

Université de Toulouse

MASTER 2 GEOMATIQUE

« **ScIences Géomatiques en environneMent et AménageMent** » (SIGMA)

<http://sigma.univ-toulouse.fr>

RAPPORT DE STAGE

**ETUDE DE L'ORGANISATION DES DONNEES
GEOGRAPHIQUES DE LA DIRECTION DE LA
TRANSITION ECOLOGIQUE DU CD 31**



© K. Lesage, 2021

LESAGE Karine

CONSEIL DEPARTEMENTAL DE LA HAUTE-GARONNE



Maître de stage : Frank ARDITE / Référent technique : Fabrice THEVENON

Enseignant-référent : Laurent JEGOU

Septembre 2021

RESUME

La Direction de la Transition Ecologique (DTE) met en œuvre les politiques environnementales du Département de la Haute-Garonne. De par ses missions, elle est à la fois utilisatrice et productrice de données géographiques. L'objectif est d'établir un état des lieux détaillé des données géographiques de la DTE, de les caractériser, puis de proposer des améliorations dans l'organisation et la structuration des données. L'utilisation de l'utilitaire DicoGIS (Isogeo) couplée à une analyse statistique avec Python (bibliothèque Pandas) et à des entretiens avec les agents de la DTE a permis de réaliser un état des lieux quantitatif et qualitatif. Une bibliothèque d'accès aux données géographiques identifiées lors de l'état des lieux a été réalisée et un besoin de structuration des données cyclables a été identifié. Les standards existants dans le domaine cyclable, le Géostandard Veloroutes et Voies Vertes (COVADIS – Vélo & Territoires, 2014) et le Schéma de données des aménagements cyclables (Transport.gouv.fr – Vélo & Territoires, 2020), ont permis la création d'un modèle conceptuel de données. Une base de données PostgreSQL / PostGIS sur les itinéraires cyclables définis dans le Schéma Directeur Cyclable du Département a ensuite été développée, en assurant sa normalisation et son interopérabilité. Avec une configuration adaptée, cette base de données a ensuite été testée sur des relevés terrain grâce à un projet QGIS et la solution mobile associée QField.

Mots-clés : état des lieux, données géographiques, itinéraire cyclable, base de données, normalisation, interopérabilité

ABSTRACT

The Ecological Transition Direction (DTE) implements the environmental policies of the Département de la Haute-Garonne. In regard of its missions, it both uses and produces geographical data. The objective is to establish a detailed analysis of the geographical data of the DTE, to characterize them and then to propose improvements to better organize and structure them. A tool named DicoGIS (Isogeo) was used, combined with a statistical analysis with Python (Pandas library) and interviews with agents from the DTE. This methodology enabled us to carry out both a quantitative and qualitative analysis. A data library listing accesses to geographical data identified during the analysis was created and a need to structure data linked to cycling paths was identified. Existing standards in the cycling field, the *Géostandard Veloroutes et Voies Vertes* (COVADIS – Vélo & Territoires, 2014) and the *Schéma de données des aménagements cyclables* (Transport.gouv.fr – Vélo & Territoires, 2020), helped to create a data conceptual model. A PostgreSQL / PostGIS database on cycling paths listed in the Cycling strategic guidelines of the Département was then developed, assuring its standardization and interoperability. With an adapted configuration, that database was then tested with a field data acquisition thanks to a QGIS project and the associated mobile solution QField.

Key words : analysis, geographical data, cycling path, database, standardization, interoperability

REMERCIEMENTS

Un grand merci :

A toute l'équipe de la DTE :

- Annick Vézier pour m'avoir accueilli dans sa direction,
- Frank Ardite pour m'avoir donné l'opportunité de réaliser ce stage au sein de la DTE. Merci à lui pour ses conseils, sa disponibilité et son suivi tout au long du stage. J'espère pouvoir revenir un jour travailler sur le modèle conceptuel de données des compétences eau !
- A la cellule Vélo : Valérie, Sylvain, Julia et Patrick pour m'avoir permis de travailler sur la base de données Vélo et m'avoir emmené pour un test terrain sur la Trans-Garona.
- A toute la DTE et plus particulièrement aux services Eau pour les pauses café, les moments de rire, leur accueil chaleureux.

A l'équipe géomatique :

- Fabrice Thevenon pour son suivi côté géomatique, sa disponibilité et sa gentillesse.
- Monique Lebarbey pour le partage de son expérience QGIS et sa disponibilité.

A Rémi Delabryère, Jérémie Ory, Clotilde Carton et Grégory Preynat pour m'avoir accordé de leur temps pour partager leur expérience.

A Laurent Jégou pour avoir accepté d'encadrer mon stage et avoir assuré un suivi régulier.

A toute l'équipe enseignante du M2 SIGMA pour la qualité de leur enseignement et toutes les compétences transmises : l'année d'études la plus dure mais la plus formatrice que j'ai pu suivre jusqu'alors.

Enfin, aux copains SIGMA : à défaut de s'être beaucoup vus en vrai, nos discussions virtuelles m'ont permis de tenir le coup pendant cette année intense.

LISTE DES FIGURES

Figure 1 Les 27 cantons de la Haute-Garonne et les chiffres clés (source : Cd 31, 2019).....	3
Figure 2 Principales compétences du Département de la Haute-Garonne (source : Cd 31, réalisation : K. Lesage, 2021 - Excalidraw)	4
Figure 3 Organigramme de la Direction de la Transition Ecologique (réalisation : K. Lesage, 2021).....	5
Figure 4 Architecture du SIG au Cd 31 (source : COPIL SIG du 6 mars 2018, retravaillé par K. Lesage, 2021).....	8
Figure 5 Diagramme de Gantt prédictif (source : K. Lesage, 2021).....	9
Figure 6 Diagramme de Gantt réel (source : K. Lesage, 2021)	10
Figure 7 Méthodologie pour la réalisation de l'état des lieux des données cartographiques.....	11
Figure 8 Schéma simplifié du serveur partagé de la DTE (source : K. Lesage, 2021)	12
Figure 9 Répartition de l'ensemble des fichiers du serveur partagé de la DTE et des fichiers cartographiques selon les répertoires	14
Figure 10 Part des fichiers cartographiques sur le serveur partagé de la DTE.....	15
Figure 11 Taille totale occupée par les répertoires sur le serveur partagé de la DTE	15
Figure 12 Répartition des fichiers cartographiques entre répertoires et entre types (en nombre).....	17
Figure 13 Répartition de l'espace occupé par les fichiers cartographiques selon les répertoires (à gauche) et le format (à droite)	18
Figure 14 Synthèse de l'analyse des doublons des fichiers vecteurs avec trois méthodes différentes	19
Figure 15 Nombre de copies des fichiers vecteurs dupliqués des méthodes n° 1 (à gauche) et n° 3 (à droite).....	19
Figure 16 Analyse des doublons des fichiers rasters.....	20
Figure 17 Analyse des doublons des fichiers CAO/DAO.....	21
Figure 18 Synthèse de l'analyse des doublons des fichiers cartographiques	21
Figure 19 Onglet « Infos » de la bibliothèque de données	25
Figure 20 Extrait de la bibliothèque d'accès aux données utilisées par tous les agents de la DTE.....	27
Figure 21 Méthodologie pour le déploiement de la base de données des itinéraires cyclables de loisirs	32
Figure 22 Schéma de principe des éléments repris dans les tables de la BD velo (K. Lesage, 2021, réalisé avec Excalidraw)	34
Figure 23 Modèle Conceptuel de données de la base de données Velo	36
Figure 24 Description des tables de la base de données Velo dans QGIS.....	41
Figure 25 Gestion des rôles de connexion et de groupes dans la base de données (source : SIGEA, 2021).....	41
Figure 26 Description d'une partie des privilèges pouvant être accordés sous PostgreSQL (source : SIGEA, 2021)	42
Figure 27 Paramétrage des relations attributaires entre les tables de la BD Velo	43
Figure 28 Exemple de l'intérêt de l'ajout des relations attributaires dans QGIS – relation portion cyclable / itinéraire.....	43
Figure 29 Extrait de la BD Velo - fond de carte : plan IGN v2.....	44
Figure 30 Extrait du résultat du test de relevés pour la BD Velo avec QField.....	45
Figure 31 Illustration des écarts inopinés observés sur la trace GPS globale (en noir) et segmentée (en rouge)	46
Figure 32 Méthodologie de catalogage des données basée sur la solution Isogeo.....	48

LISTE DES TABLEAUX

Tableau 1 Méthodes testées pour l'analyse des doublons des fichiers vecteurs avec Python	18
Tableau 2 Estimation de l'espace occupé par les fichiers cartographiques à supprimer	21
Tableau 3 Etat des lieux général - Points positifs et pistes d'amélioration.....	22
Tableau 4 Retours d'expérience pour l'organisation et la structuration des données géographiques	24
Tableau 5 Synthèse des propositions spécifiques aux services	30
Tableau 6 Droits et privilèges définis sur la BD Velo	42

SOMMAIRE

INTRODUCTION	1
PARTIE 1 - CONTEXTE ET OBJECTIFS	3
1-1- Présentation du Département de la Haute-Garonne	3
1-1-1- Présentation de la direction d'accueil.....	4
1-1-2- Présentation de l'équipe géomatique.....	7
1-2- La géomatique au Département de la Haute-Garonne	8
1-3- Objectifs du stage et organisation	9
PARTIE 2 – ETAT DES LIEUX	11
2-1- Méthodologie	11
2-1-1- Analyse du serveur partagé.....	11
2-1-2- Entretiens	13
2-2- Résultats Généraux	14
2-2-1- Analyse du serveur avec DicoGIS.....	14
2-2-2- Entretiens	22
PARTIE 3 – PROPOSITIONS D'AMELIORATIONS	23
3-1- Retours d'expérience d'autres collectivités	23
3-2- Faciliter l'usage des données utilisées par tous	24
3-3- Vers une structuration des données produites et utilisées par les services	28
PARTIE 4 – STRUCTURATION DES DONNEES DES ITINERAIRES CYCLABLES DE LOISIRS	31
4-1- Déploiement d'une base de données sur les itinéraires cyclables de loisirs	31
4-1-1- Création du modèle conceptuel de données	33
4-1-2- Création de la base de données PostgreSQL / PostGIS	37
4-1-3- Tests et vérifications réalisés.....	41
4-1-4- Gestion des droits.....	41
4-2- Fonctionnement avec QGIS	42
4-3- Relevés terrain avec QFIELD	45
4-4- Perspectives d'amélioration	47
PARTIE 5 - CATALOGAGE DES DONNEES	48
CONCLUSION ET PERSPECTIVES	49
BIBLIOGRAPHIQUE / SITOGRAPHIE	i
ACRONYMES	ii
ANNEXES	iii

INTRODUCTION

Parmi les compétences exercées par les Départements, les possibilités d'intervention en matière environnementale sont définies par la loi dans le cadre de l'aménagement durable du territoire : ceux-ci sont compétents en matière d'espaces naturels sensibles, d'espaces agricoles et naturels périurbains, ainsi que dans les domaines de l'eau et des milieux aquatiques et marins, à l'exclusion du champ de la compétence relevant de la Gestion des Milieux Aquatiques et de la Prévention des Inondations (GEMAPI).

Ils ont aussi la possibilité de mettre à la disposition des communes et des établissements publics de coopération intercommunale une assistance technique pour l'exercice de leurs compétences en matière d'assainissement, de protection de la ressource en eau, de restauration et de l'entretien des milieux aquatiques.

Ils peuvent également contribuer au financement de projets portés par les communes ou leurs groupements contribuant à la protection de l'environnement.

La Direction de la Transition Ecologique (DTE) met en œuvre les politiques environnementales du Département de la Haute-Garonne, conformément aux compétences qui lui sont attribuées (Cd 31, Intranet). Les thématiques exercées à la DTE sont diverses : élaboration du Plan départemental des itinéraires de promenade et de randonnée (PDIPR) et du Schéma directeur cyclable, gestion des Espaces Naturels Sensibles (ENS) et du conservatoire des zones humides, gestion quantitative de la ressource en eau ou encore financement des projets des collectivités en matière d'eau potable et d'assainissement. De par ses missions, la DTE est à la fois utilisatrice et productrice de données géographiques. L'utilisation, le partage et la production de données géographiques sont donc au cœur de ses activités.

Afin de gérer ses données géographiques, le Département a déployé un Système d'Information Géographique (SIG) géré par le Pôle Géomatique au sein de la Direction des Systèmes d'Information et du Numérique (DSIN) : bases de données partagées, applications métiers, logiciel QGIS, WebSIG. Une base de données partagée est centrale puisqu'elle contient les données produites par le Département : base31.

Depuis 2013, le SIG 31 a déployé Géosource¹, solution libre et gratuite développée par le Ministère de l'Ecologie à partir de GeoNetworks², solution libre et gratuite de l'*Open Source Geospatial Foundation*. Les données de base31 ont ainsi été cataloguées, conformément à la Directive INSPIRE³. La solution Géosource peinait à être maintenue et son utilisation n'a jamais été satisfaisante. Celle-ci a été migrée vers GeoNetworks en 2018, mais sans apporter de plus-value (maintenance manuelle lourde, pas de lien direct avec les bases de données ni les outils SIG). En 2021, une nouvelle solution de catalogage est en cours de déploiement, basée sur la solution OpenCatalog de Isogéo. Les gains attendus sont une maintenance facilitée (scan FME) une meilleure intégration avec les outils SIG (plugin QGIS, liens WebSIG).

Pour la diffusion externe une plateforme OpenData⁴ est déployée. Cette plateforme est alimentée par des flux WFS issus de la base31.

¹ <http://www.geosource.fr/>

² <https://www.geonetwork-opensource.org/>

³ http://cnig.gouv.fr/?page_id=1177

⁴ <https://data.haute-garonne.fr/pages/accueil/>

Cependant, malgré les réflexions entreprises et la ré-organisation du serveur de fichiers partagés par la DTE en 2016 et en l'absence d'une solution de catalogage des données géographiques satisfaisante et de responsables des données identifiés au sein de la DTE, la gestion actuelle des données géographiques de la DTE n'est pas optimale : multiplicité de couches cartographiques, doublons de données référentielles, manque de structuration des données produites et absence de suivi des mises à jour des données utilisées par tous les agents de la DTE.

Face à ces constats, mon stage au sein de la DTE avait pour objectif d'optimiser l'organisation de données de la DTE. Le stage comporte deux grandes parties :

1. Réaliser un état des lieux détaillé des données géographiques de la DTE afin de proposer des pistes d'amélioration pour l'accès aux données et la structuration des données produites
2. Mettre en œuvre une des pistes d'amélioration identifiées, en concertation avec la direction et les services

En tant que référente SIG de la direction pendant mon stage, mon rôle était également de participer aux missions des services et de répondre aux sollicitations techniques, en lien avec le Pôle Géomatique.

Dans une première partie, le contexte du stage avec les missions des services et l'organisation de l'infrastructure du système d'information géographique au Département seront détaillés, suivis par une déclinaison plus détaillée des objectifs du stage. Ensuite, la méthodologie employée pour l'état des lieux sera explicitée avec une présentation des résultats généraux et par service. Dans une troisième partie, les propositions d'amélioration pour la structuration des données produites et fréquemment utilisées seront abordées.

En quatrième partie, la mise en place d'une base de données des itinéraires cyclables de loisirs sera présentée en justifiant la démarche et les choix effectués

En dernière partie, la solution de catalogage des données Isogeo sera présentée avec un test sur les données géographiques de la DTE identifiées pendant l'état des lieux.

Enfin, une conclusion avec des perspectives viendra clore ce rapport de stage.

PARTIE 1 - CONTEXTE ET OBJECTIFS

1-1- PRESENTATION DU DEPARTEMENT DE LA HAUTE-GARONNE

J'ai effectué mon stage au siège du Conseil départemental de la Haute-Garonne (Cd 31) basé à Toulouse, du 8 mars au 27 août 2021. « *Le Conseil départemental est l'assemblée délibérante du département. Il est composé de 54 conseillers départementaux, 27 hommes et 27 femmes, qui ont été élus pour 6 ans pour représenter les 27 cantons qui composent aujourd'hui la Haute-Garonne.* » (Cd 31, site Internet). Les caractéristiques du territoire sont indiquées sur la Figure 1 ci-dessous.

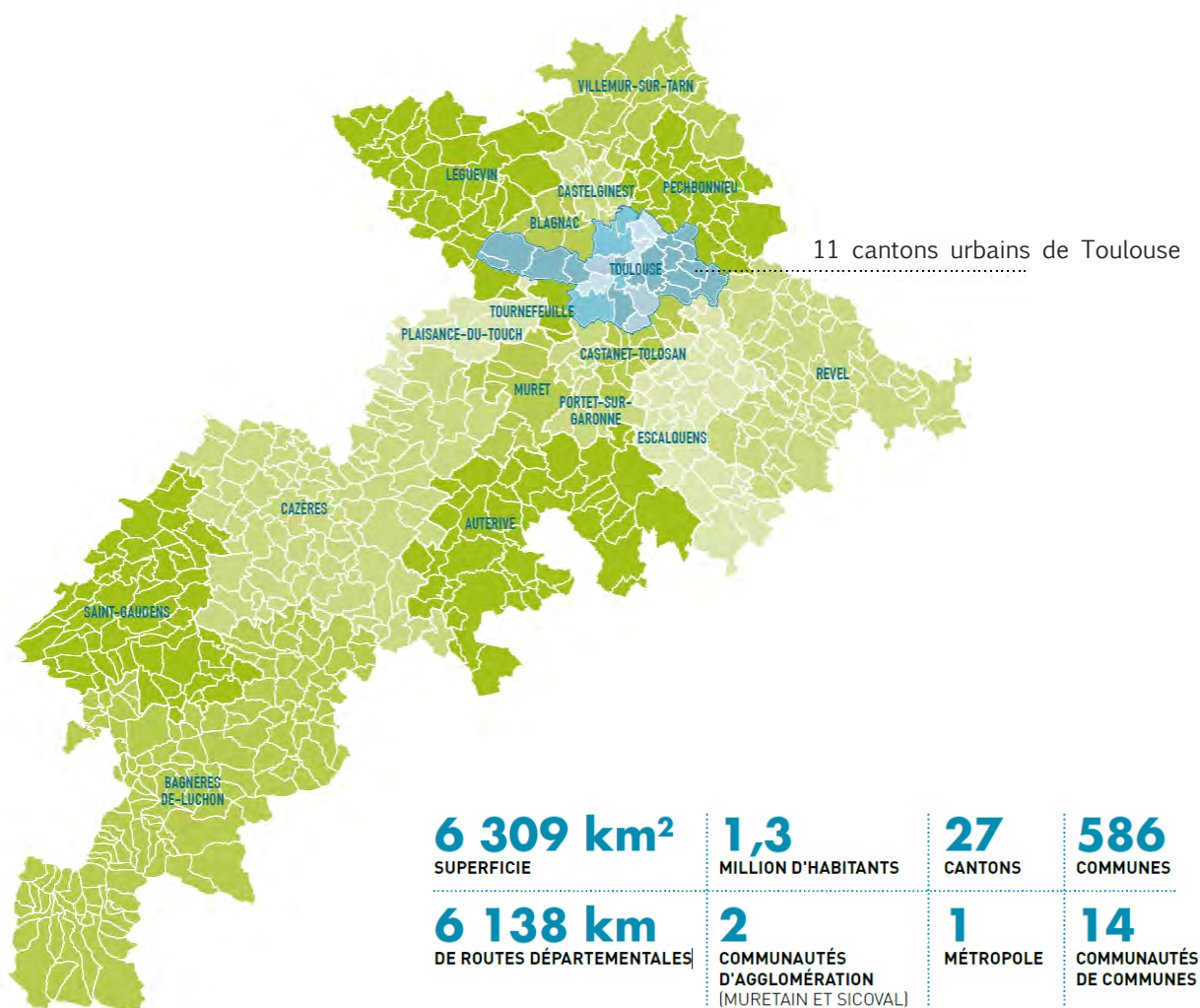


Figure 1 Les 27 cantons de la Haute-Garonne et les chiffres clés (source : Cd 31, 2019)

« *Le Département est l'échelon opérationnel de proximité et le premier acteur des solidarités humaines et territoriales, les élu.es départementaux décident de la mise en place de politiques publiques liées aux compétences de l'institution.* » (Cd 31, Intranet).

Le Département de la Haute-Garonne est divisé en sept Directions Générales Déléguées (DGD) (voir Annexe 1), organisation basée sur ses différentes compétences (Figure 2). Ces DGD sont elles-mêmes subdivisées en plusieurs directions.

Lors de mon stage j'ai travaillé au sein de la DGD « Service Opérationnels » à la Direction de la Transition Ecologique (DTE) et j'ai également été accompagnée par le pôle Géomatique intégré à la Direction des Systèmes d'Information et du Numérique (DSIN).

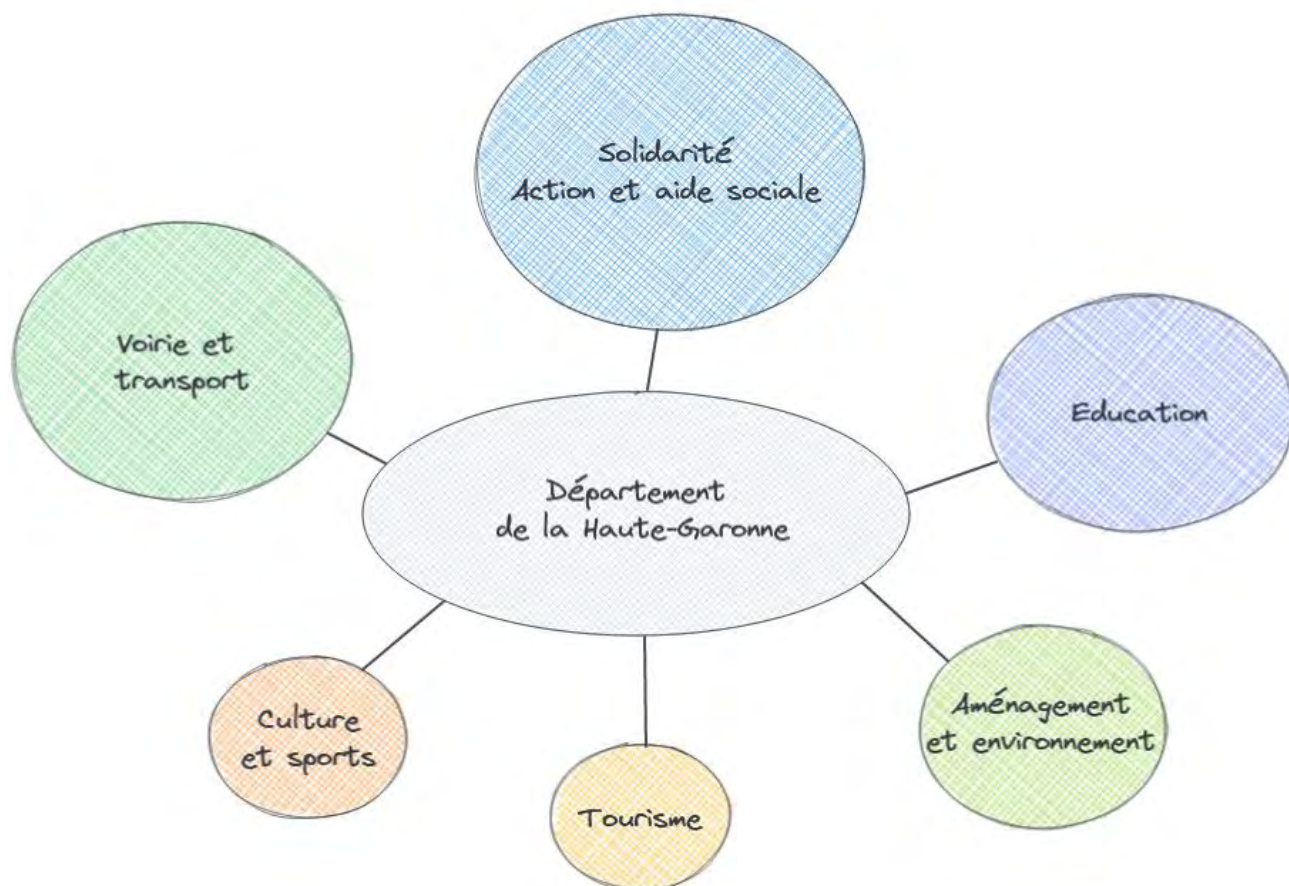


Figure 2 Principales compétences du Département de la Haute-Garonne (source : Cd 31, réalisation : K. Lesage, 2021 - Excalidraw)

1-1-1- Présentation de la direction d'accueil

La Direction de la Transition Ecologique (DTE), dirigée par Mme Annick VEZIER, s'inscrit au sein de la Direction Générale Déléguée « Services opérationnels ». La direction est constituée d'environ soixante agents répartis dans quatre services (Figure 3 ci-dessous)

« Les actions de la direction prennent leur source dans les engagements pour la transition écologique pris par le Conseil départemental en 2016, pour la période 2017-2021.

Les missions de la direction se déclinent autour de quatre grands thèmes :

- Le soutien aux collectivités, aux territoires, aux particuliers, aux associations, aux autres directions.*
- La préservation des milieux, de la biodiversité : zones humides, espaces naturels sensibles, milieux aquatiques.*
- Le développement des infrastructures : déplacements doux, eau potable et assainissement.*
- La sensibilisation de tous les publics à l'environnement et au développement durable. »*

(Cd 31, Intranet)

Organigramme de la Direction de la Transition Ecologique (DTE)

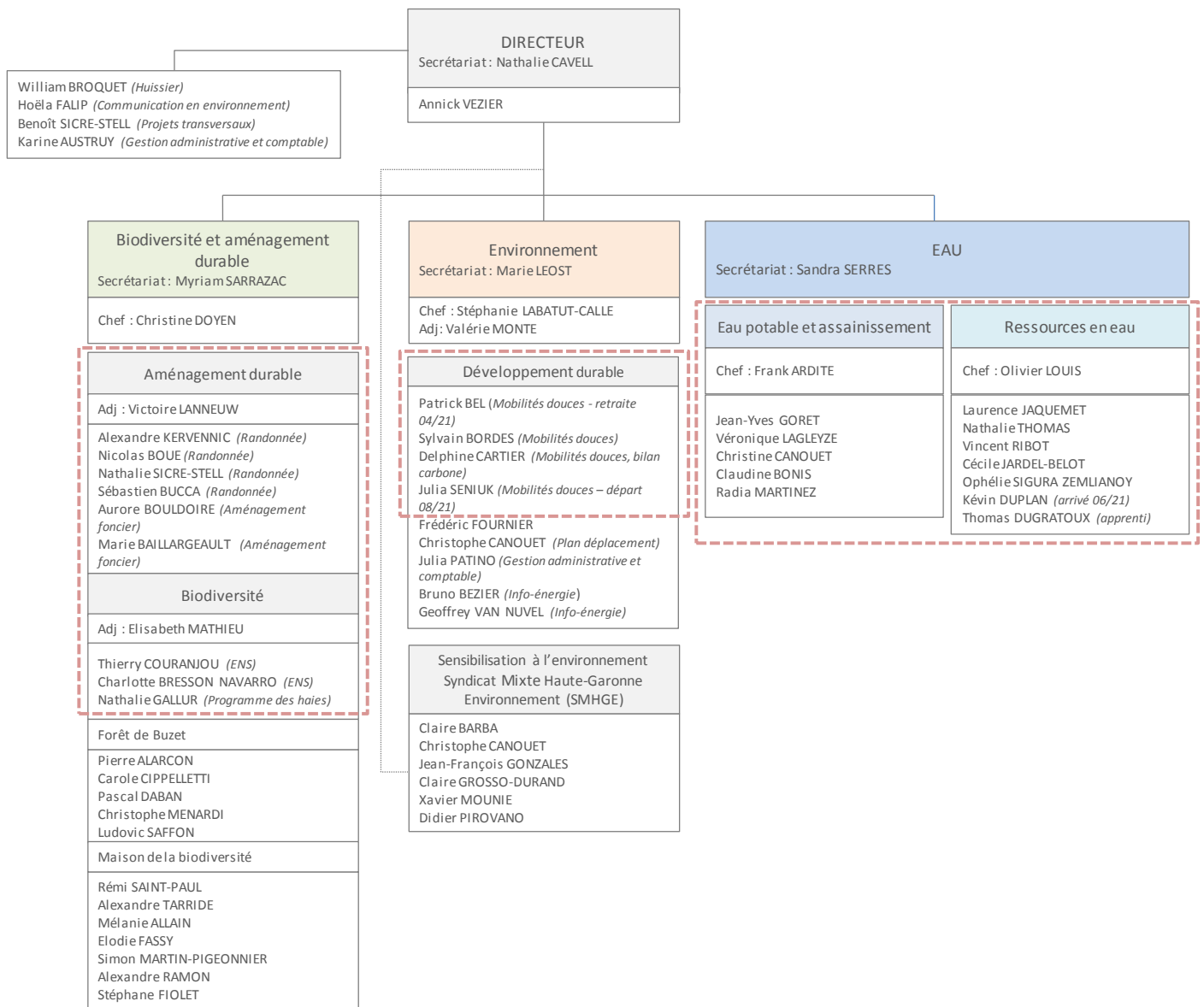


Figure 3 Organigramme de la Direction de la Transition Ecologique (réalisation : K. Lesage, 2021)

Les missions des quatre services montrent la diversité des thématiques traitées au sein de la DTE. Toutefois, toutes les missions sont liées à l'aménagement du territoire et dans le travail au quotidien au sein de leur service, les agents sont amenés à croiser leurs données métiers avec un panel de données relativement similaires dans chaque service.

Les missions principales des quatre services de la direction sont synthétisées ci-dessous :

Eau potable et assainissement

◆ **Programmation départementale des subventions**

Instruction des dossiers de demande de subventions pour les études et travaux dans le domaine de l'eau potable et l'assainissement

Avis technique sur les opérations envisagées

◆ **Expertise Eau potable et assainissement**

Le Cd 31 a établi un Schéma Départemental d'Alimentation en Eau Potable (SDAEP), il vérifie la cohérence des études et travaux du territoire avec les orientations de ce schéma Directeur (<https://www.haute-garonne.fr/service/eau-potable-un-schema-departemental>)

Ressources en eau

Missions liées au grand cycle de l'eau

◆ **Gestion et entretien des cours d'eau**

Réseau complémentaire départemental (RCD) : programmation annuelle des prélèvements pour le suivi de la qualité de l'eau et synthèse analytique

◆ **Gestion des zones humides**

Conservatoire des Zones Humides : suivi administratif et financier des espaces inventoriés

◆ **Gestion quantitative (irrigation, étiage)**

Porteur du projet de territoire Garon'Amont

Financement et suivi des Programmes d'Actions de Prévention des Inondations (PAPI)

Environnement

Missions liées à l'environnement et au développement durable

◆ **Cellule Vélo**

La cellule Vélo agit sur l'ensemble du Département, hormis le territoire de Toulouse Métropole (TM). De nombreuses concertations existent avec TM, Tisseo, la Région Occitanie, et également d'autres directions du CD 31 (Direction des Routes, Direction des Transports) qui peuvent assurer la maîtrise d'ouvrage d'une partie des itinéraires cyclables.

Itinéraires cyclables de loisirs : Schéma directeur

7 itinéraires achevés ou en cours de réalisation

Itinéraires cyclables pour le déplacement multimodal : Réseau Express Vélo (REV)

Maîtrise d'ouvrage d'une partie du REV

Instruction demandes de financement des collectivités

Subvention des aménagements cyclables après étude des tracés

◆ **Cellule Rénovation énergétique**

Conseils aux usagers pour réduire les consommations d'énergie dans l'habitat.

◆ **Sensibilisation à l'environnement**

Porté par le syndicat mixte Haute-Garonne Environnement

◆ **Autres missions**

Gestion durable des déchets, énergies renouvelables, plan de déplacement du Département

Biodiversité et Aménagement Durable (BAD)

◆ Biodiversité

Espaces naturels sensibles

- Mise en œuvre de la politique du CD 31 liée aux Espaces Naturels Sensibles (ENS) :
 - Gestion des ENS propriétaires du Département
 - Accompagnement des porteurs de projet issus des initiatives territoriales
- Animation du réseau des ENS
- Suivi des zones humides (avec le service Ressources en Eau)

Programme de plantation de haies

Aide à la plantation de haies sur les parcelles agricoles privées (100 à 1000 m de linéaire). Les plantations et leurs protections sont financées par le Département. Une visite terrain est effectuée pour marquer l'emprise de la future haie sur la parcelle.

◆ Aménagement durable

Atlas des paysages

Sur demande du Ministère de l'Ecologie, les Départements avaient pour objectif de valoriser les grands aspects du paysage actuel de leur territoire sous forme d'atlas. En Haute-Garonne, 24 unités paysagères ont été identifiées (<https://paysages.haute-garonne.fr/>).

Aménagement foncier agricole

Sur demande de la commune, le Département instruit et finance les opérations d'aménagement foncier agricole en tant que maître d'ouvrage. Cela comprend différentes procédures de regroupement parcellaire.

Le travail de classement des terres et de cartographie est réalisé par un Géomètre expert, le Département est récepteur et propriétaire des documents réalisés.

Randonnée

- Gestion du Plan Départemental d'Itinéraires de Promenades et de Randonnée (PDIPR) :
- Réception des demandes d'inscription des itinéraires de randonnée des collectivités au PDIPR
 - Visite terrain : éligibilité de l'itinéraire fourni au PDIPR (intérêt du sentier, sécurité du randonneur)
 - Etude foncière

1-1-2- Présentation de l'équipe géomatique

La partie géomatique est intégrée à la Direction Adjointe Système d'Information (DASI) / Service Transverses et Experts au sein de la Direction des Systèmes d'Information et du Numérique (DSIN).

La DSIN, dirigée par M. Didier LANGOLFF, s'inscrit au sein de la Direction Générale Déléguée « Ressources financières et numériques – Commande Publiques – Pilotage et Méthode ».

Deux agents, Fabrice THEVENON et Monique LEBARBEY, sont en charge du pilotage et de l'assistance technique géomatique au Cd 31, leurs missions sont les suivantes (Cd 31, COPIL SIG, 2018) :

- « Favoriser l'usage et le partage de l'information géographique au Cd31
- Animer le projet SIG transversal (Copil, site intranet SIG, réseau des référents...)
- Gérer les référentiels communs (IGN, DGFiP, etc.) et les BD partagées
- Aider les directions opérationnelles à réaliser leurs projets SIG (Méthodologie, conseils, outils, etc.)

- Développer les applications métiers GEO & Aigle
- Assister les utilisateurs, formations QGIS et GEO
- Suivre les marchés (CIRIL Group, Veremes, EMC3, IGN)
- Favoriser la maîtrise d'ouvrage au sein des directions
- Aider la montée en charge des logiciels métiers »

Le service SIG transversal est de petite taille, le but est de développer la maîtrise d'ouvrage dans les directions.

La gestion de l'OpenData est directement rattachée à la Direction Générale Adjointe.

1-2- LA GEOMATIQUE AU DEPARTEMENT DE LA HAUTE-GARONNE

Le schéma ci-dessous présente les liens entre les composantes du SIG du Cd 31 (Figure 4). Cette infrastructure combine des solutions libres et gratuites comme QGIS et PostgreSQL / PostGIS ainsi que des solutions payantes proposées par des prestataires : Business Geographic (GEO), Veremes (FME) et Isogeo (OpenCatalog). L'élément majeur sur lequel repose l'organisation du SIG est la base de données partagée base31 qui regroupe les données produites au Cd 31. Cette base de données est utilisée en direct pour les applications internes et répliquée chaque soir sur un serveur externe pour les applications et services liés à GEO SAAS.

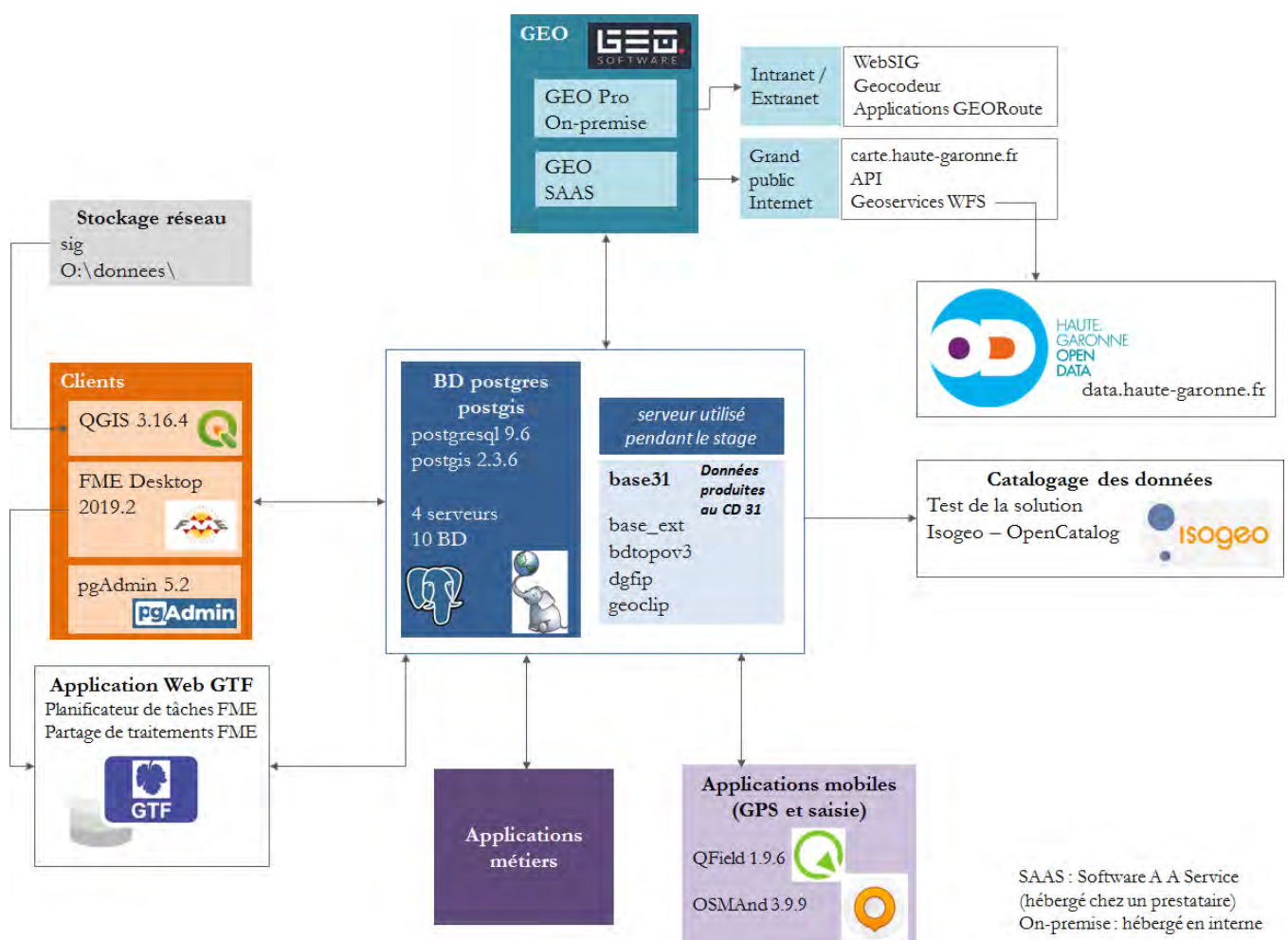


Figure 4 Architecture du SIG au Cd 31 (source : COPIL SIG du 6 mars 2018, retravaillé par K. Lesage, 2021)

1-3- OBJECTIFS DU STAGE ET ORGANISATION

Le stage s’est déroulé en suivant l’organisation définie dans la fiche de stage, tout en adaptant les missions réalisées selon les observations réalisées pendant la première phase d’état des lieux.

En définitive, les objectifs du stage sont les suivants :

- 1- Réaliser un état des lieux quantitatif des données géographiques présentes sur le serveur partagé (J :) de la DTE, à l’aide de l’utilitaire DicoGIS. Ce serveur stocke l’ensemble des répertoires utilisés par les agents au quotidien. Les données PostGIS intégrées à base31 sont hors étude sur ce premier objectif.
- 2- Réaliser un état des lieux qualitatif des données géographiques utilisées à la DTE en identifiant les données produites, référentiels et re-traitées (issues de l’agrégation de différentes sources)
- 3- En fonction de l’état des lieux, proposer des améliorations dans l’utilisation des données et dans leur structuration
- 4- Structurer les données des itinéraires cyclables de loisirs en garantissant leur interopérabilité
- 5- Tester la solution IsoGeo pour le catalogue des données

En complément, j’ai occupé le rôle de référent SIG pour la DTE pendant la durée de mon stage, j’ai donc répondu à des sollicitations techniques relatives à QGIS ou à l’utilisation de données géographiques.

La fiche de stage prévoyait une organisation en cinq phases, présentée globalement dans le diagramme de Gantt prédictif (Figure 5).

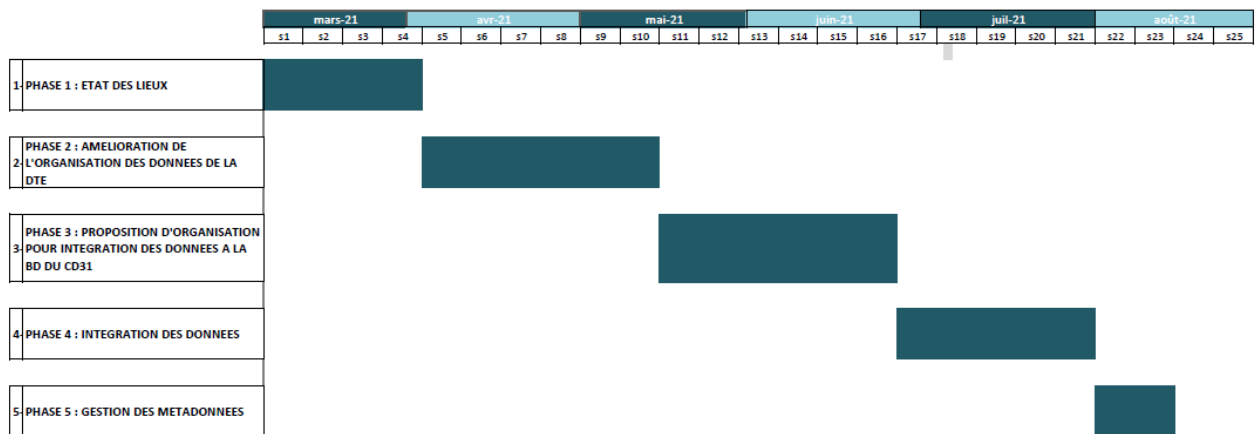


Figure 5 Diagramme de Gantt prédictif (source : K. Lesage, 2021)

Le diagramme de Gantt réel présente l’organisation du stage tel qu’il s’est déroulé (Figure 6).

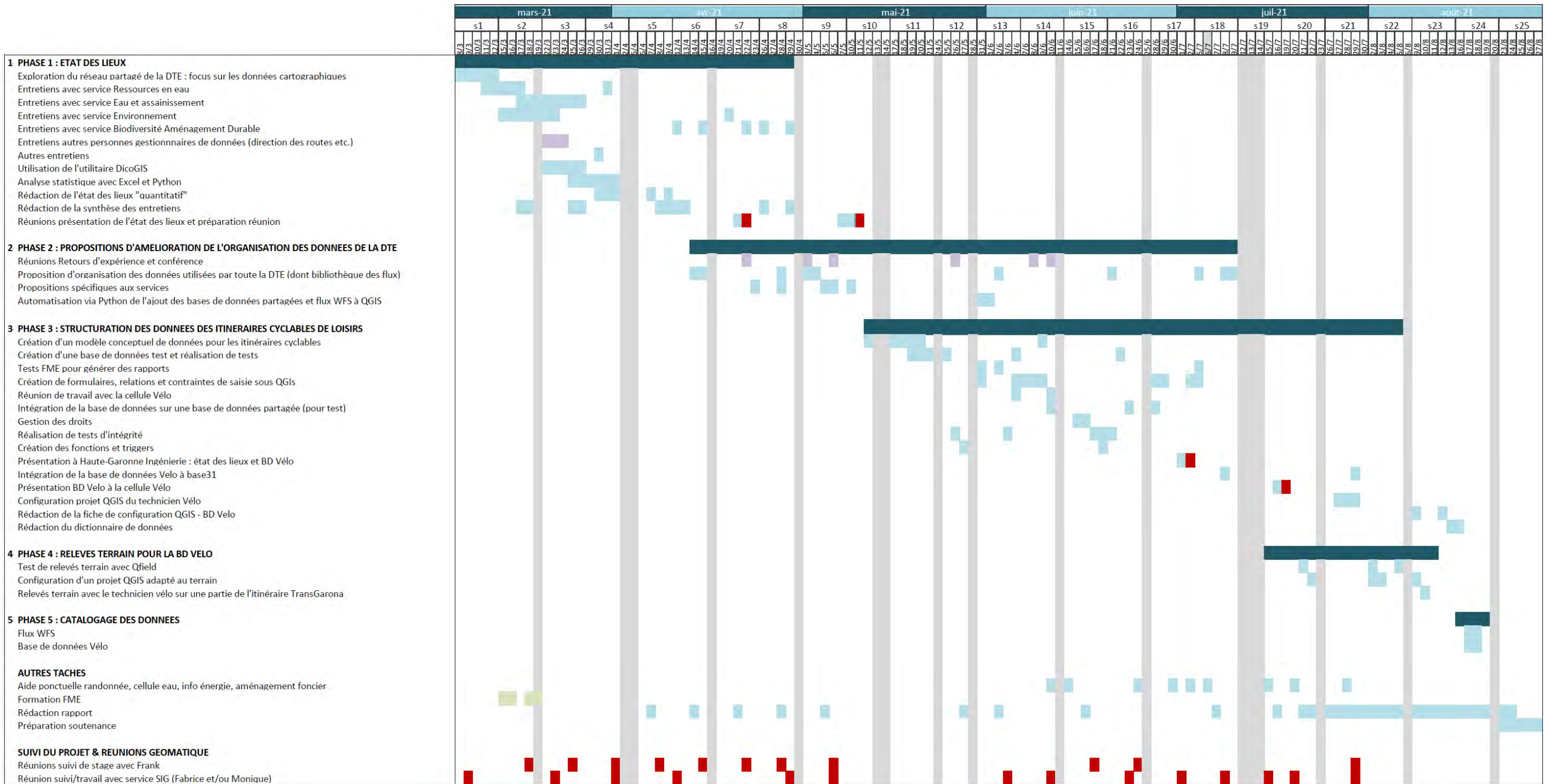


Figure 6 Diagramme de Gantt réel (source : K. Lesage, 2021)

Légende :

- formation
- retours d'expérience
- réunions
- travail autonome / entretiens

PARTIE 2 – ETAT DES LIEUX

Rappel de l'objectif :

Etablir un état des lieux des principales données cartographiques utilisées par les services de la DTE

2-1- METHODOLOGIE

La méthodologie générale employée pour l'état des lieux est présentée dans le schéma ci-dessous (Figure 7).

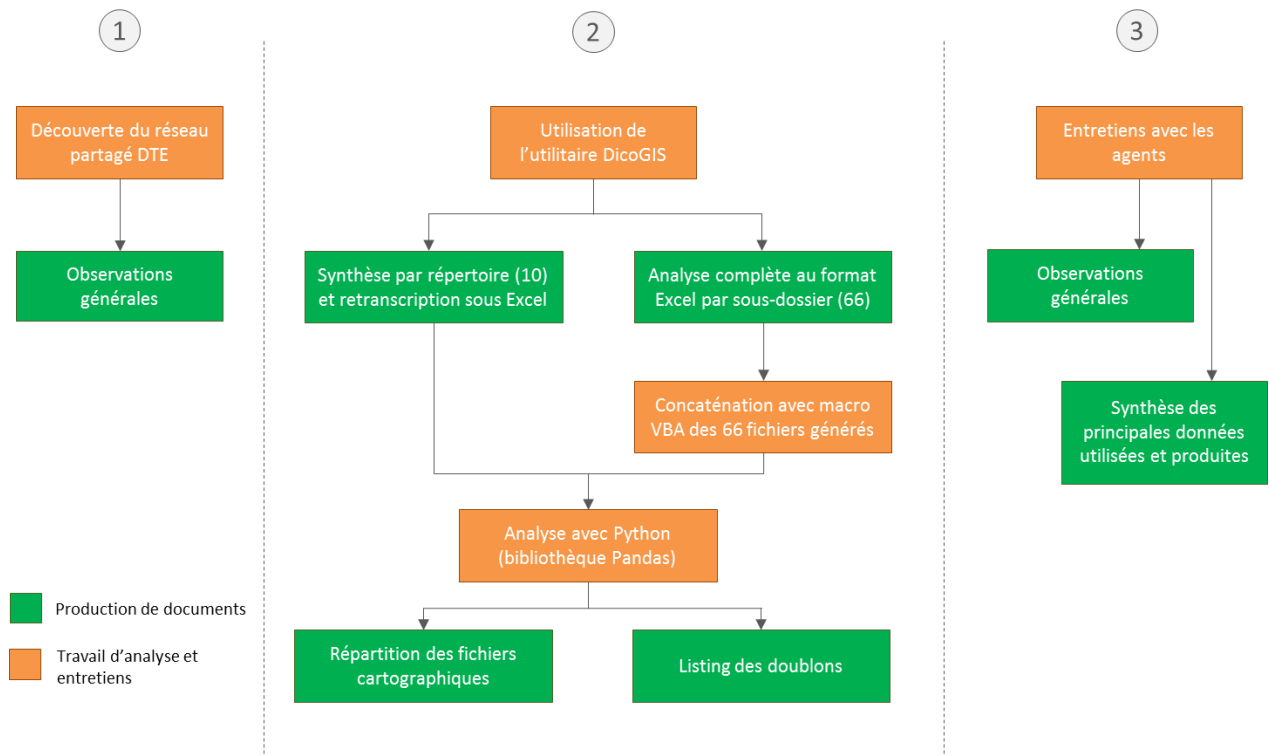


Figure 7 Méthodologie pour la réalisation de l'état des lieux des données cartographiques

Chaque partie est détaillée dans les paragraphes ci-après.

2-1-1- Analyse du serveur partagé

2-1-1-2- Présentation et découverte exploratoire

La DTE travaille sur un serveur partagé (J :). L'arborescence de ce serveur a été revue en 2016 pour mieux structurer l'organisation des fichiers. Lors de la ré-organisation il était prévu de regrouper toutes les informations cartographiques (données, projets QGIS, cartes PDF) dans le répertoire *05_INFORMATIQUE/02_CARTO_SIG*. En 2021, la présence de nombreux fichiers cartographiques est constatée dans les dossiers *08_BAD*, *09_EAU* et *10_ENV*, ce qui entraîne notamment une redondance d'information concernant les données référentielles (topographique ou thématique) de la direction.

Un schéma simplifié du serveur partagé de la DTE est représenté sur la Figure 8 ci-dessous.

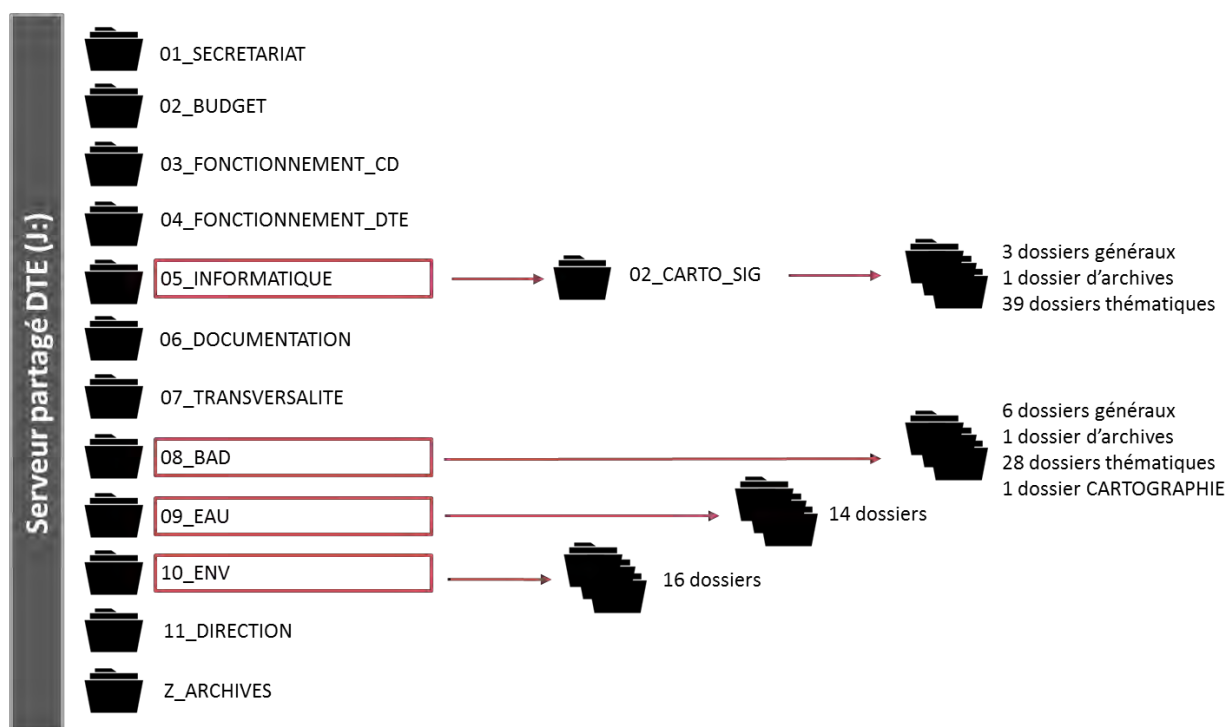


Figure 8 Schéma simplifié du serveur partagé de la DTE (source : K. Lesage, 2021)

Quelques premières observations sont effectuées pendant l'exploration du serveur :

- Des données de référence (IGN, occupation du sol) semblent dupliquées dans plusieurs dossiers
- La distinction entre les données externes (reçues ou téléchargées) et les données produites par les services n'est pas évidente
- Des fichiers cartographiques portent le même nom sans mention des différences entre fichiers
- Peu de fiches de métadonnées sont présentes sur le serveur : la provenance des données utilisées et le suivi des traitements ne sont pas tracés
- Le suivi des données externes n'est pas assuré (notamment pour la mise à jour)
- Les cartographies PDF semblent dispersées sur le serveur

2-1-1-3- Analyse statistiques des fichiers cartographiques du serveur à l'aide de DicoGIS

DicoGIS a été introduit à la DSIN lors de la présentation de la solution Isogeo pour le catalogage des données. Cette solution, actuellement en cours de test, sera détaillée dans la partie 5 sur le catalogage des données.

DicoGIS permet d'analyser le contenu d'un dossier spécifié afin de détecter la présence de données géographiques brutes. Différents formats sont recherchés : shp (ESRI), tab (MapInfo), kml, gml, geojson, geospatial PDF, rasters (ecw, tif, jp2), Esri FileGDB, CAO/DAO (type dxf / dwg pour AutoCAD), SpatialLite.

Et l'analyse des formats cartographiques contenant plusieurs fichiers ?

Pour les formats type ESRI shapefile ou Mapinfo TAB, DicoGIS ne comptabilise qu'un seul fichier mais prend en compte l'espace occupé par l'ensemble des fichiers associés.

Une colonne « *Dependencies* » permet de lister tous les fichiers associés (ex : pour les shapefiles : prj, dbf etc. / pour les TAB : DAT, MAP etc.)

Les listings de fichiers générés par DicoGIS ont ensuite été traités sur Excel et sur Python avec la bibliothèque Pandas. Une notice détaillée de l'analyse effectuée est présente en Annexe 2, avec les différents scripts utilisés.

Certaines limites de l'outil ont été identifiées :

Limites directes liées au développement de l'outil :

- Les formats géographiques suivant ne sont pas pris en compte : geopackage, GeoConcept (IDX, GCM, GCR), shapefile zippé
- Au niveau de la taille du fichier, lorsque l'utilitaire trouve une erreur sur le fichier, il ne peut pas l'analyser et donc la taille n'est pas indiquée
- DicoGIS n'est plus maintenu à jour par Isogeo
- La date de création du fichier est souvent postérieure à la date de dernière mise à jour.

La non prise en compte du format geopackage est la limite la plus problématique. Ce format semble actuellement peu utilisé par les agents des services Eau et Environnement mais certains agents du service Biodiversité et Aménagement Durable les utilisent lors de leurs sorties terrain avec QField, la solution mobile de QGIS. Ce format tend à être de plus en plus utilisé au sein de la DTE. DicoGIS n'étant plus maintenu à jour, ce format ne sera pas intégré à court terme.

Quelques limites « indirectes » :

- L'enregistrement des rasters a été désactivé sur 5 dossiers car cela empêchait la sauvegarde du fichier Excel. Toutefois dans les cas constatés il n'y avait que très peu de « rasters » et ceux-ci étaient en fait de simples fichiers images (type capture d'écran) et non de vrais rasters.
- Au vu du nombre important de PDF présents dans les dossiers, ceux-ci n'ont pas été inscrits dans le fichier Excel car le processus n'arrivait pas à terme. Toutefois l'outil est sensé recenser les PDF géographiques (geospatial PDF) et non les PDF classiques. D'après mes entretiens avec les agents de la direction et mes observations, les PDF géographiques ne sont pas utilisés par la direction.

2-1-2- Entretiens

Des entretiens ont été menés avec les agents de la DTE identifiés comme utilisateurs de QGIS en début de stage. Dans un premier temps, l'agent présente sa mission et ensuite des questions sont posées afin de déterminer quelle est son utilisation des données géographiques.

Les questions abordées sont les suivantes :

- Quelles bases de données sont utilisées ? Quelles données à l'intérieur de la base ?
- D'où proviennent les données (flux WMS/WFS, BD partagée, téléchargement, envoi) ?

Le but de cette question est notamment d'attribuer une typologie à la donnée : issue d'un référentiel topographique ou thématique, partenaire (obtenue par envoi, convention, spécifique à une étude), retraitée (extraction d'une partie de la donnée, agrégation de différentes sources etc.), produite par le service (via une étude ou du terrain), donnée de travail

- Qui est le responsable de la donnée au sein de la direction ?
- Quelle est la dernière année disponible ?
- Quelle est la fréquence de mise à jour ?

- Quel est l’emplacement de la couche cartographique sur le serveur partagé ?
- Quels traitements sont réalisés sur ces données ?
- Une cartographie est-elle réalisée ? Quel est son usage ? (service, direction, Département ou externe)

En complément, suite aux premiers entretiens :

- Les données sont-elles sensibles (confidentialité) ?
- Existe-t-il une convention pour le partage de données ?
- La convention de nommage de fichiers est-elle appliquée ? (date au format américain, pas d’espaces, pas d’accents etc.)

Des entretiens complémentaires sont menés avec les chefs de service pour avoir une vision plus générale de la gestion des données géographiques dans leur service. Quelques entretiens ont également été réalisés avec des agents d’autres directions pour échanger autour de leur gestion et leurs problématiques.

2-2- RESULTATS GENERAUX

2-2-1- Analyse du serveur avec DicoGIS

2-2-1-1- Analyse générale, tous fichiers confondus

Une analyse générale du contenu des dossiers du serveur partagé a été effectuée (nombre de fichiers et espace occupé par les fichiers cartographiques et non cartographiques), croisée avec une première analyse des seuls fichiers cartographiques avec DicoGIS (données synthétiques).

NB : Le dossier 11_DIRECTION est soumis à des droits très restreints, son contenu n’est donc pas accessible. Le dossier Z_ARCHIVES ne contient qu’un seul dossier avec 5 fichiers. Ces deux dossiers ont donc été écartés de l’analyse.

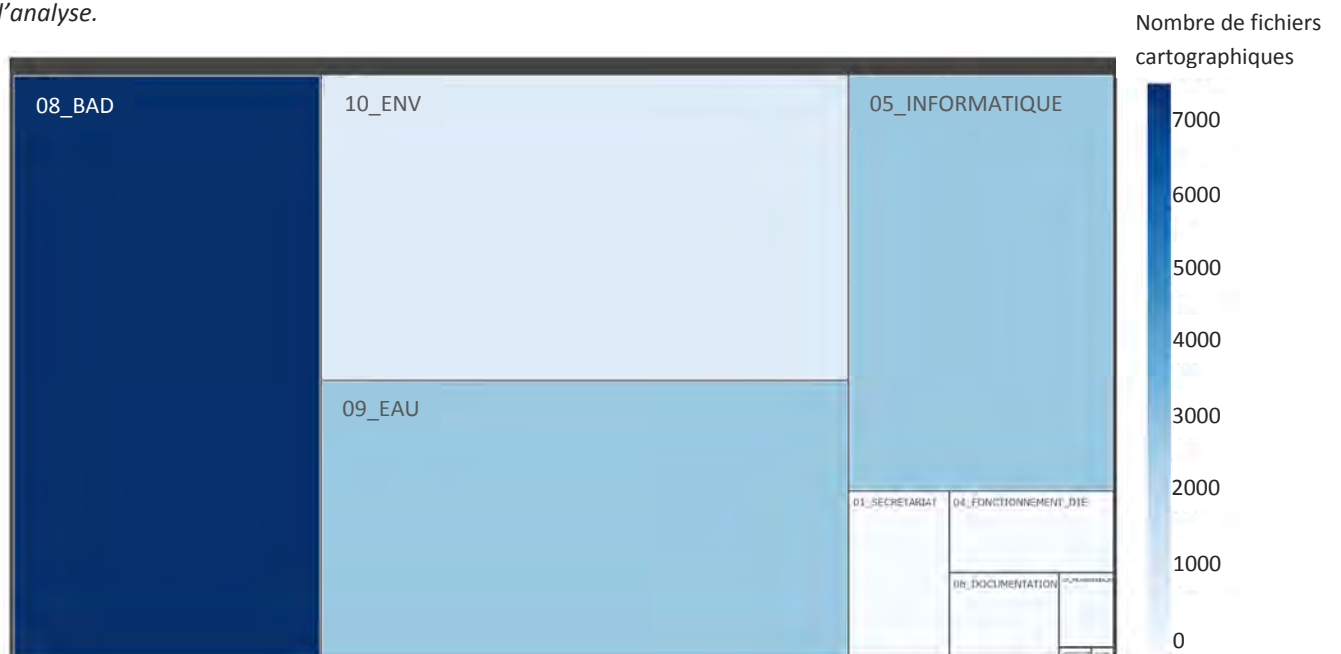
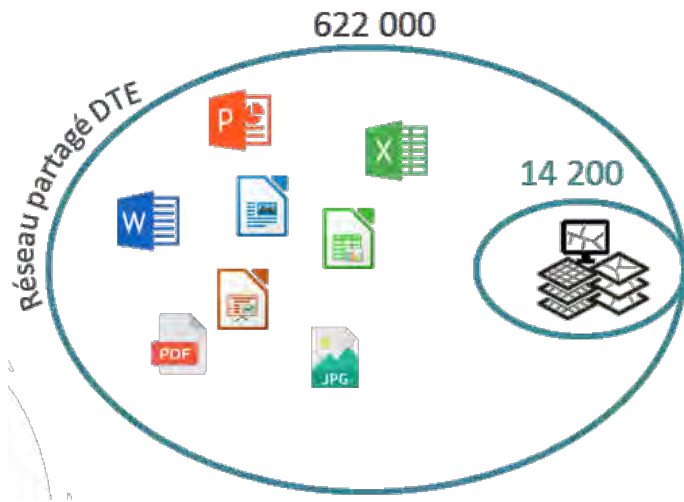


Figure 9 Répartition de l’ensemble des fichiers du serveur partagé de la DTE et des fichiers cartographiques selon les répertoires

La Figure 9 ci-dessus montre la répartition des fichiers (taille de la boîte) du serveur partagé selon 10 répertoires et le nombre de fichiers cartographiques (couleur) de chaque répertoire.

Les principaux contributeurs des 621 970 fichiers du serveur sont les répertoires *BAD*, *ENV*, *EAU* et *INFORMATIQUE* qui représentent à eux quatre 93 % du nombre total de fichiers. D'après la couleur des boîtes, ces répertoires sont également les principaux contributeurs en nombre de fichiers cartographiques (environ 14 200).



En terme d'espace sur le serveur partagé (Figure 11), sur environ 410 Go, le répertoire *BAD* représente 50 % de l'espace occupé. Ce sont ensuite les répertoires *ENV* (25 %) et *EAU* (11%) qui occupent le plus d'espace sur le serveur.

Les mêmes dossiers occupent le plus d'espace disque et contiennent le plus de fichiers, cartographiques et non cartographiques.

Figure 10 Part des fichiers cartographiques sur le serveur partagé de la DTE

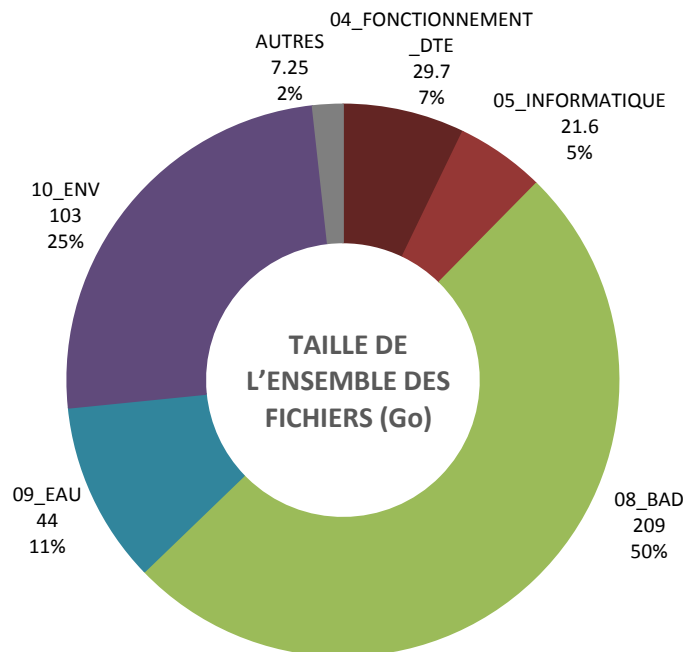


Figure 11 Taille totale occupée par les répertoires sur le serveur partagé de la DTE

2-2-1-2- Analyse des fichiers cartographiques

Pour la suite de l'analyse, seuls les répertoires contenant un nombre significatif (> 20) de fichiers cartographiques seront étudiés : 08_BAD, 09_EAU, 10_ENV, 05_INFORMATIQUE et 07_TRANSVERSALITE. A savoir que dans le répertoire 05_INFORMATIQUE, seul le répertoire 02_CARTO_SIG contient des fichiers cartographiques. Lors de l'évocation du répertoire 05_INFORMATIQUE, on se réfère en fait à l'analyse du sous-répertoire 02_CARTO_SIG.

Généralités

Les graphes sur la Figure 12 ci-dessous montrent la répartition des fichiers cartographiques selon les répertoires identifiés ci-dessus. Le nombre total de fichiers cartographiques (tous types confondus – vecteurs, rasters et CAO/DAO) présents sur le serveur est de 14 135. Les limites de l'outil listées dans la méthodologie sont à prendre en compte pour nuancer ce résultat.

56 % des fichiers sont situés dans le répertoire BAD, 19 % dans le répertoire INFORMATIQUE, 17 % dans le répertoire ENV, 7 % dans le répertoire EAU et 1% dans TRANSVERSALITE.

La majorité des fichiers cartographiques présents sur le serveur sont des fichiers vecteurs (74 %), puis des fichiers rasters (19 %) et le restant des fichiers CAO/DAO (7 %).

Les trois derniers graphiques montrent la répartition entre les répertoires pour chaque type de fichier. Ce qui ressort est la présence de plus de 50 % des fichiers vecteurs et rasters dans le répertoire BAD. Ensuite les répertoires EAU et INFORMATIQUE réunissent la majorité des fichiers vecteurs (respectivement 17 % et 19 %) et rasters restant (respectivement 19 % et 34 %).

Au niveau des fichiers CAO/DAO, qui restent minoritaires, 71 % des fichiers sont présents dans le répertoire EAU. Ces fichiers sont majoritairement issus de la modélisation des réseaux d'eau potable dans le cadre du Schéma Directeur d'Alimentation en Eau Potable (SDAEP).

Concernant la taille des fichiers cartographiques (Figure 13), la répartition entre répertoires est plus contrastée. Sur 119 Go de fichiers, 40 % sont présents dans le répertoire de la BAD, 19 % dans le répertoire INFORMATIQUE, 17 % dans le répertoire ENV, 14 % et 10 % dans les répertoires INFORMATIQUE et EAU.

85 % de l'espace occupé correspond à des fichiers vecteurs. Toutefois, comme mentionné dans les limites de l'outil DicoGIS, l'espace occupé par les fichiers rasters et CAO/DAO est sous-estimé puisque certains formats ne sont pas correctement analysés par l'utilitaire.

Avec une analyse plus fine, on obtient que ce sont respectivement 1,4 % des fichiers vecteurs et 9,2 % des fichiers rasters qui occupent 80 % de l'espace pour chaque type de fichiers. Certains projets sont donc très volumineux.

Que ce soit en termes de nombre de fichiers cartographiques ou d'espace occupé, on remarque que la part imputable au répertoire INFORMATIQUE ne dépasse pas 20 %. L'organisation espérée en 2016 avec une centralisation des données cartographiques dans le répertoire INFORMATIQUE/CARTO_SIG ne semble pas avoir été appliquée par une majorité des agents. C'est un élément à garder en considération si l'on souhaite à nouveau re-centraliser une partie des données cartographiques.

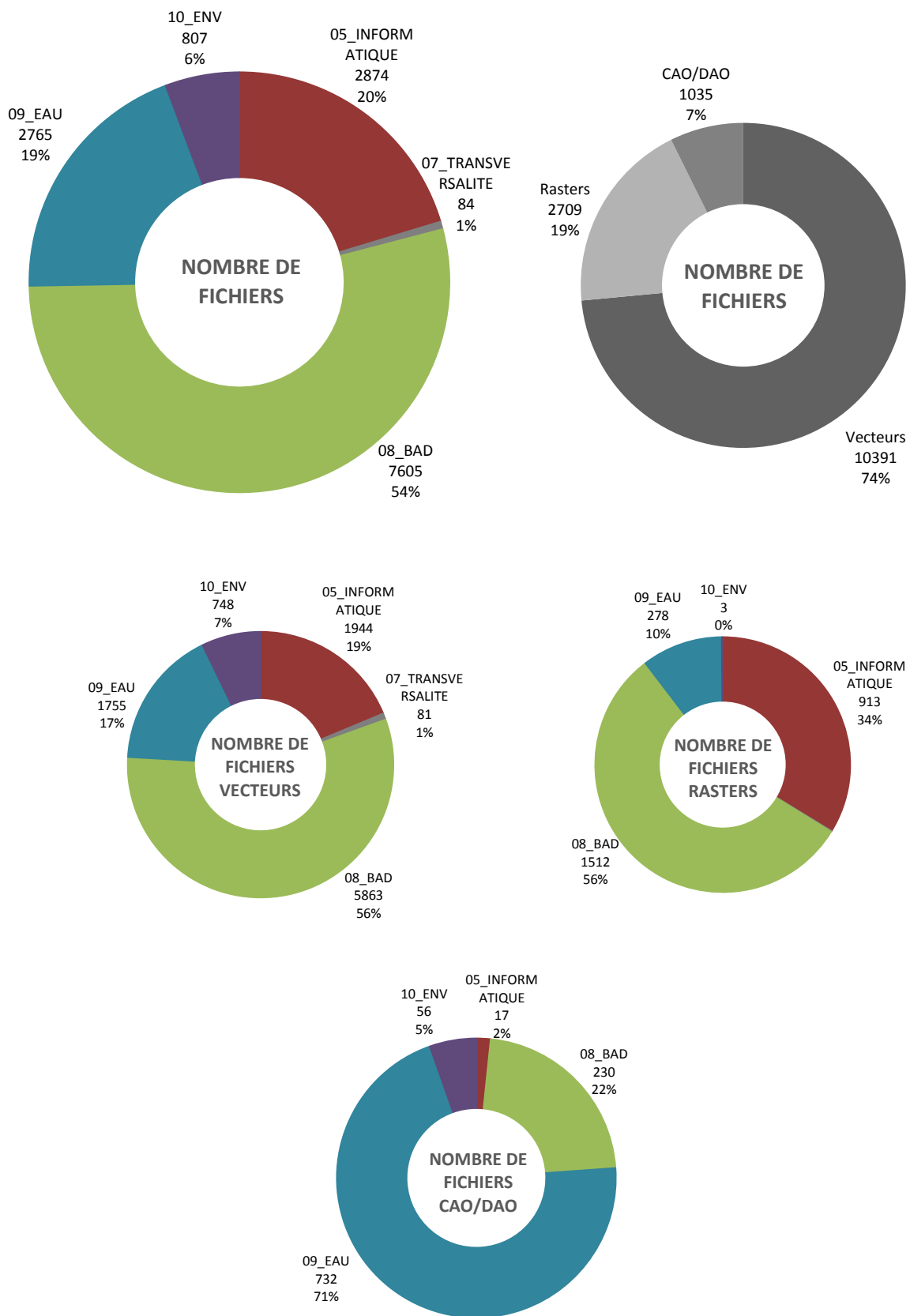


Figure 12 Répartition des fichiers cartographiques entre répertoires et entre types (en nombre)

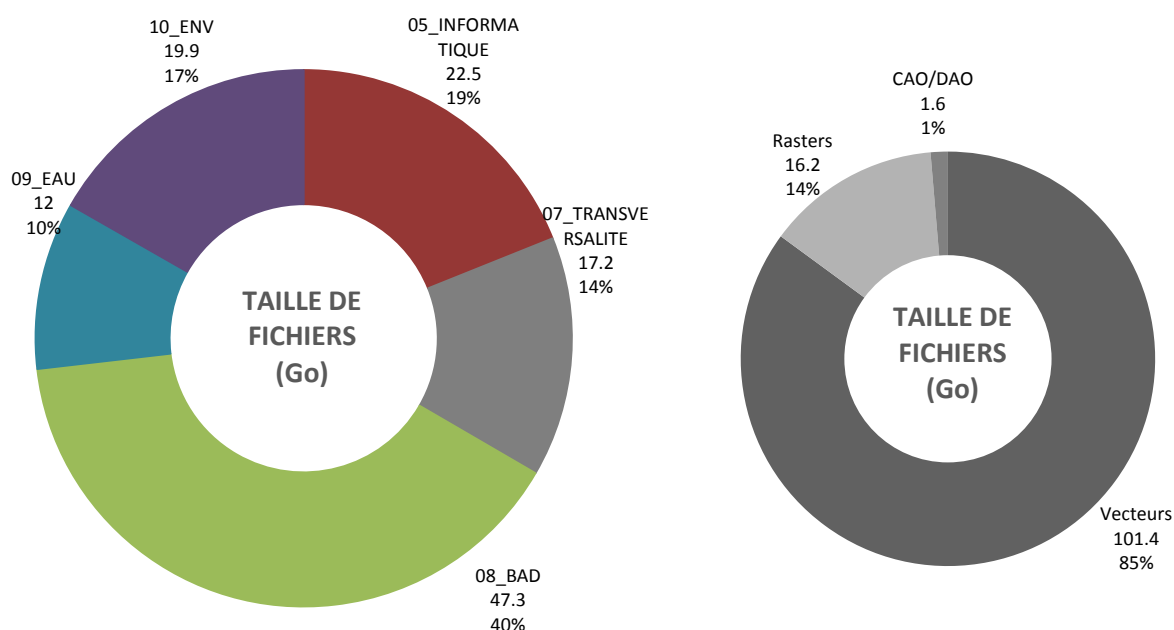


Figure 13 Répartition de l'espace occupé par les fichiers cartographiques selon les répertoires (à gauche) et le format (à droite)

Analyse des doublons

L'analyse des doublons est réalisée en se basant sur les listes de fichiers générées par DicoGIS puis avec un script programmé sous Python (Annexe 2).

Données vectorielles

Au niveau des données vectorielles, trois méthodes ont été testées pour l'analyse des doublons (Tableau 1).

	Nom du fichier	Nom du fichier en minuscule	Nombre d'objets	Nombre de champs
Méthode n° 1	X			
Méthode n° 2	X		X	
Méthode n° 3		X	X	X

Tableau 1 Méthodes testées pour l'analyse des doublons des fichiers vecteurs avec Python

Les résultats obtenus avec chaque méthode sont présentés sur la Figure 14 ci-dessous.

A noter qu'avec les méthodes 2 et 3, 474 fichiers vecteurs se trouvent exclus de l'analyse : ce sont des fichiers pour lesquels le fichier de projection n'a pas été trouvé par DicoGIS ou pour lesquels aucun objet n'est présent dans le fichier. Ces fichiers pourraient a priori être supprimés du serveur après vérification par les agents du service.

Les diagrammes ci-dessous (Figure 15), réalisées pour les méthodes n° 1 et n° 3, montrent le nombre de copies présentes sur le serveur partagé de la DTE pour les fichiers vecteurs présents au moins deux fois sur le serveur (non uniques). La majorité des fichiers sont présents deux ou trois fois.

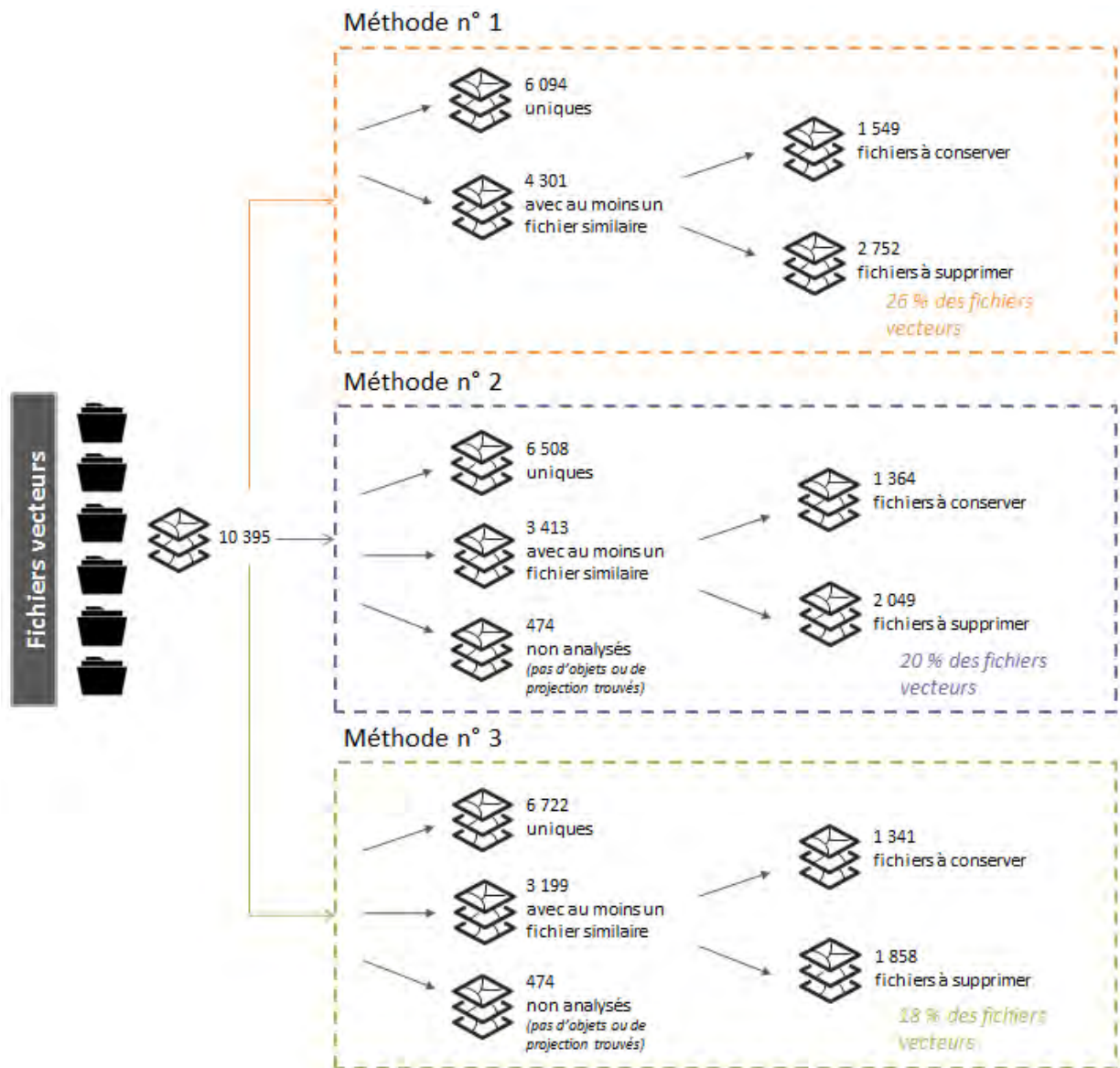


Figure 14 Synthèse de l'analyse des doublons des fichiers vecteurs avec trois méthodes différentes

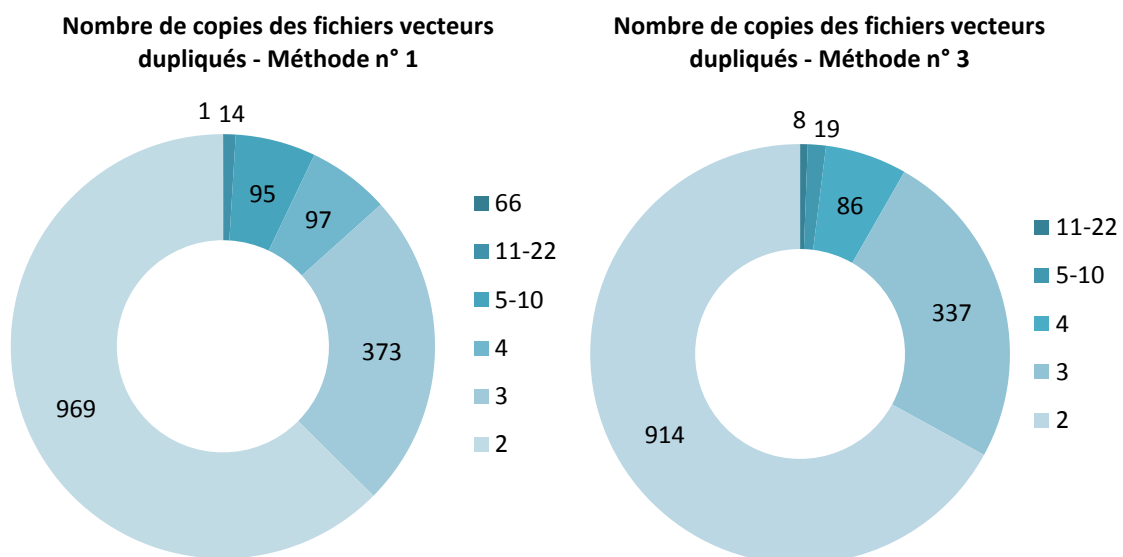


Figure 15 Nombre de copies des fichiers vecteurs dupliqués des méthodes n° 1 (à gauche) et n° 3 (à droite)

S'il ne fallait retenir qu'une seule méthode d'analyse des doublons, ce serait la méthode n° 3 qui est la plus restrictive. Toutefois, au vu des incertitudes et des résultats obtenus avec les trois méthodes, le nombre de fichiers vecteurs à supprimer est estimé à environ 20 %. Avant suppression, il faudra s'assurer que ces fichiers ne sont plus utilisés dans des projets QGIS ou qu'ils peuvent être substitués par leur copie.

Données rasters

Dans le cas des rasters il paraît plus pertinent d'effectuer une analyse des doublons uniquement sur le nom du fichier. Les noms de fichiers sont préalablement transformés en minuscules.

Plusieurs raisons expliquent ce choix :

- DicoGIS ne parvient pas à analyser les fichiers d'extension ecw : seuls le nom du fichier et le chemin sont indiqués, il est donc impossible de faire l'analyse des doublons sur le nom du fichier et le nombre de lignes et de colonnes si l'on souhaite conserver ces fichiers.
- Ce problème existe aussi sur certaines images géo-référencées au format .tif (115 fichiers au total).
- De plus, certains fichiers rasters sont des images géo-référencées pas nécessairement enregistrés dans la même orientation, ainsi les nombres de colonnes et de lignes sont inversés alors que le contenu du fichier est similaire.

Les résultats sont indiqués sur la Figure 16 ci-dessous.

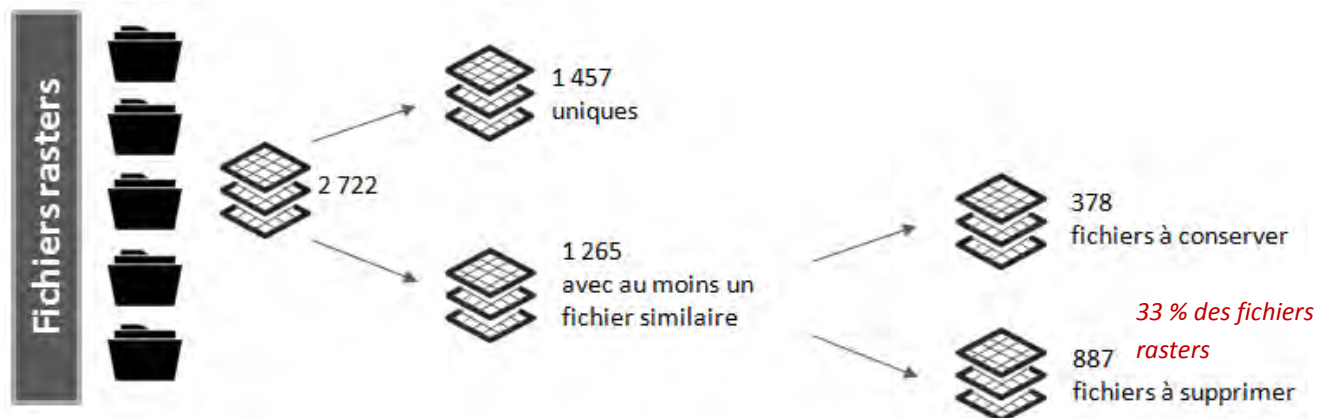


Figure 16 Analyse des doublons des fichiers rasters

Données CAO/DAO

Les fichiers de CAO/DAO type Autocad sont minoritaires sur le serveur partagé, et pas nécessairement bien analysés par DicoGIS. Toutefois pour obtenir un ordre de grandeur des doublons, l'analyse est également conduite sur ces fichiers. Seul le nom du fichier est considéré, après transformation en minuscules.

Les résultats sont indiqués sur la Figure 17 ci-dessous.

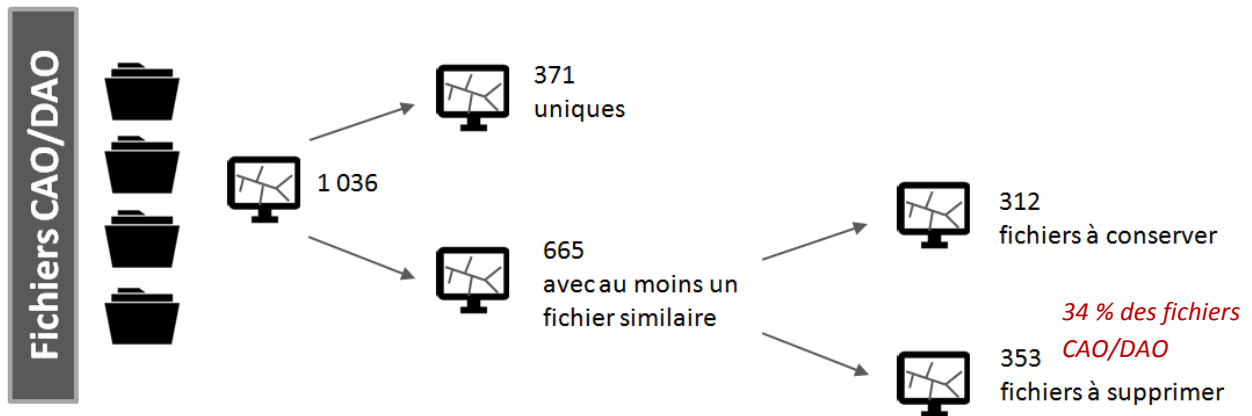


Figure 17 Analyse des doublons des fichiers CAO/DAO

2-2-1-3- Synthèse de l'analyse

Cette analyse a permis de quantifier, en tenant compte de toutes les limites listées dans la partie méthodologie, le nombre de fichiers cartographiques et l'espace occupé par ces fichiers sur le serveur partagé de la DTE. L'analyse a été effectuée par répertoire et par type de fichier (vecteur, raster et CAO/DAO).

Dans un second temps, suivant une méthodologie choisie, une analyse des doublons potentiels a été réalisée. Un extrait de listing des doublons obtenu pour chaque type de fichiers est présent en Annexe 3 (pour le type vecteur).

La Figure 18 ci-dessous permet de synthétiser les informations apportées par l'analyse des doublons.

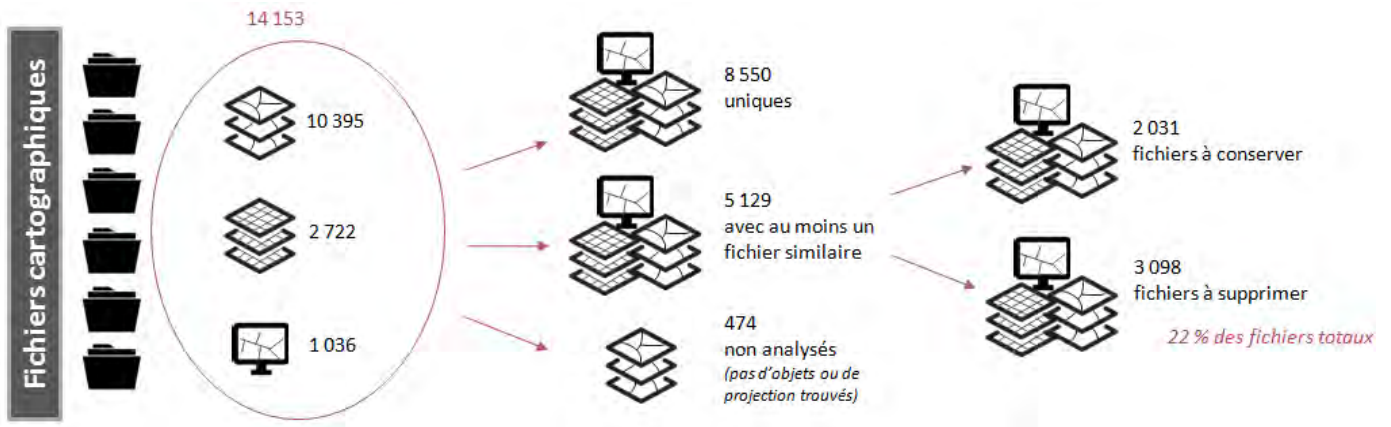


Figure 18 Synthèse de l'analyse des doublons des fichiers cartographiques

Une estimation de l'espace occupé par les fichiers à supprimer a également été effectuée (Tableau 2 ci-dessous).

	Taille estimée des fichiers à supprimer	Nombre de fichiers à supprimer	Nombre de fichiers à supprimer pour libérer 80 % de l'espace occupé par les copies
Vecteurs	25 Go	1 858	36
Rasters	984 Mo	887	24
CAO / DAO	Non représentatif	353	Non représentatif

Tableau 2 Estimation de l'espace occupé par les fichiers cartographiques à supprimer

Note : la taille des fichiers de CAO/DAO au format .dwg n'est pas donnée par DicoGIS, il est donc difficile d'estimer l'espace occupé par ces fichiers

En conclusion de l'analyse quantitative, on remarque que la problématique de la DTE concerne plutôt la dispersion des données cartographiques et la multiplicité des couches sur le serveur qu'une saturation de l'espace réseau du fait de fichiers cartographiques trop volumineux. Un tri ciblé permettrait toutefois de gagner environ 25 Go d'espace disque.

2-2-2- Entretien

Seize agents ont été rencontrés, répartis sur les quatre services de la DTE. Ces entretiens ont permis l'élaboration d'un tableau synthétique des données utilisées et produites par service. Ce tableau sert de base aux propositions effectuées pour l'amélioration de l'organisation des données de la DTE.

L'outil de catalogage GeoNetwork, en place avant le test de la solution Isogeo, n'est actuellement pas utilisé par les agents du fait de son manque d'ergonomie et de lisibilité (par exemple au niveau de la description des champs) pour les usagers de QGIS.

Lors des échanges, les agents mettent en avant les points suivants :

- La difficulté d'identifier les couches de travail et les couches validées
- Le doute sur l'utilisation de données à jour
- La difficulté de savoir quelles données sont présentes sur le serveur partagé
- La difficulté de trouver les données recherchées
- La difficulté pour trouver les cartes PDF produites et validées

A noter que le niveau d'utilisation des QGIS est disparate dans la direction, notamment car certains agents en ont une utilisation très occasionnelle. Ce sera un point à prendre en compte pour les propositions d'organisation et de structuration des données (développement d'interfaces de saisie, priorisation des actions).

Le Tableau 3 ci-dessous reprend les points positifs observés sur l'ensemble ou une partie des services ainsi que des pistes d'amélioration communes.

Points positifs observés	Pistes d'améliorations
<ul style="list-style-type: none">+ Utilisation des données sur le serveur partagé sig (O :) géré par la DSIN+ Utilisation de certaines données des bases de données partagées+ Volonté de structuration dans les dossiers thématiques et projets QGIS+ Utilisation des flux WFS/WMS par certains agents	<ul style="list-style-type: none">± Regroupement des données de mêmes caractéristiques (ex : bassins versants, canaux)± Suivi des données utilisées qui ne sont pas forcément à jour± Documentation de la donnée (externe et interne), création de fiches de métadonnées± Centralisation des données thématiques utilisées par toute la direction± Séparation des données géographiques des autres fichiers

Tableau 3 Etat des lieux général - Points positifs et pistes d'amélioration

Les observations par service sont détaillées en Annexe 4.

PARTIE 3 – PROPOSITIONS D'AMELIORATIONS

Rappel de l'objectif :

Par rapport aux observations réalisées pendant l'état des lieux l'objectif est double : identifier les données géographiques utilisées par tous les services afin de les centraliser et proposer des actions spécifiques à chaque service

3-1- RETOURS D'EXPERIENCE D'AUTRES COLLECTIVITES

Afin d'avoir une vision de l'organisation d'autres collectivités, j'ai organisé trois entretiens et suivi plusieurs webinaires sur l'architecture SIG de collectivités.

Entretiens :

- 22/04 : Rémi Delabruyère, Chargé d'études mobilités douces au service Environnement au Département du Calvados
- 03/05 : Clotilde Carton, Chef du service SIGMA (Service information géographique et modernisation des applications) et Grégory Preynat, Responsable des données environnement au Département de la Loire
- 26/05 : Jérémie Ory, Chef du pôle Système d'Information Géographique au Département du Calvados

Webinaires Isogeo :

- 06/05 : Céline Polisset, Administratrice SIG au Département du Loiret
- 06/05 : Jérémie Ory, Ingénieur en géomatique au Département du Calvados
- 10/06 : Christine Lenoir, Technicienne SIG et Jérémy Fleurisson, Responsable SIG à la Région Normandie

Les retours d'expérience des intervenants mettent en avant les points suivants :

Organisation	<ul style="list-style-type: none">- Réaliser un Schéma directeur de l'information géographique pour avoir une vue globale des données dans toute la collectivité et pas uniquement une direction.- Ne pas négliger la dimension animation / accompagnement du personnel- S'appuyer sur un agent avec une appétence pour le SIG dans chaque service quand cela est possible- Travailler sur le nettoyage de patrimoine de données, notamment l'archivage : nécessite une connaissance des données
Structuration	<ul style="list-style-type: none">- Intervenir en amont des projets des directions pour structurer les données et ainsi faciliter leur intégration- Fournir les modalités de structuration des données aux bureaux d'études quand la collectivité est Maître d'ouvrage- Créer une seule base de données centralisée avec les données externes et produites :<ul style="list-style-type: none">o Facilite le croisement de donnéeso Meilleure valorisation car les données sont décloisonnéeso Les agents ont une base de données unique à laquelle se référer

	<ul style="list-style-type: none"> – Figurer la structure de la base de données partagée, indépendamment de l'organisation de la collectivité en directions et services (celle-ci étant amenée à évoluer). Certaines collectivités se basent sur l'organisation des thèmes INSPIRE – Créer un serveur de fichiers dédié aux données cartographiques. Possibilité de créer un dossier commun à tous les services et des dossiers dédiés aux données métiers avec accès restreint – Dédier un serveur de fichiers avec accès restreint aux données cartographiques. Stocker les projets QGIS sur ce serveur permet également aux agents du pôle géomatique d'accéder aux projets des agents pour leur apporter une assistance technique à distance
Catalogage des données	<ul style="list-style-type: none"> – Ne pas sous-estimer le temps nécessaire à la documentation des données. La donnée est dynamique, la métadonnée également. Pour information, la Région Normandie indique avoir passé environ 2 mois à plein temps pour la mise en œuvre du catalogage des données et y consacre désormais environ un jour par mois – Cataloguer les données facilite l'intégration des agents : cela donne une vision des données utilisées et produites dans la collectivité – Identifier un agent ressource, dédié au catalogage des données – Organiser des journées de « meta-data party » pour sensibiliser les agents et créer ou mettre à jour les fiches de métadonnées

Tableau 4 Retours d'expérience pour l'organisation et la structuration des données géographiques

En complément, la Région Normandie a indiqué avoir créé un guide de saisie des métadonnées, ils sont enclins à le partager et à y apporter des améliorations.

Ces retours d'expérience ont pour but de mettre en avant quelques bonnes pratiques relevées lors des différents échanges. Toutefois, l'organisation de la structure SIG en collectivité est dépendante de nombreux facteurs, notamment la taille du service SIG, les moyens affectés et la présence du service SIG dans une direction métier ou dans un service transversal.

3-2- FACILITER L'USAGE DES DONNEES UTILISEES PAR TOUS

L'état des lieux a mis en avant plusieurs besoins :

Besoins généraux :

- la centralisation des données référentiel métier de la DTE
- l'identification d'un responsable pour chaque donnée référentiel, assurant un suivi des mises à jour
- l'utilisation de services de données (flux WFS/WMS) pour les données le permettant
- le catalogage des données de la DTE

Besoins liées aux services :

- l'utilisation des données suivies par la DSIN dans les BD partagées
- l'utilisation des données suivies par les directions dans base31
- le regroupement de certaines données dans une seule couche cartographique
- la structuration des données spécifiques en bases de données

Face à ces constats et en concertation avec les chefs de service, j'ai créé une bibliothèque des données au format Excel afin de recenser les principales données utilisées dans la direction et les possibilités d'accès à une donnée à jour. Ce travail a été réalisé sous Excel pour pouvoir être pris en main directement par les agents de la DTE, dans la phase de transition entre GeoNetwork et la solution Isogeo. Il n'est pas prévu de mise à jour à court terme de ce fichier, en l'absence d'un responsable des données géographiques identifié dans la direction.

Ce travail a eu lieu en deux temps :


- identifier les données à intégrer dans la bibliothèque de données, en croisant les informations recueillies dans chaque service de la DTE
- trouver l'accès possible à la donnée, en privilégiant un accès direct aux services du producteur de données. Cette phase a été particulièrement chronophage (environ 10 jours)

Le but de la bibliothèque de données étant de recenser les données maintenues à jour, les flux WFS et WMS ainsi que les données stockées dans les bases de données partagées du Cd 31 et maintenues à jour par la DSIN ou les services producteurs ont été privilégiées. Cependant pour les données sur les cours d'eau par exemple, la centralisation des données avec un responsable identifié serait plus rationnelle car le temps de chargement des entités via les flux WFS peut être important.

L'onglet d'informations de la bibliothèque de données est présent ci-dessous (Figure 19).

Informations utiles

Définition des colonnes de l'onglet "acces_donnees"

Provenance	Provenance de la donnée : externe (directement extraite d'une source externe), re-traitée (provenant de sources externes mais travail de concaténation ou de synthèse) ou produite au CD31
Thématique	Grande thématique de la donnée
Donnée	Nom de la donnée
Echelle	Echelle : France, Occitanie, Haute-Garonne, Bassin Adour-Garonne
Type	Typologie : référentiel topographique, thématique, donnée partenaire (envoi ou convention pour accès aux données)
Accès à la donnée	Quel est le moyen d'accès à la donnée ?
Stockage réseau interne	- Accès sur le réseau interne (réseau J: de la DTE ou réseau O: SIG)
Flux externe	- Accès via un flux distant externe (WFS - vecteur ou WMS - raster)
Base de données CD31	- Accès via une base de données partagée du CD 31
	 indique un accès restreint à la donnée
Lieu où trouver la donnée	Où retrouver l'accès à la donnée
Lien (si flux)	Si l'accès est via un flux externe, le lien à entrer dans QGIS est indiqué
Emplacement réseau ou base de données	Si l'accès est via le réseau interne ou une base de données partagée, le chemin d'accès est indiqué
Fréquence de mise à jour	Si l'accès est via le réseau interne, indication de la fréquence de mise à jour de la donnée (si définie)
Responsable	Responsable de la donnée (pour la mise à jour sur le réseau ou la complétion des tables de la base partagée)
Commentaires	Informations complémentaires à consulter

Ressources utiles

Ajout des flux WFS/WMS à QGIS (WebServices) https://noova.cd31.fr/intranet/hgi_130500/sig?histstate=6&portlet=hgi_130509

Ajout des bases de données partagées à QGIS (PostGis) https://noova.cd31.fr/intranet/hgi_201519/ajouter-des-donnees-postgis

L'espace collaboratif SIG (https://noova.cd31.fr/intranet/hgi_130500/sig) et notamment l'Espace documentaire regorge d'informations, n'hésitez pas à le consulter lorsque vous avez une interrogation !

Figure 19 Onglet « Infos » de la bibliothèque de données

Un extrait de la bibliothèque de données est présenté en page suivante (Figure 20). La bibliothèque de données complète est présentée en Annexe 5.

Réflexion sur l'utilisation des flux WFS/WMS

Les flux WFS et WMS permettent d'accéder à des données vectorielles ou raster stockées sur des serveurs distants. L'utilisation des flux se pose pour faciliter la maintenance des données et l'utilisation de données à jour.

Les limites de leur utilisation sont :

- la nécessité d'une connexion Internet,
- la fiabilité des serveurs sur lesquels sont hébergés les flux,
- le temps de chargement pour les couches comportant un nombre d'objets vectoriels important.

MAJ : 15/07/2021 (KL)							Accès à la donnée							
Provenance	Thématique	Donnée	Echelle	Type	Stockage réseau interne	Flux externe	Base de données CD31	Lieu où trouver la donnée	Lien du flux externe (si applicable)	Emplacement réseau ou base de données	Fréquence de mise à jour	Responsable	Commentaires	
Externe	Masses d'eau	Rapportage 2010 - Masse d'eau cours d'eau	France	référentiel thématique				SANDRE https://www.sandre.eaufrance.fr/atlas/srv/fre/catalog_search#/metadata/ae19118c-a4a6-4c51-bcc9-3146d336fc65	flux WFS : Filtre possible sur le code du bassin (Adour-Garonne = F ou code 05) https://services.sandre.eaufrance.fr/geo/sandre Couches : MasseDEauRiviere_VEDL2019_FXX MasseDEauRiviere_VRAP2016_FXX MasseDEauRiviere_VEDL2013_FXX MasseDEauRiviere_VRAP2010_FXX MasseDEauSouterraine_VEDL2019_FXX MasseDEauSouterraine_VRAP2016_FXX MasseDEauSouterraine_VEDL2013_FXX MasseDEauSouterraine_VRAP2010_FXX	Envisager de centraliser sur le réseau car le chargement des entités peut être long				
Externe		Rapportage 2010 - Masse d'eau souterraine												
Externe		Etat des lieux 2013 - Masse d'eau cours d'eau												
Externe		Etat des lieux 2013 - Masse d'eau souterraine												
Externe		Rapportage 2016 - Masse d'eau cours d'eau												
Externe		Rapportage 2016 - Masse d'eau souterraine												
Externe		Etat des lieux 2019 - Masse d'eau cours d'eau												
Externe		Etat des lieux 2019 - Masse d'eau souterraine												
Externe	Référentiel hydrographique à grande échelle - BD Topage	Cours d'eau	France	référentiel thématique			SANDRE https://www.sandre.eaufrance.fr/ > BD Topage	flux WFS : Filtre possible sur le code du bassin (Adour-Garonne = F ou code 05) https://services.sandre.eaufrance.fr/geo/topage?	Envisager de centraliser sur le réseau car le chargement des entités peut être long				La BD Topage est le nouveau référentiel hydrographique, plus précise que la BD Carthage (précision 1 m contre 30 m)	
Externe		Tronçons hydrographiques élémentaires												
Externe		Bassins versants topographiques												
Externe		Plans d'eau												
Externe		Noeuds hydrographiques												
Externe		Limites terre-mer												
Externe		Bassins hydrographiques												
Externe	Référentiel hydrographique à moyenne échelle - BD Carthage	Cours d'eau	France métropolitaine	référentiel thématique			SANDRE https://www.sandre.eaufrance.fr/ > BD Carthage	flux WFS : Seuls les flux Métropole sont fonctionnels. Faire un filtre "métropole" dans la barre de filtre https://services.sandre.eaufrance.fr/geo/eth	Envisager de centraliser sur le réseau car le chargement des entités peut être long					
Externe		Eléments hydrographiques de surface												
Externe		Laisses des basses eaux												
Externe		Laisses des eaux												
Externe		Laisses des hautes eaux												
Externe		Noeuds hydrographiques												
Externe		Plans d'eau												
Externe		Points d'eau isolés												
Externe		Régions hydrographiques												
Externe		Secteurs hydrographiques												
Externe		Sous secteurs hydrographiques												
Externe		Tronçons hydrographiques élémentaires												
Externe		Zones d'hydrographie de texture												
Externe	Zones hydrographiques													
Externe	Zones inondables	Zones inondables hors PPRi en Haute-Garonne	Haute-Garonne	référentiel thématique				flux WFS: http://ogc.geo-ide.developpement-durable.gouv.fr/wxs?map=/opt/data/carto/geoide-catalogue/1.4/org_38004/0f29ca4-9d97-4c1f-af77-211a5479bc44.internet.map						
Externe		Aléas Inondation des PPRN prescrits en Haute-Garonne	Haute-Garonne	référentiel thématique				flux WFS : http://ogc.geo-ide.developpement-durable.gouv.fr/wxs?map=/opt/data/carto/geoide-catalogue/1.4/org_38004/87b85aa5-3358-4023-8e68-7f07ac04b606.internet.map						
Externe		Aléas Inondation des PPRN approuvés en Haute-Garonne	Haute-Garonne	référentiel thématique				http://catalogue.geo-ide.developpement-durable.gouv.fr/catalogue/srv/fre/catalog_search#/search?resultType=details&from=1&to=120&orgName=DDT%2031&type=dataset&sortBy=title&sortOrder=reverse&any=ppr&fast=index&_content_type=json	flux WFS : http://ogc.geo-ide.developpement-durable.gouv.fr/wxs?map=/opt/data/carto/geoide-catalogue/1.4/org_38004/c48fca66-5508-4cc1-afa2-e51aefa3a093.internet.map					
Externe		Cotes des Plus Hautes Eaux Connues (isocotes) en Haute-Garonne	Haute-Garonne	référentiel thématique					flux WFS : http://ogc.geo-ide.developpement-durable.gouv.fr/wxs?map=/opt/data/carto/geoide-catalogue/1.4/org_38004/afa15fbc-4876-41d9-8805-6d965aa4252a.internet.map					
Externe		Zonages réglementaires des PPRN approuvés en Haute-Garonne	Haute-Garonne	référentiel thématique					flux WFS: http://ogc.geo-ide.developpement-durable.gouv.fr/wxs?map=/opt/data/carto/geoide-catalogue/1.4/org_38004/5768ff38-3dcb-4c04-a863-27f93aab2ef7.internet.map					
Externe	Réseau d'observation - gestion quantitative	Stations hydrométriques	France	référentiel thématique			SANDRE https://www.sandre.eaufrance.fr/atlas/srv/fre/catalog_search#/metadata/13f9ce88-7764-494e-ae8f2d737b5eaa8	flux WFS: http://services.sandre.eaufrance.fr/geo/hydratable : StationHydro_FXX_shp						
Externe	Réseau hydrographique	Canaux	Bassin Adour-Garonne	partenaire			Partenaire : RESEAU 31	Pas de flux	J:\05_INFORMATIQUE\02_CARTO_SIG\08_HYDROGRAPHIE\HYDROGEO\01_RIVIERES_CANAU\CANAU	Envisager de centraliser sur le réseau car donnée utilisée par de nombreux agents	A définir	A définir		

Figure 20 Extrait de la bibliothèque d'accès aux données utilisées par tous les agents de la DTE

3-3- VERS UNE STRUCTURATION DES DONNEES PRODUITES ET UTILISEES PAR LES SERVICES

Cette partie se concentre sur les possibilités d'amélioration de la structuration des données produites ou utilisées par chaque service afin de pouvoir les intégrer dans la BD partagée base31. Cela permet de partager les données avec QGIS, les applications du WebSIG, de diffuser des cartes sur le site Internet institutionnel et enfin les partager sur l'OpenData (Tableau 5).

Ces propositions ont été revues avec le service Géomatique et mon tuteur de stage puis présentées en réunion d'encadrement le 11 mai 2021. Suite à celle-ci, la priorité a été donnée à la structuration en base de données des données des itinéraires cyclables de loisirs du Schéma Directeur Cyclable, produites par la cellule Vélo du service Environnement.

Donnée	Proposition	Opportunité	Contrainte	Temps estimé	Priorité
Zones humides	Déjà structuré dans base31				
Itinéraires cyclables de loisirs (Schéma Directeur Cyclable) et points d'intérêt associés	Structurer les itinéraires cyclables dans une base de données opérationnelle pouvant alimenter base31	Existence d'un géostandard COVADIS (2016) Véloroutes et voies vertes et d'un schéma de données aménagements cyclables (2021) en vue de la constitution d'une Base nationale des aménagements cyclables		2-3 mois	
	Itinéraires du Réseau Express Vélo	Structurer les itinéraires du REV (au moins ceux sous MO du Département)	Adaptation des schémas de données existant pour la partie loisirs	1 mois	
Espaces Naturels Sensibles (ENS)	Déjà structuré dans base31				
ENS Forêt de Buzet	1- Créer un modèle de données pour la structuration des données liées à la forêt de Buzet 2- Créer une base de données (ou geopackage dans un 1 ^{er} temps)		Données existantes très dispersées, demande un inventaire détaillée des données à intégrer	2-3 mois	
Inventaires faunistiques et floristiques	Créer une base de données d'inventaires	Données actuellement bien structurées sous format Excel, facilite la création d'un modèle de données	Avoir une vision sur la mise à jour de la base	1 mois	

Donnée	Proposition	Opportunité	Contrainte	Temps estimé	Priorité
Plan Départemental d'itinéraire Pédestre et de Randonnée (PDIPR)	Si toutes les informations sont présentes dans la table existante : Créer une vue sur base31 comportant les informations à diffuser sur l'OpenData	- Table déjà existante dans base31 - Demande de l'OpenData pour la diffusion des données	Validation de la donnée diffusée avec la date de CP (implique passage en OpenData, ré-utilisation par d'autres directions)	< 1 mois	
	Si structuration à prévoir : - Créer un modèle de données - Créer une vue			1 mois	
Suivi annuel de la qualité des eaux	Créer un modèle de données pour les stations de suivi		Capacité à tenir à jour et croiser les données	1 mois	
Bassins versants	Intégrer en base de données		Définir l'échelle d'agrégation des masses d'eau élémentaires	1 mois	
Périmètres de gestion : Syndicats de rivière, eau potable et assainissement	- Créer un modèle de données pour les syndicats et leurs compétences - Créer des synthèses pour utilisation par les agents		Capacité à tenir à jour et croiser les données	2-3 mois	
Installations eau potable	Intégrer les captages, stations de production et réservoirs en base de données		- Données propriété des gestionnaires de l'eau, et bancarisées par l'ARS > convention de partage à actualiser - Capacité à tenir à jour et croiser les données	< 6 mois	
Installations assainissement	Intégrer les stations d'épuration en base de données		- Données bancarisées par l'AEAG - Capacité à tenir à jour et croiser les données	< 6 mois	

Tableau 5 Synthèse des propositions spécifiques aux services

PARTIE 4 – STRUCTURATION DES DONNEES DES ITINERAIRES CYCLABLES DE LOISIRS

Rappel de l'objectif :

En se basant sur les standards existants, créer une base de données pour la cellule Vélo du service Environnement pouvant être intégrée à la base de données des données produites par le Cd 31 : base31.

4-1- DEPLOIEMENT D'UNE BASE DE DONNEES SUR LES ITINERAIRES CYCLABLES DE LOISIRS

Rappel des objectifs

Dans le cadre de sa Stratégie des mobilités cyclables, le Département a pour objectif de développer le réseau cyclable du quotidien et de consolider le réseau cyclable touristique et de loisirs :

- ✓ Maître d'ouvrage des Réseau Express Vélo (REV) en cours d'étude hors du territoire de Toulouse Métropole
- ✓ Maître d'ouvrage d'un réseau d'itinéraires cyclables structurants de loisirs dans le cadre du Schéma directeur des itinéraires cyclables initié en 1997
- ✓ Financement des projets d'aménagements cyclables des collectivités

Source : Stratégie des mobilités cyclables en Haute-Garonne, décembre 2020

En vue de faciliter la valorisation des actions menées et le partage des données avec les autres acteurs du territoire sur la thématique de la mobilité (Région Occitanie, Toulouse Métropole, Tisséo), j'ai réalisé une base de données sur les aménagements cyclables, en priorisant les itinéraires cyclables de loisirs, au cœur de la Maitrise d'ouvrage départementale depuis 1997.

La méthodologie générale employée est présentée dans le schéma ci-dessous (Figure 21).

Cette méthodologie se base à la fois sur de la recherche documentaire, de la conception technique, des réunions de travail avec la cellule Vélo et enfin sur une phase de formation avec le technicien de la cellule Vélo qui reprendra ce travail à la fin de mon stage.

Les objectifs de cette base de données sont multiples : centraliser les données de la cellule Vélo pour capitaliser la connaissance technique et la développer, permettre la diffusion des données sur les aménagements cyclables dans le respect des schémas de diffusion, faciliter les échanges avec les partenaires et maintenir la disponibilité actuelle de la donnée sur les itinéraires cyclables de loisirs en interne, sur le WebSIG et en Open Data.

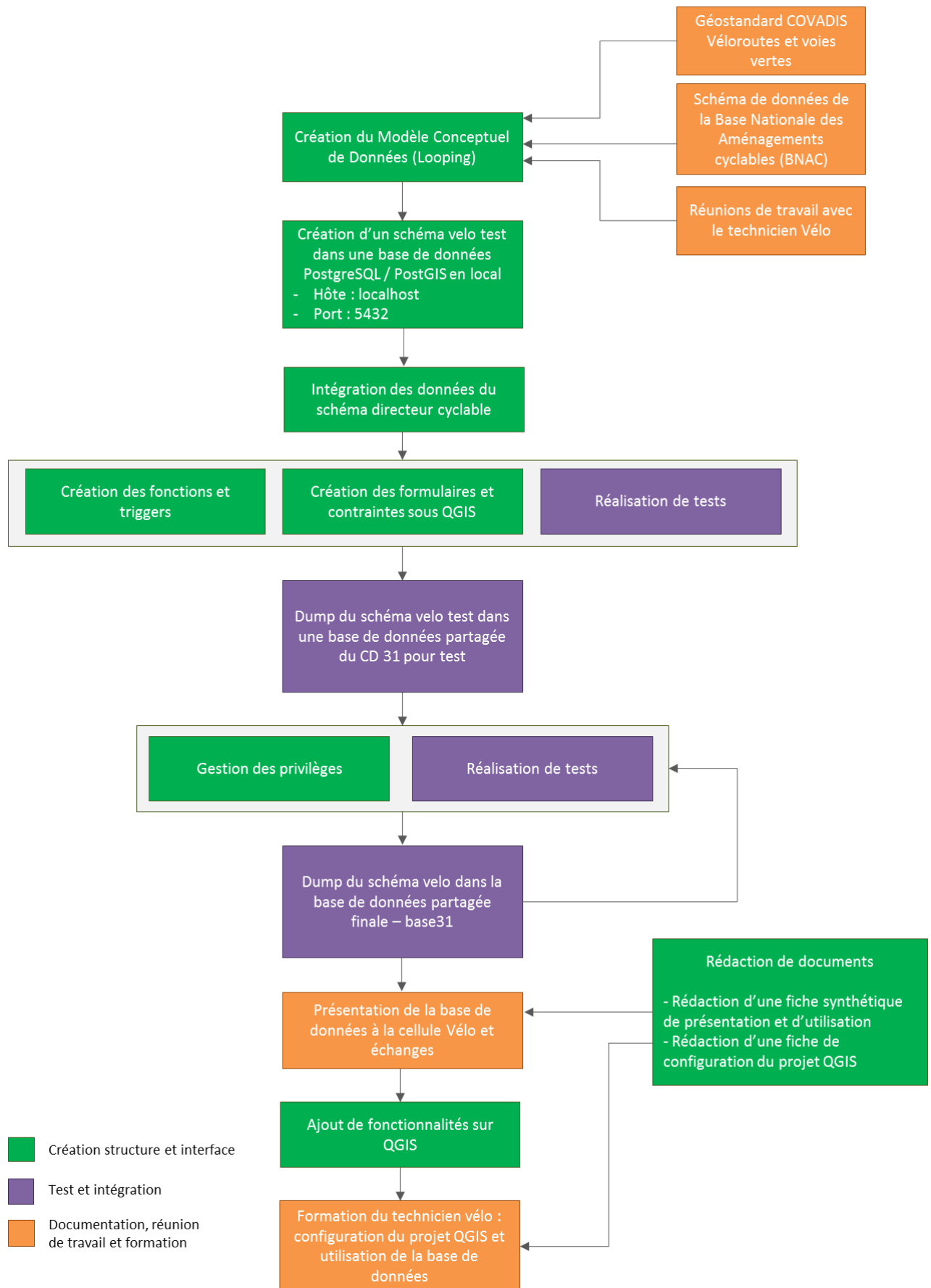


Figure 21 Méthodologie pour le déploiement de la base de données des itinéraires cyclables de loisirs

4-1-1- Création du modèle conceptuel de données

Les mobilités douces sont des thématiques en pleine croissance et portées par de nombreux acteurs, la première étape a été de chercher s'il existait déjà des schémas de données existants sur la thématique cyclable.

Deux schémas sont ressortis :

- Le modèle de données Véloroutes et voies vertes lié au géostandard Covadis (2014), développé dans le cadre de l'Observatoire National des Véloroutes et Voies Vertes (ON3V)⁵
- Le modèle de données des aménagements cyclables (2020) développé par le Ministère des transports et Vélo & Territoires en lien avec des collectivités aménageurs cyclables en vue de la constitution d'une Base nationale des aménagements cyclables (BNAC)⁶

Les objectifs des deux schémas sont différents, d'après Covadis, « *Le géostandard Véloroutes et voies vertes propose un noyau d'information décrivant le réseau français des véloroutes et voies vertes. Son objectif est d'apporter un vocabulaire et une structure de données aux acteurs impliqués dans la planification et la réalisation du réseau des véloroutes et voies vertes* » tandis que le schéma de données sur les aménagements cyclables a été réalisé afin d'« *harmoniser la description de ces derniers et faciliter la publication de ces informations en OpenData* », en vue de la diffusion des données au Point d'Accès National (voir encadré Point réglementaire).

Le géostandard a donc davantage une visée opérationnelle avec une description plutôt exhaustive des véloroutes et voies vertes, tandis que le schéma sur les aménagements cyclables vise l'harmonisation du modèle de diffusion des données sur les aménagements cyclables (sans tenir compte pour le moment de l'intégration de ces aménagements dans un schéma touristique ou de mobilité plus global).

Le modèle conceptuel de données (MCD) a été réalisé en se basant sur le géostandard covadis et en intégrant les attributs nécessaires à la diffusion des aménagements cyclables sur la BNAC.

Point réglementaire

« L'ouverture des données sur les aménagements cyclables nécessaires à l'information voyageur est une obligation légale, définie par le règlement délégué (UE) 2017/1926 concernant la mise à disposition de services d'informations sur les déplacements multimodaux.

Le règlement statue la création d'un Point d'Accès National par pays membre ainsi que la mise à disposition des données nécessaires à l'information voyageur sur celui-ci [...] à échéance du 1er décembre 2019. Ces obligations sont précisées en droit français par la loi d'orientation des mobilités (LOM). Elle désigne les collectivités territoriales comme étant responsables de la mise à disposition des données sur la plateforme transport.data.gouv.fr [...] »

Source : transport.data.gouv.fr

Le MCD (voir Annexe 6) présent dans le géostandard covadis est une base de travail. Dans ce géostandard, le modèle est également transcrit pour être déployé sous Mapinfo. C'est une adaptation de ces deux modèles qui est utilisée pour le MCD réalisé.

⁵ Détails du géostandard : <https://www.velo-territoires.org/observatoires/observatoire-national-des-veloroutes-et-voies-vertes/geostandard-velo/>

⁶ Détails de la démarche BNAC : <https://www.velo-territoires.org/observatoires/observatoire-national-des-veloroutes-et-voies-vertes/schema-de-donnees-amenagements-cyclables/>

Afin de mieux comprendre les interactions entre les différents éléments du géostandard véloroutes et voies vertes, j'ai réalisé un schéma de principe (Figure 22), reprise dans une fiche synthétique de présentation de la base de données vélo. Ce schéma permet de mieux appréhender la lecture du MCD.

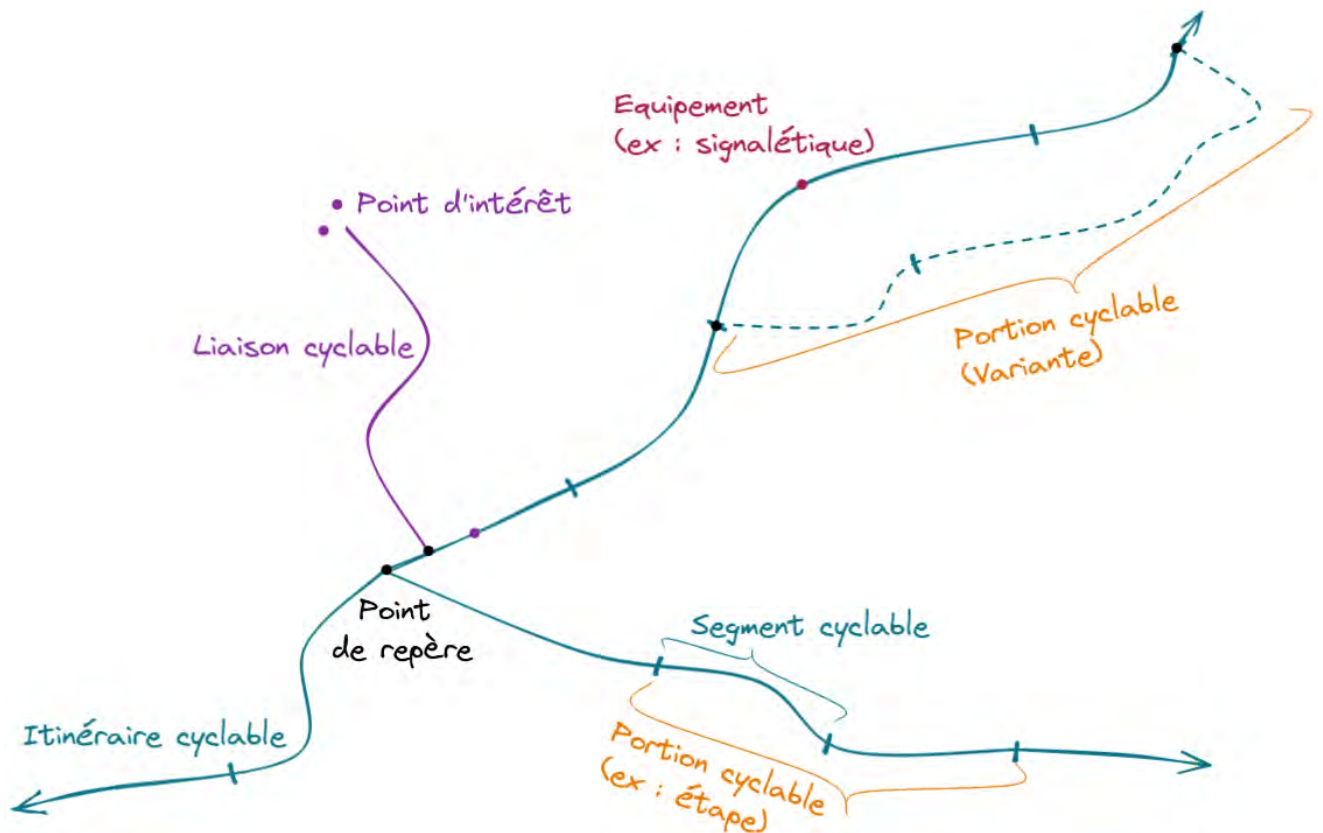


Figure 22 Schéma de principe des éléments repris dans les tables de la BD velo (K. Lesage, 2021, réalisé avec Excalidraw)

Les principales notions sont explicitées ci-dessous :

Segment cyclable

- élément le plus fin du réseau
- section homogène d'un itinéraire

Caractéristiques :

- un **état d'avancement*** (projet, tracé arrêté, travaux en cours, ouvert),
- un **revêtement** (lisse, meuble, rugueux),
- un **statut foncier** (route départementale, privée, etc.),
- un **statut réglementaire*** (voie verte, piste cyclable, etc.),
- n'appartient qu'à **une seule commune***

* indique un critère obligatoire

- Géostandard véloroutes et voies vertes
- Base nationale des aménagements cyclables
- Recherche additionnelle

Portion cyclable

- agrégat de segments cyclables
- types possibles : étape, variante, portion provisoire, portion observée

Un même segment peut appartenir à **plusieurs portions** cyclables.

Itinéraire cyclable

- composé d'une ou plusieurs portions
- agrégat des segments qui composent les portions auxquelles il est associé

Une portion ne peut appartenir qu'à **un seul itinéraire**.

Liaison cyclable

- possède sa **propre géométrie** (n'est pas un agrégat de segments)
- permet d'accéder à des points d'intérêt en dehors de l'itinéraire (ex : une gare ferroviaire, un site touristique, des sanitaires etc.)

En complément des deux standards, les documents issus de leur déploiement ont été utilisés : les notices de numérisation des véloroutes et voies vertes (Départements et Régions cyclables, 2014) et des aménagements cyclables (Vélo & territoires, 2020) ainsi que la présentation du schéma de données des aménagements cyclables de janvier 2021 (Vélo & territoires, 2021). Ces sources sont détaillées dans la BIBLIOGRAPHIQUE / SITOGRAPHIE.

Les principales réalisations sont les suivantes :

- Mutualisation des informations du géostandard véloroutes et voies vertes et du schéma de données sur les aménagements cyclables : l'élément commun est l'aménagement cyclable (aussi appelé segment cyclable)
- Ajout d'une table pour les équipements des itinéraires cyclables type signalétique ou borne : actuellement ces équipements ne sont pas recensés mais ils le seront à court ou moyen terme avec possiblement des photos géo localisées
- Adaptation des relations attributaires entre les tables : priorité à la relation spatiale quand cela est possible afin de simplifier le modèle de données (exemple : pas de relation attributaire entre le point d'intérêt et la portion cyclable)
- Création d'une seule table pour les points d'intérêt avec l'ajout d'un attribut « categorie_poi » au lieu de trois tables accès intermodal, service et tourisme

Les différentes sources d'informations consultées ainsi que la réflexion sur les améliorations possibles en concertation avec la cellule Vélo ont permis d'établir le modèle conceptuel de données en page suivante (Figure 23).

Le MCD est réalisé avec Looping⁷, ce qui permet de générer le script SQL pour intégration à la base de données PostgreSQL / PostGIS.

Un dictionnaire de données simplifié est présent en Annexe 9.

⁷ Looping est un logiciel de modélisation conceptuelle de données gratuit et libre d'utilisation <https://www.looping-mcd.fr/>

- Géostandard véloroutes et voies vertes
- Schéma de données des aménagements cyclables
- Informations additionnelles
- Liaison spatiale obligatoire
- - - Liaison spatiale facultative

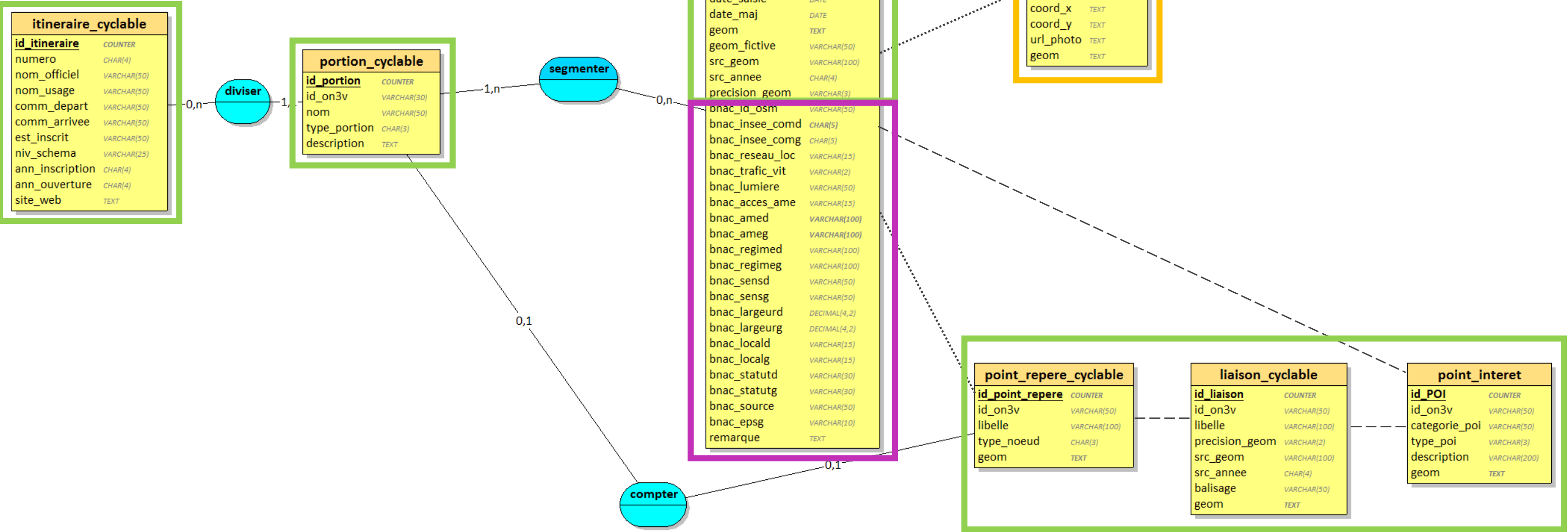


Figure 23 Modèle Conceptuel de données de la base de données Velo

4-1-2- Création de la base de données PostgreSQL / PostGIS

Pour créer la base de données à partir du MCD, le choix s'est porté automatiquement sur PostgreSQL avec la cartouche spatiale PostGIS. En effet, ce sont les technologies actuellement utilisées au Cd 31, il était donc nécessaire de s'appuyer sur celles-ci pour intégrer la base nouvellement créée à l'infrastructure existante. La nouvelle base de données sera intégrée à un nouveau schéma nommé « velo » dans la base de données partagée base31. Une fois les tables finalisées, les métadonnées associées seront créées via l'outil de catalogage d'Isogeo (voir PARTIE 5 - CATALOGAGE DES DONNEES).

Les versions utilisées sont : postgresql 9.6 et postgis 2.3.6

Qu'est-ce que PL/pgSQL?

PL signifie « Procedural language », langage de procédures. C'est un langage associé à PostgreSQL qui permet d'ajouter au SQL des fonctions plus complexes comme des boucles ou des instructions conditionnelles. Il peut être utilisé pour créer des fonctions et des triggers.

Création des listes de valeurs

Que ce soit dans le géostandard COVADIS ou dans le schéma de la BNAC, des listes de valeurs sont définies pour la valeur de certains attributs. Dans le géostandard COVADIS, des types énumérés sont définis alors que dans le schéma de la BNAC des tables sont créées pour chaque liste de valeurs.

Les types énumérés permettent de limiter la liste des valeurs d'une colonne à celles définies lors de la création du type. Ensuite ce type est affecté à la colonne, comme il serait fait avec un type *varchar*.

Exemple de requête :

```
CREATE TYPE velo.bnac_reseaulocval
as ENUM ( 'REV', 'STRUCTURANT', 'AUTRE' );
```

```
bnac_reseauloc velo.bnac_reseaulocval
```

Dans mon cas, j'avais dans un premier temps envisagé de créer des types énumérés. En effet, cela compliquait la structure de la base de créer a minima huit tables avec uniquement un identifiant et un libellé. Toutefois, les types énumérés ont des désavantages. D'après la documentation de PostgreSQL 9.6, « *les types énumérés sont principalement destinés à des listes de valeurs statiques. Il est possible d'ajouter de nouvelles valeurs aux types ou de les renommer, mais il n'est pas possible d'en supprimer* ».

Les types énumérés compliquent donc l'administration de la base de données. Pour l'implémentation de notre base il a été choisi de fonctionner avec une liste des valeurs directement dans QGIS, cela permettra de garantir l'intégrité de la base tout en allégeant l'administration des données, puisque les agents de la cellule Vélo pourront eux-mêmes modifier les listes.

Les valeurs sont saisies sous Excel, à partir des documents disponibles et les fichiers CSV générés à l'aide d'une macro Excel en VBA (Annexe 10).

Il est choisi de fonctionner directement avec les libellés plutôt qu'avec les codes pour faciliter l'exploitation de la base par les agents. Un traitement FME permettra de changer les libellés en code avant export des données vers les partenaires extérieurs.

Création des tables

Le script de création des tables est généré en partie automatiquement à partir du MCD créé sur Looping. Les modifications principales concernent la définition des géométries.

Les tables sont également créées sans contraintes autres que les clés primaires dans un premiers temps, les autres contraintes sont ajoutées après intégration des données.

Exemple de requête pour la création de la table `itineraire_cyclable` :

```
CREATE TABLE velo.itineraire_cyclable(  
  id_itineraire SERIAL,  
  numero CHAR(4),  
  nom_officiel VARCHAR(50),  
  nom_usage VARCHAR(50),  
  comm_depart VARCHAR(50),  
  comm_arrivee VARCHAR(50),  
  est_inscrit VARCHAR(50),  
  niv_schema VARCHAR(25),  
  ann_inscription CHAR(4),  
  ann_ouverture CHAR(4),  
  site_web TEXT,  
  PRIMARY KEY(id_itineraire)  
);
```

Création des contraintes

Des contraintes de clés étrangères sont créées sur les tables le nécessitant. L'ajout des instructions `ON UPDATE CASCADE` et `ON DELETE CASCADE` est essentielle pour ne pas engendrer une erreur lors de la suppression de l'entité parente. Ainsi l'entité enfant est supprimée simultanément.

Exemple de requête :

```
ALTER TABLE velo.rel_portion_segment  
ADD CONSTRAINT fk_rel_portion_segment_id_segment FOREIGN KEY(id_segment)  
REFERENCES velo.segment_cyclable(id_segment)  
ON UPDATE CASCADE  
ON DELETE CASCADE;
```

Création des index

Des index spatiaux sont créés sur les tables présentant une géométrie ligne / polygones ainsi que sur les vues spatiales. Des index sont automatiquement créés par PostgreSQL sur les contraintes de clés primaires. Pour le moment aucun autre index n'a été créé mais à l'usage si certains champs font l'objet de recherches régulières, il faudra les indexer également pour accélérer les recherches.

Exemple de requête :

```
CREATE INDEX idx_geom_segmentcyclable  
ON velo.segment_cyclable USING gist(geom) ;
```

Création des fonctions et triggers

Dans la base de données, certaines actions doivent être déclenchées lors de l'ajout, la modification ou la suppression d'une donnée. Dans ces cas des fonctions sont créées, associées à des « triggers » pour que les fonctions se déclenchent automatiquement. Les fonctions sont créées à l'aide du langage de procédure PL/pgSQL.

Des fonctions avec trigger sont créées pour :

- Compléter l'attribut « categorie_poi » des points d'intérêt en fonction du type de point d'intérêt
- Ajouter le code EPSG sur le segment cyclable pour inclure automatiquement l'information sur la référence spatiale demandée dans les standards

NB : ici on utilise la fonction ST_SRID(geom) mais on aurait pu accéder à la projection via la colonne « srid » de la vue spatiale geometry_columns stockée dans le schéma public

Exemples de requête :

```
-- Remplissage automatique de la colonne categoriePOI en fonction du type
de POI
CREATE OR REPLACE FUNCTION velo.update_categorie_poi() RETURNS trigger AS
$$
    BEGIN
        UPDATE velo.point_interet
        SET categorie_poi = (CASE
                                WHEN type_poi in ('OFFICE DE
TOURISME', 'HEBERGEMENT', 'SITE TOURISTIQUE') THEN 'TOURISME'
                                WHEN type_poi in ('GARE
ROUTIERE', 'GARE FERROVIAIRE', 'GARE MARITIME',
'GARE FLUVIALE', 'GARE AEROPORTUAIRE', 'STATION DE TRANSPORT URBAIN') THEN
'ACCES INTERMODAL'
                                WHEN type_poi in
('RESTAURATION', 'MERCHE', 'SANITAIRES', 'RELAIS INFO SERVICE',
'RELAIS VELO', 'STATION DE GONFLAGE', 'REPARATION, LOCATION DE VELO',
'AIRE DE PIQUE-NIQUE', 'POINT D'EAU', 'BANC', 'PARKING', 'POUBELLE') THEN
'SERVICE'
                                END);
        RETURN NULL; -- pour les trigger de type 'After'
    END;
$$ LANGUAGE plpgsql;

DROP TRIGGER IF EXISTS update_categorie_poi ON velo.point_interet ;
CREATE TRIGGER update_categorie_poi AFTER INSERT OR UPDATE OF type_poi ON
velo.point_interet
FOR EACH STATEMENT EXECUTE PROCEDURE velo.update_categorie_poi();
```

NB : à partir de PostgreSQL v11, il existe une instruction « execute function », ici on utilise PostgreSQL v9.6 et donc l'instruction « execute procedure ».

Création des vues spatiales

Les portions cyclables et itinéraires cyclables ne possèdent pas de géométrie propre, celle-ci est issue de la géométrie des segments cyclables. Cependant, il est nécessaire aux agents de pouvoir visualiser le tracé des itinéraires et des portions.

Vue spatiale ou vue matérialisée ?

Une **vue spatiale** n'est pas stockée sur le disque, elle est dynamique, actualisée à chaque interrogation. Pour des bases de données de taille importante (plusieurs milliers d'entités au moins), le temps d'actualisation peut donc être important. Pour limiter ce temps, il est nécessaire d'indexer les colonnes utilisées.

Une **vue matérialisée** est écrite sur le disque (comme une table classique). Pour être mise à jour, elle doit être rafraîchie via une instruction `REFRESH MATERIALIZED VIEW`. Pour les bases de données de taille importante, cela permet d'avoir directement accès à la vue, sans actualisation à chaque interrogation et donc améliorer la performance. Il faut ensuite mettre à jour la vue matérialisée selon une fréquence définie (ex : chaque soir) ou via des trigger. Niveau gestion des droits, seul le propriétaire de la vue peut la rafraîchir.

Dans notre cas, la base de données n'est pas de taille importante, des vues spatiales ont été créées.

Exemple de requête :

```
-- Vue pour les itinéraires cyclables (pour la géométrie)
DROP VIEW IF EXISTS velo.vs_itineraire_cyclable ;
CREATE VIEW velo.vs_itineraire_cyclable AS
    SELECT itineraire_cyclable.id_itineraire, numero, nom_officiel,
nom_usage, comm_depart, comm_arrivee,
    est_inscrit, niv_schema, ann_inscription, ann_ouverture, site_web,
    st_multi(st_union(segment_cyclable.geom)) as geom_union,
    round(st_length(st_multi(st_union(segment_cyclable.geom)))) as
longueur_iti
    FROM velo.itineraire_cyclable
    LEFT JOIN velo.portion_cyclable ON
itineraire_cyclable.id_itineraire = portion_cyclable.id_iti
    LEFT JOIN velo.rel_portion_segment ON portion_cyclable.id_portion =
rel_portion_segment.id_portion
    LEFT JOIN velo.segment_cyclable ON rel_portion_segment.id_segment =
segment_cyclable.id_segment
    WHERE segment_cyclable.avancement = 'OUVERT'
    GROUP BY itineraire_cyclable.id_itineraire
    ORDER BY numero ASC ;
```

Commentaires des tables et des colonnes de chaque table

Dès qu'il est possible de fournir de l'information associée aux données directement dans la base de données, il est nécessaire de la renseigner. Chaque table de la base de données a été commentée ainsi que chaque colonne de chaque table.

Cela permet :

- D'avoir une description des tables lors de l'ajout d'une couche PostgreSQL (Figure 24)
- Que tous les commentaires soient repris dans QGIS lors de l'ajout des tables de la base de données
- Le pré-remplissage des fiches de métadonnées lors de l'utilisation d'un outil de catalogage

Dans les colonnes, le commentaire indiqué sert comme alias dans QGIS. Un script Python (Annexe 11) a été utilisé via la console Python de QGIS pour utiliser les commentaires des colonnes en tant qu'alias.

velo		
velo	equipement	Equipements associés aux segments d'itinéraires cyclables de loisirs : signalétique, borne etc.
velo	itineraire_cyclable	Itinéraires cyclables de loisirs du schéma directeur cyclable : seules les parties en service sont indiquées.
velo	liaison_cyclable	Liaisons cyclables associées aux itinéraires : permettent de rejoindre des points d'intérêt (tourisme, service, transports)
velo	point_interet	Points d'intérêt (tourisme, service, transports) se situant sur l'itinéraire cyclable ou à proximité
velo	point_repere_cyclable	Points structurant du réseau des itinéraires cyclables de loisirs (début/fin d'étape, intersection d'itinéraires, bifurcation etc.)
velo	portion_cyclable	Portion cyclable : division de l'itinéraire cyclable pouvant correspondre à une étape, une variante de l'itinéraire principale, une portion ...
velo	rel_portion_segment	Table de relation entre les segments cyclables et les portions cyclables. Un segment doit obligatoirement être associé à une portion po...
velo	segment_cyclable	Element le plus fin de l'itinéraire cyclable de loisirs : correspond à une section homogène d'un itinéraire et possède une géométrie.
velo	vs_bnac_aménagement_cyclable	Vue spatiale pour le format de diffusion de la Base Nationale des Aménagements Cyclables (transport.gouv.fr / Velo & territoires)
velo	vs_itineraire_cyclable	Vue spatiale pour la visualisation des itinéraires cyclables de loisirs ouverts
velo	vs_portion_cyclable	Vue spatiale pour la visualisation des portions cyclables

Figure 24 Description des tables de la base de données Velo dans QGIS

4-1-3- Tests et vérifications réalisés

De nombreux tests ont été réalisés pour arriver à une version finalisée de la base de données :

- Fonctionnement attendu des fonctions et triggers
- Incrémentation des séquences
- Longueurs de champs adéquates : notamment lors du passage du fonctionnement par code au fonctionnement par libellé, certaines longueurs n'étaient plus adaptées
- Prise en compte des contraintes définies
- Relations entre les tables correctement paramétrées (détails dans la partie 4-2-)

4-1-4- Gestion des droits

Un des atouts du fonctionnement en base de données est la possibilité d'avoir une gestion fine des droits d'accès aux données intégrées à la base.

D'après la documentation de PostgreSQL, « le standard SQL définit les concepts d'utilisateurs et de rôles mais les considère comme des concepts distincts et laisse la spécification des commandes de définition des utilisateurs à l'implémentation de chaque moteur de bases de données ». Dans PostgreSQL, « les utilisateurs et les rôles sont unifiés au sein d'une même entité ».

Pour s'approcher du standard, l'idée est tout de même de créer des rôles de groupe et des rôles de connexion (comme présenté dans la Figure 25 ci-dessous). Ce sera ainsi le rôle de groupe qui aura les droits sur les schémas, tables et vues et le rôle de connexion sera affecté à un ou plusieurs rôles de groupe.



Figure 25 Gestion des rôles de connexion et de groupes dans la base de données (source : SIGEA, 2021)

Au niveau du code SQL, la différence se traduit ainsi sur un exemple issu de la gestion des droits de la base de données Velo :

Rôle de connexion « adm_velo »	Rôle de groupe « sig31_maj_ecologie_itineraire_cyclable »
<pre>CREATE ROLE adm_velo WITH LOGIN NOSUPERUSER INHERIT NOCREATEDB NOCREATEROLE</pre>	<pre>CREATE ROLE sig31_maj_ecologie_itineraire_cyclab le WITH NOLOGIN NOSUPERUSER INHERIT</pre>

<pre>NOREPLICATION ENCRYPTED PASSWORD 'xxxx' VALID UNTIL 'infinity'; GRANT sig31_consultation, sig31_consultation_dadre, sig31_maj_dadre_velo, sig31_maj_ecologie_itineraire_cyclab le TO adm_velo;</pre>	<pre>NOCREATEDB NOCREATEROLE NOREPLICATION VALID UNTIL 'infinity';</pre>
--	--

L'utilisation de la clause « INHERIT » (valeur par défaut) permet d'indiquer que le rôle « *pourra utiliser tout droit dévolu par un rôle dont il est membre direct ou indirect* ».

La BD Velo est intégrée à base31, j'ai donc utilisé les rôles de connexion et les rôles de groupe existant dans base31 pour gérer les droits de la base de données Velo, intégrée au schéma « velo » nouvellement créé.

Nom	Description
ALL	Autorise tous les privilèges (selon le type de l'objet)
SELECT	Permet la sélection sur tout ou partie des colonnes d'une table (requête SELECT)
INSERT	Permet d'insérer un enregistrement dans une table
UPDATE	Permet la mise à jour des champs d'une table.
DELETE	Permet d'effacer un enregistrement de la table
REFERENCES	Droit requis pour création de clé étrangère (où il faut faire référence à une table tierce)
TRUNCATE	Permet d'effacer une table ou un ensemble de tables (récupération immédiate de l'espace disque sans VACUUM) Truncate = Delete+Vacuum pour faire simple.
TRIGGER	Permet de créer de nouveaux déclencheurs associés à la table
TEMPORARY	Permet de créer des tables temporaires
EXECUTE	Permet d'exécuter une fonction (précise)

Figure 26 Description d'une partie des privilèges pouvant être accordés sous PostgreSQL (source : SIGEA, 2021)

Les privilèges suivants sont définis :

	Super utilisateur	Rôle de groupe	
	postgres	sig31_maj_ecologie_itineraire_cyclable rôle de connexion : adm_velo	sig31_consultation_ecologie sig31_public_ecologie rôle de connexion : invite
Schéma velo	ALL	USAGE	USAGE
Tables (x 8)	ALL	DELETE, INSERT, SELECT, UPDATE	SELECT
Vues (x 3)	ALL	USAGE	USAGE
Séquences (x 7)	ALL	USAGE	/
Trigger fonctions (x 2)	EXECUTE	EXECUTE	/

Tableau 6 Droits et privilèges définis sur la BD Velo

NB : ne pas oublier d'accorder le droit d'accès au schéma

La description complète de la gestion des droits peut être consultée sur la documentation de PostgreSQL. En Annexe 12, deux tableaux issus de la documentation résument les abréviations et privilèges d'accès selon les types d'objets.

4-2- FONCTIONNEMENT AVEC QGIS

Paramétrage des relations

La première chose à paramétrer lors de la configuration du projet QGIS sont les relations attributaires entre les tables, conformément au modèle conceptuel de données.

Pour ce faire dans QGIS : *Projet > Propriétés > Relations* (voir Figure 27 ci-dessous)

Nom	Couche de référence	Champ de référence	Couche de référencement	Champ référençant	Id	Force
base31_rel_point_repere_portion_obs	portion_cyclable	id_portion	point_repere_cyclable	id_portion	point_repe_id_po...	Association
base31_rel_portion_itineraire	itineraire_cyclable	id_itineraire	portion_cyclable	id_iti	portion_cy_id_iti...	Association
base31_rel_portion_segment_portion	portion_cyclable	id_portion	rel_portion_segment	id_portion	rel_portio_id_port...	Association
base31_rel_portion_segment_segment	segment_cyclable	id_segment	rel_portion_segment	id_segment	rel_portio_id_seg...	Association

Figure 27 Paramétrage des relations attributaires entre les tables de la BD Velo

Paramétrer les relations permet de faciliter la saisie des clés étrangères dans les couches de référencement (ou couches « enfant ») via un menu déroulant dans la saisie en mode formulaire (voir Figure 28 ci-dessous).

The screenshot shows a form titled 'Identification de la portion' with a sub-tab 'Segments cyclables associés (relation)'. Fields include 'Identifiant de la portion' (15), 'Nom' (V81 - Projet Mane), 'Description' (NULL), and 'Type de portion' (PORTION PROVISoire). A dropdown menu for 'Itinéraire associé' is open, showing options like 'V81 Piémont Pyrénéen', 'V80 Canal des deux mers', 'V83 Trans Garona', 'V84 Rigole de la plaine', 'V85 Véloroute Vallée du Tarn-Causse-sud Cévennes', and 'Vallée de l'Arize'. There is also a 'facultatif' section with 'Identifiant ON3V' set to NULL.

Figure 28 Exemple de l'intérêt de l'ajout des relations attributaires dans QGIS – relation portion cyclable / itinéraire

Paramétrage des champs

En complément de la définition des types de champs dans la base de données, certaines fonctionnalités ont été intégrées via QGIS pour plus de souplesse dans l'administration de la base, tout en garantissant son intégrité et une facilité d'utilisation pour les agents.

Quelques exemples de paramétrage :

- Les champs remplis via des triggers dans la base de données sont paramétrés comme « non éditable ».
- Valeur par défaut pour la date de saisie (pour la date de mise à jour, cocher la case *Appliquer la valeur par défaut sur la mise à jour*) :

The screenshot shows the 'Défauts' dialog box. The 'Valeur par défaut' field contains the SQL expression 'to_date(now())'. Below it, the 'Aperçu' (Preview) shows '2021-07-30'. There is a checkbox labeled 'Appliquer la valeur par défaut sur la mise à jour' which is currently unchecked.

- Valeur par défaut pour le code INSEE de la commune en fonction de la localisation de la numérisation (basé sur la couche commune de la BD Topo V3)

```

aggregate (
layer:= 'commune_b4aec260_dfa7_4870_9662_ad39d7cd4222',
aggregate:='majority',
expression:=code_insee,
concatenator:=', ',
filter:=intersects($geometry, geometry(@parent))
)

```

- Valeur par défaut pour l'auteur de la numérisation :

@user_full_name

- Liste de valeurs sur certains champs en fonction des valeurs définies dans le geostandard COVADIS et le schéma de la BNAC. Celles-ci sont reprises des documents associés aux standards ou saisies sur Excel puis exportées en csv par macro VBA (voir Annexe 10)

Pour la définition d'une valeur par défaut en fonction du contenu d'une colonne devant être remplie par l'utilisateur, il n'est pas possible de définir cette valeur via l'instruction *DEFAULT* dans PostgreSQL. Un essai a été réalisé via un trigger, mais l'utilisateur ne pouvait pas modifier la valeur imposée par le trigger. Il est donc plus simple de paramétrer la valeur par défaut dans QGIS.

Exemple pour le champ du statut d'avancement des aménagements gauche et droit :

```

CASE
WHEN "avancement" = 'OUVERT' THEN 'EN SERVICE'
WHEN "avancement" = 'PROJET' THEN 'EN PROJET'
WHEN "avancement" = 'TRAVAUX EN COURS' THEN 'EN TRAVAUX'
WHEN "avancement" = 'TRACE ARRETE' THEN 'PROVISOIRE'
ELSE NULL
END

```

Des formulaires ont également été créés pour chaque table via le *Drag and Drop Designer* afin de catégoriser les champs et faciliter le remplissage.

Options d'accrochage

Un segment cyclable ne doit appartenir qu'à une seule commune, l'accrochage a donc été défini pour que les segments s'accrochent aux limites communales. Globalement, l'accrochage est défini sur toutes les couches au niveau des sommets et des segments et l'auto-accrochage est désactivé.

Sémiologie



Sur les conseils de mon référent technique, une symbologie particulière est affectée aux segments cyclables afin de visualiser le début (point) et la fin du segment (flèche) ainsi que le sens de numérisation (sens de la flèche)(Figure 29).

Pour la symbologie des itinéraires j'ai repris celles définies par la cellule Vélo.

Figure 29 Extrait de la BD Velo - fond de carte : plan IGN v2

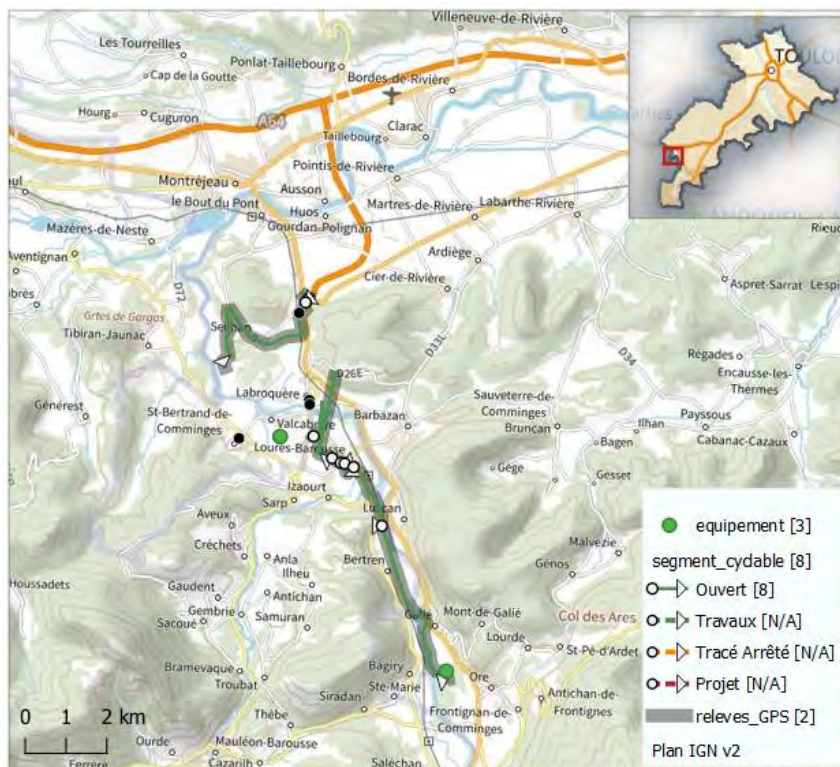
4-3- RELEVÉS TERRAIN AVEC QFIELD

Lors de la présentation de la BD Velo à la cellule Vélo, la question de pouvoir remplir la base de données directement sur le terrain a été évoquée par les agents. Actuellement, les agents enregistrent les traces GPS des itinéraires cyclables avec une tablette avec l'application OSMAnd, mais n'effectuent pas de relevés d'éléments attributaires.

L'application QField⁸ est utilisée par d'autres directions dans le cadre de numérisation sur le terrain. QField est une solution mobile libre et gratuite développée par l'entreprise suisse OPENGIS.ch. Tous les formulaires de saisie et la symbologie ont déjà été paramétrés via QGIS, une simple adaptation pour le terrain de ces formulaires permet une utilisation avec QField, c'est donc cette solution qui a été retenue pour effectuer quelques tests.

Les tests ont été effectués un mois avant la fin de mon stage. Par conséquent il a été décidé de fonctionner via un geopackage plutôt qu'avec une connexion directe à la base de données PostGIS, afin de s'exonérer des problèmes de synchronisation avec la base Velo. Les données sur les ponctuels pourront être intégrées à base31 en copiant les données du geopackage à l'issue des relevés terrain dans la couche issue de base31, en mode administrateur. Les noms de colonnes étant les mêmes, le test réalisé a été concluant. Pour les données des segments cyclables, il faudra effectuer un retraitement (voir les perspectives d'amélioration en page suivante) avant de les intégrer de cette même façon.

Un projet QGIS spécifique pour le terrain a été réalisé. Seules les couches utiles ont été sélectionnées et les formulaires ont été simplifiés pour s'adapter au terrain. La priorité pour ce test était de relever le statut des segments cyclables (voie verte, piste cyclable, route etc.) et leur revêtement (lisse, meuble ou rugueux).



Le test a été réalisé mardi 10 août 2021 sur l'itinéraire Trans-Garona de Gourdan-Polignan aux environs de Bagigny, à l'aide d'une tablette Samsung Galaxy Tab A – modèle SM-T515.

Figure 30 Extrait du résultat du test de relevés pour la BD Velo avec QField

⁸ <https://qfield.org/>

Les relevés réalisés sont les suivants :

- Equipement type jalonnement ou panneaux pédagogiques avec une photo géolocalisée associée
- Trace GPS de l'ensemble du parcours
- Trace GPS segmentée

A l'enregistrement d'un suivi, QField demande de paramétrer un intervalle de temps ou de distance, ou les deux. Un point sera enregistré selon ces paramètres. Le test est réalisé avec l'enregistrement d'un point toutes les 2 secondes.

Observations à l'issue du test :

- Sur une partie du parcours, QField a planté et la trace ne s'est pas enregistrée
- La numérisation des équipements a été effectuée de façon aisée, et les résultats peuvent être directement intégrés à base31
- La numérisation de la trace GPS du parcours dans son ensemble est satisfaisante, même si quelques écarts inopinés sont observés
- La numérisation de la trace GPS segmentée est de faible qualité topologique, avec de nombreux écarts inopinés (Figure 31). Cela pourrait être dû à des pertes de signal et aux enregistrements successifs avec arrêt et reprise du suivi GPS
- Le technicien en charge des relevés terrain doit bien s'approprier les listes de caractéristiques des différents champs et si nécessaire définir un ensemble de règles pour savoir quelle caractéristique s'applique



Figure 31 Illustration des écarts inopinés observés sur la trace GPS globale (en noir) et segmentée (en rouge)

De plus, QField impose de rentrer les paramètres de suivi GPS à chaque nouveau suivi. Lors de l'enregistrement de nombreuses traces, cette façon de fonctionner n'est pas optimale. Ce sujet a été évoqué via la déclaration d'un problème sur le GitHub de QField⁹ le 27 juillet 2021 et l'équipe de développement a indiqué avoir réalisé cette amélioration le 20 août 2021 pour une mise en œuvre dans la version bêta de QField 2.0.

⁹ <https://github.com/opengisch/QField/issues/2060>

Tests supplémentaires à réaliser :

- Modifier les paramètres de suivi pour essayer d'améliorer la trace GPS
- Réaliser un test sur une zone plus complexe avec davantage de changements de caractéristiques

4-4- PERSPECTIVES D'AMELIORATION**Partie base de données :**

- Echanger avec les partenaires aménageurs cyclables (Toulouse Métropole, Tisseo etc.) sur les bases de données de chacun
- Développer des vues spatiales en lien avec les besoins d'autres directions (Transport, Routes)
- Créer une table liée à la maintenance. La direction de routes a en charge la maintenance des itinéraires cyclables (panneaux et revêtements). La cellule Vélo aimerait pouvoir signaler dans une table les besoins en maintenance lors des sorties terrain afin de les partager à la direction des routes. Ce serait une table de gestion uniquement, accessible aux agents concernés.

Partie en lien avec les relevés terrain :

- Améliorer le lien entre les relevés terrain sur les segments cyclables et les informations à intégrer dans la base de données.

En effet, un segment est une section homogène de l'itinéraire mais sur le terrain seules quelques informations sont relevées pour le moment : le statut et le revêtement. Il faudra a minima re-segmenter ces segments par rapport aux limites communales et possiblement selon d'autres caractéristiques comme le propriétaire ou le gestionnaire. De plus, la trace GPS nécessite des corrections topologiques.

Il faut donc envisager la chaîne de traitement dans son ensemble, en envisageant une couche de travail distincte de la couche finale de base31.

- Travailler avec une numérisation des segments sur plan pour éviter les nombreuses corrections topologiques et relever les caractéristiques des segments sur le terrain via des ponctuels pour ensuite segmenter la numérisation en fonction de ces caractéristiques.

PARTIE 5 - CATALOGAGE DES DONNEES

Rappel de l'objectif :

Documenter les tables de la base de données Velo nouvellement créée en utilisant l'outil de catalogage en cours de déploiement au Cd 31.

Après une phase de test, le Département est entré en phase de déploiement de la solution de catalogage basée sur la solution d'Isogeo en juillet 2021.

La méthodologie générale pour le catalogage des données est présentée dans la Figure 32 ci-dessous. Elle est basée sur deux produits Isogeo : le Scan FME¹⁰ et l'OpenCatalog¹¹. En complément, un plugin pour QGIS¹² permet d'effectuer une recherche de données dans les catalogues créés et d'ajouter directement les données souhaitées au projet.

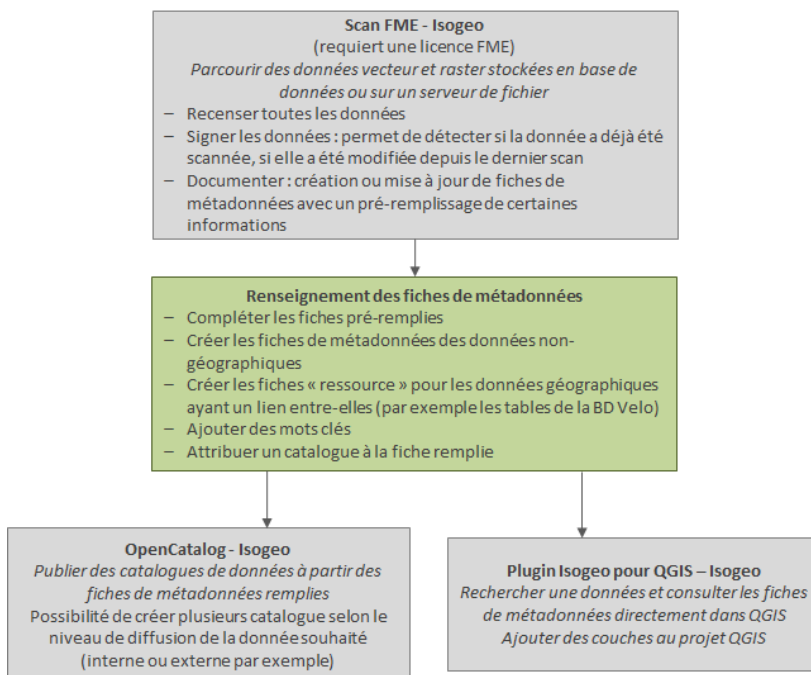
J'ai suivi cette méthodologie pour cataloguer les tables et vues spatiales de la BD Vélo possédant une géométrie. Pour les tables sans géométrie, il faut créer manuellement les fiches de métadonnées (la donnée n'est donc pas liée à la fiche). La création d'une fiche « ressource » permet d'associer les différentes tables entre-elles.

Point réglementaire

« La convention d'Aarhus (1998), puis la directive INSPIRE (2007) et un chapitre du code de l'environnement (2011) imposent aux autorités publiques de :

- rendre leurs données environnementales géographiques accessibles au public en publiant sur Internet ces données et les métadonnées correspondantes,
- les partager entre elles »

Source : cours d'Emilie LERIGOLEUR – master SIGMA
– 15 décembre 2020



■ Utilisation d'un produit Isogeo

■ Actions complémentaires à réaliser

L'Annexe 14 montre quelques exemples issus du catalogage.

Les avantages par rapport à l'utilisation de DicoGIS seul sont nombreux : la création automatique de fiches de métadonnées pré-complétées, l'identification de données modifiées ou ajoutées et l'intégration à l'infrastructure globale du SIG.

Figure 32 Méthodologie de catalogage des données basée sur la solution Isogeo

¹⁰ <https://www.isogeo.com/nos-produits/Scan>

¹¹ <https://www.isogeo.com/nos-produits/OpenCatalog>

¹² <https://www.isogeo.com/nos-produits/Plugins-Widgets>

CONCLUSION ET PERSPECTIVES

L'état des lieux des données géographiques de la Direction de la Transition Ecologique (DTE) du Conseil départemental de la Haute-Garonne (Cd 31) a permis de quantifier les données présentes sur le serveur partagé grâce à l'utilisation de l'utilitaire DicoGIS (Isogeo) puis d'une analyse sous Excel et Python (bibliothèque Pandas). Les résultats obtenus mettent en avant la dispersion des données sur le serveur et la présence de copies de l'ordre de 25 % des fichiers totaux.

En complément, des entretiens ont été menés avec les agents utilisateurs de QGIS afin d'établir une bibliothèque d'accès aux données géographiques principales utilisées et produites par la DTE. Les accès privilégiés aux données sont les flux distants WFS et WMS, les données stockées sur le serveur partagé dédié au SIG et les données stockées dans les bases de données partagées du Département. La base de données partagée centrale pour le Cd 31 est base31, dédiée aux données produites en interne. Ce choix a été fait afin de recenser uniquement les données maintenues à jour.

Une évolution de cette bibliothèque de données serait l'ajout de données thématiques stockées sur le serveur avec des responsables identifiées au sein de la DTE. Ma présence au sein de la DTE a permis d'insuffler un nouvel élan à la problématique d'organisation des données géographiques mais un investissement plus important de la DTE est nécessaire pour maintenir une dynamique afin d'assurer le maintien d'une organisation pérenne optimale. Il n'est actuellement pas prévu de faire évoluer l'organisation de la DTE au niveau de la gestion des données géographiques.

L'état des lieux et les discussions avec les agents m'ont permis de proposer dix améliorations dans la structuration de données utilisées ou produites, plus ou moins longues et complexes à mettre en œuvre. Une réunion de direction a permis de prioriser ces actions et le choix a été fait de travailler sur la structuration des données de la cellule Vélo sur les itinéraires cyclables de loisirs, intégrés au Schéma directeur cyclable du Cd 31.

En me basant sur le Géostandard Véloroutes et voies vertes (COVADIS – Vélo & Territoires, 2014), le Schéma de données des aménagements cyclables (Transport.gouv.fr – Vélo & Territoires, 2020) et des réunions de travail avec la cellule Vélo, j'ai pu concevoir un modèle conceptuel de données puis créer une base de données PostgreSQL / PostGIS. Cette base de données a été intégrée à base31.

La difficulté a été de trouver les bons compromis pour garantir l'intégrité de la base de données, tout en minimisant le temps d'administration et en assurant une facilité d'utilisation pour les agents, non géomaticiens. En complément de la configuration de la base PostGIS, des formulaires de saisie avec la définition de valeurs par défauts, de contraintes de champs puis des symbologies particulières ont été définis sous QGIS pour permettre une utilisation la plus aisée possible à l'agent qui reprendra la base de données.

Le sujet a pu être finalisé en configurant un projet QGIS dédié aux relevés terrain et à une utilisation avec la solution mobile de QGIS, QField. Par manque de temps pour gérer les potentiels problèmes de synchronisation, le choix a été fait de fonctionner avec le format geopackage plutôt que la base PostGIS directement. Un test terrain a été réalisé avec le technicien Vélo sur une partie de l'itinéraire cyclable Trans-Garona, prochainement inaugurée par le Cd 31.

Une des perspectives pour cette base de données est d'échanger avec d'autres aménageurs cyclables en Haute-Garonne : Toulouse Métropole, Tisséo et autres EPCI partenaires. En effet une harmonisation des bases de données des partenaires avait été envisagée en 2020 mais suite à

l'avancement du schéma de données de la Base Nationale sur les Aménagements Cyclables et à sa sortie imminente (fin 2020), cette démarche a été abandonnée. Ce travail pourrait intéresser les partenaires car il permet de mutualiser deux standards et de répondre à la fois aux besoins de structuration des données des itinéraires cyclables de loisirs et du Réseau Express Vélo.

D'un point de vue personnel, ce stage m'a beaucoup apporté, que ce soit techniquement ou au niveau des relations humaines.

Durant les deux à trois premiers mois de stage j'ai travaillé sur un état des lieux où j'ai dû développer une méthodologie alliant analyse technique et entretiens avec les agents. Même si le volet technique a demandé du travail, le plus difficile a été de mener les entretiens avec les agents, de poser les bonnes questions et d'extraire les informations pertinentes de nos entrevues. De plus, cette phase s'est étalée dans le temps car les agents n'étaient pas nécessairement disponibles. Il a également fallu adapter mon discours selon le niveau de pratique des agents en SIG et réexpliquer les notions clés. Toutefois j'ai toujours été bien accueillie et j'ai également apprécié découvrir les missions de chacun.

J'ai ensuite travaillé sur la thématique qui me tenait à cœur : la conception et l'administration d'une base de données. Grâce aux compétences acquises en bases de données pendant la formation, j'ai pu être directement autonome sur les notions de base et les approfondir grâce à la lecture de la documentation en ligne et aux explications de mon encadrant technique. J'ai pu faire mes propres choix, tout en comptant sur le suivi et la disponibilité de mes encadrants.

Le côté Référent SIG pour la DTE m'a également beaucoup plu puisque j'ai pu être le support technique des agents pendant ces six mois de stage. Cela m'a permis de travailler sur des problématiques auxquelles sont confrontés les agents au jour le jour et d'acquérir des connaissances supplémentaires dans le but de répondre à leurs questionnements.

Au final, j'ai également appris que dans une structure avec peu de ressources humaines allouées à la géomatique, il faut faire des compromis entre la technique et la facilité d'utilisation pour les agents tout en maintenant l'intégrité de l'infrastructure SIG.

BIBLIOGRAPHIQUE / SITOGRAPHIE

Boutry M., *Gestion des itinéraires cyclables*, [En ligne] 2018,

<https://prezi.com/fkx5czdpudyd/gestion-des-itineraires-cyclables/>

Commission de validation des données pour l'information spatialisée, 2014 (version 1), *Standard de données Véloroutes et voies vertes*, 57 p

Conseil Départemental de la Haute-Garonne, 2020, *Rapport d'activités 2019*, 153 p

Conseil Départemental de la Haute-Garonne, 2018, *Comité de pilotage SIG élargi*, 11p

Conseil Départemental de la Haute-Garonne, *Le Conseil départemental : mode d'emploi*, [en ligne]

<https://www.haute-garonne.fr/cantons>

Conseil Départemental de la Haute-Garonne, *Missions du Cd31*, [En ligne]

https://noova.cd31.fr/intranet/hgi_53840/missions-du-cd31

Conseil Départemental de la Haute-Garonne, *Direction de la transition écologique*, [En ligne]

https://noova.cd31.fr/intranet/hgi_27279/direction-de-la-transition-ecologique

Départements et Régions cyclables, 2014, *Les véloroutes voies vertes et itinéraires cyclables – Notice de numérisation terrain*, 19 p

Documentation PostgreSQL :

- Gestion des droits : <https://docs.postgresql.fr/9.6/ddl-priv.html>,
<https://docs.postgresql.fr/13/ddl-priv.html>, <https://docs.postgresql.fr/9.6/sql-creatorole.html>
- Types énumérés : <https://www.postgresql.org/docs/9.6/datatype-enum.html>
- Type serial : <https://www.postgresqltutorial.com/postgresql-serial/>
- PL/pgSQL : <https://www.postgresql.org/docs/current/plpgsql.html>

Etalab, *Schéma de données des aménagements cyclables*, [En ligne] 2020,

<https://schema.data.gouv.fr/etalab/schema-amenagements-cyclables/latest.html>

Etalab, *Github - Schéma de données des aménagements cyclables*, [En ligne] 2020,

<https://github.com/etalab/schema-amenagements-cyclables>

Région Rhône-Alpes, 2010, *Système d'Information Géographique (SIG) – Les enjeux de la société de l'information pour le développement des territoires rhônalpins*, 47 p

SIGEA, *Administration de PostgreSQL / PostGIS – La gestion des droits*, [En ligne] 2021,

https://data.sigea.educagri.fr/download/sigea/supports/PostGIS/distance/perfectionnement/M02_administration_gen_web/co/M02_administration_4.html

The QField Project / OPENGIS.ch, *QField - la solution [Q]GIS mobile*, [En ligne]

<https://qfield.org/docs/fr/>

Vélo & territoires, 2020, *Notice de numérisation terrain – Aménagements cyclables*, 19 p

Vélo & territoires, 2021, *Le schéma de données des aménagements cyclables (présentation du 29 janvier 2021)*, 25 p

ACRONYMES

BAD	Biodiversité et Aménagement Durable
BD	Base de Données
BNAC	Base Nationale des Aménagements Cyclables
Cd 31	Conseil Départemental de la Haute-Garonne
CDRP	Comité Départemental de Randonnée Pédestre
CZH	Conservatoire des Zones Humides
DGD	Direction Générale Déléguée
DSIN	Direction des Systèmes d'Information et du Numérique
DTE	Direction de la Transition Ecologique
ENS	Espaces Naturels Sensibles
EPCI	Etablissement Public de Coopération Intercommunale
GEMAPI	GEstion des Milieux Aquatiques et Prévention des Inondations
INSEE	Institut National de la Statistique et des Etudes Economiques
MCD	Modèle Conceptuel de Données
PDIPR	Plan Départemental d'Itinéraires de Promenade et de Randonnée
PIR	Pôle Infrastructures et Réseaux (ancien nom du service Eau et Assainissement)
REV	Réseau Express Vélo
SAGE	Schéma d'Aménagement et de Gestion des Eaux
SDAEP	Schéma Directeur d'Alimentation en Eau Potable
SDAGE	Schéma Directeur d'Aménagement et de Gestion des Eaux
SGBD	Système de Gestion de Base de Données
SMEA	Syndicat Mixte de l'Eau et de l'Assainissement de Haute-Garonne (ancien nom de RESEAU 31)