



Atelier collaboratif de création d'un schéma de données pour les itinéraires de randonnées

Compte-rendu - 16/03/2021

Rédaction :

Lise Le Lann

Participants :

- Apidae Tourisme : Serge Bregliano et Pierre Granger
- Apitux : Jean-Christophe Becquet
- Auvergne-Rhône-Alpes Tourisme : Dominique Bambier
- Conseil départemental des Côtes d'Armor : Eric Ravenet
- Conseil départemental des Hautes-Alpes : Claire Lang
- Conseil départemental du Vaucluse : Fabrice Braverel
- DATAtourisme / ADN Tourisme : Pascale Vinot
- Etalab / DINUM : Geoffrey Aldebert
- IGN : Stephanie Barriere-Riffort, Herve Quinquenel et Etienne Tricoire
- Parc national des Écrins : Pierrick Navizet, Camille Monchicourt, Lise Le Lann, Cendrine Hoarau
- Parc national des Cévennes : Amandine Sahl et Idrissa Djepa Creutz
- Parc naturel régional du Luberon : Arnoul Hamel
- Région Ligurie : Giovanni Rocca
- Région SUD : Céline Coulomb, Joanne Musset

Objet :

Définir la pertinence de mettre en place un schéma de données des itinéraires de randonnées

Supports des présentations :

- <https://geotrek.ecrins-parcnational.fr/ressources/gt/10-schema-donnees-randonnees/2021-03-16-atelier-schema-rando.pdf>
- <https://geotrek.ecrins-parcnational.fr/ressources/gt/10-schema-donnees-randonnees/>
- DATAtourisme : <https://geotrek.ecrins-parcnational.fr/ressources/gt/10-schema-donnees-randonnees/2021-03-15-DATAtourisme-presentation.pdf>
- schema.data.gouv.fr : https://geotrek.ecrins-parcnational.fr/ressources/gt/10-schema-donnees-randonnees/2021-03-16-schema-data-gouv_presentation.pdf

Ordre de jour :

1. Tour de table des participants
2. Présentation du contexte, des ressources et des objectifs
3. Analyse de la pertinence et de la méthodologie
4. Analyse des ressources
5. Suites à donner

1. Tour de table des participants

Cette séance de travail en visioconférence a réuni 21 participants venant principalement du secteur publique et du domaine de l'open data.

2. Contexte

2.1. Présentation du contexte

Le Parc national des Écrins organise cet atelier de travail sur la pertinence de la création d'un schéma de données des itinéraires de randonnées pour anticiper les besoins de la communauté

de Geotrek sur la publication en open data des données et pour donner suite au travail réalisé dans le cadre du PITEM outdoor data.

Dans le cadre du programme européen Alcotra (PITEM MITO), le Parc national des Écrins est mandaté afin de réaliser un standard d'échange de données entre les différents partenaires français et italiens. Pour cela, il a réalisé une analyse du socle commun de données permettant de définir certaines activités outdoor, dont la randonnée. Fin 2020, ce standard a été validé par les partenaires du projet Européen. Afin d'apporter une valeur ajoutée à ce projet, le Parc national souhaite proposer la création d'un "schéma de données" concernant les itinéraires de randonnée à partir de ce socle commun. À terme, il est envisagé que ce schéma puisse enrichir les schémas disponibles sur le site <https://schema.data.gouv.fr>, ce qui permettrait de publier facilement des données standardisées et interopérables en open data, notamment sur le site <https://www.data.gouv.fr>.

Les utilisateurs de Geotrek commencent à publier sur différents portails d'open data (portail national ou régional). ême si les données publiées proviennent du même outil, leur publication est globalement semblable, mais la dénomination des champs publiés varie.

Les différentes réflexions sur le développement de Geotrek et maintenant les dialogues entre les membres de la communauté font ressortir des éléments communs sur la définition d'un itinéraire de randonnées. Il est aussi constaté que les itinéraires ont généralement comme géométrie des linéaires.

2.2. Présentation des ressources

Geotrek

Geotrek est une application Web et Mobile open source de gestion et de valorisation de l'offre de randonnée. Elle a été développée initialement par les Parcs nationaux des Écrins, du Mercantour et Alpi maritime en 2012. Depuis elle a été transférée à d'autres espaces naturels et à des départements de manière libre et gratuite. Elle s'enrichit progressivement de nouvelles fonctionnalités, et une communauté s'est construite au fil des années. Aujourd'hui, la communauté de Geotrek comprend une centaine de structures. La communauté se réunit régulièrement lors des rencontres Geotrek, un événement organisé une fois par an ou tous les deux ans.

Lors des rencontres de 2017, le Parc national des Cévennes et Apitux ont sensibilisé la communauté à l'open data et a fait émerger une dynamique.

En 2020, certains membres de la communauté de Geotrek ont contribué à la mission Bothorel sur la partie consacrée au logiciel libre. La communauté souhaite être active dans cette démarche d'ouverture des données.

L'atelier d'aujourd'hui permet de prendre les devants afin d'éviter la publication divergente de données liées à la randonnée.

Ces dernières années Geotrek s'est ouvert à d'autres outils afin de partager des données.

Échanges :

- le département du Vaucluse cherche à mutualiser les données vélo et randonnée voir l'outdoor, même s'il cible principalement les sports de déplacement. Est-ce la volonté de ce projet ?
 - Oui , il est souhaité de travailler sur un socle commun pour la randonnée mais pas uniquement sur la randonnée pédestre. On souhaite dans un premier temps travailler sur un socle commun en se concentrant sur les points communes entre toutes les pratiques. L'objectif est de réfléchir sur les champs qualifiant la randonnée et de se réinterroger sur le socle. Si le temps le permet, nous commencerons ce travail en fin d'atelier. Les attributs fondamentaux pressentis sont : le nom, la pratique, la durée, la difficulté, le parcours, le tracé (linéaire)

IGNrando'

Le portail web Espace Loisirs de l'IGN a vu le jour en 2014, offrant au grand public une offre d'itinéraires et aux professionnels qui le souhaitent un outil de saisie et de diffusion de leur offre. L'enjeu pour l'IGN est de valoriser son travail de création de données certifiées, de produire des topos et d'autres services tout en proposant une aide aux politiques publiques. En 2017, l'Espace Loisirs a été renommé IGNrando' dont l'objectif est de fédérer les producteurs et les consommateurs de données de randonnée autour d'une donnée structurées. Aujourd'hui l'IGN travaille sur une version enrichie de l'outil.

Les données de certains Geotrek sont déjà publiées dans le portail IGNrando'. Cette démarche s'est faite dans une logique de complémentarité entre les outils de diffusion territoriaux et nationaux.

Apidae

Apidae est un entrepôt de données touristiques collaboratif initialement développé par AURA Tourisme. La plateforme Apidae répond aux besoins des acteurs du tourisme. Elle évolue fréquemment en fonction des besoins et des attentes du secteur. Actuellement, un travail est mené sur une nouvelle définition de la structuration des données.

Apidae propose principalement des objets touristiques, mais également des équipements comprenant les randonnées. Lors de la création de la passerelle entre Apidae et Geotrek, un premier travail de correspondance des champs a été réalisé.

Il existe une passerelle entre Geotrek et Apidae. Les contenus touristiques d'Apidae remontent dans Geotrek et les randonnées de Geotrek remontent dans Apidae.

OpenStreetMap

OSM est de plus en plus présent dans les applications de randonnées et des pratiques Outdoor.

OSM s'est développé ces dernières années grâce à une communauté internationale composée de contributeurs qui enrichissent les contenus et les données. La communauté fait évoluer les objets existants en validant des étiquettes et des tags en fonction des besoins. Récemment, un travail sur les aménagements cyclables a été mené. Si le schéma est bien réalisé, il pourra venir enrichir l'objet randonnée d'OSM.

Le conseil départemental des Alpes-de-Haute-Provence a déjà publié des randonnées dans OSM.

Échanges :

- Le département du Vaucluse s'interroge sur le suivi et la fiabilité à long terme des données publiées par une institution.
 - Pour OSM, les territoires ont toutes les compétences pour être des contributeurs de l'outil et assurer le suivi de leurs données (logique de vérification des données pour être sûr qu'elle soient correctement modifiées). Certains territoires comme les villes d'Orange et de Montpellier le font déjà.
 - Il faut d'abord savoir si l'on souhaite contribuer ou non, puisque les données des structures sont dans d'autres outils et leurs modifications sur OSM n'impactent pas la qualité de la donnée dans les BDD de la structure.
- Le département des Hautes-Alpes se demande si un export automatique de Geotrek vers OSM est déjà mis en place ou est planifié .
 - Non

Autres opérateurs

Pour ce premier atelier le PNE a réuni une sphère d'acteurs publics pour travailler sur ce projet, mais d'autres acteurs / outils existent comme Visorando et la FFRP.

DATAtourisme

DATAtourisme est la plateforme gouvernementale, nationale et locale permettant d'agréger les données des SIT pour leur publication en open data. Elle s'organise autour d'une ontologie sur les objets du tourisme. Le travail d'ontologie a commencé en 2016. Elle a commencé à proposer une brève de standardisation des randonnées où les itinéraires sont un POI (point d'intérêt). Ce projet peut venir enrichir le modèle déjà existant de DATAtourisme mais il faut chercher la cohérence entre les deux schémas.

Sur le territoire national, il existe plus de 40 BDD touristiques institutionnelles (départementales et régionales) avec des vocabulaires différents. Le rôle de DATAtourisme est de réunir ces différentes BDD pour les publier de manière standardisée en open data. Pour ce faire, DATAtourisme travaille sur un écosystème national répondant aux normes européennes.

Aujourd'hui, 83 objets sont homogénéisés par DATAtourisme et sont publiés sur data.gouv.fr.

Échanges :

- DATAtourisme a un potentiel de traduction.
- Pour la construction du schéma, il faut prendre en compte l'ordre des informations principalement pour les étapes (itinérance)
- DATAtourisme s'inscrit dans une logique de complémentarité avec les SIT territoriaux. La plateforme ne permet pas et ne souhaite pas permettre la saisie de données. Elle a pour objectif de réceptionner des flux issus des BDD locales.
 - Par exemple, des données d'Apidae remontent dans DATAtourisme.

3. Les Schémas de données

3.1. Méthodologie

Data.gouv.fr propose pour la création de schéma de données la méthodologie suivante :

1. Phase d'investigation : envisager de créer un schéma de données ;
2. Phase de concertation : rassembler plusieurs parties prenantes pour créer un schéma de données ;
3. Phase de construction : implémenter le schéma de données obtenu après la phase de concertation.

Lors de cet atelier nous tentons de répondre aux deux premières phases.

Ressources : <https://guides.etalab.gouv.fr/producteurs-schemas/#a-qui-s-adresse-ce-guide>

3.2. La démarche de mise en place d'un schéma : présenté par Etalab

Aujourd'hui, il est constaté une hétérogénéité des données issues des différents producteurs de données. La plateforme vise à référencer ces schémas et à accompagner leur schématisation, afin de faire monter en qualité les données publiées en open data. Des communautés se créent autour de ces schémas et les font évoluer dans le temps. Les schémas sont hébergés dans des forges

GIT. Ils sont alors récupérés par la plateforme schema.gouv.fr. De manière générale, les outils sont déjà développés et prennent la forme de tableur. Pour la randonnée ce format ne correspond pas au besoin. Il faudra sûrement travailler sur un format GeoJSON.

Les métadonnées permettent de lier une donnée à un schéma, ce qui permet de centraliser les données.

Le référencement d'un schéma s'effectue par Etalab. La validation du schéma s'affecte par un test sur un jeu de données. Les données sont ensuite publiées sur data.gouv.fr en indiquant un schéma de référence.

Si nous décidons de créer un schéma de données pour la randonnée, il faudra travailler sur un Json schema publié sur une forge Git. Ce début de schéma pourra alors être référencé.

Échanges :

- Le PNE s'interroge sur les questions suivantes :
 - Est-ce que la création d'un schéma rajoute des contraintes pour les producteurs de données ne l'appliquant pas ?
 - L'existence d'un schéma ne contraint pas les autres producteurs à y faire référence.
 - Si une donnée comprend plus de champs que ceux définis par le socle, peut-elle être rattachée à ce schéma ?
 - Oui, il y a une mention indiquant qu'elle correspond partiellement au schéma, mais si tous les champs du schéma ne sont pas remplis alors cette donnée publiée n'est pas conforme.
- Le travail pour créer un schéma est d'établir une liste des champs avec une typologie et un caractère obligatoire ou non et de mentionner dans l'idéal les spécifications relatives à la donnée elle-même.
- Un schéma peut être la somme de toutes les spécificités de chacun ou le minimum commun. Le socle commun est identifié comme obligatoire. Les spécificités sont généralement optionnelles. Il est possible d'inscrire des spécificités afin de les standardiser par anticipation.

4. La pertinence de mettre en place un Schéma pour la randonnée : recueil des avis

- IGN : la démarche est plutôt bonne mais il nécessite un gros travail sur un modèle commun. La publication d'un schéma est bien si chacun y trouve son compte.
- CD05 : pertinence du projet. Il faudrait voir pour inclure le pôle national des sports de pleine nature au projet.
- PNC : veiller à avoir un retour d'expérience et échanger avec les structures ayant travaillé sur le projet du vélo afin d'avoir une cohérence entre les deux schémas. Analyser le travail réalisé sur les voies vertes pour avoir une cohérence avec celui-ci.
- Région SUD : lors de la réflexion des champs, il serait intéressant d'intégrer la notion d'accessibilité.

5. Conclusion

- Intérêt du collectif pour la démarche

- Intérêt du collectif pour le renforcement et la standardisation des échanges de données
- Bon moment car volonté collective d'enrichir les données de randonnées
- Un socle commun existant entre les différentes BDD et outils, malgré des vocabulaires parfois un peu différents
- Contexte favorable à la mise en place d'un schéma de données de randonnées

6. Suite à mener

1. Travailler sur une V0 du schéma afin de faciliter les échanges pour un prochain atelier de validation. Le travail s'effectuera par retour d'emails ou de tickets dans le Github.
 - a) Geoffrey (Etalab – DINUM) a créé un ticket Github pour que nous puissions échanger entre les différentes réunions. <https://github.com/etalab/schema.data.gouv.fr/issues/150>
 - b) Cet outil est à utiliser de préférence afin de centraliser nos échanges.
2. Une première proposition de champs pouvant composer le schéma est déjà présente dans un tableau partagé : <https://calc.ecrins-parcnational.fr/schema-rando>
 - a) Il comprend actuellement : les champs identifiés dans le cadre du PITEM, les champs mis en correspondance avec ceux existants dans IGNrando', Apidae, DATAtourisme et OSM
3. Un document en ligne est disponible afin de compléter ce CR. Il est modifiable par tous.
 - a) <https://pad.ecrins-parcnational.fr/p/schema-rando>
4. Une V0 du schéma au format Table Schema a été créée sur le GitHub des Parcs nationaux : https://github.com/PnX-SI/schema_randonnee
5. Planifier un prochain atelier pour travailler sur la v0 du schéma et discuter de ses différents champs