

ACTION 18 DU PNSE4

## **GD4H - Green Data for Health**

**Pour la mobilisation et valorisation, par les acteurs de la recherche et expertise, des données environnementales au service de la santé environnement.**

- **Historique du projet**
- **Architecture technique**
- **FOCUS: Fonctionnalités coeur du catalogue**

# ● Historique technique du projet



# Green Data for Health

Identifier les effets des facteurs  
environnementaux sur la santé par la donnée

Santé-Environnement

Promotion 5

## Les défis techniques

La solution :

- décrira les données environnementales pertinentes et disponibles pour être croisées avec des données de santé  
CATALOGUE DE DONNES en ligne spécifique à la Santé/Environnement:  
Un modèle de données dynamique et flexible
- facilitera la repérabilité et l'accès aux données environnementales  
MOTEUR DE RECHERCHE ET FILTRES adaptés au domaine
- améliorera l'interopérabilité des données environnementales entre elles et avec des données de santé  
IMPORT/EXPORT aux formats standards
- prendra en compte les enjeux juridiques de mise à disposition et de sécurité des données  
QUALIFICATION JURIDIQUE + module d'aide juridique
- favorisera la synergie entre acteurs de l'écosystème recherche et l'expertise en Santé-Environnement

PLATEFORME COLLABORATIVE permettant la montée  
en qualité des données et l'échange entre les parties  
prenantes autour des ressources

## ● Historique et enseignements

— GD4H

### Brique recensement et outillage:

Améliorer la repérabilité et l'accessibilité des bases de données environnementales et assurer leur interopérabilité à travers un catalogue de données en ligne

#### Septembre 2021

Recensement initial auprès des experts métiers (excel)

Prototype CKAN par Wavestone

#### Janvier 2022

##### 3 grands enseignements:

- Spécificité Santé/Environnement: modèles du catalogue sur mesure
- Catalogue avec de multiples parties prenantes: collaboratif
- Gestion du recensement et du catalogue: à outiller

#### Février 2022

##### Développement d'une plateforme de catalogage

- collaborative
- adaptée aux exigences de qualité et interopérabilité\*
- flexible et sur mesure
- > DEMO d'un prototype de catalogue le 24 février

\* Suivre les principes [FAIR](#)

Adaptation au schema [dcat-ap](#)

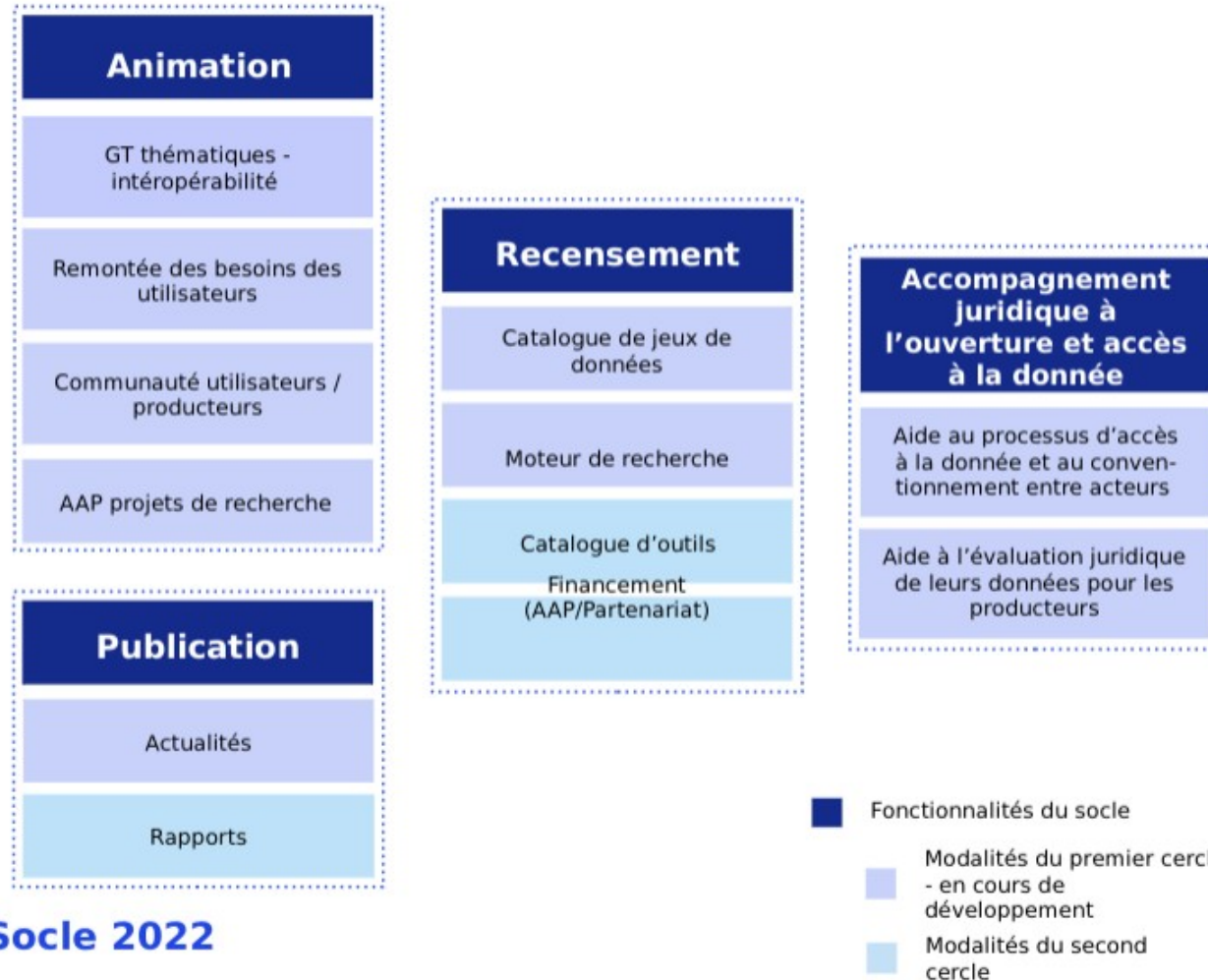
+ alignement avec [catalogue interministériel](#)

+ alignement avec [catalogue INSPIRE](#)

# Périmètre de l'offre de service en développement en 2022

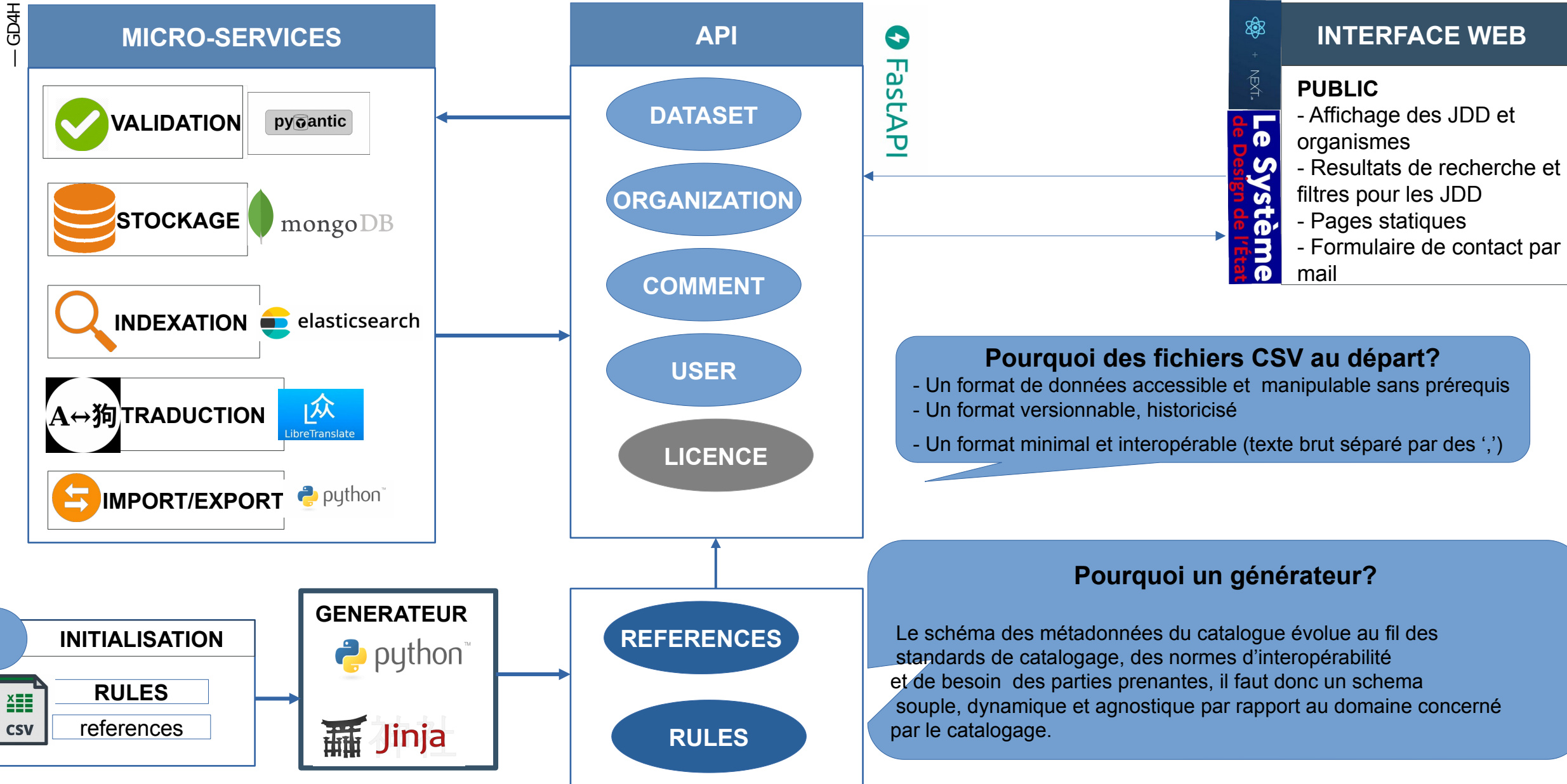
## Éléments techniques découlant de l'offre de service :

- Plateforme en ligne
- Développement d'un catalogue (pris en charge par la data-ingénierie)
- Développement d'un moteur de recherche, affichage résultats, fiches descriptives des JDDs
- Module de remontée de commentaires généraux sur le catalogue et plus précis sur les valeurs descriptives des données
- Comptes administrateurs (plusieurs niveaux) incluant tableau de suivi de l'activité, des commentaires, ajout/mise à jour de JDDs, modification des méta-données des JDDs etc.
- Module FAQ dynamique (arborescence de décisions)
- Pages diverses (accueil, qui sommes-nous, contact, pages de contenu et légales etc.)



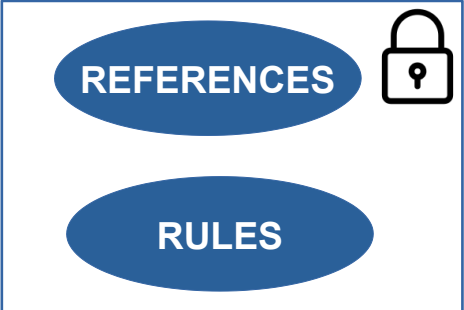
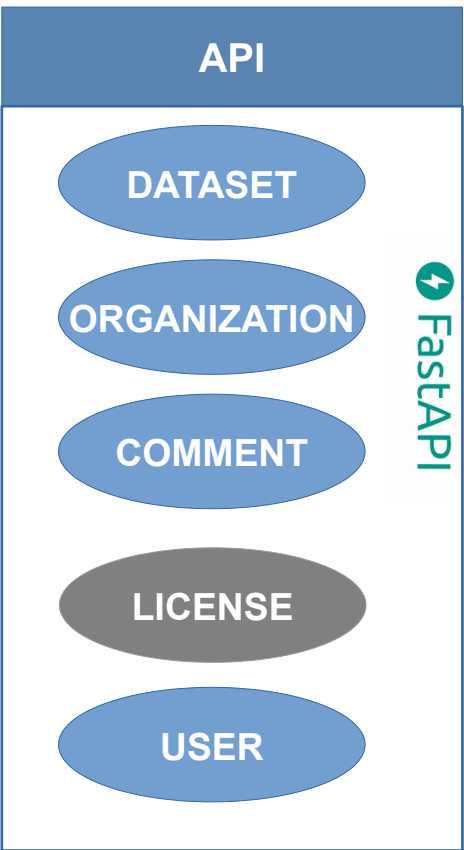
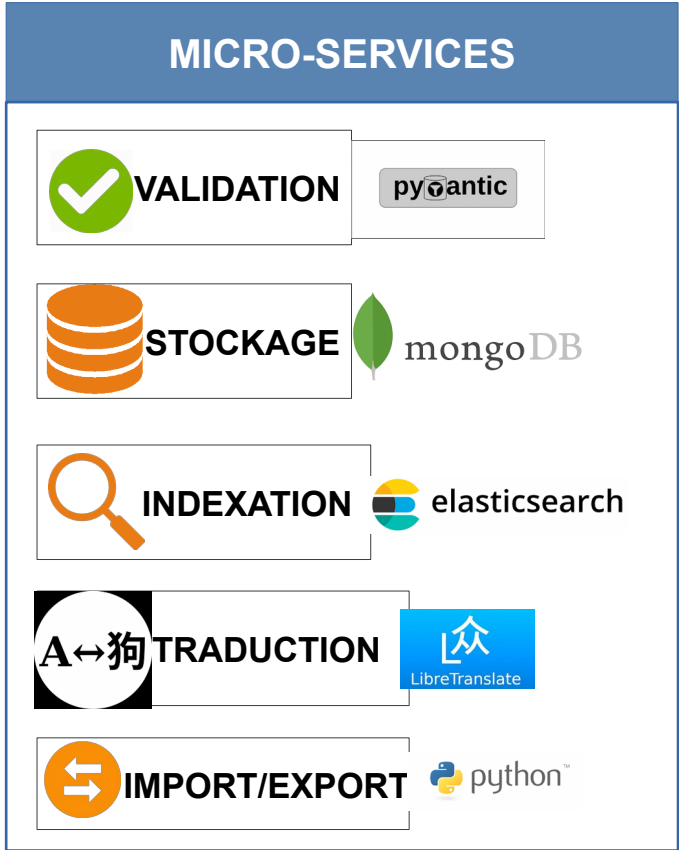
# ● Architecture technique

# 1 Initialisation du système de catalogage

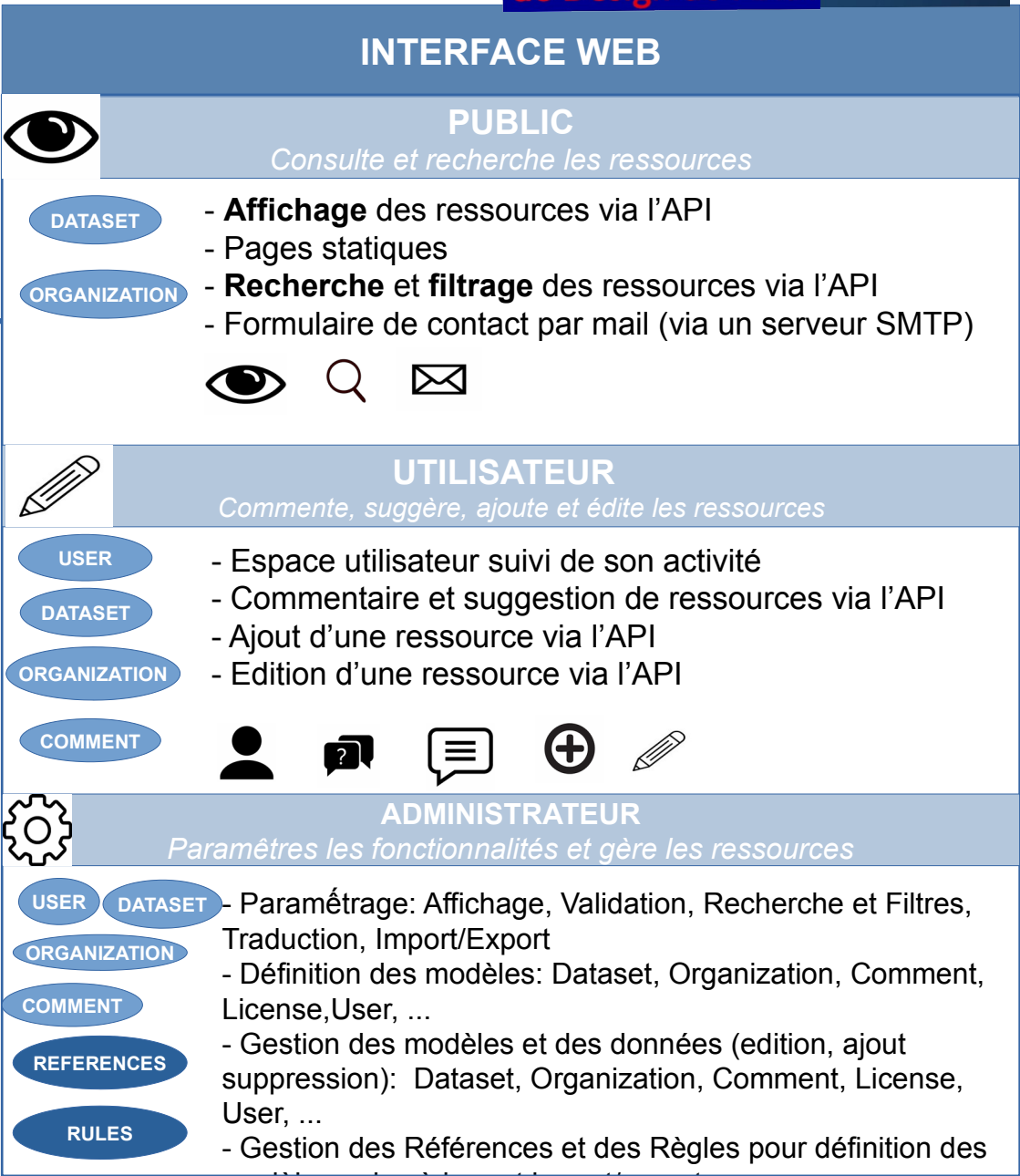




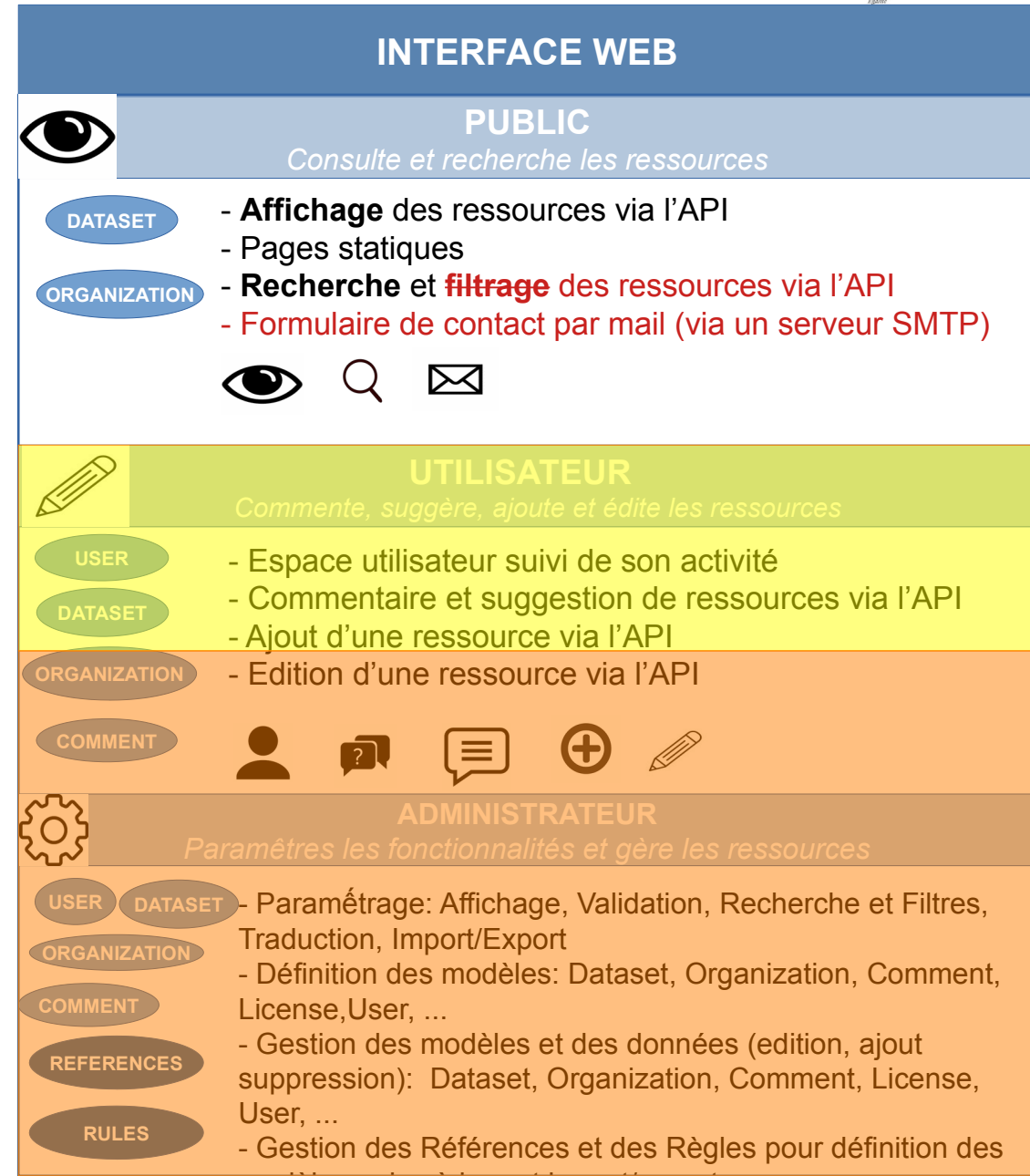
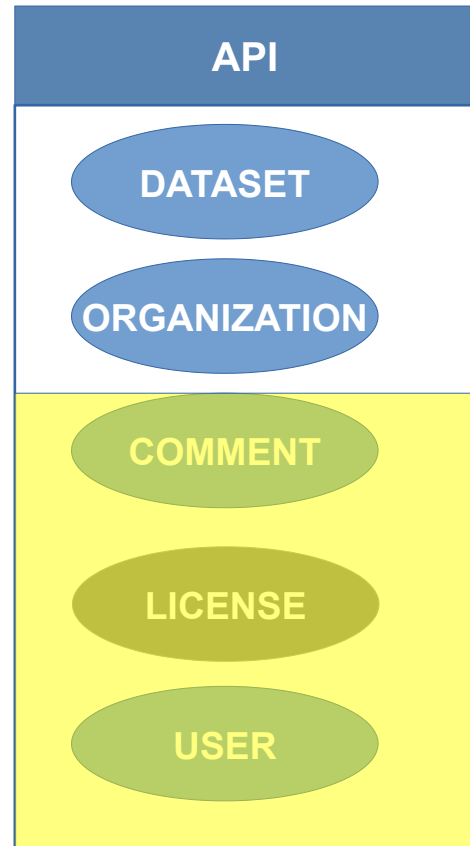
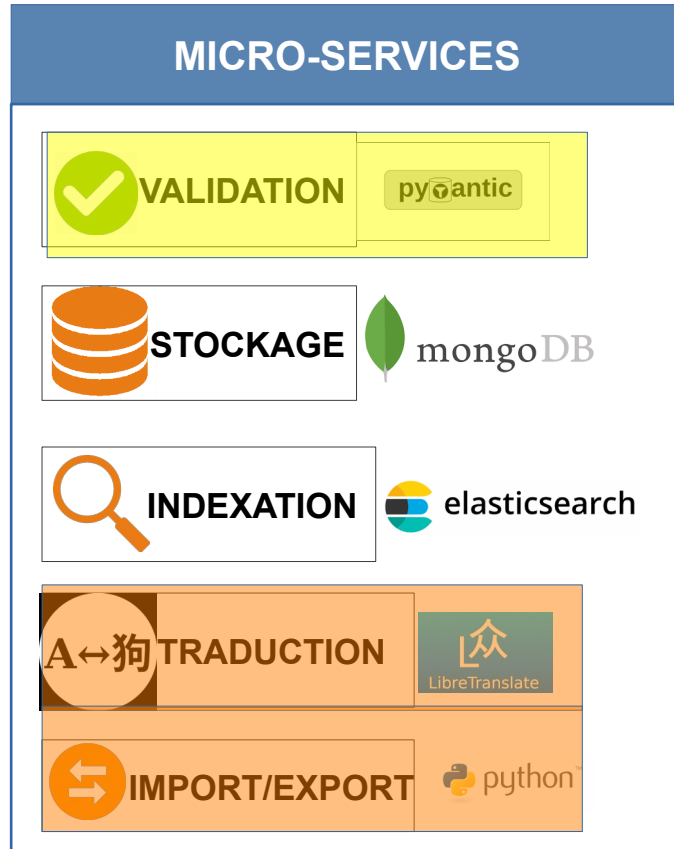
● 2. Gestion du catalogage au long cours



Une fois le projet initialisée, la gestion (paramétrage, ajout, édition et mises à jours...) se fait via l'API et idéalement depuis une interface d'administration .



### 3. Aperçu du périmètre fonctionnel



Fonctionnalités planifiées pour la V2

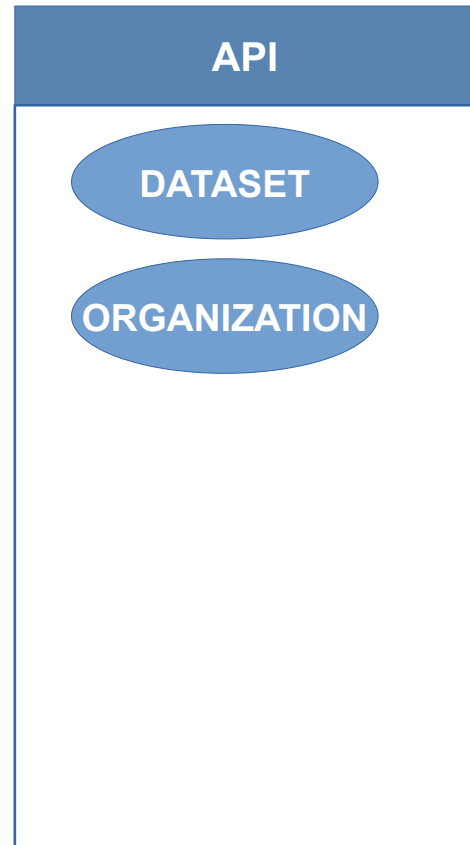
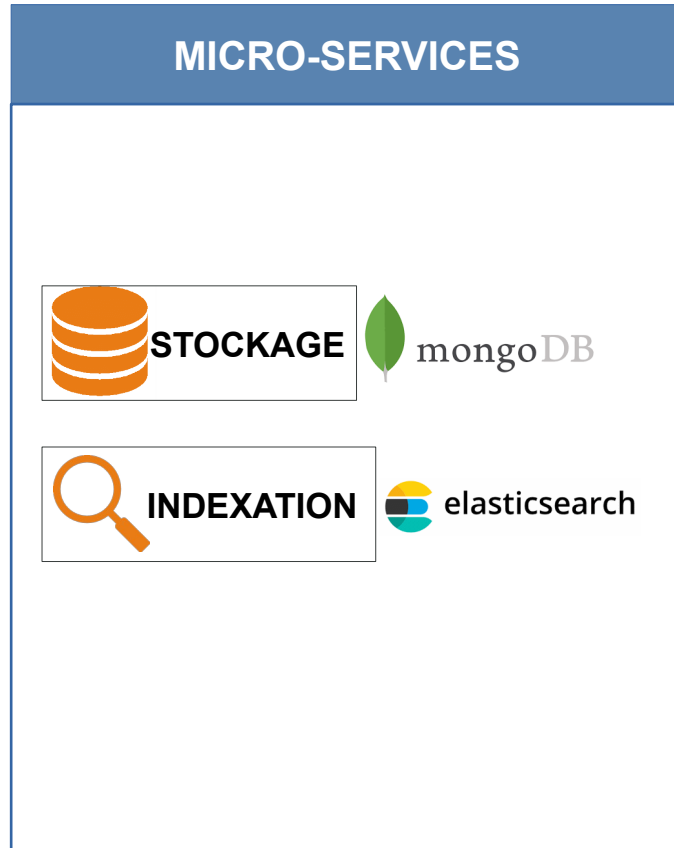
Fonctionnalités à arbitrer

REFERENCES

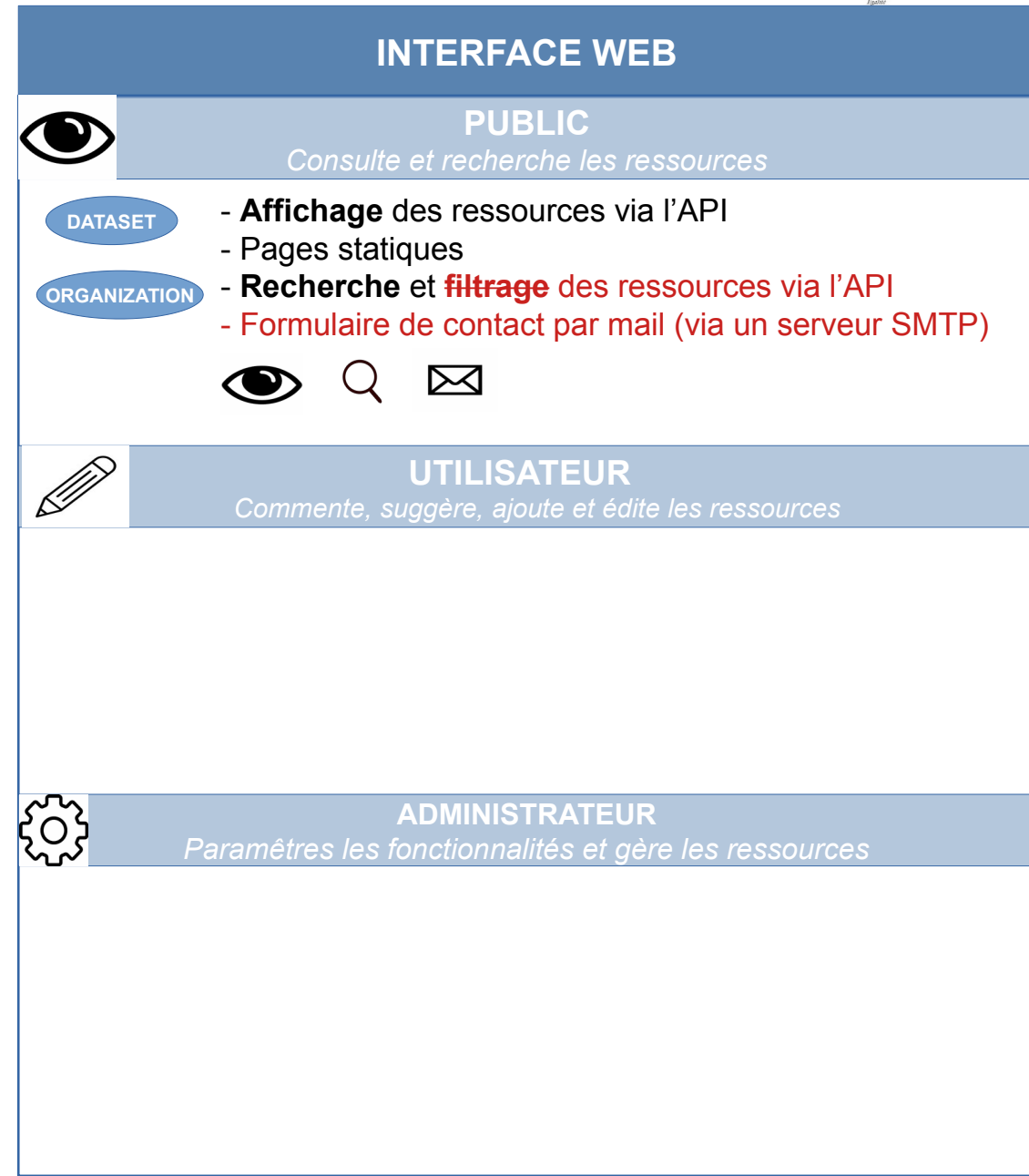
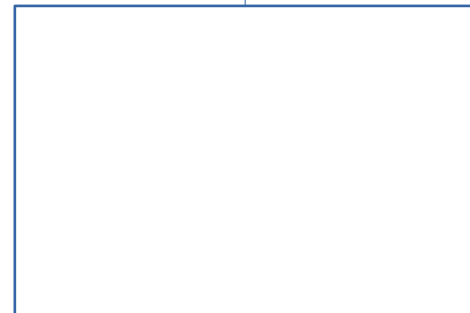
RULES

## 4. Périmètre fonctionnel de la V1

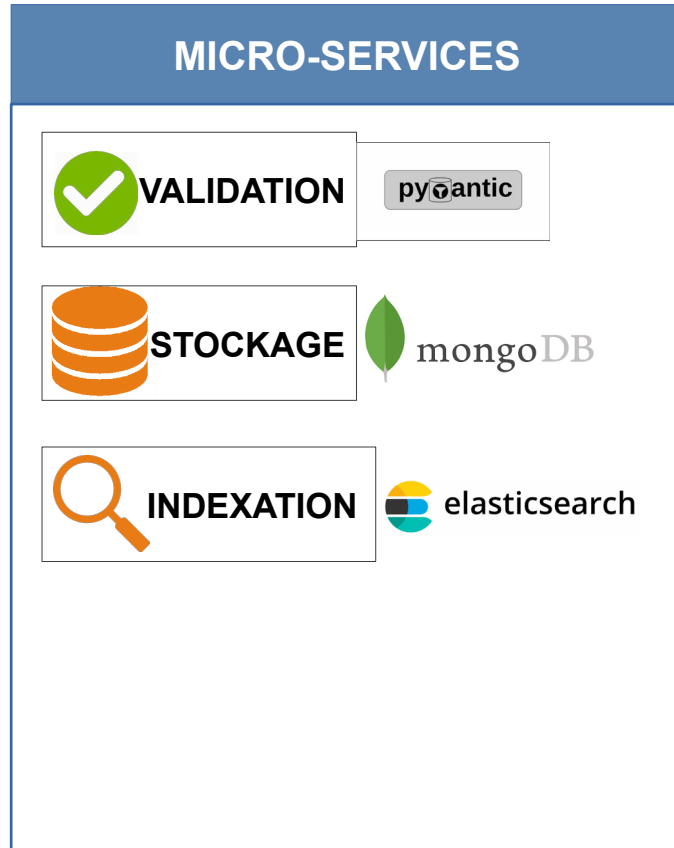
— GD4H



PROTOTYPE V1 livré le 29/04/2022



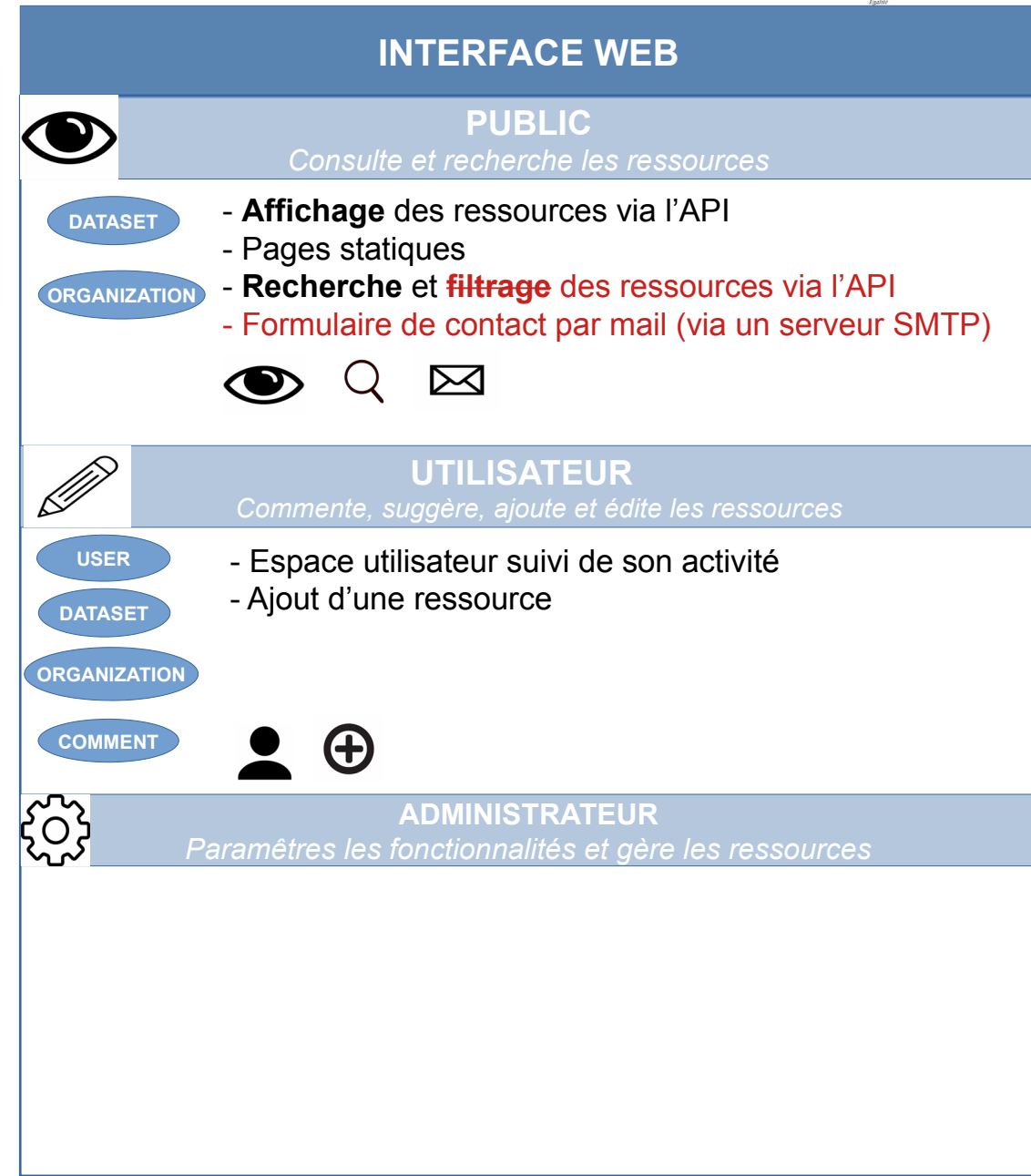
## 5. Périmètre fonctionnel pour la V2



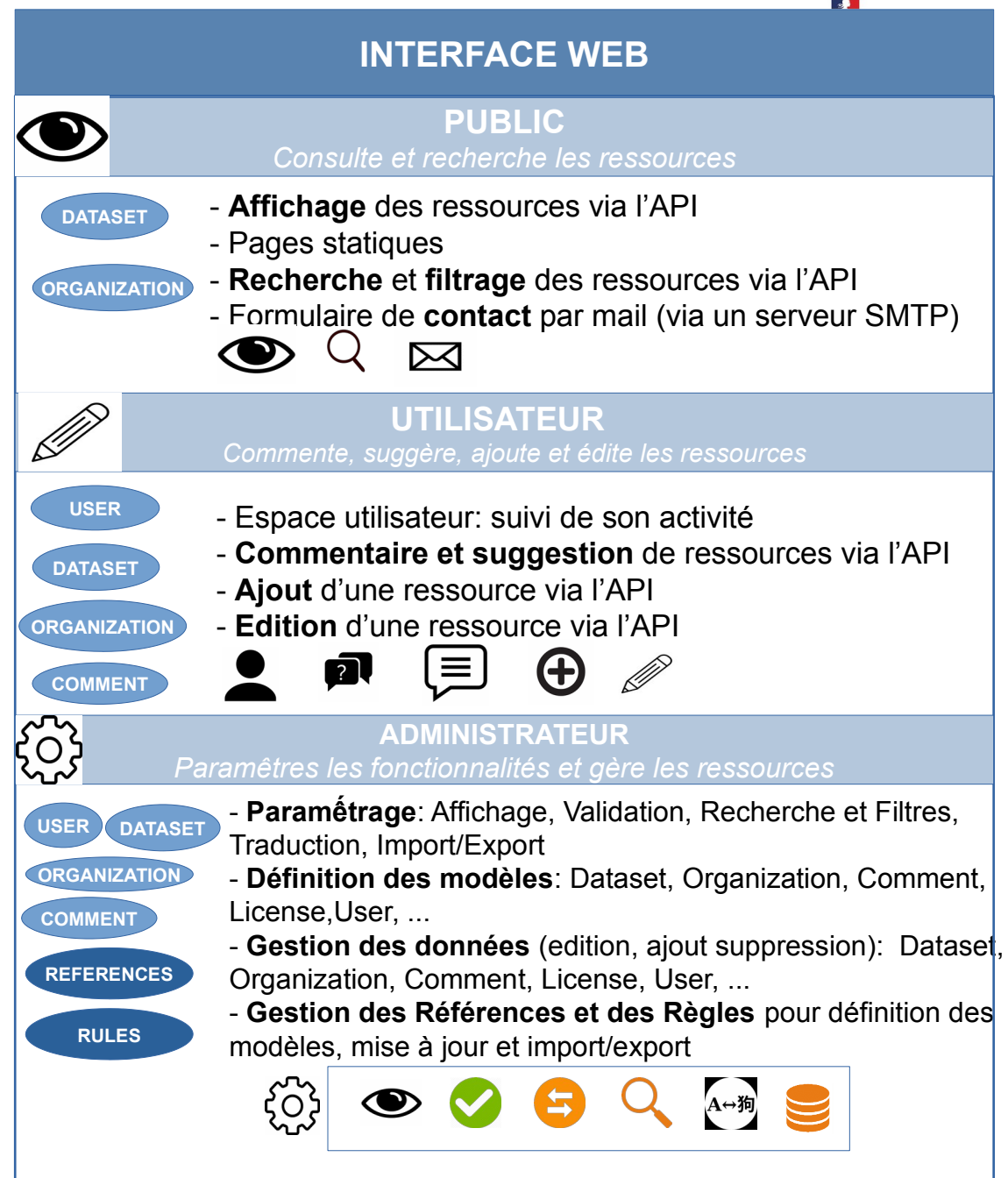
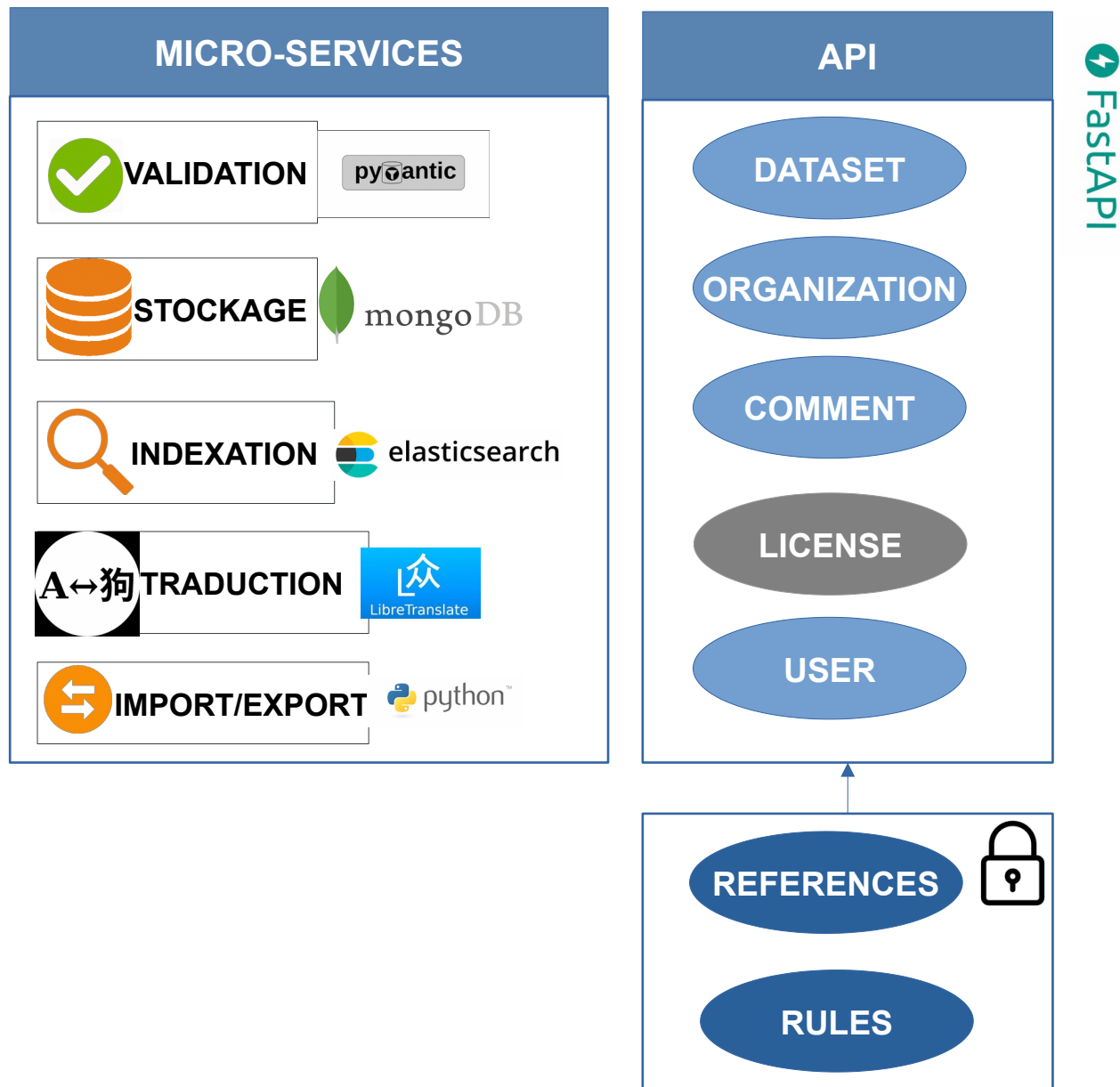
FastAPI



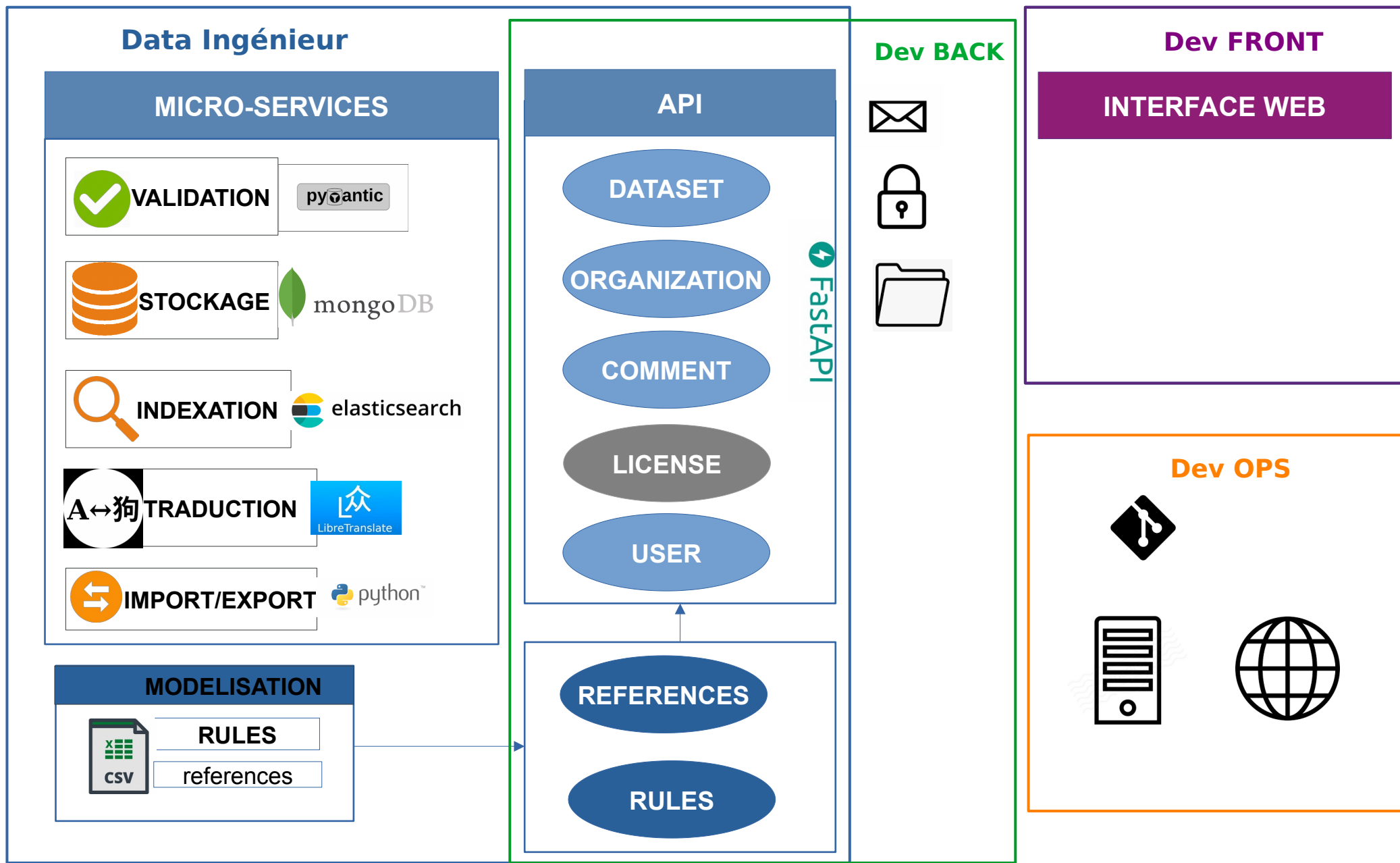
PROTOTYPE V2 livraison au XXXX?



## 6. Périmètre fonctionnel idéal



## 5. Périmètre d'intervention des spécialités techniques



## 6. Status des briques techniques

— GD4H

### DEVOPS:

Procédure de sauvegarde  
de la BDD et de l'index?

### En suspens:

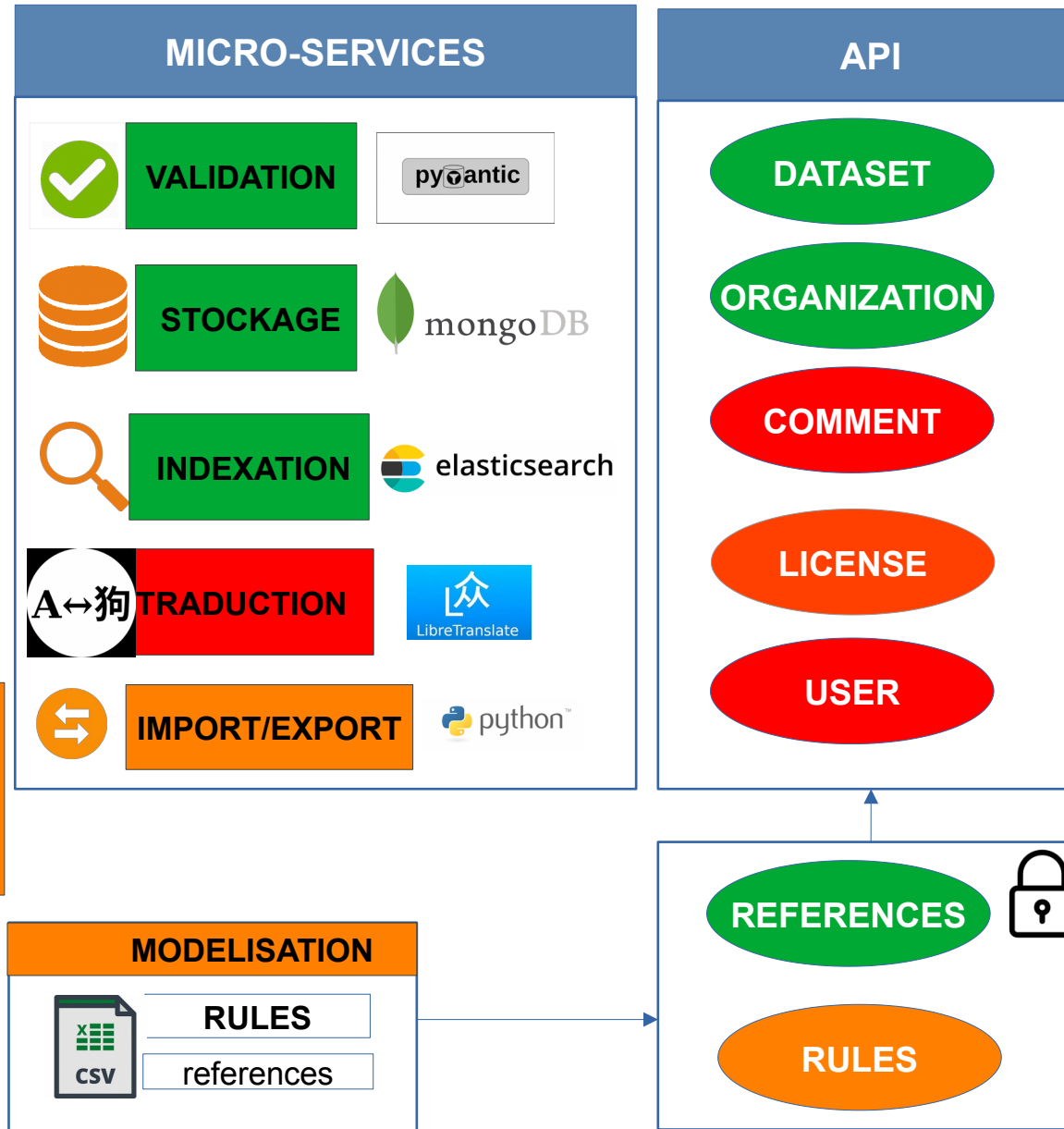
Module de traduction  
désinstallé

### Partiellement en suspens:

- export DCAT AP
- import multiple

Tous en attente d'une  
modélisation stabilisée

En cours



### En suspens:

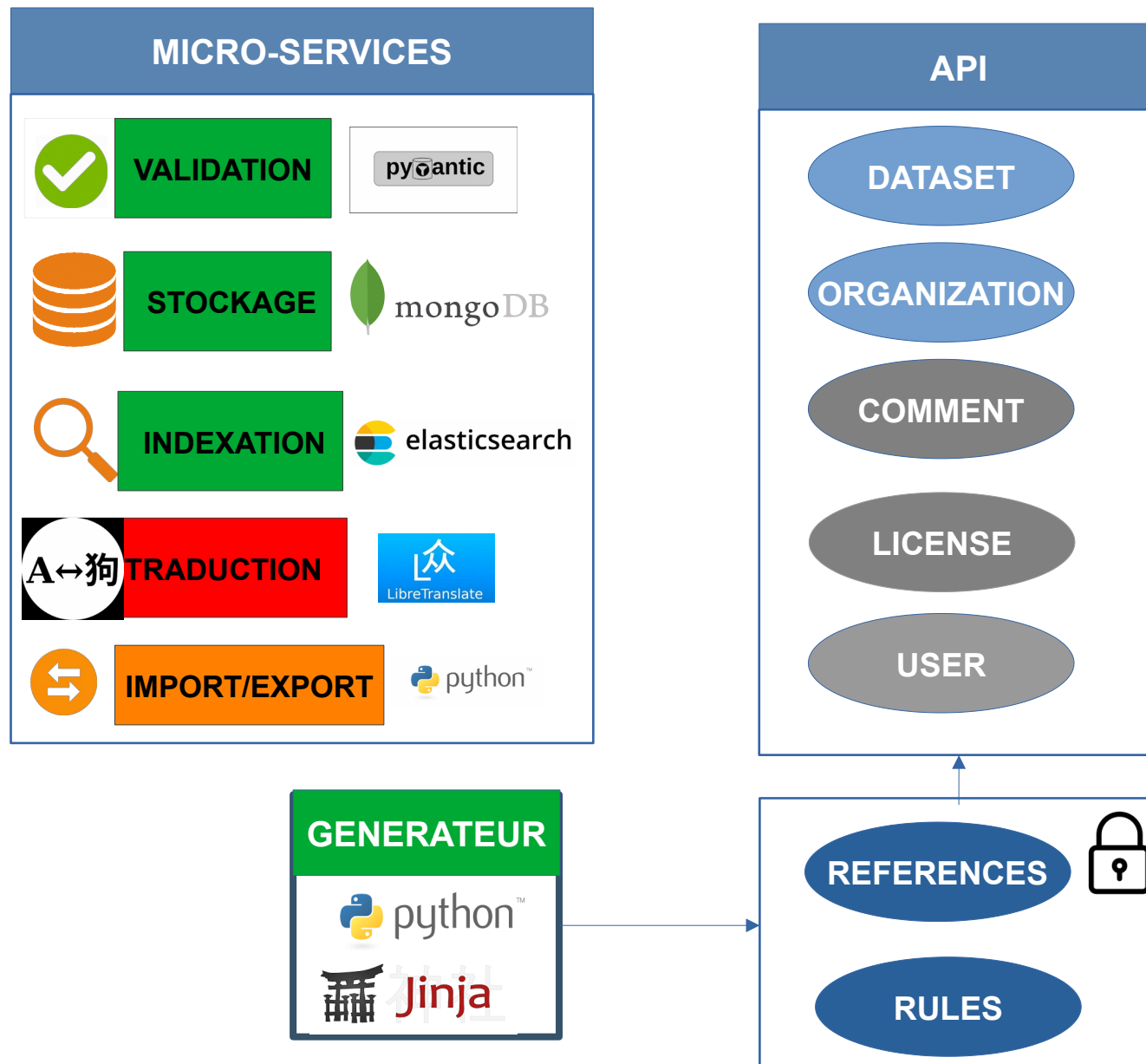
Modèles et endpoints  
disponibles mais non utilisés  
en attente de définition

### Reste à faire:

Synchronisation du schéma  
avec les données dans la base  
et dans l'index.

# PLANS DE TEST

Objectif: couverture de tests  $\geq 50\%$



## TESTS UNITAIRES:

- DATASET et ORGANIZATION:
  - **disponibilité du point API:**  
get, put, post, delete, search, describe, filters
- REFERENCE et RULE:
  - **disponibilité du point API:**  
get, put, post, delete, export

## TESTS d'INTEGRATION

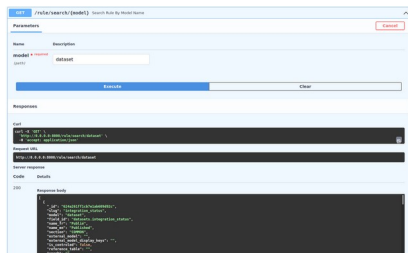
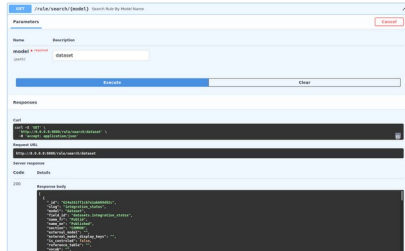
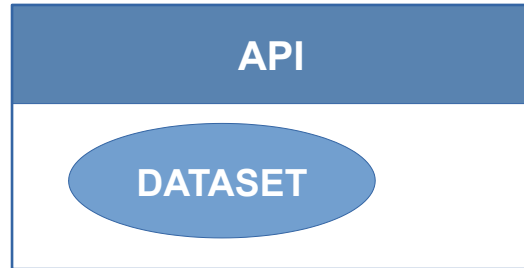
- **consistance** entre les règles/références et les modèles
- **consistance** entre les données et le schéma de données
- **import de données**



● **FOCUS:**

**Fonctionnalités coeur du catalogue**

## FOCUS: CONSULTER UNE RESSOURCE



### FONCTIONNALITE V1

1. L'interface web demande à l'API **tous les champs disponibles du jeu de données:** GET *dataset/describe*, elle génère un gabarit HTML

2. L'interface web demande **tous les jeux de données disponibles** à l'API: GET */dataset/* et remplit les gabarits HTML avec les résultats  
**La page des jeux de données s'affiche**

Les champs descriptifs d'un jeu de données sont paramétrables depuis l'API et dans l'idéal dans une interface d'administration Sans qu'il n'y ait rien à modifier dans l'interface web

## INTERFACE WEB PUBLIQUE (v.1)

Affichage de la liste des JDD

Affichage de la fiche du jeu de données

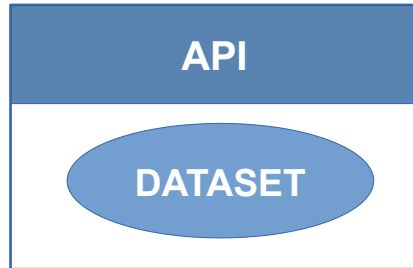
IMT Nord Europe   Données non ouvertes   Non téléchargeable en ligne		
Milieux :	Couverture spatiale	Couverture temporelle
Dépôts atmosphériques (composés inorganiques majeurs, métaux lourds, HAP), les composés gazeux (O3, NO2, Composés Organiques Volatils) et particulaires (gravimétrie des PM10 et PM2.5, métaux lourds et HAP dans les PM10, spéciation des PM2.5) et les paramètres météorologiques	Granularité de la couverture spatiale	N/A
	Données géospatialisées	Pas de temps
	Non	N/A

Airparif   Données non ouvertes   Non téléchargeable en ligne		
Milieux :	Couverture spatiale	Couverture temporelle
Moyennes annuelles des concentrations de HAM-HAP et métaux par station	Granularité de la couverture spatiale	N/A
	Données géospatialisées	Pas de temps
	Non	N/A

Airparif   Données non ouvertes   Non téléchargeable en ligne		
Milieux :	Couverture spatiale	Couverture temporelle
Moyennes annuelles des concentrations de HAM-HAP et métaux par station	Granularité de la couverture spatiale	N/A
	Données géospatialisées	Pas de temps
	Non	N/A

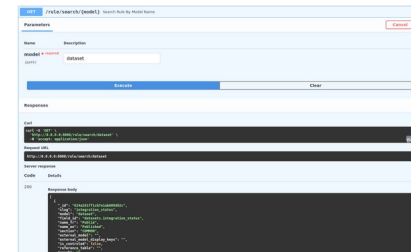
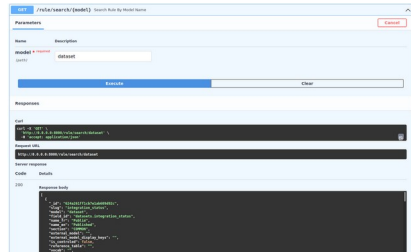
Santé publique France   Données non ouvertes   Non téléchargeable en ligne		
Milieux :	Couverture spatiale	Couverture temporelle
"Association des activités professionnelles et de l'exposition aux métaux avec deux maladies neurodégénératives à partir du Système National des Données de Santé"	Granularité de la couverture spatiale	N/A
	Données géospatialisées	Pas de temps
	Non	N/A

## FOCUS: RECHERCHER UNE RESSOURCE



Rechercher

Rechercher



1. L'interface web expose une barre de recherche sur la page des jeux de données, lorsque l'utilisateur appuie sur le bouton rechercher, l'interface demande à l'API les résultats de la requête: POST dataset/search? query="Chlordécone"

2. L'API lui renvoie les résultats correspondants à la recherche

3. L'interface charge le gabarit HTML en appelant l'API /describe? (Voir le slide précédent: Consulter une resource)

4. L'interface remplit le gabarit avec les résultats qui s'affichent: les résultats de la recherche s'affichent

## FONCTIONNALITE V1

Les champs disponibles à la recherche plein texte sont paramétrables dans l'API et idéalement via une interface ADMIN.

Cela garantit de pouvoir synchroniser les champs du modèle de données (ici Dataset) avec l'index et donc de mettre à jour automatiquement le moteur de recherche

## INTERFACE WEB

Barre de recherche  
Affichage de la liste des résultats

Jeux de données environnementales

métaux lourds

Rechercher

(0 résultat)

Version beta - Cette plateforme est en cours de création

Dans une logique d'amélioration en continu nous sollicitons la communauté des producteurs de données environnementales et des chercheurs en santé environnement dans le but de proposer des améliorations sur la plateforme en ligne et le catalogue. Vous pouvez participer à cette démarche en cliquant ci-dessous pour faire partie des groupes de travail et devenir beta-testeur

Devenir beta-testeur

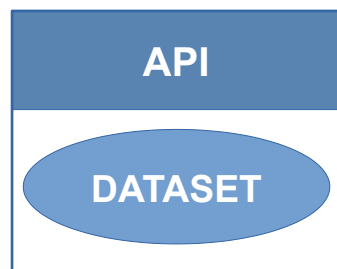
IMT Nord Europe - Données non ouvertes | Non téléchargeable en ligne

Milieu :	Couverture spatiale	Couverture temporelle
Dépôts atmosphériques (composés inorganiques majeurs, métaux lourds, HAP), les composés gazeux (O3, NO2, Composés Organiques Volatils) et particulaires (gravimétrie des PM10 et PM2.5, métaux lourds et HAP dans les PM10, spéciation des PM2.5) et les paramètres météorologiques	Granularité de la couverture spatiale	Pas de temps
	Données géospatialisées	N/A
	Non	

Airparif - Données non ouvertes | Non téléchargeable en ligne

Milieu :	Couverture spatiale	Couverture temporelle
Moyennes annuelles des concentrations de HAP-HAP et métaux par station	Granularité de la couverture spatiale	Pas de temps
	Données géospatialisées	N/A
	Non	

# FOCUS: FILTRER UNE RESSOURCE

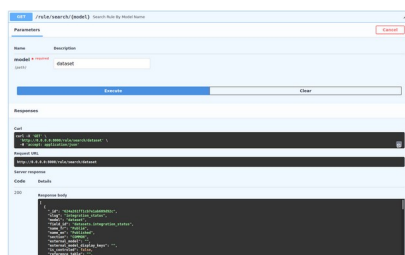
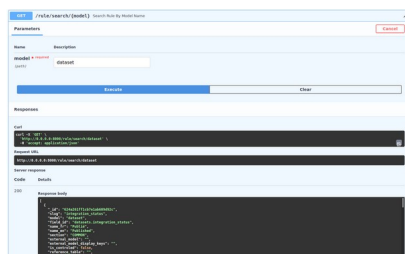


1. L'interface web propose des **filtres pour un jeu de données** en demandant à l'API les filtres disponibles: GET /filters
2. L'utilisateur remplit la barre de recherche et/ou le menu de filtres et clique sur le **bouton filtrer** l'interface demande à l'API POST /search

3. L'API exécute la requete renvoie **les résultats de recherche**

4. L'interface construit **le gabarit HTML** d'affichage en demandant les valeurs à l'API via GET /describe

- 5 Les résultats de la recherche sont insérés dans le gabarit HTML: **les résultats sont affichés sur la page web.**

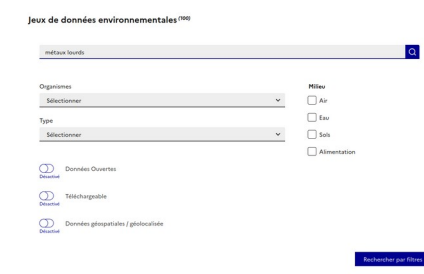


**FONCTIONNALITE à arbitrer**

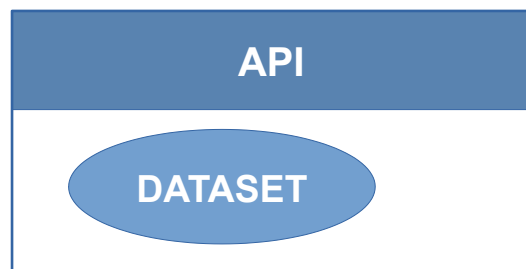
Les champs descriptifs d'un jeu de données et les filtres sont paramétrables depuis l'API et dans l'idéal dans une interface d'administration les changements sont automatiquement répercutés sur l'interface

## INTERFACE WEB

- Affichage du menu des filtres
- Résultats de recherche et filtres



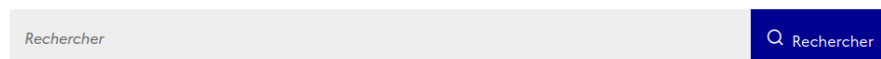
## FOCUS: RECHERCHER ET FILTRER UNE RESSOURCE



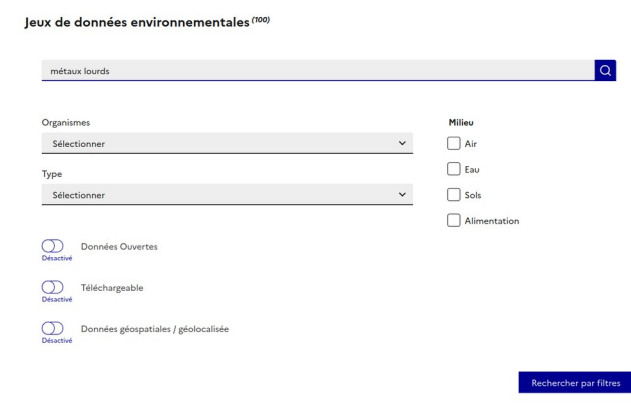
### INTERFACE WEB

- Barre de recherche + filtre
- 1 seul **bouton**
- Résultats de recherche et filtres

L'interface met à disposition **un seul bouton** rechercher qui à la soumission prene les valeurs de la barre de recherche **et** les valeurs du menu de filtres. Le fonctionnement est le même qu'à la slide précédente "Filtrer une ressource"



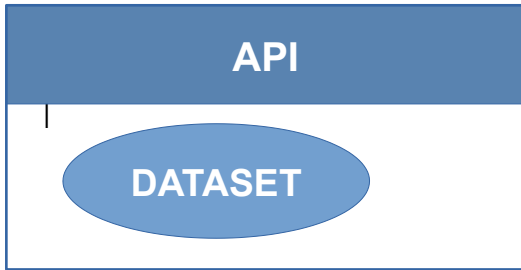
+



### FONCTIONNALITE à arbitrer

Les champs descriptifs d'un jeu de données sont paramétrables depuis l'API et dans l'idéal dans une interface d'administration Sans qu'il n'y ait rien à modifier dans l'interface web

## ● FOCUS: AJOUTER UNE RESSOURCE



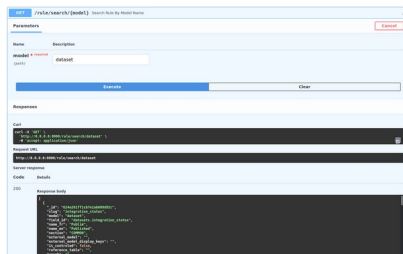
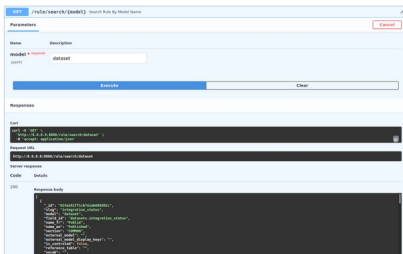
1. L'interface web demande à l'API tous les **champs descriptifs d'un jeu de données**: GET *dataset/describe*, elle génère un gabarit HTML de formulaire

2. L'interface web remplit ensuite le gabarits HTML du formulaire avec les résultats: **un formulaire d'ajout de JDD s'affiche**

3. Quand l'utilisateur appuie sur ajouter/soumettre le jeu de données, **l'interface envoie à l'API la demande d'ajout POST /dataset**

4. L'API vérifie la validité du jeu de données via l'API:  
=> VALIDE 5. Le Jeu de données est inséré dans la **BDD, traduit, indexé et puis disponible dans l'interface**

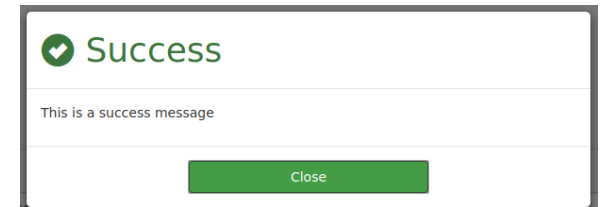
=> INVALIDE 6. Le Message d'erreur s'affiche dans le formulaire pour corriger et soumettre à nouveau  
Le Jeu de données est inséré dans la **BDD "en attente de validation"** et s'affiche dans l'interface administrateur pour traitement



## INTERFACE WEB

### Formulaire d'ajout

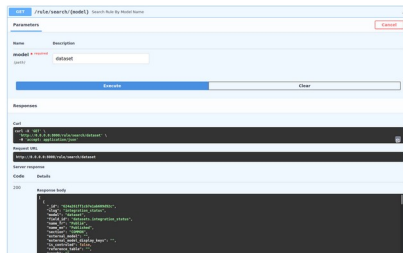
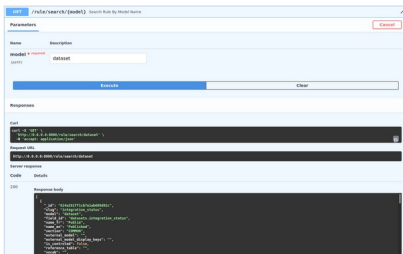
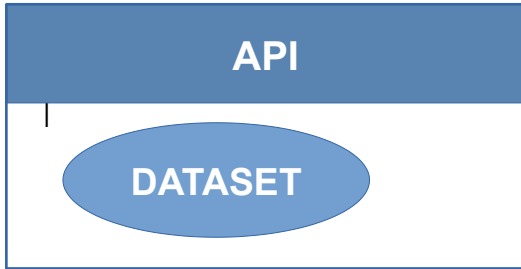
The screenshot shows a web form titled 'Jeux de données environnementales'. It includes a search bar, a dropdown menu for 'Organismes', a dropdown menu for 'Type', and a section for 'Métadonnées' with checkboxes for 'Air', 'Eau', 'Sol', and 'Alimentation'. There are also links for 'Données Ouvertes', 'Statistiques', and 'Données géographiques / géoservices'. A 'Soumettre' button is at the bottom right.



The screenshot shows a 'Sample Form' with two input fields. The 'Fax Number' field contains '111-111-1111' and has a red border with the error message 'Invalid input. Please enter the fax number in the form xxx-xxx-xxxx'. The 'Phone Number' field contains '111 111 1111' and has a yellow border with the error message 'Invalid input. Please enter the phone number in the form (xxx) xxx-xxxx'. A 'Submit' button is at the bottom.

Les champs descriptifs d'un jeu de données  
sont paramétrables depuis l'API  
et dans l'idéal dans une interface d'administration  
Sans qu'il n'y ait rien à modifier dans l'interface web

## ● FOCUS: MODIFIER UNE RESSOURCE



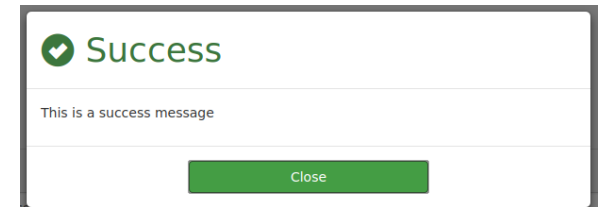
1. L'interface web demande à l'API tous les **champs descriptifs d'un jeu de données**: GET *dataset/describe*, elle génère un gabarit HTML de formulaire
2. L'interface demande ensuite à l'API **les informations sur le jeu de données**: GET *dataset/<id>*
3. L'interface web remplit ensuite le gabarits HTML du formulaire avec les données du JDD: **un formulaire d'édition de JDD s'affiche**
3. Quand l'utilisateur appuie sur ajouter/soumettre le jeu de données, **l'interface envoie à l'API la demande d'ajout POST /dataset**
4. **L'API vérifie la validité du jeu de données via l'API:**  
=> VALIDE
5. Le Jeux de données est mis à jour dans la **BDD, traduit, indexé et puis affiché dans l'interface**  
=> INVALIDE
6. Un Message d'erreur s'affiche dans le formulaire pour corriger et soumettre à nouveau  
Le Jeux de données est inséré dans la **BDD "en attente de validation"** et s'affiche dans l'interface administrateur pour traitement

FONCTIONNALITE à arbitrer

Les champs descriptifs d'un jeu de données  
sont paramétrables depuis l'API  
et dans l'idéal dans une interface d'administration  
Sans qu'il n'y ait rien à modifier dans l'interface web

## INTERFACE WEB

### Formulaire d'édition



# ● FOCUS

## Fonctionnalités avancées du catalogue

**A VENIR ...**