





### Rapport de stage

Master 1 - Géographies Numériques Promotion 2020/2021



Auteure: Cendrine HOARAU

Tuteur professionnel : Camille MONCHICOURT

Tuteur académique: Thierry JOLIVEAU



# MISE A JOUR, STRUCTURATION ET DIFFUSION DES DONNEES DU PARC NATIONAL DES ECRINS

Stage effectué entre mars et juillet 2021



## **SOMMAIRE**

SOMMAIRE	2
REMERCIEMENTS	4
GLOSSAIRE	5
INTRODUCTION	6
CONTEXTE	<u>7</u>
Le Parc national des Écrins	7
LA LOGIQUE DE CHAÎNE DE TRAVAIL AU PNE	
LES OBJECTIFS DU STAGE	
STRUCTURATION ET MODERNISATION DU SIG	12
Création d'une BDD des capteurs du PNE	12
Co-construction du Modèle de données de la BDD	
Insertion des données dans PostGreSQL	14
Calculs spatiaux automatisés avec PostGIS	16
Mise à jour des données de référence du PNE	18
Actualisation des formats vecteurs	19
Actualisation des formats rasters	23
MISE EN FORME DES DONNÉES ET FACILITATION TECHNIQUE DES AGENTS DU PNE	24
Fusion et découpage des fonds de carte	24
Mise à jour des projets-type QGIS	27
Création d'un profil utilisateur QGIS	28
Montée en compétences des agents en gestion de l'information géographique	30
DÉPLOIEMENT D'OUTILS DE GESTION ET DE CONSULTATION DES DONNÉES	32
Création d'un Web SIG avec Lizmap	32
PUBLICATION D'UNE CARTE DANS UNE PAGE WEB AVEC LEAFLET	34
OUVERTURE ET DIFFUSION DES DONNEES	37
Les enjeux de l'open data au PNE	38
Définition et objectifs de l'open data	38

L'OPEN DATA ET LE CATALOGAGE	39
Publication de jeux de données	41
Publier sur une plateforme open data	41
Référencer sur un catalogue et renseigner les métadonnées	47
Favoriser l'interopérabilité des données	47
Automatiser la publication des données	49
BILAN	50
Résultats et pistes d'amélioration	50
Montée en compétences	51
Perspectives	52
BIBLIOGRAPHIE	52
WEBOGRAPHIE	52
INDEX DES FIGURES	54
INDEX DES ANNEXES	54

#### REMERCIEMENTS

En premier lieu, je tiens à remercier vivement l'encadrant de ce stage, **Camille Monchicourt**, pour son accueil, sa bienveillance et sa confiance. Le partage de son expertise et de ses connaissances ainsi que son investissement m'ont été d'une aide très précieuse.

Je remercie également les membres du pôle SI, **Théo Lechémia** et **Vincent Pietri** pour leur disponibilité et leur apport d'expertise technique tout au long du stage.

Merci à **Clotilde Sagot** pour sa collaboration, sa gentillesse et la confiance qu'elle m'a accordée.

Un grand merci à tous les collègues avec qui j'ai partagé le bureau des « geeks » et le Good Morning SI : Elie Bouttier, Adrien Pajot, Baptiste Garcin, Mathilde Leclerc, Corentin Lange et Jean-Paul Milcent, pour votre soutien, votre partage de connaissances et surtout votre bonne humeur qui a contribué à rendre ce stage stimulant!

Merci aussi à l'ensemble du service scientifique, particulièrement **Richard Bonet**, chef du service, pour m'avoir accueilli au sein de cette équipe dynamique et motivée.

Je souhaite remercier l'ensemble du personnel du Parc du siège et des secteurs que j'ai pu rencontrer et notamment le groupe de « jeunes et moins jeunes du PNE » avec qui j'ai partagé de bons moments de convivialité. Ce stage au sein du Parc national des Écrins a été une expérience très enrichissante et formatrice tant sur le point personnel que professionnel.

Je tiens également à remercier **Amandine Sahl** et **Idrissa Djepa** au Parc national des Cévennes pour le temps qu'ils m'ont accordé afin que je progresse techniquement en géomatique.

Je tiens également à remercier l'équipe encadrante du master et notamment **Thierry Joliveau** pour leur disponibilité et leur suivi tout au long de l'année et notamment dans cette période exceptionnelle de pandémie mondiale. Enfin, merci à **Jeremy Kalsron** pour son aide précieuse.

#### **GLOSSAIRE**

**PNE**: Parc national des Ecrins

**OGC**: Open Geospatial Consortium, consortium international qui développe et promeut des standards ouverts pour faciliter l'interopérabilité des contenus, des services et des échanges dans le domaine de la géomatique et de l'information géographique.

**SIG** : Système d'Information Géographique. Système d'information conçu pour recueillir, stocker, traiter, analyser, gérer et présenter tous les types de données spatiales et géographiques.

Base de données spatiale : Base de données permettant de stocker des informations géolocalisées (point, ligne, polygones) grâce à un type de données nommé « Geometry ».

**Capteur**: dispositif qui détecte un phénomène déterminé. Ce phénomène est le plus souvent d'ordre physico-chimique (ex : température, pression...) mais peut aussi être un mouvement physique. Les valeurs observées par les capteurs sont transformées en un signal électrique et transmises à un système de traitement des données.

**Station météo** : ensemble de capteurs qui enregistrent et fournissent des mesures physiques et des paramètres météorologiques liés aux variations du climat. Ces capteurs sont placés dans un boîtier faisant office d'abri météorologique. Les données de ces capteurs sont stockées dans une centrale d'acquisition, se trouvant également sur la station.

**Open data** : donnée numérique ouverte dont l'accès et l'usage sont laissés libres aux usagers.

**Métadonnées** : information descriptive servant à définir, décrire ou qualifier une donnée.

**PostgreSQL** : système de gestion de base de données relationnelle (SGBD) distribué sous licence libre.

**PostGIS**: extension du SGBD PostgreSQL, permettant le stockage et la manipulation d'informations géolocalisées.

#### INTRODUCTION

Ce rapport vise à rendre compte des travaux effectués lors de mon stage de M1, réalisé dans le cadre du Master de géomatique « Géographies Numériques » en partenariat avec l'Université Lumière Lyon 2, l'Université Jean Monnet à Saint-Étienne et l'École Normale Supérieure de Lyon et qui s'est déroulé au siège du Parc national des Ecrins dans la ville de Gap, dans les Hautes-Alpes.

J'ai effectué ce choix de stage dans la mesure où il constituait pour moi l'occasion de conforter mon projet professionnel. En effet, je souhaite d'une part garder un lien avec le secteur de l'environnement dans mon métier de géomaticienne, et d'autre part travailler dans un contexte collaboratif et moderne. Il m'a donc paru essentiel d'avoir une expérience professionnelle qui puisse regrouper ces éléments. C'est dans cette optique que j'ai candidaté au Parc national des Ecrins qui représente pour moi un lieu idéal d'exercice de mes futures fonctions.

Durant ce stage, j'ai pu me rendre compte de la place que tient un géomaticien au sein d'un établissement public tel qu'un parc national. Cela s'est révélé formateur dans la mesure où j'ai développé la polyvalence que nécessite le métier. Au PNE, le géomaticien est l'intermédiaire entre d'un côté les agents de terrain et chargés de mission, et de l'autre, les informaticiens. Pour répondre au mieux aux besoins et compétences de chacun, le pôle SI (Système d'Informations) du PNE a modernisé sa chaîne de travail, ce qui lui a permis de développer des outils open source aujourd'hui reconnus et largement utilisés à l'échelle du territoire national<sup>1</sup>. C'est dans ce cadre que j'ai pu approfondir ma connaissance du métier autant sur l'aspect technique et théorique, que sur la posture à adopter pour devenir une géomaticienne avertie.

Afin de pouvoir avoir un aperçu global du métier de géomaticien, il a été convenu avec mon tuteur de stage que je travaille sur plusieurs missions au cours du stage qui permettent chacune d'approfondir un aspect du métier. De ce fait, j'ai articulé mon rapport de stage autour de ces missions. J'exposerai dans un premier temps le contexte du stage en apportant des éléments d'information sur le PNE et sa logique de chaîne de travail ainsi que les objectifs du stage. Puis, je développerai les travaux de structuration des données et de modernisation du SIG du PNE effectués durant le stage qui ont permis par la suite leur ouverture et leur diffusion sur internet via le développement d'outils web.

<sup>1</sup> Biodiversité : une plateforme et des données présentées à la secrétaire d'Etat, février 2021, <a href="https://www.ecrins-parcnational.fr/actualite/biodiversite-plateforme-donnees-presentees-secretaire-etat">https://www.ecrins-parcnational.fr/actualite/biodiversite-plateforme-donnees-presentees-secretaire-etat</a>

### **CONTEXTE**

Les avancées technologiques récentes ont conduit à l'explosion de la quantité de données numériques produites. Ce nombre croissant de données induit une réflexion nécessaire autour de leur structuration et de leur stockage. La production de données s'accompagne également depuis une quinzaine d'années d'une réflexion autour de leur partage et leur diffusion, dans une dynamique européenne en faveur de l'open data. C'est dans ce cadre que le PNE a fait évoluer son système d'informations pour favoriser le partage de connaissances<sup>2</sup> auprès du plus grand nombre, notamment en utilisant et partageant des outils "libres".

## Le Parc national des Écrins

#### ◆ <u>Un parc national</u>



Figure 1 : Les 11 Parcs nationaux de France, <a href="http://www.parcsnationaux.fr/">http://www.parcsnationaux.fr/</a>

La France compte aujourd'hui onze parcs nationaux, couvrant près de 9,5 son territoire. Ces aires d'exception, créées à partir d'espaces terrestres ou maritimes, reconnues, aussi bien à l'échelle nationale qu'internationale, en raison de leur patrimoine naturel, culturel et paysager remarquable. Les objectifs de ces institutions sont la gestion et la protection des richesses naturelles (la faune et la flore essentiellement), ainsi que la sensibilisation du public à la découverte et au respect patrimoine<sup>3</sup>.

Historique de la création des onze parcs nationaux : la Vanoise (1963), Port-Cros (1963), les Pyrénées (1967), les Cévennes (1970), les Ecrins (1973), le Mercantour (1979), la Guadeloupe (1989), La Réunion (2007), la Guyane (2007), les Calanques (2012) et le Parc national de forêts (2019).

<sup>2</sup> Le Parc national des Écrins "libère" des données !, décembre 2018, https://www.ecrins-parcnational.fr/actualite/parc-national-ecrins-libere-données

<sup>3</sup> Les parcs nationaux : 11 espaces naturels protégés d'exception, 2021, <a href="http://www.parcsnationaux.fr/fr/des-decouvertes/les-parcs-nationaux-de-france/les-parcs-nationaux-11-espaces-naturels-proteges">http://www.parcsnationaux.fr/fr/des-decouvertes/les-parcs-nationaux-de-france/les-parcs-nationaux-11-espaces-naturels-proteges</a>

Les parcs nationaux sont rattachés à l'Office français de la biodiversité (OFB) afin de renforcer l'action collective pour préserver la biodiversité. Les missions de gestion et de protection du milieu naturel amènent les agents des parcs nationaux à collecter des centaines de milliers de données naturalistes : ils réalisent en permanence l'inventaire et le suivi des espèces se trouvant sur leur territoire. Historiquement, ces données contribuaient à alimenter des études scientifiques et bénéficiaient d'un usage principalement interne aux parcs. Depuis l'instauration de la « charte » des parcs nationaux en 2006, une dynamique d'ouverture et de diffusion des données a vu le jour. Ces données doivent non seulement être mises à disposition des partenaires institutionnels (ministères, régions, etc.), mais également être accessibles au grand public.

#### ♦ Le Parc national des Écrins

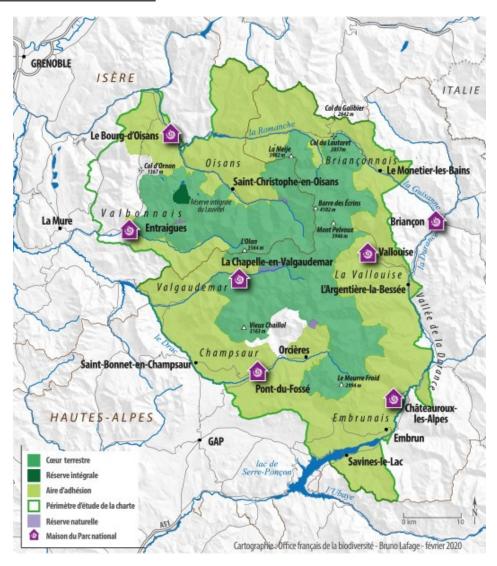


Figure 2 : Cartographie du Parc national des Ecrins © Bruno Lafage / OFB

Le Parc national des Écrins (PNE) se compose du siège se trouvant à Gap (Charance) et de maisons du parc se trouvant dans les différentes vallées qui composent son territoire. Le territoire du PNE s'étend sur les vallées de l'Oisans et du Valbonnais pour sa partie Iséroise et du Valgaudemar, du Champsaur, de l'Embrunais, de la Vallouise et du Briançonnais dans le département des Hautes-Alpes. Le cœur du Parc national des Ecrins est une zone de très haute montagne avec plus de 150 sommets dépassant 3000 mètres d'altitude et la Barre des Ecrins comme point culminant à 4102 mètres.

L'équipe permanente du PNE est constituée d'environ 90 personnes réparties entre le siège, à Gap, et les 7 secteurs historiques qui couvrent l'ensemble du territoire du parc (figure 2). Le siège est composé de quatre services : le service scientifique, le service aménagement, le service communication et le service général.

Mon stage se déroule au sein du service scientifique du PNE. Ce service est lui-même divisé en deux pôles :

- Le pôle Connaissance qui travaille sur la mise en place de protocoles de suivi scientifiques (faune, flore, géologie et mesures physiques).
- Le pôle Système d'Informations (SI), dans lequel je me trouve, qui s'occupe de la géomatique et de l'informatique. Il est constitué d'un chargé de mission base de données et développement web, d'un chargé de mission administration réseau, téléphonie et informatique, et d'un géomaticien, chef du pôle.

#### Les missions du SI

Le pôle SI occupe une position transversale puisqu'il est amené à travailler avec tous les services du Parc. Il assiste aussi bien le service scientifique dans la mise en place de protocoles de suivi faune-flore, que le service aménagement dans le suivi du patrimoine bâti et de l'agriculture, ou encore le service communication dans le mise en place d'outils de mise en valeur des sentiers de randonnées et l'animation du site web.

De part ces missions de protection de la faune et de la flore, le Parc national est amené à collecter des quantités importantes de données spatialisées. Le rôle du SI au sein du parc est donc d'organiser et de faciliter la collecte de ces données, de les gérer mais également de créer des outils pour les analyser. Une grande composante métier du SI tient donc dans l'administration de base de données.

## La logique de chaîne de travail au PNE

Le PNE a été novateur dans la mise en place de la collecte et le stockage des données sur informatique et possède aujourd'hui une architecture de base de données et des outils structurés. Dès 2002, le pôle SI a mis en place une stratégie géomatique qui rompt avec la logique de SIG traditionnelle (SIG bureautique et bases de données Access). Le PNE déploie une politique globale, reposant sur des bases de données spatiales, le développement informatique et la migration des données sur un serveur commun. En 2006, le pôle SI se

tourne vers des outils open Source pour gérer, stocker et traiter ses données, tels que QGIS, PostgreSQL et PostGIS.

La figure 3 ci-dessous résume la modernisation de la stratégie générale du SI, du recueil de la donnée jusqu'à son traitement et sa consultation. Les nouveaux modules du système apparaissent en vert, tandis que les anciens sont en blanc.

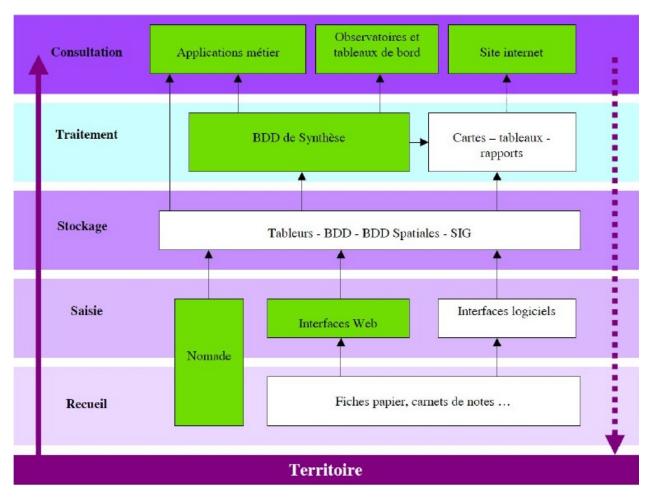


Figure 3 : La chaîne de travail du SI, source PNE (2008)

La collecte sur le terrain, anciennement effectuée sur papier est depuis plusieurs années récoltée avec des outils nomades (applications mobiles sur tablette). Originellement, la saisie des données recueillies au format numérique se faisait à l'aide d'interfaces de logiciels clients lourds (Excel, Access, SIG bureautiques), dorénavant cette tâche s'effectue sur des interfaces web pour les données alphanumériques et WebSIG pour les objets géographiques.

Si les outils de stockage sont encore multiples (tableurs, SIG, BDD...), la tendance est à la centralisation des informations. Par exemple, une BDD de Synthèse a été créée pour agréger toutes les données faune, flore et fonge. De manière générale pour les protocoles importants, l'ensemble des données sont centralisées dans des bases de données PostgreSQL avec extension PostGIS pour les besoins spatiaux. Le stockage d'informations géographiques dans une base de données offre alors un large panel de fonctions de traitements (calculs d'aire, de

buffer, de distance, intersections...) que l'on peut intégrer directement dans les requêtes SQL. Les BDD spatiales occupent donc une place centrale pour la structuration et la diffusion des données du parc.

Connectées à ces BDD, des applications web de consultation métier ou grand public sont développées soit directement par le SI dans la majorité des cas, soit par des prestataires extérieurs.

## Les objectifs du stage

C'est dans cette logique de travail déjà bien établie que je suis arrivée en stage en mars 2021. Son fonctionnement ayant fait ses preuves avec le déploiement de Geotrek<sup>4</sup> en 2011 et de GeoNature<sup>5</sup> en 2014, le pôle SI détenait déjà des compétences solides sur la gestion de projets d'envergure et sur les outils informatiques qui les permettent.

Depuis quelques années, le PNE dispose de capteurs et stations météo permettant de suivre l'évolution des phénomènes physico-chimiques et climatiques qui s'opèrent sur son territoire. Les informations récoltées étaient jusqu'à présent recueillies, saisies et stockées par une même personne, en charge des mesures physiques du Parc, majoritairement sous forme de tableurs. Cette organisation ainsi que l'intérêt que représentent ces données ont fait naître un besoin de structuration et de centralisation des informations. L'idée de création d'une BDD « capteurs » émergeait donc comme prochain projet du pôle SI.

En parallèle, des besoins réguliers en SIG comme la mise à jour des données géographiques de références (fonds IGN), des projets-types QGIS destinés aux agents ainsi que leur formation QGIS devaient être maintenus, tout comme la diffusion des données produites en open data.

C'est en réponse à ces besoins que mon stage s'est articulé autour de trois missions :

- Contribuer à l'actualisation du SIG du PNE en mettant à jour les données géographiques de références (fonds IGN notamment) et en ayant une réflexion sur les formats de fichiers géographiques afin de proposer aux agents un projet-type QGIS et des données à jour. Être en appui à la formation QGIS des agents du Parc.
- Contribuer à la diffusion des données produites par le PNE en mettant à jour les jeux de données publiés sur la plateforme open data, data.gouv.fr, et en les référençant sur le catalogue des Parcs nationaux. Sur ce volet, la recherche d'une solution pour automatiser la publication de données sera effectuée.

<sup>4</sup> Geotrek, dans les coulisses des portails de randonnée, Novembre 2019,

https://www.ecrins-parcnational.fr/actualite/geotrek-coulisses-portails-randonnee

<sup>5</sup> Les geeks de GeoNature carburent, Octobre 2018, <a href="https://www.ecrins-parcnational.fr/actualite/geeks-geonature-carburent">https://www.ecrins-parcnational.fr/actualite/geeks-geonature-carburent</a>

• Concevoir et mettre en place une base de données des capteurs et stations météo du PNE et développer des applications de saisie (WebSIG) pour soulager la chargée de mission mesures physiques et une page web de consultation des capteurs (avec Leaflet), connectés à cette nouvelle base de données.

## STRUCTURATION ET MODERNISATION DU SIG

## Création d'une BDD des capteurs du PNE

### Co-construction du Modèle de données de la BDD

#### ◆ Contexte

De par ses missions et ses travaux d'observation à long terme, le Parc national des Écrins contribue à la veille écologique de nombreux écosystèmes. Au sein de plusieurs réseaux et en partenariat avec des équipes de recherche, il participe à une meilleure compréhension des évolutions en cours liées aux changements climatiques.

Depuis une vingtaine d'années, le Parc dispose sur son territoire des stations météorologiques et des capteurs qui suivent ces évolutions et met en place des programmes pluridisciplinaires sur les lacs d'altitude, les alpages et quelques espèces sentinelles. Les mesures physiques constituent le premier axe de travail, qu'elles soient réalisées par des appareils automatiques ou par les agents du Parc national.

C'est de l'accumulation des données recueillies par les agents du Parc et l'envie de les valoriser qu'est né le besoin de construire une base de données. Afin d'organiser la conception de cette BDD, une première réunion avec les membres du pôle SI et la chargée de mesures physiques a permit de définir les besoins, les premières définitions, les objectifs fonctionnels et visuels finaux (outils), ainsi que d'élaborer un premier modèle de donnée et de définir un calendrier des avancements.

A l'issue de cette réunion, il est convenu que je propose lors de mon stage une première base de données spatiale des capteurs du Parc avec quelques données associées et que je développe des outils de gestion et visualisation web connectés à cette base de données (voir Annexe 1). En parallèle, un autre stagiaire, Baptiste Garcin, explorera le potentiel des capteurs connectés au réseau LoraWan et nous connecterons ses données à la BDD. Ce

travail préliminaire sera prolongé et complété par le développeur et gestionnaire de base de données du Parc, Théo Lechémia, qui pérennisera ce projet en développant des outils de plus grande envergure.

#### Méthode

C'est en collaboration avec Clotilde Sagot, la chargée des mesures physiques du PNE, que j'ai par la suite peaufiné les modèles de données. Dans un premier temps, elle m'a fourni ses jeux de données, la plupart sous format tabulaire et certains en format SIG (shape). Après les avoir analysés et en partant du premier modèle de données schématisé en réunion, j'ai entrepris des modifications afin d'être certaine que celui-ci réponde bien à tous les besoins métiers (voir Annexe 2).

Cette phase a duré plusieurs semaines car il était nécessaire de concilier d'une part les besoins métiers, qui révélaient la complexité du modèle de données, et d'autre part, les réalités fonctionnelles de logique de base de données qui nécessitent une certaine simplification des concepts théoriques. C'est en faisant plusieurs allers-retours entre thématicienne et gestionnaire de base de données que nous avons co-construit le modèle de données de la BDD des capteurs du PNE (voir Annexe 3).

#### Consolidation

Afin de mieux comprendre le fonctionnement des capteurs du Parc, une sortie terrain a été organisée pour récolter les données des capteurs situés dans un alpage du Lautaret. L'équipe du pôle SI et la chargée des mesures physiques du Parc étaient présents ainsi que la gestionnaire de données des stations de recherche du Jardin du Lautaret, Lucie Liger. Cette rencontre était également l'occasion d'échanger au sujet des capteurs connectés au réseau LoraWan, en phase de test au Lautaret et au siège du Parc.



Figure 4 : Relevé d'un capteur Alpages sentinelles

C'est lors de cette rencontre que je me suis rendue compte de l'importance de « standardiser » les éléments de la base de données des capteurs. Après avoir présenté mon modèle de données (Annexe 3) à Lucie Liger ainsi que nos objectifs de mise à disposition des données, celle-ci nous a fait part d'initiatives similaires et notamment des travaux entrepris par l'infrastructure de recherche OZCAR qui œuvre pour un système d'information commun recensant et valorisant les données in situ de grands observatoires internationaux. Le Jardin du Lautaret, situé dans le cœur du PNE, étant un des

observatoires appartenant au projet Ozcar, il est apparu nécessaire de faire concorder nos modèles de données afin que les données issues des capteurs puissent être interopérables.

Suite à cette prise de conscience, j'ai effectué des recherches sur le modèle de données utilisé par Ozcar ainsi que sur les normes et standards préconisés par l'OGC pour construire une base de données de capteurs in situ et connectés. Le modèle de données a encore une fois subi plusieurs modifications car il a fallut trouver un équilibre entre conformité de la BDD aux standards et besoins, moins complexes, du Parc (voir Annexe 4).

Enfin, après avoir créé la base de données sur PostgreSQL, ce dernier modèle de données a été validé avec Baptiste Garcin, en charge des capteurs connectés au réseau LoraWan, puisque nous avons réussi à connecter les données issues de ces capteurs, à la base de données.

Les requêtes SQL qui ont permis de créer la BDD ont été publiées sur un dépôt Github <sup>6</sup> où mon tuteur a pu suivre les modifications apportées. Cette méthode de travail m'a permis de monter en compétence en gestion de projet développement, compétence essentielle pour mon futur métier.

Pour résumer, la base de données des capteurs du Parc repose sur un modèle de données qui prend en compte trois facteurs clés : les besoins métiers, l'aspect fonctionnel et les standards existants. Une fois créée, une phase de test commence par l'insertion des données à disposition.

### Insertion des données dans PostGreSQL

Afin de s'assurer que les informations dont nous disposons trouvent chacune leur place dans la base de données, il est nécessaire d'y implémenter des données et de procéder à des tests pour vérifier la fiabilité des relations entre les tables.

Plusieurs méthodes ont été utilisées pour insérer des données dans la BDD :

#### En passant par QGIS

Cette méthode est utile si on souhaite importer des données spatiales au format SIG dans une BDD, et les visualiser sur le logiciel SIG. En se connectant à une base de données PostgreSQL sur QGIS, il est possible d'ouvrir des tables d'une BDD dans un projet via le gestionnaire de base de données. Afin de pouvoir copier/coller les données d'une couche SIG dans une table de la BDD, il est cependant nécessaire de disposer des mêmes noms de champs. En utilisant cette méthode, je suis partie de plusieurs couches spatiales de localisation de différents capteurs dans le Parc et je les ai agrégées dans une nouvelle

<sup>6</sup> Dépôt github Geosensor <a href="https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql">https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql</a>

couche, unique, comportant les mêmes champs que la table de destination dans la BDD. Ainsi, en ouvrant cette table dans mon projet QGIS et en la passant en mode édition, j'ai pu sélectionner tous les objets de ma couche et les copier/coller dans la table et l'enregistrer. Cette méthode est efficace si l'on n'a pas beaucoup d'entités à importer car un logiciel SIG n'est pas adapté à la gestion de grandes BDD (plusieurs milliers de données).

#### Avec des INSERT INTO

Cette méthode est utile si l'on veut ajouter quelques entités à une table, au coup à coup. Pour ce faire, il faut effectuer une requête SQL dans le gestionnaire de BDD, ici PostgreSQL. C'est le « INSERT INTO » suivi du nom de la table (et du schéma si il existe), puis des noms des champs à compléter ainsi que des données à insérer qui permet l'implémentation de données. J'ai utilisé cette méthode pour ajouter les cinq capteurs LoRa, en phase de test au siège du Parc.

Exemple de requête pour ajouter un capteur dans la table « sensor » du schéma « geosensor » :

```
INSERT INTO geosensor.sensor
(id_sensor,name,sensortype,manufacturer,brand,model,serialnumber,inform
ation,owner)
VALUES
(37,'LairdTH_35FC_Vlb','capteur de temp et
hum','Laird','Laird','Sentrius RS1xx','0025ca0a000035fc','capteur
outdoor de temp et hum sans sonde externe localise a Entraigues','PNE')
```

#### Avec COPY CSV

Cette méthode est utile pour importer une quantité importante de données. Cependant pour pouvoir l'utiliser il faut être super utilisateur de sa base de données PostgreSQL et il faut que le fichier à copier présente les mêmes champs que la table de destination, ainsi que le même nommage. Cette méthode a été efficace pour insérer des informations déjà contenues dans un tableur et présentant beaucoup de texte.

Exemple de requête pour importer des données issues d'un fichier csv dans la table « intervention » :

```
COPY intervention FROM '/.../intervention.csv' DELIMITER ',' CSV HEADER
```

Après avoir inséré les informations sur les capteurs dans la base de données, j'ai obtenu une première validation de mon modèle de données puisque celui-ci permettait bien de

contenir toutes les métadonnées relatives aux capteurs. Les informations géographiques de localisation des capteurs ont quant à elles été ajoutées dans un second temps.

## Calculs spatiaux automatisés avec PostGIS

Afin d'exploiter les potentialités d'une BDD spatiale et d'éviter les erreurs et mises à jour régulières, il est possible de mettre en place des triggers pour calculer automatiquement certains champs. J'ai mis en place des triggers dans la BDD pour récupérer la géométrie des capteurs, à partir des tables du référentiel géographique, qui constitue à lui seul un schéma de la BDD, contenant toutes les limites administratives du PNE ainsi qu'un MNT (Modèle Numérique de Terrain). En pratique, les triggers créés exécutent une intersection entre d'un côté la table des capteurs du schéma « geosensor » et de l'autre, les tables contenant les communes et le MNT du schéma « ref\_geo ». Ainsi, à chaque création ou modification d'un point de la table des capteurs, le trigger s'exécute et les champs « commune » et « altitude » sont calculés automatiquement.

#### Trigger pour récupérer les communes des capteurs

```
CREATE OR REPLACE FUNCTION geosensor.update_commune()
RETURNS trigger
LANGUAGE plpgsql
AS $function$
DECLARE
geom change boolean; --variable qui déterminera si la géométrie de 'station' a changé
BEGIN
geom_change = false;
IF(TG_OP = 'UPDATE') THEN --si une station est mise à jour dans la table 'station' alors,
  SELECT INTO geom_change NOT ST_EQUALS(OLD.geom, NEW.geom); --vérifie si la géométrie
de la station a changé de géométrie
END IF;
IF(TG_OP='INSERT' OR (TG_OP='UPDATE' AND geom_change)) THEN --si il y a création ou
modification d'une station et que sa géométrie a changé
 UPDATE geosensor.station SET commune=ref_geo.l_areas.area_name --met à jour le champ
'commune' des stations en te basant sur le champ 'area_name'
       FROM ref_geo.l_areas --de la table 'l_areas' du schéma ref_geo
       WHERE ref_geo.l_areas.id_type=25 and
ST_Within(geosensor.station.geom,ref_geo.l_areas.geom)--en prenant uniquement en compte
les communes (id_type=25) ET fait ceci lorsqu'une station se trouve dans une commune de la
table 'l areas'
       and geosensor.station.id_station=new.id_station; --ET ne fait ceci que pour la station
ajoutée ou modifiée
END IF;
RETURN NEW;
END;
$function$;
```

CREATE TRIGGER update\_commune AFTER -- met en place le trigger 'update\_commune' APRES

insert or update of geom --la création d'une station ou la modification de sa géométrie on geosensor.station for each row execute procedure geosensor.update\_commune(); -- éxecute ce trigger pour chaque ligne de la table 'station'

• Trigger pour récupérer l'altitude des capteurs

```
CREATE OR REPLACE FUNCTION geosensor.update_alti()
RETURNS trigger
LANGUAGE plpgsql
AS $function$
DECLARE
geom_change boolean; --variable qui déterminera si la géométrie de 'station' a changé
BEGIN
geom_change = false;
IF(TG_OP = 'UPDATE') THEN --si une station est mise à jour dans la table 'station' alors,
  SELECT INTO geom_change NOT ST_EQUALS(OLD.geom, NEW.geom); --vérifie si la
géométrie de la station a changé de géométrie
END IF;
IF(TG OP='INSERT' OR (TG OP='UPDATE' AND geom change)) THEN --si il y a création ou
modification d'une station et que sa géométrie a changé
 UPDATE geosensor.station SET
elevation=ST_Value(ref_geo.dem.rast,1,geosensor.station.geom) --met à jour le champ
'elevation' des stations en récupérant la valeur de la bande 1 du raster de la station
       FROM ref_geo.dem --depuis la table 'dem' qui contient le raster (MNT)
      WHERE st_intersects(geosensor.station.geom,ref_geo.dem.rast) -- au niveau du point
d'intersection entre le MNT et la géométrie de la station
       and geosensor.station.id station=new.id station; --ET ne fait ceci que pour la station
ajoutée ou modifiée
END IF;
RETURN NEW;
END;
$function$
CREATE TRIGGER update alti AFTER -- met en place le trigger 'update alti' APRES
insert or update of geom --la création d'une station ou la modification de sa géométrie
on geosensor.station for each row execute procedure geosensor.update_alti(); --éxecute ce
trigger pour chaque ligne de la table 'station'
```

Les requêtes SQL permettant de créer ces triggers sont également renseignées dans le dépôt Github dédié<sup>7</sup>. L'ensemble des requêtes SQL qui permettent de créer la BDD capteurs constituent le « dump » qui permet à tout usager de pouvoir créer la même base de données. Ce dump a été annoté afin d'expliciter la création des requêtes, et faciliter la réutilisation de ce travail.

Après avoir effectué un premier travail de structuration et de stockage des données dans une BDD spatiale, je me suis tournée vers la mise à jour des outils SIG à disposition des agents du Parc, afin qu'ils puissent bénéficier d'une gestion optimisée de leurs données spatialisées. Le PNE a adopté depuis une dizaine d'années une stratégie de démocratisation des outils SIG, permettant à un nombre croissant d'agents de monter en compétences sur QGIS, entre autres, et de pouvoir être autonomes en production de cartes. Pour rendre cela possible, la cellule géomatique a effectué un travail de réorganisation des données géographiques, en les regroupant dans un dossier commun placé sur un serveur commun à tous les agents du PNE : la « Geobase ». Cette Geobase est structurée de manière thématique (ex : limites administratives, MNT, hydrologie...) et permet de stocker les données de référence. La centralisation ainsi que la mise à disposition de ces données contribue fortement à l'évolution du rôle du géomaticien au PNE, celui-ci n'étant plus celui qui fait les cartes, mais celui qui accompagne et organise la création de cartes. Son rôle s'articule aujourd'hui essentiellement autour de la mise à jour et la mise en forme des données de références, l'optimisation des outils SIG (facilitation technique) et l'accompagnement des agents au travers de formations. Le métier de géomaticien reposant sur des outils informatiques en constante évolution, son rôle est également d'effectuer une veille technologique pour anticiper les futurs besoins de sa structure.

## Mise à jour des données de référence du PNE

Afin d'encadrer au mieux la production de cartes des agents du PNE, le géomaticien est chargé de mettre à jour les données de références contenues dans la Geobase. Pour cela, il s'informe des fonds de cartes et données à actualiser et effectue une veille technologique afin d'optimiser les ressources qu'il met à disposition des agents. Certains formats de fichiers deviennent obsolètes et de nouveaux émergent avec de nouvelles potentialités. Il en est de même pour les extensions et bibliothèques associées à QGIS qui évoluent constamment et offrent de nouvelles opportunités. C'est en se tenant informé de ces évolutions, en les testant et en portant un regard critique sur ces avancées que le géomaticien prend la décision de faire évoluer sa méthode de travail et l'utilisation des outils SIG à disposition des agents.

<sup>7</sup> Dépot Github de la BDD capteurs : <a href="https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql">https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql</a>

## Actualisation des formats vecteurs

#### Du Shapefile au Geopackage

A mon arrivée en stage, une réflexion sur les formats de fichiers vecteurs avait déjà été entamée à l'occasion d'un stage précédent, avec Raphaël Bres<sup>8</sup>. Cette réflexion porte sur le format shapefile, alors exclusivement utilisé par le Parc. Le format shapefile étant un format propriétaire est d'autres formats étant plus performants, une réflexion sur une potentielle conversion des données de la Geobase dans un nouveau format de fichier s'impose.

Le format shapefile est le format vecteur le plus démocratisé de nos jours dans l'utilisation des SIG. Il a pendant très longtemps été la recommandation principale de l'OGC. Cependant, ce format dispose de plusieurs inconvénients comme la création de plusieurs fichiers pour une même couche (SHP, DBF, SHX, PROJ), ce qui implique un volume de stockage conséquent. De plus, une couche shapefile est limitée à 2 Go de stockage. D'autres caractéristiques du shapefile sont également limitantes comme le fait qu'on ne puisse pas combiner plusieurs géométries dans une même couche ou encore la limitation du nom des champs à 10 caractères.

Un autre format de fichier pour le stockage des données vectorielles a été testé par Raphael Bres, le GeoJSON. C'est un format non préconisé par l'OGC mais largement utilisé en développement web avec par exemple les librairies Leaflet et Openlayers. Après avoir testé la performance de ce format de fichier sur des jeux de données de volume différents, il s'est avéré très performant pour des petits jeux de données (> 1Mo) mais moins intéressant lorsque le nombre de données augmente.

Une réflexion sur le format GeoPackage a également été initiée avant mon arrivée en stage et j'ai approfondi à mon tour les avantages et inconvénients de ce format, récemment recommandé par l'OGC pour les traitements SIG. Format de données par défaut dans QGIS3, il est basé sur une base de données SQLite, ce qui permet d'accélérer les requêtes spatiales et le temps d'affichage de la couche, principalement lorsqu'on a de gros volumes de données. C'est un format non propriétaire qui se présente sous la forme d'un fichier unique, pouvant contenir plusieurs types de géométries. A contrario du shapefile, le geopackage peut stocker le style des couches qu'il contient ainsi que le système de projection.

Afin de vérifier la performance du Geopackage, j'ai effectué plusieurs tests de comparaison avec le shapefile. Le volume de stockage ainsi que le temps d'ouverture de fichiers légers et volumineux dans QGIS ont été comparés.

<sup>8</sup> GeoJSON VS Shapefile VS GeoPackage <a href="https://si.ecrins-parcnational.com/blog/2020-02-geojson-shapefile-geopackage.html">https://si.ecrins-parcnational.com/blog/2020-02-geojson-shapefile-geopackage.html</a>

Le fichier léger comprend une cinquantaine de points avec une table attributaire peu conséquente tandis que le fichier volumineux comprend près de 700 000 polygones avec une table attributaire fournie.

#### <u>Tableau comparatif du volume de stockage des fichiers</u>

	SHP	GPKG
Fichier léger	23 Ko	96 Ko
Fichier volumineux	627 Mo	351 Mo

En terme de volume de stockage, il est plus avantageux d'utiliser le format Geopackage si l'on dispose majoritairement de gros fichiers de données, ce qui est le cas du PNE. Par contre le format shapefile est intéressant à utiliser pour stocker des petits jeux de données.

Pour comparer le temps de chargement d'une couche sur QGIS, j'ai créé un script python qui permet de calculer le temps d'ouverture en secondes, que j'ai lancé sur la console python de QGIS.

Exemple de script pour calculer le temps d'ouverture du shape « communes » dans QGIS :

```
from datetime import datetime
start_time = datetime.now()
iface.addVectorLayer("/Users/cendr/Desktop/Stage/QGIS/test
gpkg/bati.shp", "bati", "ogr")
end_time = datetime.now()
print('Duration: {}'.format(end_time - start_time))
```

J'ai réalisé des tests sur deux systèmes d'exploitation différents (Windows et Ubuntu) ainsi qu'en testant le chargement d'un même fichier en local, et depuis le serveur commun. Il est important de noter que les temps présentés peuvent varier d'une machine à l'autre et en fonction du moment de la journée où est lancé le script, les capacités du serveur étant dépendants de la qualité de la connexion internet. Pour affiner cette comparaison, il aurait fallu lancer le script une dizaine de fois pour chaque scénario, et à des horaires différents afin d'abaisser la marge d'erreur du temps d'ouverture.

#### Tableaux comparatifs des temps d'ouverture (en secondes) des couches dans QGIS

Taille du fichier Local/Serveur Format			Fichier	· léger	
		Local		Serveur	
		SHP	GPKG	SHP	GPKG
Ubuntu	Ouverture	0,01	0,01	0,17	0,13
Windows	Ouverture	0,2	0,2	0,34	0,31

Taille d	u fichier		Fichier vo	lumineux	
Local/Serveur Extension		Local		Serveur	
		SHP	GPKG	SHP	GPKG
Ubuntu	Ouverture	0,9	0,04	3,7	0,69
Windows	Ouverture	1,34	0,8	7,1	5,12

Travailler avec le format Geopackage sur un logiciel SIG offre un gain de temps à l'ouverture des fichiers, comparé au shapefile. Celui-ci est d'autant plus intéressant lorsqu'il est utilisé en local, sous Ubuntu. C'est d'ailleurs cette configuration qui est utilisée par le géomaticien au PNE. C'est donc le format Geopackage qui offre de meilleures performances pour le PNE.

Une solution alternative pour remplacer l'ensemble des couches vectorielles du PNE , celle de recenser toutes les données de référence dans une BDD spatiale, n'est pas envisageable compte tenu de l'éclatement des huit sites de travail du PNE et de leur débit internet faible et irrégulier.

La décision a donc été prise avec le géomaticien du Parc de convertir les données de la Geobase en Geopackage. Pour cela, il est apparu nécessaire de trouver une manière de convertir toutes les données en lot, sans passer par l'interface graphique de QGIS.

#### Utilisation d'OGR

Afin de trouver une solution pour convertir en masse tous les fichiers shape de la Geobase en geopackage, mon tuteur m'a préconisé de me renseigner auprès de l'administratrice des données du Parc national des Cévennes, Amandine Sahl, qui avait déjà entrepris la manœuvre. Celle-ci m'a dirigé vers l'utilisation de GDAL/OGR.

OGR est un sous-ensemble de la bibliothèque libre GDAL (Geospatial Data Abstraction Library), qui permet de lire et de traiter un très grand nombre de format d'images géographiques (raster). OGR permet de faire de même avec des formats de données vectorielles. Cette bibliothèque est d'ailleurs intégrée à QGIS sous forme de géotraitements dans la boîte à outils, mais fonctionne également en ligne de commande. Cette possibilité permet alors d'effectuer des traitements de données lourds avec des fichiers de grande taille, puisqu'on n'a pas besoin d'afficher les données sur un logiciel SIG.

Un des outils d'OGR, ogr2ogr<sup>9</sup>, permet de convertir des fichiers de données vectorielles. C'est cet outil que j'ai utilisé pour convertir, en ligne de commande, toutes les données de la Geobase en geopackage, en quelques minutes.

<sup>9</sup> Documentation ogr2ogr <a href="https://gdal.org/programs/ogr2ogr.html">https://gdal.org/programs/ogr2ogr.html</a>

Pour cela, il est nécessaire d'installer la bibliothèque GDAL/OGR sur son ordinateur. A savoir que celle-ci est incluse dans la distribution OSGeo4W, souvent utilisée pour obtenir QGIS sous Windows. Travaillant sur un poste Linux, voici comment j'ai procédé pour l'installation en ligne de commande <sup>10</sup>:

```
$ sudo apt-get install python3.6-dev
$ sudo add-apt-repository -y ppa:ubuntugis/ppa
$ sudo apt-get update
$ sudo apt-get install gdal-bin
$ sudo apt-get install libgdal-dev
$ pip3 install GDAL
```

Une fois GDAL installé, j'ai effectué une copie de la Geobase en local sur mon ordinateur, puis j'ai repris les fichiers ressources placés sur le dépôt Github des Parc nationaux permettant de convertir en masse des shape en geopackage<sup>11</sup>.

Les données de référence contenues dans la Geobase existant maintenant en geopackage, la décision de basculer cette nouvelle version sur le serveur commun a été laissée au géomaticien du Parc, celui-ci devant communiquer à tous les agents du PNE ce changement de format de données.

En effet, ce changement implique un usage différent de QGIS, notamment au niveau des projets-types à destination des agents, basés sur la Geobase en shapefile. Les projet-types ont été créés par le géomaticien du Parc afin de faciliter la production de carte des agents. Il contient toutes les couches de référence (limites administratives et du Parc, fonds de carte...) centrées sur le territoire du PNE, et des paramètres pré-enregistrés (projection, demande de SCR à l'import de couche...).

Afin d'accompagner le renouvellement de la Geobase, j'ai créé un nouveau projet-type QGIS, à partir des fichiers de données en geopackage (voir p.27).

#### Récupération de la BD TOPO en GPKG

Le format geopackage présentant de nombreux avantages, l'IGN a proposé en téléchargement libre sur Geoservices, en mai 2021, une nouvelle version de la BD TOPO, dans ce format <sup>12</sup>. Ainsi, toutes les données de la BD TOPO sont contenues dans un seul et même fichier, les couches sont organisées en sous-groupes et tous les styles sont

<sup>10</sup> Installation ogr2ogr <a href="https://mothergeo-py.readthedocs.io/en/latest/development/how-to/gdal-ubuntu-pkg.html">https://mothergeo-py.readthedocs.io/en/latest/development/how-to/gdal-ubuntu-pkg.html</a>

<sup>11</sup> Ressources techniques ogr2ogr\_shp2gpkg

https://github.com/PnX-SI/Ressources-techniques/tree/master/Python/Shp2GPKG/ogr2ogr\_shp2gpkg 12 BD TOPO IGN en Geopackage <a href="https://geoservices.ign.fr/la-bd-topor-au-format-geopackage">https://geoservices.ign.fr/la-bd-topor-au-format-geopackage</a>

enregistrés avec un affichage dépendant de l'échelle. Cette version graphique de la BD TOPO a permis d'enrichir le projet-type du PNE.

### Actualisation des formats rasters

#### Téléchargement des fonds de carte IGN

Afin de proposer un nouveau projet-type aux agents du PNE avec les données de référence à jour, j'ai également actualisé les fonds de carte IGN du PNE. L'Institut national de l'information géographique et forestière (IGN) a pour mission d'assurer la production, l'entretien et la diffusion de l'information géographique de référence en France.

Depuis le 1er janvier 2021, l'IGN rend libres et gratuitement accessibles ses données publiques. L'ouverture des données IGN sous licence ouverte Etalab 2.0, implique un accès et un usage gratuit pour tous. Il n'y a donc plus de droits de reproduction ni de diffusion à régler, sauf concernant les produits Scans (ex : Scan25) et cartes papiers IGN qui restent payant-e-s.

Cette décision s'inscrit dans une logique d'ouverture des données publiques, de modernisation de l'action de l'Etat et de révolution des usages numériques. Dans son rapport<sup>13</sup> sur les données géographiques souveraines diffusée en 2018, la députée Valéria Faure-Muntian recommandait de « diffuser à terme l'ensemble des données géographiques souveraines sous licence ouverte type Etalab, et n'autoriser l'IGN à recourir à la licence payante qu'à titre transitoire et au plus tard jusqu'en 2022 ».

C'est dans ce cadre que l'IGN a entrepris une refonte de sa plateforme de diffusion, Geoservices<sup>14</sup>, pour simplifier l'accès à son offre de données et aux services web destinés aux professionnels. Fonctionnelle depuis juillet 2021, elle était encore en développement lors de mon stage ce qui m'a confronté à des difficultés pour télécharger les fonds de carte.

En effet, les liens de téléchargement des données fonctionnaient de manière aléatoire et les données n'étaient pas toutes disponibles encore. De plus, il était difficile d'avoir une vision d'ensemble des données disponibles par territoire, puisque les données étaient présentées sous forme de longue liste, classées par type fond de carte puis par département, ce qui rendait la recherche fastidieuse, notamment en travaillant sur un territoire à cheval entre plusieurs départements (Isère et Hautes-Alpes).

<sup>13</sup> Les données géographiques souveraines – Rapport <a href="https://www.vie-publique.fr/rapport/37550-les-donnees-geographiques-souveraines-rapport-au-gouvernement">https://www.vie-publique.fr/rapport/37550-les-donnees-geographiques-souveraines-rapport-au-gouvernement</a>

<sup>14</sup> Geoservices IGN <a href="https://geoservices.ign.fr/">https://geoservices.ign.fr/</a>

Pour parer à ces difficultés, j'ai trouvé deux alternatives me permettant de télécharger les fonds de carte IGN: OpendatArchives et IGN2Map. OpendatArchives<sup>15</sup> est un projet initié par Christian Quest en 2019, fervent défenseur de l'open data et de l'open source en France. Il permet un accès libre aux archives des données publiques publiées en opendata. Ainsi, j'ai pu récupérer des liens de téléchargement fonctionnels pour les fonds de carte IGN. De son côté, l'équipe de Geotribu, site collaboratif sur la géomatique libre et ouverte, propose en février 2021 IGN2map<sup>16</sup>, une solution cartographique alternative pour accéder aux données libres publiées par l'IGN. Leurs créateurs ont récupéré les liens de téléchargement se trouvant sur le site de l'IGN en utilisant la méthode du web-scraping, afin de les répertorier sur une carte dynamique Leaflet. C'est d'ailleurs cette première offre visuelle des données IGN qui a été reprise par le Geoservice de téléchargement de l'IGN.

#### Stockage des données

Après avoir résolu le problème de téléchargement des données en ligne, j'ai été confrontée au problème du stockage des données. En effet, le PNE souhaitait récupérer la nouvelle BD ORTHO Haute Résolution, une collection de mosaïques (tuiles) d'orthophotographies numériques en couleur (images aériennes), d'une précision pixellaire de 20 cm. Afin de couvrir le territoire du PNE, il me fallait stocker toutes les tuiles des départements de l'Isère et des Hautes-Alpes, ce qui représentait un volume de stockage de plus de 60 Go. Afin de ne pas surcharger les ordinateurs et le serveur du PNE, il a été décidé que tous les fonds de carte seraient stockés sur un disque dur externe, puis une fois découpés à l'échelle du territoire et leur volume optimisé, placés sur le serveur commun.

## Mise en forme des données et facilitation technique des agents du PNE

## Fusion et découpage des fonds de carte

Afin d'optimiser le volume de stockage des fonds de carte IGN, ces derniers ont été découpés en fonction du territoire du PNE, et les tuiles ont été fusionnées afin de n'obtenir qu'un seul fichier en sortie. A cette étape, j'ai été confrontée aux limites de l'utilisation de QGIS pour effectuer des géo-traitements avec des fichiers rasters lourds.

A plusieurs reprises, lorsque je souhaitais fusionner des tuiles de la BD ORTHO par exemple, ma fenêtre QGIS plantait et je n'arrivais pas à bout du traitement. Pour résoudre

<sup>15</sup> OpendatArchives <a href="https://www.opendatarchives.fr/">https://www.opendatarchives.fr/</a>

<sup>16</sup> IGN2Map https://geotribu.github.io/ign-fr-opendata-download-ui/index.html

ce problème, j'ai utilisé plusieurs alternatives : IGNMap, GDAL, le fichier VRT et le Geopackage.

#### • Découverte d'IGNMap

IGNMap est un logiciel gratuit développé par l'IGN qui permet de lire et d'afficher leurs données géographiques. IGNMap n'est pas un système d'information géographique, c'est plutôt un visualiseur de données muni de quelques fonctions spécifiques, comme la reprojection de données, la fusion des tuiles et le découpage des fonds raster. IGNMap permet également de consulter des portails web géographiques comme le GéoPortail et Google Maps.

Ainsi, pour tous les fonds de carte IGN téléchargés (BD ORTHO HR, Scan25 et Plan IGN), je suis passée par IGNMap. Malgré la taille importante des fichiers, IGNMap a été performant et m'a permis de récupérer un fichier unique en sortie. Il présente également l'avantage de pouvoir définir le format et la résolution de l'image en sortie (en pixels), et le taux de compression. Ces fonctionnalités ont été utiles pour optimiser le volume de stockage de la BD ORTHO HR qui était trop lourde en résolution pixellaire de 20 cm, et que j'ai dû dégrader pour arriver à une résolution de 50 cm.

Le géomaticien se doit d'optimiser l'espace de stockage des données spatiales, notamment celles des images raster. Ce sont elles qui occupent le volume de stockage le plus important. Il est alors nécessaire de s'adapter au possibilités du parc informatique de sa structure.

#### Utilisation de GDAL

Une autre solution alternative à QGIS pour effectuer des geo-traitements avec des fichiers de données volumineux est d'utiliser la bibliothèque GDAL<sup>17</sup>. Faire appel aux fonctionnalités de GDAL en ligne de commande réduit considérablement le temps de traitement des images rasters puisqu'elles n'ont pas besoin d'être affichées. Les possibilités de re-traitement des images raster avec GDAL sont multiples, mais nécessitent beaucoup de temps de documentation et d'essais. Lors de mon stage j'ai passé un certain temps à tester les fonctionnalités de GDAL pour parvenir à réduire au maximum le volume de stockage des fonds de carte, tout en perdant au minimum en qualité d'image. S'est alors posé la question du format de fichier raster qui permettrait de répondre au mieux à ce besoin.

Le format des tuiles raster IGN sont en JPEG2000 (JP2), format déjà légèrement compressé permettant une taille de fichier assez réduite et performante à l'affichage. Un autre format raster, le GeoTIFF (TIF), offre quant à lui une résolution d'image sans perte de qualité mais

<sup>17</sup> Documentation GDAL https://gdal.gloobe.org/gdal/index.html

entraîne un stockage plus conséquent. C'est avec ce dernier format que fonctionne GDAL par défaut.

Après plusieurs essais de conversion des tuiles rasters en GeoTIFF, avec différents niveaux de dégradation de la qualité d'image, les résultats ne sont pas concluants. La décision est prise de rester sur le format JPEG2000. Hors, GDAL ne supporte pas par défaut ce format de données, il faut installer un « driver » (librairie) adéquat. Le support du JPEG2000 par GDAL est alors obtenu après installation de la librairie Jasper (libjasper)<sup>18</sup>:

\$ sudo apt-get install libjasper-dev

#### Le fichier VRT

Pour faciliter l'affichage des tuiles de la BD Ortho et du Scan 25 de l'IGN, il est également possible de créer un raster virtuel (VRT), qui va réaliser un assemblage de tuiles et se comporter comme une couche unique. Ce fichier est en fait un fichier XML (Extensible Markup Language) qui stocke les caractéristiques des tuiles et les chemins d'accès vers ces tuiles. C'est un fichier qui ne comporte que du texte, donc très léger, mais qui nécessite de disposer d'un dossier avec toutes les tuiles à afficher, et donc d'effectuer un tri. L'avantage est que son affichage dans un logiciel SIG est rapide. Un fichier VRT est d'ailleurs maintenant fourni par l'IGN lors du téléchargement des données raster.

La fonction GDAL, gdalbuidvrt, permet de créer facilement un fichier VRT :

```
$ gdalbuildvrt output.vrt [chemin d'accès des tuiles]/*.jp2 - addalpha -a_srs EPSG:2154
```

Il est également possible de passer par la boîte à outils de QGIS pour créer un VRT, mais le traitement est plus long. Malgré sa facilité de création, je n'ai pas réussi à afficher le VRT de la BD Ortho dans QGIS, alors que cela a fonctionné pour le Scan 25. N'ayant pas réussi à identifier la cause de problème, c'est un point que j'aimerai résoudre dans le futur.

#### Le Geopackage

Le Plan IGN est fourni en plusieurs dalles pour 11 niveaux de zoom. Afin de reconstituer un fichier unique, plus pratique à intégrer au projet-type QGIS, j'ai entrepris de fusionner les dalles pour chaque niveau de zoom avec IGNMap puis de les regrouper dans un geopackage. Les différents niveaux de zoom ont ensuite été prédéfinis dans le style du geopackage. Après avoir affiché le geopackage dans QGIS, je trouve finalement que cette méthode n'est pas adaptée pour ce fond de carte et que le passage d'un niveau de zoom à un autre n'est pas fluide.

<sup>18</sup> Installation de libjasper <a href="https://stackoverflow.com/questions/44468081/unable-to-locate-package-libjasper-dev">https://stackoverflow.com/questions/44468081/unable-to-locate-package-libjasper-dev</a>

Après avoir testé plusieurs méthodes pour mettre en forme les fonds de carte IGN, j'ai obtenu des résultats plus ou moins satisfaisants. Le logiciel QGIS n'est pas adapté au traitement de fichiers raster lourds, ou alors il est nécessaire de disposer d'une bonne machine. La meilleure méthode selon moi est de passer par IGNMap pour fusionner et découper les tuiles de la BD Ortho. On obtient ainsi un fichier unique en format JP2, qui pèse néanmoins 8 Go pour couvrir l'intégralité du territoire du PNE. De ce fait, l'affichage dans QGIS prend quelques secondes mais il est possible de pallier à ce problème en définissant un niveau de zoom maximum pour l'affichage de la couche. Concernant le Scan25, étant donné qu'il y a moins de tuiles à stocker car la résolution est moins fine, l'option du fichier VRT, en passant par GDAL, est la plus satisfaisante. Enfin, pour le Plan IGN, la solution de passer par le flux WMS de l'IGN est la plus acceptable, même si cela implique de disposer d'un certain débit internet, souvent faible dans les antennes des secteurs du Parc.

## Mise à jour des projets-type QGIS

Une fois les données de la Geobase converties en geopackage et les fonds de cartes IGN actualisés et mis en forme, je dispose de tous les éléments pour mettre à jour le projet-type QGIS à destination des agents du Parc. A l'ouverture, le projet est centré sur le territoire du Parc et le système de projection du projet est prédéfini. Des mises en pages pré-remplies au format A4 et A3, orientées paysage et portrait sont également enregistrées.

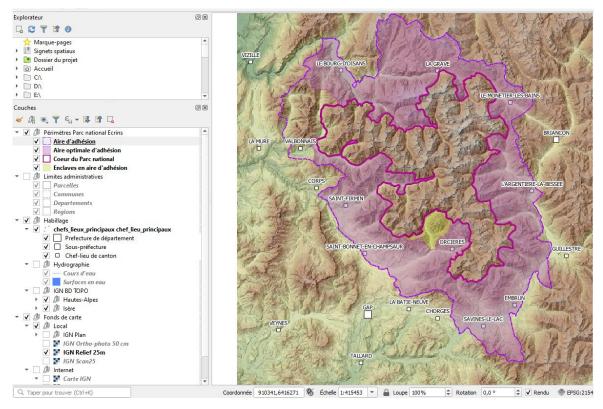


Figure 5: Projet-type QGIS du PNE

J'ai créé deux versions similaires de ce projet-type QGIS, une pour les ordinateurs du Parc disposant d'un système d'exploitation Linux, et une autre pour les ordinateurs sous Windows. Lorsque je suis arrivée en stage, il n'y avait qu'un projet-type QGIS qui fonctionnait sous Windows. Le bureau du pôle SI étant pourvu de machines Linux pour des besoins de développement et de gestion des serveurs, j'ai été confrontée à un projet-type où aucune couche ne s'affichait. Pourtant, toutes les couches sont stockées sur le serveur commun. Pour cause, les chemins d'accès qui diffèrent entre les deux systèmes d'exploitation. En effet, sous Linux, les noms des répertoires et éventuels fichiers sont séparés par un slash '/ ', alors que sous Windows, c'est un antislash '\'. Ainsi, les deux projets-type sont maintenant sur le serveur commun, à disposition des agents.

## Création d'un profil utilisateur QGIS

C'est Amandine Sahl, du Parc National des Cévennes, qui m'a introduit à la mise en place de profils utilisateurs sur QGIS. Par défaut, une installation QGIS ne contient qu'un seul profil utilisateur appelé « par défaut », mais il est possible d'en créer autant qu'on le souhaite en accédant au menu  $Préférences \rightarrow Profils utilisateurs \rightarrow Nouveau profil.$  Le dossier racine du profil peut être ouvert depuis QGIS en utilisant Ouvrir le dossier du profil actif, par exemple sous Windows il se trouve ici :  $AppData \land Roaming \land QGIS \land QGIS \land Profiles$ .

Chaque profil utilisateur contient des paramètres, des plugins et un historique qui lui sont propres. Un profil utilisateur permet d'enregistrer dans un même dossier les configurations suivantes :

- Tous les paramètres définis dans les *Options (Préférences)* : les paramètres généraux permettant de définir le projet à l'ouverture, le système de projection, les échelles prédéfinies...
- Les configurations et personnalisations de l'interface : raccourcis claviers, les menus, panneaux et barres d'outils
- Les plugins installés et leur configuration
- Les réglages des traitements, scripts et modèles

En créant un profil utilisateur il est alors possible de configurer une interface QGIS en fonction des besoins des agents, afin de rendre le logiciel plus ergonomique et encadrer son utilisation. J'ai utilisé cette fonctionnalité de QGIS pour créer un profil pour la chargée des mesures physiques du Parc, avec qui nous avons mis en place la BDD Capteurs.

J'ai personnalisé l'interface en ne laissant que les boutons nécessaires à l'édition des couches dans la barre d'outils. Lorsque l'on choisit de basculer sur ce profil, le projet QGIS adéquat s'ouvre. La connexion à la BDD Capteurs est enregistrée et les couches du projet sont issues de la BDD. Ainsi, toute modification est directement effectuée dans la BDD PostGIS. J'ai créé des formulaires d'attributs avec des listes déroulantes pour faciliter l'édition de ces couches et éviter les erreurs de frappe. Afin d'accompagner au mieux la saisie des données, j'ai créé un tutoriel avec des captures d'écran qui guide l'utilisateur pour ajouter, modifier ou supprimer une entité.

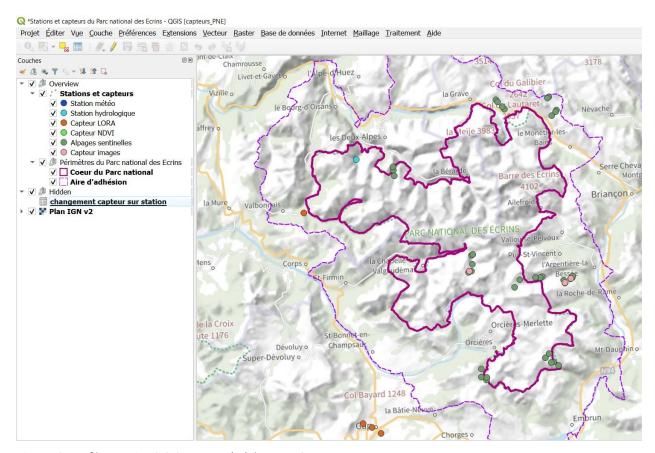


Figure 6: Profil et projet QGIS connectés à la BDD Capteurs

## Montée en compétences des agents en gestion de l'information géographique

Mettre en place des outils pour assurer l'autonomie des agents fait partie du rôle de géomaticien au PNE. En outre, il doit également veiller aux bonnes pratiques des utilisateurs quant à la gestion des outils et données spatiales. Depuis une quinzaine d'années, la cellule géomatique a adopté une stratégie d'accès différenciés à l'information géographique. Les utilisateurs sont positionnés selon trois niveaux représentés dans la figure 7 ci-dessous :

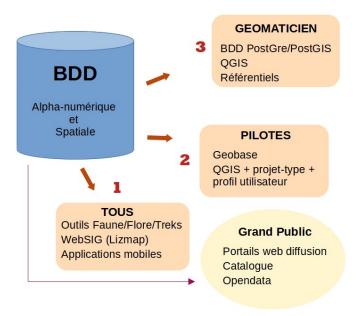


Figure 7: Les trois niveaux d'accès à l'information géographique

#### 1: Un niveau pour tous

Les utilisateurs disposent d'outils "clique-bouton" basés sur des interfaces intuitives, développés par le pôle SI. Deux outils principaux ont été déployés pour la saisie des données Faune, Flore et Trek : Geonature et Geotrek. D'autres interfaces graphiques simples ont été conçues pour saisir des données liées à d'autres protocoles, comme les WebSIG Geoarcheo (inventaire des sites archéologiques) et Geoconsat (inventaire des constats d'attaques de loups), conçus avec Lizmap. Ces outils opens source sont directement disponibles depuis l'intranet et l'extranet du PNE. Des applications mobiles ont également été développées pour permettre aux agents de saisir les observations directement sur le terrain et d'alimenter les différentes bases de données (ex : Recherche Flore, Mortalité, Invertébré...).

A ce niveau, les utilisateurs peuvent essentiellement consulter et saisir des données.

#### 2 : Un niveau pour les pilotes

Les pilotes sont des chargés de mission et gardes-moniteurs en charge d'une thématique spécifique. Ceux-ci ont des besoins auxquels les outils de niveau 1 ne

peuvent pas répondre. Pour cela, le géomaticien organise des formations QGIS, collectives ou individuelles, pour répondre à leur besoins métier. La Geobase, qui contient toutes les données géographiques de référence, leur permet de concevoir des projets QGIS, à partir du projet-type, et de manipuler leurs données en autonomie. De plus, le géomaticien met à disposition des tutoriels et guides d'utilisation pour QGIS afin d'assurer une bonne pratique du logiciel. Il peut également créer un profil utilisateur QGIS pour répondre aux besoins spécifiques d'un agent.

A ce niveau, les utilisateurs peuvent effectuer des traitements simples des données.

#### 3 : Un niveau géomaticien

Ce niveau est celui du chargé de mission géomatique qui, avec le soutien du gestionnaire de base de données et développeur du Parc, administre les bases de données spatiales, les données de référence ainsi que les outils SIG. Le géomaticien accompagne et encadre la gestion de ces outils via des formations et manuels d'utilisation. Il est en charge de cataloguer et publier les données géographiques produites en open data et de les valoriser sur les portails web de diffusion du Parc. A ce niveau, le géomaticien est garant des bons usages pour la collecte, le traitement, l'analyse et la diffusion des données géographiques.

Un autre niveau est également identifié, celui du grand public. Celui-ci peut consulter les données du Parc sur le Catalogue et les portails web de diffusion (ex : Bouquetins Ecrins<sup>19</sup>, Rando Ecrins<sup>20</sup>, Biodiv'Ecrins<sup>21</sup>) et les télécharger sur les plateformes open data (ex : data.gouv).

A tous les niveaux, les utilisateurs qui produisent de la donnée géographique nourrissent les BDD spatiales via les outils de saisie développés. Les portails web de diffusion interrogent ces BDD et les données publiées sur le catalogue et en open data en sont issues. Ainsi, l'information géographique est centralisée.

Après avoir mis à jour et structuré l'information géographique, le pôle SI a misé sur des outils web simples pour les gérer et les consulter. Dans un premier temps l'objectif est de pouvoir rendre accessible la gestion des données aux thématiciens à partir, par exemple, d'un Web SIG permettant l'édition des données via une interface et des formulaires simplifiés. Dans un second temps, le pôle SI développe des applications web de diffusion et de valorisation de ses données produites, comme par exemple un Atlas de biodiversité (Biodiv'Ecrins) ou un Portail de randonnées (Rando Ecrins), avec comme point central une carte interactive. Tous les outils web utilisés font le pont entre d'un côté la BDD et de l'autre, les agents du PNE et le grand public.

<sup>19</sup> Bouquetins Ecrins <a href="http://bouquetins.ecrins-parcnational.fr/">http://bouquetins.ecrins-parcnational.fr/</a>

<sup>20</sup> Biodiv'Ecrins <a href="https://biodiversite.ecrins-parcnational.fr/">https://biodiversite.ecrins-parcnational.fr/</a>

<sup>21</sup> Rando Ecrins <a href="https://rando.ecrins-parcnational.fr/">https://rando.ecrins-parcnational.fr/</a>

## DÉPLOIEMENT D'OUTILS DE GESTION ET DE CONSULTATION DES DONNÉES

## Création d'un Web SIG avec Lizmap

Afin de faciliter la consultation et la saisie des données liées aux capteurs du PNE, j'ai déployé une application web cartographique avec Lizmap, depuis QGIS.

Lizmap est une application WebSIG open source qui fonctionne avec un plugin QGIS, pour publier des cartes en ligne, à partir d'un projet QGIS. La préparation des données, la configuration et la publication des projets/cartes sont les trois étapes pour réaliser un projet Lizmap. La figure ci-dessous présente l'architecture de Lizmap :

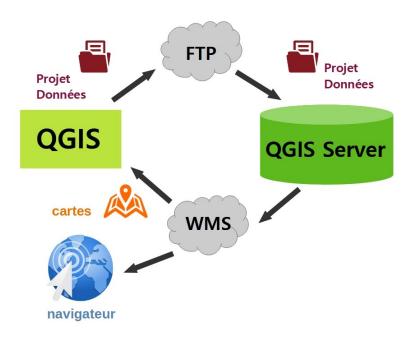


Figure 8: Architecture de Lizmap, source 3Liz

Pour fonctionner, il est nécessaire d'installer au préalable le client web LizMap sur un serveur. Puis on procède par étapes :

- 1. **Préparer le projet QGIS** en important les couches, en définissant leurs styles et en créant des formulaires d'attributs, puis en paramétrant QGIS Server (dans QGIS).
- 2. **Configurer l'affichage du projet avec le plugin** pour modifier les options générales de la carte et celles liées aux couches.
- 3. Publier son projet en synchronisant les données du projet entre le poste de travail et le serveur (protocole FTP). Lorsque la synchronisation est terminée, le projet/carte QGIS peut être consulté simultanément sur Internet. Il est disponible dans l'application Lizmap Web Client avec un navigateur Web.

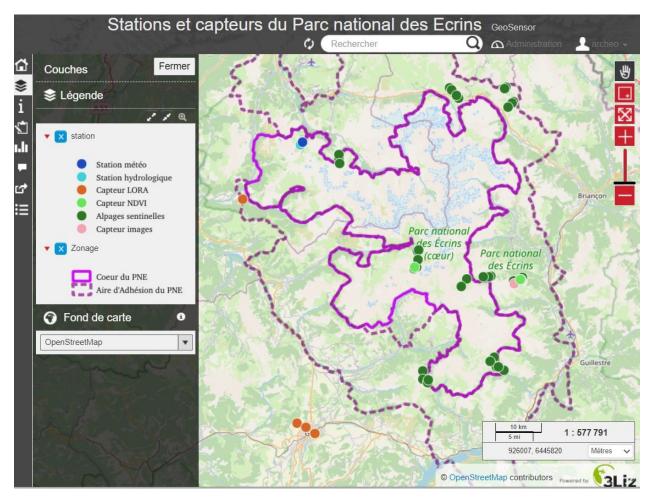


Figure 9: WebSIG des capteurs du PNE réalisé avec Lizmap

Pour construire ce WebSIG, j'ai utilisé le projet QGIS connecté à la BDD Capteurs (voir figure 5), que j'avais mis en place pour la chargée de mission mesures physiques. Elle bénéficie à présent d'une interface web plus ergonomique et simple pour consulter et saisir ses données.

Le pôle SI a fait le choix de n'employer ce type de solution WebSIG que dans des cas d'utilisation restreints. Côté client, le rendu ergonomique et le fonctionnement similaire à QGIS rend pratique son utilisation pour des utilisateurs initiés, et pour un nombre de données limitées. Cependant, côté configuration de l'application, cette solution n'offre pas une grande possibilité de personnalisation et il paraît complexe de configurer un projet QGIS avec beaucoup de couches.

D'autres solutions de cartographie web, comme Leaflet, offrent de meilleures potentialités en terme d'affichage et de fonctionnalités.

## Publication d'une carte dans une page web avec Leaflet

Afin de valoriser les données produites par le Parc, le pôle SI a misé sur l'utilisation de Leaflet, une bibliothèque JavaScript libre de cartographie en ligne. Utiliser des outils libres permet de bénéficier de communautés très actives qui garantissent la pérennité des projets au niveau de leur maintenance et de leur futures évolutions. Le choix du pôle SI d'opter pour des outils open source s'inscrit dans une démarche à la fois éthique et stratégique. Éthique parce que cela permet de contribuer à un bien public et de faire profiter aux autres utilisateurs des améliorations et développements, et stratégique parce que l'on peut s'approprier le code source, réutiliser des applications pour en développer d'autres et ainsi ne pas s'enfermer dans une solution propriétaire, dépendant d'une seule société et ne répondant pas toujours aux normes d'interopérabilité.

L'avantage de la bibliothèque Leaflet est qu'elle peut être étendue avec de nombreux plugins ce qui permet une personnalisation à la carte. De plus, elle est supportée par la plupart des navigateurs web bureautiques et mobiles. Cependant, elle nécessite des compétences en développement et un certain temps de mise en œuvre.

La carte Leaflet ci-dessous (p.36) a été réalisée avec les données de référence du PNE pour les limites du Parc et les données de la BDD Capteurs. Le code HTML source est disponible sur le dépôt github dédié<sup>22</sup>. A partir d'un éditeur de texte (ex : Notepad++), voici comment j'ai procédé :

1. Inclure les fichiers CSS (fichiers de style) de Leaflet dans la section head du document

2. Inclure le fichier JavaScript contenant Leaflet

```
<script src="https://unpkg.com/leaflet@1.7.1/dist/leaflet.js"
integrity="sha512-
XQoYMqMTK8LvdxXYG3nZ448h0EQiglfqkJs1N0QV44cWnUrBc8PkA0cXy20w0vlaXaVUear
IOBhiXZ5V3ynxwA=="
    crossorigin=""></script>
```

<sup>22</sup> Dépôt github dédié au code html de la carte Leaflet : <a href="https://github.com/PnEcrins/GeoSensor/blob/main/leaflet/geosensor.html">https://github.com/PnEcrins/GeoSensor/blob/main/leaflet/geosensor.html</a>

3. Inclure la librairie Jquery pour ouvrir des fichiers GeoJSON

```
<script
    src="https://code.jquery.com/jquery-3.2.1.js"
    integrity="sha256-DZAnKJ/6XZ9si04Hgrsxu/8s717jcIzLy3oi35EouyE="
    crossorigin="anonymous">
    </script>
    </head>
```

4. Dans la section body, placer un élément div (un bloc) avec un certain id où l'on souhaite placer la carte. Définir une hauteur et une largeur pour la carte.

5. Introduire la carte et définir son zoom et son centrage

```
var map = L.map('map').setView([44.75, 6.25], 9.5);
```

6. Importer un fond de carte (l.tilelayer)

```
var osmLayer = L.tileLayer('http://{s}.tile.osm.org/{z}/{x}/{y}.png',
{
         attribution: '@ OpenStreetMap contributors',
     });
```

Note: Il existe depuis peu une extension Géoportail Leaflet qui permet d'importer les fonds de carte IGN<sup>23</sup>.

7. Déclarer la carte

```
map.addLayer(osmLayer);
```

8. Charger les données sur la carte à partir d'un fichier GeoJSON :

```
$.getJSON('data/stations.geojson', function(geojson) {
    var stations = L.geoJSON(geojson,
    {
        style: function (feature) {
            return feature.properties && feature.properties.style;
        },
        onEachFeature: function(feature, layer) {
```

<sup>23</sup> Extension Geoportail Leaflet <a href="https://geoservices.ign.fr/documentation/services/utilisation-web/extension-pour-leaflet">https://geoservices.ign.fr/documentation/services/utilisation-web/extension-pour-leaflet</a>

```
if(feature.properties && feature.properties.name) {
             layer.bindPopup(feature.properties.name);
Infobulles
       pointToLayer: function (feature, latlng) {
             return L.circleMarker(latlng, {
                   radius: 6,
                   fillColor: "#ff7800",
Marqueurs
                   color: "#000",
circulaires
                   weight: 1,
                   opacity: 1,
                   fillOpacity: 0.8
             }); }});
         stations.addTo(map); });
  </script>
```

## Stations et capteurs du Parc National des Ecrins

Cendrine HOARAU - Parc national des Ecrins / Juillet 2021

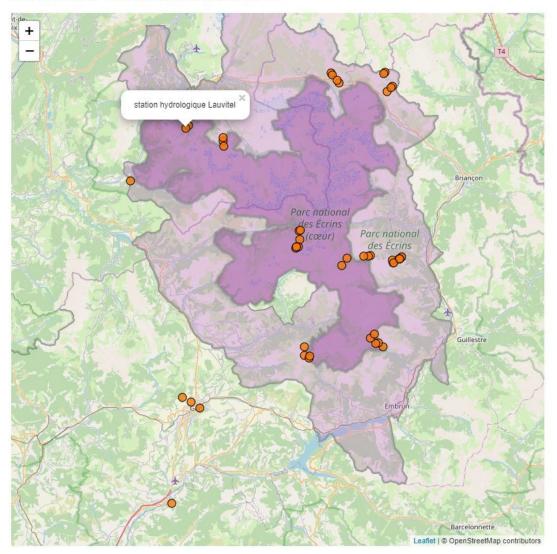


Figure 10: Carte Leaflet connectée à la BDD Capteurs

Sur cette carte interactive Leaflet, les données sont issues de la BDD Capteurs et sont mises à jour automatiquement (voir p.49). L'idée est de développer par la suite cette première carte afin de pouvoir l'insérer dans une page web du PNE, dédiée à la présentation des stations et capteurs du parc, avec la possibilité de visualiser en plus, les données des capteurs sous forme de graphiques.

Au final, le choix d'une solution de webmapping doit être adaptée à l'objectif : il ne sert à rien de développer quelque chose de très lourd pour une carte simple, et, à l'inverse, une lourde phase de programmation en amont peut alléger bien des opérations en aval. Ainsi le WebSIG est destiné à des besoins métiers peu complexes, nécessitant peu de données géographiques à traiter. Tandis que le développement d'une carte interactive avec la bibliothèque Leaflet offre de fortes potentialités cartographiques permettant de configurer l'affichage et de requêter un grand nombre de données. Disposant des compétences d'un développeur, le PNE a donc privilégié cette dernière solution pour ses portails web de diffusion Rando-Ecrins et Biodiv'Ecrins, et pour le futur portail web de diffusion des stations et capteurs du parc.

Une fois que les agents ont pris en main les outils de consultation et de saisie développés, ils participent activement à la production de données. Dans le cadre de la réglementation française et européenne qui s'inscrivent depuis plusieurs années dans un mouvement de libération des données publiques, le PNE doit publier ses données géographiques sur des plateformes open data et renseigner leur métadonnées sur des catalogues afin qu'elles soient accessibles et réutilisables. Leur degré d'interopérabilité influençant fortement leur réutilisation, il est nécessaire de passer par une phase de standardisation des données (OGC), voir même de création de schémas de données. Afin de faciliter la publication des données en ligne, il est possible d'automatiser leur mise à jour en générant automatiquement le fichier téléchargeable à partir d'une BDD.

# **OUVERTURE ET DIFFUSION DES DONNEES**

Une de mes missions au cours du stage étant de contribuer à la publication des données produites par le PNE sur la plateforme open data gouvernementale (data.gouv.fr), la question de la mise à jour des jeux de données déjà publiés est apparue comme une tâche pertinente à automatiser, pour des raisons de gain de temps et d'actualisation des données. Étant donné que la documentation de data.gouv.fr n'était pas renseignée au sujet de la mise à jour dynamique d'une ressource distante (sur un serveur), je me suis penchée sur la question dans l'objectif de proposer et publier une solution sur le blog du SI du PNE<sup>24</sup> et de la diffuser aux géomaticiens des autres parcs nationaux, entre autres.

# Les enjeux de l'open data au PNE

Le droit d'accès des citoyens à l'information publique émerge en 1978 avec la loi dite CADA, et les mesures prises pour ouvrir les données publiques en France s'accélèrent depuis 2015 avec la loi dite « Valter » qui établit le principe de gratuité pour la réutilisation des données publiques. Puis l'année suivante, en 2016, la loi pour une République Numérique dite « Lemaire » instaure l'obligation pour les établissements publics de publier en ligne dans un standard ouvert leurs principaux documents, y compris leurs codes sources, ainsi que leurs bases de données et les données qui présentent un intérêt économique, social, sanitaire ou environnemental. Récemment, le 27 avril 2021, une nouvelle circulaire du Premier Ministre a vu le jour pour donner suite au rapport « Bothorel », positionnant la politique d'ouverture des données publiques comme une priorité stratégique de l'État.

Au niveau européen, c'est la directive INSPIRE qui depuis 2007, impose aux autorités publiques de rendre accessibles au public, en les publiant sur Internet, les données géographiques environnementales afin de favoriser la protection de l'environnement.

C'est dans ce cadre que les missions du géomaticien au PNE évoluent, puisqu'il doit publier de manière régulière ses données produites : faune, flore, itinéraires de randonnée et prochainement les données des capteurs du parc. D'où la nécessité d'alléger cette tâche en l'automatisant.

# Définition et objectifs de l'open data

L'open data, qu'est-ce que c'est ?

L'open data ou donnée ouverte est une donnée numérique dont l'accès et l'usage sont laissés libres aux usagers. Elle est diffusée de manière structurée selon une méthode et une licence ouverte garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière.

Une donnée publique est une donnée produite ou reçue dans le cadre d'une mission de service public.

• L'open data, quels objectifs?

L'open data s'inscrit dans une tendance qui considère l'information publique comme un bien commun dont la diffusion est d'intérêt public et général.

La publication en open data vise à :

- Renforcer la transparence de l'action publique
- Faciliter le travail des services publics
- Améliorer la connaissance et la recherche en valorisant la diffusion et la réutilisation des données
- Permettre la création de nouveaux services numériques et de nouveaux usages des données recueillies par les établissements publics en générant de l'activité économique autour de ces données

# L'open data et le catalogage

#### Les métadonnées

Les métadonnées sont des données décrivant d'autres données. Un enjeu majeur de l'open data pour la réutilisation des données publiées est la description standardisée de celles-ci. Les métadonnées identifient l'auteur, renseignent sur la production du jeu de données et définissent les usages de ce dernier. En facilitant la recherche des données et en aidant a la compréhension des jeux de données, elles permettent de préparer leur partage et leur interopérabilité tout en participant à leur pérennisation. Les métadonnées sont présentes soit de manière embarquée (ex : dans un fichier informatique : logiciel, document XML...), soit de manière externe (ex : dans un catalogue).

La norme reconnue et utilisée pour la rédaction des métadonnées, est la norme ISO19115 de l'Organisation Internationale de Normalisation. Son utilisation est par ailleurs préconisée par la Directive INSPIRE qui insiste particulièrement sur l'importance de disposer de métadonnées harmonisées, libres d'accès, afin de faciliter l'usage des données géographiques.

#### Différencier publication et catalogage

Les recommandations et obligations liées à l'open data et au catalogage étant complexes à cerner, j'ai effectué un travail de documentation afin d'éclaircir ces points, que j'ai synthétisé dans le document ci-dessous (figure 11).

## **OUVERTURE DES DONNEES PUBLIQUES**

1789	Déclaration des Droits de l'Homme et du Citoyen « La société a le droit de demander compte à tout agent public de son administration »	F R A
1978	Loi CADA  Veille à la liberté d'accès du public aux documents administratifs et s'assure de la bonne réutilisation des informations publiques	N C E
2007	Concerne la diffusion des données géographiques environnementales Les acteurs publics doivent dorénavant les: CATALOGUER - PARTAGER - HARMONISER Pour cela ils devront: - Rédiger des fiches de métadonnées - Fournir un lien de téléchargement + une possiblité de visualisation - Utiliser un modèle de données commun	E U R O P E
2015	Loi dite Valter Instaure la gratuité de l'information publique et de sa réutilisation	
2016	Loi pour une République Numérique L'ouverture des données concerne également : - Les codes sources - Les algorithmes - Les bases de données	F R A
27 avril 2021	Circulaire du Premier Minsitre sur l'opendata  La "Politique publique de la donnée, des algorithmes et des codes sources" devient une priorité stratégique de l'État en s'inscrivant dans les missions des ministères	C E

#### **OPENDATA**

#### Concerne toutes les données

Publication des jeux de données d'intérêt public non sensibles:

ex: data.gouv.fr

- Titre spécifique et précis
- Description brève --> contenu, utilité
- Fréquence de mise à jour
- URL de téléchargement

Associer un sigle

Ajouter des mots-clés

Préciser la couverture temporelle et spa-

Préciser la granularité spatiale

Données tabulaires : CSV

Données géographiques

- Vecteurs : GeoJSON, GeoPackage

- Rasters : JPEG2000, GeoTIFF

#### CATALOGUE

#### Concerne les données géographiques

### Obligations

Possibilités

Répertoire organisé de publication des métadonnées répondant à la norme ISO 19115 des données géographiques produites

ex : geocatalogue.fr

Moissonner les données vers :

- data.gouv.fr (que les données ouvertes, géographiques ou non)
- geo.data.gouv.fr (que les données géographiques, opendata ou non)

Fournir une URL de téléchargement

Formulaire en ligne pour les métadonnées (nécessité d'avoir un compte)

Guide du CNIG pour créer ce formulaire : http://cnig.gouv.fr/?page\_id=9108

#### Licence Ouverte 2.0

**Formats** 

Autorisation de réutilisation des données sous réserve de mentionner la source et la dernière date de mise à jour

#### Licence oDbl

Autorisation de réutilisation des données sous réserve de mentionner la source + Obligation de réciprocité quant au partage des données

Cendrine HOARAU - Parc National des Ecrins - Juin 2021

Figure 11: Ouverture des données publiques, Open data et catalogage

# Publication de jeux de données

Après avoir pris connaissance des obligations liées à la diffusion des données publiques en open data, l'étape suivante consiste à créer le jeu de donnée qui sera publié.

## Publier sur une plateforme open data

Les données, pour être accessibles par le plus grand nombre, sont publiées sur des plateformes open data qui les agrègent. Un jeu de données est publié sous forme de fichier téléchargeable qui peut soit être chargé directement sur la plateforme, soit être publié sous forme de lien vers un serveur distant. Dans tous les cas, le fichier doit être mis à jour de manière régulière afin que le public ait accès à une information récente et pertinente.

Le PNE a publié sur la plateforme gouvernementale data.gouv.fr quatre jeux de données<sup>25</sup>, dont deux qui concernent ses itinéraires de randonnées et ses patrimoines associés. Ce sont ces deux jeux de données que j'ai actualisé.

En parallèle de mon stage, Camille Monchicourt et Idrissa Djepa, en stage au Parc national des Cévennes, ont travaillé sur un schéma de données pour les itinéraires de randonnées (voir p.48), afin de favoriser et standardiser les échanges de ces données entre structures productrices et utilisatrices (communautés de communes, parcs naturels, départements...). Plusieurs réunions ont eues lieux avec les producteurs et gestionnaires de ces données, ce qui a donné lieu à une V0 (première version) d'un schéma de données, en vue d'une publication future sur le site gouvernemental de référencement des schémas de données publiques en France<sup>26</sup>. La mise à jour du jeu de données des itinéraires de randonnée du PNE en open data reprend donc la V0 du schéma de données.

Les parcs nationaux disposent de bases des données spatiales pour stocker les informations de biodiversité et d'itinéraires de randonnées sur leur territoire (respectivement GeoNature et Geotrek). Afin d'extraire les données, il faut mettre en forme les informations à publier dans une vue de la base de données. Une fois la vue créée, il faut en exporter le contenu dans un format de données qui facilitera sa réutilisation (ex : CSV, JSON...).

<sup>25</sup> Jeux de données publiés par le PNE <a href="https://www.data.gouv.fr/fr/organizations/parc-national-des-ecrins/">https://www.data.gouv.fr/fr/organizations/parc-national-des-ecrins/</a> 26 Schémas de données publiques en France <a href="https://schema.data.gouv.fr/">https://schema.data.gouv.fr/</a>

#### Méthode de création des vues dans la BDD Geotrek

Les itinéraires de randonnée du PNE et leurs POI (Point of Interest) associés sont stockés dans une même base de données spatiale PostGIS, sur un serveur distant. Des vues ont été créées dans cette BDD Geotrek, afin de sélectionner et de rassembler des informations provenant de plusieurs tables. Mon objectif, est de partir de la vue « v\_treks » déjà existante et de créer une nouvelle vue qui tienne compte des champs de la V0 du schéma de données. Afin de ne pas altérer la BDD Geotrek, une BDD Geotrek de demo a été mise en place, sur laquelle j'ai pu m'exercer sans risque.

L'export SQL permettant de créer les vues est disponible sur le dépôt github dédié<sup>27</sup>.

La BDD Geotrek contenant des itinéraires de randonnées de différentes sources (ex : Offices du tourisme, PNE...), des tronçons non publiés, ainsi que des étapes d'itinérances ne constituant pas un itinéraire de randonnée en soi, j'ai appliqué plusieurs filtres à la fin de ma requête SQL pour ne récupérer que les itinéraires de randonnée à la journée déjà publiées et dont la source est le PNE :

```
WHERE t.published = true AND t.structure_id = 1 AND t.route_id <> 5 AND
e.deleted = false
```

Pour récupérer les noms des communes parcourues par chaque itinéraire, j'ai utilisé la fonction *ST\_Intersects* qui permet d'opérer une intersection spatiale entre d'un côté le tronçon de randonnée et de l'autre, la table contenant la géométrie des communes :

```
SELECT t.topo_object_id AS tid, string_agg(z.name, ', ') AS communes
    FROM public.zoning_city z, v_treks t
    WHERE ST_intersects(t.geom, z.geom)
    GROUP BY t.topo_object_id
```

Afin de récupérer les informations qui n'étaient pas contenues dans la vue initiale « v\_treks », j'ai effectué des requêtes WITH, également connues sous le nom de Common Table Expressions (CTE). Ces requêtes peuvent être vues comme des tables temporaires qui n'existent que pour une requête. L'intérêt de SELECT dans WITH est de diviser des requêtes complexes en parties plus simples. Voici un exemple :

<sup>27</sup> Export SQL rando opendata <a href="https://github.com/PnX-SI/Ressources-techniques/blob/master/Geotrek/export-rando-opendata.sql">https://github.com/PnX-SI/Ressources-techniques/blob/master/Geotrek/export-rando-opendata.sql</a>

Pour créer la vue des patrimoines (POI) associés aux itinéraires de randonnée publiés, j'ai dû créer des sous-requêtes SQL (*SELECT* imbriqué dans un *SELECT*), dans des requêtes *WITH*, afin de faire coïncider les géométries des points avec le système de référencement linéaire (segmentation dynamique) des tronçons :

```
WITH topo_poi AS (
SELECT *
    FROM core_pathaggregation
    WHERE topo_object_id IN (
        SELECT t.id
        FROM core_topology t
        JOIN v_pois i ON i.id = t.id
        WHERE kind = 'POI' AND deleted = false AND i.published = true
    )
)
SELECT
p.geom
FROM topo_trek t
JOIN topo_poi ON topo_poi.path_id = t.path_id
    AND (
     (t.start_position <= t.end_position AND topo_poi.start_position
between t.start_position AND t.end_position)
     (t.start_position > t.end_position AND topo_poi.start_position
between t.end_position AND t.start_position)
JOIN v_pois p ON p.id = topo_poi.topo_object_id
```

La segmentation dynamique<sup>28</sup>, ou système de référencement linéaire, est une notion de géomatique qui consiste à localiser des objets à partir d'un réseau de lignes. Cette méthode a été utilisée par le PNE pour stocker ses tronçons de randonnées dans la BDD Geotrek. Les géométries sont décrites à l'aide d'une mesure le long des lignes du réseau, et lorsqu'un segment du réseau est modifié, celles-ci sont mises à jour automatiquement.

<sup>28</sup> La segmentation dynamique <a href="https://makina-corpus.com/blog/metier/2014/la-segmentation-dynamique">https://makina-corpus.com/blog/metier/2014/la-segmentation-dynamique</a>

Ainsi, les équipements, les chantiers, les statuts fonciers, les droits de passage, la nature des chemins, le patrimoine, les itinéraires de randonnées, ou encore la signalétique sont modélisés en fonction des chemins qui composent le réseau de tronçons.

Procéder comme tel présente plusieurs avantages :

- Aucune redondance des géométries, seule celle du réseau est saisie, les autres sont calculées
- Association d'attributs à la totalité ou une partie d'un linéaire
- Liens automatiques entre les objets grâce à leur superposition vis-à-vis du réseau
- Garantie de la cohérence géométrique et topologique entre les tronçons et tous les objets

#### Extraction des vues en GeoJSON

Il y a plusieurs moyens de récupérer les vues créées et de les enregistrer au format souhaité. Il est possible notamment de passer par un logiciel SIG (ex : QGIS) en se connectant à la base de données PostGIS et en ouvrant le contenu de la vue afin de l'enregistrer sous forme de fichier dans un format choisi (Geopackage, GeoJSON...). Cependant cette méthode n'est pas la plus efficace en terme de temps. Il existe une méthode plus rapide, qui nécessite de passer en ligne de commandes avec la librairie GDAL/OGR (voir p.21). Une des commandes d'OGR, ogr2ogr, permet de convertir des données vecteurs vers divers formats de fichiers, tout en réalisant d'autres opérations pendant le processus comme des sélections spatiales et/ou attributaires ou encore la définition du système de coordonnées en sortie.

Pour récupérer une vue créée dans une base de données avec *ogr2ogr*, il est nécessaire de renseigner les informations de connexion à la base de données ainsi que la requête SQL qui permet de sélectionner les données souhaitées. Il faudra également renseigner le chemin de sortie du fichier, qui se trouve dans notre cas sur un serveur web distant. De cette manière, on dispose d'une URL qui sera notre lien de téléchargement pour le jeu de données publié sur la plateforme open data.

Exemple de conversion de fichier en GeoJSON:

```
ogr2ogr -f "GeoJSON" home/pne/files/randos_pne.geojson PG:"host=****
user=**** dbname=****** password=****" -sql "select * from
public.v_opendata_treks"
```

Le format privilégié par le PNE pour publier des données spatiales est le GeoJSON car c'est un format ouvert et simple d'échange de données géographiques. Le PNE a fait le choix de publier également ses données en format CSV, format tabulaire, plus facilement réutilisable par des utilisateurs qui ne disposent pas d'outils cartographiques pour traiter les données en GeoJSON.

Exemple de conversion de fichier en CSV (itinéraires de randonnées) :

```
ogr2ogr -t_srs EPSG:4326 -f CSV -lco GEOMETRY=AS_WKT -lco SEPARATOR=SEMICOLON /home/pne/Data/files/randos_pne.csv PG:"host=**** user=**** dbname=****** password=****" -sql "select * from public.v_opendata_treks"
```

#### • Rendre le fichier téléchargeable

Pour que ce fichier soit accessible sur une URL en HTTP, il faut ensuite créer une configuration Apache<sup>29</sup>.

Un serveur HTTP permet à un site web de communiquer avec un navigateur en utilisant le protocole HTTP(S). Apache est probablement le serveur HTTP le plus populaire. C'est donc lui qui met à disposition la plupart des sites Web du WWW. C'est un logiciel libre fourni sous la licence spécifique Apache.

On utilise généralement Apache en conjonction avec d'autres logiciels, puisqu'il permet d'interpréter du code et d'accéder à des bases de données. Un seul serveur Apache permet de déployer simultanément plusieurs sites et services qu'il faut configurer individuellement.

Lorsqu'il démarre, Apache charge les fichiers de configuration et se met en attente de requêtes sur les interfaces réseaux. On dit qu'il écoute certains ports.

Lorsqu'on utilise un navigateur web, que l'on clique sur un lien ou qu'on rentre directement une URL dans la barre d'adresse, on effectue une requête :

- Le navigateur résout le nom de domaine (il obtient l'adresse IP du serveur).
- Il envoie une requête HTTP avec la méthode GET à l'IP du serveur sur le port 80 pour lui demander de retourner un contenu particulier.
- Le serveur HTTP reçoit la requête, et en fonction de divers paramètres (URL appelée, configuration du serveur, etc.), va chercher un contenu dans un fichier ou lance un script qui va générer un contenu.
- Le serveur renvoie ce contenu à l'IP du navigateur.

<sup>29</sup> Documentation ubuntu Apache <a href="https://doc.ubuntu-fr.org/apache2">https://doc.ubuntu-fr.org/apache2</a>

Exemple de configuration pour le domaine http://data.ecrins-parcnational.fr:

Dans le fichier "/etc/apache2/sites-available/data.conf" :

#### • Publier le jeu de donnée sur data.gouv.fr

Il ne reste maintenant plus qu'à plus publier le jeu de données en open data. Il existe de nombreuses plateformes open data où il est possible de déposer un jeu de données. La plateforme qui a le plus de visibilité et qui recense le plus de données au niveau national est data.gouv.fr<sup>30</sup>, c'est pourquoi le PNE a fait le choix de publier ses jeux de données sur cette plateforme.

On commence par créer et décrire le jeu de données, puis à y ajouter une ressource, sous forme d'URL vers notre fichier distant, plutôt qu'en chargeant un fichier statique :



Figure 12: Capture d'écran de l'interface de publication d'un lien vers un fichier distant sur data.gouv

<sup>30</sup> Publications du PNE en open data <a href="https://www.data.gouv.fr/fr/organizations/parc-national-des-ecrins/">https://www.data.gouv.fr/fr/organizations/parc-national-des-ecrins/</a>

# Référencer sur un catalogue et renseigner les métadonnées

Une fois le jeu de données publié, il est aussi possible de le référencer sur un catalogue de données, en renseignant ses métadonnées. Le catalogue représente l'emplacement centralisé où sont regroupées les informations sur les données, les métadonnées. Le choix du catalogue de référencement appartient au producteur des données. Il est préférable d'importer l'information sur un seul et unique catalogue qui sera moissonné par la suite, par le biais d'un mécanisme permettant d'en récupérer le contenu pour le publier sur une autre plateforme.

Dans le cadre de la directive européenne Inspire, les parcs nationaux de France ont mis en place une infrastructure d'informations géographiques. Il s'agit d'un ensemble de services Internet qui permettent la diffusion et le partage de leurs données géographiques<sup>31</sup>. Chaque parc national dispose de son propre catalogue de données et il est possible de parcourir directement les fiches de métadonnées existantes par thème INSPIRE.

La création d'une fiche de métadonnées nécessite l'identification du créateur à l'interface. L'interface dispose d'un système de création de fiche, basé sur la complétion de champs, répondant aux normes ISO19115. Après avoir renseigné tous les champs, il est possible de renseigner l'URL de publication du jeu de données afin de le moissonner sur le catalogue.

Le remplissage de tous les champs pour créer la fiche de métadonnées est assez fastidieux et même s'il permet de renseigner de manière complète un jeu de données, cette tâche a tendance à être délaissée par les gestionnaires de données, par manque de temps. Favoriser interopérabilité des données ne devrait pas être une contrainte, le risque est que cette étape de catalogage soit négligée et que ceci constitue un frein à la réutilisation des jeux de données. Une piste d'amélioration pour alléger cette procédure pourrait être par exemple de réduire la saisie manuelle à l'essentiel : titre, résumé, mots-clés, champs, résolution.

# Favoriser l'interopérabilité des données

#### Standards de l'OGC

Favoriser l'interopérabilité des données est un enjeu majeur de l'ouverture des données, et c'est une des missions de l'OGC. L'Open Geospatial Consortium (OGC), est une organisation internationale à but non lucratif fondée en 1994 pour répondre aux problèmes d'interopérabilité des systèmes d'information géographique (SIG). Les missions de ce consortium sont de regrouper tous les acteurs concernés afin de développer et

<sup>31</sup> Catalogue des parcs nationaux <a href="https://catalogue.parcnational.fr/">https://catalogue.parcnational.fr/</a>

promouvoir des standards ouverts garantissant l'interopérabilité dans le domaine de la géomatique et de l'information géographique et de favoriser la coopération entre développeurs, fournisseurs et utilisateurs.

Les standards ouverts de l'OGC<sup>32</sup> fournissent des modèles de données uniformes et interexploitables permettant de créer, de reproduire, de mettre à jour, de stocker des données et de mettre en place des services géographiques. Ils sont présentés sous forme de documents techniques où sont détaillés les interfaces et les codages. Les concepteurs de logiciels et les producteurs de données les utilisent pour créer des interfaces et des codages libres dans leurs produits et services. L'utilisation de ces standards encourage l'innovation, favorise la transparence des organismes producteurs de données et sont un indicateur de qualité.

Lors de la construction du MCD de la BDD Capteurs (voir p.13), il a été nécessaire d'engager une phase de documentation au sujet des standards de l'OGC existants liés aux capteurs et aux mesures physiques. Les recherches ont été synthétisées dans un document PDF<sup>33</sup> que j'ai présenté aux géomaticiens des autres parcs nationaux à l'occasion du café géomatique, point régulier de présentation des avancements de chaque parc.

Le standard Sensor Web Enablement (SWE)<sup>34</sup> est un ensemble de services permettant aux développeurs de faire en sorte que tous les types de capteurs, de transducteurs et de dépôts de données des capteurs puissent êtres investigués, accessibles et utilisés via le Web. En 2016, l'OGC a accepté la spécification standard de l'API SensorThings, qui fournit un cadre ouvert et unifié pour interconnecter les dispositifs de détection de l'internet des objets (IoT), les données et les applications sur le Web. C'est sur le modèle de données de cette API (voir Annexe 5) que je me suis référée pour concevoir mon MCD.

#### • Schémas de données

Les schémas de données permettent également de décrire des modèles de données : les différents champs, comment sont représentées les données, quelles sont les valeurs possibles, etc. Les schémas permettent entre autres de valider qu'un jeu de données se conforme à un schéma, de générer de la documentation automatiquement ou encore de proposer des formulaires de saisie standardisés. Fonctionnant à la manière de gabarits, ils permettent une montée en qualité des données proposées en open data,

<sup>32</sup> Standards OGC <a href="https://www.ogc.org/docs/is">https://www.ogc.org/docs/is</a>

<sup>33</sup> Document Creation BDD Capteurs et standards

surtout lorsque plusieurs producteurs de données produisent un même jeu de données, comme c'est le cas pour les données d'itinéraires de randonnée.

Ainsi, afin de favoriser les échanges entre structures productrices et utilisatrices de ce type de données, un schéma a été travaillé par un groupe de travail composés d'acteurs du tourisme pédestre (entre autres l'IGN, APIDAE, DATATourisme, Région SUD...). L'évolution et les propositions issues de ce projet commun ont été recensées dans un dépôt github mis en place par Idrissa Djepa<sup>35</sup>.

Lors d'une première réunion avec ce groupe de travail, un document commun a été créé pour identifier les champs du schéma et leurs caractéristiques<sup>36</sup>, constituant la première version (V0) du schéma. C'est à partir de cette version que j'ai créé ma vue des itinéraires de randonnées, issue de la BDD Geotrek.

# Automatiser la publication des données

Tout le processus de publication des données en open data étant assez long, une solution pour automatiser la mise à jour des jeux de données en ligne issues d'une BDD, a été mise en place.

Après avoir créé une vue des données à extraire de la BDD, il faut mettre mettre en place une tache planifiée sur un serveur pour automatiser la génération du fichier à intervalle régulier et le rendre accessible sur une URL fixe.

Sous le système d'exploitation Linux, c'est l'outil Crontab<sup>37</sup> qui permet de lancer des tâches de façon automatisée et régulière. Il suffit de définir la temporalité à laquelle on souhaite régénérer le fichier et insérer la commande ogr2ogr.

Dans l'exemple suivant, nous créons une tache Cron pour l'utilisateur "root" qui s'exécutera sur le serveur tous les jours à 3 heures du matin. Pour le créer on exécute la commande "sudo crontab -e" puis on y inclut la commande :

```
0 3 * * * ogr2ogr -f "GeoJSON" /home/pne/files/randos_pne.geojson
PG:"host=**** user=**** dbname=**** password=****" -sql "select * from
public.v_opendata_treks"
```

Le chemin de sortie du fichier que nous avons renseigné se trouve sur un serveur distant. De cette manière, on dispose d'une URL qui sera notre lien de téléchargement du fichier pour le jeu de données publié sur la plateforme open data.

<sup>35</sup> Dépôt github Schéma données <a href="https://github.com/ldrissaD/schema randonnee/">https://github.com/ldrissaD/schema randonnee/</a>

<sup>36</sup> Schéma V0 https://calc.ecrins-parcnational.fr/schema-rando

<sup>37</sup> Documentation Crontab <a href="https://www.linuxtricks.fr/wiki/cron-et-crontab-le-planificateur-de-taches">https://www.linuxtricks.fr/wiki/cron-et-crontab-le-planificateur-de-taches</a>

Le fichier des randonnées du PNE, mis à jour quotidiennement est ainsi accessible à l'URL permanente : <a href="http://data.ecrins-parcnational.fr/files/randos\_pne.geojson">http://data.ecrins-parcnational.fr/files/randos\_pne.geojson</a>.

#### BILAN

## Résultats et pistes d'amélioration

Mon stage s'est déroulé autours de 3 missions principales :

- Contribuer à l'actualisation du SIG du PNE
- Contribuer à la diffusion des données produites par le PNE
- Concevoir et mettre en place une base de données des capteurs et stations météo du PNE et développer des applications web cartographiques de consultation et de saisie de données

La première mission aura nécessité plus de temps que je ne m'étais imaginé. En effet, le téléchargement des fonds raster IGN a été fastidieux et leur mise en forme a pris du temps. Je me suis beaucoup documentée pour trouver des solutions performantes afin de fusionner les tuiles sans trop altérer la résolution d'origine tout en essayant d'optimiser leur espace de stockage. L'échec du fichier VRT pour l'affichage de la BD ORTHO sur QGIS demeure incompris, c'est pourquoi j'aimerai trouver une solution alternative au fichier unique en JP2. Une piste non explorée par manque de temps, aurait été d'envisager la construction des fonds de carte raster avec MBTiles<sup>38</sup>, déjà utilisé par les applications mobiles GeoNature et Geotrek. Par ailleurs, le transfert de la Geobase en geopackage ainsi que la création d'un nouveau projet-type utilisant les données de référence sous ce format est une réussite puisque cela permet d'optimiser l'espace de stockage du serveur et de simplifier l'utilisation des couches SIG (fichier unique).

Pour la deuxième mission, j'ai découvert une facette du métier de géomaticien que j'avais peu exploré lors du M1. Je me suis rendue compte à quel point le catalogage pouvait être chronophage pour un géomaticien, mais également combien les métadonnées et les standards pouvaient être utiles pour l'échange de données géographiques. Je suis fière d'avoir proposé une solution facilitant la publication des données en open data et d'avoir rédigé un article présentant cette méthode, en HTML. Je me suis rendue compte à l'issue de cette publication, de l'importance de coopérer et de partager ses avancées en ligne afin qu'elles profitent à toute la communauté géomatique (via Github, Georezo mais aussi les réseaux sociaux professionnels comme Linkedin).

<sup>38</sup> Wiki MBTiles <a href="https://wiki.openstreetmap.org/wiki/MBTiles">https://wiki.openstreetmap.org/wiki/MBTiles</a>

Enfin, la dernière mission a constituée la majeure partie du stage. En effet, le projet de construction d'une BDD spatiale des capteurs et stations météo du parc étant au commencement à mon arrivée en stage, l'organisation des premières réunions et la compréhension des besoins pour construire le MCD ont pris du temps. J'ai apprécié collaborer avec tous les acteurs qui ont gravité autour de ce projet et j'ai énormément appris à leur côté. Je me suis également sentie fière d'avoir développé deux solutions de webmapping, toutes deux reliées à la BDD Capteurs que j'ai mis en place. Enfin, nous avons réussi avec Baptiste Garcin, un autre stagiaire qui était en charge de tester la mise en place de capteurs reliés au réseau LoraWan, à connecter les données issues de ces capteurs à la BDD.

## Montée en compétences

De part la diversité des missions du stage, et grâce à un accompagnement de qualité du personnel du pôle SI, j'ai pu monter en compétences sur de nombreux points.

La mission de structuration et de modernisation du SIG du parc a été l'occasion pour moi de progresser en technicité. En effet, la découverte de GDAL/OGR pour effectuer des géotraitements en ligne de commande me permettra à l'avenir de gagner du temps pour traiter des fichiers lourds, sans passer par l'interface graphique d'un logiciel SIG. Travaillant sur un poste Linux, j'ai également découvert l'usage du terminal pour installer des logiciels, me déplacer dans les fichiers du répertoire de documents et effectuer des commandes spatiales. La construction d'une BDD avec PostGre/PostGIS a été l'occasion d'approfondir les compétences en SQL acquises en M1. Enfin, la découverte des profils utilisateurs sur QGIS ainsi que la participation aux sessions de formation QGIS des agents, m'ont permis d'élargir mes usages du logiciel.

La phase de déploiement d'outils de webmapping m'a également fait fortement progresser en technicité puisque j'ai appris un nouveau langage de programmation. La création d'une carte interactive avec la bibliothèque Leaflet a été l'occasion de m'initier au langage JavaScript.

En général, j'ai développé des compétences sur une multitude d'outils open source : QGIS, GDAL/OGR, PostGre/PostGIS, Linux, Libre Office, Lizmap et Leaflet.

J'ai également acquis de nombreuses connaissances théoriques sur les standards de l'OGC et l'open data. D'autres connaissances transversales sur le fonctionnement d'une architecture client-serveur, du réseau LoraWan, et des capteurs et stations météo, pourront certainement me servir dans mon avenir professionnel.

Côté relationnel, j'ai développé mon réseau professionnel et j'ai acquis une certaine confiance pour présenter mes travaux à un collectif, confiance que j'ai également développé en participant activement à la vie du parc.

Pour résumer, j'ai approché tous les aspects du rôle de géomaticien dans un parc national. Je me sens chanceuse d'avoir effectué ce stage en première année de Master puisqu'il a permis de mettre en pratique et compléter un large panel des enseignements déjà reçus.

## Perspectives

Ce stage a été pour moi l'occasion d'approfondir non seulement mes connaissances et compétences en géomatique, mais aussi de découvrir les acteurs qui préservent le milieu montagneux dans lequel je vis depuis peu, les Ecrins. Le PNE étant un des établissements publics pionniers en matière d'innovation structurelle, ce stage a largement participé à mon ouverture au monde de l'open source et de l'open data et à mon appétence pour le travail collaboratif. J'ai également eu le goût de participer à la formation des agents du parc et d'élargir mes connaissances environnementales.

A l'avenir, j'aimerai rester dans un environnement professionnel aussi dynamique et moderne que le PNE et approfondir mes compétences sur les outils open source. J'aimerai également en apprendre davantage en développement informatique, au service des SIG.

#### **BIBLIOGRAPHIE**

LECHÉMIA, Théo, 2016. Création d'un atlas dynamique de la faune et de la flore au Parc national des Écrins. Grenoble : Université Grenoble-Alpes ; Parc national des Écrins

MONCHICOURT, Camille, 2008. La diffusion des données localisées au Parc national des Ecrins. Parc national des Écrins.

DIDIER, Marie, 2020. Inventorier, structurer et organiser les données d'un lac d'altitude : le lac de la Muzelle. Parc national des Écrins.

## **WEBOGRAPHIE**

Biodiversité : une plateforme et des données présentées à la secrétaire d'Etat (2021, février). <a href="https://www.ecrins-parcnational.fr/actualite/biodiversite-plateforme-donnees-presentees-secretaire-etat">https://www.ecrins-parcnational.fr/actualite/biodiversite-plateforme-donnees-presentees-secretaire-etat</a>

Le Parc national des Écrins « libère » des données ! Ecrins Parc National (2018, décembre). <a href="https://www.ecrins-parcnational.fr/actualite/parc-national-ecrins-libere-donnees">https://www.ecrins-parcnational.fr/actualite/parc-national-ecrins-libere-donnees</a>

Les parcs nationaux : 11 espaces naturels protégés d'exception (2021). <a href="http://www.parcsnationaux.fr/fr/des-decouvertes/les-parcs-nationaux-de-france/les-parcs-nationaux-11-espaces-naturels-proteges">http://www.parcsnationaux.fr/fr/des-decouvertes/les-parcs-nationaux-de-france/les-parcs-nationaux-11-espaces-naturels-proteges</a>

Geotrek, dans les coulisses des portails de randonnée (Novembre 2019) <a href="https://www.ecrins-parcnational.fr/actualite/geotrek-coulisses-portails-randonnee">https://www.ecrins-parcnational.fr/actualite/geotrek-coulisses-portails-randonnee</a>

Les geeks de GeoNature carburent (Octobre 2018) <a href="https://www.ecrins-parcnational.fr/actualite/geeks-geonature-carburent">https://www.ecrins-parcnational.fr/actualite/geeks-geonature-carburent</a>

Dépôt Github Geosensor : <a href="https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql">https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql</a>

Dépôt Github de la BDD capteurs: https://github.com/PnEcrins/GeoSensor/blob/main/data/geosensor.sql

Blog du SI du PNE : <a href="https://si.ecrins-parcnational.com/blog/">https://si.ecrins-parcnational.com/blog/</a>

Documentation ogr2ogr <a href="https://gdal.org/programs/ogr2ogr.html">https://gdal.org/programs/ogr2ogr.html</a>

Guide d'installation ogr2ogr <a href="https://mothergeo-py.readthedocs.io/en/latest/development/how-to/gdal-ubuntu-pkg.html">https://mothergeo-py.readthedocs.io/en/latest/development/how-to/gdal-ubuntu-pkg.html</a>

Ressources techniques ogr2ogr\_shp2gpkg https://github.com/PnX-SI/Ressources-techniques/tree/master/Python/Shp2GPKG/ogr2ogr\_shp2gpkg

BD TOPO IGN en Geopackage https://geoservices.ign.fr/la-bd-topor-au-format-geopackage

Les données géographiques souveraines – Rapport <a href="https://www.vie-publique.fr/rapport/37550-les-donnees-geographiques-souveraines-rapport-au-gouvernement">https://www.vie-publique.fr/rapport/37550-les-donnees-geographiques-souveraines-rapport-au-gouvernement</a>

Geoservices IGN https://geoservices.ign.fr/

OpendatArchives <a href="https://www.opendatarchives.fr/">https://www.opendatarchives.fr/</a>

IGN2Map <a href="https://geotribu.github.io/ign-fr-opendata-download-ui/index.html">https://geotribu.github.io/ign-fr-opendata-download-ui/index.html</a>

Documentation GDAL <a href="https://gdal.gloobe.org/gdal/index.html">https://gdal.gloobe.org/gdal/index.html</a>

Installation de la bibliothèque Libjasper <a href="https://stackoverflow.com/questions/44468081/unable-to-locate-package-libjasper-dev">https://stackoverflow.com/questions/44468081/unable-to-locate-package-libjasper-dev</a>

Dépôt github dédié au code HTML de la carte Leaflet <a href="https://github.com/PnEcrins/GeoSensor/blob/main/leaflet/geosensor.html">https://github.com/PnEcrins/GeoSensor/blob/main/leaflet/geosensor.html</a>

Jeux de données publiés par le PNE https://www.data.gouv.fr/fr/organizations/parc-national-des-ecrins/

Export SQL des itinéraires de randonnée <a href="https://github.com/PnX-SI/Ressources-techniques/blob/master/Geotrek/">https://github.com/PnX-SI/Ressources-techniques/blob/master/Geotrek/</a> export rando opendata.sql

La segmentation dynamique https://makina-corpus.com/blog/metier/2014/la-segmentation-dynamique

Documentation Ubuntu Apache <a href="https://doc.ubuntu-fr.org/apache2">https://doc.ubuntu-fr.org/apache2</a>

Catalogue des parcs nationaux <a href="https://catalogue.parcnational.fr/">https://catalogue.parcnational.fr/</a>

Document Creation BDD Capteurs et standards (Cendrine HOARAU) <a href="http://data.ecrins-parcnational.fr/documents/technique/si/2021-07-BDD-Capteurs-PNE-standards-OGC.pdf">http://data.ecrins-parcnational.fr/documents/technique/si/2021-07-BDD-Capteurs-PNE-standards-OGC.pdf</a>

Standards OGC <a href="https://www.ogc.org/docs/is">https://www.ogc.org/docs/is</a>

Dépôt github Schéma de données <a href="https://github.com/ldrissaD/schema randonnee/">https://github.com/ldrissaD/schema randonnee/</a>

V0 du schéma de donnée https://calc.ecrins-parcnational.fr/schema-rando

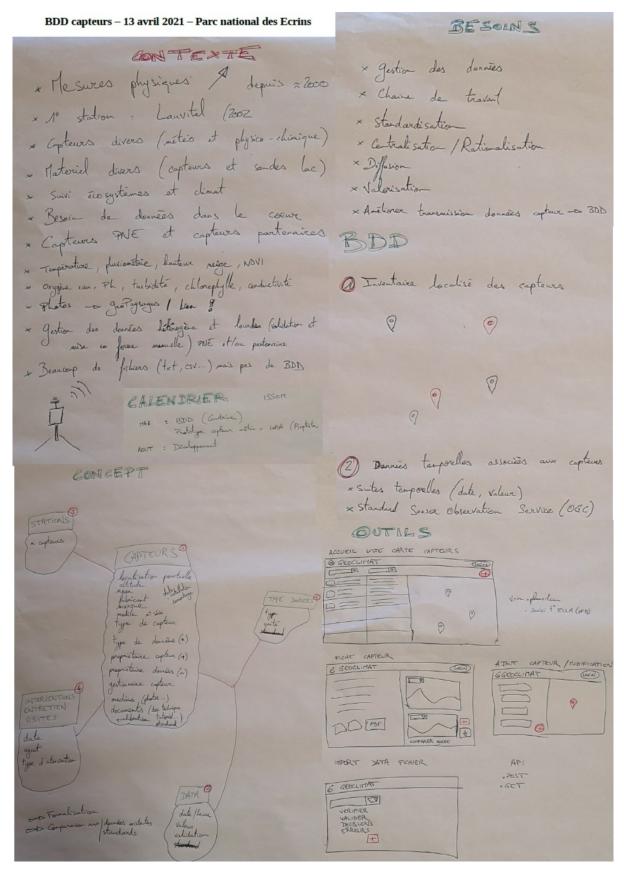
# **INDEX DES FIGURES**

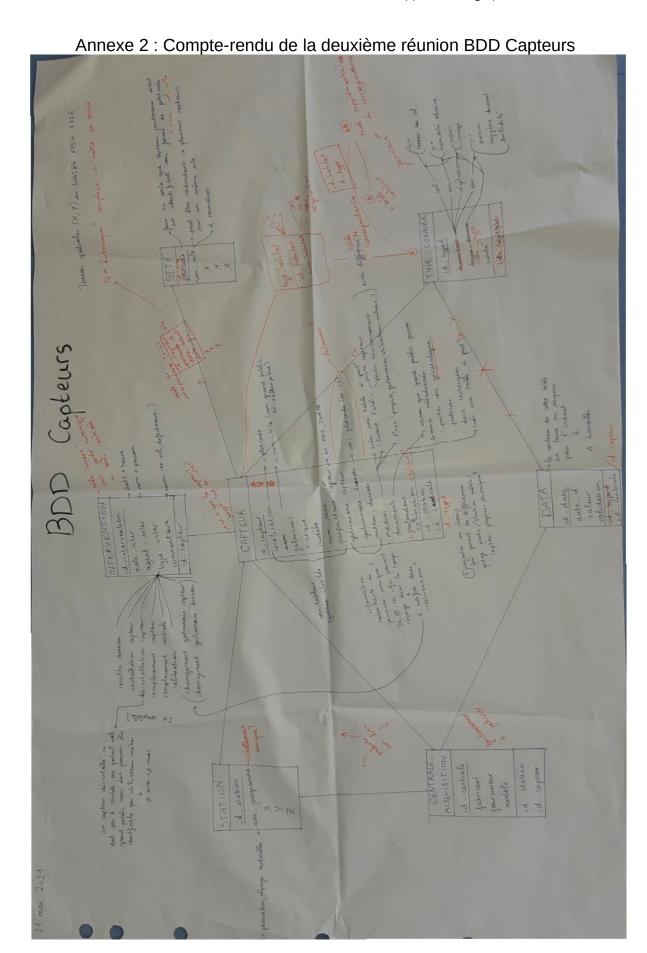
Figure 1: Les 11 Parcs nationaux de France, http://www.parcsnationaux.fr/	7
Figure 2: Cartographie du Parc national des Ecrins © Bruno Lafage / OFB	8
Figure 3: La chaîne de travail du SI, source PNE (2008)	10
Figure 4 : Relevé d'un capteur Alpages sentinelles	13
Figure 5: Projet-type QGIS du PNE	27
Figure 6: Profil et projet QGIS connectés à la BDD Capteurs	
Figure 7: Les trois niveaux d'accès à l'information géographique	30
Figure 8: Architecture de Lizmap, source 3Liz	32
Figure 9: WebSIG des capteurs du PNE réalisé avec Lizmap	33
Figure 10: Carte Leaflet connectée à la BDD Capteurs	36
Figure 11: Ouverture des données publiques, Open data et catalogage	40
Figure 12: Capture d'écran de l'interface de publication d'un lien vers un fichier dis	tant sur
data.gouvdata.gouv	46

## **INDEX DES ANNEXES**

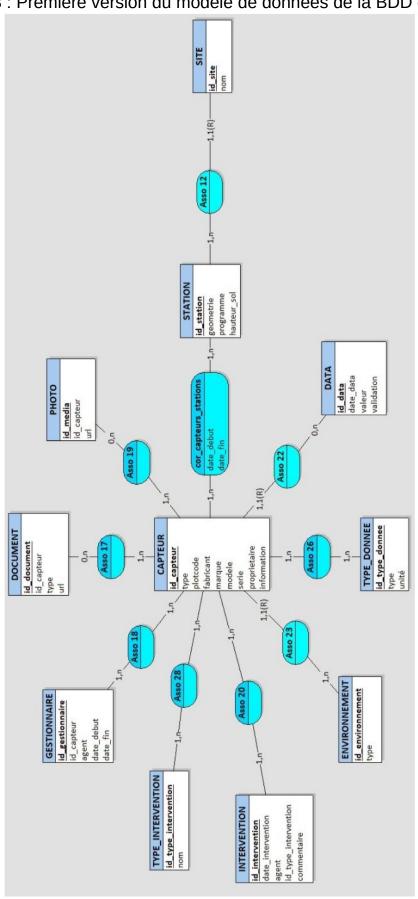
Annexe 1 : Compte-rendu de la première réunion BDD Capteurs	55
Annexe 2 : Compte-rendu de la deuxième réunion BDD Capteurs	56
Annexe 3 : Première version du modèle de données de la BDD capteurs	57
Annexe 4 : Dernière version du modèle de données de la BDD capteurs	58
Annexe 5 : Modèle de données de l'API SensorThings de l'OG	59

Annexe 1 : Compte-rendu de la première réunion BDD capteurs

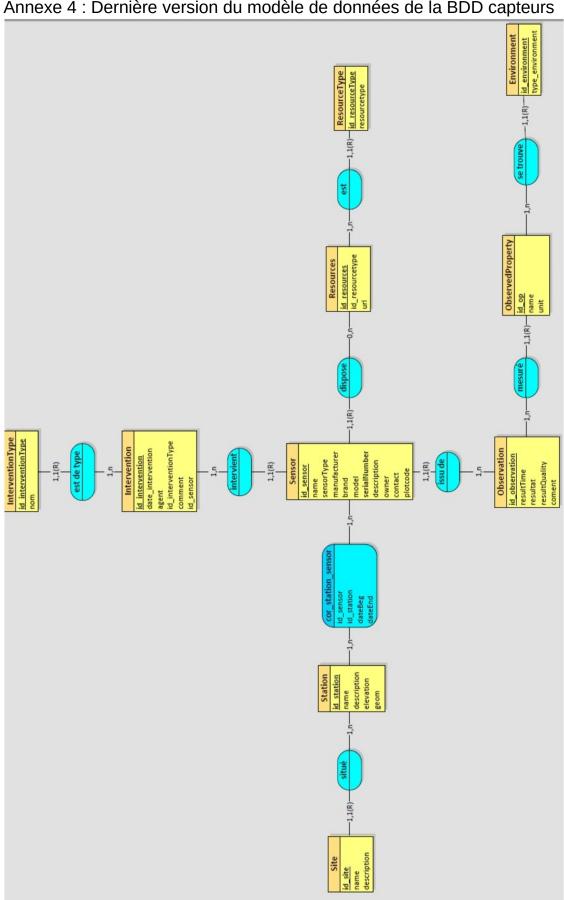




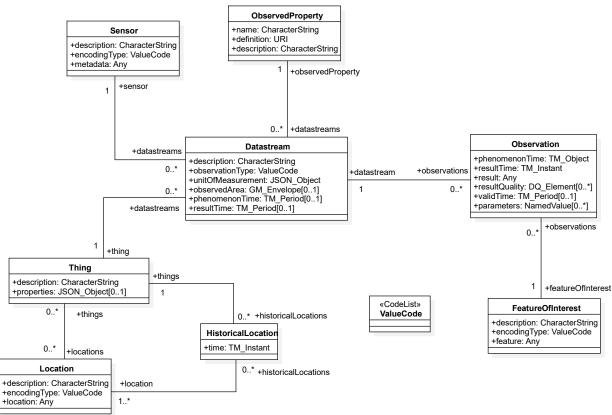
Page 56 sur 59



Annexe 3 : Première version du modèle de données de la BDD capteurs



Annexe 4 : Dernière version du modèle de données de la BDD capteurs



Annexe 5: Modèle de données de l'API SensorThings de l'OGC