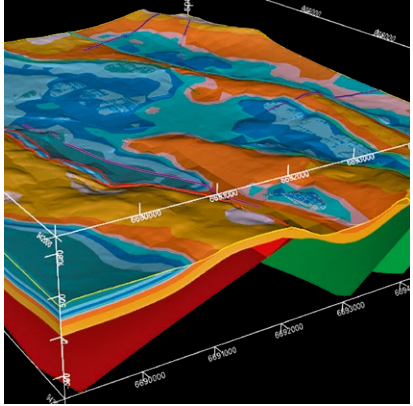


SERVICE
GÉOLOGIQUE
NATIONAL

LA POLITIQUE
DE SCIENCE
OUVERTE AU
BRGM
2022



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*



Géosciences pour une Terre durable

brgm

INTRODUCTION

Le mouvement pour la science ouverte s'inscrit dans un contexte international de la science depuis plus de vingt ans. Il connaît une accélération sans précédent à l'échelle mondiale depuis quelques années avec la mise en place de cadres de politiques scientifiques et juridiques à l'échelle nationale et internationale favorisant l'ouverture systématique de données et connaissances acquises sur fonds publics. Dans ce contexte, la décision toute récente des États-Unis de rendre public et accessible sans période préalable d'embargo l'intégralité des études financées sur fonds publics et de nature à accélérer à l'échelle mondiale la transition en cours.

La digitalisation des données et des résultats scientifiques et leur diffusion sur internet bouleversent fondamentalement la transmission de la connaissance qui devient un bien commun mondial aisément accessible et pour l'essentiel gratuitement. Cela induit une transformation des pratiques de la science dans toutes ses dimensions depuis la recherche jusqu'à l'expertise publique ou le développement d'innovations. La science ouverte permet l'accès aux données, le partage des savoir-faire, la confrontation des résultats et bien évidemment la diffusion des connaissances, désormais à l'échelle globale. La science ouverte permet également d'améliorer l'avancement de la connaissance, d'accélérer l'établissement du consensus scientifique et d'en renforcer l'impact. Elle favorise les relations entre différents domaines scientifiques et la production de nouvelles connaissances visant des problématiques complexes et transdisciplinaires.

L'ouverture des processus scientifiques ne se limite pas à la recherche et s'inscrit dans un cadre politique plus large comme l'énonce la circulaire du Premier ministre du 27 avril 2021 relative à une politique de la donnée, des algorithmes et des codes sources. Ces impulsions politiques et institutionnelles reposent sur le constat de l'importance stratégique de l'ouverture, de l'exploitation et du partage de la donnée, des algorithmes et des codes sources publics, et de leur meilleur croisement avec les données des acteurs privés. Elles resituent ces enjeux dans le contexte ambitieux des initiatives lancées au niveau européen sur la transition numérique, la donnée et l'intelligence artificielle.

L'ouverture à d'autres acteurs, du secteur privé ou de la société civile, permet de combiner des données et des savoirs de différentes natures générant de nouvelles connaissances et des innovations en particulier dans le numérique. C'est donc aussi de nature à faire émerger de nouveaux besoins et de nouveaux marchés.

La science ouverte propose un cadre de référence pour établir un lien entre les pratiques scientifiques en partageant les données, les méthodes et les codes numériques.

Il y a d'autre part, une exigence de plus en plus grande de la société et des pouvoirs publics vers plus de transparence et d'impact en matière d'élaboration et de diffusion des connaissances scientifiques. La science ouverte vise aussi à renforcer ce lien entre les scientifiques et la société. Elle permet d'expliquer, d'informer largement et se doit d'être accessible à différents types de publics.

Dans ce contexte la science ouverte propose un cadre de référence pour établir un lien entre les pratiques scientifiques en partageant, les données, les méthodes, plus récemment les codes. En Europe, la Commission a inscrit de façon volontariste depuis le programme H2020 en 2016, sa politique de soutien à la recherche dans un cadre de science ouverte « Open Innovation, Open Science, Open to the World¹ », favorisant l'innovation et ouvrant la science sur la société, en la mettant en œuvre via les programmes-cadres (H2020, Horizon Europe) et en développant une politique de management des données et d'infrastructures via EOSC² (European Open Science Cloud).

La politique publique de science ouverte se gère toutefois selon le principe « aussi ouvert que possible, aussi fermé que nécessaire » et il convient dans le schéma de gouvernance de bien intégrer la primauté des cadres juridiques de protection intellectuelle et industrielle, voire de protection des intérêts stratégiques de la nation dans le cadre des réflexions sur le renforcement de la souveraineté.

LA POLITIQUE DE SCIENCE OUVERTE AU BRGM

Sur le plan législatif, la loi pour une République numérique de 2016³ favorise le dépôt en archives ouvertes : « Lorsqu'un écrit scientifique issu d'une activité de recherche financée au moins pour moitié par des dotations de l'État, ... est publié dans un périodique ..., son auteur dispose, même après avoir accordé des droits exclusifs à un éditeur, du droit de mettre à disposition gratuitement ..., sous réserve de l'accord des éventuels coauteurs, la version finale de son manuscrit acceptée pour publication, dès lors que l'éditeur met lui-même celle-ci gratuitement à disposition par voie numérique ou, à défaut, à l'expiration d'un délai ... au maximum de six mois pour une publication dans le domaine des sciences, de la technique et de la médecine et de douze mois dans celui des sciences humaines et sociales ».

En France, le premier Plan national pour la science ouverte⁴ a été lancé, en juillet 2018, dans le but de « généraliser l'accès ouvert aux publications, de structurer et ouvrir les données de la recherche » et d'inscrire la recherche française « dans une dynamique durable, européenne et internationale » et d'initier une gouvernance des données dans les établissements de recherche. Le second Plan national⁵ initié en juillet 2021 étend son périmètre aux codes sources issus de la recherche, multiplie les leviers de transformation afin de généraliser les pratiques de science ouverte. Il propose également des déclinaisons disciplinaires et thématiques.

En ce qui concerne les politiques de l'environnement, et au titre des obligations qui pèsent sur les administrations en matière de politique de la donnée, la transparence et le partage de l'information environnementale sont des acquis issus de deux textes prépondérants : la convention d'Aarhus - accès à l'information, participation du public au processus décisionnel, accès à la justice en matière d'environnement (1998) - et la directive européenne INSPIRE - qui vise à établir une infrastructure d'information géographique dans l'Union européenne pour favoriser la protection de l'environnement (2007). Ce contexte, dans lequel le BRGM agit au titre de ses actions d'appui aux politiques publiques, a permis l'ouverture depuis longtemps de l'essentiel des données et résultats scientifiques acquis dans ce schéma.



Exemple d'une visualisation de carte géologique du BRGM sur le portail Infoterre.
© BRGM

Cette politique générale s'inscrit pleinement dans le cadre fixé par la feuille de route de la donnée, des algorithmes et des codes sources publiée en 2021 par le ministère en charge de la transition écologique.

¹ Commission européenne, Direction générale de la recherche et de l'innovation, Open innovation, open science, open to the world : a vision for Europe, Publications Office, 2016, <https://data.europa.eu/doi/10.2777/061652>

² <https://eosc-portal.eu/>

³ Loi pour une république numérique du 7 octobre 2016, article 30. https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000033202841

⁴ <https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte/>

⁵ <https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/>

Les ambitions du BRGM au travers de cette politique générale

La science ouverte est au cœur de la stratégie scientifique et vise à positionner le BRGM comme centre national de référence pour les données du sol et du sous-sol



Exemple d'une visualisation de carte géologique du BRGM sur le portail Infoterre.

© BRGM

Le BRGM est un acteur de premier plan national et européen en matière de sciences et d'informations sur le sol, le sous-sol et ses ressources. Son action de recherche et d'expertise pour les pouvoirs publics, et notamment sa mission de service géologique national, implique la collecte, la valorisation et la diffusion d'un grand nombre de données et connaissances sur le sous-sol. Cela place la science ouverte au cœur de sa stratégie scientifique. Il s'agit de positionner le BRGM comme centre national de référence pour les données du sol et du sous-sol.

Cette politique a pour ambition de structurer les pratiques à l'échelle de l'établissement et d'accélérer le processus de la science ouverte dans toutes les dimensions scientifiques portées par le BRGM en s'appuyant sur cinq axes stratégiques. Elle sera mise en œuvre au travers d'actions intégrées au contrat d'objectifs, de moyens et de performance 2023-2027.

La politique de science ouverte du BRGM couvre l'ensemble des connaissances et données scientifiques

produites par le BRGM sur fonds publics pour son compte ou celui de ses partenaires publics ou donneurs d'ordre, ainsi que celles qui lui sont confiées, quel que soit le cadre juridique dans lequel ces données ont été produites ou collectées, à l'exception des données rentrant dans le champ de la confidentialité pour les besoins de la Défense, du secret des affaires ou pour préserver la souveraineté nationale. L'ouverture des données scientifiques et savoirs générés ou possédés par le BRGM doit être considéré comme la pratique de référence, cependant elle est encadrée par des conditions spécifiques :

- Le respect du règlement général sur la protection des données (RGPD) lorsque ces données comportent des données à caractère personnel ;
- Le respect de la propriété des données non publique et des droits relatifs à la confidentialité, au secret des affaires et

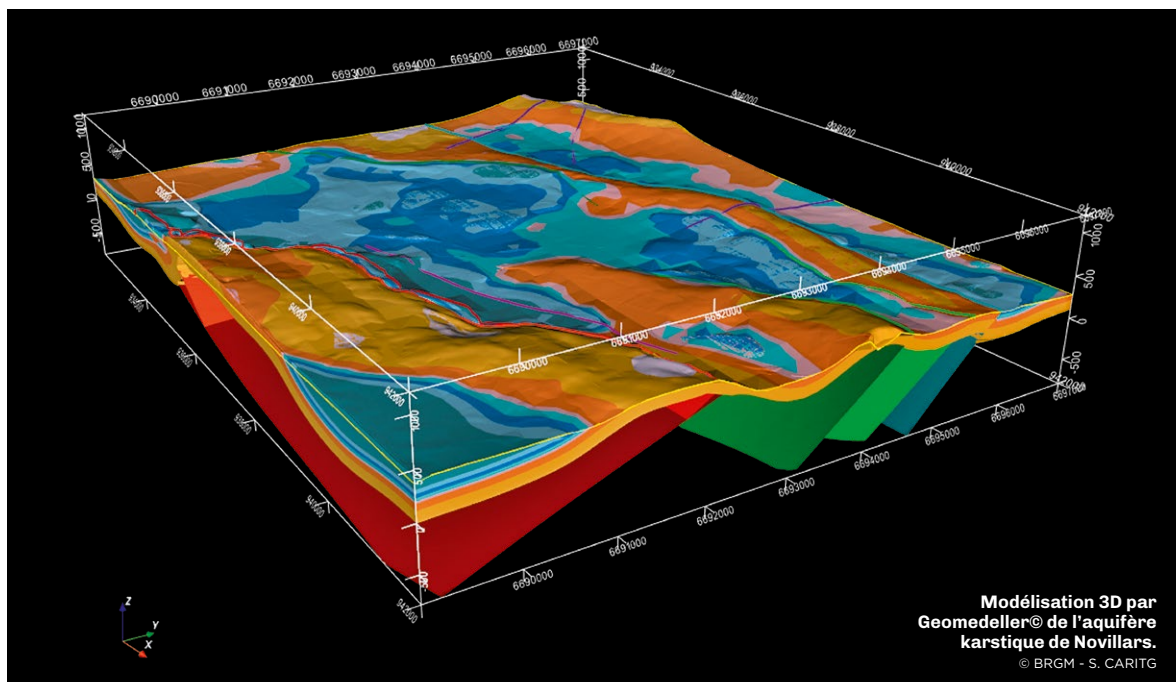
- à la préservation de la souveraineté, notamment dans les actions de recherche ou de développement dont le financement public est inférieur à 50% ;
- Le respect du droit des affaires du BRGM et notamment de sa politique de développement de la propriété intellectuelle (brevet, licence...);
- Les données et savoirs mettant en péril la sécurité et la défense de la nation ;
- Les données et savoirs inachevés, c'est-à-dire devant être utilisés, vérifiés ou remaniés avant d'être diffusés publiquement (e.g. article en cours de rédaction).

La politique de science ouverte du BRGM vise plusieurs objectifs :

- Assurer le contrôle de notre production scientifique produite sur fonds publics et aboutir à terme à 100% des publications du BRGM (articles scientifiques, cartes et rapports publics) en accès ouverts ;
- Systématiser à terme, la mise en place de plans de gestion des données (PGD) pour l'ensemble des projets de l'établissement. Utiliser les standards ouverts des différentes disciplines scientifiques et contribuer à leur élaboration internationale si nécessaire. Garantir la pérennisation des données scientifiques des activités de recherche, d'expertise et leur libre accès ;
- Faire évoluer les processus internes et les compétences en matière d'acquisitions de données et de publications afin de développer une culture de la gestion des données chez tous les acteurs du cycle de vie de la donnée et basée sur la mise en œuvre des principes FAIR (Faciles à trouver, Accessibles, Interopérables et Réutilisables) ;

- Développer et promouvoir des infrastructures, des outils et des standards permettant les requêtes et l'analyse des contenus scientifiques en toute indépendance ;
- Inscrire cette politique d'établissement dans un cadre de partenariats au niveau national et international avec d'autres établissements publics, producteurs de connaissances et données, afin de fédérer ces sources de données. Il s'agira notamment de participer activement au déploiement des infrastructures de données européennes (EPOS, eLETER, ECCSEL) et nationales (OZCAR, RESIF, DataTerra) dédiées aux thématiques sol, sous-sol et ressources associées ainsi que celles permettant une interopérabilité et des mises en accès plus larges et extra-communautaires (Data Terra, ENVRI, EOSC).
- Croiser les données du BRGM avec des données externes afin de générer des opportunités scientifiques, pour produire de la science nouvelle et développer des services à plus forte valeur ajoutée en réponse aux défis scientifiques et sociétaux de notre temps.

Le BRGM a également des missions de recherche industrielle et d'activité commerciale, il s'agit également de veiller à préserver ses savoir faire, ses découvertes et innovations selon le principe « la science doit être aussi ouverte que possible et aussi fermée que nécessaire » afin de protéger les intérêts de l'établissement et respecter les politiques de souveraineté de la France ou des pays ou partenaires pour lesquels le BRGM effectue des travaux scientifiques. La politique de science ouverte du BRGM vise à maîtriser ces risques afin de gérer la propriété intellectuelle des résultats acquis, tout en examinant les possibilités de partage des données une fois la protection des résultats et innovations acquise.



⁶ <https://www.ecologie.gouv.fr/feuille-route-donnee-des-algorithmes-et-des-codes-sources>

Les fondements de la politique de science ouverte du BRGM

La politique de science ouverte est basée sur les principes FAIR et couvre l'ensemble du cycle de vie des données pour garantir une diffusion des données et résultats à long terme

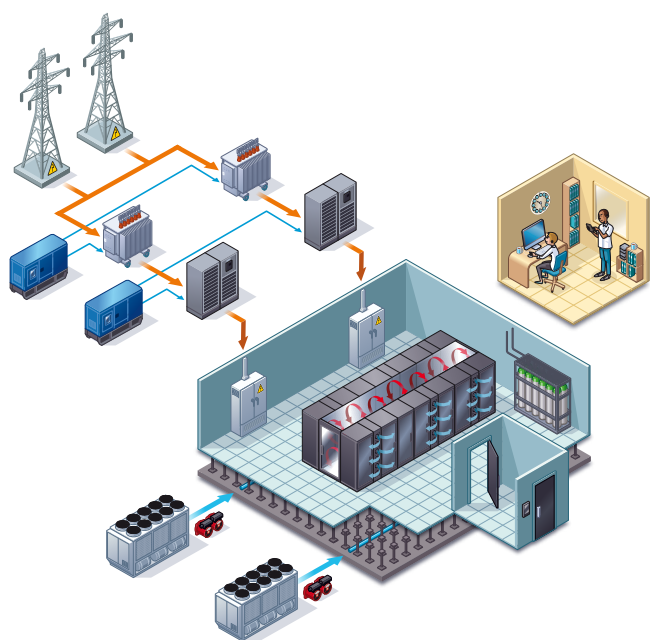
Garantir le respect de la réglementation et de la politique nationale en matière de sciences et données ouvertes

Le BRGM a joué un rôle de pionnier dans le mouvement d'ouverture des données et de leur interopérabilité en France depuis plus de 20 ans. Conformément à la réglementation française et européenne, l'établissement garantit que les résultats, les codes sources et données de son action scientifique acquises avec des financements publics sont mis à disposition, dans un souci de transparence de l'action publique et dans l'espoir d'aider au développement d'une activité économique directe ou indirecte.

En tant qu'acteur économique, le BRGM pourra lui aussi participer, pour son propre compte, à la mise au point de services numériques payants qui exploiteront des données et codes gratuits de provenance très variée.

Garantir le respect de la propriété intellectuelle des résultats et données scientifiques

L'objectif de la politique de données et des codes du BRGM est de rendre les résultats scientifiques, les codes sources et données aussi ouverts que possible conformément aux dispositions de la science ouverte, mais également aussi de les protéger autant que nécessaire, lorsque les règles de propriété intellectuelle, de confidentialité, de souveraineté ou d'enjeux économiques l'imposent. Le BRGM s'engage à identifier le cadre juridique des données, codes et résultats scientifiques qu'il produit, qui lui sont confiés par des partenaires ou dont il fait l'acquisition. Cela implique la prise en compte, dans le système d'information, la gestion



Infrastructure du DataCentre BRGM, permettant de stocker, sauvegarder et traiter les données.

© BRGM

L'objectif est de rendre les données, les résultats scientifiques, les codes sources aussi ouverts que possible.



Un levé géologique dans
le cadre du RGF Alpes.

© BRGM - F. CAGNARD

des méta-données des données, des codes (embargo inter/externe, cause, durée) permettant de déterminer les droits d'accès appropriés.

Gérer les données scientifiques sur l'ensemble de leur cycle de vie en appliquant les principes FAIR

L'ensemble des connaissances et données acquises ou produites par le BRGM pour des actions de recherche ou d'expertise, devra être géré suivant les principes FAIR à l'aide d'un processus décrit dans le système qualité du BRGM. Appliquer les principes FAIR sur le cycle des données depuis l'acquisition des données brutes sur le terrain ou expérimentales et analytiques, jusqu'aux données transformées et résultats numériques. L'ensemble des données produites doit systématiquement faire l'objet d'une description et d'une sauvegarde dans les plateformes numériques du BRGM afin d'en permettre des réutilisations ultérieures.

Adopter autant que possible les standards de la profession pour gérer les données et codes scientifiques et assurer leur future interopérabilité et donc leur ré-utilisabilité.

Garantir la préservation et la diffusion des données et résultats à long terme

Établissement public de recherche, le BRGM a la responsabilité de garantir la préservation et l'absence de compromission des données sur le long terme. Au-delà d'un objectif réglementaire, l'enjeu est de permettre de capitaliser à l'échelle nationale la connaissance du sous-sol qui relève d'une démarche agrégative, d'aider au développement des nouveaux usages, mais aussi de disposer et mettre à disposition des séries longues sur les données du sous-sol et de l'environnement.

Les axes de la politique de science ouverte

La politique de science ouverte se décline autour de cinq axes : l'ouverture des processus de recherche et d'expertise ; la gestion des données numériques ; la gouvernance et la montée en compétences ; l'ouverture au grand public et l'appui aux dynamiques européennes et internationales



Le prélèvement, l'étude et le stockage (ici à la géothèque du BRGM) de carottes géologiques est une source d'informations et de données géoscientifiques.

© BRGM - C. BOUCLEY

Ouvrir les processus de la recherche et de l'expertise

Les publications scientifiques externes et les résultats d'expertise produits par les scientifiques du BRGM et financés sur fonds publics doivent être accessibles et réutilisables dans le respect des droits d'auteurs et de la propriété intellectuelle parallèlement engagés.

Garantir le libre accès aux publications scientifiques

Le BRGM a développé un portail spécifique d'archives ouvertes⁷ en 2010 sur la plateforme nationale HAL en tant que réservoir institutionnel permettant de garantir l'archivage pérenne et la visibilité internationale des publications scientifiques du BRGM. L'ensemble des publications scientifiques du BRGM doit donc y être déposé via le portail HAL-BRGM avec le référencement aux normes (DOI, adresses ORCID des auteurs...).

Le dépôt concerne le texte intégral de la version acceptée ou publiée, de la publication, selon le contrat signé avec l'éditeur. Le dépôt dans HAL-BRGM doit se faire dès la récupération de cette version auteur en indiquant, le cas échéant dans HAL, l'embargo imposé par l'éditeur. Le délai d'embargo ne devrait pas excéder 6 mois, sauf exception, notamment pour certaines disciplines (ex : SHS) pouvant s'étendre à un an.

Le BRGM respectera autant que possible les préconisations du ministère en charge de la recherche, en matière de paiement de frais de publication supplémentaires (APC) pour le libre accès dans des revues sous abonnement et recommande le modèle « green open access⁸ ». Ce modèle autorise l'auteur à réaliser l'archivage en accès libre et gratuit des articles publiés par ailleurs dans un journal

Les publications scientifiques du BRGM seront déposées dans le portail HAL-BRGM avec le référencement aux normes (DOI, adresses ORCID des auteurs, etc.)

à souscription. L'auteur doit respecter la période d'embargo éventuelle (< 6mois) et préciser le type de licence ouverte désirée. Le modèle « gold open access⁹ » étant étudié au cas par cas via un processus centralisé. Dans ce cas, l'auteur doit payer à la revue concernée les frais de mise à disposition libre et gratuite de l'article. Le paiement d'APC dans des revues nativement en libre accès est recommandé. Le BRGM déconseille le paiement de frais de publications pour rendre libre l'accès à un article dans les revues dites hybrides qui ne sont pas vertueuses puisqu'on opère de fait un double paiement entre l'abonnement et le service de mise en accès libre. La publication dans de telles revues, sans paiement de frais de publication, accompagnée d'un dépôt dans HAL préserve la liberté de choix par les chercheurs de la voie de publication. Une centralisation de paiement d'APC sera examinée.

Renforcer la visibilité de l'ensemble des productions scientifiques du BRGM

Le BRGM produit à la fois des publications externes dans des revues scientifiques et des actes de colloques, mais également des documents scientifiques et techniques édités dans un cadre interne (rapports, cartes, guides techniques, etc.). Les nouvelles technologies de référencement en interne (DOI) et la généralisation des standards d'identification des auteurs (ORCID, IdHAL et ResearcherID¹⁰) améliorent très fortement la visibilité et donc l'accès des productions scientifiques et révolutionnent les processus de l'information scientifique et technique. Ces outils permettent de rendre plus visible l'ensemble des travaux menés par le BRGM et notamment ceux issus des processus internes de littérature dite « grise » (rapports BRGM, cartes, guides techniques...). Il s'agira de rapprocher et généraliser les pratiques de référencement et d'accessibilité entre tous les types de publications scientifiques. Une politique de soutien à la généralisation de ces outils de référencement sera mise en place.

L'ensemble de la production scientifique interne et externe et mise à disposition par le BRGM devrait à terme être aisément accessible et partageable via un lien de type DOI¹¹. Cela facilitera également la citation ouverte de données et informations scientifiques figurant dans les rapports du BRGM et utilisées par des tiers.

L'identification des auteurs de publications scientifiques internes et externes dans un contexte de science ouverte sera également généralisée via l'usage des identifiants d'auteurs (ORCID, IdHAL et ResearcherID) afin

⁷HAL-BRGM <https://hal-brgm.archives-ouvertes.fr/>

⁸ La voie verte ou green open access est la voie de l'auto-archivage ou dépôt par l'auteur dans une archive ouverte. L'auteur respecte le délai d'embargo imposé par l'éditeur.

⁹ - Le modèle « gold open access » concerne des revues ou ouvrages nativement en open access, dès leur publication. L'accès au document est libre pour le lecteur mais l'auteur ou son institution paye des frais appelés Article processing charges (APC) pour financer la publication.

¹⁰ Cf. glossaire

¹¹ Cf. glossaire



Un contrôle piézométrique pour le monitoring d'une nappe souterraine.

© BRGM

d'améliorer la visibilité des auteurs et de leur affiliation et permettre de rattacher de façon certaine un auteur à une publication.

Le BRGM privilégie l'accès à sa production scientifique via des plateformes d'archives ouvertes et essentiellement ses plateformes de référence HAL-BRGM et InfoTerre. L'usage des réseaux sociaux scientifiques ou professionnels de type (Researchgate, Academia, LinkedIn...) peuvent être utilisés avec discernement par les scientifiques du BRGM à titre personnel, à des fins de communication et mise en valeur de résultats scientifiques et non comme des outils d'archivage.

Gérer les données scientifiques sur leur cycle de vie en appliquant les principes FAIR

L'ensemble des connaissances et données acquises ou produites par le BRGM, devra être géré suivant les principes FAIR, suivant un processus décrit dans le système qualité du BRGM. Le partage des données issues des projets financés sur fonds publics sera l'approche par défaut en

respectant les règles d'ordre juridiques, éthiques et contractuels applicables aux données concernées.

Toutefois, sont exclues, ou différées de la diffusion publique des données :

- les données confidentielles, notamment celles servant à protéger les savoir-faire et innovations du BRGM notamment en vue d'une protection par brevet ;
- les données à caractère personnel ;
- les données classifiées ou les données rendues confidentielles par la loi ou le règlement ;
- plus généralement, toute donnée dont la divulgation pourrait compromettre la stratégie ou les intérêts du BRGM ou ayant un caractère de souveraineté nationale, notamment dans le domaine des ressources.

Faciliter l'accès et la traçabilité des données brutes

Les données brutes sont les données non interprétées issues directement de mesures ou d'observations sur site ou expérimentales dans le cadre des activités de recherche ou d'expertise. Leur traçabilité et leur accessibilité sont la garantie d'une science ouverte reproductible et vérifiable.

Un investissement significatif sera mené pour établir une typologie et les outils de gestion associés à l'échelle de l'établissement des différents types de données générées au sein du BRGM avec la mise en place de standards et procédures internes autour du mode d'acquisition (observation, expérimentation, ...) et la nature de la donnée brute (observations de terrain, eau, sol, roches, gaz, mesures de terrain, cahiers de laboratoires...). Il s'agira d'uniformiser des protocoles de gestion et traçabilité des données brutes pour chaque catégorie identifiée dans la typologie ainsi que des règles permettant le transfert d'une catégorie à une autre.

Cela implique également d'initier un plan sur le référencement et la conservation patrimoniale à long terme des échantillons collectés par le BRGM ou qui lui sont confiés par des tiers.

Mettre en place un plan de gestion des données et des codes pour l'ensemble des projets

En matière de gestion des données de la recherche, le BRGM vise la généralisation, pour l'essentiel des projets, de la rédaction d'un Plan de gestion de données (PGD) dans l'esprit de "FAIR data" qui formalise la manière dont seront gérées les données selon différents points de vue : stockage, permission d'accès, sécurité etc. Le PGD est un formalisme qui sera intégré au système de management de la qualité, certifié ISO 9001, en lien avec le processus de réalisation des projets. Dans un souci de cohérence, un modèle de PGD générique sera développé et mis à disposition des différents projets...

Dans le cas spécifique des codes sources, un plan de gestion de logiciel (SMP) sera adapté pour prendre en compte les spécificités des codes sources scientifiques : pérennité de l'usage, obsolescence technologique, ergonomie, design, niveau de TRL (Technology readiness level), financement...

Faciliter la réutilisation des données

Issues d'observations, d'expérimentations ou modélisations/simulations, produites dans le cadre de projets ponctuels ou récurrents, l'ensemble des données produites doit systématiquement faire l'objet d'une description et d'une sauvegarde dans les plateformes numériques du BRGM afin d'en permettre des réutilisations ultérieures. Ces réutilisations sont à la fois source d'économies et d'amélioration de la connaissance (machine learning...).

S'inscrire dans la démarche de plateforme numérique intégrative

Cette stratégie s'appuie notamment sur le centre de données régional DataCentre dont le BRGM assure le pilotage, afin de constituer une infrastructure unique de plateformes numériques intégratives permettant de capitaliser l'ensemble des données de l'établissement et de certains partenaires.

Le BRGM favorisera le dépôt conjoint des données aux publications scientifiques externe ou interne, notamment par l'intégration de cette démarche dans le cadre du processus qualité de réalisation des projets. Il accompagnera également les scientifiques dans l'utilisation des outils de gestion de données.

La mise en œuvre de cette politique générale s'appuiera sur une plateforme numérique intégrative unique. Des actions de refonte et d'associations de bases de données existantes seront réalisées pour intégrer l'ensemble de nos connaissances (PGD) générées dans les projets scientifiques. Ces bases seront outillées afin de respecter les principes FAIR. Ces données seront soit hébergées par l'infrastructure DataCentre, soit référencées dans le cata-

Le BRGM investira dans le développement de nouvelles approches et outils de la data science, s'appuyant également sur son expertise de modélisation numérique traditionnelle, pour produire de nouvelles données et services au service de ses missions.

logue BRGM de métadonnées et accessibles via leur DOI.

En matière d'outils de simulation, un environnement de développement numérique unique (forge BRGM) sera développé afin de mettre à disposition un ensemble d'outils génériques (open source et/ou propriétaires) et facilitant l'interface avec les données référencées dans le catalogue.

Une mise en cohérence dans un outil unique sera mis en place pour appuyer la création de modèles numériques (conventionnels comme IA), en intégrant progressivement l'ensemble des fonctionnalités et s'appuyant autant que possible sur du code open source.

En matières de services numériques (API), leur développement se fera en conformité avec les standards édités, afin de garantir leur généricité (réutilisabilité) mais en adaptant leur contenu à chaque besoin.

Un travail de transformation numérique des plateformes expérimentales et des laboratoires est également à mener afin de faciliter la production de données respectant les objectifs de cette politique.





Data Terra est une e-infrastructure de recherche de données et services pour le système Terre, le BRGM acteur des données des géosciences est fortement impliqué sur les pôles de données Form@Ter et Theia.

© BRGM

Développer les outils et la pratique de valorisation des données

Le BRGM investira dans le développement de nouvelles approches et outils de la data science, s'appuyant également sur son expertise de modélisation numérique traditionnelle, pour produire de nouvelles données et services au service de ses missions.

S'inscrire dans une démarche de sobriété numérique

Le BRGM est engagé dans une démarche de réduction de l'impact énergétique de son infrastructure DataCentre. Une démarche raisonnée de gestion des actifs de données (éviter la duplication, détruire des données non pertinentes ou obsolètes, développer des services numériques éco-conçus...) permettra de contribuer à la sobriété numérique. Le BRGM s'engagera également dans une démarche de réduction de l'impact énergétique de son infrastructure DataCentre.

Transformer la gouvernance et les pratiques de la science ouverte et maîtriser les risques potentiels à l'ouverture

La gouvernance

La politique de science ouverte implique à la fois des évolutions techniques, des pratiques de l'action scientifique et des changements culturels. Cela nécessite pour le BRGM de structurer un « écosystème » d'acteurs pour prendre en charge cette problématique dans sa globalité et sa mise en œuvre.

Aux côtés d'autres partenaires, au niveau national ou international, le BRGM participe au développement des infrastructures numériques indispensables à la gestion, au partage et à l'analyse des données et des connaissances.

Un administrateur de la science ouverte et des données scientifiques (Chief data officer) sera nommé et en charge de la mise en place de la politique de science ouverte et de management des données scientifiques. Il aura également une mission de coordination et de suivi sur l'ouverture et la protection des données.

Il s'appuiera sur une cellule de gouvernance des données scientifiques regroupant l'ensemble des directions concernées de l'établissement et traitera à la fois des données brutes et des données interprétées. Cette cellule composée de référents « données scientifiques » pour chaque direction sera constituée en proximité directe des équipes de recherche. Elle aura pour rôle la mise en place des processus internes, la rédaction de fiches techniques, les actions de formation et sensibilisation.

Un comité de pilotage de la politique de science ouverte, présidé par la Direction générale, sera le garant de la mise en œuvre de cette politique. Toutes les directions impli-

quées par la politique de science ouverte rapporteront sur une base annuelle des actions mises en place.

Une procédure de suivi des manuscrits avant soumission sera mise en place à la fois pour accompagner les auteurs en matière de choix pertinent des revues scientifiques au regard des critères de la politique de science ouverte, pour valider la pertinence du projet de publication eu égard aux enjeux de confidentialité ou de souveraineté, et pour assurer un bon suivi des préprints dans HAL BRGM. Il s'agira également d'analyser en concertation avec les auteurs de la maîtrise des risques en matière de propriété intellectuelle.

Faire évoluer l'analyse bibliométrique

Le suivi de la production scientifique couvrira à la fois les publications scientifiques externes, mais également les publications internes (rapports, etc.) et l'évolution des modalités d'archivage.

Il s'agira de faire évoluer l'analyse bibliométrique dans le sens d'une meilleure prise en compte des indicateurs de science ouverte (taux de publication en accès ouvert, taux de dépôt en archive ouverte, en les alignant notamment sur les critères du baromètre de science ouverte publié par le ministère en charge de la recherche et en lien avec les standards internationaux sur le sujet (Manifeste de Leiden...)).

Un baromètre bibliométrique du BRGM intégrant les principaux indicateurs de science ouverte sera publié annuellement et repris dans le rapport d'activité de l'établissement. En cohérence avec le spectre d'activité du BRGM, il couvrira à la fois les activités de recherche, mais aussi d'expertise en appui aux politiques publiques.

La formation et les compétences

L'adoption de la politique de science ouverte du BRGM et des pratiques associées sera atteint si l'ensemble des personnels concernés est accompagné et aidé dans cette démarche. La compréhension des enjeux de la science ouverte et des bénéfices collectifs associés est nécessaire. Il s'agira de développer une culture et une pratique de la gestion et du partage des données chez tous les acteurs du cycle de vie de la donnée scientifique (chercheurs, ingénieurs, techniciens, informaticiens, documentalistes, ...).

La mise en place de la politique de science ouverte à l'échelle de l'établissement va nécessiter d'importants efforts de sensibilisation et de formation de tous les personnels scientifiques, en particulier les doctorants et les métiers supports concernés. Cela constituera une composante importante du plan de formation de l'établissement dans le cadre du contrat d'objectif, de moyens et de performance (2023-27). Il s'agira d'accompagner les chefs de projets et les ingénieurs-chercheurs en matière de gestion de données scientifiques, d'analyse et de fouille de données.

Des actions de développement des compétences en matière de recherche ouverte, y compris les compétences

en intégrité scientifique seront également mises en place avec l'appui d'établissements partenaires.

Des actions spécifiques de communications interne et externe illustrant les bonnes pratiques, et la dynamique de l'entreprise sur ce thème seront également renforcées.

Renforcer l'ouverture de la science du BRGM vers la société civile et le grand public

L'ouverture des résultats de la science et des processus de recherche à tous les acteurs de la société est un élément déterminant pour le BRGM dans un contexte où il est nécessaire de renforcer la confiance dans la science et de contribuer au débat de société sur les problématiques relatives au sol et au sous-sol en se positionnant sur le champ de l'état des connaissances scientifiques. La politique de science ouverte du BRGM doit aussi viser à faciliter la diffusion de données, d'informations et de connaissances et des données issues de la recherche et de l'expertise dans des formats accessibles aux non spécialistes, vers les acteurs de la société civile et le grand public.

Le BRGM développera son action de communication scientifique et de médiation sur l'ensemble des enjeux de la stratégie scientifique avec un focus sur les thématiques à fort enjeux sociétal comme les risques naturels, la gestion de l'eau et des environnements pollués, les ressources minérales ou encore le sous-sol dans la transition énergétique.

Une démarche visant à encourager des projets pilotes de science participative sera mise en place.

S'inscrire dans une dynamique durable, européenne et internationale

Aux côtés d'autres partenaires, au niveau national ou international, le BRGM participe au développement des infrastructures numériques indispensables à la gestion, au partage et à l'analyse des données et des connaissances. Il s'associe aux démarches nationales : HAL pour les publications et les catalogues de données scientifiques (data.gouv.fr). Plus spécifiquement, le BRGM prend une part active dans la montée en puissance du hub national Data-Terra¹² en tant que centre de référence pour les données du sous-sol.

Sur le plan européen, le BRGM est membre de l'infrastructure de recherche européenne EPOS¹³ rassemblant les données et métadonnées d'observation de la Terre solide de 14 pays européens. Il est notamment l'un des deux centres d'hébergement des données et contribue au développement des services associés.

Le développement des standards de données numériques issues des géosciences pour favoriser notamment le catalogage et l'interopérabilité avec d'autres secteurs scientifiques et industriels sera l'un des axes de développement en matière de recherche et développement.

¹² DataTerra <https://www.data-terra.org/>

¹³ EPOS <https://www.epos-eu.org/des-algorithmes-et-des-codes-sources>

RÉFÉRENCES

The Open Science Training Handbook (2018)

- <https://book.fosteropenscience.eu/>

Premier Plan National pour la Science Ouverte (2018-2021)

- https://www.enseignementsup-recherche.gouv.fr/sites/default/files/content_migration/document/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

Deuxième Plan National pour la Science Ouverte (2021-2024)

- https://www.enseignementsup-recherche.gouv.fr/sites/default/files/content_migration/document/MEN_brochure_PNSO_web_1415209.pdf

Leiden Manifesto for research Metrics (2015)

- <http://www.leidenmanifesto.org/>

Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France (2018)

- https://www.ouvrirelascience.fr/wp-content/uploads/2018/11/Guide_Juridique_V2.pdf

GLOSSAIRE

DOI

Digital Object Identifier, littéralement « identifiant numérique d'objet ») est un mécanisme d'identification de ressources. Le but des DOI est de faciliter la gestion numérique à long terme de toute chose en associant des métadonnées à l'identifiant de la chose à gérer. Les métadonnées peuvent évoluer au cours du temps, mais l'identifiant reste invariant. C'est une alternative aux URI. Depuis 2012, le système d'identifiant numérique d'objet a été normalisé sous la forme de la norme ISO 263241. - <https://www.doi.org/>

ORCID

Open Researcher and Contributor ID soit « identifiant ouvert pour chercheur et contributeur ») est un code alphanumérique non-propretaire, qui permet d'identifier les chercheurs et auteurs de contributions académiques et scientifiques. L'ORCID donne une identité pérenne aux auteurs, selon une méthode semblable à celle utilisée pour les contenus avec les identifiants numériques d'objet - <https://orcid.org/>

idHAL

L'idHAL est un identifiant unique géré dans HAL. Il permet à un utilisateur authentifié, connu comme auteur dans HAL, de regrouper toutes ses publications, quelles que soient les différentes formes sous lesquelles son nom a pu être saisi et d'en choisir une par défaut - <https://doc.archives-ouvertes.fr/identifiant-auteur-idhal-cv/>

Embargo

Période pendant laquelle le document déposé dans une archive ouverte ne peut être en libre accès : l'éditeur se réserve l'exclusivité de la diffusion pendant cette période. Si l'éditeur impose une durée d'embargo, il est possible de déposer le fichier dès qu'on le souhaite mais il faudra respecter la durée demandée et l'appliquer dans HAL : la référence bibliographique sera visible en consultation mais le fichier ne le sera qu'à l'expiration du délai fixé.

APC

Article Processing Charges sont des Frais de Publication demandés par les éditeurs pour financer la diffusion en open access des articles scientifiques. Ils s'ajoutent aux autres frais de publication (nombre de pages, pages en couleur, figures, frais de soumission...) hérités de l'ère imprimée. Les APC constituent un modèle économique, également connu sous le nom de modèle «auteur-payeur», utilisé par certaines revues en open access.

HAL

Hyper Article en Ligne. L'archive ouverte pluridisciplinaire HAL, est l'infrastructure nationale destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche, publiés ou non, et de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. Administrée par le Centre pour la Communication Scientifique Directe (CCSD), HAL est le portail d'archives ouvertes le plus important en France. - <https://hal.archives-ouvertes.fr/>

Publication scientifique interne

Publication à contenu scientifique de recherche ou d'expertise revue dans le cadre du contrôle qualité interne au BRGM, référencé et diffusé par les moyens propres du BRGM via le portail InfoTerre (rapports référencés, cartes, guides techniques, etc.).

Publication scientifique externe

Publication scientifique revue par les pairs et éditée dans le cadre d'un processus externe au BRGM, communément : journaux scientifiques, cartes, actes de colloques, etc.

Plan de gestion des données (PGD)

Document formalisé, établi au démarrage d'un projet de recherche, et qui décrit la façon dont les données produites seront obtenues, documentées, utilisées, gérées et disséminées, pendant la phase de recherche et une fois le projet terminé.

Revue hybride

Revue accessible que sous abonnement, mais l'auteur est libre d'y publier ses travaux en open access, moyennant versement de frais de publication supplémentaires APC.

Green Open Access

Ou "voie verte" est la voie de l'auto-archivage ou dépôt par l'auteur dans une archive ouverte - <https://scienceouverte.couperin.org/la-voie-verte-2/>

Gold Open Access

Ou "voie dorée" concerne des revues ou ouvrages nativement en open access, dès leur publication - <https://scienceouverte.couperin.org/la-voie-doree-2/>

EPOS ERIC

European Plate Observing System - European Research Infrastructure Consortium est une infrastructure de recherche européenne distribuée en charge de l'intégration de données, produits de données, services et installations techniques issus de communautés thématiques multiples des Sciences de la Terre, à partir d'infrastructures de recherche en Europe. La mission d'EPOS est d'intégrer ces infrastructures de recherche nationales et transnationales existantes et de tirer parti des nouvelles opportunités offertes par l'informatique pour surveiller et comprendre le système dynamique et complexe de la Terre solide. - <https://www.epos-eu.org/epos-eric>

DataTerra

Infrastructure de recherche dédiée aux données d'observation de la Terre. Créée en 2016, elle est fondée sur quatre pôles de données et services correspondant à chacun des grands compartiments du système Terre : AERIS - Atmosphère, ODATIS - Océan, THEIA - Surfaces continentales, FORMATER - Terre solide - <https://www.data-terra.org/>



Géosciences pour une Terre durable

brgm

SIÈGE - CENTRE SCIENTIFIQUE ET TECHNIQUE

3, avenue Claude-Guillemain
 BP 36009 - 45060 Orléans Cedex 2 - France
 Tél.: (33) 2 38 64 34 34 - Fax: (33) 2 38 64 35 18

www.brgm.fr

