



I-SITE CLERMONT

Clermont Auvergne Project

CEBA

Cloud Environnemental au Bénéfice de l'Agriculture

Francis Ogereau, Jérémy Mezhoud

CEBA - Qu'est-ce que c'est ?

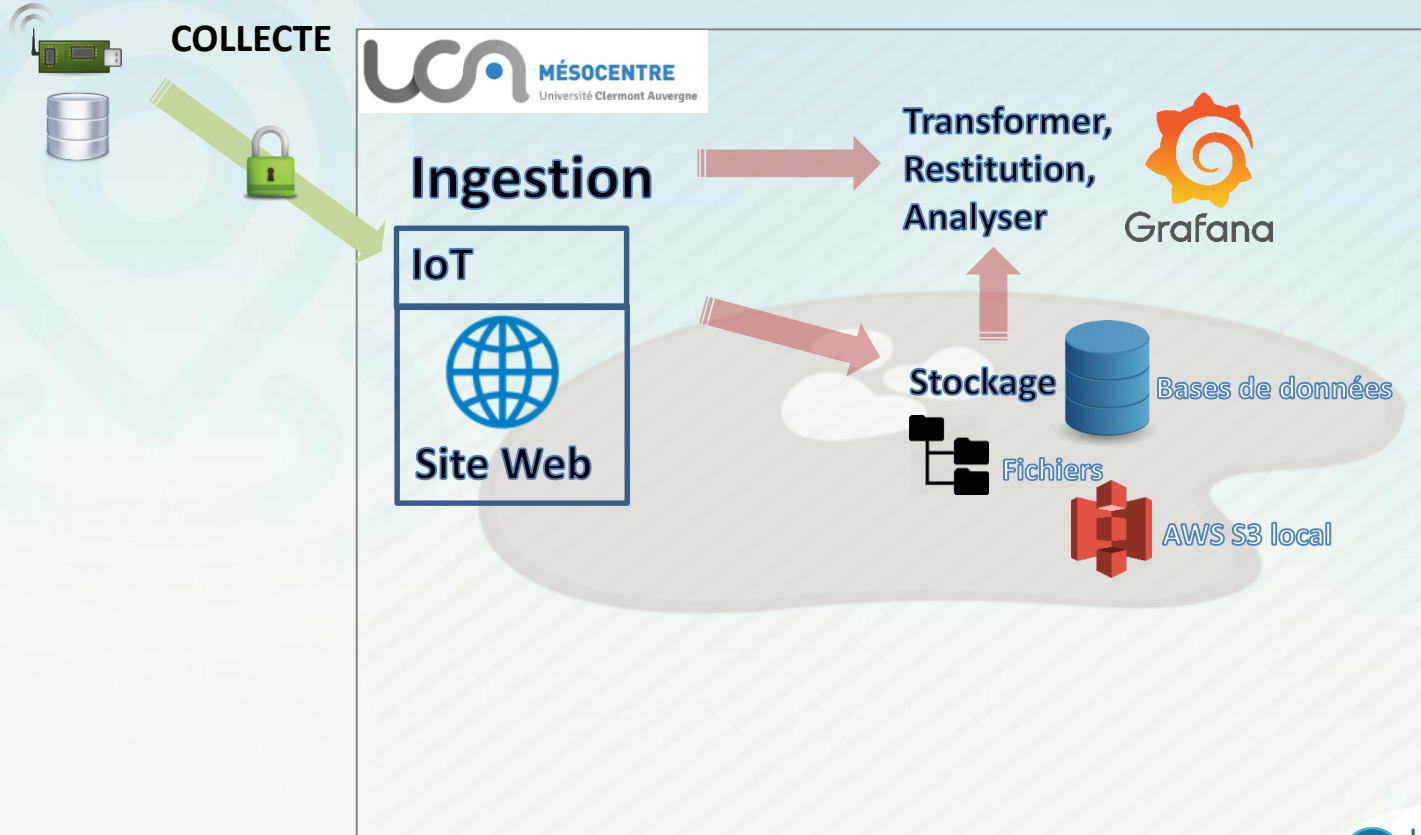
<https://ceba.uca.fr>

Délivrable du CIR1 de l'I-Site sous la thématique "les Agro-écosystèmes durables dans un contexte de changement global"

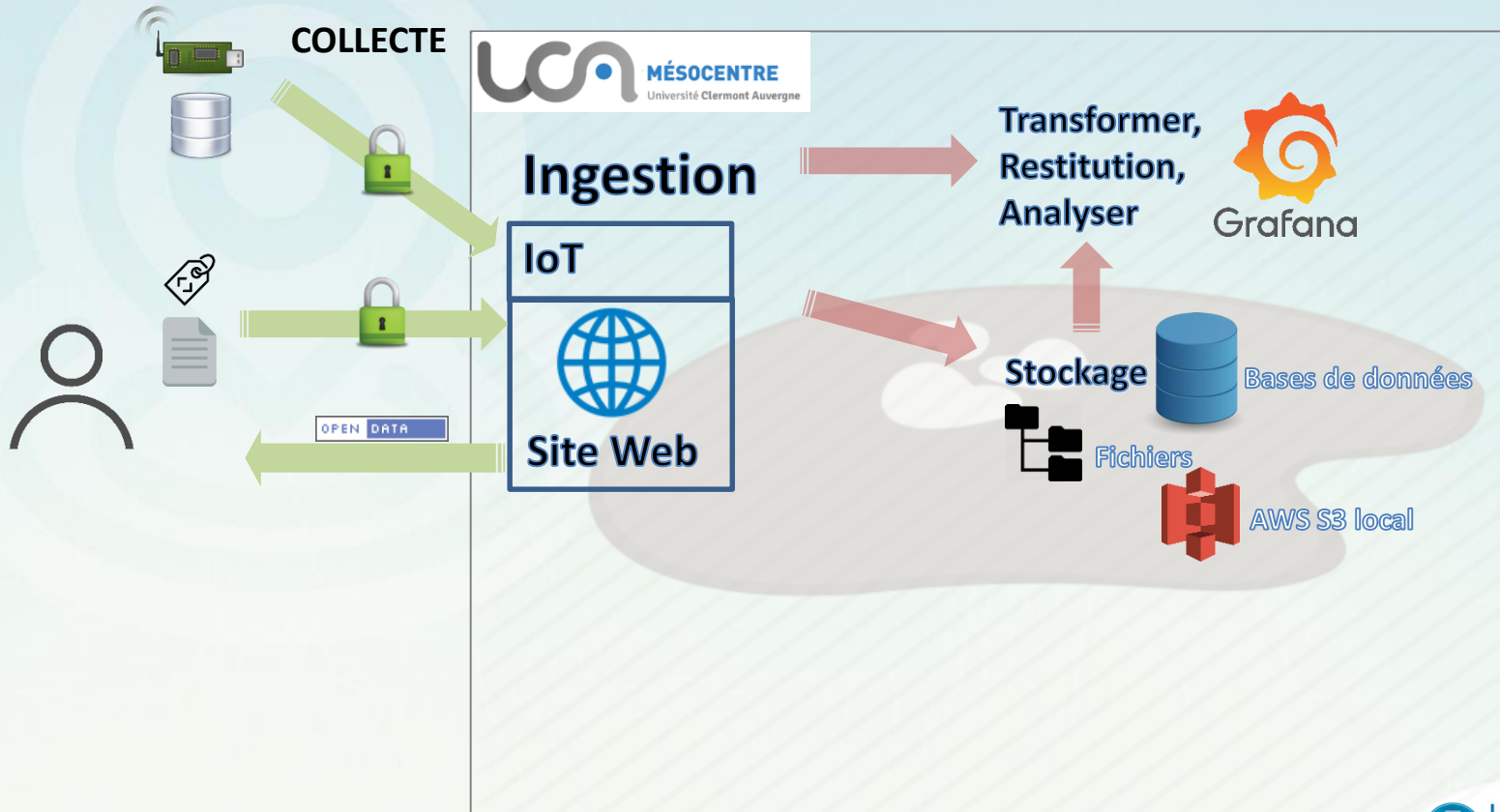
Notre but est de créer un grand observatoire numérique interconnecté en Auvergne. Les objectifs visés par le livrable sont :

- ➔ - **Concevoir** un environnement numérique interconnecté valorisant les données environnementales existantes ou à acquérir
- ➔ - **Automatiser** le monitoring de données à distance par la mise au point de nouveaux systèmes de capteurs sans fil
- **Accroître** notre compréhension des compartiments environnementaux face au changement global
- **Comprendre** l'impact de ces compartiments sur les agroécosystèmes et réciproquement
- **Conceptualiser** les interrelations entre les compartiments pour comprendre, modéliser et prédire le comportement des agroécosystèmes
- **Accompagner** l'aide à la décision (agriculteurs, entreprises, collectivités territoriales)

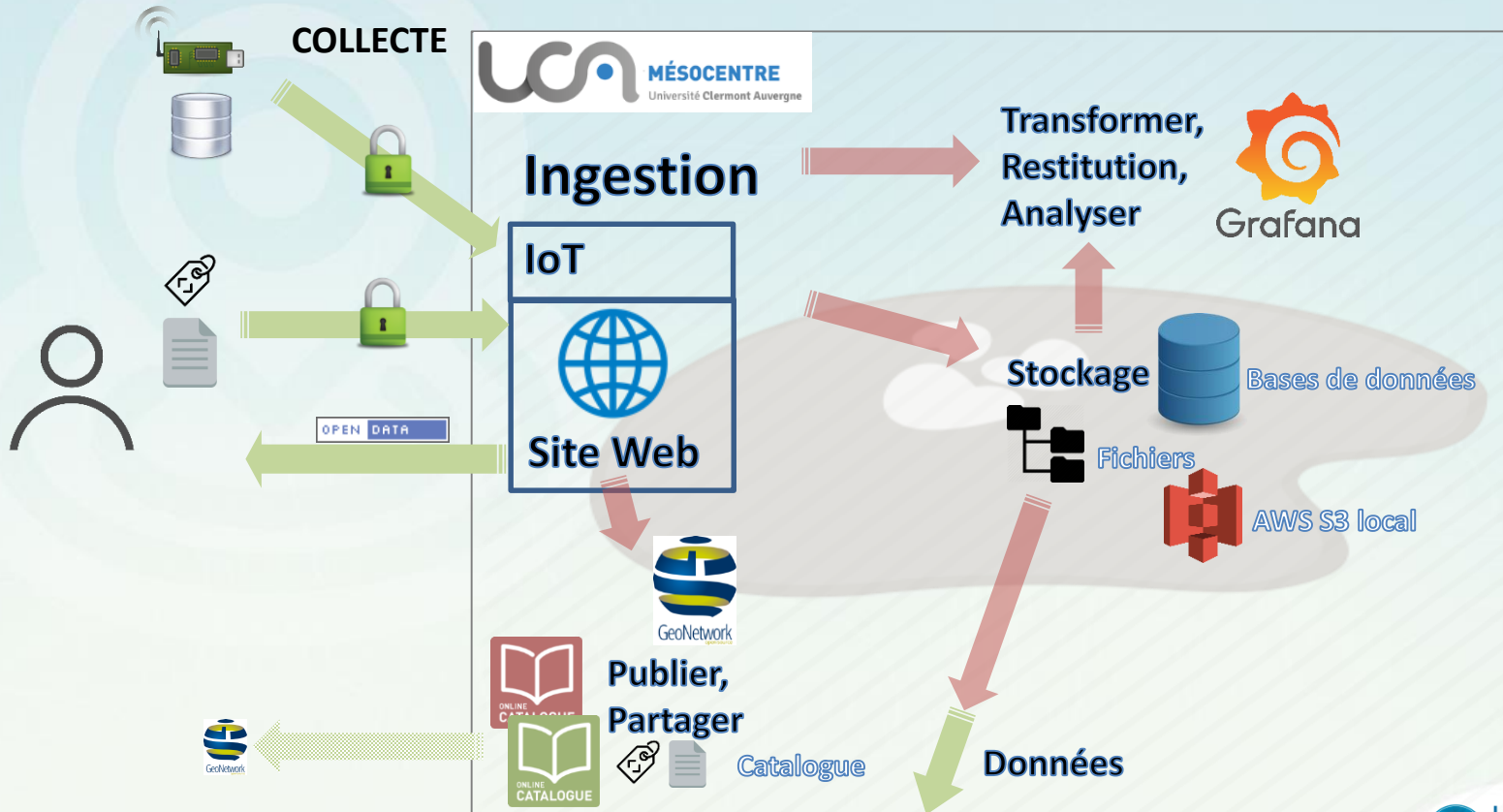
CEBA - Architecture



CEBA - Architecture



CEBA - Architecture



CEBA - Le portail Web

<https://ceba.uca.fr>



- **Point d'accès pour les utilisateurs**
- **Framework PHP : Symfony 4.4**
 - Compétences au sein de l'université
 - MVC
 - Système d'authentification Guard
 - Migration des versions de Symfony
 - ...

Le portail Web - Accès



- **Sécurisé**

- Comptes locaux
 - Gestion en base de données
- Comptes Fédération d'identité (RENATER)
 - Shibboleth + Symfony
 - Création d'un compte sur la bdd à la première connexion
- **Gestion des droits**
 - » User
 - » Admin
 - » Super Admin



Shibboleth.



- **Public**

- Jeu de données et données opendata visibles / accessibles

Le portail Web - La notion de projets



- **Approche FAIRisation des projets dans le CEBA pour permettent de :**
 - Rendre visible un projet de recherches existant (description, image, contacts)
 - Gérer les droits sur les membres du projet
 - ➔ – Associer des **jeux de données**
 - Partager les « données » (fiches métadonnées + fichiers de données)
 - ➔ – Associer / accéder aux **réseaux de capteurs**
 - Décrire la donnée avant de pouvoir la déposer

Le portail Web - La notion de projets




ZATU

...


Date de création du Projet 2021-06-24

La ZATU développe une démarche pluridisciplinaire confrontant les points de vue de biologistes, chimistes, géographes, écologues, géologues, médecins, physiciens, sociologues sur les questions que pose la vie dans des environnements caractérisés par une irradiation chronique d'origine naturelle ou naturelle renforcée. Directeurs : David BIRON, Gilles MONTAVON, Patrick CHARDON

Jeux de données au sein du projet

Le symbole  indique une métadonnée opensource

ZATU SIG Donnees PhysicoChim Plan Compteur 

ZATU SIG Donnees PhysicoChim Sediments 

ZATU SIG Donnees Sociales 

ZATU SIG Fond Carto Images Aeriennes 

ZATU SIG Fiches de prelevements 

Mesure Colibri SN0647 mai 2016 ZATU fin de l'embargo au 19-11-2022

Mesure Colibri SN0641 aout 2015 - Série 2 ZATU 

Sols ZATU 

Nichoires de mésanges Rophin 2017 ZATU

Réseau de capteurs au sein du projet

réseau de capteurs ZATU

Réseau de capteurs

Site au sein du projet

Roffin

Ajouter un nouveau site

Point de contact

Porteur de projet

dasarram@uca.fr

Site web

https://zatu.org

Organisme porteur

INEE

Le portail Web - Jeux de données



• Création

- Par :
 - Fichier excel (plusieurs)
 - Formulaire (1 par 1)
- Métadonnées :
 - ISO 19115, 19139
 - INSPIRE
 - Thésaurus
- Fichiers de données



• Pipeline R :

- Site Web -> CSV -> geometa -> XML -> CSW -> Enregistré dans le Geonetwork

• Lecture / Édition / Suppression

- API REST Geonetwork

<https://github.com/eblondel>

| resource_identifier | parent_iden | status | title | abstract |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Définition : Ce champ est rempli par l'administrateur. Il s'agit d'un code permettant d'identifier de manière unique la fiche de métadonnées. Ce code n'est jamais modifié, même lors d'une mise à jour. | Définition : Il s'agit du code identifiant le jeu de données dont dépend le jeu de donnée actuel | Définition : status du jeu de données ou état d'avancement | Définition : Il s'agit du titre correspondant à un ensemble d'informations permettant d'identifier et de qualifier rapidement la donnée. C'est généralement l'information qui apparaît en premier à la suite des recherches effectuées pour trouver une donnée. Ce titre doit être explicite et définir clairement la donnée s'y rapportant. Ce nom doit être caractéristique et unique. Les acronymes sont déconseillés. | Définition : Le résumé permet de décrire la donnée en lui attribuant une définition officielle, quand elle existe, ou, le cas échéant, une définition commune afin de la rendre compréhensible par l'utilisateur. Il s'agit en particulier de définir au mieux l'information ou le phénomène représenté. C'est un élément indispensable de la fiche de description. |

Le portail Web - Les réseaux de capteurs 1/2



- **Connexion des réseaux de capteurs à l'infrastructure CEBA**
- **Format JSON recommandé**
- **Plusieurs moyens de visualiser la donnée**
 - Grafana
 - Requêtes stockées génériques prédéfinies
 - Appel de la requête comme une fonction
 - Aucun changement à faire côté web si mise à jour
- **Requêtes disponibles via une interface graphique**
- **Extraction des données en CSV ou en JSON possible**

Le portail Web - Les réseaux de capteurs 2/2



réseau de capteurs ZATU

⋮ GRAFANA

ACCÉDER À L'INTERFACE GRAFANA

Accéder à grafana pour créer des tableaux de bords et visualiser vos données

Grafana

⋮ INTERACTION BASE DE DONNÉES

Statut

Connexion à la base de données ok

RÉCUPÉRER TOUTES LES DONNÉES

Sélectionner une plage temporelle :

Date début

Date Fin

Envoyer

JSON CSV

RÉCUPÉRER LES DONNÉES D'UNE MESURE

Tous les types de mesures ne sont pas forcément liés à votre réseau de capteurs

Préciser la mesure

Sélectionner une plage temporelle :

Date début

Date Fin

JSON CSV Envoyer

RÉCUPÉRER LES DONNÉES D'UN DEVEUI

Préciser le devEUI

JSON CSV Envoyer

RÉCUPÉRER LES DONNÉES ENTRE QUATRE POINTS GÉOGRAPHIQUES

Préciser les points géographiques :

Pour réaliser une emprise, renseigner les points dans l'ordre (A,B,C,D) du polygone.

latitude 1 longitude 1

latitude 2 longitude 2

latitude 3 longitude 3

latitude 4 longitude 4

JSON CSV

Envoyer

Le portail Web - Exploration



- **Par projet**
- **Dans les métadonnées (simple)**
 - Recherche par mot clés dans le nom, le sujet, les thésaurus, qualité de la donnée, depuis une barre de recherche
- **Dans les métadonnées (plus poussée avec des filtres)**
 - Recherche par filtre grâce à des boutons html radio sur les Topic Categories (norme INSPIRE), les formats, les dates, les Thésaurus INSPIRE / GEMET / Autres



Le portail Web - La base de données WEB

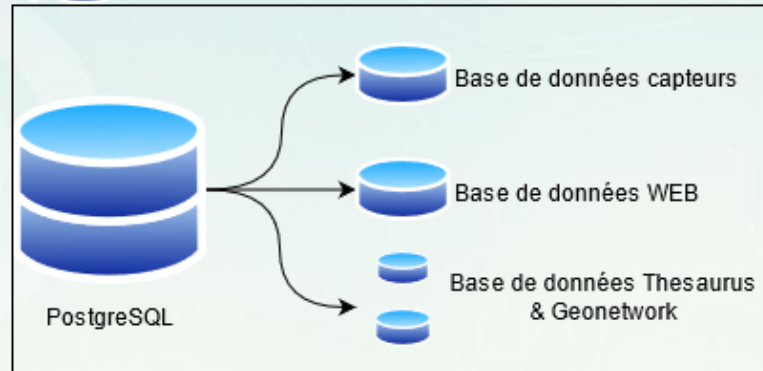


- Contient toutes les données permettant le bon fonctionnement du site web
- Contient toutes les métadonnées
- Connexion à plusieurs bases de données



BDD - L'existant

Serveur de base de données



BDD - L'existant



- **Actuellement, 2 types de réseaux de capteurs connectés**
 - Délimité & Connecsens, INRAE à venir
- **Stockage en JSON**
 - Enrichissement possible lors de l'insertion
- **Plusieurs moyens de visualiser la donnée**
- **Requêtage via emprise géographique**
 - Ajout de coordonnées fixes pour les systèmes qui ne transmettent pas d'informations GPS (WGS 84)

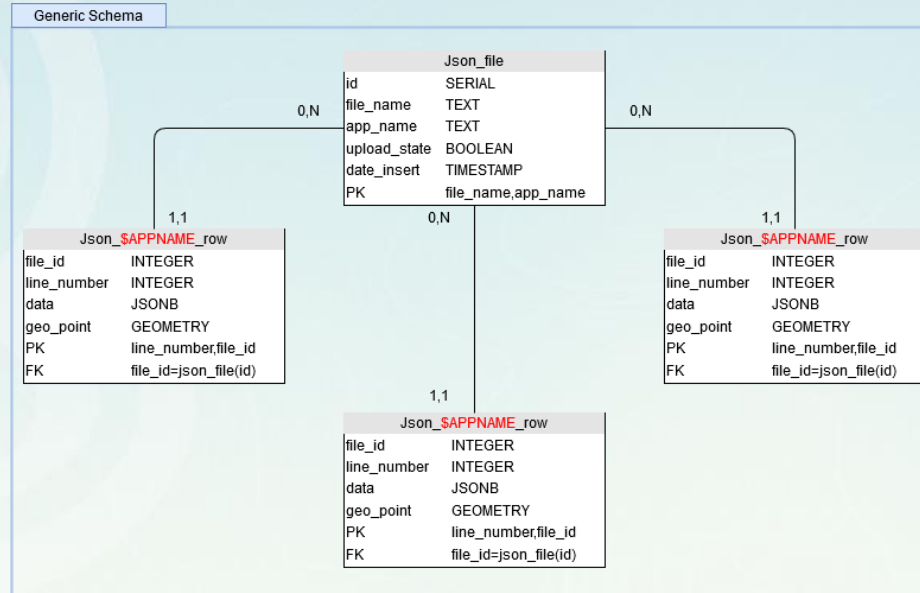
BDD - La base de données capteurs

- **Pourquoi stocker du JSON**
 - Permet des requêtes SQL directement sur le contenu des lignes JSON
 - Stockage de données brutes sous un seul format
- **Utilisation de l'attribut JSONB sous Postgres**
- **Plusieurs avantages:**
 - Permet des recherches précises et complexes
 - JSONB plus rapide que le JSON pour le requêtage
 - JSONB supporte l'indexation (sur une clé JSON)
 - Export possible des fichiers en JSON
- **Couche géographique disponible via l'extension PostGIS**

BDD - Structure de la base de données capteurs

- **Structuration sous forme de schémas Postgres**
 - Un schéma est l'équivalent d'une base
 - Permet la réutilisation des noms de tables
- **Séparation par type de réseaux de capteurs**
 - Un schéma pour les réseaux de capteurs ConnecSens
 - Un schéma pour les réseaux de capteurs créant des fichiers avec délimiteurs
 - ...

BDD - Structure de la base de données capteurs



| Schema_jsonfields | |
|-------------------|-----------------------|
| schema | TEXT |
| spec | TEXT |
| jsonfield | TEXT |
| PK | schema,spec,jsonfield |

| Schema_sensors | |
|----------------|--------------------|
| schema | TEXT |
| spec | TEXT |
| measure | TEXT |
| PK | schema,spec,mesure |

| Field_descriptor | |
|------------------|--------------------|
| schemaname | TEXT |
| appname | TEXT |
| datefield | TEXT |
| geo_long_field | TEXT |
| geo_lat_field | TEXT |
| geo_long_value | DOUBLE PREC |
| geo_lat_value | DOUBLE PREC |
| delimiter | TEXT |
| FK | schemaname.appname |

BDD - Exemple de requête SQL

SELECT

```
file_id,
data ->> 'Data-Type' as "sensor-type",
data ->> 'Data-Value' as "sensor-value",
data ->> 'Date' as "date"
data ->> 'WinterTimeStamp' as "hour"
```

Clés JSON

FROM

```
delimitedtojson.json_aydatbouee_row
```

WHERE

```
file_id between 29500 and 30000
```

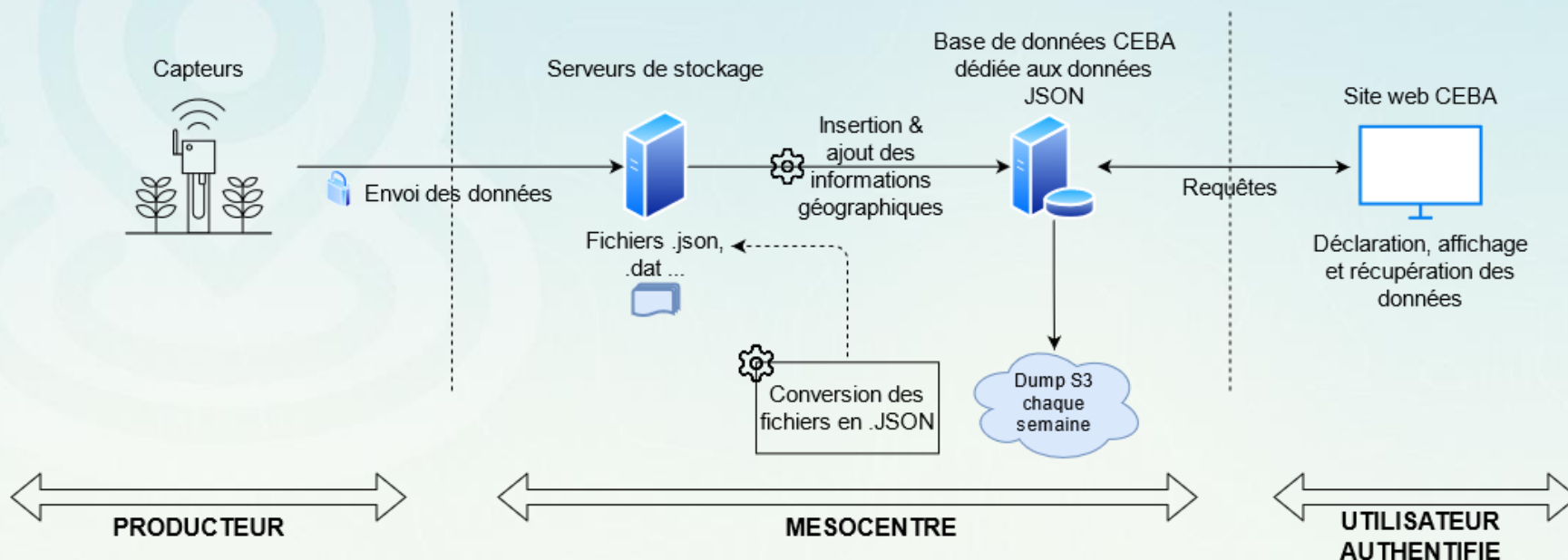
AND

```
data ->> 'Data-Type' = 'Temp'
```



| file_id | sensor-type | sensor-value | date | hour |
|---------|-------------|--------------|----------|--------|
| 29501 | Temp | 12,50 | 20211015 | 070000 |
| 29502 | Temp | 12,50 | 20211015 | 080000 |
| 29502 | Temp | 12,50 | 20211015 | 090000 |
| 29507 | Temp | 13,20 | 20211015 | 140000 |
| 29508 | Temp | 13,10 | 20211015 | 150000 |
| 29509 | Temp | 13,20 | 20211015 | 160000 |
| 29513 | Temp | 13,10 | 20211015 | 200000 |
| 29514 | Temp | 13,00 | 20211015 | 210000 |
| 29515 | Temp | 12,90 | 20211015 | 220000 |
| 29519 | Temp | 12,50 | 20211016 | 020000 |
| 29520 | Temp | 12,40 | 20211016 | 030000 |
| 29521 | Temp | 12,30 | 20211016 | 040000 |

BDD - Ingestion des données provenant de réseaux de capteurs



BDD - Ingestion des données provenant de réseaux capteurs

Fichier .dat

```
0000000067;0001;20211124;070000;7,40
0000000067;0002;20211124;070000;155,9
0000000067;0003;20211124;070000;33,6
0000000067;0004;20211124;070000;3,39
0000000067;0005;20211124;070000;0,13
0000000067;0006;20211124;070000;2,13
0000000067;0007;20211124;070000;63,16
0000000067;0008;20211124;070000;0,01385
0000000067;0009;20211124;070000;0,01253
0000000067;0010;20211124;070000;12,8
```



Transformation avec
enrichissement de la
donnée si possible

Fichier .json

```
{ "StationNumber": "0000000067", "SensorNumber": "0001", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "7,40", "Data-Type": "Temp", "Data-Unit": "°C" }
{ "StationNumber": "0000000067", "SensorNumber": "0002", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "155,9", "Data-Type": "SpConductivité", "Data-Unit": "µS/cm" }
{ "StationNumber": "0000000067", "SensorNumber": "0003", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "33,6", "Data-Type": "O2", "Data-Unit": "%" }
{ "StationNumber": "0000000067", "SensorNumber": "0004", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "3,39", "Data-Type": "O2", "Data-Unit": "Mg/l" }
{ "StationNumber": "0000000067", "SensorNumber": "0005", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "0,13", "Data-Type": "TensionSonde", "Data-Unit": "V" }
{ "StationNumber": "0000000067", "SensorNumber": "0006", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "2,13", "Data-Type": "ChloroA", "Data-Unit": "µg/l" }
{ "StationNumber": "0000000067", "SensorNumber": "0007", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "63,16", "Data-Type": "Phyco", "Data-Unit": "µg/l" }
{ "StationNumber": "0000000067", "SensorNumber": "0008", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "0,01385", "Data-Type": "ChloroA", "Data-Unit": "V" }
{ "StationNumber": "0000000067", "SensorNumber": "0009", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "0,01253", "Data-Type": "Phyco", "Data-Unit": "V" }
{ "StationNumber": "0000000067", "SensorNumber": "0010", "Date": "20211124", "WinterTimeStamp": "070000", "Data-Value": "12,8", "Data-Type": "Tension", "Data-Unit": "V" }
```

Application concrète

Requêtage/extraction depuis le site web



ACCÉDER À L'INTERFACE GRAFANA

Accéder à grafana pour créer des tableaux de bords et visualiser vos données

Grafana

INTERACTION BASE DE DONNÉES

Statut Connexion à la base de données ok

RÉCUPÉRER TOUTES LES DONNÉES

Sélectionner une plage temporelle :

Date début Date Fin Envoyer

JSON CSV

RÉCUPÉRER LES DONNÉES D'UNE MESURE

Tous les types de mesures ne sont pas forcément liés à votre réseau de capteurs

Préciser la mesure

Sélectionner une plage temporelle :

Date début Date Fin

JSON CSV Envoyer

- Requête directement la base de données
- Création d'un fichier JSON ou CSV lors de l'extraction

```
"txInfo-frequency": 868100000, "data-radioactivity": 103, "servtimestampUTC": "2020-02-14T00:18:28.623627", "data-CNSSRFData": 0, "txInfo-frequency": 868100000, "data-radioactivity": 0, "servtimestampUTC": "2020-02-13T23:48:59.897722", "data-CNSSRFData": 0, "txInfo-frequency": 868500000, "data-radioactivity": 0, "servtimestampUTC": "2020-02-13T23:32:38.008047", "data-CNSSRFData": 0, "txInfo-frequency": 868100000, "data-radioactivity": 206, "servtimestampUTC": "2020-02-13T23:17:57.524887", "data-CNSSRFData": 0, "txInfo-frequency": 868300000, "data-radioactivity": 0, "servtimestampUTC": "2020-02-13T22:35:02.882650", "data-CNSSRFData": 0, "txInfo-frequency": 868300000, "data-radioactivity": 103, "servtimestampUTC": "2020-02-13T22:35:02.882650", "data-CNSSRFData": 0, "txInfo-frequency": 868500000, "data-radioactivity": 310, "servtimestampUTC": "2020-02-13T22:05:08.052619", "data-CNSSRFData": 0, "txInfo-frequency": 868500000, "data-radioactivity": 0, "servtimestampUTC": "2020-02-13T22:05:08.052619", "data-CNSSRFData": 0, "txInfo-frequency": 868300000, "data-radioactivity": 206, "servtimestampUTC": "2020-02-13T21:34:48.708835", "data-CNSSRFData": 0
```

BDD - Sauvegarde & sécurité

- **Sauvegardes des bases sur le S3**
- **Réplication active du serveur principal sur un serveur secondaire**
- **Multiples utilisateurs en base pour chaque tâche**
 - Chaque utilisateur possède un accès limité
 - Un compte en base par tâche
- **Volumétrie actuelle**
 - 3M5 enregistrements pour ConnecSenS
 - 224k provenant de la bouée Aydat
 - Représente 5,2 Go au total

CEBA - Ses interactions

- **Local/Régional/National :**
 - Forte collaborations avec les autres équipes/laboratoires de l'UCA (Geoloab, LGME, Inrae, OPGC, FRE...)
 - Impliqué dans plusieurs PIA
 - inLIT : données privées, LIT Grandes Cultures en Auvergne (non démarré)
 - Terra-forma : CEBA identifié comme une solution innovante à déployer sous la forme de fédération de plateforme
 - Impliqué dans l'action collective « Bancarisons Ensemble les données » (BED) du RZA (contact région sud ouest)
 - Héberger les données
 - Du projet Nolimit – Pack Ambition Recherche
 - De la ZATU : Zone Atelier Territoires Uranifères - INEE
 - Du projet NEEDS INSPECT - CNRS
 - Projet Radonorm – Europe
- **Restitution de données pour événements : Aydat 15 octobre 2021**
 - Permet l'accès à la donnée depuis plusieurs sources
- **Accompagnement des utilisateurs pour la saisie des métadonnées**

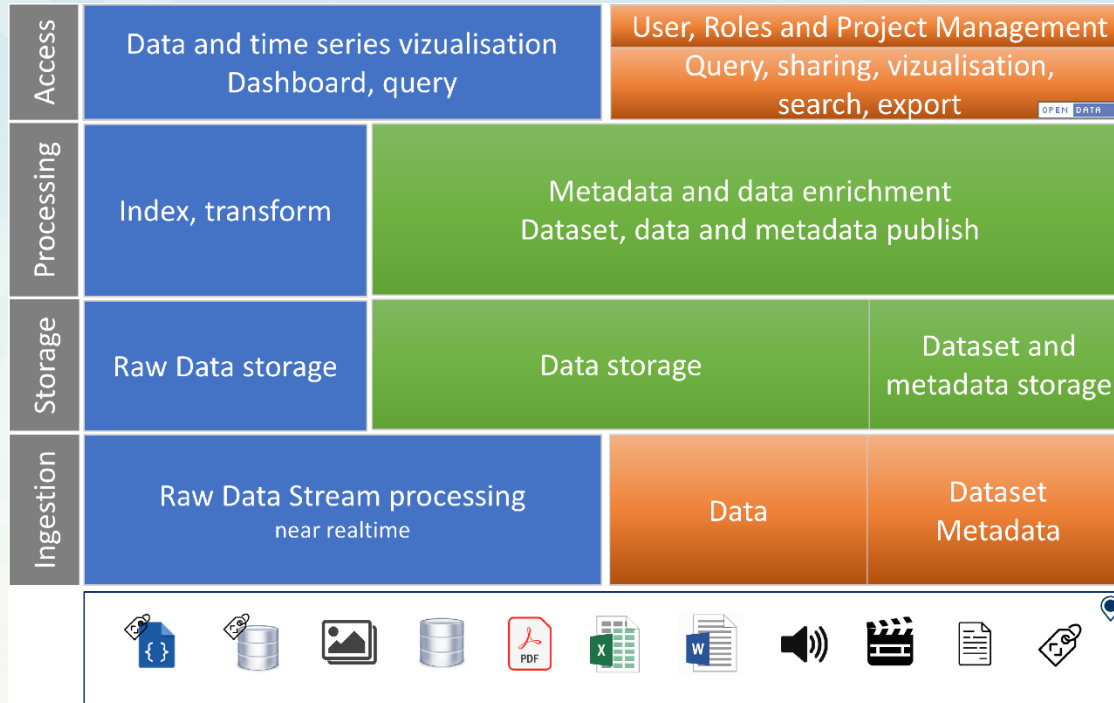
CEBA - Les évolutions

- **DOI (Digital Object Identifier)**
- **Moissonnage**
- **Performance énergétique**
- **Intégration de nouveaux réseaux de capteurs**
- **Mise à jour des outils**



Merci de votre attention !

CEBA - Qu'est-ce que c'est ?



Heterogenous sources : structured, semi-structured, unstructured

Équipe

- Vincent Breton
- Alexandre Claude
- Gilles Mailhot
- Jérémy Mezhoud
- Francis Ogereau
- David Sarramia
- Jean-Claude Bailly
- Emma-Laure Gros

- Estelle Théveniaud
- Antoine Mahul
- David Grimbichler
- Alexis Ozwald

Coordination
Suite elastic & architecture
Coordination
Base de données
Web, data-catalog
Coordination / Suite elastic
Infrastructure, système (1 an)
Chargé de communication (2 mois)

Geolab
Mésocentre
Mésocentre
DOSI, système



I-SITE CLERMONT

Clermont Auvergne Project



Un projet I-Site du Programme Investissement d'Avenir II