)
- 3) API web service (sans)
- 4) Observatoire virtuel (sans)

« Tous ces éléments sont des Indicateurs de la mise en œuvre de l'interopérabilité »

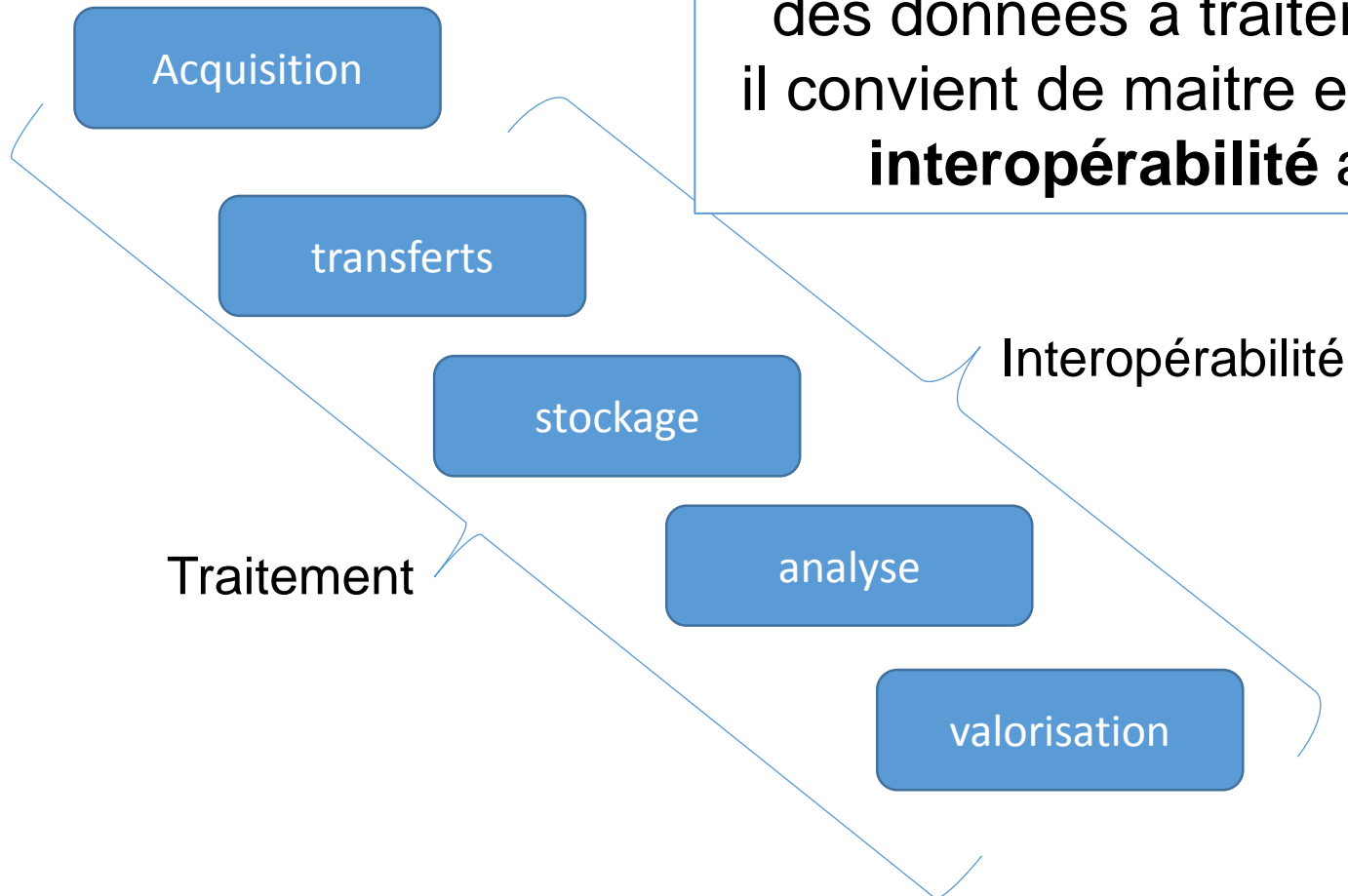
Infrastructure opérationnelle



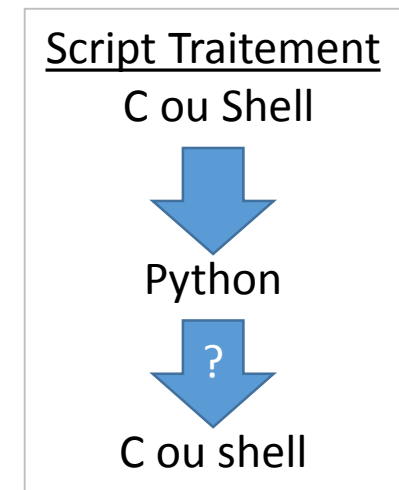
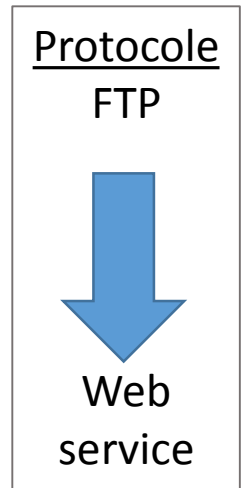
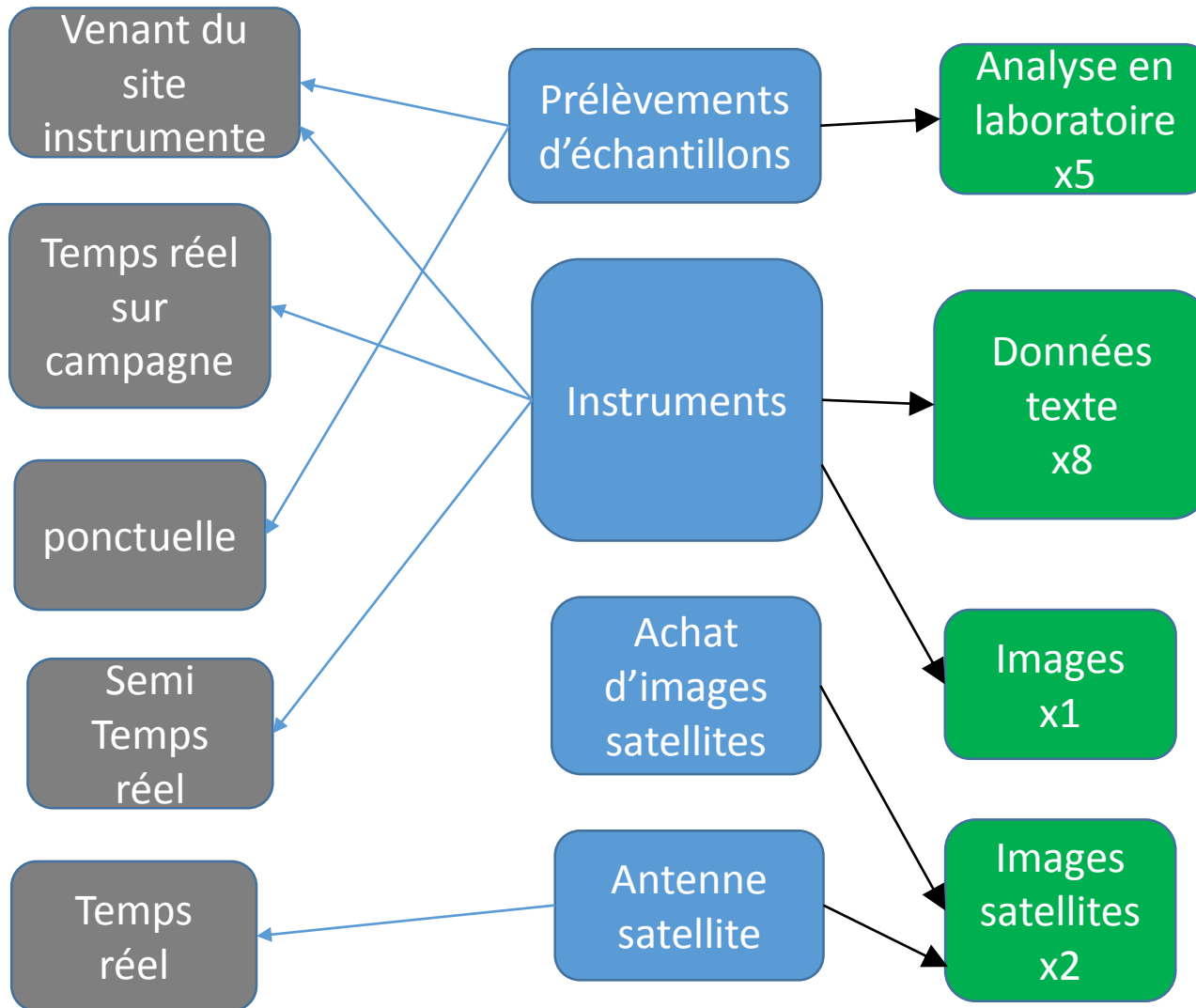
200 To utiles
RAID 5 ou 6 !!!

Processus

Les processus du **traitement** de l'information deviennent aujourd'hui de plus en plus complexes. En effet, devant l'accroissement considérable des données à traiter (**BigData**), il convient de mettre en œuvre une **interopérabilité** adaptée.

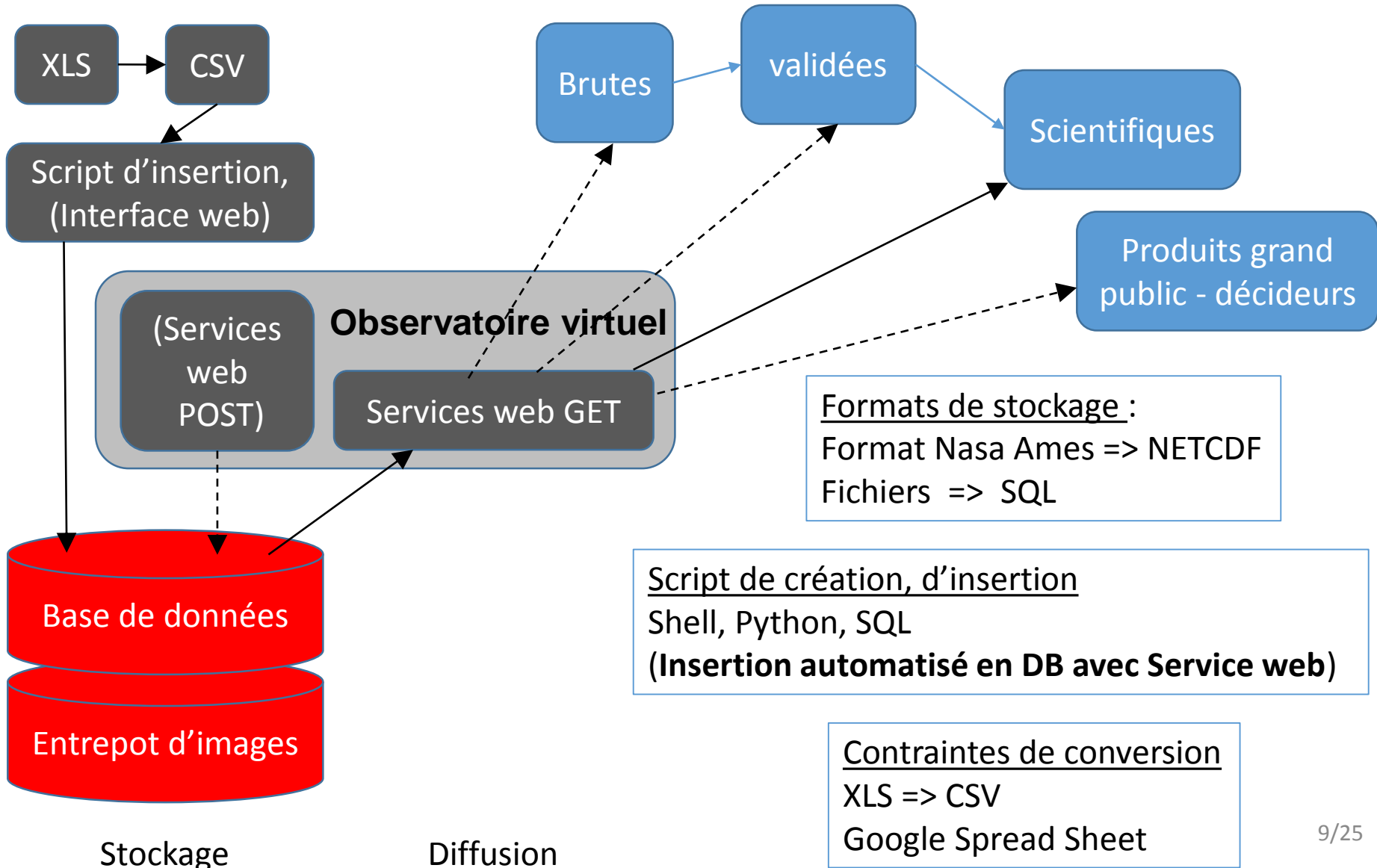


Acquisition, transfert

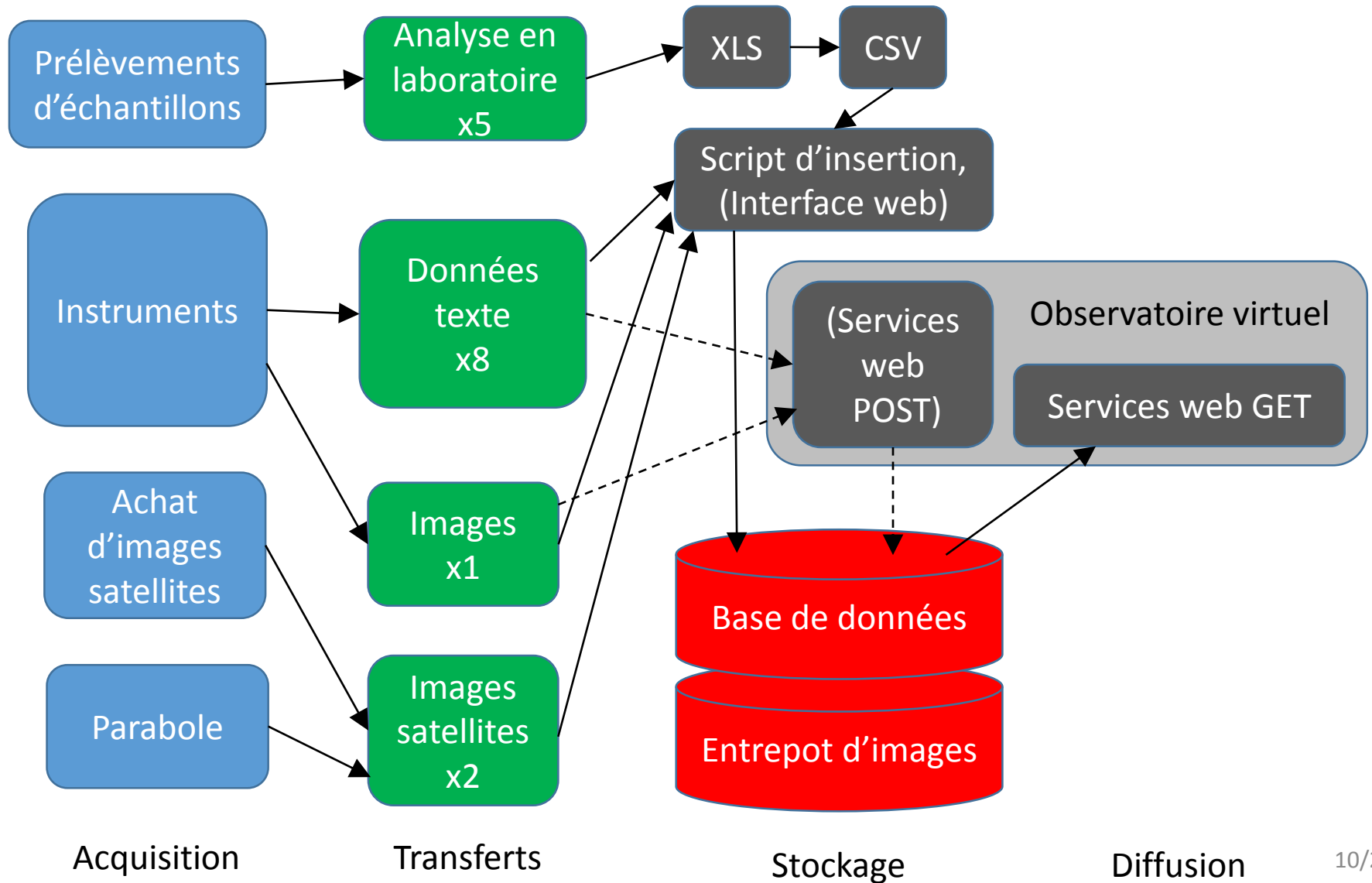


Matlab...

Stockage, diffusion



Flux de données : de l'acquisition à la valorisation



Bases de données

Choix de PostgreSQL

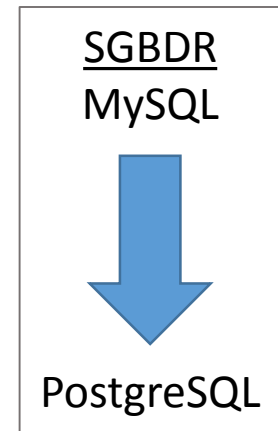
- Préconisé par le socle interministériel des logiciels libres
- **PostGIS**
- **JSON**, BSON(JSONB) sont un type pris en charge depuis 9.4
- Gère bien les grands volumes de données (MongoDB en NoSQL est orienté document BSON)
- Utilisé de plus en plus (<http://db-engines.com/en/ranking>) et dans notre communauté
- MySQL racheté par Oracle (ou utiliser MariaDB)

Réorganisation des bases de données (pas de relation dans les anciennes DB)

⇒ Modification des bases => **modification des IHM**

Rq: **L'observatoire virtuel** est une opportunité pour opérer une **nouvelle organisation** des données

Bases de données => métadonnées => web services (API)



metadata	
id	int
name	varchar(150)
Description	varchar(200)
authorization_	smallint
date_b	date
time_b	time
date_e	date
time_e	time
latitude	float
longitude	float
elevation	integer
epsg	integer
acquisition	smallint
format	smallint
+	
Champs spécifiques...	11/25

Métadonnées

Format : **JSON** ou XML?

JSON ↔ XML

Standard :

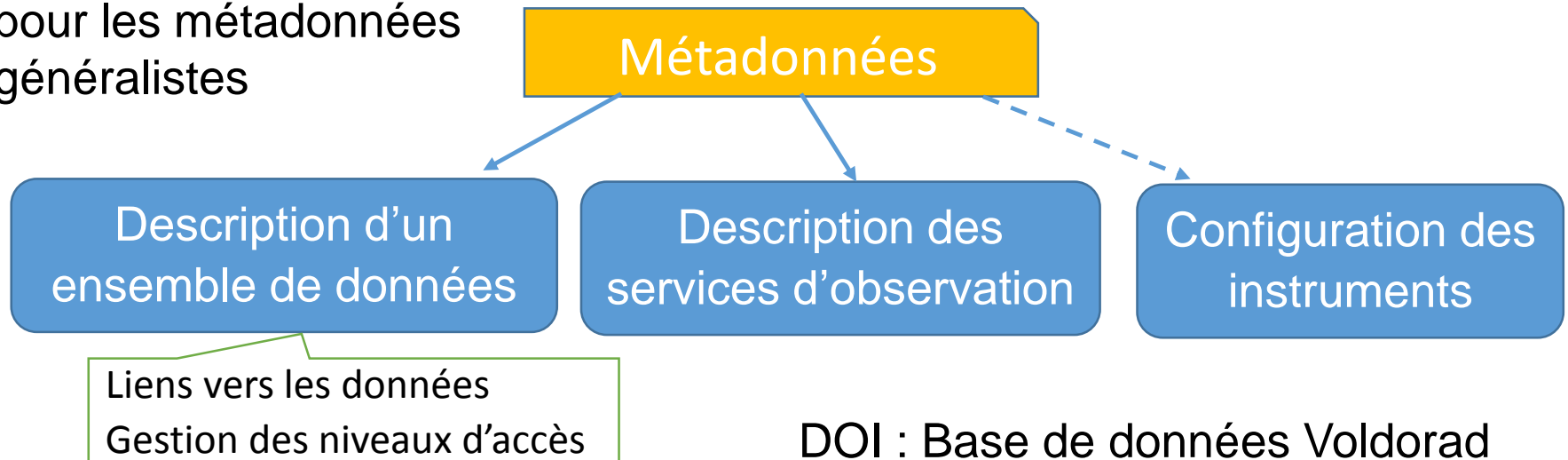
NETCDF, Opensearch pour l'écriture des recherches, WPS standard de l'OGS pour les données géo-spatiales

Normes:

ISO 19115 schéma requis pour décrire de l'information géographique,
ISO 19139, DublinCore pour les métadonnées généralistes

Granularité :

BIOCAP => évènement (plusieurs prélèvements),
DYNVOLC => échantillon,
AUVERWATCH => point de captage,
OI2 => Image ...



DOI : Base de données Voldorad

Bases de données => **métadonnées** => web services (API)

API (Web service)

REST (Architecture retenue)

- **Simple** à mettre en œuvre,
- Utilise l'URI comme représentation de ses ressources (flexibilité)
- Uniquement HTTP

SOAP

- Client et serveur étroitement liés (états) => problème d'interopérabilité
- XML => lourd
- Bénéficie d'un meilleur soutien des normes
- Moins de code pour la sécurisation

API REST en PHP : **Slim** (Tonic, Zend Framework...)

GET : récupérer les données d'une ressource

POST : envoyer des données dans une requête

(pas de PUT ou DELETE)

Exemple de GET : <http://OPGC.fr/vobs/rest2/req.php/dynvolc/data/id/12>

Sécurité:

HTTPS,

login/**mot de passe** (encode, hash, salt),

IP **white listées**

Envoie d'images :

POST (Content-Type : image/png),

FTP

Rq: [Plupload: Multi-runtime File-Uploader](#)

décompose les fichiers lourds et envois sur HTTP

Bases de données => métadonnées => **web services (API)**



Site instrumenté COPDD :

Station de recherche atmosphérique
Cézeaux - Opme - Puy de dôme

Météo

Radar X



COPDD Quicklooks

Pôle de Télédétection des
Volcans

Activité volcanique en direct

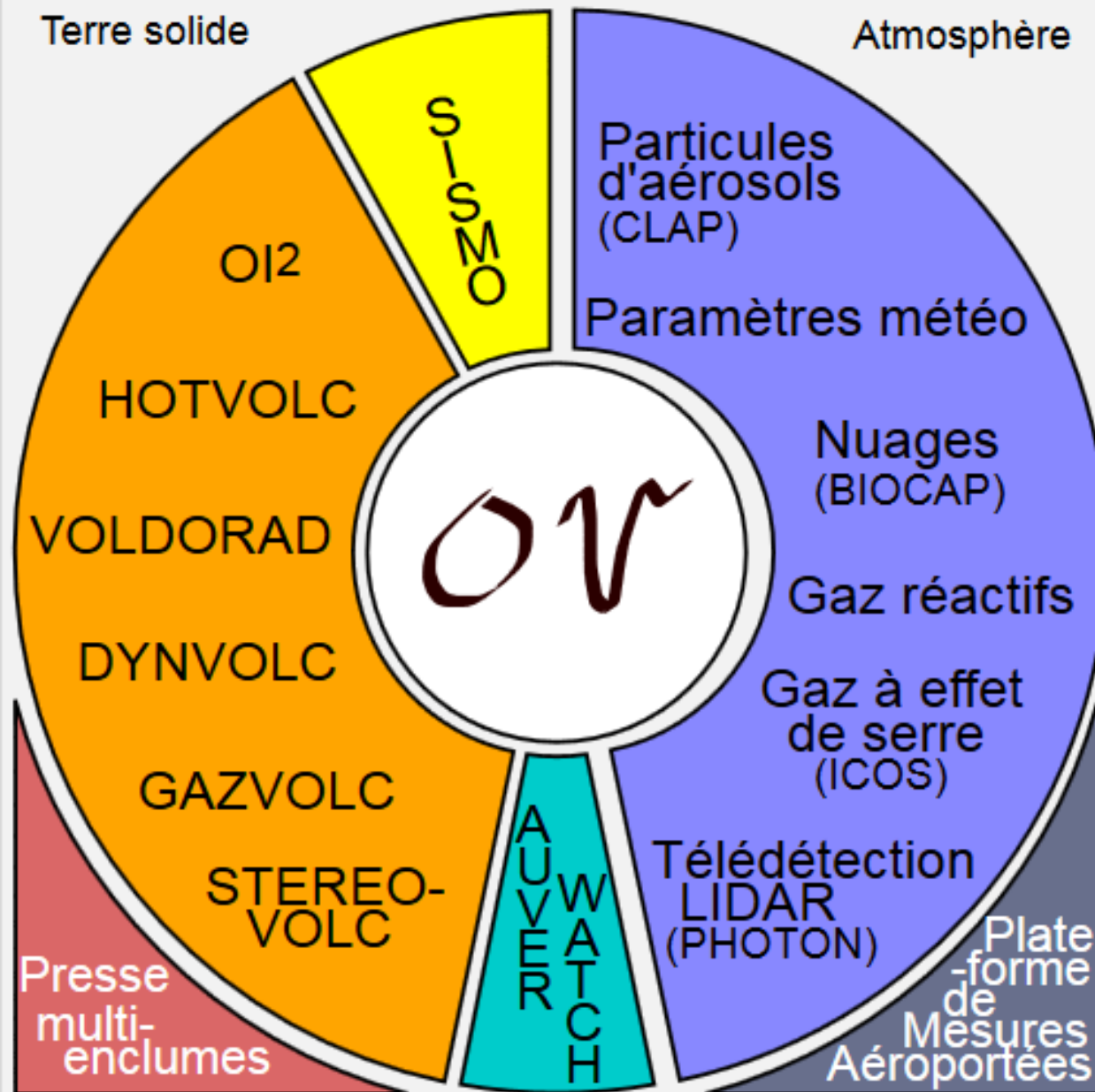


Instruments nationaux :

- Plateforme de Mesures Aéroportées
- Presse multi-enclumes

Terre solide

Atmosphère





Particules d'aérosols

(CLAP)

Particules d'aérosols (CLAP)



Gaz réactifs

Gaz réactifs



Nuages (BIOCAP)

Nuages (BIOCAP)



Gaz à effet de serre

(ICOS)

Gaz à effet de serre (ICOS)



Téledétection LIDAR

(PHOTON)

Téledétection LIDAR (PHOTON)



SISMO

SISMO



AUVERWATCH

AUVERWATCH



EMG

EMG



DYNVOLC



GAZVOLC



VOLDORAD



OI2

DYNVOLC

Observation system within France's National Service for Volcanological Observation (SNOV). The system contributes complete, integrated, data sets for eruptive products from : field studies (i.e., field mapping, outcrop and sample descriptions), textural analysis (i.e., particle size and size distribution, morphology, density, porosity, connectivity, permeability and vesicle and crystal size, distribution, shape and orientation), and chemical data (i.e., bulk and glass chemistry).



Données :

Acquisition : real-time soil
niveau de données : raw
format : .csv
Volume annuel: 42
Volume total: 42

[Accéder aux données](#)

Responsable(s) scientifique(s) : Lucia Gurioli (Physicien-adjoint OPGC)

Contributeur(s) scientifique(s): Ivan Vlastelic (CR1-CNRS, LMV), Patrick Bachelery (Prof, UBP-UFRST, LMV), Andrea Di Muro (PhysAdj-OVPPF, IPG), Andrew Harris (Prof, UBP-UFRST, LMV), Jean Luc Le Pennec (DR-IRD, LMV).

Instruments utilisés :

AUCUN

Logiciels utilisés :

AUCUN

Mots-clés :

[VOLCAN](#) [PARTICLES](#) [TEXTURE](#) [MAGMA](#)



Atmosphère

Terre solide

dynvolc X

auverwatch X

blocap X

Localisation :

Lieu

Coordonnées



Dates :

EXC 04/02/15

EXC 12/06/06

EXC 05/04/15

EXC 12/06/06

Type d'acquisition :

-

Format des données :

Tous les formats

Rechercher

ALUVERWATCH : Interface de données



Nombre de résultats:7

Pages: 1

metadata 1:Data

metadata 2:Data

metadata 3:Data

metadata 4:Data

metadata 5:Data

metadata 6:Data

metadata 7:Data

BIOCAP : Interface de données



Nombre de résultats:109

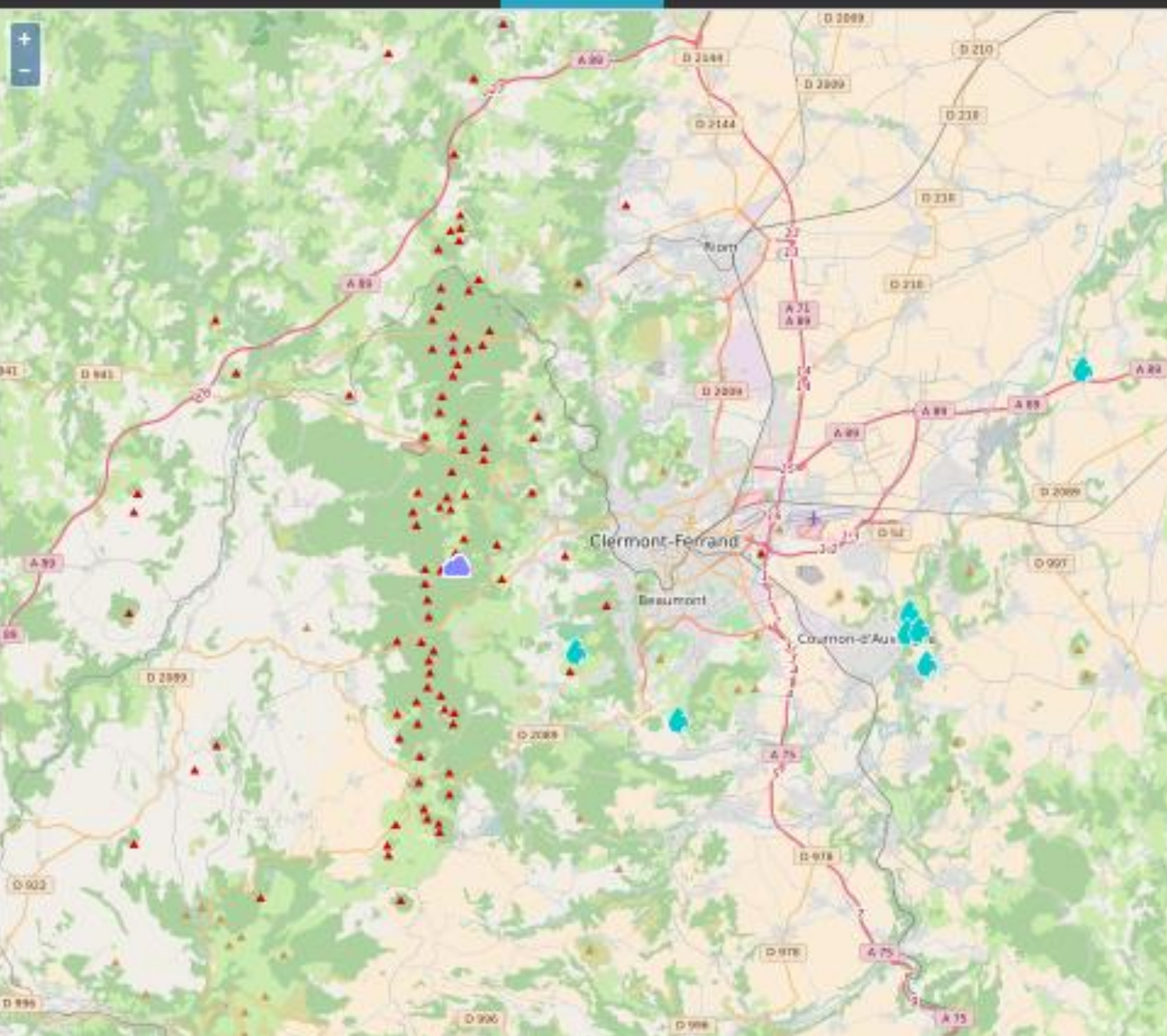
Pages: 1,2,3,4,5,6,7,8,9,10,11

DYNVOLC : Interface de données



Nombre de résultats:31

Pages: 1,2,3,4



Auverwatch

Nom : D13

Localisation : D13

Elevation : 322 m

Position: Non communiquée
Service d'observation destiné à mesurer les paramètres quantitatifs et qualitatifs des masses d'eau caractéristiques en Auvergne afin d'observer l'évolution à long terme des hydrosystèmes en fonction des modifications anthropiques et climatiques



bulkchemistry

Partout, sans intervalle de temps, tout format, dans la table de données bulkchemistry

Rechercher

Tableau : ☯

◀ Page 1/7 ▶

Nombre de données par page: 5

id	sample name	occured	collected	texture aka	note on samples	vertical location	spatial location	latitude	longitude	elevation
1	REU140621_1	2014-06-21	2014-06-21	lava	bad quenching in water at 4h38 UTM	s/s upper crust	null	-21.265917	55.718111	1983
2	REU140624_1	2014-06-21	2014-06-24	aphyric, glassy large scoria	null	null	main vent	-21.250444	55.195556	2362
3	REU140624_3	2014-06-21	2014-06-24	scoria lapilli	hot deposit (405°C)	top of 10 cm thick deposit	main vent	-21.251056	55.719667	2336
5	REU140624_6	2014-06-21	2014-06-24	coarse aphyric scoria	null	null	main vent	-21.251056	55.719667	2336
4	REU140624_4	2014-06-21	2014-06-24	scoria lapilli	hot deposit (405°C)	null	main vent	-21.251056	55.719667	2336

CSV XML JSON

Graphique : ☯

Axe des X :

sample_name

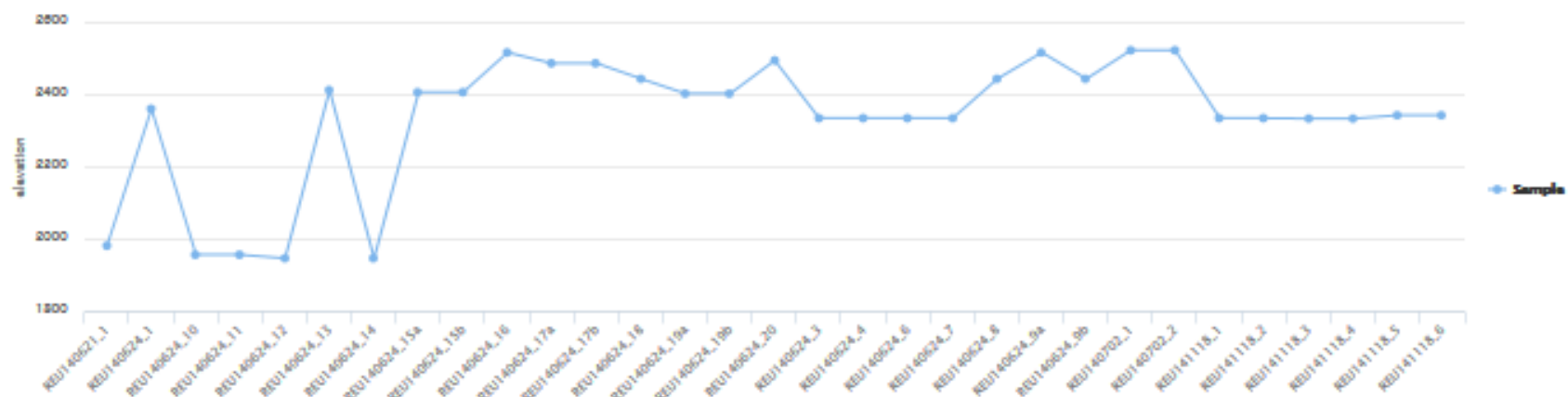
Axe des Y :

elevation

Type de graphiques :

courbe

sample_name



Highcharts.com

Carte : ☯

Etat d'avancement (fin 2016)

L'architecture de l'observatoire virtuel permet d'intégrer tous types de données (SQL, Images) et tous types d'acquisition.

Nombre de SO	Acquisition		Données		Intégrés à l'OV
	Temps Réel	Ponctuel	SQL	Images	
6		×	×		3
7	×		×		
1		×	×	×	
2	×		×	×	
Totaux	16	9	7	4/16 pg	3

Partiellement ou entièrement

La mise en production des données « ponctuelles » est efficace.

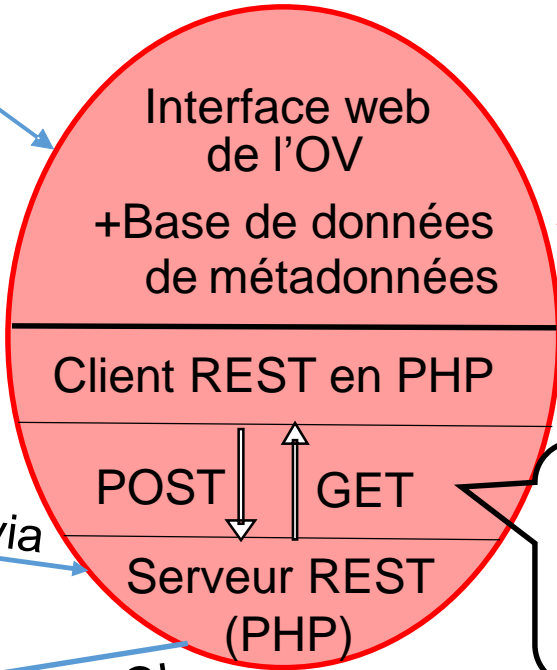
Futurs développements:

- Mise en œuvre **API méthode POST** (pour l'intégration des données temps réel),
- **Intégration** des SO « images » et « temps-réel »,
- Logiciels autour des données,
- Interface web d'administration, interface tableau de bord...

Accès aux données

Navigateur web

Application Internet



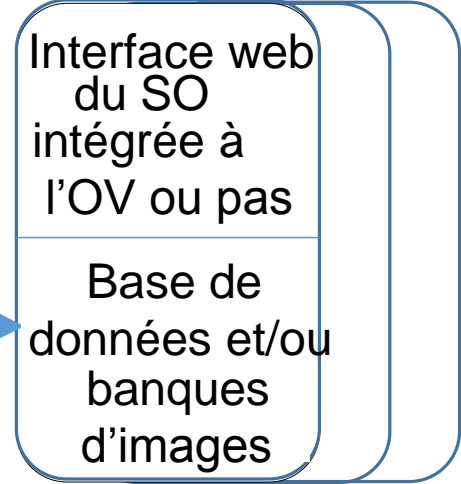
Observatoire virtuel

Web service avec Architecture REST
Format d'échange métadonnée JSON

Accès avant l'OV

Accès direct via une URI

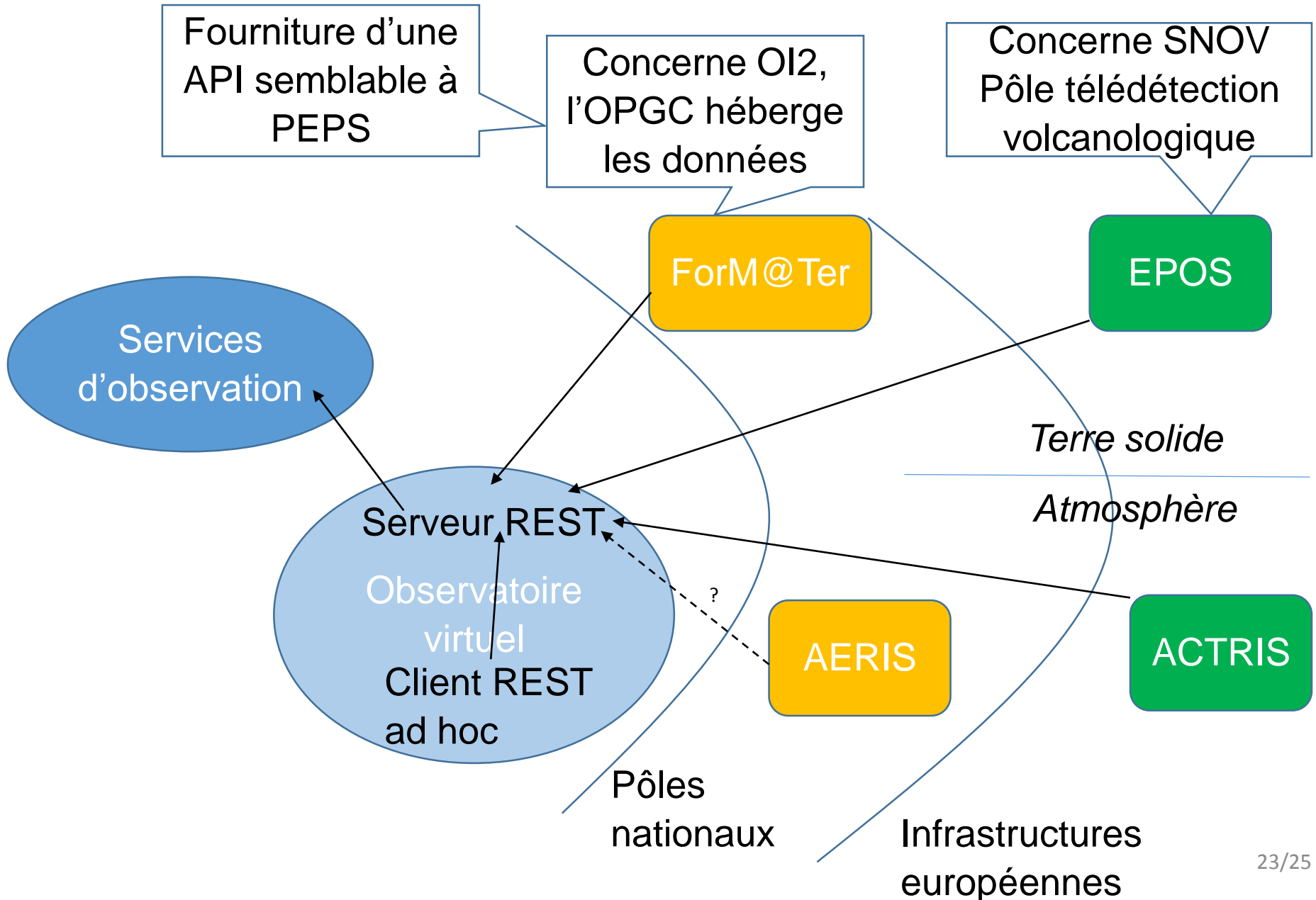
Requête d'accès SQL



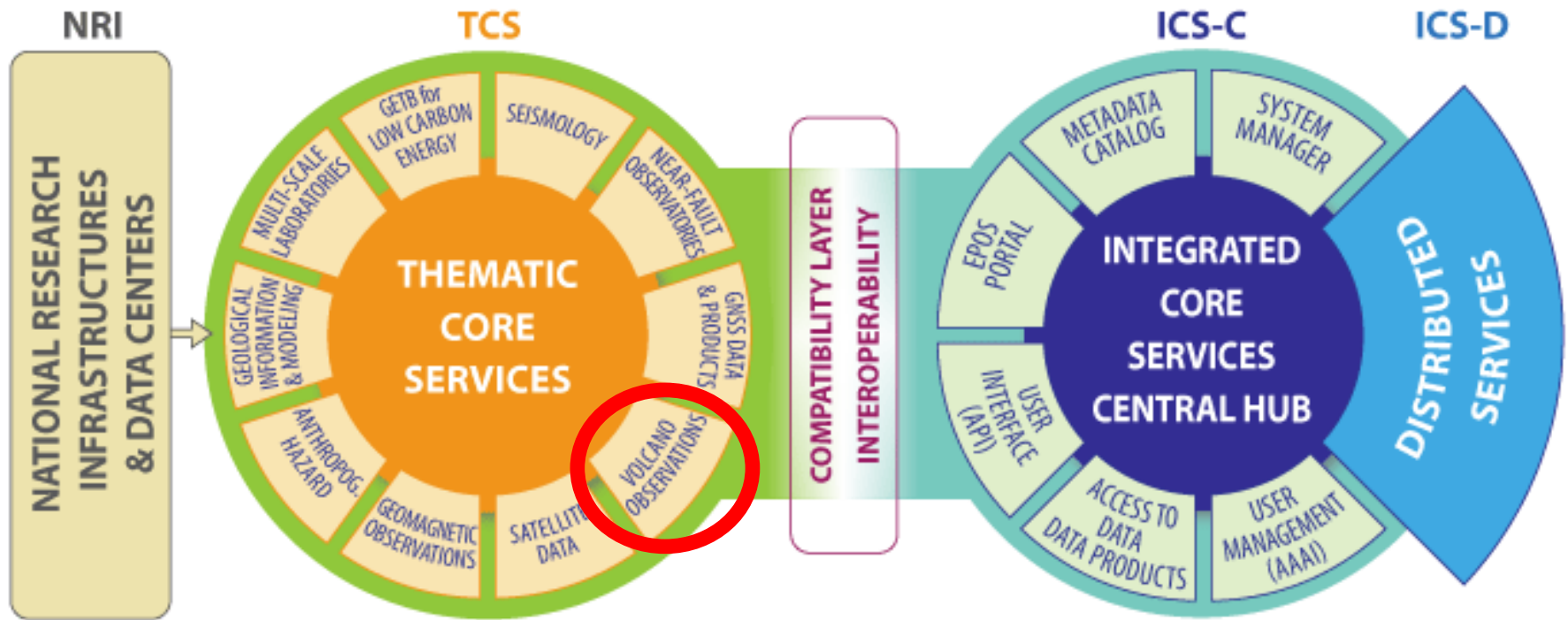
x15

Vers le mise en œuvre de l'interopérabilité

Interopérabilité avec l'extérieur



Interopérabilité EPOS



+



⇔



Bases de données => métadonnées => web services (API)

Interopérabilité ForM@Ter

Une **solution** technique mise en œuvre par le **CNES** pourraient être **déployée** à l'**OPGC**. Cette solution permettrait à ForM@Ter d'interroger les catalogues de l'OPGC et du CNES et d'afficher les résultats depuis la même interface

L'OPGC et le CNES
hébergent leurs données

ForM@Ter



Observatoire de Physique du Globe de Clermont-Ferrand



Interopérabilité AERIS et ACTRIS

AERIS est constitué autour de 4 centres de données:
ICARE, ESPRI, SEDOO, SATMOS => hébergement des données



EBAS est une base atmospheric chemical composition and physical properties
=> Export data in specific formats, calculation of statistics

L'OPGC a des
Problématiques
Communes avec

