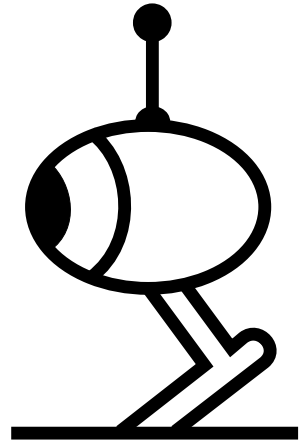


# TRANSPORTS & DONNÉES OUVERTES

TRANSFORMER DES DONNÉES BRUTES EN INFOS UTILES

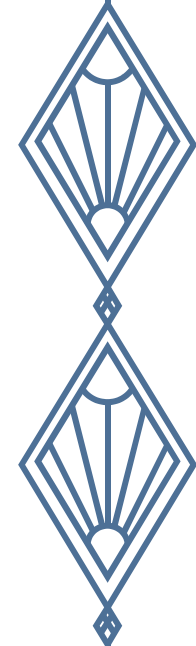




**FRÉDÉRIC BISSON**

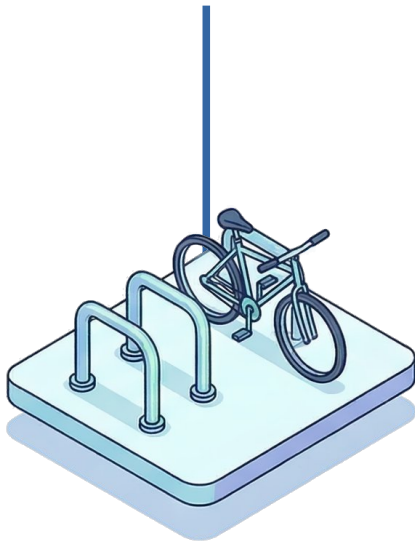
Développeur web à Rouen

Amateur de vieilleries

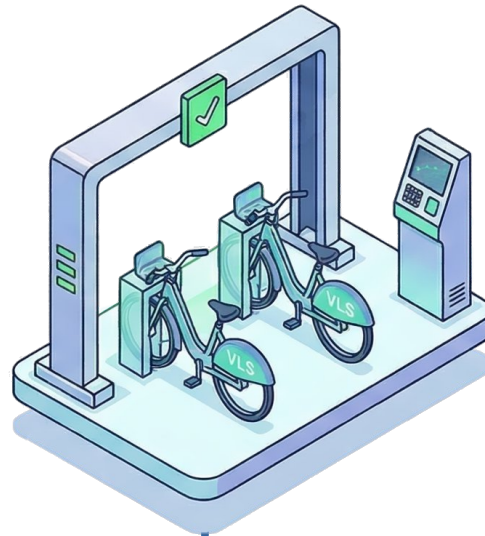
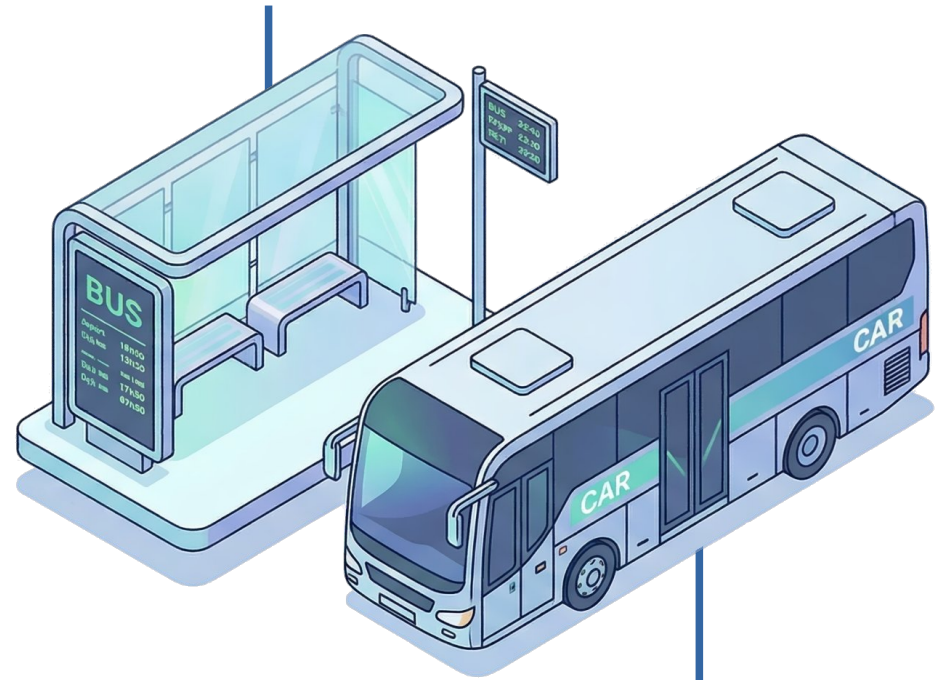


# INDIQUER LA PROXIMITÉ DE...

Arceaux vélo



Arrêts de bus



Vélos en libre-service

Lignes de bus/car



# BUT

## ACCÈS TRANSPORTS

À moins de 300 mètres :

- Lignes 35, 529, 26, F4, T2, T3, T1 (arrêt Mont-Riboudet)
- Lignes 511, 530, 526 (arrêt Rouen - Mont Riboudet - Kindarena)



# BÉNÉFICES ATTENDUS

- Mises à jour globales
- Homogénéité des informations affichées
- Exhaustivité des facilités
- Facilité d'insertion des informations

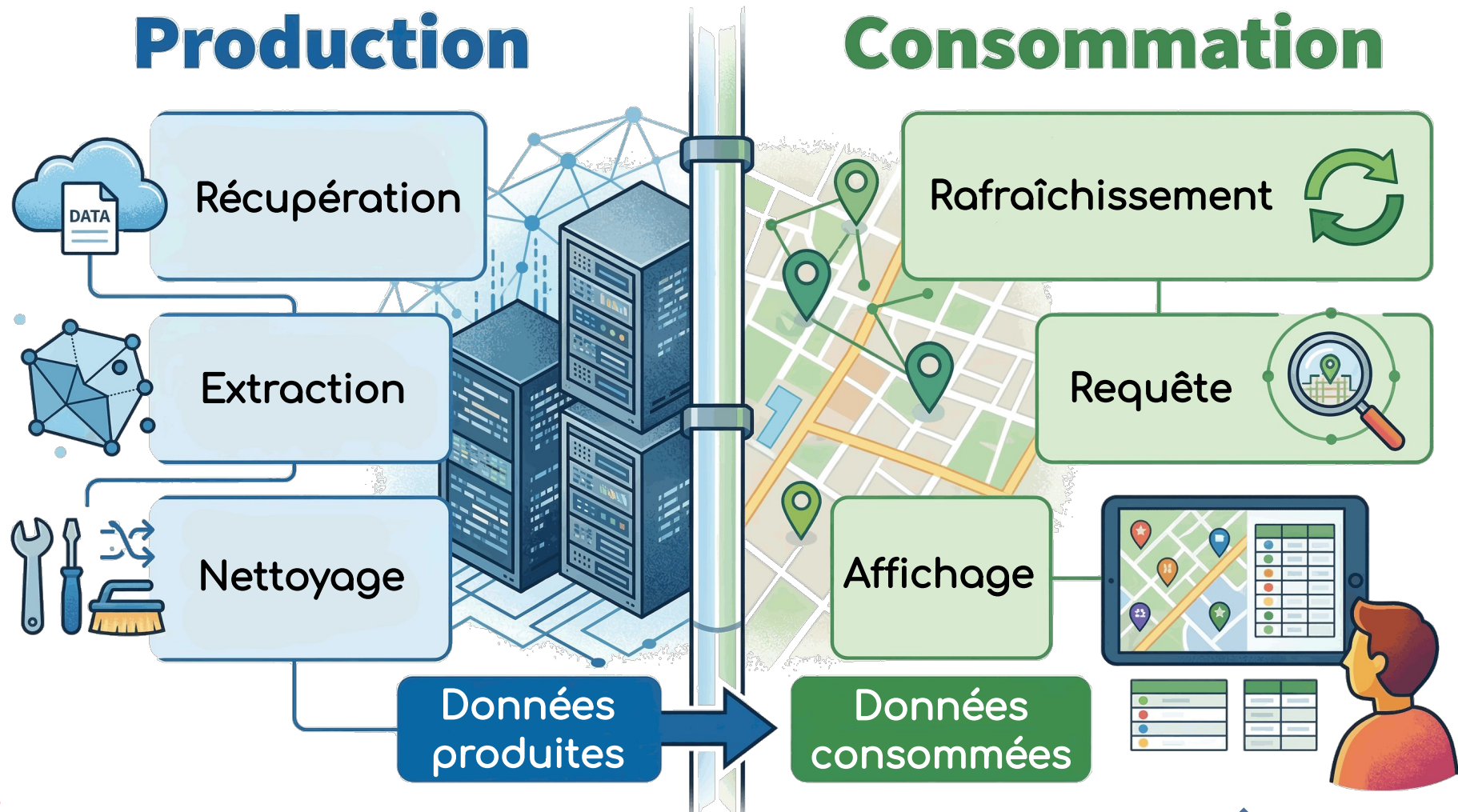


# OBLIGATION DE PUBLICATION

- Règlement européen 2017/1926
  - Mise à disposition des données de mobilités via un PAN
  - Dans des formats normalisés ou interopérables
- Loi LOM (2019)
  - Précise les acteurs soumis à l'obligation
  - Étend l'ouverture à d'autres catégories de données
- Décret n°2020-183
  - [transport.data.gouv.fr](https://transport.data.gouv.fr) est le PAN



# DU PRODUCTEUR AU CONSOMMATEUR



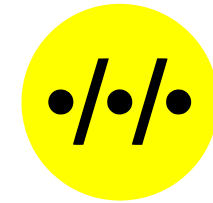
ÉTAPE 1

# PRODUCTION DES DONNÉES

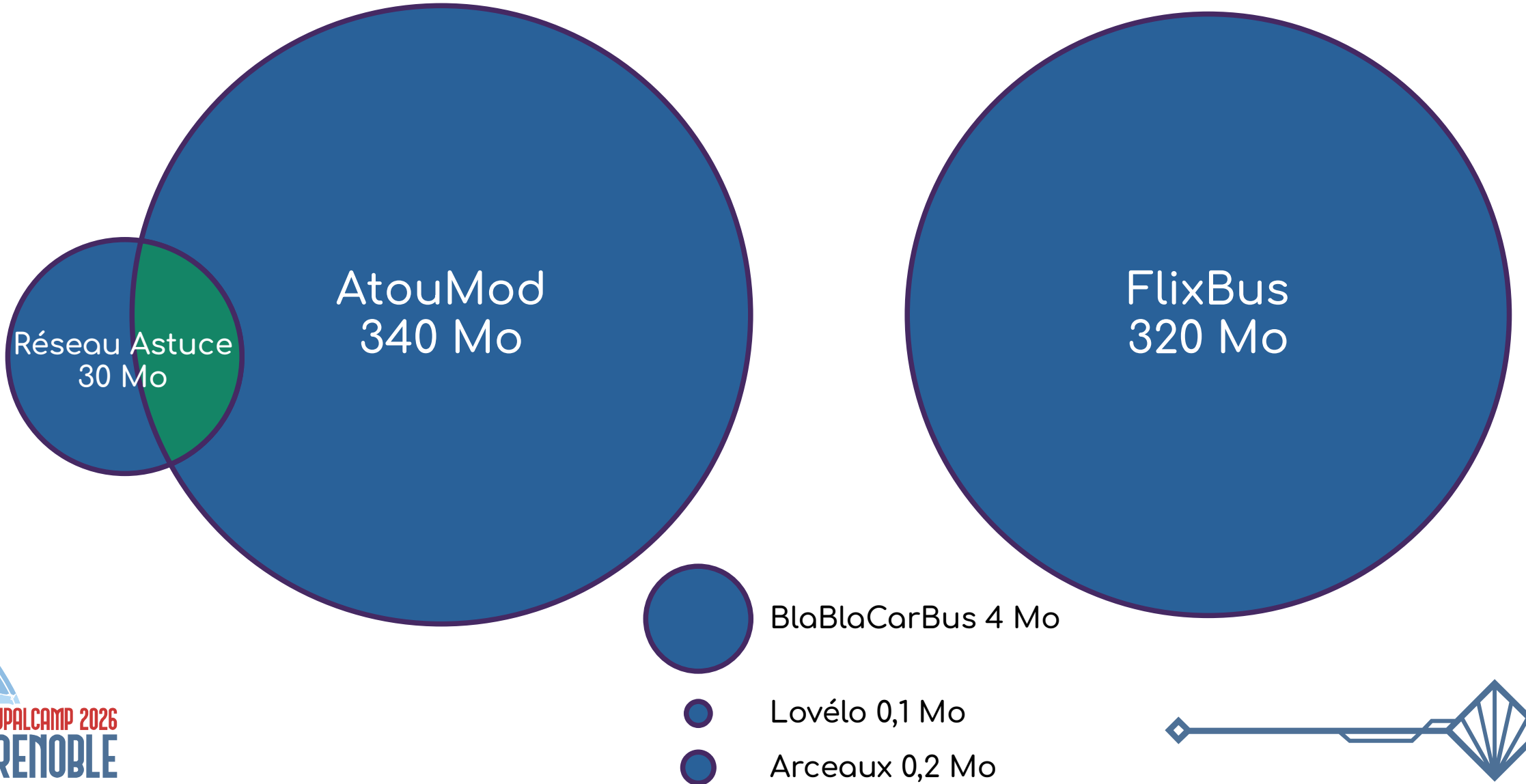


# FORMATS DE DONNÉES

- **GTFS**  
General Transit Feed Specification  
Fichiers CSV dans un ZIP (et un peu de JSON)
- **GBFS**  
General Bikeshare Feed Specification  
Fichier JSON
- **CSV**  
Comma Separated Value  
Pour les arceaux vélo



# VOLUME DES DONNÉES BRUTES



# DONNÉES ORIENTÉES « TRAJETS »

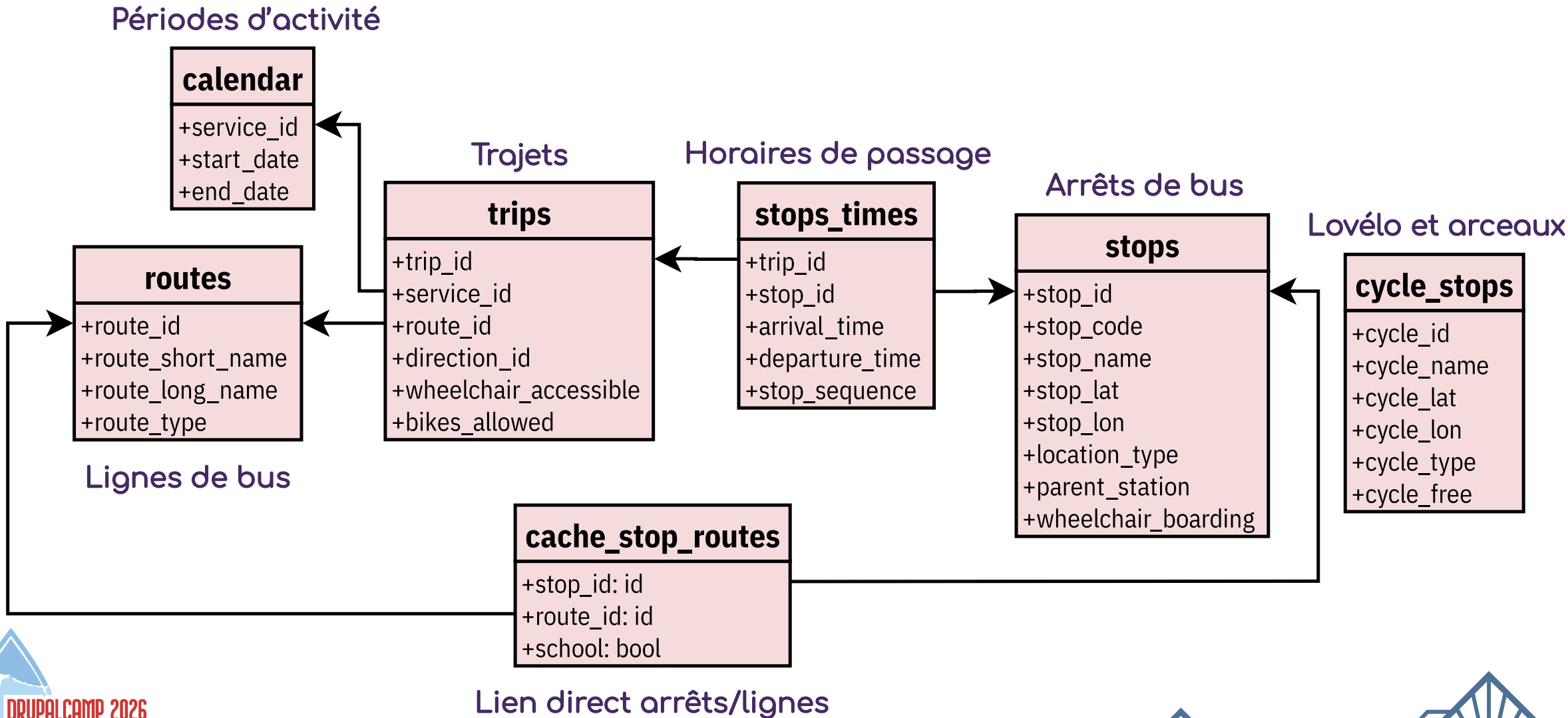


LIGNE	DATE	HEURE	N°	STATION
222	10/04/26	7h55	1	Route de Yainville
222	10/04/26	7h58	2	École de Jumièges
222	10/04/26	8h02	3	Rue Fessard
222	10/04/26	8h04	4	Rue du Quesney
222	10/04/26	8h06	5	La Petite Charbonnière
222	10/04/26	8h07	6	L'Essartel
222	10/04/26	8h09	7	Haute Ville
222	10/04/26	8h14	8	Le Vide du Bois
222	10/04/26	8h16	9	Ancienne École
222	10/04/26	8h18	10	La Trimarée
222	10/04/26	8h19	11	Rue Saint-Jean
222	10/04/26	8h23	12	Abbaye de Jumièges
222	10/04/26	8h25	13	École de Jumièges





# SIMPLIFICATION DU SCHÉMA

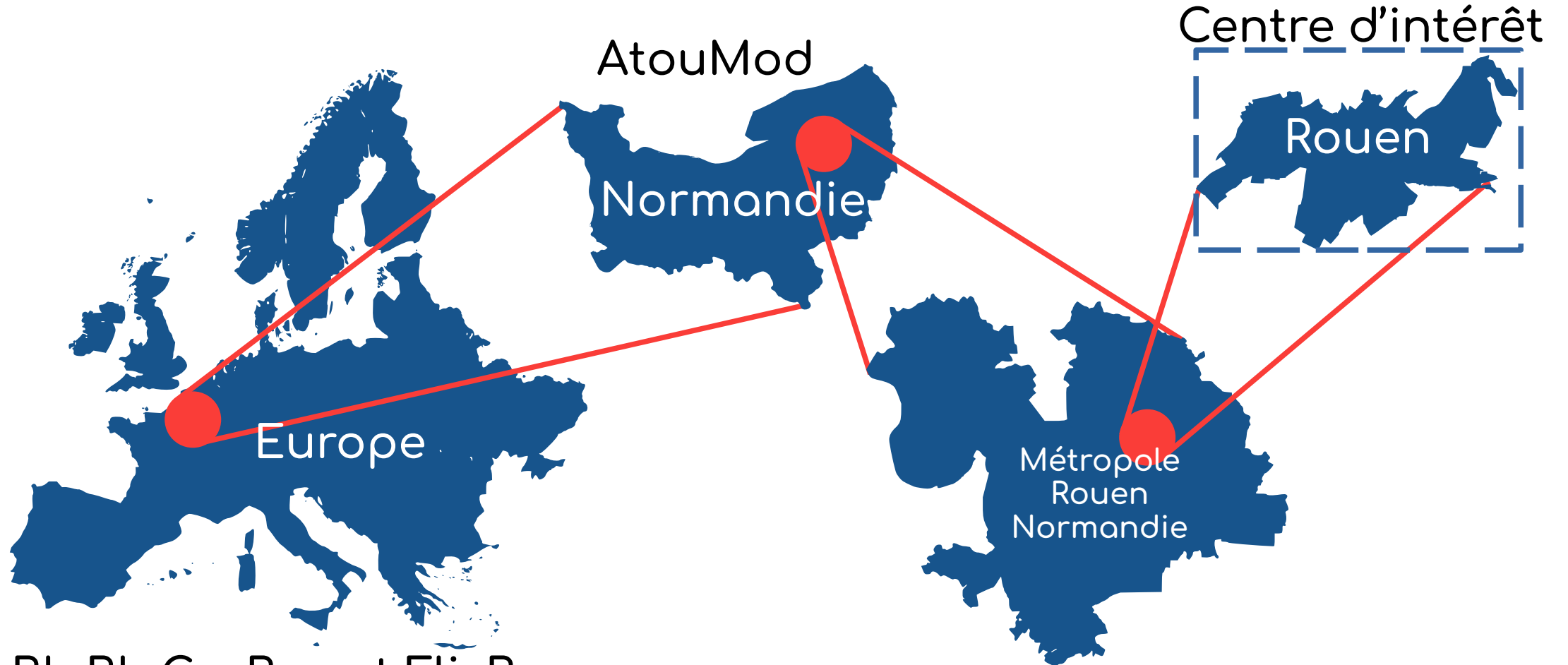


# QUELQUES « SUBTILITÉS » DE GTFS

- Une information → plusieurs formats autorisés  
Par exemple avec les tables `calendar` et `calendar_dates`
- Une station → plusieurs arrêts  
La différence se fait sur les coordonnées GPS
- Des journées de plus de 24 heures  
24:10:00 ou 25:30:00 sont des horaires valides
- Pas d'identifiant universel  
Les identifiants doivent être réécrits pour la fusion



# RESTREINDRE À UNE ZONE GÉOGRAPHIQUE



BlaBlaCarBus et FlixBus

Réseau Astuce et Lovélo

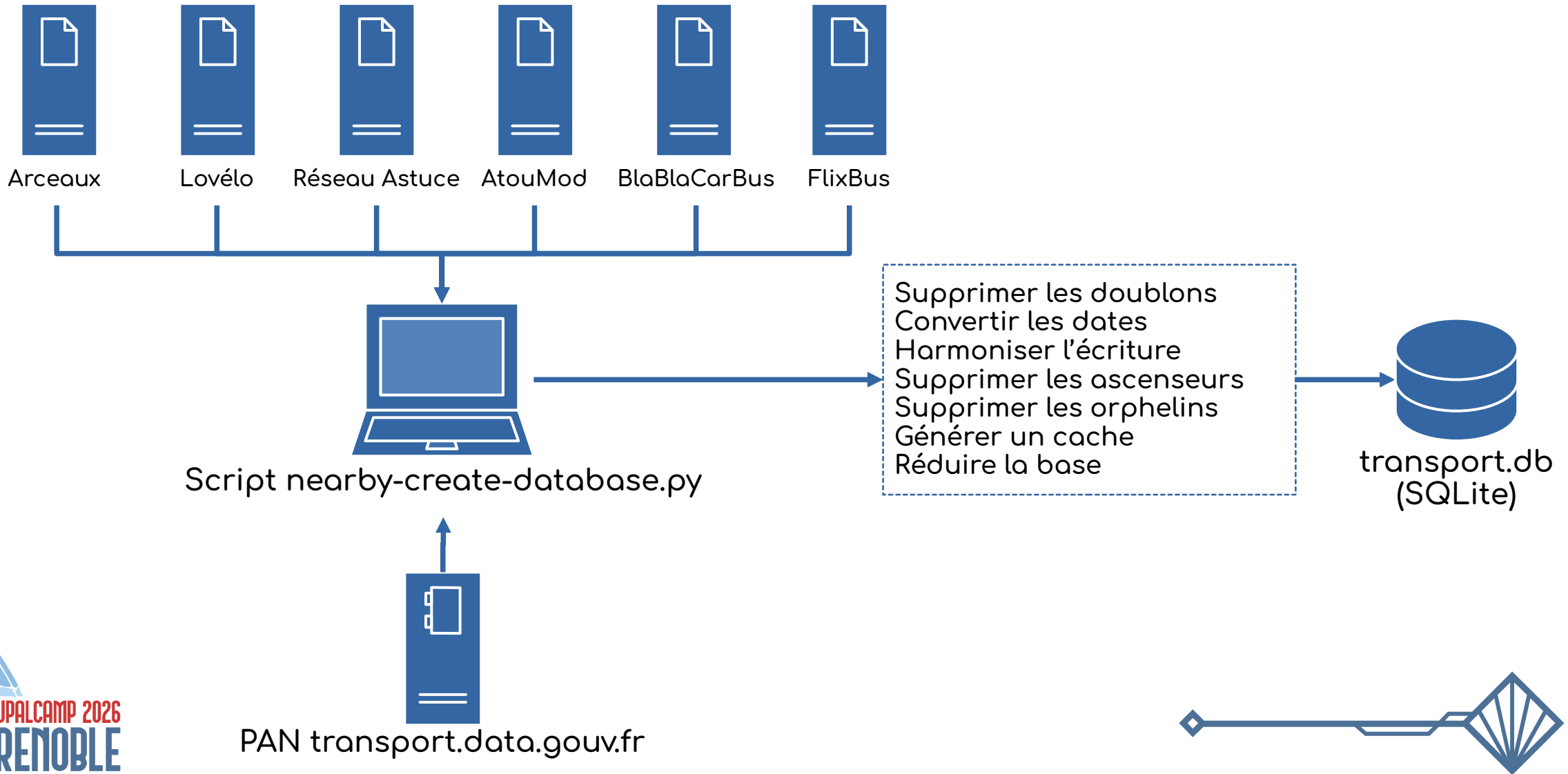


# CORRECTIONS TYPOGRAPHIQUES

- *Salle N. Armstrong* ⇒ Salle Neil Armstrong  
Développement d'abréviations
- *Pont S.N.C.F.* ⇒ Pont SNCF  
Suppression des points dans les sigles
- *VICTOR HUGO* ⇒ Victor Hugo  
Capitales abusives
- *MIN Lecoeur* ⇒ MIN Lecœur  
Accents et ligatures



# PRÉPARATION DES DONNÉES HORS DRUPAL



# NEARBY-CREATE-DATABASE.PY EN ACTION

```
fred@prezigion:~/Programmation/transport-nearby/nearby-create$
```

Tâche automatisable

Facilité de débogage

Peu de dépendances

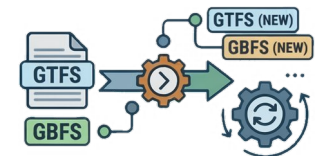
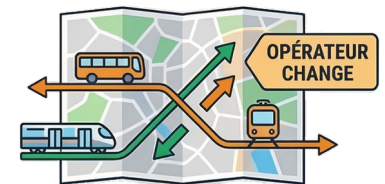
Téléchargement: 27 s

Création de la base: 34 s



# RISQUES LIÉS À LA RÉCUPÉRATION

- Indisponibilité d'une ressource  
Connexion réseau, indisponibilité, droits...
- Intégrité des données  
Cas imprévu, fichier corrompu, données manquantes...
- Instabilité de l'offre de transport  
Création/disparition, changement d'opérateur...
- Évolution des formats utilisés  
GTFS a évolué en septembre 2025 et GBFS fin 2024



# SAVOIR S'ADAPTER



TYPE	ASTUCE	ATOUMOD	FLIXBUS	BLABLACARBUS
Erreurs	0	4 660	0	1
Warnings	102 686	113 617	109 049	7 874
Notice	0	0	4	0

- Valideur GTFS
  - <https://gtfs-validator.mobilitydata.org>
  - Limité à 100000 messages/type




## ÉTAPE 2

# CONSOMMATION DES DONNÉES

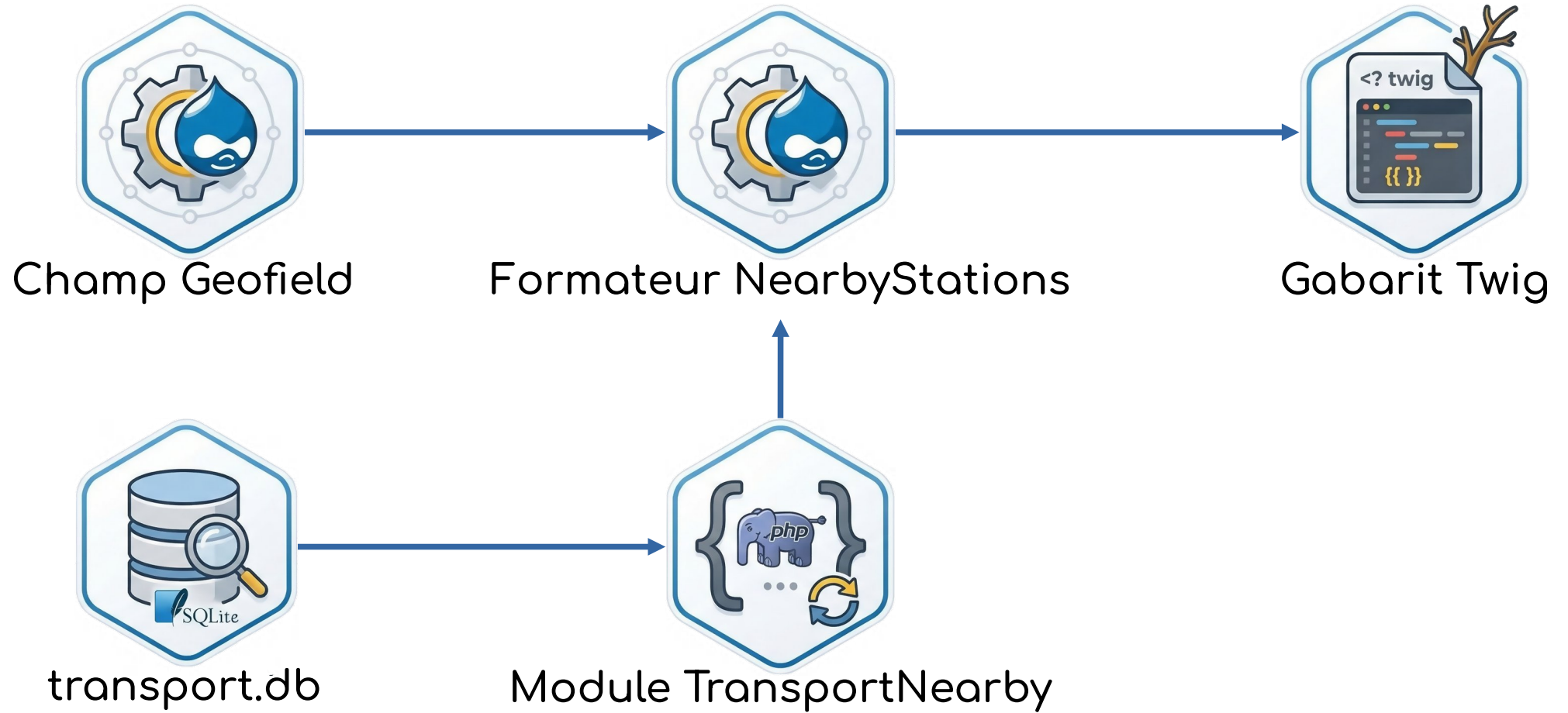


# LIVRAISON DES DONNÉES

- Push ou pull
- Manuel ou automatisé
- Base de données  SQLite = ~10 Mo
- Mode lecture seule
- Rafraîchissement du cache



# UTILISATION DES DONNÉES SOUS DRUPAL



# POURQUOI UN FORMATEUR?



OPTION	AVANTAGES	LIMITES
Bloc	Simple à ajouter	Contexte moins naturel
View	Idéal avec des données Drupal	Peu adapté à notre source
Preprocess	Rapide	Peu configurable par UI
Champ calculé	Élégant sur le papier	Rigide
Formateur	Natif, réutilisable, configurable	Nécessite un peu de code



# CONFIGURATION : CHAMP GEOFIELD

**Gérer les champs** [+ Créer un nouveau champ](#) [+ Ré-utiliser un champ existant](#)

Modifier **Gérer les champs** Gérer l'affichage du formulaire Gérer l'affichage Traduire types de paragraphes

Champ	Type de champ	Actions
<b>Coordonnées de référence</b> <small>field_reference_point</small> <input type="button" value="Single"/> <input type="button" value="Requis"/>	Geofield	<input type="button" value="Modifier"/> ▼
<b>Titre</b> <small>field_title</small> <input type="button" value="Single"/>	Texte (brut)	<input type="button" value="Modifier"/> ▼



# CONFIGURATION : DÉTAILS DU CHAMP

Paramètres du champ Coordonnées de référence pour Transports à proximité ✕

**Libellé \***

Coordonnées de référence

**Stockage du champ**

Ces paramètres s'appliquent au champ *Coordonnées de référence* partout où il est utilisé. Certains ont également un impact sur la manière dont les données sont stockées et ne peuvent plus être modifiés une fois que des données ont été créées.

**Storage backend**

Default (WKT) ▼

Choisissez le backend pour stocker les données de champ. Les suivants sont disponibles:

- Default (WKT): Default Geofield Backend for storing values in WKT Format
- PostGIS Geometry: Geofield Backend storing values in EWKB Format, suitable for PostgreSQL/PostGIS (needs PostGis enabled)

**Nombre de valeurs autorisées**

Limité ▼ 1 ▲

[Enregistrer les paramètres](#) [Supprimer](#)



# CONFIGURATION : FORMATEUR

## Gérer l'affichage

Modifier

Gérer les champs

Gérer l'affichage du formulaire

Gérer l'affichage

Traduire types de paragraphes

Champ	Nom système	Libellé	Format	<a href="#">Afficher le poids des lignes</a>
<b>Content</b>				
⇅ Titre	field_title	- Non affiché -	Titre de niveau 2	⚙️
⇅ Coordonnées de référence	field_reference_point	- Non affiché -	Nearby stations	⚙️
<b>Désactivé(e)s</b>				
<i>Aucun champ n'est caché.</i>				



# CONFIGURATION : PARAMÈTRES

## Gérer l'affichage

Modifier

Gérer les champs

Gérer l'affichage du formulaire

Gérer l'affichage

Traduire types de paragraphes

Champ

Nom système

Libellé

Format

[Afficher le poids des lignes](#)

### Content

Titre

field\_title

- Non affiché -

Titre de niveau 2

Coordonnées de référence

field\_reference\_point

- Non affiché -

Paramètres de format: **Nearby stations**

Distance maximum en mètres

300

Mettre à jour

Annuler

### Désactivé(e)s

Aucun champ n'est caché.



# ÉDITION : CRÉATION D'UN PARAGRAPHE

Hôtel de Ville Publié Enregistrer Aperçu

Media small\_deficients-moteur-cmjn\_0\_0.jpg, left Modifier

Texte mis en forme <p><img alt="small\_personnes-a-mobilite-red... Modifier

**Transports à proximité** Accès transports Modifier

Texte mis en forme Contact adapté aux personnes sourdes et mal... Modifier

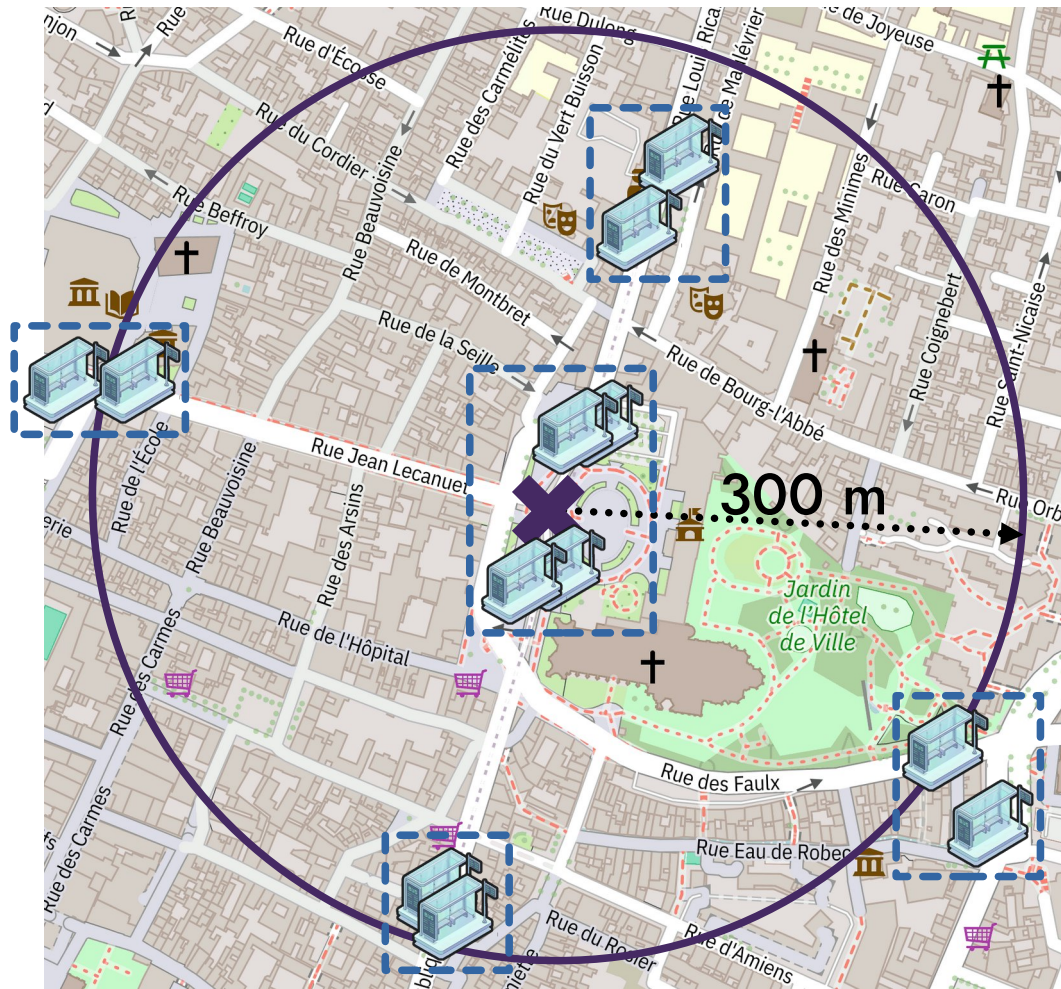
Lien bouton En savoir plus et utiliser le service, \_blank Modifier

Ajouter Média à Contenu





# SOUS LE CAPOT



- Rayon = 300m
- 12 arrêts
- 5 stations
  - Lycée Corneille F1 20 305
  - Beaux-Arts F2 F7 11 15 22 98
  - Hôtel de Ville F1 F2 F7 11 15 20 22 98 305
  - Saint-Vivien F2 15 20 22 305
  - Saint-Nicolas F1 F7 11 98
- Après filtrage
  - Hôtel de Ville F1 F2 F7 11 15 20 22 98 305

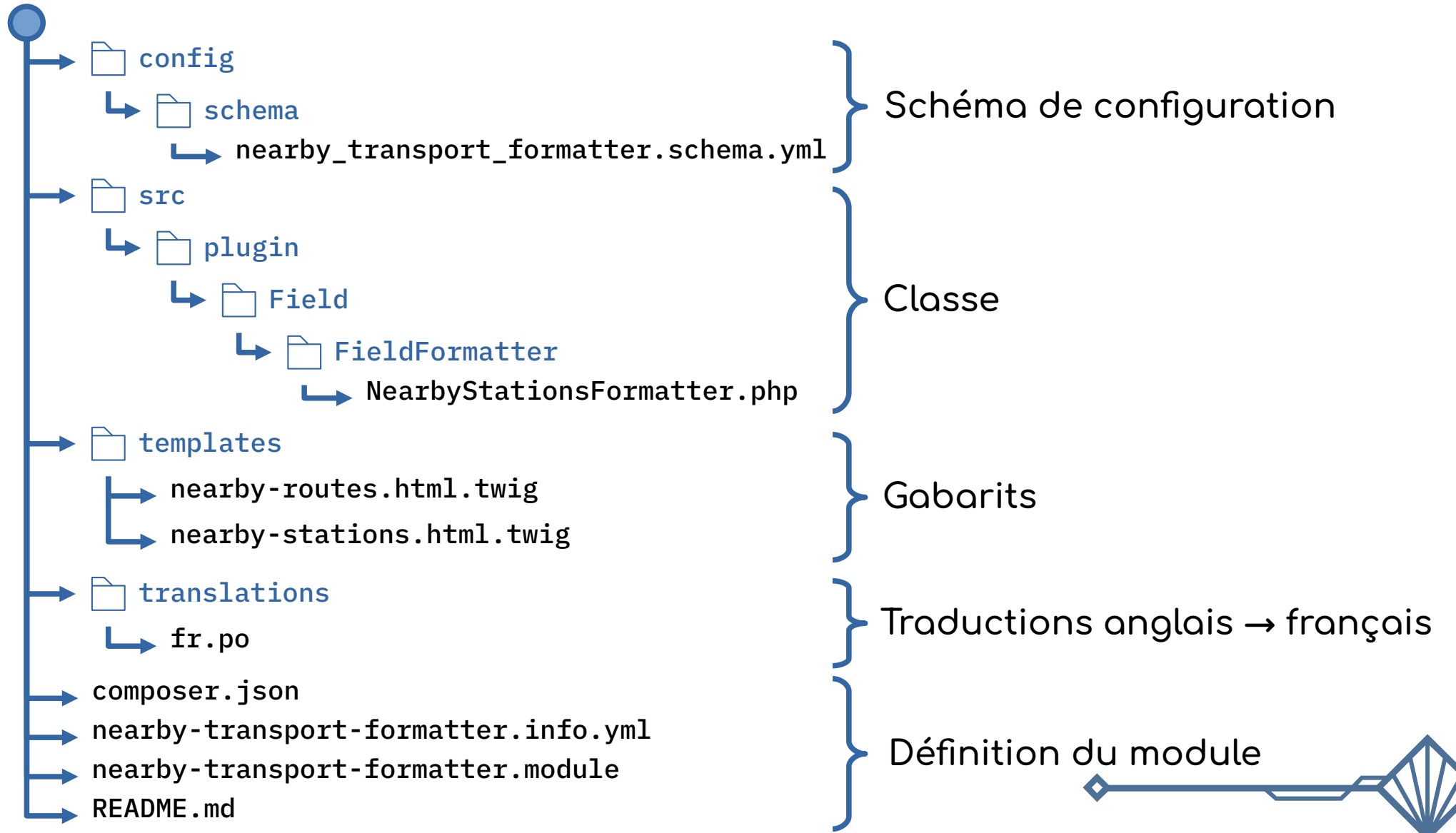


# PERFORMANCE ET CACHE DANS DRUPAL

- Base SQLite en lecture seule
  - Lecture uniquement lors du rendu du champ
  - Mise à jour de transport.db simple et rapide
- Gestion du cache
  - Contexte de l'entité parente
  - Invalidation par changement de configuration



# ORGANISATION DU FORMATEUR



# ET APRÈS ?



# FONCTIONNALITÉS SUPPLÉMENTAIRES

- Noms de lignes plus complets (A11Y)
- Accessibilité aux fauteuils roulants
- Accessibilité aux vélos
- Gestion du cache

- Support de zones
- Support d'autres villes
- Carte interactive
- Rayon de proximité individualisé



# QUELQUES ENSEIGNEMENTS

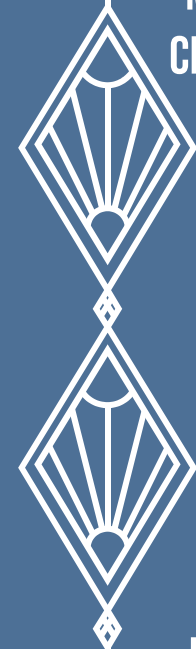
- Un standard ne garantit pas la cohérence  
<https://gtfs-validator.mobilitydata.org>
- Nettoyer est plus long qu'exploiter
- Une architecture simple est souvent suffisante
- L'open data nécessite de la résilience



# MERCI À VOUS!



MERCI D'ÉVALUER  
CETTE CONFÉRENCE



MERCI À TOUS  
LES SPONSORS

