
DMP du projet "SNO KARST"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

Renseignements sur le plan

Titre du plan	DMP du projet "SNO KARST"
Langue	fra
Date de création	2021-04-22
Date de dernière modification	2021-07-14
Identifiant	SNOKARST_v1

Renseignements sur le projet

Titre du projet SNO KARST

Résumé

Le SNO KARST vise essentiellement à caractériser l'état qualitatif et quantitatif de la ressource en eau des hydrosystèmes karstiques, et à prévoir son évolution en réponse à des forçages à différentes échelles temporelles. Cette problématique se décline en trois questions scientifiques et challenges :

- Les mécanismes de transfert et de transport en milieu karstique
- Les liens entre structure géologique et écoulement
- L'évolution des flux d'eau et de matière en réponse aux changements globaux

Les observatoires du service constituent un panel représentatif des différents types d'environnements carbonatés karstiques. Leur utilisation comme objet d'études est antérieure à la labellisation du service, permettant d'inscrire les observations dans des séries temporelles longues.

La notion d'hydrosystème implique l'existence d'un bassin d'alimentation, entité géographique tri-dimensionnelle au sein de laquelle circulent des flux (massiques, chimiques) qui convergent vers un ou plusieurs exutoires. Le service développe ainsi une stratégie d'observation intégratrice, pertinente pour la compréhension des processus physiques régissant les transferts de pression et solutés à l'échelle de l'hydrosystème karstique :

- caractérisation particulièrement fine des réponses hydrodynamiques et hydrogéochimiques des exutoires (sources pérennes et temporaires) au moyen de mesures in situ à haute fréquence temporelle des débits et paramètres physico-chimiques (conductivité électrique, température, fluorescence naturelle en lien avec la matière organique dissoute, turbidité) associées à la mesure en laboratoire de paramètres biogéochimiques de prélèvements d'eau
- suivi temporel de la dynamique hydrologique (débit, piézométrie), physico-chimique, biogéochimique des différents compartiments

du karst (sources, conduits karstiques et fractures ouvertes accessibles via les aven ou forages, matrice carbonatée peu karstifiée et peu fracturée accessibles via les forages, pertes)

- déploiement d'expériences (essais hydrauliques, campagnes de mesure hydrogéophysiques) sur des sites instrumentés au sein de l'hydrosystème.

Des variables météorologiques et relatives à la qualité hydrochimique des précipitations sont acquises sur certains sites. Ces séries temporelles sont complétées par des données acquises lors de campagnes géochimiques ou géophysiques qui permettent de décrire plus précisément les mécanismes caractéristiques du fonctionnement des aquifères karstiques en fonction de différents forçages à l'échelle du site expérimental.

Dans le cadre des activités du réseau, une attention particulière est portée sur le lien données-modèles afin, d'une part, de mieux comprendre la physique des transferts dans le karst et, d'autre part, d'enrichir la capacité des modèles à reproduire les variations de flux et de matières. Le SNO KARST a pour ambition d'assurer le rayonnement scientifique de ces activités à l'échelle internationale, en synergie avec les actions menées dans le cadre des Infrastructures de Recherche (IR) OZCAR et eLTER, tout en assurant la pérennité des mesures sur le long terme et leur sauvegarde.

Produits de recherche :

1. Données d'observation du SNO KARST (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Hervé Jourde		<ul style="list-style-type: none">• Coordinateur du projet• Personne contact pour les données
Juliette Fabre		<ul style="list-style-type: none">• Responsable du plan

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

DMP du projet "SNO KARST"

1. Description des données et collecte ou réutilisation de données existantes

Données produites par le SNO KARST

Sites de mesure

L'acquisition des données sur les différents sites du SNO KARST est assurée par **9 observatoires** qui s'appuient sur plusieurs laboratoires, OSU ou EPIC :

- observatoire *MEDYCYSS* : UMR HSM / OSU OREME
- observatoire *Fontaine de Vaucluse* : UMR EMMAH (AU / INRAE)
- observatoire *Val d'Orléans* : UMR ISTO / OSUC
- Observatoire *Karst de la Craie* : UMR M2C / UMR METIS / OSU Ecce Terra / SFR 4116 Scale
- Observatoire du *Baget* : UMR GET / UMR ECOLAB / OMP / BRGM
- Observatoire *Port-Miou* : UMR CEREGE / OSU PYTHEAS
- Observatoire *Fontaine de Nîmes* : BRGM
- Observatoire *Jurassic Karst* : UMR Chrono-Environnement / OSU THETA
- Observatoire *Karsts Aquitains* : UMR I2M / OASU

Variables mesurées

L'ensemble des données potentiellement acquises est répartie en 7 catégories :

- climatologie / météorologie / atmosphère : précipitation, température de l'air, vitesse et direction du vent, pression atmosphérique, etc ;
- hydrologie / hydrogéologie : débit, niveau d'eau (altitude ou profondeur de la nappe), pression de l'eau, etc ;
- pédologie : température et humidité du sol ;
- chimie : pH, conductivité, turbidité, température de l'eau, ions majeurs, éléments traces, isotopes, terres rares, etc ;
- bactériologie : coliformes totaux, *Escherichia coli* ;
- hydrobiologie : Chlorophylle, densité de cyanobactéries ;
- chimie des précipitations atmosphériques : pH, conductivité, matières en suspension, ions majeurs, etc.

La base de données pourra évoluer pour accueillir des variables géophysiques (ex : sondages et tomographies électriques, cartes de résistivité, polarisation spontanée, sondages RMP, etc).

Observations communes à tous les sites

L'observé actuellement mesuré au sein de chacun des observatoires concerne principalement les variables hydrodynamiques (niveaux, débits) et physico-chimiques en entrée, au sein du et en sortie du karst. Des variables météorologiques sont acquises sur la plupart des sites.

Ces paramètres sont mesurés à haute fréquence au moyen de sondes automatiques autonomes ou nécessitant une alimentation électrique, avec ou majoritairement sans télétransmission des données acquises :

- Les pressions acquises peuvent alimenter :
 - Directement une série temporelle de pression barométrique ;
 - Après compensation barométrique, une série temporelle de hauteurs d'eau. La compensation barométrique peut utiliser une série acquise au point même de mesure des hauteurs d'eau ou à plusieurs kilomètres de distance.
- Les séries temporelles des débits sont dérivées des séries temporelles de hauteurs d'eau :
 - La relation hauteur-débit est établie au moyen de jaugeages. Ceux-ci peuvent être de plusieurs types : micro-moulinet, sel, fluorescéine, radar...
 - La relation hauteur-débit est susceptible d'évoluer au fil du temps : modification du cours du lit pour certains cours d'eau, réalisation de jaugeages supplémentaires permettant de préciser les caractéristiques de cette relation.
 - La relation hauteur-débit la plus complète est utilisée pour dériver la série temporelle.
- Les températures et conductivité de l'eau sont usuellement acquises concomitamment. Les sondes de conductivité font l'objet d'un étalonnage au moyen de solutions étalon. L'étalonnage constructeur est conservé pour les sondes de température.
- L'étalonnage constructeur est conservé pour les variables météorologiques.
- Les fluorescence naturelle et turbidité de l'eau sont acquises concomitamment. L'étalonnage constructeur est conservé pour la turbidité. La fluorescence naturelle fait l'objet d'une calibration.

Pour des raisons de compatibilité avec la cinétique de réactivité des hydrosystèmes karstiques, ces variables sont mesurées selon un pas d'échantillonnage court (intervalle de mesures fixé à entre 15 à 30 minutes en fonction de la réactivité du site, pouvant être temporairement réduit dans le cas d'objectifs spécifiques).

Observations spécifiques à certains sites

Des variables géochimiques (ions majeurs, en trace, isotopes) et microbiologiques résultent de la réalisation de prélèvements et de leur analyse en laboratoire. Il s'agit d'une démarche analytique plus lourde. Ces variables font donc l'objet de campagnes ciblées portant sur des périodes guidées par les mesures hydrologiques de base et la saisonnalité. Certains observatoires effectuent également des analyses au laboratoire sur des eaux échantillonnées manuellement à basse fréquence, ou par préleveur automatique à intervalle fixe tout au long de l'année (ex. ions majeurs bien solubles).

Données produites par les partenaires du SNO KARST

Des sites partenaires ont également rejoint le SNO KARST :

- Parc Naturel Régional des Causses du Quercy
- Parc Naturel Régional des Grands Causses
- Parc National des Cévennes

Ces sites partenaires réalisent l'acquisition de variables hydrodynamiques (niveaux, débits), physico-chimiques et géochimiques des sources karstiques dans l'emprise de leur territoire. Ils les bancarisent dans les banques de données nationales ADES (pour les variables géochimiques) et HYDRO (pour les débits). La bancarisation des variables géochimiques est satisfaisante. La bancarisation des débits se fait au pas journalier, tandis que les séries sont acquises au pas de temps de la demi-heure. Une partie de l'information est ainsi perdue, tandis que les données brutes ne sont pas conservées. Les variables physico-chimiques ne sont pas bancarisées.

Pour ces sites, le SNO KARST réalise donc la bancarisation et diffusion des séries de pression, hauteur d'eau, débit, conductivité et température au pas de temps d'acquisition. Les procédures d'acquisition et validation mises en place sont homogènes avec celles utilisées dans les observatoires labellisés du SNO KARST.

Données acquises par des tiers

Le SNO KARST bancarise également des séries chronologiques de débit et niveau d'eau produites par les DREAL et distribuées par la [banque Hydro](#), pour les stations de mesure d'intérêt pour les scientifiques du SNO KARST. L'accès à ces données est restreint aux membres du réseau et permet un accès aisé à ces chroniques, en comparaison de leur extraction via la banque hydro.

Documentation de la provenance des données

Les informations relatives aux organismes produisant et distribuant la donnée, et aux méthodes d'acquisition, instruments de mesure et procédures de prélèvement sont associées à chaque série de mesure (cf **2. Documentation et qualité des données**).

Type, format et volume des données

Les données consistent en des séries chronologiques numériques (date * paramètre * valeur) et sont téléchargeables au format CSV (format ouvert).

La volumétrie de la base de données est de 2.6 Go en 2021.

L'ensemble de ces données est pertinent pour apprécier l'évolution des flux hydrogéochimiques à l'exutoire des hydrosystèmes karstiques en lien avec les changements globaux ce qui justifie une conservation à long terme.

2. Documentation et qualité des données

Métadonnées et standards

Les métadonnées peuvent être regroupées en plusieurs catégories :

- description de la thématique des données du SNO KARST ;
- description des observatoires ;
- description des personnes ;
- description des stations de mesure ;
- description des paramètres mesurés ;
- description des instruments de mesure ;
- description des procédures de prélèvement pour les données de chimie, hydrobiologie, bactériologie et chimie des précipitations.

On utilise tant que possible des référentiels ou des vocabulaires contrôlés pour renseigner les métadonnées.

Les descripteurs et les référentiels associés sont listés ici :

Catégorie	Descripteur	Standard ou référentiel utilisé
Description de la thématique	Mots-clés thématiques	GEMET , AGROVOC , EnvThes , INSPIRE
Description des observatoires	Nom de l'observatoire	
	Personnes associées	
	URL de description sur DEIMS-SDR	DEIMS-SDR
Description des personnes	Identifiant ScanR	ScanR
	Nom, organisation et adresse mail	
Description des stations de mesure	Identifiant ORCID	ORCID
	Nom de la station	

	URL éventuelle de la station	
	Coordonnées de la station (longitude, altitude)	EPSG:4326
	Altitude de la station	
	Commune et département de la station	BD GEOFLA de l'IGN
	Pays de la station	Natural Earth
	Nom et code de l'entité hydrographique associée à la station	BD CARTHAGE de l'IGN
	Bassin hydrographique / hydrogéologique associé à la station	
	Date d'installation et de fermeture de la station	
	Photos de la station	
Description des paramètres	Catégorie du paramètre	
	Nom du paramètre	
	Nom du paramètre dans le thésaurus OZCAR-Theia	Thésaurus OZCAR-Theia
	URI du paramètre dans le thésaurus OZCAR-Theia	Thésaurus OZCAR-Theia
	Catégories du paramètre dans le thésaurus OZCAR-Theia	Thésaurus OZCAR-Theia
	Catégories du paramètre dans le thésaurus GCMD	Thésaurus du GCMD
	Fraction d'échantillonnage	Vocabulaire contrôlé OZCAR-Theia
Description des instruments de mesure	Nom de l'instrument	
	Numéro de série, modèle et fabricant de l'instrument	
	Observatoire associé à l'instrument	
Description des procédures de prélèvement et d'analyse	Flaconnage : nettoyage, matériel (PP, HDPE, LDPE, verre ...)	
	Seuil de coupure, composition du filtre et nom du fabricant	
	Conditionnement de l'échantillon (acidification, température de stockage, délais avant analyse)	
	Méthode analytique	
Description des séries chronologiques	Identifiant et nom	
	Station de mesure associée	
	Paramètre mesuré associé	
	Unité de mesure	
	Niveau de données (Raw data, Quality-controlled data, Derived products)	Vocabulaire contrôlé OZCAR-Theia
	Dates de début et de fin	
	Fréquence de mesure	
	Méthode d'acquisition	
	Suivi d'événement de crue ou non	
	Instrument de mesure associé	
	Procédure de prélèvement associée	

Observatoire associé

Organisation gestionnaire des données

Organisation distributrice des données

URL du DOI associé

Série publique ou non

Tableau 1 : Les descripteurs et référentiels de la base de données du SNO KARST

Diffusion des métadonnées sur le portail de données SNO KARST

L'intégralité des métadonnées sont associées aux données publiques lors de leur téléchargement depuis le [portail des données](#) du SNO KARST. Elles sont exportées dans un fichier CSV associé aux fichiers de données.

Diffusion des métadonnées sur le catalogue de l'OSU OREME

Les métadonnées qui concernent les données produites par les observatoires et partenaires du SNO KARST sont également consultables et téléchargeables sur le [catalogue des données de l'OSU OREME](#). Dans le catalogue, les métadonnées sont agrégées à 2 niveaux de granularité :

- pour **l'ensemble** des séries de données produites par les observatoires et les partenaires du SNO KARST ;
- pour chaque **jeu de données** produit par les observatoires et les partenaires du SNO KARST : un jeu de données correspond à l'ensemble des séries mesurées sur un bassin hydrologique / hydrogéologique donné, pour une catégorie de paramètres donnée. Exemple : *données chimiques mesurées sur le bassin du Lez (observatoire MEDYCYSS)*.

Les fiches de métadonnées du catalogue sont générées automatiquement à partir des métadonnées stockées dans la base de données : un [script R disponible en ligne](#) extrait les métadonnées, les formate et les envoie au catalogue (librairies R [geometa](#) et [geonapi](#)). Le catalogue est mis à jour régulièrement (au moins 1 fois par an).

Les métadonnées du catalogue (outil GeoNetwork) respectent la norme ISO 19115 et la directive européenne INSPIRE, et sont diffusables et interrogeables selon le standard CSW (Catalogue Service for the Web) de l'Open Geospatial Consortium.

Un moissonnage du catalogue de l'OSU OREME par le portail [data.gouv.fr](#) est en cours de réalisation.

Indexation des métadonnées sur Datacite

Enfin, les métadonnées de chaque jeu de données produit par les observatoires et les partenaires du SNO KARST sont associées aux DOI de ces jeux, sous le format [Datacite](#) (voir **5. Partage des données et conservation à long terme**).

Mesure et contrôle de la qualité lors de la collecte des données produites par les observatoires du SNO KARST

Comme il n'est pas possible de suivre l'ensemble des proxies à très haute fréquence, l'un des intérêts du SNO KARST est de proposer des stratégies d'observation qui permettent la meilleure optimisation coût - fréquence - qualité des mesures en essayant d'agréger autour de l'observation une vue cognitive "amont".

Pour les mesures réalisées à partir d'échantillons d'eau, un enjeu important concerne la mise en place de protocoles de prélèvement et de conditionnement, standardisés et communs aux différents sites, de façon à ce que l'inter-comparaison des données entre les différents sites ne soit pas entachée de biais liés à une différence dans les protocoles de mesure.

Le SNO propose une stratégie d'acquisition des données (homogénéisation des fréquences d'échantillonnage notamment) pour les variables hydrologiques de base ainsi que pour les modalités de prélèvement et de conservation des échantillons.

Le matériel utilisé pour l'acquisition (sondes, préleveur automatique, etc.) en cours d'utilisation ne peut pas être uniformisé.

Contrôle de la qualité des données et métadonnées lors de leur insertion dans la base de données

Des interfaces web CRUD (Create Read Update Delete), destinées aux scientifiques et aux administrateur-rices de la base de données, permettent de saisir la quasi intégralité des métadonnées (stations de mesure, instruments, séries chronologiques, etc). Ces interfaces proposent des menus déroulants et mettent en œuvre des règles de contrôles qui permettent d'assurer l'unicité et la cohérence des informations (noms des stations, des paramètres, etc) et leur intégrité.

La commune, le département et le pays associés aux stations de mesure sont calculés automatiquement à partir des coordonnées de la station par croisement avec la couche COMMUNE de la BD GEOFLA et la couche ne_50m_admin_0_countries de Natural Earth.

Une interface web permet aux scientifiques de générer des squelettes (templates) Excel de données. Pour chaque série de données, un en-tête fournit l'identifiant de la série, la station, le paramètre, l'unité et le niveau de données associés. Ces informations permettent aux scientifiques et au script d'insertion d'identifier la série sans ambiguïté. Les scientifiques y copie les données associées à chaque série. Les fichiers Excel de données sont ensuite déposés par les scientifiques sur un serveur FTP, dans le répertoire correspondant à leur observatoire / site partenaire.

Une interface web permet enfin de déclencher l'insertion des nouveaux fichiers dans la base de données. Un script R parcourt le FTP, et vérifie pour chaque fichier le format des données : cohérence des métadonnées (dans l'en-tête), données de type numérique, format des dates. Les données valides sont insérées dans la base, et un mail récapitulatif est envoyé au-à la scientifique et aux administrateur-rices de la base de données.

Cas particulier des données de la banque Hydro

Les données de débit et de niveau d'eau des stations de la banque Hydro qui présentent un intérêt pour les scientifiques du SNO KARST sont automatiquement récupérées toutes les 6h avec l'[API Hub'Eau hydrométrie](#) et insérées dans la base de données du SNO KARST via un script R.

3. Stockage et sauvegarde pendant le processus de recherche

Stockage et organisation des données

Les données et les métadonnées sont stockées dans une base de données relationnelle PostgreSQL / PostGIS. Cette base de données est stockée en mode RAID6 sur une baie de stockage localisée dans les locaux de Géosciences Montpellier.

Le serveur de base de données sur lequel est installé le SGBD PostgreSQL est hébergé sur un serveur physique (via une machine virtuelle), lui-même hébergé dans les locaux de Géosciences Montpellier.

Sauvegarde des données

Un robot de sauvegarde sur bandes LTO7 délocalisé sur le campus de la DR Occitanie-Est du CNRS prend en charge la sauvegarde de données et de la machine virtuelle. La sauvegarde est journalière pendant 2 mois, puis hebdomadaire et une copie sur bande est extraite annuellement. Le fait que le robot soit localisé sur un autre site que les baies de stockage permet de se prémunir de la perte des données en cas de vol ou destruction physique de tout ou partie des baies de stockage.

Déménagement de l'OSU OREME

L'OSU OREME déménage en 2022 sur un autre campus de l'Université de Montpellier. Le stockage du serveur physique et sa sauvegarde devront donc être redéfinis. Il est envisagé de passer par l'offre de service du mésocentre Meso@LR et/ou du DataCentre Occitanie qui est en ce moment à l'étude (possibilité de réplication des données entre Montpellier et Toulouse).

Récupération des données en cas d'incident

En cas d'incident, la base de données est récupérée par la plateforme SI de l'OSU OREME à partir de la dernière sauvegarde.

Accès aux données et authentification

L'accès aux données s'effectue par le seul biais du portail de données. Les autres types d'accès ne sont permis que par le personnel du Système d'Information.

Seules les données produites par les observatoires du SNO KARST et éventuellement ses partenaires sont ouvertes et accessibles au public. Les données produites par des tiers (DREAL etc), et éventuellement les données des partenaires du SNO KARST, sont uniquement accessibles aux scientifiques du SNO KARST et sont soumises aux conditions d'utilisation définies par les organismes produisant, gérant et distribuant ces données.

La saisie des données et métadonnées est limitée aux scientifiques du SNO KARST et aux administrateur·rices de la base de données, tout comme la consultation des données non encore ouvertes ou produites par des tiers. Ces pages sont protégées et nécessitent une authentification sur le portail des données de l'OSU OREME.

Le système d'authentification s'appuie sur la [Fédération d'Identité Enseignement Supérieur et Recherche](#) opérée par le RENATER et mis en œuvre à l'aide du logiciel shibboleth. Ainsi la gestion des identifiants/mot de passe et l'authentification est déléguée aux organismes affiliés (CNRS, Universités, ...). Des identifiants/mots de passe sont cependant gérés localement pour les scientifiques appartenant à un organisme non affilié à la Fédération d'Identité et l'authentification est alors effectuée localement. L'utilisateur·rice connecté·e dispose de différents rôles qui lui donnent accès ou non à certaines pages protégées. L'attribution des droits est effectuée manuellement et à la demande des scientifiques du SNO Karst par le personnel en charge du Système d'Information.

La consultation des données ouvertes et produites par le SNO KARST et éventuellement de ses partenaires, et de l'ensemble des métadonnées, est accessible sans authentification.

Données sensibles

Les données du SNO KARST ne revêtent pas de caractère sensible, et ont vocation à être ouvertes. Seules les informations des points de contact (noms, prénoms, identifiants Orcid et adresse mail des contributeur·rices), ainsi que les comptes locaux sont considérés comme des données à caractère personnel et seront traitées comme telles (voir **4. Exigences légales et éthiques, codes de conduite**).

4. Exigences légales et éthiques, codes de conduite

Les noms, prénoms, identifiants Orcid et adresses mail des contributeur·rices du SNO KARST sont stockés dans la base de données et associés aux données ouvertes comme points de contact.

Un message sera envoyé régulièrement (1 fois par an par exemple) à l'ensemble de ces personnes pour leur rappeler que ces informations sont stockées et dans quel but, et les informer que ces données peuvent être supprimées sur demande.

Au-delà de la gestion des données du SNO KARST, une politique et un plan d'action RGPD vont être mis en place au niveau de l'ensemble des données scientifiques et administratives de l'OSU OREME.

L'ensemble de ces mesures permettra de mettre en conformité les données du SNO KARST avec le RGPD.

Données produites par les observatoires du SNO KARST

Les données produites par les observatoires du SNO KARST étant financées par des subventions publiques, selon la [Loi pour une](#)

[République Numérique](#), elles ont l'obligation d'être ouvertes et accessibles.

Ces données sont donc ouvertes dès que possible (voir **5. Partage des données et conservation à long terme**), et aucune authentification n'est nécessaire pour accéder aux données publiques.

Les données sont diffusées sous [Licence Ouverte 2.0](#), compatible avec la [Licence Creative Commons Attribution 4.0 \(CC BY 4.0\)](#) : les données sont librement ré-utilisables, sous condition de citer leurs auteur-es.

Par ailleurs, en cas d'utilisation des données du SNO KARST, il est demandé :

- de contacter la personne ressource du site d'observation, de préférence dans le cadre d'une collaboration de recherche ;
- de fournir à la personne ressource du site d'observation l'opportunité de donner leur expertise sur les résultats obtenus à partir de l'utilisation du jeu de données ;
- si l'utilisation de ces données donne lieu à une publication, de communiquer le DOI et de remercier le SNO KARST via une phrase type.

Données produites par les partenaires du SNO KARST

Les partenaires du SNO KARST peuvent ouvrir leurs données s'ils le souhaitent. Elles sont alors diffusées sous [Licence Ouverte 2.0](#) : les données sont librement ré-utilisables, sous condition de citer leurs auteur-es.

Données produites par des tiers

Les données produites par des tiers (DREAL etc) sont uniquement accessibles aux scientifiques du SNO KARST et sont soumises aux conditions d'utilisation définies par les organismes produisant, gérant et distribuant ces données.

Les données du SNO KARST ne soulèvent aucune question d'éthique.

5. Partage des données et conservation à long terme

Les données d'observation produites par les observatoires du SNO KARST et ses partenaires sont diffusées sur le [portail des données du SNO KARST](#) :

- visualisation des stations de mesure et de leurs informations : [carte interactive](#) ;
- [recherche multi-critères et téléchargement](#) des données, avec leurs métadonnées et les DOI associés ;
- recherche et visualisation des données : [graphiques interactifs](#).

Les données sont indexées dans le [catalogue des données](#) de l'OSU OREME et les métadonnées sont diffusées via un flux standard CSW (voir **2. Documentation et qualité des données**).

Les données cartographiques (couche spatiale des stations de mesure) sont diffusées sur le serveur cartographique de l'OSU OREME (outil GeoServer) dans les flux standards WFS (https://sig.oreme.org/geoserver/sf/wfs/snokarst_station) et WMS (https://sig.oreme.org/geoserver/sf/wms/snokarst_station).

Le SNO KARST étant partie intégrante de l'IR OZCAR et de l'ALTER-RI, les données et les métadonnées produites par ses observatoires (à l'exclusion de celles de ses partenaires) sont exportées régulièrement (au moins 1 fois par an) vers le [portail Theia - OZCAR](#), sur lequel elles sont visualisables. Elles sont également accessibles dans la base de données [DEIMs](#), utilisée comme base de connaissance des observations disponibles dans eLTER-RI, sous le référencement OZCAR-RI SNO KARST *nom de l'observatoire* (exemple pour Jurassic Karst : <https://deims.org/6544abc3-1994-4ccd-9b85-95acba02712>).

Ouverture et ré-utilisation des données produites par les observatoires du SNO KARST et ses partenaires

Le délai de diffusion des données après acquisition est variable selon les séries et types de sondes. De façon générale :

- Les variables *hauteur d'eau* et *débit* aux exutoires karstiques ainsi que les variables météorologiques sont diffusées dès insertion dans la base de données.
- Les variables *hauteur d'eau* et *débit* au sein des hydrosystèmes karstiques sont également insérées dès insertion dans la base de données. Cependant pour ces variables les fréquences de relevé peuvent être basses : c'est en particulier le cas des points d'observations nécessitant un accès spéléologique, sur lesquels les capteurs peuvent être laissés en autonomie durant plusieurs années.
- Les variables physico-chimiques et les espèces ioniques majeurs (variable géochimique) sont soumises à embargo de 24 mois après validation.
- Les autres variables sont insérées après avoir fait l'objet d'une publication scientifique.

Données produites par des tiers

Les données produites par des tiers (DREAL, etc) ne sont pas diffusées. Seules les stations de mesure et les métadonnées associées sont consultables sur la [carte interactive du SNO KARST](#).

Utilisation prévisible des données

Les données produites par les observatoires et partenaires du SNO KARST seront utilisées par :

- les scientifiques membres du réseau et rattachés à l'observatoire producteur (la base de données servant de support d'archivage) ;

- les scientifiques membres du réseau hors observatoire producteur (la base de données servant de support à la transmission des données validées) ;
- les citoyen.nes en recherche d'informations relatives à la ressource en eau (avec un accès facilité à la visualisation des données via les graphiques de présentation des données) ;
- les scientifiques hors réseau SNO KARST, au sein de la communauté nationale et internationale.

L'accès aux données produites par des tiers est limité aux seuls scientifiques du SNO KARST.

Préservation des données

Les données d'observation du SNO KARST sont des données chaudes et n'ont pas vocation à être archivées à court ou moyen terme.

A la fin de l'existence du SNO KARST, le maintien de la diffusion de ces données sur le portail de l'OSU OREME devra être discuté, sachant que les données produites par les observatoires du SNO KARST seront disponibles sur le portail Theia - OZCAR.

L'archivage pérenne des données sera envisagé sur le long terme, en s'appuyant sur les centres de données et les offres de service du moment (exemple : [offre d'archivage du CINES](#)). Seules les données produites par les observatoires et les partenaires du SNO KARST seront préservées. Les séries calculées simplement à partir d'autres séries (ex: moyenne journalière à partir de données horaires) pourront ne pas être archivées. On pourra également sélectionner seulement certains niveaux de données (ex : Raw data, ou Quality-controlled data).

Les données et métadonnées sont téléchargeables au format CSV et sont donc ré-utilisables sans outil spécifique. L'accès aux métadonnées est rendu possible via le protocole CSW à l'aide de l'outil de catalogue GeoNetwork. La diffusion des données via le protocole SOS et/ou SensorThings est à l'étude.

On définit la notion de **jeu de données**, qui correspond à l'ensemble des séries de données mesurées sur :

- sur un bassin hydrologique / hydrographique donnée
- et pour une catégorie de paramètres donnée.

Chaque jeu de données ouvert produit par les observatoires du SNO KARST et éventuellement ses partenaires est **identifié** par un DOI (Digital Object Identifier) fourni par l'organisation Datacite, pour laquelle l'OSU OREME est centre de données. Ces jeux de données sont donc catalogués sur [Datacite](#).

6. Responsabilités et ressources en matière de gestion des données

L'OSU OREME, qui est l'OSU porteur du SNO KARST, est responsable de la gestion des données du SNO KARST. La plateforme Système d'Information (SI) de l'OREME, constituée d'Olivier Lobry (responsable) et de Juliette Fabre, est en charge de cette gestion pour les aspects suivants :

- le stockage et la sauvegarde des données ;
- la production des métadonnées ;
- la diffusion des données et des métadonnées ;
- l'archivage des données ;

Les scientifiques du SNO KARST sont responsables de :

- la saisie des données et des métadonnées ;
- la qualité des données.

Le Plan de Gestion des Données est rédigé conjointement par la plateforme SI de l'OSU OREME et les scientifiques du SNO KARST, notamment Hervé Jourde et Naomi Mazzilli. La plateforme SI de l'OSU OREME est responsable de sa mise en œuvre et de sa révision régulière (1 fois par an).

La plateforme SI de l'OSU OREME est responsable de la gestion FAIR des données du SNO KARST. Elle y consacre le temps nécessaire, en fonction des autres contraintes de la plateforme. Elle dispose de 2 ingénieur.es dédié.es à la gestion de l'ensemble des données d'observation dont elle a la charge. Elle possède sa propre infrastructure matérielle hébergée à Géosciences Montpellier et sur le campus CNRS de la Délégation Régionale Occitanie-Est.

La plateforme SI possède un budget propre pour la gestion des données dont elle est responsable, qu'il s'agisse du SNO KARST ou d'autres services d'observation. Ce budget couvre :

- la maintenance de l'infrastructure matérielle (serveurs, robot de sauvegarde, baie de stockage, onduleurs) ;
- les licences (système d'exploitation, logiciel de sauvegarde) ;
- la cotisation auprès de Datacite (attribution de DOI).

Mutualisé à l'ensemble des données d'observation de l'OREME, il couvre intégralement les besoins budgétaires de la gestion des données du SNO KARST.

L'infrastructure matérielle est garantie jusqu'à fin 2022. Au-delà, trois solutions sont à l'étude pour chaque équipement :

- le prolongement de garantie de l'équipement sur des crédits de fonctionnement ;
- la jouvence de l'équipement sur des crédits d'équipement ;
- le non remplacement de l'équipement via la location de ressources sur des Cloud recherche, via des crédits de fonctionnement.

Cette réflexion, globale au niveau de la plateforme, se fait conjointement avec celle relative au déménagement de la plateforme sur un autre site de l'Université de Montpellier prévu à l'automne 2022.

Les scientifiques du SNO KARST consacrent une partie de leur temps pour collecter les données et les insérer dans la base de données.

Les coûts d'archivage devront être estimés et pris en charge par l'OSU OREME et le SNO KARST.