

schema.data.gouv.fr

Mars 2022

Equipe data.gouv.fr

☀️ Le programme

Les missions

Les produits

Le fonctionnement

Les standards



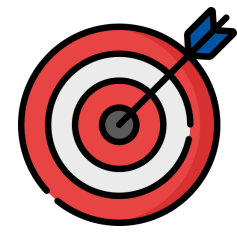
Définition

Lexique : Schémas de données

Les schémas de données permettent de décrire la structure d'un jeu de données. Ils indiquent clairement quels sont les différents champs, comment sont représentées les données, quelles sont les valeurs possibles etc.

Synonymes : *modèle de données, modèle logique de données, schéma.*

Les missions



Les missions

L'initiative schema.data.gouv.fr est portée par l'équipe data.gouv.fr d'**Etalab**

Accompagnement et animation de communautés

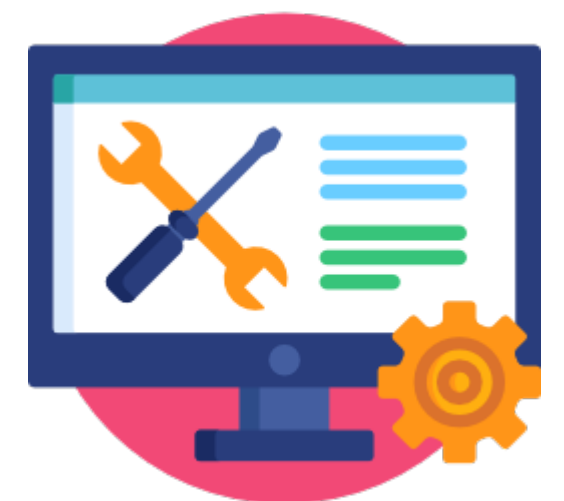
Nous accompagnons les **producteurs de schémas** dans la conception de schémas de données.

Nous accompagnons les **producteurs de données** dans la production de leurs données de qualité



Développement et maintien d'outils

Nous développons un certain nombre d'outils (schema.data.gouv.fr, publier.etalab.studio, data.gouv.fr ...) permettant l'amélioration de la qualité des données publiées en open data.



Les produits



Les produits

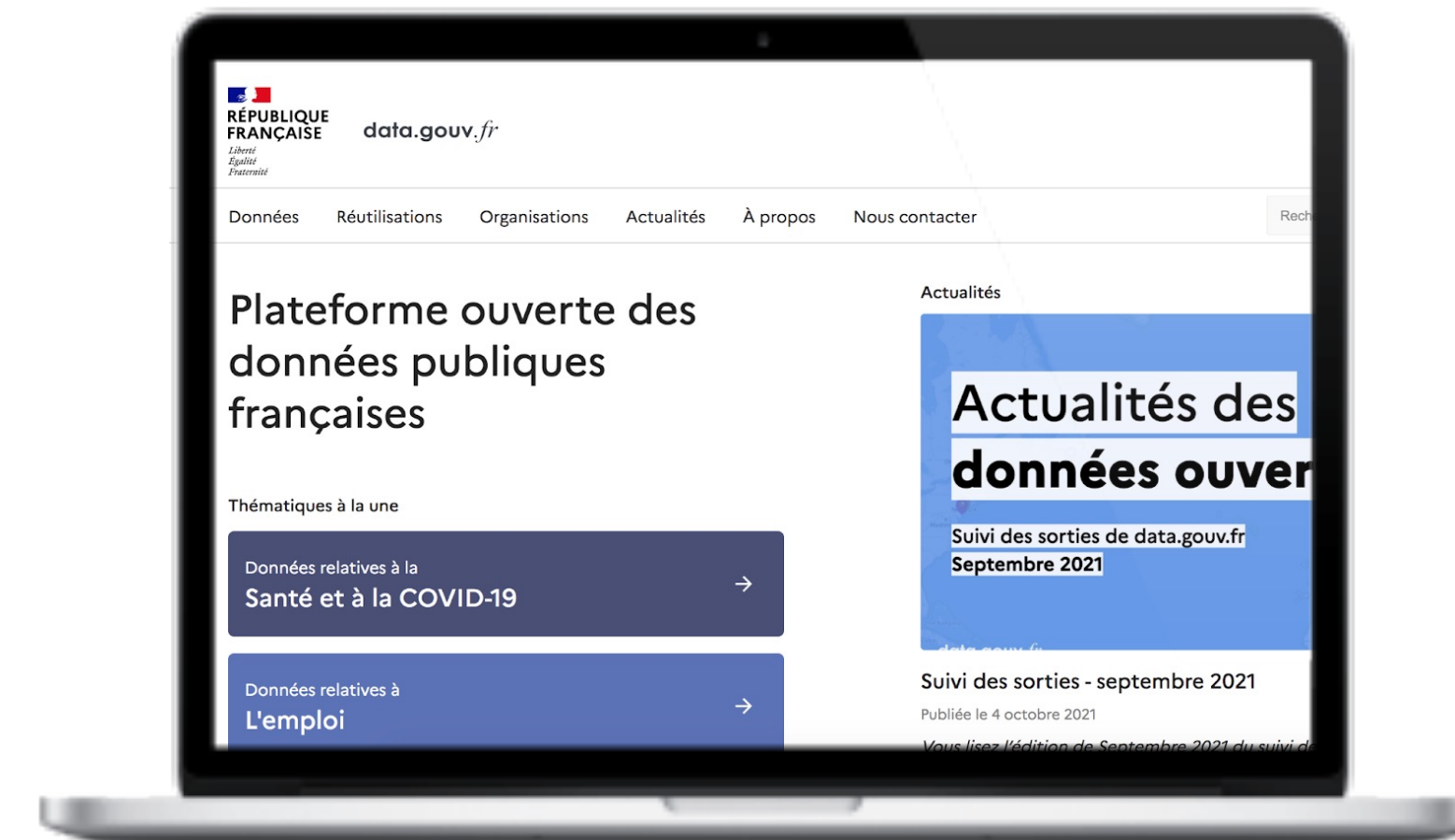
schema.data.gouv.fr



publier.etalab.studio



data.gouv.fr



Producteurs de schémas



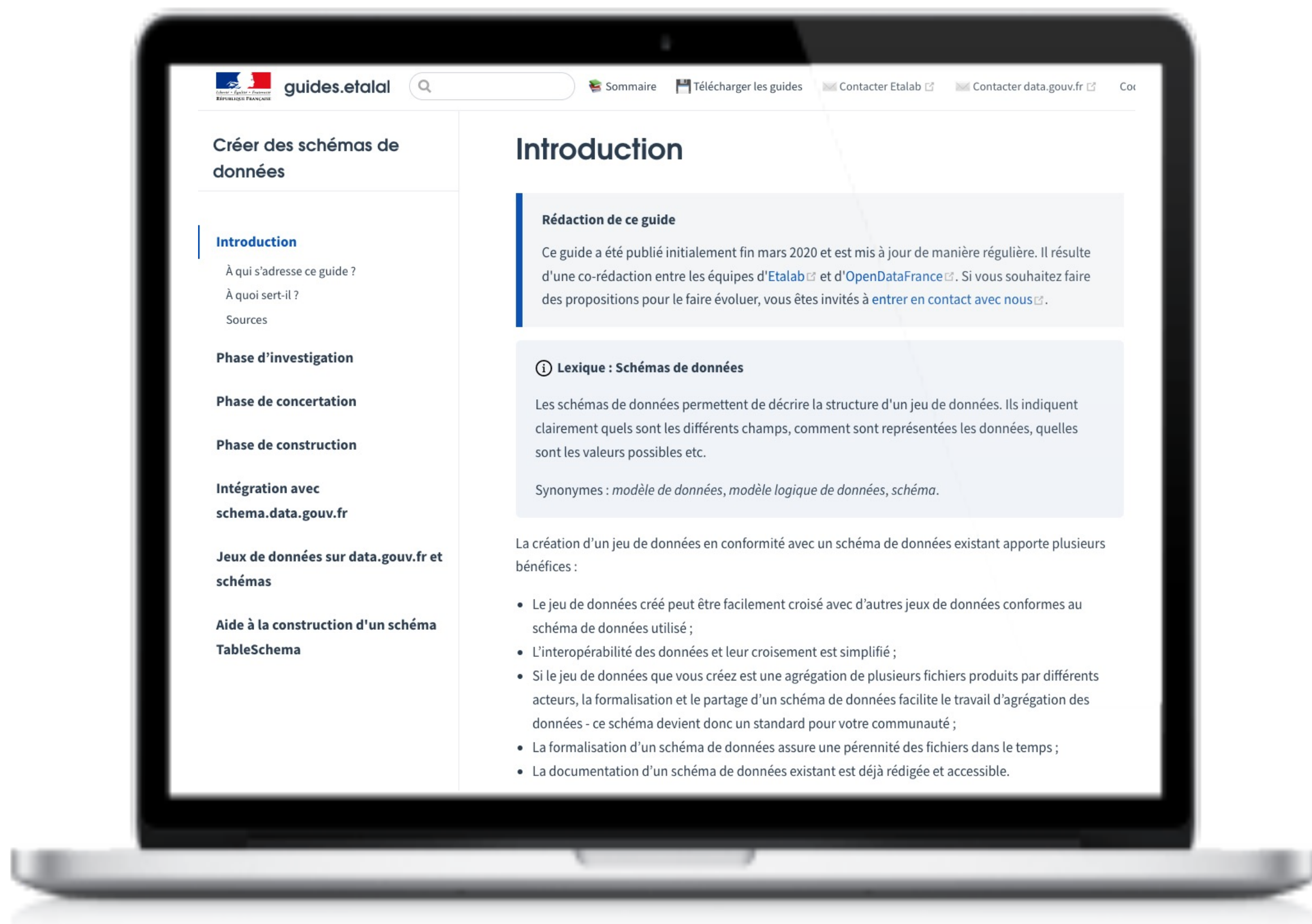
Producteurs de données



Utilisateurs open data



schema.data.gouv.fr



Guide pour accompagner une organisation dans la conception d'un nouveau schéma de données.

Co-rédigé avec OpenDataFrance, ouvert à la contribution.

<https://guides.etalab.gouv.fr/producteurs-schemas/>



schema.data.gouv.fr



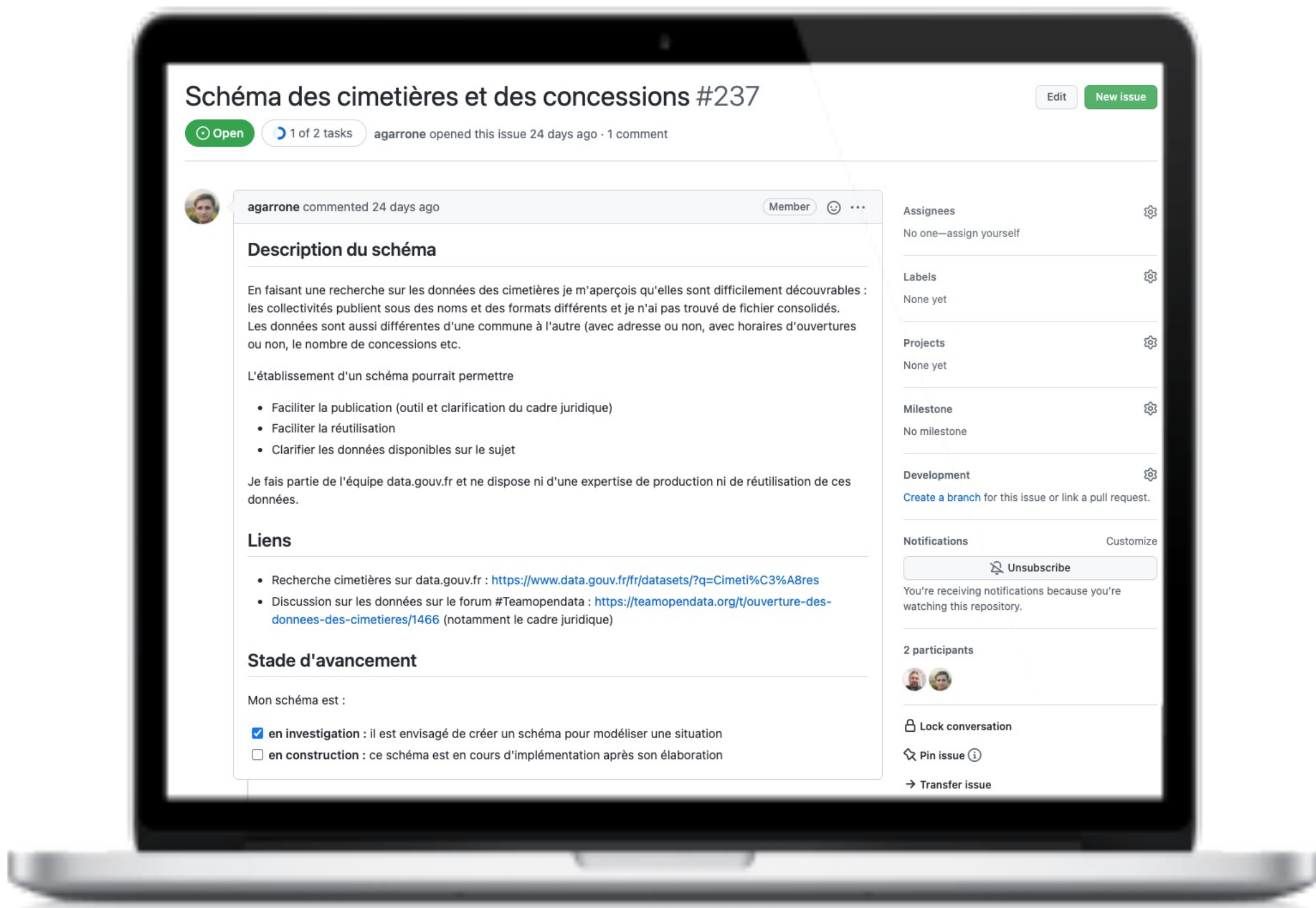
Page « Contribuer » sur schema.data.gouv.fr

Permet d'accompagner un producteur de schéma vers la ressource la plus adéquate en fonction de l'étape de maturité du schéma sur lequel il travaille.

<https://schema.data.gouv.fr/contribuer.html>



schema.data.gouv.fr



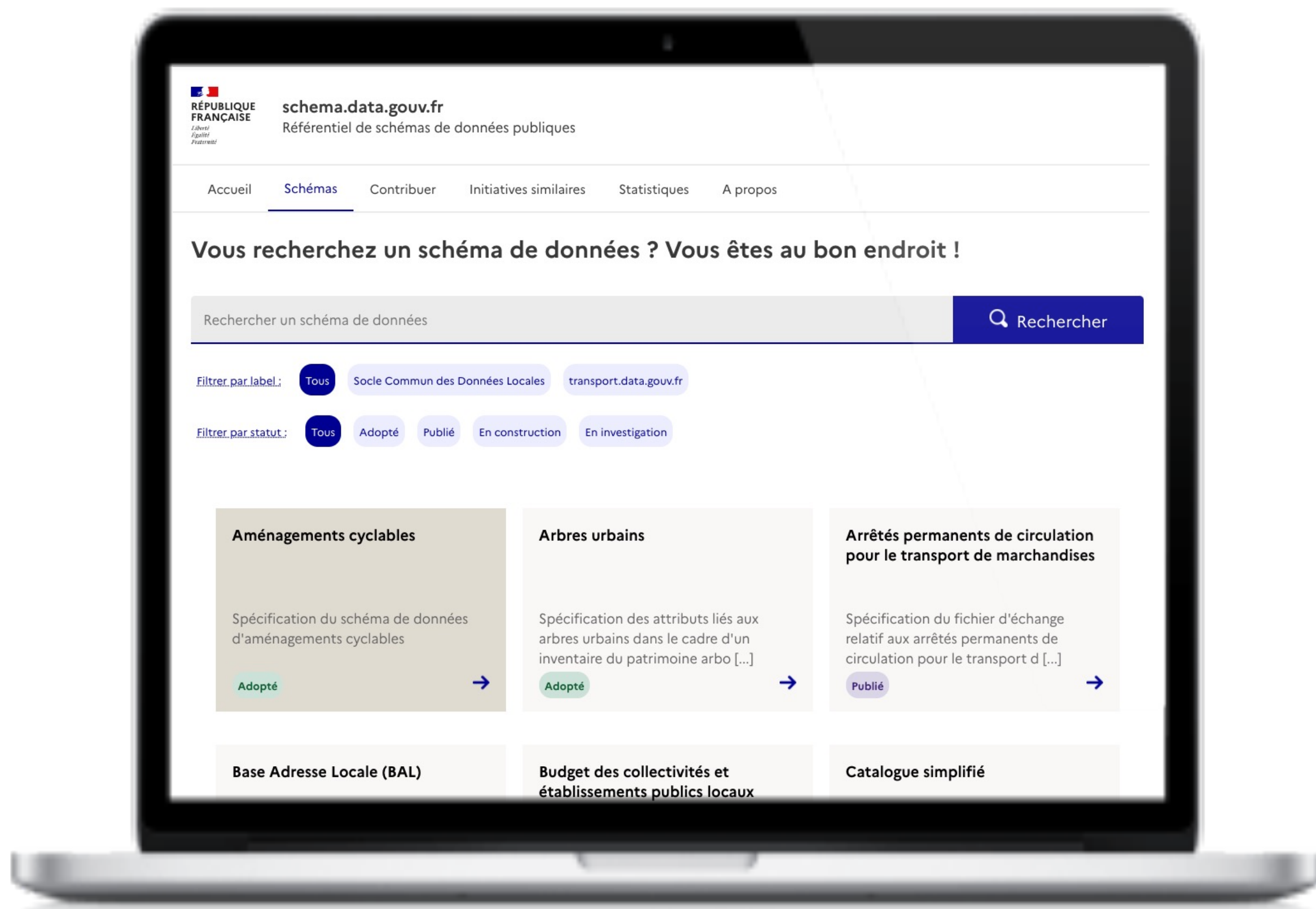
Espace de discussion sur le repo du code source de schema.data.gouv.fr

Lieu où se déclare l'idée d'un nouveau schéma, et où la communauté peut réagir et indiquer son souhait de faire partie ou non du groupe de travail pour un schéma donné.

<https://github.com/etalab/schema.data.gouv.fr/issue>



schema.data.gouv.fr



Page « Schémas » sur schema.data.gouv.fr

Visualisation de tous les schémas référencés et possibilité d'accès à la documentation pour chacun d'entre eux.

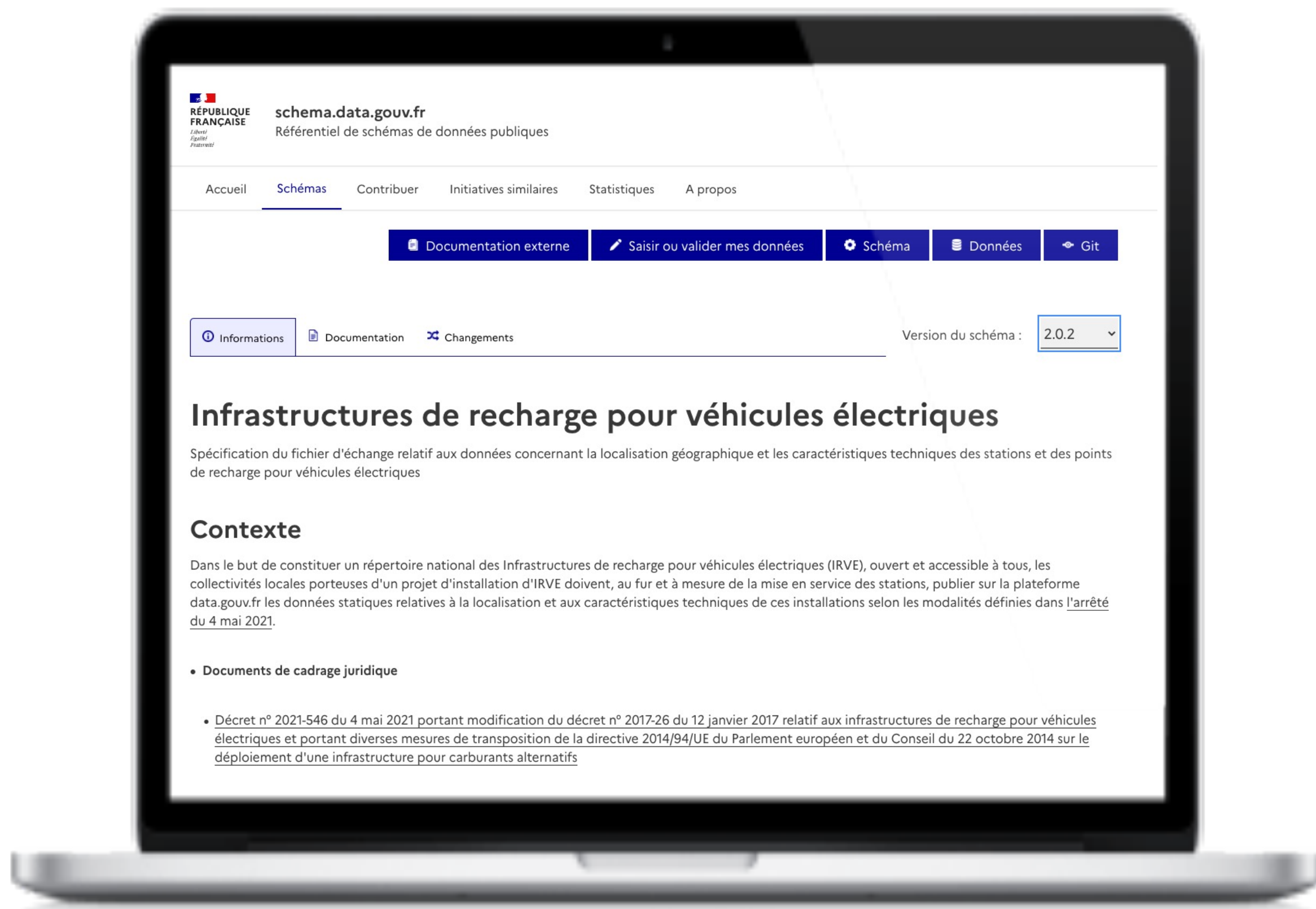
Possibilité de trier par statut du schéma.

Possibilité de filtrer par « label » (regroupement d'une sous-partie des schémas référencés sur un thème donné, ex: les transports, les collectivités locales...)

<https://schema.data.gouv.fr/schemas.html>



schema.data.gouv.fr



<https://schema.data.gouv.fr/schemas.html>

Documentation pour un schéma spécifique.

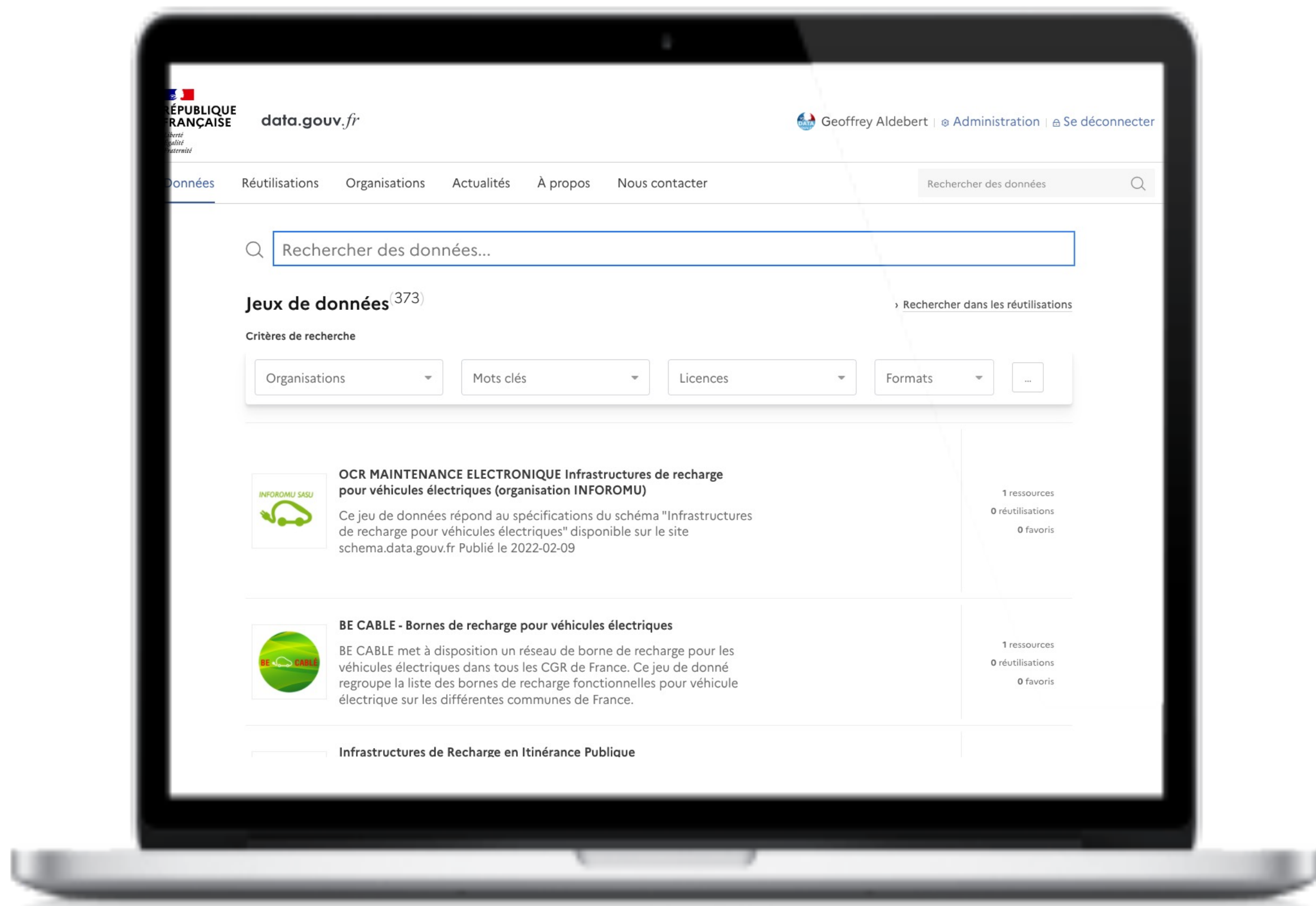
Accès à une page de présentation du schéma et à sa documentation technique (description champs par champs).

Accès aux différentes versions du schéma qui évoluent au fil du temps.

Boutons d'accès à une éventuelle documentation externe, à des outils de saisies de données (s'ils existent), aux données obéissant à ce schéma sur data.gouv.fr ou au repo Git source du schéma.



schema.data.gouv.fr

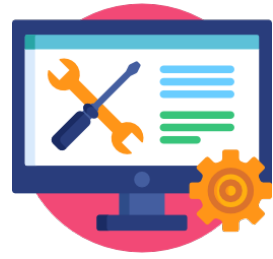


Consultation des jeux de données obéissant à un schéma sur data.gouv.fr

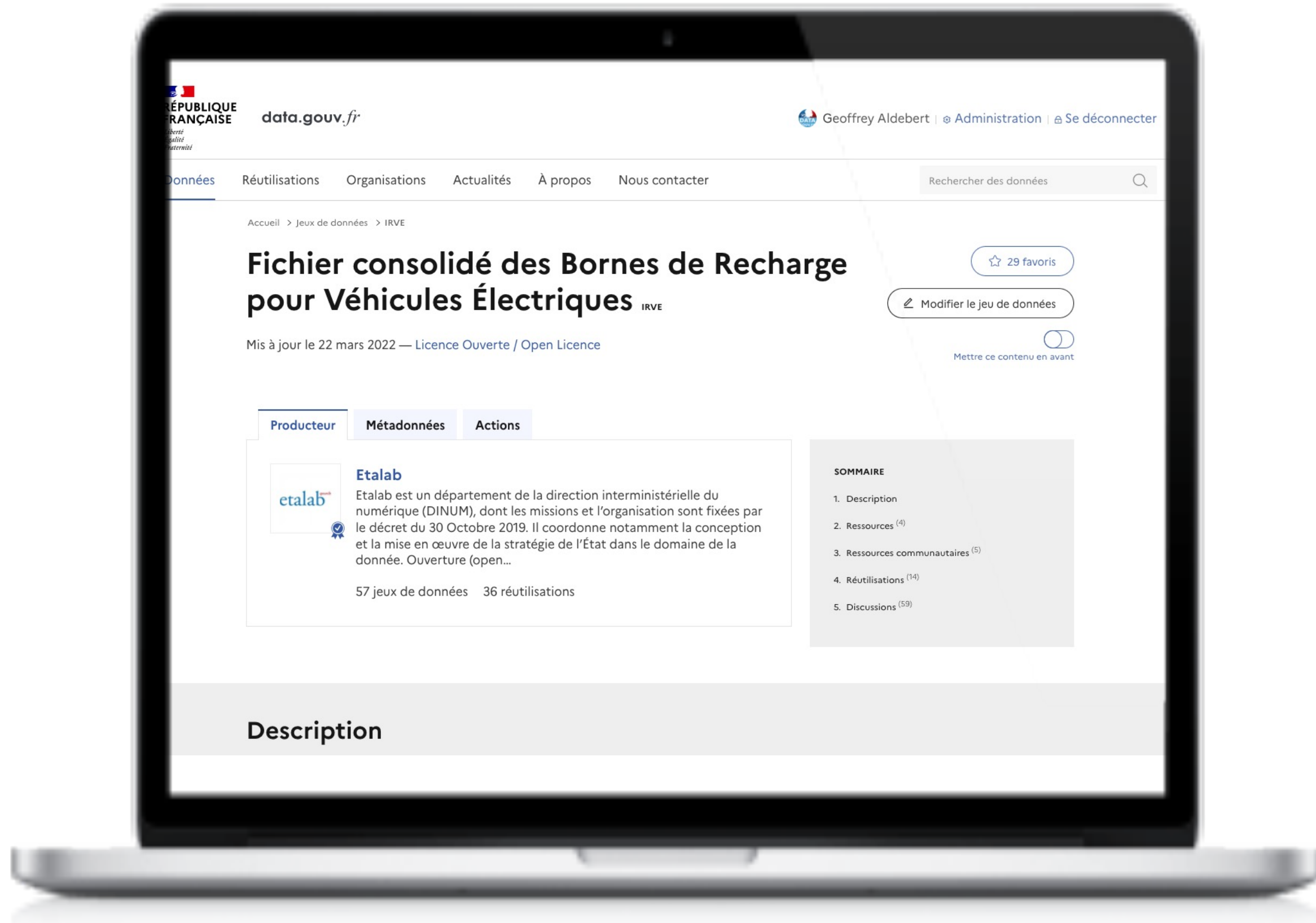
Les données publiées sur data.gouv.fr peuvent être liées aux schémas référencés sur schema.data.gouv.fr

Les analyses, tris, recherches parmi les données obéissant à un schéma s'en trouvent facilités

<https://www.data.gouv.fr/fr/datasets/?schema=etalab%2Fschema-irve>



schema.data.gouv.fr



Consolidation automatique des données obéissant à un schéma de donnée.






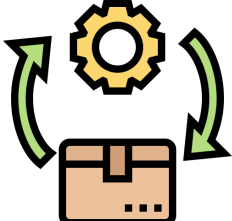
Cela permet de constituer une base nationale regroupant l'ensemble des données obéissant à un schéma et publiées sur data.gouv.fr

<https://www.data.gouv.fr/fr/datasets/fichier-consolide-des-bornes-de-recharge-pour-vehicules-electriques/>

Le fonctionnement



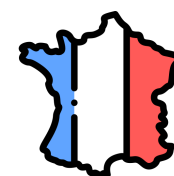
Les étapes pour le référencement d'un nouveau schéma

- 1 Un producteur nous contacte avec une idée plus ou moins mature d'un schéma 
- 2 Une annonce est publiée auprès de la communauté concernant la conception d'un nouveau schéma 
- 3 Des groupes de travail s'organisent avec plusieurs acteurs pour la conception du schéma 
- 4 L'implémentation technique du schéma est réalisée sur un repo Git publique 
- 5 Le schéma est référencé sur schema.data.gouv.fr 
- 6 Le schéma suit sa propre gouvernance avec des montées de versions lorsque nécessaire 



Les prérequis pour le référencement d'un schéma

Le schéma doit décrire des données publiques



Le schéma doit avoir été soumis à la communauté via un message sur notre espace de discussion



Le schéma doit avoir été co-conçu avec plusieurs acteurs du domaine auquel il fait référence



Le schéma doit avoir été conçu selon une spécification technique autorisée et doit être valide



Si aucune donnée n'est encore produite à partir de ce schéma, la version de celui-ci doit être inférieure à la v1.0.0

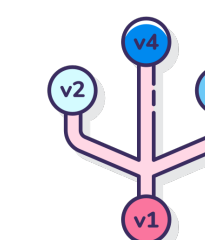




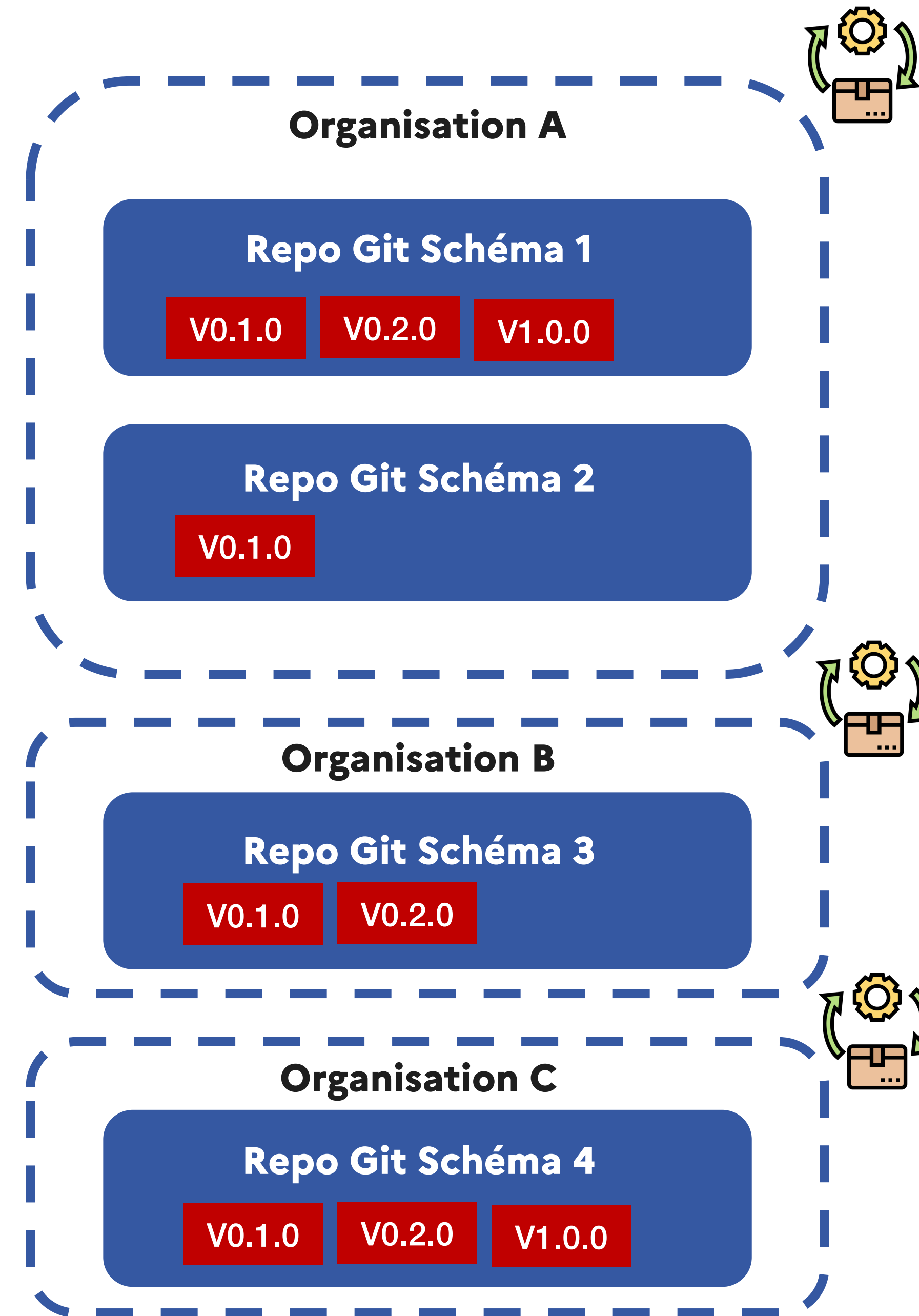
Schéma de fonctionnement

schema.data.gouv.fr



Tous les soirs, schema.data.gouv.fr :

- Récupère l'ensemble des releases publiées des schémas référencés
- génère une documentation automatisée
- met à jour le site



Les standards



Les standards préconisés



**FRICTIONLESS
DATA**

Pour les données tabulaires, le standard tableschema de frictionless est préconisé. Ce type de schéma est particulièrement adapté pour l'open data.

Si un schéma respecte ce standard, l'équipe data.gouv.fr offre des outils complémentaires pour aider à la production de données, la validation ou encore améliorer la documentation.

{JSON}

Pour les données représentant des zones ou points géographiques, le format jsonschema peut être utilisé.



Le format XML avec une spécification XSD peut également être envisagé

...

D'autres formats spécifiques sont envisageables, à discuter au cas par cas avec l'équipe schema.data.gouv.fr.



Les schémas complexes

Jusqu'à présent, c'est en grande majorité des schémas « simples » qui sont référencés sur schema.data.gouv.fr (ex: table d'une base de données)

Pour référencer des schémas complexes (ex: base de données contenant différentes tables), Etalab expérimente actuellement le standard **datapackage** de frictionlessdata permettant de lier plusieurs tables (fichiers csv) entre elles via des clés primaires ou étrangères.

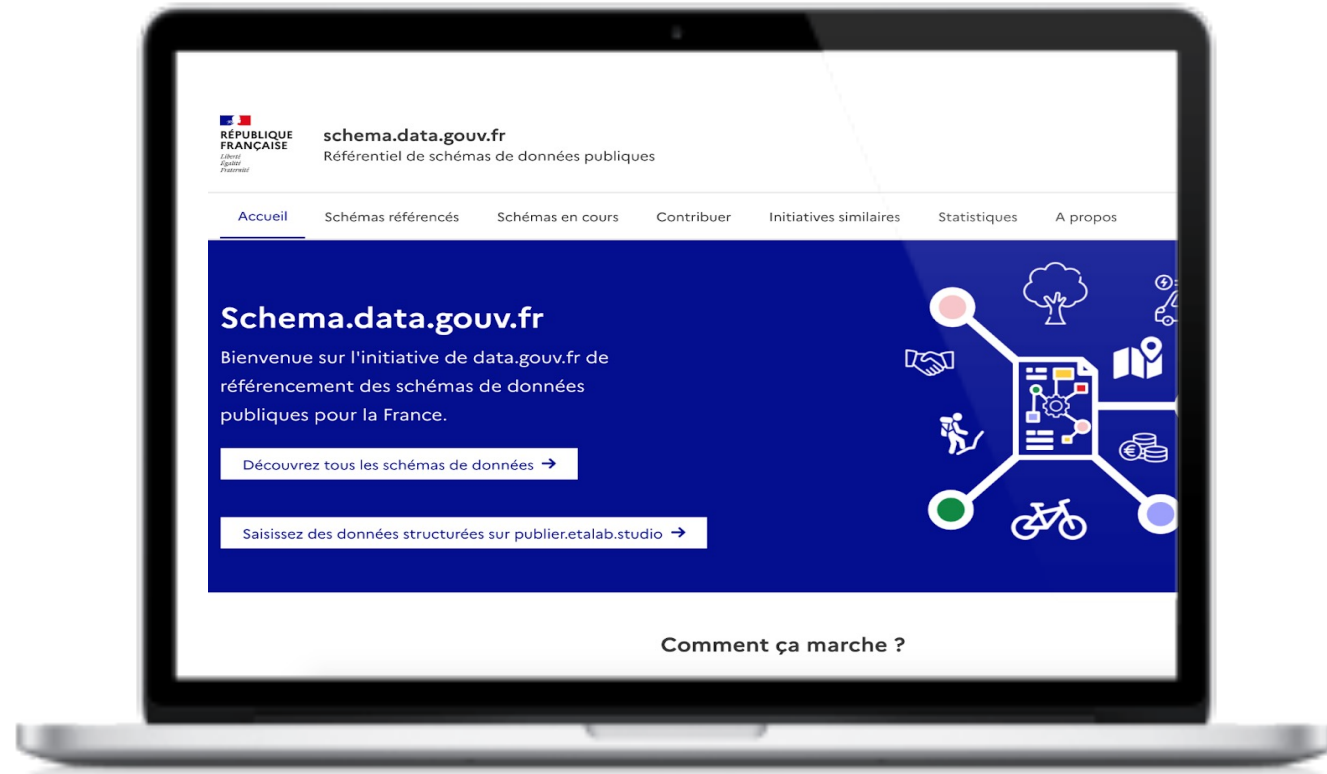


FRICTIONLESS
DATA

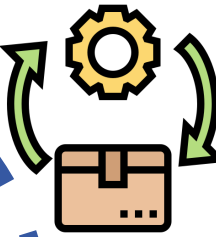


Schéma de fonctionnement data package

schema.data.gouv.fr



Organisation A



Repo Git Schéma Data Package

Table 1 (tableschema)

Table 2 (tableschema)

Table 3 (tableschema)

V0.1.0

V0.2.0

V1.0.0

Vous recherchez un schéma de données ? Vous êtes au bon endroit !

Rechercher un schéma de données

Filtrer par label:

Filtrer par statut:

Schéma Data Package

Spécification du schéma de données d'aménagements cyclables

Adopté



Arbres urbains

Spécification des attributs liés aux arbres urbains dans le cadre d'un inventaire du patrimoine arbo [...]

Adopté



Arrêts permanents de circulation pour le transport de marchandises

Spécification du fichier d'échange relatif aux arrêts permanents de circulation pour le transport d [...]

Publié

