



Analyse d'opportunités dans la diffusion des données de sol

Janvier - Mai 2023

Coraline SOUL, Lola DÉHO, Mathis ROTTIER, Quentin DELABARRE

Sommaire

- 01.** Introduction de l'atelier
- 02.** État des lieux dans la diffusion des données de sol
- 03.** Préconisations et mises en application
- 04.** Conclusion

01. Introduction

- 1.1** Les enjeux autour des données de sol et de leur diffusion
- 1.2** Présentation du commanditaire
- 1.3** La commande et les livrables

1.1 Les enjeux autour des données de sol et de leur diffusion

En France, et plus largement en Europe, la norme Inspire rend obligatoire la mise à disposition de données environnementales et ce librement. Pourtant, les oubliées de ce programme sont bien les données de sol. Un constat survient quant aux données de sol : elles sont aujourd'hui peu accessibles et peu diffusées malgré leur importance. Le sol, sur lequel on marche tous les jours mais auquel on ne fait pas attention, remplit diverses fonctions : support physique de nos productions agricoles mais aussi de nos aménagements, pouvoirs épurateurs et régulateurs des eaux, stockage du carbone et habitat de biodiversité... C'est donc l'interface entre divers milieux : l'air, l'eau, la roche et le vivant. Il s'agit d'un compartiment environnemental complexe dont le rôle est prépondérant pour la gestion durable des territoires, ce qui rend son analyse indispensable dans divers domaines. La connaissance précise de la qualité et des caractéristiques des sols est indispensable pour prendre des décisions éclairées en matière d'aménagement du territoire et de protection de l'environnement. Sa mise à connaissance et sa diffusion de ces données représentent alors un enjeu important.

Il existe pourtant plusieurs programmes de cartographie des sols et qui constituent des sources de données sur les sols à diverses échelles. Ces documents cartographiques associés à des bases de données peuvent constituer des outils d'aide à la décision de gestionnaires locaux pour répondre à de nombreux enjeux de territoire. Cependant, la diffusion des données de sol soulève également des questions de propriété intellectuelle et de protection des données personnelles. C'est pourquoi la mise en place d'un cadre réglementaire et éthique approprié est nécessaire pour garantir l'accès à l'information tout en préservant les droits et les intérêts des différents acteurs impliqués. Dans ce contexte, la diffusion des données de sol en France représente un défi complexe qui requiert une approche pluridisciplinaire et collaborative

1.2 Présentation du commanditaire

Partant du constat que les données de sols ne sont pas facilement accessibles et encore peu diffusées malgré leur importance, le projet Solid'R intervient. Lancé en 2022 pour 18 mois, il est porté par Bordeaux Sciences Agro (BSA) et est financé par le Ministère de l'Agriculture et de l'Alimentation et est coordonné par l'INRAE (Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement). Le projet vise à favoriser l'accès et l'utilisation des données de sol par des tiers et de rendre les producteurs de données de sol plus visibles, ainsi que structurer et harmoniser la diffusion des données RRP.

Bordeaux Sciences Agro est une école d'ingénieurs spécialisée dans les sciences agronomiques, agroalimentaires et environnementales. En tant que productrice de connaissances sur les sols, l'école est fortement engagée dans la recherche et la formation en matière de gestion des sols. Elle a développé une expertise reconnue dans l'analyse et la caractérisation des sols, ainsi que dans la modélisation de leur comportement et de leur évolution. Mais BSA n'est pas seulement productrice de connaissances mais aussi de données de sol, sujet de l'atelier. L'école est impliquée dans de nombreux projets de recherche en collaboration avec des partenaires nationaux et internationaux, visant à mieux comprendre les mécanismes qui régissent la dynamique des sols et à proposer des solutions innovantes pour préserver leur qualité et leur fertilité. Bordeaux Sciences Agro est également membre de l'Observatoire National des Sols (ONS), qui a pour mission de surveiller l'état des sols en France et de sensibiliser les acteurs concernés aux enjeux de leur préservation.

1.3 La commande et les livrables

Solid'R s'est associé en janvier 2023 aux M2 SIGAT de l'université de Rennes 2 pour chercher des opportunités autour de ces enjeux. L'objet de ce travail est donc d'étudier les opportunités qui existent pour structurer et dynamiser la diffusion de ces données. Stéphanie Jalabert, notre référente pour cet atelier, est venue à nous avec deux pistes d'entrée sur le sujet : la diffusion des métadonnées associées aux données de sols et la création de lien entre producteur et diffuseur. L'atelier a pour objectifs d'approfondir ces deux pistes pour proposer des opportunités de diffusions.

Pour répondre aux objectifs, il a été nécessaire d'établir un état des lieux de la diffusion qui prend en compte aussi les producteurs et les utilisateurs et non pas seulement les diffuseurs. Une donnée fortement demandée peut obtenir plus d'opportunités de diffusion, c'est pourquoi une partie de l'atelier s'attardera à proposer des préconisations et à explorer des réutilisations possibles et la forme qu'elles peuvent prendre.

Le présent rapport d'étude est le livrable attendu dans le cadre de cette demande. Il comprend donc une analyse des besoins, les résultats d'une enquête par questionnaire auprès des infrastructures de données géographiques, les conclusions tirées des entretiens réalisés avec certains producteurs. Enfin, il présente quelques exemples d'utilisation de la donnée de sol.

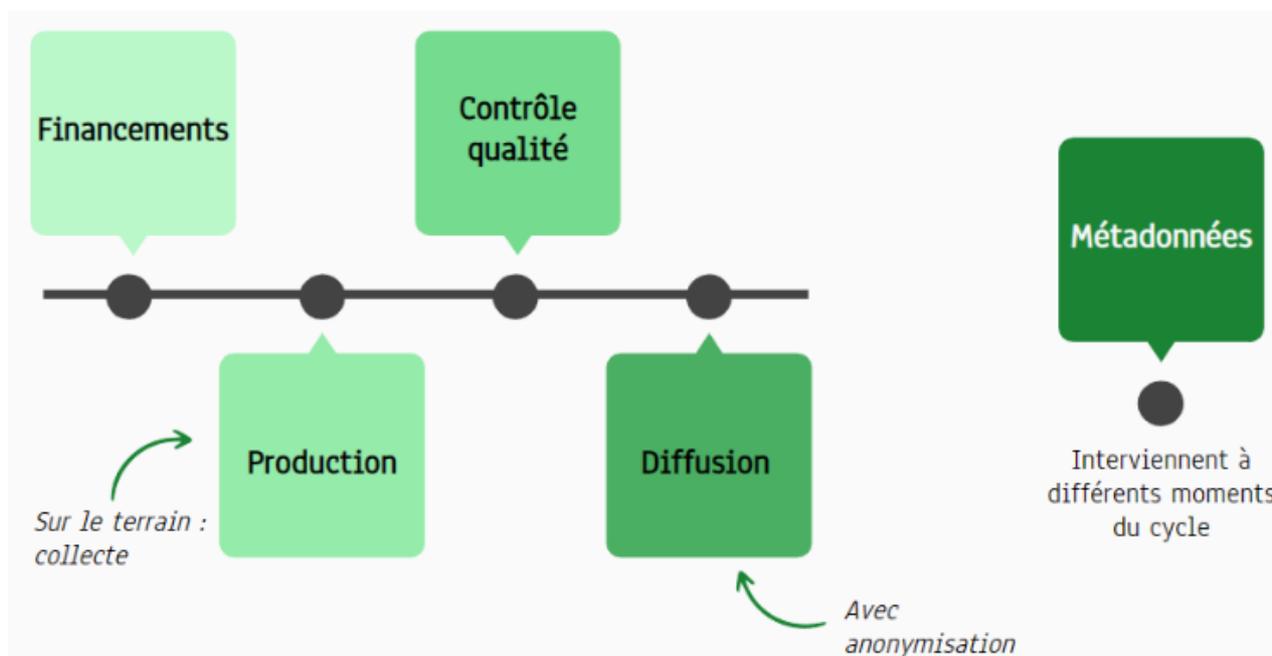
02. État des lieux dans la diffusion des données de sol

- 2.1** Les différents acteurs impliqués
- 2.2** La diversité des données et le référentiel régional pédologique
- 2.3** Le Gis Sol et les supports techniques
- 2.4** Les enquêtes Géoportail
- 2.5** Le questionnaire et les entretiens
- 2.6** États des lieux : enjeux et limites dans l'organisation de la communauté

02. État des lieux dans la diffusion des données de sol

Les conclusions auxquelles était arrivé Solid'R en lançant l'atelier étaient que si les données de sol peinent à être diffusées, c'est à cause du manque de métadonnées qui permettent de porter à connaissance une donnée ainsi que du manque de lien entre les producteurs des données et les diffuseurs. En partant d'un regard neuf et naïf, nous avons tenu à vérifier ces hypothèses lors de notre état des lieux de la situation. Progressivement, nous nous en sommes affranchis. Le diagnostic qui suit rend compte des différentes pistes que nous avons tenu à analyser.

Pour appuyer notre démarche, nous sommes partis du cycle de diffusion d'une donnée :



Cycle de diffusion d'une donnée (simplifié)

Nous avons exclu de la trame les métadonnées puisque nous considérons qu'elles peuvent intervenir à différents moments du cycle de diffusion : que ce soit lors de la production, du contrôle qualité ou lors de la diffusion. La production d'une donnée sur le terrain implique déjà d'avoir réfléchi à cette donnée, les informations qu'elle contient, son échelle... Une ébauche de métadonnées est donc déjà pensée voire rédigée comme fiche de route. Par contrôle qualité, nous entendons les processus qui permettent de s'assurer que la donnée est conforme aux usages que l'on souhaite en faire. Cela signifie que l'on peut commencer déjà à renseigner une fiche d'informations relatives à la donnée. Lors de la diffusion, les métadonnées sont omniprésentes : pour pouvoir mettre à disposition la donnée, il faut pouvoir renseigner sur ce qu'elle est. Certaines plateformes de diffusion (comme les Infrastructures de données géographiques, IDG) mettent à disposition des protocoles de saisie pour remplir les métadonnées et standardiser l'affichage d'une fiche sur le site.

Dans ce cycle, peu importe la donnée concernée, ce sont donc plusieurs acteurs qui sont impliqués : ceux qui financent, ceux qui produisent, ceux qui diffusent et ceux qui réutilisent.

2.1 Les différents acteurs impliqués

Lorsque l'on pense au sol, à qui peut-on penser ? Quelle est l'institution que l'on associe à la thématique ? Si pour la géologie, une discipline cousine de la science des sols, il vient presque naturellement de penser au BRGM, pour la pédologie la réponse est moins évidente. En effet, la communauté pédologique compte de nombreux partis et acteurs. Si l'on retraçait tout le processus qui amène à la diffusion d'une donnée de sol, nombreux sont ceux qui interviendraient, tant dans la création et la diffusion de la donnée que dans la coordination des projets et leurs financements. Qu'il s'agisse de travail terrain, d'expertise sur le sol, d'expertise sur le sol dans un domaine d'application précis (sylviculture, agriculture), de professorat, de recherche, les compétences en pédologie sont semblables à celles d'une autre discipline de sciences naturelles mais une des différences réside dans l'impact de l'organisation multiscale. On trouve des acteurs à différentes échelles : nationale, régionale, départementale et autres. On compte aux échelles nationales et régionales différentes agences et instituts qui opèrent à ces niveaux (pour les appels à projet, les commandes, les financements), ainsi que le Gis Sol, coordonnateur de programmes à l'échelle nationale. La pédologie étant une discipline dont l'objet d'étude est commun avec d'autres, on peut alors trouver l'INRAE, l'OFB, l'IGN ou encore le Ministère de la transition écologique comme parties prenantes. C'est à partir du niveau régional qu'il est plus pertinent pour les producteurs de commencer à opérer, bien qu'on les retrouve généralement à l'échelon du département. Certains acteurs s'affranchissent d'une limite administrative comme la Société du Canal de Provence, qui agit à l'échelle du canal.

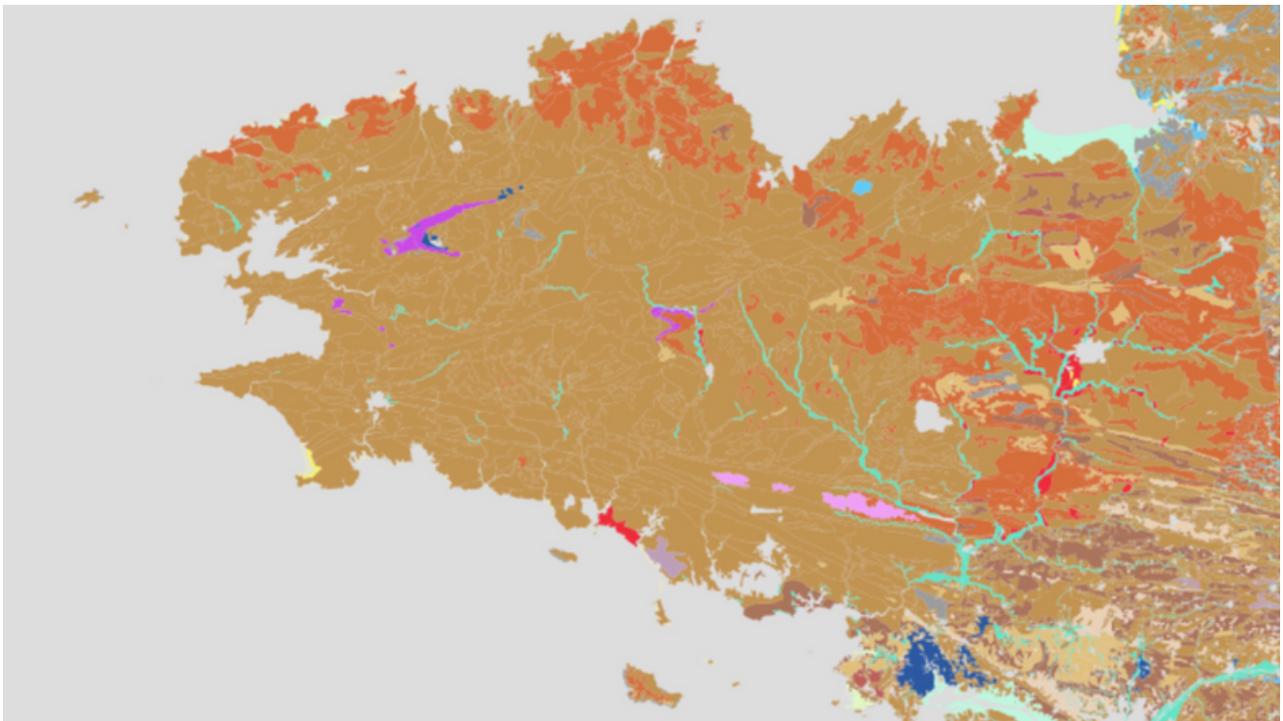
Bien que pour chaque discipline l'organisation peut être la même lorsque l'on creuse un peu le sujet, pour la pédologie cela peut poser problème. Pour un.e non expert.e, l'impossibilité d'identifier un ou des acteurs centraux peut être une question importante notamment lorsqu'elle influe sur la visibilité de la discipline auprès de potentiels consommateurs et du grand public. Le jour du lancement de l'atelier en est témoin : nous avons été très perdus face à la multitude d'acteurs, de programmes, de supports (avec chacun leurs acronymes) qui gravitent autour des données de sol. Un glossaire nous a été nécessaire avant de lancer réellement l'atelier. Toute une étude pourrait être menée sur l'origine de cette complexité mais dans ce rapport nous choisissons de prendre la situation telle qu'elle est.

Puisqu'aucun acteur central n'existe et que la visibilité de la donnée de sol et de la pédologie ne peut venir que d'une échelle plus locale, il est propre à chacun de se mettre en avant et de se faire connaître. Pourtant, la compétence pédologique n'est plus forcément mise en avant dans tous les services. Départs à la retraite, perte de la compétence pédologique dans les structures dont les domaines qui font intervenir une expertise sur les sols (agriculture, agronomie, urbanisme...), manque de compétences autres que pédologiques mais qui lui serviraient... Conséquemment, les effectifs baissent et il n'est pas toujours possible de monter de nouveaux projets et de proposer une grande diversité de services. La diversité de ces acteurs la communauté pédologique crée donc une communauté éclatée.

Un dernier acteur est crucial à étudier puisque sans lui la donnée n'a pas de sens : l'utilisateur des données de sol. Les différents profils d'utilisateur sont étudiés lors du questionnaire auprès des IDG.

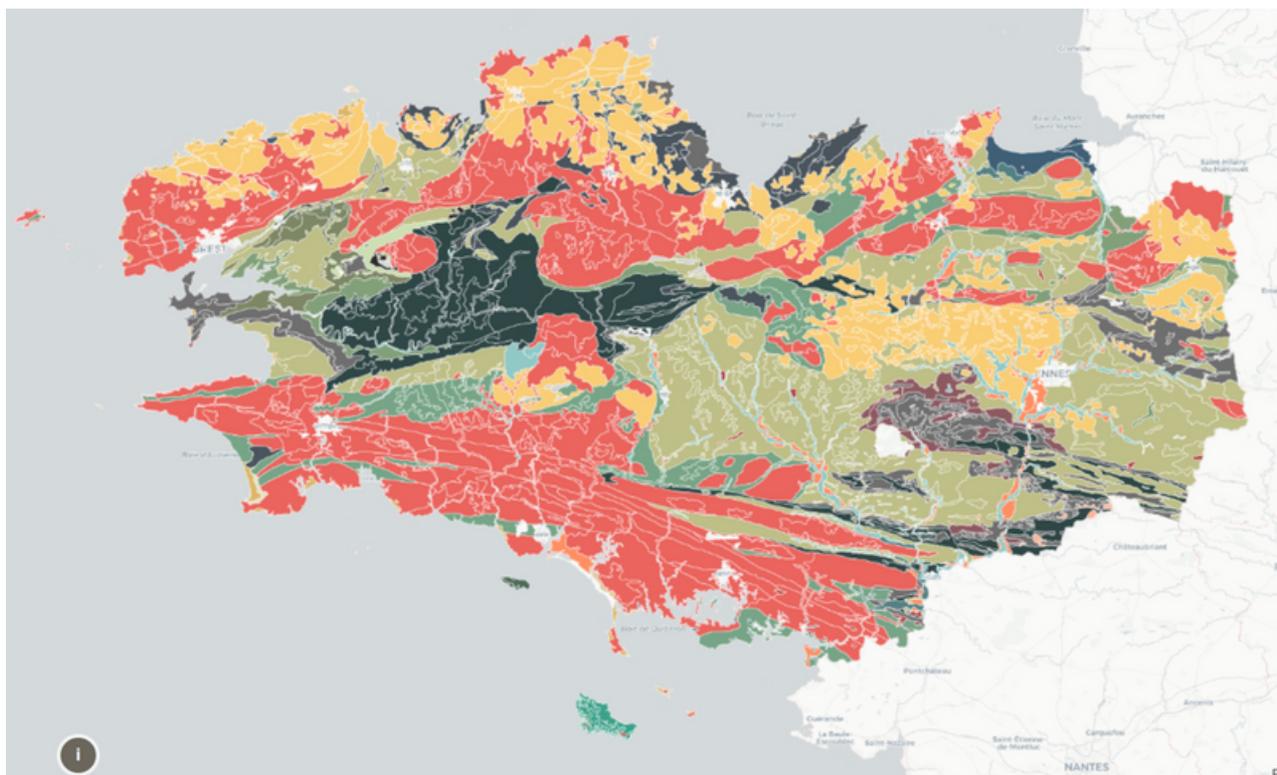
2.2 La diversité des données et le référentiel régional pédologique

La description d'un sol ne peut tenir en une phrase. Diverses informations permettent de caractériser un sol : sa profondeur, sa porosité, son acidité, sa teneur en argile, limon, sable... Pour illustrer ce propos, on peut prendre l'exemple de la carte des sols en fonction du matériau dominant. Il n'existe pas une seule façon de représenter les sols, en voici donc trois en Bretagne.



Types de sols dominants en Bretagne (Géoportail)

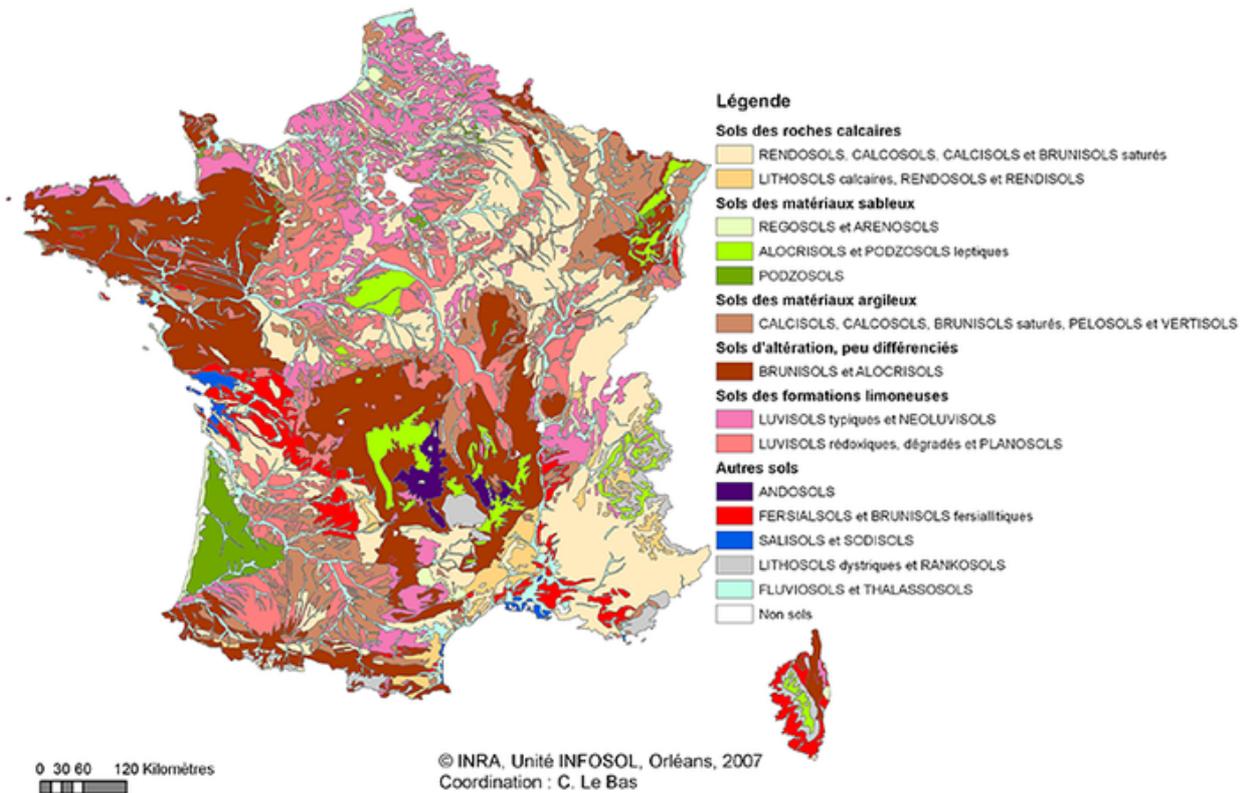
Cette première carte est issue du Géoportail et les données proviennent du Référentiel Régional Pédologique (RRP). La carte, au 1/250 000, est bientôt complétée à l'échelle de la France. Elle représente le type de sols dominants.



Matériau parental dominant en Bretagne (GéoSAS)

La deuxième est issue de Sols de Bretagne par GéoSAS et représente le matériau parental dominant. À savoir que l'on trouve sur le site de Sols de Bretagne une carte qui elle aussi représente le type de sols dominants, similaire à celle sur Géoportail.

Carte des sols dominants de France

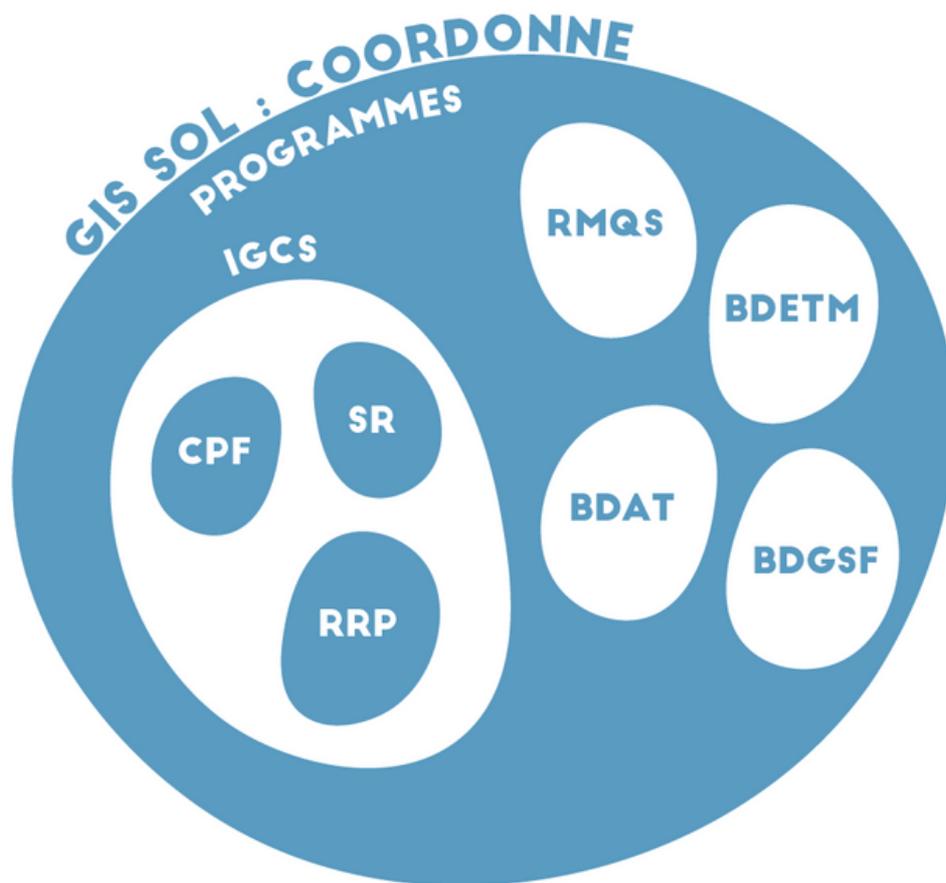


Sols dominants en France (BDGSF)

Enfin, la troisième carte est issue de la Base de données Géographiques des Sols de France (BDGSF). Avec une échelle au 1/1 000 000, elle a été produite par l'INRAE et c'est aujourd'hui la seule à couvrir entièrement le territoire métropolitain en matière de cartographie des sols. On semble reconnaître en Bretagne des formes similaires à la carte du Géoportail mais bien plus simplifiées.

Bien que différentes, ce sont toutes des cartes de sols, autant valables les unes que les autres. Ces exemples montrent que la question de la représentation des sols s'accompagne forcément de choix et regroupements de caractéristiques à faire en amont : qu'est-ce que l'on souhaite mettre en avant ? À quel niveau de détail peut-on aller (détail sémantique et échelle) ? Cette première complexité est inhérente aux sols.

Pour arriver à cerner l'entière des sols et réussir à constituer un système d'information des sols, il existe divers programmes.



Les différents programmes d'inventaire des sols

Ils ambitionnent à traduire ces caractéristiques en données tangibles utilisables et sont regroupés en deux catégories : les programmes de connaissance et les programmes de surveillance. Ces derniers, constitués du RMQS et du BDAT, n'ont pas été étudiés ou mobilisés dans le cadre de cet atelier au contraire des programmes de connaissance avec l'IGCS et la BDGSF.

Voici un tableau récapitulatif des caractéristiques de chaque programme réalisé par le RMT dans son guide.

		Base graphique	Données sémantiques	Base de données	Accessibilité	Métadonnées
IGCS	Inventaire	Polygone	Données surfaciques : description des UCS, UTS, Strates. Pour chaque strate à <i>minima</i> pierrosité, taux de chaque fraction granulométrique, teneur en matière organique, capacité d'échange cationique, pH eau, teneur en calcaire et texture (GEPPA, Aisne), couleur, taille des sables, dimension, nom, abondance des taches d'oxydo-réduction, effervescence, porosité, type de structure, contraintes, salure.		DonesolWeb3 (DW3), Outil d'export de DW3, Websol (à paraître), Cas par cas/propriétaire.	Refersols, Dictionnaire D3 en ligne, Modèle D3 en ligne <i>Applicasol</i>
		Coordonnées X, Y	Données ponctuelles : environnement du profil et données synthétiques, description d'horizons, analyses (paramètres pédologiques classiques (granulométrie, PH, cations, C, N,...)).		Cas par cas/propriétaire.	
BDGSF	Inventaire	Polygone	Description UCS et UTS Pas de données ponctuelles.	Format propre UCS/UTS	InfoSol signature d'une licence	Dictionnaire Envoi par <i>mail</i>

Synthèse des programmes d'inventaire des sols

L'IGCS est notamment très intéressant puisque c'est dans ce cadre que l'on retrouve le Référentiel Régional Pédologique (RRP) qui a été au cœur de notre étude. Le RRP est une base données qui renseigne sur les types de sols et les caractéristiques pédologiques (texture, pH, carbone, hydromorphie...). Elle permet entre autres une cartographie au 1/250 000e mais son utilisation principale reste des renseignements tirés de la base de données aux échelles de la strate et de l'UTS (la combinaison de diverses strates aux caractéristiques similaires). La base de données diffusable est constituée de 9 tables relationnelles : codage, codage_strate_qual, code, etude, l_ucs_uts, strate, strate_qual, strate_quant, ucs et uts. Ces 9 tables comptabilisent entre elles 238 champs. Le nombre d'informations disponibles est donc assez massif.

En juin 2022, sur les 101 départements, on comptait 64 RRP disponibles à la diffusion et 4 en cours de constitution (en Corse, Haute-Loire et dans le Cantal). Les premiers résultats de l'enquête Géoportail vont permettre de rentrer plus dans les détails des spécificités territoriales et des conditions d'accès du RRP. On retient néanmoins du RRP qu'il s'agit d'une base de données nationale, plus précise que la BDGSF et qui est riche d'informations.

2.3 Le Gis Sol et les supports techniques

Le Gis Sol, mentionné plus tôt comme acteur national en pédologie est un groupement d'intérêt scientifique autour des sols créé en 2001 et géré par l'unité Info&Sols de l'INRAE. Il regroupe les intérêts de divers participants comme les ministères de l'agriculture, de l'écologie, de l'INRAE, de l'IGN, de l'Ademe... L'objectif du Gis Sol est assez important puisqu'il cherche à constituer et gérer un système d'information sur les sols de France, ce qui n'existe pas encore totalement aujourd'hui. Le Gis Sol a permis de redynamiser le programme IGCS et aujourd'hui il pilote d'autres programmes de surveillance comme le RMQS, BDAT et BDETM. En plus de cela, ses missions sont entre autres de contrôler la qualité des données produites, gérer l'accès aux demandeurs et contribuer à la diffusion et valorisation des bases.

Le Gis Sol possède un site web pour se visibilité. Nous avons choisi de nous intéresser à celui-ci pour d'une part apprivoiser la thématique et d'autre part comprendre le fonctionnement du Gis Sol et la façon dont il répond à ses objectifs. Selon les recherches effectuées sur un moteur de recherche, le site apparaît plus ou moins facilement dans les résultats. Des tests ont été effectués sur Google, Ecosia et Bing et on notifie que le référencement du site varie entre les trois. Le tableau ci-dessous indique si on trouve le Gis Sol dans la première page de résultats du moteur de recherche.

Terme recherché	Google	Ecosia	Bing
Gis Sol	Oui	Oui	Oui
Réserve utile	Oui	Oui	Oui
Stockage carbone	Non	Non	Non

Erosion des sols	Non (2e page)	Non	Non
Les sols dans l'adaptation au changement climatique	Oui	Oui	Non
Artificialisation et imperméabilisation des sols	Oui	Oui	Oui
Artificialisation des sols	Non	Non	Non
Imperméabilisation des sols	Non	Non	Non
Pertes en sol	Oui	Oui	Oui
Propriétés des sols	Oui	Oui	Non (2e page)
Données de sol	Oui	Oui	Oui
Done Sol	Non	Oui	Oui
Référentiel Régional Pédologique	Oui	Oui	Oui
RRP	Non	Non	Non
RMQS	Oui	Oui	Oui
IGCS	Oui	Oui	Oui

Bien que l'on trouve quelques différences, les résultats sont les mêmes selon les moteurs de recherche. Certaines recherches étaient plus importantes que d'autres, comme "RRP", "Artificialisation des sols" et l'on remarque que ces termes ne sont pas retrouvés. Globalement, c'est lorsque l'on utilise des termes que l'on trouve dans les titres des pages du site que celui-ci apparaît. Bien que dans ce petit test ce sont majoritairement des "Oui" qui ressortent (ce qui peut signifier qu'il fait office de référence dans la communauté non experte en pédologie), cela montre aussi les limites de son référencement : on ne peut pas trop s'écarter des termes. Cela se joue au niveau des métadonnées et des mots clés d'un site web.

En faisant les recherches, on peut se rendre compte également que le site de l'INRAE apparaît régulièrement, tout comme celui de la FAO. Parfois certaines chambres d'agriculture apparaissent aussi (notamment celle du Grand-Est). Cela témoigne d'un bon travail de la part de ces autres plateformes pour porter à connaissance les sols mais tend aussi à rendre confus quelqu'un qui ne serait pas expert de la thématique. La plus grande activité dans la communauté pédologique provient néanmoins de l'INRAE et notamment de l'unité Info&Sols (qui coordonne le Gis Sol). Les cartes et rares données que l'on retrouve sur le site du Gis Sol proviennent exclusivement de cette unité. En cela résident deux problèmes : le non référencement des données produites par d'autres producteurs et l'impossibilité de réutiliser ces cartes. Le Gis Sol, seul acteur national dont la vocation est uniquement tournée vers la pédologie, peut faire figure d'acteur central. Qu'il en ait l'ambition ou non, le référencement sur Internet qu'a le site aurait l'avantage de pouvoir donner à voir également les producteurs isolés qui sont actuellement très peu mis en avant sur le site. Il faut éplucher diverses pages ou se rendre en bas de page dans la rubrique "Plan du site" pour voir apparaître le nom d'acteurs régionaux et locaux. Les données qui sont disponibles actuellement sont soit la forme de cartes, soit sous la forme de liens renvoyant à la plateforme de téléchargement entrepot.recherche.data.gouv.fr. À moins que le sujet de la carte soit en parfait accord avec l'objet d'étude d'un réutilisateur potentiel, le problème de ces cartes est qu'elles ne sont pas réemployables. Diffusées au format image, elles permettent de montrer la diversité des informations que l'on peut tirer de l'analyse des sols mais postées ainsi sur le web, elles ne peuvent pas avoir d'autre utilisation car fixes et finies.

Le Gis Sol propose également divers outils qui, pour des non connaisseurs, peuvent s'avérer utiles. Cependant, ces outils sont nombreux, divers, parfois obsolètes et peu mis en avant sur le site. Lorsque l'on rentre dans la rubrique principale "Outils", plusieurs noms d'outils s'affichent. Il faut ouvrir la page de chacun pour obtenir, après une description de l'outil, le lien vers l'application. Ces liens, colorés en marron, ne sont pas très identifiables rapidement.

Application	Adresse	Description	État
Geosol	https://webapps.gissol.fr/geosol/	Visualise en cartes différentes analyses de sols agricoles	Obsolète (aucune carte n'apparaît)
Refersols	http://refersols.gissol.fr/	Recense les études de sol réalisées en France	Fonctionnel mais pas exhaustif
Appicasol	https://annuaires.gissol.fr/appicasol	Recense les applications thématiques	Fonctionnel mais pas exhaustif
Repedo	https://annuaires.gissol.fr/repedo	Recense les organismes nationaux	Fonctionnel mais pas exhaustif

Synthèse des différents outils proposés sur le site du Gis Sol

Ce sont diverses applications qui permettent de montrer la multiplicité des études qui existent dans la thématique. Refersols en particulier offre une approche intéressante qui passe par une carte pour visualiser les espaces concernés par une étude. Toutes les applications sont disponibles (bien que l'onglet "Outil" ne renvoie que vers trois pages sur les quatre applications), ce qui pêche avec c'est qu'elles manquent de nombreuses informations actuelles et qu'elles forcent le contact vers Info&Sols pour obtenir les coordonnées des structures porteuses de la donnée.

On retient alors du site du Gis Sol qu'il agit surtout d'un site encyclopédique, contenant de nombreuses informations et explications sur le sol. Au-delà de cet aspect, ses fonctions sont peu définies puisqu'il propose diverses fonctionnalités mais qu'aucune n'est réellement approfondie.

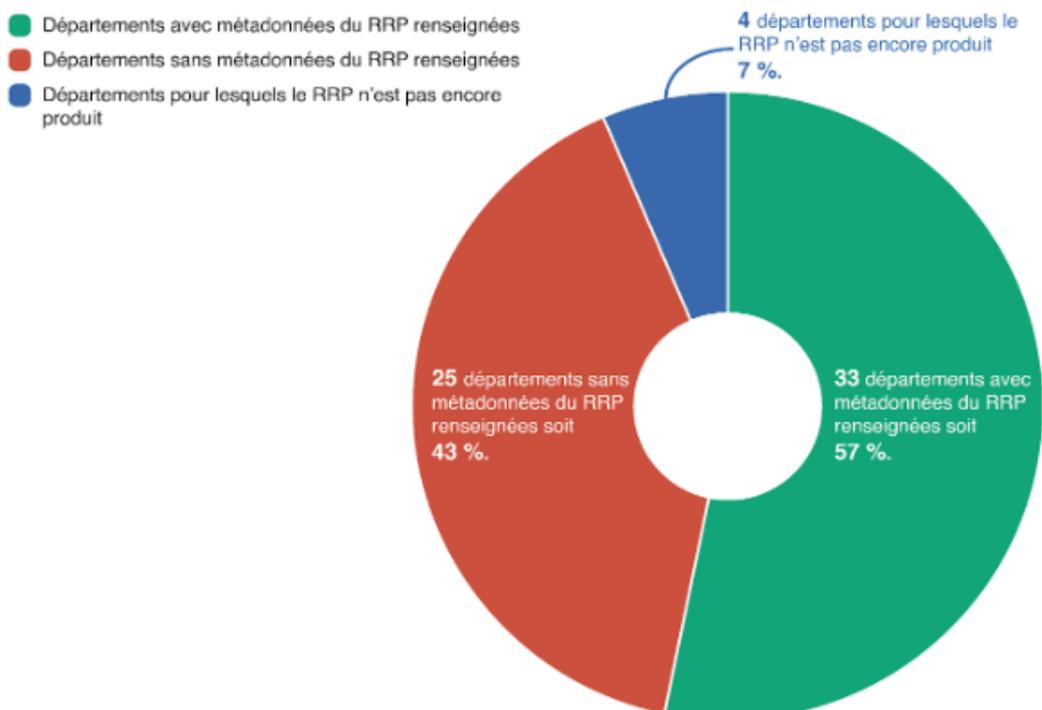
2.4 Les enquêtes Géoportail

2.4.1 Les enseignements de l'enquête

L'enquête Géoportail est une démarche entreprise par Info&sols dans le but d'appréhender l'état de la diffusion du RRP en mettant à jour les contacts sur les fiches du Géoportail et en jugeant l'avancement du remplissage des métadonnées et la modalité d'accès à la donnée. Une fiche de questions a été envoyée à chaque département. Les analyses ont donc porté sur les départements plus que sur les producteurs eux-mêmes pour donner une idée de l'état de la diffusion du RRP d'un point de vue territorial. Effectivement, certains producteurs ont été chargés de la production et de la diffusion du RRP de plusieurs départements ou régions, cela rendrait l'interprétation complexe. Nous allons proposer quelques graphiques d'interprétation. Une approche critique sera effectuée par la suite.

Renseignement des métadonnées du RRP par département :

Seuls les producteurs ayant répondu sont pris en compte.



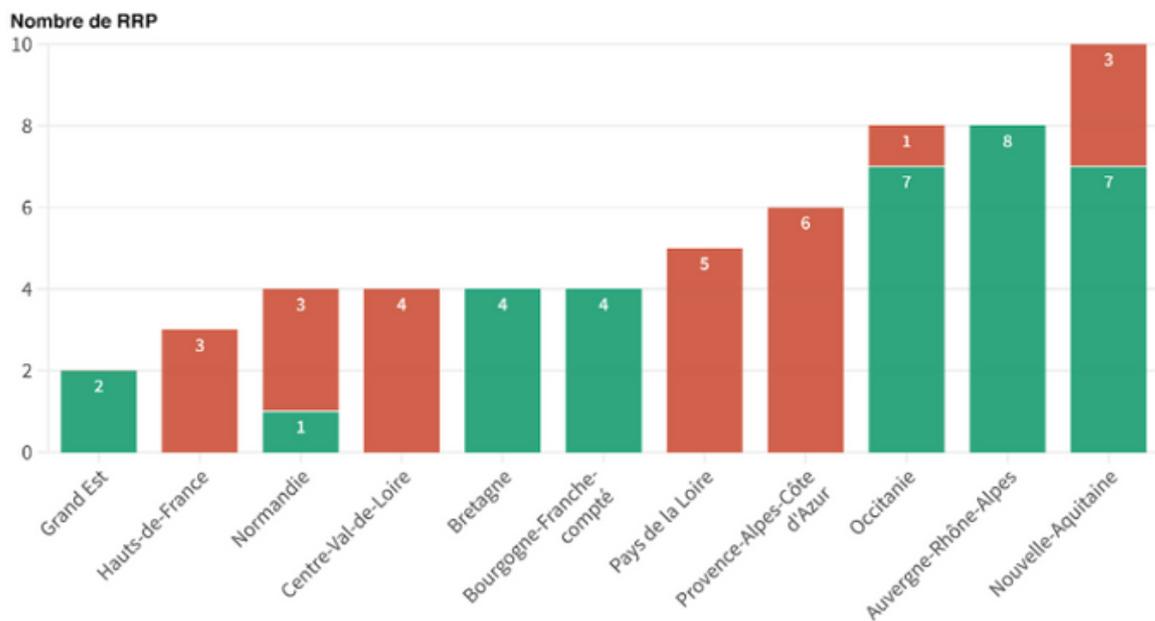
Source: SOLIDR, enquête geoportail / RRP métadonnées
Traitements des résultats : le 22/04/2023 par M2 SIGAT 2022/2023

En sachant qu'au moment de l'analyse nous n'avions pas à notre disposition toutes les réponses, on peut voir sur ce premier graphique que le remplissage des métadonnées du RRP est un chantier encore d'actualité avec 43 % des départements qui n'en disposent pas pour leur RRP. Les métadonnées sont pourtant un maillon essentiel à la diffusion de la donnée, on a donc un potentiel frein à celle-ci qui est identifié ici.

Renseignement des métadonnées du RRP par région :

Seuls les producteurs ayant répondu sont pris en compte.

- Départements avec métadonnées du RRP renseignées
- Départements sans métadonnées du RRP renseignées



Source: SOLIDR, enquête geoportail / RRP métadonnées

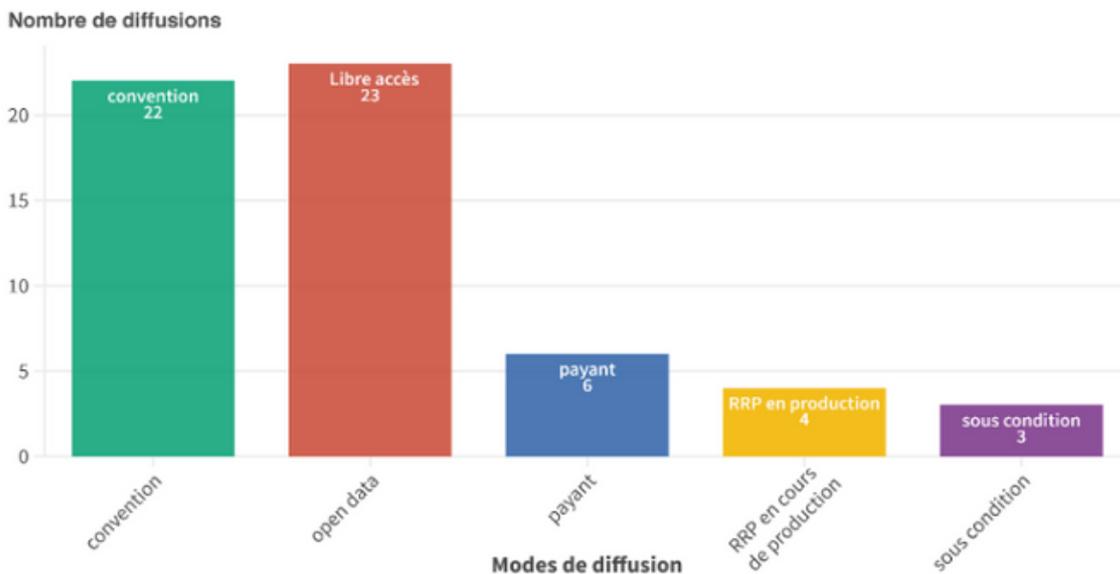
Traitements des résultats : le 22/04/2023 par M2 SIGAT 2022/2023

Ici on s'interroge plus particulièrement sur le renseignement des métadonnées par région. Ce graphique montre la part des départements pour lesquels les métadonnées sont renseignées par rapport à la part des départements pour lesquels elles ne sont pas renseignées, le tout regroupé par région. Ici, il ne faut pas tenir compte du nombre brut de RRP par région puisque certaines sont plus étendues que d'autres et contiennent par conséquent plus de départements et de RRP.

De plus, comme évoqué plus haut, tous les départements n'ont pas répondu à l'enquête, ce qui crée un biais. Certaines régions sont représentées avec moins de départements qu'elles n'en ont dans la réalité. Par exemple, la région Grand-Est est composée en réalité de dix départements contre seulement deux affichés ici.

Les différents modes de diffusion du RRP :

Seuls les producteurs ayant répondu sont pris en compte.



Source: SOLIDR, enquête geoportail / RRP métadonnées

Traitements des résultats : le 22/04/2023 par M2 SIGAT 2022/2023

Ici, on s'intéresse plus particulièrement aux méthodes de diffusion employées par les producteurs du RRP qui ont répondu à l'enquête. On remarque deux modes de diffusions principaux que sont la diffusion par convention et la diffusion en libre accès (souvent via des portails open data).

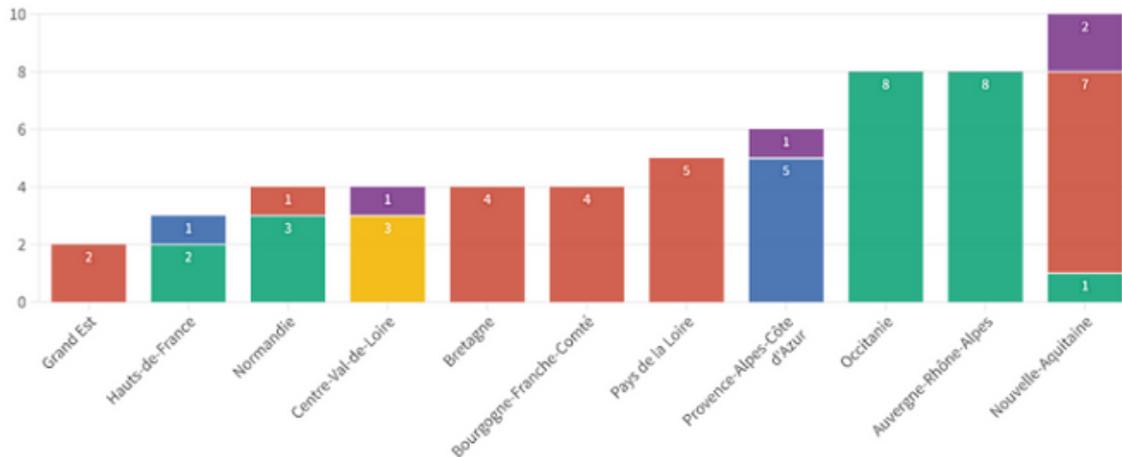
Il est assez intéressant de noter ces deux modes qui sont en concurrence et dont aucun n'émerge particulièrement de leur fréquence d'usage. Ce sont deux modes relativement opposés dans leur approche. Le libre accès correspond à une volonté de mettre à disposition de manière libre, gratuite et sans contrainte la donnée alors que la diffusion par convention témoigne d'une volonté de contrôle sur la diffusion et les réutilisations du RRP. Il y a donc deux paradigmes différents qui ressortent ici de cette étude en matière de diffusion, avec deux volontés de diffusion différentes.

Les différents modes de diffusion du RRP par région :

Seuls les producteurs ayant répondu sont pris en compte.

Conventions Libre accès Payant Sous conditions RRP en cours de production

Nombre de diffusions



Source: SOLIDR, enquête geoportail / RRP méladonnées

Traitements des résultats : le 22/04/2023 par M2 SIGAT 2022/2023

On s'interroge ici sur la proportion de chaque type de diffusion par département, le tout regroupé par région. On observe ici de manière assez claire les régions pour lesquelles un producteur unique applique la même méthode de diffusion sur l'ensemble des RRP départementaux, là où pour d'autres régions des dynamiques plus complexes se décrivent.

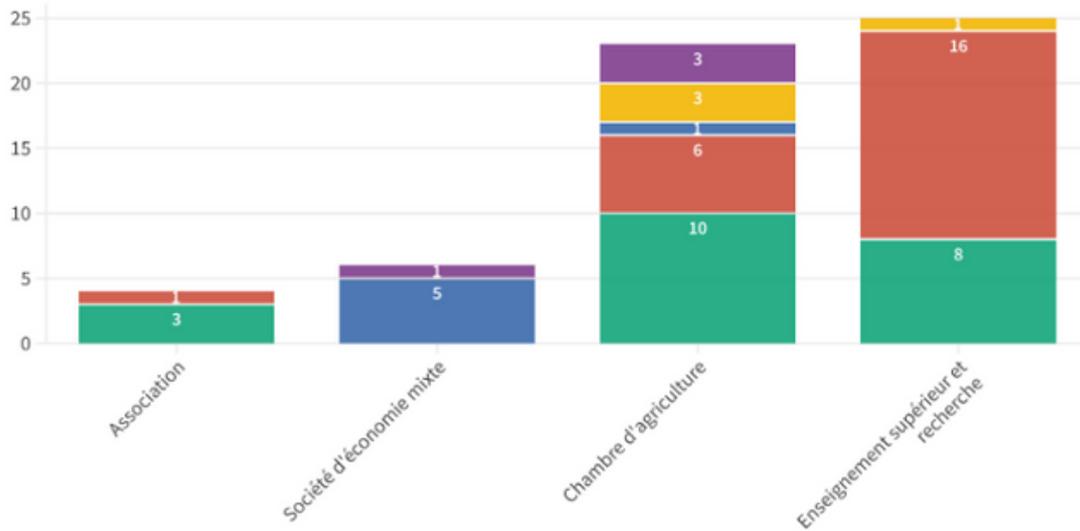
Il convient cependant d'alerter ici sur un biais de compréhension qui ressort des résultats de cette question, il semblerait que la différence entre certaines modalités de la question n'ait pas été comprise de manière uniforme par les producteurs. Nous pensons qu'il y a eu une confusion entre la modalité "libre accès" et "convention" pouvant amener des producteurs à se tromper lorsqu'ils répondent. Une donnée sous convention a pu être interprétée comme donnée libre, notamment en Bretagne.

Les différents modes de diffusion du RRP par type de producteur :

Seuls les producteurs ayant répondu sont pris en compte.

● Conventions ● Libre accès ● Payant ● Sous conditions ● RRP en cours de production

Nombre de diffusions



Source: SOLIDR, enquête geoportail / RRP métadonnées

Traitements des résultats : le 22/04/2023 par M2 SIGAT 2022/2023

Le but ici est de donner à voir les différences dans les pratiques de diffusion en fonction des producteurs. On voit ici la part de chaque modalité de diffusion par type de producteur. Des dynamiques différentes en fonction des types de producteurs ressortent ici avec des chambres d'agriculture qui privilégient une diffusion par convention, là où les organismes d'enseignement et de recherche ont davantage une culture de l'open data et diffusent plus largement la donnée en libre accès. Cependant, il faudrait vérifier cas par cas que tous ont bien la même notion de "libre accès". En effet, on observe encore une fois le biais de compréhension évoqué précédemment. Il est particulièrement frappant pour les chambres d'agriculture. On trouve énormément de modalités de diffusion dans leurs pratiques, nous sommes donc en droit de nous demander si certains producteurs ne se sont pas trompés dans leurs réponses en raison d'une incompréhension des différentes modalités contenues dans le questionnaire.

2.4.2 Les limites et conclusions de l'enquête

Comme évoqué précédemment, les limites identifiées de ce questionnaire se manifestent principalement par une non exhaustivité des réponses et un biais de compréhension des questions qui influence les résultats et interprétations du questionnaire. L'interprétation des différentes modalités par les producteurs semble assez floue et est source d'erreurs. Certains producteurs confondent la diffusion en libre accès et la diffusion par convention. Cela est compréhensible du fait que la convention permette en théorie de diffuser à qui le demande, cependant cela ne rejoint pas la diffusion du libre accès qui est de diffuser de telle sorte qu'un utilisateur puisse se procurer la donnée de manière gratuite, libre et sans contrainte, la convention ne correspond pas à ces deux dernières notions (libre et sans contrainte).

Cela nous empêche de prendre ses résultats au pied de la lettre. La prudence est de mise quant à l'interprétation des résultats. Cette limite, alors identifiée, peut être corrigée dans un futur questionnaire dont les résultats deviendront alors beaucoup plus fiables que ceux-ci.

Pourtant, l'enquête Géoportail nous renseigne sur des pratiques globales que l'on peut observer, à savoir des différences dans la saisie des métadonnées entre les départements et les régions, ce qui conduit à une forme d'inégalité et d'hétérogénéité territoriale. Des différences résident aussi dans les pratiques de diffusion différentes du RRP ce qui conduit là encore à une hétérogénéité et une inégalité d'accès entre utilisateurs (certains pouvant se procurer la donnée gratuitement, d'autres devant payer). Les pratiques sont différentes aussi par type de producteurs, avec certains qui ont une appétence plus marquée pour le libre accès quand d'autres préfèrent le contrôle offert par une convention.

Tous ces résultats et conclusions doivent être faites avec de grandes précautions en gardant un tête que les biais de compréhension des interrogés quant à certaines modalités de questions fragilisent la robustesse des résultats.

2.5 Le questionnaire et les entretiens

Le cycle de vie d'une donnée comporte deux grandes étapes au cœur de notre problématique : la production et la diffusion. À notre connaissance, aucune étude à large échelle ne permet de qualifier l'état des lieux de la diffusion (là est d'ailleurs l'une des missions de Solid'R). Les questionnaires Géoportail permettent une première entrée en la matière en quantifiant l'état de la diffusion mais ils comportent trop de biais pour être utilisables. Comme il sera précisé dans les entretiens, certaines structures mènent individuellement leurs études pour comprendre comment fonctionne la diffusion des données de sol mais pas systématiquement. C'est pourquoi nous avons choisi de partir sur des formats de questionnaires et d'entretiens pour obtenir des connaissances préliminaires et plus détaillées.

Le cycle de vie de la donnée se termine aussi par une partie d'utilisation à laquelle nous nous sommes attelés de façon indirecte. Au commencement de l'atelier, il nous fallait cerner qui pouvaient être les cibles et utilisateurs des données de sol. De ce fait, le questionnaire et les entretiens n'ont pas été dirigés vers des utilisateurs précis mais ont plutôt été enrichis de questions au sujet des demandes formulées, des cas d'usage connus ainsi que des retours d'utilisateurs reçus.

2.5.1 Le questionnaire auprès des IDG

Nous avons choisi de découper le travail en fonction des différents acteurs qui interviennent dans la publication d'une donnée de sol. En tant qu'étudiants géomaticiens, nous connaissons la donnée d'abord par les plateformes et interfaces qui les mettent à disposition, c'est pourquoi l'objectif du questionnaire était de nous introduire à la thématique par le biais des infrastructures de données géographiques (IDG), les diffuseurs principaux de données dont nous sommes plus familiers. Envoyé le 24 février 2023 par Stéphanie Jalabert à des IDG territoriales (18 régionales, 10 départementales), le questionnaire a cherché à déterminer quelle est la place et le point de vue des diffuseurs dans la complexité de la diffusion des données de sol. Il s'est articulé autour de 4 grandes thématiques :

- mieux comprendre les pratiques et outils de diffusion des données,
- mieux comprendre les besoins thématiques par utilisateurs,
- situer la connaissance autour de la donnée sur les sols,
- mieux comprendre la structuration des données et métadonnées.

Dans le but d'en extraire les opportunités qui existent pour structurer et dynamiser la diffusion, les questions étaient orientées de sorte à pouvoir identifier :

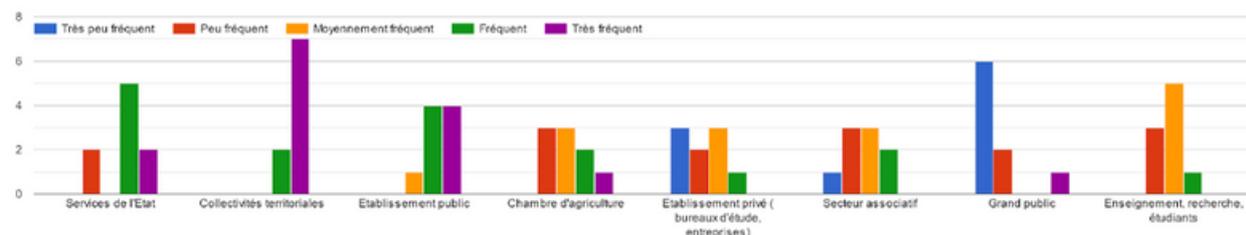
- le fonctionnement d'une IDG pour pouvoir proposer des solutions qui s'inscrivent déjà dans leurs pratiques,
- le lien entre IDG et producteurs pour comprendre comment se fait le dialogue entre eux ou pourquoi il ne se fait pas,
- les demandes et cas d'usage de la donnée de sol ainsi que les freins à la diffusion du point de vue des IDG.

Le format du questionnaire a été privilégié pour récupérer ces informations puisqu'il permettait d'interroger un nombre important d'acteurs sur un temps réduit. Outre la contrainte de temps, l'avantage du questionnaire est de permettre de quantifier les réponses reçues et de produire des analyses statistiques. Cependant, l'échantillon d'interrogés (et de répondants) reste faible, ce qui nécessite d'interpréter les résultats avec prudence. Au 27 mars 2023, le questionnaire envoyé aux IDG a reçu 9 réponses sur 28 (soit un taux de réponse de 32,14% donc un peu moins d'un tiers). Nous vous présentons ici une synthèse des réponses reçues, dont les enjeux principaux seront étudiés par la suite.

Présentation des utilisateurs

Les profils d'utilisateurs principaux des IDG sont assez variés. La différence entre chacun est la fréquence de consultation du site. Généralement, elles sont consultées et contactées le plus fréquemment par des collectivités territoriales, des établissements publics et services de l'État. Le grand public reste le moins coutumier des IDG. Entre les deux, on retrouve les établissements de recherche et d'enseignement et les étudiants, qui les consultent de temps en temps. Les IDG restent donc surtout consultées par un public déjà familier de la donnée et ayant besoin de la mobiliser.

Quels sont les profils les plus fréquents des utilisateurs de votre plateforme ?



Les collectivités territoriales consultent les IDG en majorité pour des thématiques tournées vers l'urbanisme et l'imagerie et l'occupation du sol. De manière générale, ces deux domaines, en plus de l'énergie et des réseaux ainsi que l'environnement et du climat, sont les plus demandés.

La diffusion d'une donnée se caractérise par plusieurs grandes étapes : la collecte, le contrôle qualité et la mise aux normes, la rédaction des métadonnées et la mise à disposition en choisissant la modalité d'accès : licence ouverte, OdBL, convention, redirection vers le producteur... On trouve aussi quelques variations : retravail des données si nécessaire et intégration en base de données (RGD Savoie Mont Blanc) ou valorisation des données via les réseaux sociaux (GIP ATGeRi).

Connaissance de la donnée sol et diffusion au sein de la structure

Les données de sol sont connues des IDG. Aucune n'a indiqué ne pas en connaître. Parmi les plus connues, on compte la carte des sols dominants sur Géoportail, qui est une variation du RRP (100% des réponses), et le RRP en lui-même (66%). Ce sont donc moins des données aux plus grandes échelles (comme les Typterres, le RMQS). Les données liées à des thématiques spécifiques sont moins portées à connaissance.



La moitié des IDG interrogées diffuse des données de sol (5 en diffusent et 4 non). Parmi celles qui en diffusent, seules 3 IDG diffusent les RPP (soit 1/3 des répondants totaux) alors même qu'ils connaissent tous la donnée. Les autres données diffusées sont des données locales, propres aux projets de chaque territoire. Les RRP sont disponibles en open data ou sur demande tandis qu'on trouve les autres données en open data. Cependant, comme pour les enquêtes Géoportail, il faudrait rester vigilant sur ce chaque répondant entend par "open data"

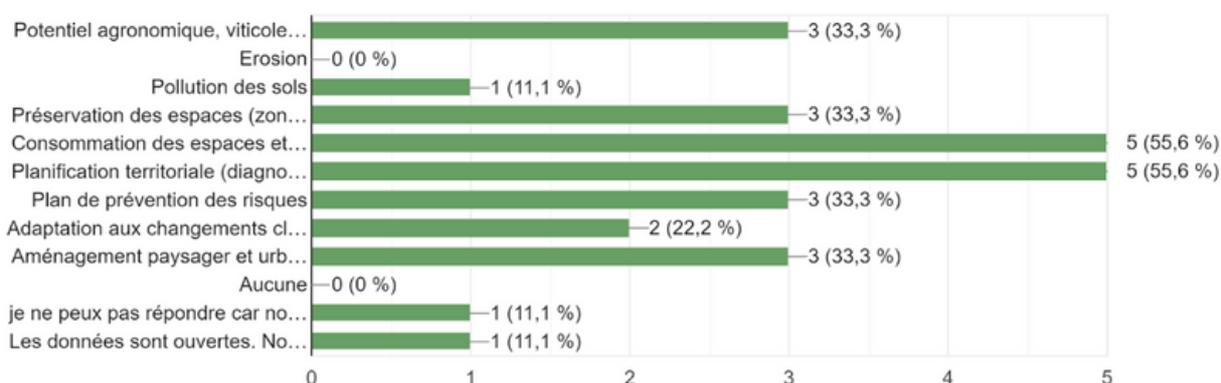
Les IDG qui ne diffusent aucune donnée de sol pointent toutes le fait qu'elles ne sont pas en contact avec les producteurs. D'autres raisons sont évoquées mais il n'est pas possible d'établir de généralités (chacune a été mentionnée une seule fois).

Les demandes liées aux données de sol

On retrouve plusieurs thématiques d'intérêt qui amènent à travailler avec des données de sol. Les thématiques qui ressortent le plus auprès des IDG sont les suivantes : planification territoriale et artificialisation des sols. On peut mettre en lien cette information avec le fait que les collectivités territoriales sont les acteurs qui fréquentent le plus leurs sites et que les demandeurs sont majoritairement des chargés d'étude en planification ou en bureau d'étude.

Quelles sont les thématiques pour lesquelles on vous sollicite le plus ?

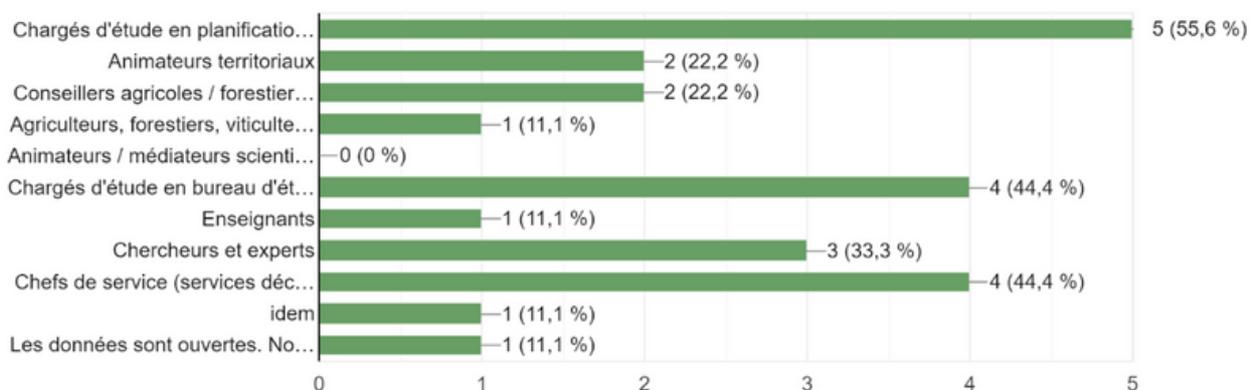
9 réponses



La fréquence de demande en données de sol varie énormément d'une structure à l'autre. 44% des répondants estiment recevoir des demandes tous les mois ou moins contre 55% à une fréquence qui se rapproche du semestre ou de l'année.

Qui sont les demandeurs ?

9 réponses

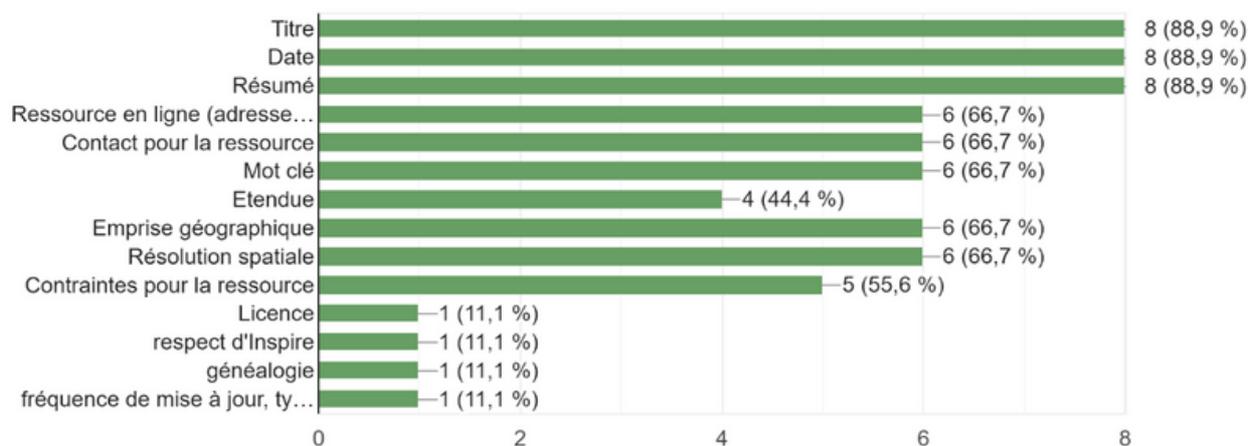


Les métadonnées

7 IDG sur 9 s'accordent à dire que des métadonnées sont obligatoires pour publier une donnée sur leur plateforme. Parmi les informations principales à renseigner, on retrouve le titre, la date de production et un résumé de la donnée. Certaines IDG proposent également un protocole simplifié de saisie de la métadonnée mais la majorité le permet.

Quelles sont les informations minimales à renseigner concernant les métadonnées ?

9 réponses



8 IDG sur 9 ont accepté d'être re-contactées dans le cadre d'études ultérieures sur les données de sol. Nous transmettrons les résultats du questionnaire à Solid'R pour que le projet puisse prendre le relais.

Critiques du questionnaire

Les résultats du questionnaire sont cependant critiquables. Les questions posées ont été faites en tout début d'atelier, lorsque nous connaissions encore que très peu le sujet. De ce fait, certaines sont assez généralistes et manquent de précision une fois l'étude avancée. Il aurait été pertinent de demander si les IDG organisaient des journées thématiques (notamment autour de la pédologie), si elles connaissaient les producteurs de leur territoire, si elles étaient en capacité de participer au financement de données...

De plus, le terme de "modalité diffusion" n'a pas été compris de la même manière pour tous. La définition que nous entendions n'a pas été précisée en amont. Par exemple, si l'on compare OpenIG, ayant indiqué diffuser des données de sol, avec GéoCentre qui indique ne pas le faire, tous deux ont mis à disposition une fiche de RRP. GéoCentre renvoie vers des liens à DoneSol et indique les contacts du producteur, OpenIG demande de s'authentifier à l'IDG pour accéder à la base DoneSol du RRP (en format Access et Excel) et impose en plus de faire la demande d'accès pour la couche graphique.

Également, l'IDG Agence Régionale du Numérique et de l'intelligence artificielle (ARNia) indique diffuser des données de sol mais celles-ci sont introuvables (tant sur le site web sur lequel trouver ça que dans la recherche par mot clé sur <https://ideo.ternum-bfc.fr/>, on ne trouve que le shp des UCS).

Du point de vue de l'échantillonnage, le questionnaire a été créé pour être représentatif avec une attention particulière à interroger des IDG régionales et départementales. Mais ce qui pose problème ici ce n'est pas le nombre d'interrogés mais bien le nombre de répondants. Nous comptabilisons 9 réponses au total (avec une IDG ayant répondu deux fois) sur 28 IDG démarchées. Ainsi seulement un tiers des IDG interrogées ont répondu à notre questionnaire ce qui, de fait, le rend très peu représentatif et très peu exploitable.

Il est donc nécessaire de vous inviter à la plus grande prudence quant à ces résultats qui sont selon nous peu représentatifs de la réalité.

Bien que certaines tendances et profils puissent être tirés de ce questionnaire, les résultats ne sont pas fiables, à terme nous recommandons de rééditer un questionnaire similaire et de la diffuser aux IDG de telle sorte à ce qu'elles répondent en plus grand nombre pour disposer de résultats fiables et exploitables. Il pourrait aussi être intéressant de mobiliser des acteurs comme l'Afigéo pour donner plus de visibilité à ce genre de questionnaire.

2.5.2 Les entretiens avec les producteurs

Les producteurs de données de sol sont identifiés mais les enjeux liés à leur diversité de statut et aux différences de moyens dont ils disposent pour la thématique pédologique restent flous. Pour mieux comprendre les enjeux liés à leur statut nous avons décidé de mener des entretiens avec plusieurs producteurs sélectionnés à partir des résultats des enquêtes Géoportail, en fonction de leurs pratiques et de leur type de structure.

Pour commencer, nous avons distingué les chambres d'agriculture et les établissements de recherche et d'enseignement (où la production de données de sol est faite par des laboratoires qui leur sont liés), qui sont les deux principaux types de producteurs. Il est à noter quelques exceptions : la société du canal de Provence qui est la seule entreprise privée à produire des données (que nous avons choisi d'interroger), et le producteur de données en Corse qui a un statut différent dû au statut de collectivité territoriale unique (que nous avons mis de côté). Les établissements publics ont relativement la même politique de diffusion des données entre eux, nous avons donc sélectionné un seul organisme de ce type. Concernant les chambres d'agriculture, nous avons dissocié deux chambres d'agriculture : l'une, en Somme, ne diffuse les données que sous convention, où l'extraction est payante et où il n'y a pas de valorisations accessibles de la donnée (chambre d'agriculture de la Somme), l'autre, en Alsace, diffuse en open data la donnée ainsi que des indicateurs issus du RRP.

Les entretiens avaient pour objectif de :

- **mieux connaître les pratiques des producteurs (production pure, valorisation, transmission des données),**
- **mieux comprendre les demandes qu'ils reçoivent (de qui et pour quelles raisons),**
- **mieux comprendre les modalités de diffusion,**
- **mieux comprendre la structuration des données et des métadonnées.**

Pour les chambres d'agriculture, nous les avons choisies de sorte à ce que l'échantillon représente au mieux les différentes pratiques mais en réalité la chambre d'agriculture d'Alsace semble être assez différente puisque le partenariat avec l'IDG Data Grand Est semble être très solide et propose plusieurs données sur l'ensemble de la région.

Ces entretiens nous ont apporté beaucoup d'éléments de réponses concernant la production des données mais les pratiques de diffusion très hétérogènes restaient un frein à l'établissement de propositions qui pouvaient être adaptées en fonction du producteur. Pour aller plus loin nous avons donc décidé de continuer les entretiens avec d'autres acteurs.

2.5.2 Autres

L'acteur prépondérant qui ressort lorsqu'il est question de données de sol est le Gis Sol. Au cours de l'étude, il semblait de plus en plus nécessaire de mieux comprendre son rôle qui nous semblait flou même après l'étude préliminaire et les entretiens avec les producteurs. Un entretien avec Bertrand Laroche, coordinateur du Gis Sol, a donc été réalisé. L'entretien avait pour objectif de mieux cerner les enjeux et contraintes du groupement d'intérêt scientifique. Les enjeux de budget et d'organisation entre producteurs ont été mentionnés. Aborder ces aspects nous permettait aussi de réfléchir à des opportunités plus en adéquation avec les réalités administratives et politiques de l'ensemble du domaine.

Pour aller plus loin, nous nous sommes également entretenus avec Emilie Tournadre, chargée de mission partage de la donnée et open data à la chambre d'agriculture France. Nous voulions mieux savoir où en étaient les différentes chambres d'agriculture quant à la diffusion des données mais aussi mieux comprendre les différences entre elles et à quoi elles étaient dûes.

2.6 État des lieux : enjeux et limites dans l'organisation de la communauté

2.6.1 Les producteurs

Les conclusions tirées de cette première étape de l'étude et principalement des entretiens sont multiples d'une part concernant la donnée elle-même et son format et d'autre part concernant l'organisation générale du domaine pédologique.

Dans un premier temps, les producteurs s'accordent à dire que la donnée de sol est très complexe et ne permet pas en l'état la prise en main par des utilisateurs non thématiques. Au-delà de la question de la diffusion, c'est bien la réutilisation qui semble être problématique pour les producteurs, soucieux que la donnée soit mal utilisée. Les RRP, donnée sur laquelle nous avons choisi d'insister lors des entretiens, semblent prometteurs en matière de diffusion et réutilisation puisqu'ils offrent un large éventail de possibilités de traitements et puisqu'ils sont disponibles sur une bonne partie du territoire métropolitain. Cependant, plusieurs contraintes ont été clairement identifiées lors de cet état des lieux. Son format (Donesol), très peu accessible, a été immédiatement identifié comme un problème. En effet, une base de données de tables relationnelles n'est pas simple à utiliser et requiert des compétences techniques relativement importantes. Un professionnel non géomaticien peut avoir du mal à s'en servir. Ce qui est apparu dans un second temps, c'est la richesse sémantique dont il dispose. Les champs dont il dispose sont très complexes et contiennent des variables difficilement compréhensibles pour un géomaticien qui ne maîtrise pas les données de sol. La cible d'utilisateurs du RRP en format Donesol est donc très réduite.

Malgré cet inconvénient majeur, les potentielles utilisations restaient multiples (notamment autour de la réserve utile en eau, du potentiel agronomique et du stockage carbone) mais ici aussi les discussions avec les producteurs ont mis en avant un point essentiel. D'après les personnes interrogées de la Société du canal de Provence et de la Chambre d'agriculture de la Somme, elles-mêmes utilisent très peu le RRP. En effet, cette donnée n'est pas assez précise pour répondre à leurs usages. La majeure partie des demandes qu'ils reçoivent nécessite des données à une échelle beaucoup plus petite, notamment pour répondre à des enjeux liés à l'urbanisme ou à l'agriculture. Les usages potentiels du RRP ne sont pas forcément identifiés par leurs producteurs, ou en tout cas par une partie d'entre eux. Nous pouvons lier cela à la différence de statut, bien que la Chambre d'agriculture du Grand-Est réemploie ces données et propose des indicateurs travaillés pour la diffusion, la valorisation et la pédagogie. En tant que chambre d'agriculture, elle semble pourtant être l'une des seules à le faire, si ce n'est la seule. Les raisons pointées du doigt dans les autres structures de ce type sont le manque de temps et des priorités différentes de celles d'un établissement de recherche par exemple. Diffuser une donnée ne fait pas partie des missions principales d'une chambre d'agriculture.

De manière plus générale, cette distinction est identifiée comme étant la conséquence de la différence dans le fonctionnement économique. Les pédologues des structures qui font payer la donnée de sol et qui ne la diffusent pas, pointent du doigt leur nécessité de s'autofinancer. Le RRP a été produit sur des fonds publics mais aujourd'hui il n'est plus financé, c'est pourquoi ils ne développent pas forcément de canaux de diffusion qui représenteraient un coût pour eux. Dans la pratique, un demandeur doit s'adresser à eux pour avoir le RRP (comme pour toutes les données de sol). L'échange fait l'objet d'une convention et certains producteurs font payer un coût d'extraction de la base de données. Ceux que nous avons interrogés mettent en avant le fait que ce n'est pas la donnée qu'ils font payer mais bien la main d'œuvre pour le temps administratif (convention) et technique (extraction) de réponse à la demande. Mettre en place un système de diffusion gratuit qui n'implique pas de temps de travail semblerait toutefois leur convenir. Leur système économique repose avant tout sur les données à grande échelle qu'ils produisent dans le cas d'étude spécifique, la mise à disposition du RRP n'implique donc pas de perte pour eux.

2.6.1 Les diffuseurs

Doit-on penser les diffuseurs comme les leviers d'action pour aider à la diffusion des données de sol ? C'est sur ce point que Solid'R misait, en plus des métadonnées. Au vu des questionnaires passés, on retient des forces et des faiblesses. Que l'IDG diffuse ou non de la donnée de sol actuellement, c'est une interface de diffusion intéressante et fiable qui peut permettre de toucher un public assez large et varié (mais déjà connaisseur de leur existence). Il existe chez les utilisateurs de données un réflexe de se rendre sur le portail d'une IDG pour pouvoir trouver des données correspondant à leurs thématiques. Les IDG agissent donc comme vitrine des données, que la donnée soit téléchargeable en un clic ou que l'on ait accès seulement à une fiche de métadonnées et des contacts (comme GéoCentre le fait pour le RRP du Loiret).

Cependant, et cela Solid'R l'avait bien identifié, il n'y a pas de lien entre producteurs et diffuseurs. Si toutes les IDG répondantes connaissaient au moins la carte des sols sur le Géoportail, il n'est pas dit qu'elles ont davantage de connaissances autour des données de sol et qu'elles soient en mesure de contacter les producteurs pour obtenir de la donnée. Cela peut venir de deux raisons : une multiplicité de producteurs à contacter sur son territoire (on compte jusqu'à 5 chambres d'agriculture départementales différentes et l'unité Info&Sols comme producteurs présents pour le GIP RECIA - Do.TeRR GéoCentre soit l'IDG du Centre Val de Loire) ainsi qu'une réticence à libérer la donnée pour certains producteurs.

L'intérêt autour de la donnée de sol est inégal. La demande autour de la donnée de sol existe, elle est parfois plus intense dans certains territoires comme dans les Hauts-de-France. Certaines IDG, d'ailleurs souvent en lien avec des producteurs, sont plus en avance concernant la diffusion des données de sol. C'est pourquoi, des journées thématiques peuvent même être organisées comme avec le CRIGE PACA et la Société du Canal de Provence.

De manière générale, les IDG n'ont pas forcément les connaissances autour des données de sol, sur les productions qui sont faites en interne, sur celles qui sont faisables, sur l'interprétation de la donnée. À moins de demandes spécifiques faites par les utilisateurs, ce ne sont pas elles qui pourront prendre les rennes pour diffuser de la donnée de sol en demandant aux producteurs de leur transmettre des données en particulier. En revanche, les IDG peuvent prendre le rôle de vitrine de la donnée de sol en mettant des fiches de données sur leurs sites. Il existe sur certaines plateformes de protocoles simplifiés de saisie des métadonnées qui permettent aux partenaires producteurs de pouvoir diffuser leurs données simplement. La difficulté réside surtout en l'identification du ou des producteurs de données pédologiques sur leur territoire ainsi que sur la modalité d'accès de la donnée. Pour qu'une IDG puisse pleinement servir de tremplin pour la diffusion d'une donnée de sol, il est nécessaire que cette donnée soit entièrement libre, c'est-à-dire qu'un ré-utilisateur puisse se procurer la donnée de manière libre, gratuite et sans contrainte (telle qu'une convention).

03. Préconisations et mises en application

- 3.1** Préconisations structurelles
- 3.2** Préconisations de valorisation
- 3.3** Une méthode automatisée : le script python
- 3.4** Un référentiel administratif pour mieux se situer

3.1 Préconisations structurelles

L'une des entraves à la bonne diffusion des données identifiées en amont est l'hétérogénéité des producteurs. Celle-ci induit des difficultés d'identification des différents acteurs et peut aboutir vers une impasse lorsque les coordonnées ne sont pas mises à jour. Actuellement, l'accès à l'information concernant l'identité d'un producteur sur son territoire se fait de trois façons : soit via l'annuaire du Gis Sol Repedo qui n'est pas très bien référencé ni exhaustif et qui nécessite une redirection de la part de l'unité Info&Sols, soit via les métadonnées disponibles sur des IDG comme Géoportail, où les coordonnées sont inscrites, ou alors par le bouche à oreille. Ces méthodes ne suffisent pas et sont un frein à la diffusion de ces données. De manière générale dans les autres disciplines, la plupart des grandes bases de données standardisées à l'échelle nationale ont un acteur qui leur est associé. Les données géologiques ont le BRGM, la BD Topo provient de l'IGN, les données démographiques viennent de l'INSEE et les données de sols pourraient avoir le Gis Sol. Le Gis Sol n'est ici pas considéré comme un producteur de données mais il est le seul acteur structuré à l'échelle nationale qui fait le lien entre les différents producteurs et peut les fédérer sous un intérêt commun. Cet acteur pourrait porter la diffusion des données à l'échelle nationale. Pour l'instant, les différents producteurs cherchent des moyens de diffusion "individuels". Nous identifions deux stratégies différentes pour se porter à connaissance : soit au travers des IDG ou du bouche à oreille (ce qui implique donc d'avoir la culture de l'IDG ou de préalablement connaître les interlocuteurs du territoire), soit au travers d'applications développées par leur propre structure et ne permettant donc de ne voir que leurs propres données. Ce dernier mode de diffusion crée donc encore plus d'hétérogénéité et ne permet pas de faire connaître les données de sols au-delà de l'échelle très locale.

Centraliser les données sur une seule interface portée par le Gis Sol pourrait ainsi permettre une meilleure identification des données de sols mais aussi de mieux visualiser le Gis Sol lui-même et donc toutes les ressources qu'il met à disposition. De plus, avoir une porte d'accès commune aux données assure d'avoir moins de distinctions entre les territoires. Le RRP est une donnée standardisée, la centralisation pourrait donc permettre le traitement de tous les RRP en même temps ce qui permettrait d'avoir les mêmes couches à disposition où que l'utilisateur soit. Aujourd'hui, il est possible de produire certaines analyses à partir du RRP, mais leur accès est inégal. Dans le cas du réservoir utile, certains producteurs (comme c'est le cas en Bretagne) l'ont déjà calculé et mis en open data, tandis que cette donnée va être payante via certaines chambres d'agriculture sur d'autres régions ou départements.

La mise à disposition du RRP en open data ou d'indicateurs calculés à partir de cette base peut aussi se défendre d'un point de vue économique. Le fonctionnement actuel de convention entraîne beaucoup de charge administrative et s'avère peu rentable d'après les entretiens menés. Avoir une interface d'accès direct permettrait donc de ne pas avoir à faire ces démarches et donc d'économiser du temps. Par extension, gagner en visibilité via le RRP promet aussi d'avoir plus de demandes. En effet, comme évoqué précédemment le RRP n'est pas assez précis pour beaucoup d'usage à l'échelle locale. Le besoin pour des EPCI ou des communes d'avoir des études pédologiques plus précises notamment pour des thématiques d'urbanisme a été remonté plusieurs fois au cours de notre étude. Ces études ne peuvent pas publier leur donnée puisqu'elles sont trop précises pour respecter le cadre de la RGPD. Ainsi les acteurs ayant un système économique nécessitant un retour financier pourraient capitaliser sur les demandes spécifiques qui seront plus nombreuses si les données de sol gagnent en visibilité grâce à la diffusion et la valorisation du RRP.

A court terme, il est possible qu'il y ait des demandes qui ne trouvent pas de réponse notamment sur les territoires où il n'y a plus de compétences pédologiques ou sur ceux où les moyens humains ne sont pas suffisants pour répondre à ce besoin. Sur ce point, nous pouvons espérer que si la demande augmente suffisamment, le domaine pédologique soit plus porté et que les demandes incitent à former plus de professionnels.

3.2 Préconisations de valorisation

3.2.1 Partager la connaissance

La conclusion directe de l'état des lieux de notre étude est l'hétérogénéité chez les producteurs dans les valorisations de données de sol et la nécessité de diffuser le RRP comme produit d'appel pour visibiliser l'ensemble des données de sol. Il serait donc intéressant de s'appuyer sur les possibilités du RRP pour montrer aux ré-utilisateurs potentiels le genre d'analyses et d'études que l'on peut faire grâce aux données de sol. Cependant, concernant le RRP, il a aussi clairement été identifié sa difficile prise en main. Il semble donc important de proposer des valorisations qui seraient susceptibles de s'adresser à un public plus large. La principale contrainte concernant la mise en place de cette proposition réside encore dans l'hétérogénéité des producteurs. En effet, certains d'entre eux ont déjà créé des indicateurs ou proposé des cartes plus accessibles, plus intelligibles ou des applications en ligne. Il s'agit principalement des pédologues issus des laboratoires de recherche qui ont un rôle de transmission de la connaissance scientifique. La plupart des autres acteurs ne proposent pas ce genre de contenu et valorisent peu leur RRP.

Actuellement, l'une des seules données simplifiées accessibles au téléchargement est la carte du type de sol diffusé sur Géoportail. L'usage de cette donnée est très limité mais l'intérêt de sa diffusion est de visibiliser le producteur. En cliquant sur l'UCS (Unité Cartographique de Sol), le nom de la structure qui a produit localement le RRP et ses coordonnées sont accessibles. Certaines cartes nationales accessibles en .jpg décrivant des indicateurs issus du RRP sont également disponibles sur le site du Gis Sol. Cependant, les possibilités de réemplois de ces données sont extrêmement limitées de par leur contenu (le type de sol est un nom mais ne donne pas beaucoup de caractéristiques sur le sol) ou de par leur format (une carte en format image n'est pas utilisable et intéresse globalement peu de monde en l'état). Néanmoins dans le cas du Géoportail, le questionnaire que nous avons mené a révélé que le RRP était connu par ce biais. Cela témoigne donc de l'avantage d'avoir rassemblé sur une même plateforme les données de quasiment toute la France.

L'objectif, dans cette lignée, serait donc d'arriver à une base commune entre tous les producteurs, où chacun aurait un fond commun de valorisations à proposer sur leurs territoires. Le format Donesol est alors un avantage considérable pour permettre d'aller plus loin dans les valorisations du RRP et pour les proposer sur l'ensemble des territoires où la base de données existe. Étant interopérable et standardisé, il permet d'envisager d'automatiser les traitements qui transforment cette donnée brute en informations plus digestes pour les usagers de la données. Pour cela, il est nécessaire pour les différents producteurs (et acteurs de manière générale) de travailler entre eux, de se partager leurs avancées en matière de valorisations. Pour s'assurer que tous puissent en bénéficier, ce partage de connaissance peut se faire par le biais de processus automatisés et de solutions open source qui permettent. Les processus automatisés (via des scripts et chaînes de traitement) assurent que même les moins connaisseurs en matière de technique peuvent bénéficier des valorisations du RRP à l'aide de quelques clics.

Quant au site du Gis Sol, celui-ci peut aussi s'inspirer de ses propres réalisations et user plus fréquemment de la carte sur ses applications qui recensent études, projets et acteurs, afin de spatialiser le contenu et de simplifier les affichages. L'utilisation de ces applications peut s'avérer plus intuitif, visuel et moderne.

3.2.2 Simplifier la donnée

Simplifier la donnée pour la rendre plus intelligible et utilisable est un enjeu important qui doit être développé en amont de la diffusion. Il existe une multitude de moyens d'y parvenir, dont la plupart ont déjà été essayés par certains producteurs. Dans un premier temps, le projet Solid'R a déjà produit des guides d'utilisations, ceux-ci ont pour objectif de permettre à des utilisateurs non experts de prendre la main sur la donnée. Ces guides fonctionnent comme des tutos, ils peuvent être utiles mais dans la pratique même avec ces guides il a été difficile pour nous de prendre en main le RRP brut. Ces guides ont l'avantage et l'inconvénient d'être très précis. Avoir besoin d'un guide pour utiliser une donnée peut aussi être source de réticence et donc décourager à utiliser la donnée. Le public qui peut réellement s'emparer de la donnée reste donc assez restreint.

Ensuite, faute de rendre plus lisible la donnée, il est possible de la transformer pour la rendre plus facilement utilisable. Le format de base de donnée relationnelle peut faire peur, c'est pourquoi le Gis Sol propose un format nommé Datamart qui condense les valeurs sémantiques dans une seule table. D'apparence plus simple, la richesse sémantique ne diminue pas pour autant et l'utilisateur se retrouve, certes, avec une seule table mais qui contient plus de 300 colonnes. Ce format nous semblait porteur initialement mais nous sommes forcés de constater que lui aussi ne simplifie pas l'utilisation du RRP.

Les deux autres moyens qui ont été développés sont quelque peu liés entre eux. Il s'agit de proposer des interfaces de géovisualisations qui permettent de voir et spatialiser des indicateurs. Ces indicateurs sont, à eux seuls, des moyens envisageables de simplifier la donnée et qui semblent tout à fait pertinents pour la diffusion. En effet, avoir une couche de données géographique avec une entité spatiale à laquelle sont attachées des valeurs est un format de diffusion beaucoup plus courant et exploité par de nombreux professionnels. Ce format permet également de garder la main sur la méthodologie employée pour calculer l'indicateur qui peut ensuite être retraité.

Cette question des indicateurs implique donc de devoir retravailler la donnée pour en extraire une autre information. Passer d'une donnée sous forme de tables à une donnée sous forme géographique induit aussi une perte de précision qu'il faut forcément accepter, sur laquelle nous reviendrons dans l'exemple concret.

Encore faut-il cependant que ces indicateurs thématiques répondent à des besoins identifiés, sans quoi, ils risqueraient d'être très peu utilisés par la suite. C'est pourquoi, au cours de notre étude nous avons identifié trois grands secteurs d'intérêt auxquels le RRP peut apporter des réponses en le simplifiant grâce à des indicateurs :

- la réserve utile,
- le stockage carbone,
- les indicateurs liés au potentiel agronomique.

3.3 Une méthode automatisée : le script python

Pour démontrer la pertinence que pourraient avoir les pré-calculs d'indicateurs et l'automatisation de traitements, nous avons choisi de réaliser un script python qui permet de calculer la réserve utile à partir d'un RRP. La méthodologie est reproductible puisqu'elle repose sur un RRP dont le format est standardisé. Le RRP à notre disposition a été celui des Pyrénées Atlantiques (64). Il est sous sa forme diffusable : toutes les tables du RRP ne sont pas présentes.

Nous avons choisi d'utiliser la plateforme Google Colab pour réaliser le script et le transmettre. Il s'agit d'une plateforme simple d'accès, collaborative (mais pas en temps réel) et qui héberge les données le temps de la session. Elle permet d'avoir à sa disposition une grande capacité de stockage et une rapidité de traitement affranchies des caractéristiques de l'ordinateur sur lequel on ferait tourner ce script. À noter que travailler sur un environnement local exigera probablement d'autres façons d'initialiser l'environnement de travail et d'importer les données (sur Colab, il suffit d'importer les données sur le panneau latéral à droite et de lancer les lignes de script qui lisent les fichiers).

Le script ne doit pas être séparé du fichier zip qui l'accompagne. Ce-dernier se compose des données qu'il faut importer, contenant divers fichiers :

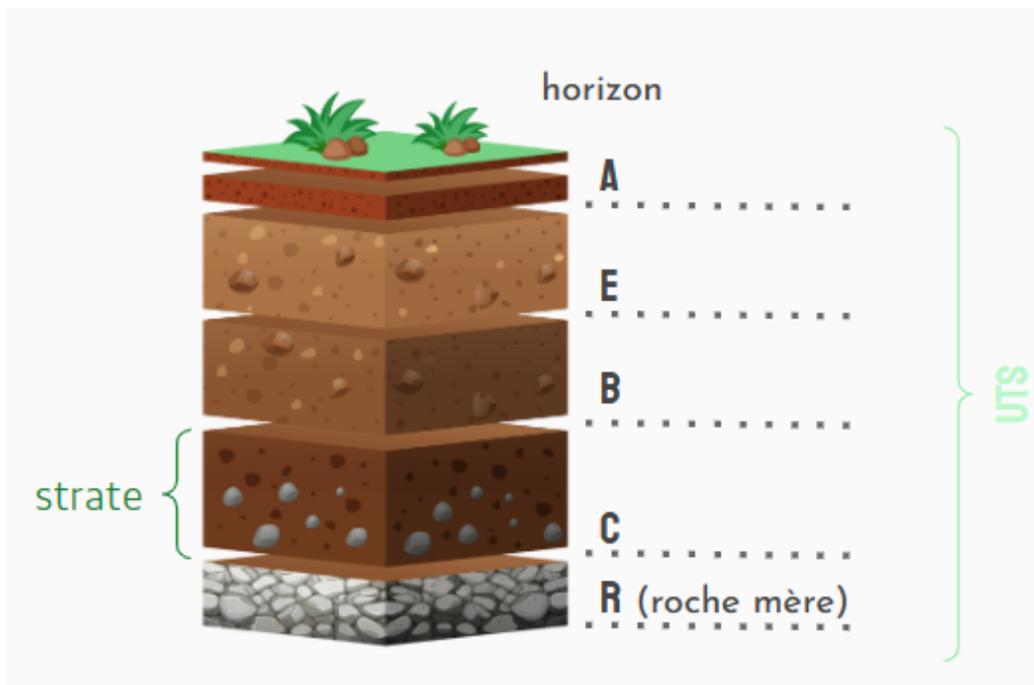
- le RRP au format mdb Access ;
- les tables au format csv de ce même RRP (I_UCS_UTS, UCS, strate, strate_qual, strate_quant) ;
- deux fichiers csv référentiels (reference_TF, reference_EG) ;
- la couche géographique des UCS au format geopackage.

Il est également important de noter que les méthodologies utilisées dans ce script relèvent de choix et décisions. Les résultats qui découlent du script sont critiquables, c'est pourquoi nous envisageons ce script avant tout comme une base à transmettre, modifiable et adaptable.

3.3.1 Calculer la réserve utile par strate et UTS

Après avoir défini la zone d'étude, c'est-à-dire le ou les RRP que l'on souhaite utiliser, le calcul se décompose en deux grandes étapes :

1. Estimer la RU par strate
2. Estimer la RU par UTS



Composition d'une unité topographique de sol (UTS)

Le graphique ci-dessus est une représentation schématique de ce qu'est une strate et de sa relation avec l'UTS. Dans le RRP, les données sont fournies par strate, ce qui signifiera donc pour nous de devoir passer de l'un à l'autre.

Dans un premier temps, il faut revenir sur les différentes étapes du script. Une réserve utile dépend de différents paramètres : la texture du sol (TF), la nature des éléments grossiers (EG) et la profondeur des sols. Le calcul de la réserve utile par strate que nous avons suivi est donc celui-ci :

$$RU_{TF} + RU_{EG}$$

RU = Réserve utile

TF = Terres fines

EG = Éléments grossiers

Il s'agit en réalité de la toute dernière étape, qui nécessite d'être décomposée.

La réserve utile pour les terres fines

Il existe différentes textures de sol et différentes natures d'éléments grossiers. C'est pourquoi, il est nécessaire de rentrer dans le détail des différentes typologies. Pour cela, ce calcul suit les méthodologies de Tétégan et al. (2011) et de Bruand et al. (2004) qui permettent d'avoir des indications plus précises sur les capacités à retenir l'eau qu'ont les différents éléments qui composent le sol. En ce qui concerne la texture des sols, nous avons suivi ce calcul qui s'appuie sur le triangle de Jamagne et la texture Aisne :

$$\begin{aligned} \text{RU TF} = \\ \text{RU TF (mm/cm)} \\ \times \% \text{ TF} / 100 \\ \times e \end{aligned}$$

e = épaisseur de la strate en cm

Dans le script qui a été constitué, le calcul de la RU TF entraîne avec lui le recalcul de la Texture Aisne pour toutes les strates du (ou des) RRP que l'on introduit dans le processus. Cela se joue au moment du calcul de la RU TF en mm/cm. Par le biais d'une requête faite de conditions, nous mobilisons les valeurs en taux de limon, de sable et d'argile pour arriver à un nouveau champ de texture. C'est à ce moment qu'intervient l'un des référentiels (reference_TF) qui, à partir de la méthode de Bruand et al. joint des valeurs de RU TF mm/cm en fonction de si la strate est de profondeur A ou E, B, C (ce que l'on trouve grâce au le champ rp_2008_nom).

Par la suite, le pourcentage de terre fine dans la strate ainsi que son épaisseur sont des informations que l'on trouve dans les tables du RRP et qui seront mobilisées pour le calcul final de RU TF.

La réserve utile pour les éléments grossiers

Nous sommes au tiers du travail puisqu'il reste la nature des éléments grossiers à retrouver pour déterminer sa capacité de restitution d'eau. En ce qui concerne la nature des éléments grossiers (classe lithologique), nous avons suivi ce calcul :

$$\begin{aligned} \text{RU EG} = \\ (\text{a3} \times \ln(\text{MVAeg}) + \text{b3}) \\ \times \text{MVAeg} \\ \times \% \text{ EG} / 100 \\ \times e \text{ (dm)} \end{aligned}$$

MVAeg = masse volumique apparente des EG
e = épaisseur de la strate en dm

Le script permet de retrouver les propriétés de rétention d'eau par classe lithologique grâce au fichier reference_EG, constitué à partir de la méthodologie de Tétégan et al. Il associe à chaque strate trouvant sa correspondance dans la liste des éléments lithologiques (actuellement uniquement constitué de silex, craie, chaille et différents types de calcaire) les valeurs a3, b3 et MVAeg. Ceux ne trouvant pas de correspondance sont considérés comme sans RU EG et seule la RU TF comptera.

Par la suite, le pourcentage d'éléments grossiers dans la strate ainsi que son épaisseur sont des informations que l'on trouve dans les tables du RRP et qui seront mobilisées pour le calcul final de RU EG.

Après avoir effectué ces deux parties de calcul, nous obtenons donc une RU TF par strate et une RU EG par strate également. Il nous est donc possible d'effectuer le calcul initial qui associe les types de RU.

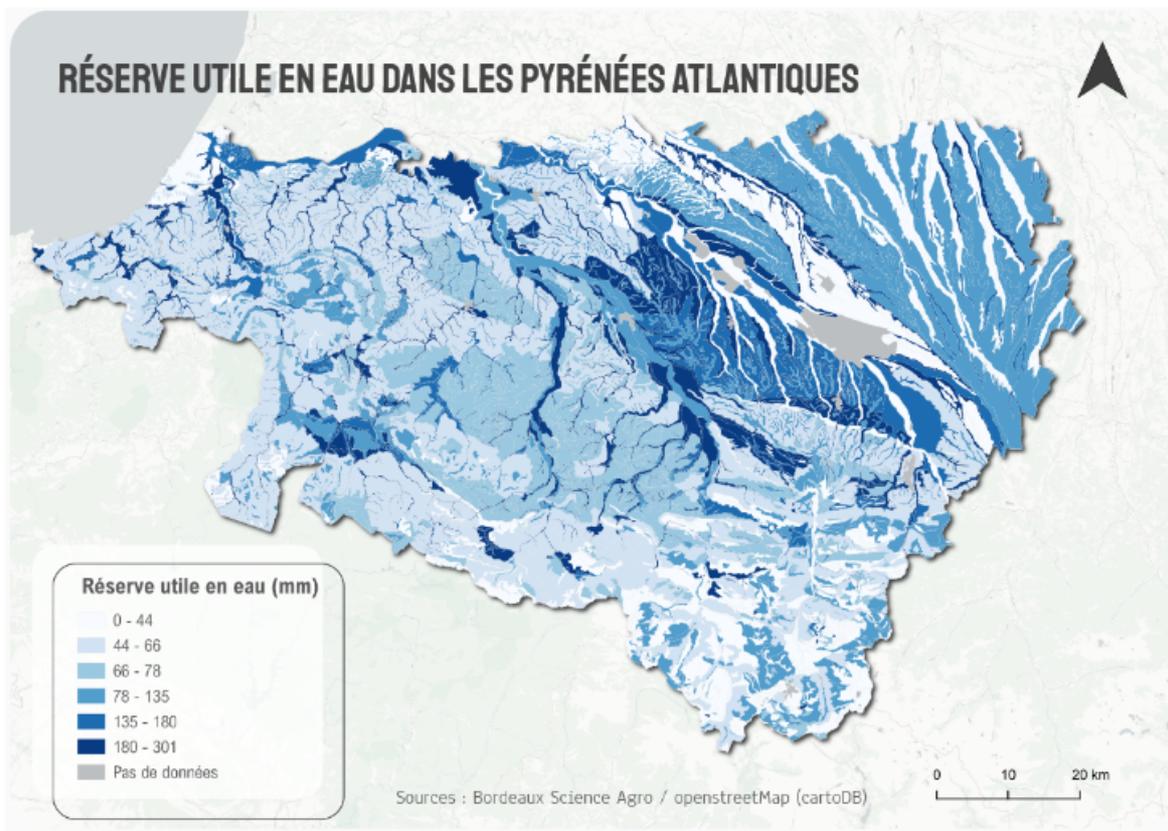
L'assemblage final

$$\text{RU UTS} = \text{RU strate 11} + \text{RU strate 12} + \dots + \text{RU strate 1n}$$

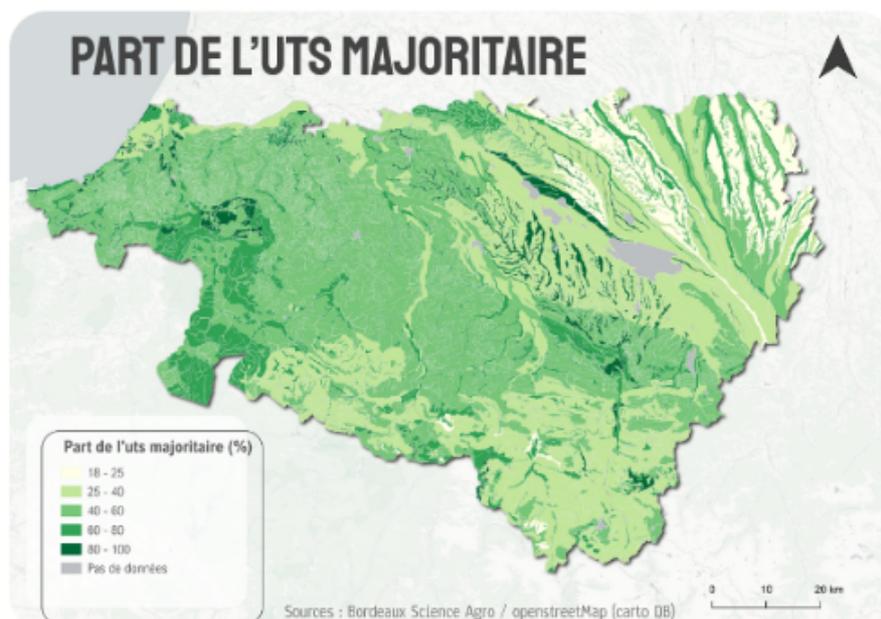
Nous obtenons alors une réserve utile par UTS.

3.3.2 Passage du UTS à UCS

Après les étapes précédentes nous avons une couche qui contient des informations sémantiques par UTS (unité typologique de sol). Ces individus correspondent à des types de sol mais ne sont pas cartographiables puisqu'ils ne correspondent pas à une entité géographique. L'unité cartographique de sol (UCS) est l'unité de référence pour la visualisation. Chaque UCS est composée de plusieurs UTS. Une agrégation de l'information est donc indispensable pour avoir une couche de donnée géoréférencé contenant dans notre cas la valeur de réserve utile en eau. Plusieurs méthodes existent pour se faire. En pédologie l'une des plus répandues est celle consistant à attribuer la valeur de réserve utile de l'uts majoritaire à l'ensemble de l'UCS.



Cette carte est complétée par une autre carte représentant la variable du pourcentage de l'UTS majoritaire par UCS pour savoir quelle est la probabilité que la réserve utile soit celle représentée sur la carte. A noter également que notre script ne calcule pas cette confiance par le biais de l'écart type mais simplement la proportion de l'UTS majoritaire. Cette carte de probabilité n'est donc pas celle couramment utilisée par les pédologues. Cette partie du traitement pourra donc être complétée en fonction des attentes de l'utilisateur.

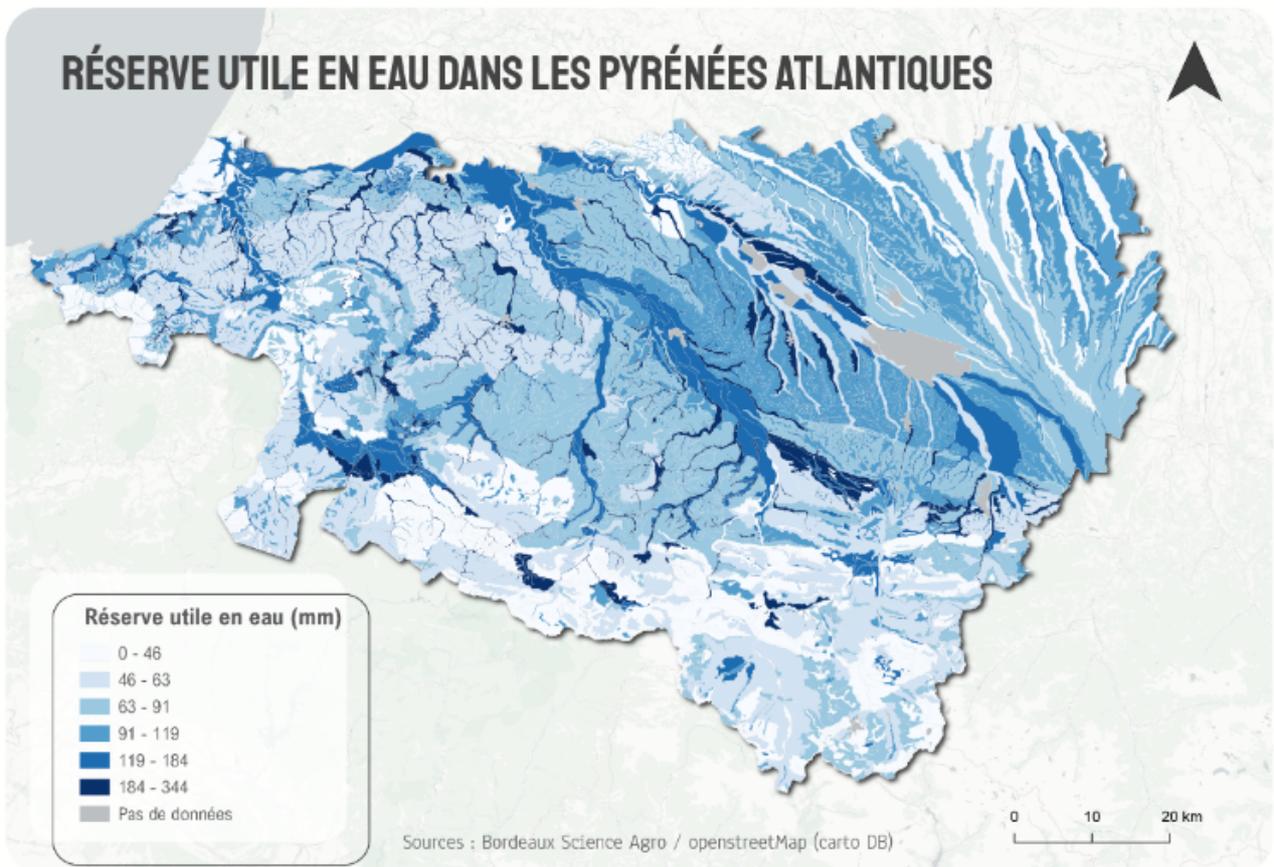


Cette méthode est considérée comme plus scientifiquement correcte par les pédologues puisqu'elle est accompagnée d'une marge d'erreur. Cependant, elle comporte le désavantage d'être assez complexe à vulgariser mais aussi à croiser avec d'autres bases de données. La couche finale est relativement complexe à traiter de manière géomatique, la diffuser en format de couche de donnée géographique n'aurait probablement pas beaucoup d'autres usages que celui de la simple visualisation.

Il existe d'autres possibilités parmi lesquelles faire une moyenne pondérée en fonction du pourcentage de chaque UTS dans l'UCS.

UTS	UCS	RU	%
22020	15994	234	70
22021	15994	191	30

UCS	RU	%
15994	$0,7 * 234 + 0,3 * 191$	70



Cette seconde méthode est une estimation de la réalité qui ne donne pas la vraie valeur associée à l'UCS. La quantification de la perte d'information est perdue puisqu'il n'est pas possible de garder la variable du poids de chaque UTS dans la couche finale. Cependant, cette couche est mobilisable dans plusieurs traitements qui correspondent à des usages différents. Cette méthode est déjà employée pour faire des modèles prédictifs croisant les données des scénarios climatiques du DRIAS avec les données de sol. Cette couche peut également être réemployée pour faire des études de recouvrement surfacique comme nous le verrons par la suite.

En bref, chaque méthode d'agrégation correspond à des usages différents, tout comme pour la méthode de calcul de la RU, il semble indispensable de parvenir à des conventions pour faire avancer la diffusion. Mettre à disposition une donnée pose toujours la question de la réutilisation. Les couches doivent donc essayer de correspondre à plusieurs usages différents pour être employés, mobilisés par des professionnels qui travaillent sur des thématiques variées.

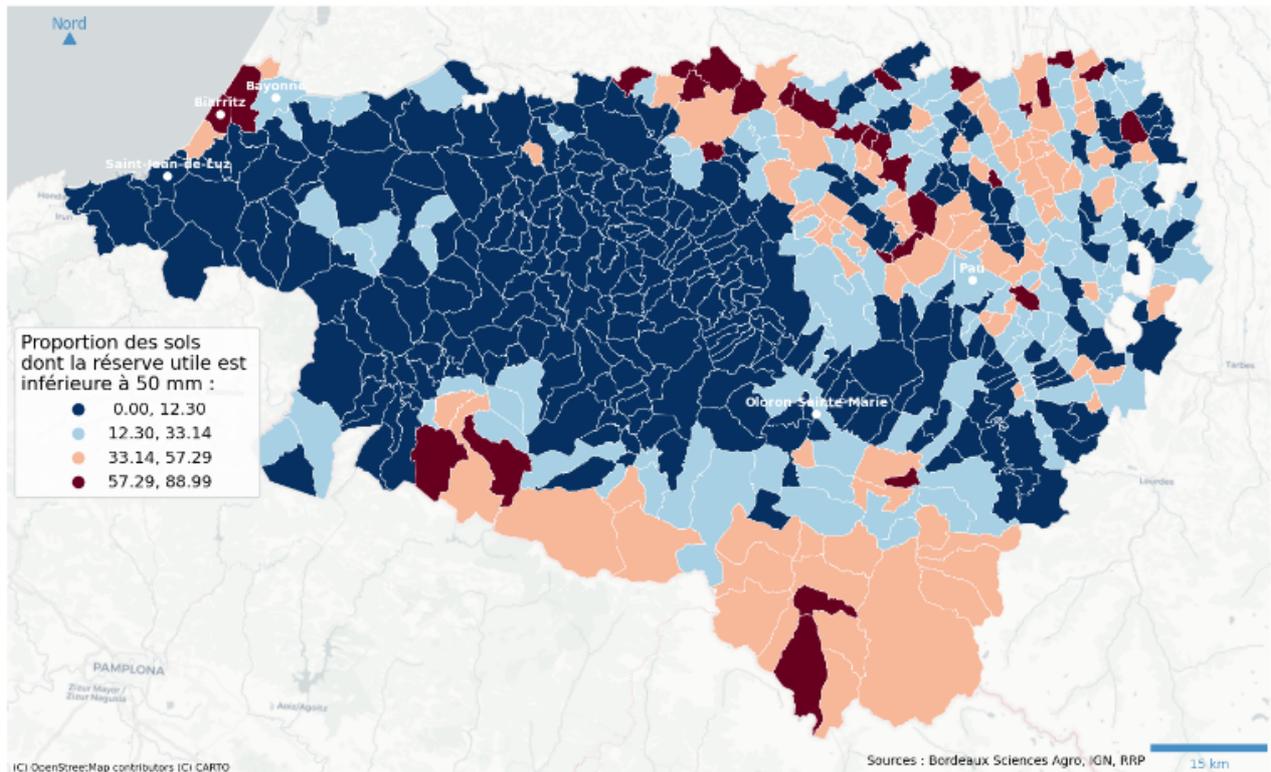
3.4 Un référentiel administratif pour mieux se situer

Les indicateurs spatialisés à l'échelle des UCS, comme cela a été créé dans l'étape précédente, ont aussi pour point faible d'être relativement peu compréhensibles par tout un chacun. C'est alors que les compétences des techniciens (cartographes, géomaticiens) entrent en jeu. Ils peuvent se saisir de ces indicateurs pour les rendre intelligibles et compréhensibles du plus grand nombre de personnes. Pour ce faire, il est nécessaire d'opérer une agrégation sur ce qu'en analyse du territoire nous appelons "un échelon administratif de référence". Il s'agit de découpages administratifs courants et généralement connus des citoyens, à savoir :

- les découpages communaux,
- les découpages des EPCI (communautés de communes, d'agglomération etc),
- les découpages départementaux,
- les découpages régionaux,
- et bien d'autres encore comme des cantons par exemple.

Une agrégation à un échelon administratif de référence a donc été opérée ici avec la réserve utile calculée précédemment grâce au script python. Cela génère une cartographie qui reste complexe mais qui nous semble compréhensible :

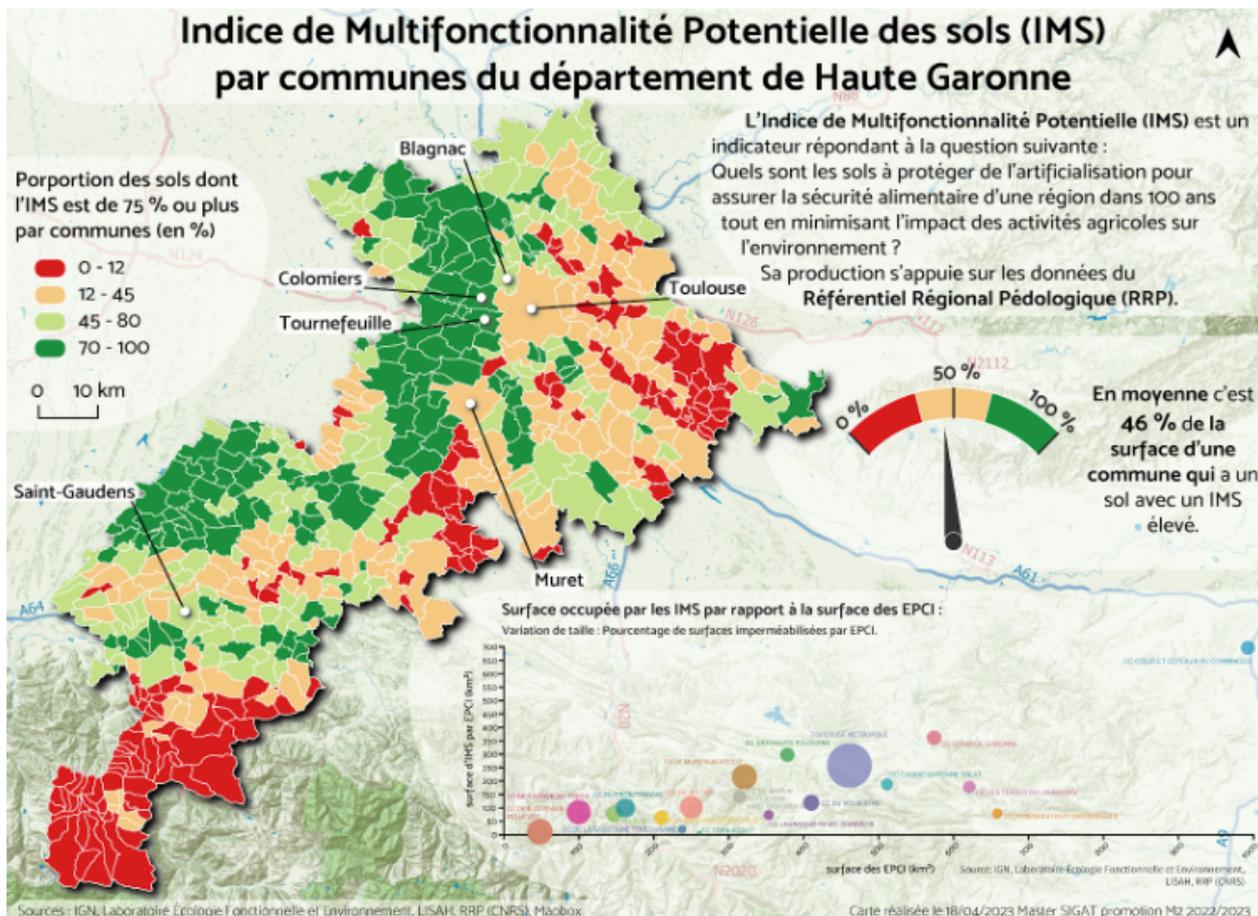
Réserve Utile des sols du département des Pyrénées Atlantiques (64):
Quelles sont les communes les plus exposées au stress hydrique ?



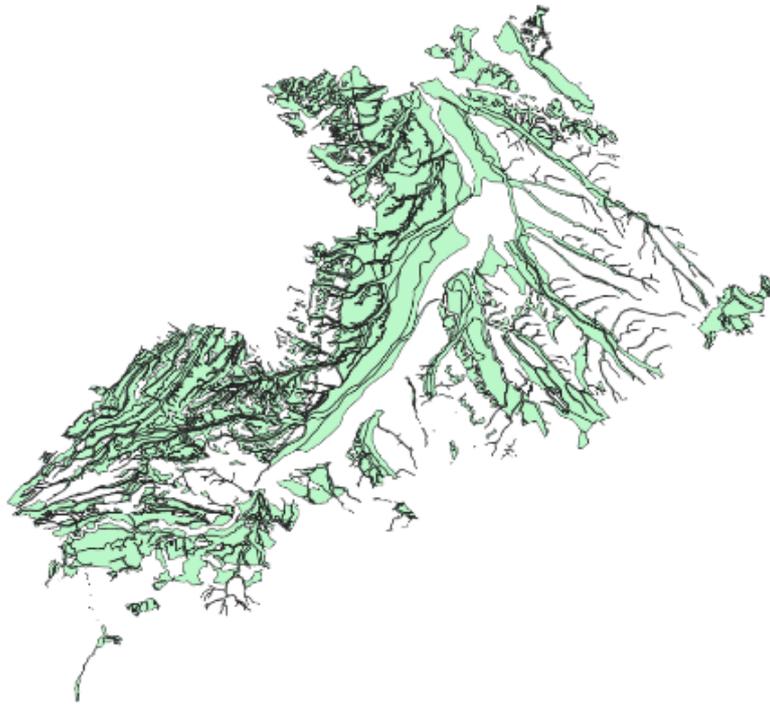
À titre de comparaison, voici la couche géographique des UCS au départ :



On se rend bien compte au travers de cet exemple que l'interprétation de la cartographie est plus complexe que pour la carte agrégée sur un échelon administratif de référence (ici les communes).



Nous avons aussi opéré la même démarche avec un indicateur diffusé en Open Data sur la plateforme OpenIG (IDG de la région Occitanie). L'Indice de Multifonctionnalité des Sols ou IMS est un indicateur prenant pour base les données présentes dans le RRP pour identifier les terres à conserver d'ici à l'année 2100 dans le but de permettre l'autonomie alimentaire de la région. Ici nous tentons de démontrer par l'exemple, le potentiel que revêt la diffusion d'indicateurs pré-calculés et spatialisés à un échelon de référence et non à celui de l'UCS comme on le voit ci-dessous :



Notre démarche pour créer la cartographie que nous avons vue plus haut a simplement été de se procurer l'indicateur d'IMS sur le portail OpenIG, de le retraiter en faisant une légère sélection attributaire, puis d'appliquer une analyse de superposition. C'est aussi cette analyse de superposition qui nous permet d'agréger la donnée à l'échelle communale puisque nous cherchons à connaître les communes qui possèdent le plus d'IMS, ou encore celles qui en sont le moins dotées. Partant de cela, il nous semble opportun d'émettre les recommandations suivantes concernant la diffusion et la réutilisation d'indicateurs pré-calculés et spatialisés issus des informations contenues dans le RRP.

Nous recommandons donc la diffusion d'indicateurs pré-calculés et spatialisés sous forme de géométries vectorielles que sont les points, les lignes ou les polygones (informations géométriques) contenant l'indicateur au sein de sa table attributaire (informations sémantiques) à l'échelon des UCS.

Cette donnée sous forme de géométrie vectorielle est un format beaucoup plus courant qui permet à d'autres professionnels de s'emparer de la donnée notamment au travers d'agrégations spatiales sur des échelons administratifs de référence. Il s'agit ici d'un exemple parmi d'autres de ceux que peut faire un géomaticien avec ce type de donnée. La diffusion à l'échelle de l'UCS permet de garder une forme qui peut s'adapter aux usages et aux thématiques. L'agrégation peut être faite aussi bien sur les EPCI que sur les communes en fonction des besoins de l'utilisateur. La diffusion sous la forme d'indicateur déjà agrégé sur une échelle administrative est une option elle aussi, elle a l'avantage de s'assurer que la donnée est traitée correctement mais elle ne permet pas de répondre à autant d'usages. Les possibilités de réutilisation sont très limitées en dehors de la cartographie.

04. Conclusion

- 4.1** Faisabilité et priorisation des préconisations
- 4.2** Perspectives du projet

4.1 Faisabilité et priorisation des préconisations

En bref, le sujet de la diffusion des données de sol semble complexe sur beaucoup d'aspects structurels, économiques et même politiques. L'étude qui fait l'objet de ce rapport examine les différents problèmes qui interviennent au cours du cycle de vie de la donnée. Ceux-ci dépassent les seules difficultés liées au manque de métadonnées. Malgré ce constat peu satisfaisant, les opportunités liées à cette thématique sont multiples. Plusieurs leviers d'actions existent pour mettre en avant cette donnée et ce domaine, nous en avons cité quelques-uns. Leur mise en place ne sera pas immédiate et certains ne sont pas simples à actionner. Il appartient désormais aux producteurs de hiérarchiser ces propositions en fonction de leur faisabilité. La communication autour des usages du RRP nous semble tout à fait prioritaire car elle semble être la première étape qui pourrait enclencher une dynamique autour des besoins sur la connaissance du sol. La mise à disposition des données de sol ne semble pas suffisante pour qu'elle soit correctement exploitée mais est nécessaire pour porter à connaissance le sol. Les utilisateurs potentiels gagneraient à avoir accès à ce que ces données permettent. Les producteurs gagneraient à leur tour en visibilité et peut-être donc en moyens humains également.

4.2 Perspectives du projet

Cette étude menée par le groupe d'étudiants SIGAT s'inscrit dans un projet plus large porté par Bordeaux Science Agro. Un stage suivra donc ce travail pour créer une plateforme de géovisualisation d'indicateurs calculés à partir du RRP (à l'image de ce que proposent Sols de Bretagne ou DataGrandEst). Ces plateformes sont développées à partir de Mviewer, une solution libre. Les codes sont donc disponibles pour les différents producteurs qui souhaiteraient aussi reprendre le concept. Celle qui sera créée dans le cadre de ce stage s'inspirera de ce qui a été fait ailleurs et pourra donc inspirer à nouveau d'autres producteurs.

Dans le même esprit, le script python pour calculer la réserve utile a pour vocation à être diffusé pour être employé sur d'autres RRP. Il devra avant cela être rectifié sur certains aspects où il était encore incomplet.