

Ouverture des données de recherche dans le domaine académique suisse : outils pour le choix d'une stratégie institutionnelle en matière de dépôt de données

[Marielle Guirlet](#), Diplômée du Master en Sciences de l'Information HEG, HES-SO (2020) et consultante

Nombre de mots:
14200

Ouverture des données de recherche dans le domaine académique suisse : outils pour le choix d'une stratégie institutionnelle en matière de dépôt de données

Introduction

Depuis les années 1990, l'apparition d'Internet permet la mise à disposition en ligne de contenus numériques sans limite matérielle de reproductibilité. L'explosion du volume des données digitales (« *data deluge* », « *digital deluge* » ; Pryor 2012 ; Pinfield, Cox et Smith 2014 ; Blumer et Burgi 2015) associée à une augmentation de leur complexité et de leur hétérogénéité pose de nouveaux défis pour leur gestion et leur conservation.

Elle s'accompagne aussi de nouvelles opportunités. Dans le domaine spécifique de la recherche scientifique, le mouvement Open Science pose les bases d'un nouveau fonctionnement de la recherche, basé sur la collaboration. D'abord concentré sur la mise à disposition gratuite et publique des publications scientifiques (la composante Open Access de l'Open Science), il englobe maintenant les autres produits de la recherche, dont les données. Les motivations sous-jacentes de l'ouverture des données sont d'améliorer l'efficacité de la recherche et d'augmenter sa transparence, de faciliter les recherches transdisciplinaires et l'innovation. Elle doit aussi permettre à tout public d'accéder à ce qui a été financé par l'argent public (Foreign Commonwealth Office 2013, Amsterdam Call for Action on Open Science 2016, The Concordat Working Group 2016, Sorbonne declaration on research data rights 2020).

Les chercheurs et les chercheuses peuvent s'engager dans ce mouvement par conviction personnelle. Ils doivent aussi se conformer à des recommandations ou des exigences de la part de leur institution de rattachement, des agences de financement de la recherche ou éventuellement des éditeurs d'articles scientifiques leur demandant de publier leurs données.

Le dépôt de données de recherche (celles-ci sont aussi parfois notées DR dans la suite) est un instrument essentiel de cette démarche. Il permet aux producteurs et productrices de données de partager celles-ci et de les archiver. Il permet à de possibles futurs utilisateurs et utilisatrices de les découvrir et d'y accéder.

Depuis l'ouverture des premiers dépôts (par exemple ICPSR pour les données quantitatives en sciences sociales, ouvert en 1962 (ICPSR [sans date])), de multiples autres dépôts se sont créés. Ils se distinguent par leur finalité principale (donner accès aux données ou préserver les données), la communauté de chercheurs et chercheuses à laquelle ils s'adressent, le fait qu'ils soient rattachés à une institution ou à d'autres formes d'organisations, les technologies sur lesquelles ils s'appuient, ou encore les services qu'ils offrent.

Devant cette diversité, et compte tenu des particularités locales de chaque institution et de ses données de recherche (la discipline de recherche, l'échelle de l'institution, l'existence d'une infrastructure de stockage ou pas), il n'existe pas de dépôt « *one size fits all* » répondant de manière certaine et exhaustive aux besoins de l'institution. Les chercheurs et les chercheuses, pour répondre correctement aux exigences qui leur sont imposées, peuvent ressentir le besoin d'être conseillé-e-s pour sélectionner le dépôt le plus adapté à leur type de données, à la culture et aux pratiques de leur discipline de recherche. Les institutions, de leur côté, doivent décider d'une stratégie : vers quel(s) dépôt(s) orienter les chercheurs et les chercheuses ? Faut-il créer un nouveau dépôt au risque de multiplier encore les offres possibles ? Faut-il améliorer un dépôt existant ? Quelles fonctionnalités le dépôt doit-il avoir, quels services doit-il proposer ?

Cette problématique a été étudiée en détail dans le cadre d'un travail de Master of Science HES-SO en Sciences de l'information, à la HEG de Genève, Haute Ecole de la HES-SO. Ce travail a été encadré par le Prof. Dr. René Schneider et a été effectué entre mars et août 2020. Ce travail avait pour objectif d'élaborer des outils pour aider les institutions suisses de la recherche publique à définir leur stratégie de soutien aux chercheurs et aux chercheuses pour le partage public de leurs données sur un dépôt. Cet article reprend certains résultats de ce travail, et à ce titre, contient des éléments, des citations et des références du mémoire correspondant (Guirlet 2020).

Après la présentation de la méthodologie et de la démarche, sont ici abordés le contexte de l'Open Science et les exigences posées aux chercheurs et aux chercheuses du milieu académique suisse pour l'ouverture de leurs données de recherche. La partie suivante est consacrée à ce qu'est un dépôt de données de recherche ouvertes, et aux concepts et aux outils importants pour la mise en œuvre de l'ouverture des données sur un dépôt. En prenant aussi en compte des critères d'évaluation de la qualité, une grille de description complète des dépôts de données de recherche ouvertes est alors élaborée. Les principaux résultats de l'étude de la stratégie de neuf institutions académiques suisses, des dépôts qu'elles utilisent et qu'elles recommandent, sont ensuite discutés. Ceci permet de dresser un panorama de ses dépôts de données de recherche ouvertes et d'identifier des bonnes pratiques à partir desquelles sont formulées des recommandations. S'appuyant sur les résultats précédents, les outils proposés aux institutions pour le choix de leur stratégie sont présentés en détail : un vade-mecum permettant le recueil d'informations utiles pour entreprendre la démarche, un guide décisionnel, et des informations complémentaires et des recommandations.

Méthodologie de recherche

Dans un premier temps le travail s'est fait sous forme de recherche théorique, avec la revue de littérature à propos des concepts (données de recherche, dépôts) et des aspects-clés du sujet : le contexte d'ouverture des données de recherche et le rôle des dépôts dans le processus de la recherche. Dans le but de caractériser un dépôt de DR ouvertes et d'élaborer un modèle de

description, cette revue de littérature a également couvert les fonctionnalités et les services usuels d'un dépôt, la mise en pratique des services FAIR et les approches d'évaluation de la qualité d'un dépôt.

Lui a succédé une phase de recherche exploratoire et descriptive, pour la revue de l'existant, afin d'identifier et d'analyser les cas à étudier (les institutions académiques et les dépôts utilisés et recommandés). Cette phase s'est appuyée sur la consultation des sites Internet des dépôts et des pages sur la GDR des sites institutionnels.

La troisième phase a englobé la réflexion pour l'analyse et la comparaison de cet existant, la synthèse des résultats et la production des outils décisionnels.

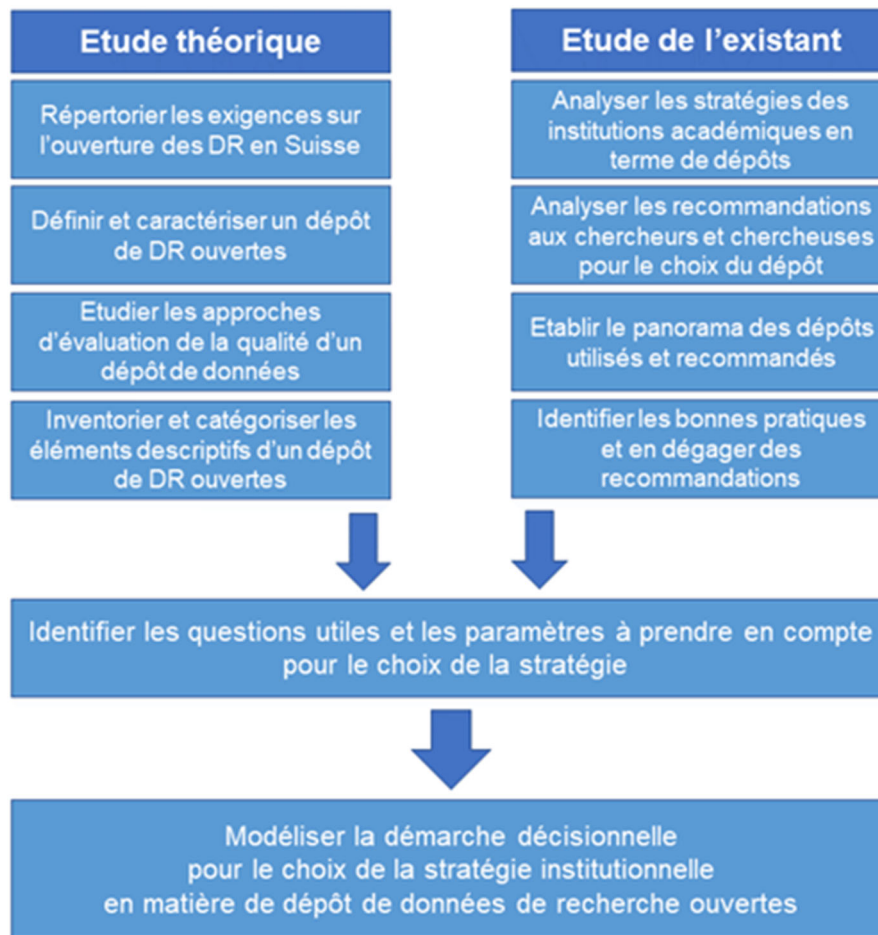
De l'étude théorique à la production des outils finaux, la progression du travail a été jalonnée de plusieurs sous-objectifs associés à des questions de recherche et donnant lieu à des résultats intermédiaires (Tableau 1).

Tableau 1 : Objectifs de recherche, questions de recherche et résultats obtenus

Objectif	Question de recherche associée	Résultat
Répertorier les exigences sur l'ouverture des données de la recherche publique en Suisse	Quelles exigences ont les principaux financeurs de la recherche suisse sur l'ouverture des données ? Quelles autres directives peuvent exister ?	Traduction de ces exigences extérieures en contraintes sur les dépôts de données
Définir un dépôt de données de recherche ouvertes et identifier ses fonctionnalités les plus importantes	Qu'est-ce qu'un dépôt de données de recherche ouvertes ? Comment se caractérise-t-il ? Quelles fonctionnalités peut-il avoir ?	Inventaire des outils et services d'un dépôt nécessaires pour des données de recherche ouvertes
Etudier les approches d'évaluation de la qualité d'un dépôt	Quels sont les critères de qualité possibles pour un dépôt de données de recherche ?	Analyse des critères de qualité d'un dépôt de données de recherche
Etablir un modèle de description des dépôts de données de recherche ouvertes et catégoriser les éléments descriptifs	Quels sont les éléments descriptifs pouvant décrire de façon exhaustive un dépôt de données de recherche ouvertes ?	Mise au point d'une grille de description d'un dépôt de données de recherche ouvertes pour l'évaluation et la spécification d'un dépôt
Analyser et faire un état des lieux de la stratégie actuelle des institutions académiques suisses en matière d'ouverture des données et de choix de dépôt (dépôts utilisés et recommandés, bonnes pratiques)	Quelle est la stratégie actuelle des institutions académiques suisses en matière d'ouverture de leurs données de recherche ? Quelles bonnes pratiques peuvent être retenues ?	Production du panorama des dépôts utilisés et recommandés par neuf institutions académiques suisses, identification de bonnes pratiques et formulation de recommandations
Identifier les questions utiles et les paramètres pour une institution de recherche impactant le choix de sa stratégie en matière de dépôt de données : politique institutionnelle et engagement OS, budget, infrastructures de stockage et services existants, ...	De quelles informations une institution de recherche a-t-elle besoin pour choisir le dépôt de données de recherche ou la stratégie qui répondra au mieux à ses besoins et à ses contraintes ?	Elaboration du vademecum
Concevoir le processus de choix de la <u>stratégie</u> ; imaginer les étapes et leur enchaînement logique	Sur quelles décisions une institution doit-elle baser sa stratégie en termes de dépôt de données de recherche ?	Elaboration du guide décisionnel

La démarche globale est schématisée sur la Figure 1.

Figure 1 : Schématisation de la démarche



Contexte

Les deux principaux financeurs de la recherche académique suisse, la Commission Européenne et le FNS, imposent des conditions sur les données de recherche des projets qu'ils financent. Dans le cadre du projet pilote ORD Pilot (Open Research Data (ORD) – the uptake in Horizon 2020 2016), pour certaines thématiques de recherche pendant la période 2014-2016, les chercheurs et chercheuses postulant pour le financement de projets de recherche auprès de la Commission Européenne doivent fournir un plan de gestion des données de la recherche (Data Management Plan ou DMP). Ils doivent s'engager dans ce document à partager autant que possible les données issues du projet ou associées aux publications, et y spécifier les conditions de conservation, de documentation et de partage de ces données (European Commission, Directorate-General for Research & Innovation 2016, p.6). Ces obligations sont étendues en 2017 à tous les projets (European Commission, Directorate-General for Research & Innovation 2017, p.8). En pratique, les données de recherche doivent être versées sur un dépôt, de préférence certifié, et être conformes aux principes FAIR (Wilkinson 2016 ; European Commission, Directorate-General for Research & Innovation 2016, p.7).

Le FNS, quant à lui, demande à ce que les données de recherche des projets qu'il finance soient partagées, à moins de « clauses légales, éthiques, de copyright, de confidentialité ou autres », et qu'elles soient déposées avec des métadonnées dans des archives publiques, « dans des formats accessibles et réutilisables sans restriction » (FNS [sans date]a). Les données

doivent aussi être rendues conformes aux principes FAIR. Le dépôt utilisé doit être non-commercial (FNS 2020). Comme pour H2020, les demandes de financement de projets auprès du FNS doivent s'accompagner d'un DMP détaillant en particulier les conditions de partage des données. Ce DMP a été rendu obligatoire par le FNS en octobre 2017 (FNS 2017).

Certains éditeurs scientifiques, de leur côté, exigent que les données sous-tendant les publications soient versées sur un dépôt, dans un but de transparence de la recherche publiée (voir par exemple Nature 2016, PLOS ONE [sans date]a, Springer Nature [sans date]a). Sur certains dépôts, ces données sont dans un premier temps partagées seulement avec les reviewer et les éditeurs pendant la phase de revue de la publication, puis elles sont rendues accessibles publiquement une fois que la publication est acceptée (PLOS ONE [sans date]a).

Dans ce contexte, les institutions de recherche accompagnent leurs chercheurs et chercheuses dans la gestion de leurs données pendant l'ensemble du processus de recherche, depuis la planification et l'écriture du DMP jusqu'au versement des données sur un dépôt. Elles encouragent ainsi l'usage de bonnes pratiques, favorisent la mise en pratique de leur politique institutionnelle pour la GDR et leur engagement en Open Science, et aident leurs chercheurs et chercheuses à se conformer aux exigences des financeurs de leur recherche et des éditeurs de leurs publications. En ce qui concerne les dépôts de données de recherche, elles émettent des recommandations sur le choix du dépôt le plus adapté sur lequel verser les données.

Données de recherche ouvertes et dépôts de DR ouvertes

Des données sont ouvertes si on peut y accéder librement et si elles peuvent être utilisées, modifiées et partagées librement, dans n'importe quel but et par n'importe qui, sous condition d'en attribuer l'origine à leurs auteur-e-s (The Concordat Working Group 2016 ; Hodson, Jones et al. 2018).

Selon le rapport Science as an Open Enterprise (The Royal Society 2012), les données de recherche ouvertes doivent être « *assessable and intelligible* » : on doit pouvoir évaluer leur qualité, leur pertinence et leur utilité pour envisager de les réutiliser. On doit pouvoir les interpréter et les comprendre pour les réutiliser correctement. Cette notion d'interprétation contextuelle est d'ailleurs présente dans la définition des données de recherche par The Consultative Committee for Space Data Systems (2012, p. 1–10) :

« A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. »

Ces données ouvertes sont accessibles publiquement, généralement sur un dépôt de données, à partir duquel on peut les chercher, les extraire, et les télécharger (Johnston et al. 2017, p.3), sans restriction de copyright, de droits de brevets ou d'autres mécanismes de contrôle (Jong et al. 2020).

D'après l'Université de Genève ([sans date]) :

« un dépôt de données (data repository) est un terme général utilisé pour désigner un lieu pour le stockage des données ».

L'Université de Boston citant le Registry of Research Data Repositories ([Re3data](#)) donne la définition suivante du dépôt de données de recherche (Boston University Data Services [sans date]):

« a subtype of a sustainable information infrastructure which provides long-term storage and access to research data that is the basis for a scholarly publication »

mentionnant ainsi les deux objectifs de ce type d'infrastructure, le stockage et l'accessibilité des données de recherche sur le long terme.

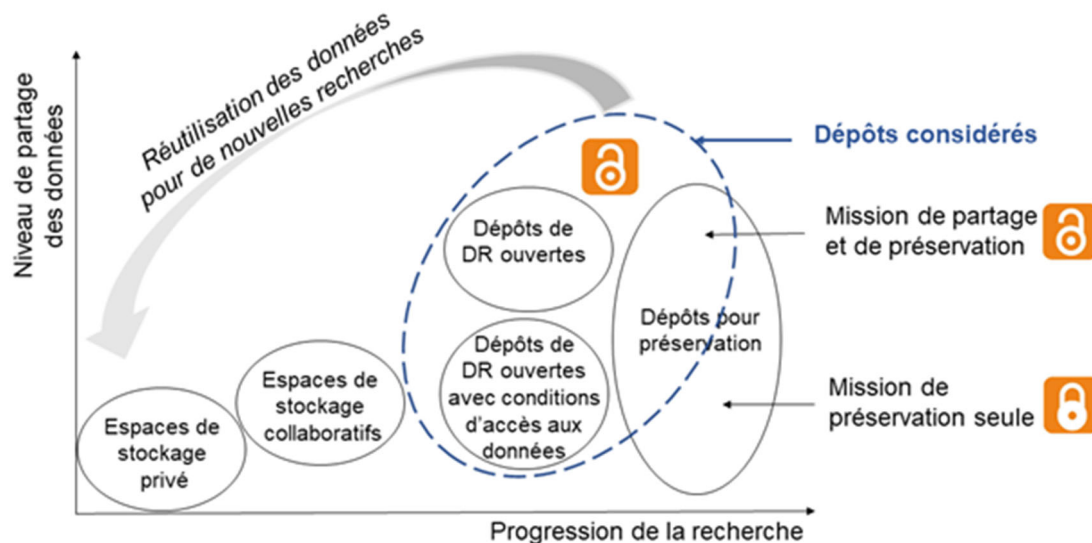
Avec sa définition du dépôt, le Data Curation Network souligne aussi l'importance de la notion de services (Johnston et al. 2016) :

« A digital archive that provides services for the storage and retrieval of digital content »

Selon le modèle du Data Curation Continuum de Treloar, Groenewegen et Harboe-Ree (2007), le dépôt intervient à différents moments du processus de la recherche. Un dépôt est utilisé pour partager les données d'un projet de façon publique vers la fin de celui-ci. Il est aussi utilisé pour préserver les données à long terme après la fin du projet (Treloar, Groenewegen et Harboe-Ree 2007 ; Treloar 2012).

Les dépôts ayant pour mission principale le partage des données répondent à l'objectif de donner accès à celles-ci le plus vite possible (approche plus orientée Open Science). Les dépôts ayant pour mission principale la préservation des données répondent à l'objectif de préserver celles-ci le plus longtemps possible (approche plus orientée archivage), tout en assurant pour certains d'entre eux l'accessibilité des données. Dans la suite, on prendra en compte les deux types de dépôts, pour autant que les dépôts de type préservation garantissent aussi l'accès aux données. Certains dépôts mettent à disposition les données de façon ouverte par défaut, mais permettent aussi aux dépositaires de restreindre l'accès à leurs données (par exemple Zenodo ([zenodo \[sans date\]](#)) et Harvard Dataverse ([Dataverse project \[sans date\]](#)). On prend également en compte ce type de dépôt. La Figure 2 illustre les dépôts de données utilisés au cours du processus de recherche et ceux auxquels on s'intéresse dans le cadre de cette étude.

Figure 2 : Types d'espaces de stockage et de dépôts de données, et niveau de partage selon la progression de la recherche



Les dépôts de données prennent diverses formes selon cette mission principale (accessibilité à court terme ou préservation sur le plus long terme), mais aussi selon la communauté à laquelle ils s'adressent (dépôt disciplinaire ou généraliste), leur entité de rattachement (dépôt institutionnel ou pas), leur business model (commercial ou à but non lucratif), le type de technologies qu'ils utilisent (open source ou propriétaires). Comme on va le voir dans la suite, les dépôts se distinguent aussi les uns des autres selon les fonctionnalités et les services qu'ils fournissent, et la façon dont ceux-ci sont déclinés. On peut évoquer par exemple le niveau de curation appliqué aux données déposées, le type d'identifiant pérenne attribué aux données, les standards utilisés pour les métadonnées, la possibilité de mettre les données sous embargo, et évidemment le niveau d'ouverture des données. Certains de ces services peuvent dépendre éventuellement de la catégorie de dépôt (disciplinaire ou généraliste).

Modèle de description d'un dépôt de DR ouvertes

Outils et services

Selon les définitions données plus haut, le dépôt de DR peut donc se définir comme une infrastructure informatique qui stocke des données et les rend accessibles sur le long terme. Ces deux finalités font partie des principes FAIR sur les données (Wilkinson et al. 2016), qui, on le rappelle, selon le FNS et H2020, doivent être appliqués par les dépôts utilisés pour partager ses données.

Les principes A (Accessibilité) et R (Réutilisation) de FAIR correspondent directement aux objectifs du dépôt de rendre et de maintenir les données accessibles et réutilisables. Pour appliquer le principe F (Findability ou Découvrabilité), le dépôt doit rendre les données recherchables et trouvables par des humains et par des machines. L'application du principe I (Interopérabilité) permet l'échange efficace de contenu (données, métadonnées) entre chercheurs et chercheuses, entre institutions, entre systèmes, et par des machines aussi bien que par des humains, pour une utilisation la plus large possible (Swiss National Science Foundation [sans date]).

Comme noté par Hodson, Jones et al. (2018), les principes FAIR et l'ouverture des données (aboutissant à des données librement utilisables, modifiables et partageables, selon les définitions vues plus haut) ont en commun l'objectif ultime de contribuer à la réutilisabilité des données. Néanmoins, des données accessibles selon le principe A de FAIR ne sont pas forcément ouvertes. Pour autant que cela n'aille pas à l'encontre de restrictions légitimes à leur ouverture (concernant en particulier les données personnelles et les données sensibles), les données hébergées par les dépôts, en plus d'être FAIR, doivent donc être rendues ouvertes.

Les quatre grands principes FAIR et l'ouverture des données sont mis en pratique par les dépôts à l'aide de services et d'outils spécifiques (FNS [sans date]^b, Swiss National Science Foundation [sans date], Perini 2019).

Identifiants uniques pour les données : un identifiant unique permet de trouver, de citer et de tracer les données auxquelles il est assigné. On peut l'utiliser afin de citer les données utilisées pour obtenir les résultats présentés dans une publication, afin de se référer aux données d'origine lorsqu'on décrit des données secondaires, ou encore dans la liste de ses jeux de données ou sur son profil ORCID pour la description de ses activités de recherche. Les identifiants uniques pour les données doivent être pérennes (PID) et globaux (c'est-à-dire non internes au dépôt). Le DOI, délivré contre rétribution par Datacite (Datacite [sans date]) est couramment utilisé. Les principaux autres identifiants uniques sont les suivants : ARK (utilisé par le dépôt DaSCH ([sans date])), Handle ou hdl (utilisé par B2SHARE EUDAT; Re3data.org 2017), PURL, URN et RRID (Digital Preservation Coalition [sans date], Swiss National Science Foundation [sans date]).

Métadonnées : la description des données et de leur contexte, essentielle comme on l'a vu pour une possible réutilisation des données, est assurée en grande partie par les métadonnées. En plus du contexte et de la provenance des données, celles-ci détaillent la structuration et le contenu des jeux de données (EPFL Library, Research Data Library Team [sans date]). Elles aident ainsi à découvrir les données, à y accéder, à connaître leurs conditions de réutilisation (informations utiles pour les utilisateurs et utilisatrices), et à les gérer (pour le dépôt).

La structuration des métadonnées sous une forme interprétable par les machines permet l'automatisation de leur traitement (lecture, recherche, extraction) (Johnston et al. 2018). Les dépositaires doivent alors fournir ces métadonnées par le biais de formulaires conformes à des schémas standards. On utilise fréquemment les schémas Dublin Core (Dublin Core 2020, Dublin Core Metadata Initiative 2020) et DataCite Metadata (DataCite Metadata Working Group 2019, DataCite [sans date]) pour les métadonnées descriptives. Dans le cas de dépôts spécialisés, des métadonnées supplémentaires peuvent être fournies suivant le schéma couramment utilisé dans le domaine de recherche ou dans la discipline (schéma DDI pour les sciences sociales par exemple). Certains dépôts (ou solutions techniques) proposent d'utiliser son propre standard pour les métadonnées ou d'étendre celui proposé par défaut (Dryad, B2SHARE EUDAT, OLOS, Figshare ; Guirlet 2020, Annexes 4 et 5).

Formats des fichiers : l'utilisation de formats conformes à des standards ouverts, disponibles publiquement et non propriétaires pour les fichiers de données et de métadonnées permet une utilisabilité à la fois par un plus grand nombre de chercheurs et chercheuses mais aussi sur le plus long terme. Les dépôts publient souvent à l'intention des dépositaires des listes de formats recommandés et acceptables⁽¹⁾. Certains dépôts assurent la migration des formats lorsque celui des fichiers versés n'est pas pérenne. Dans le cas d'une préservation à long terme, une veille régulière du risque d'obsolescence des formats déclenche le cas échéant des

actions préventives, dont cette migration des formats (Rosenthaler, Fornaro et Clivaz 2015 ; L'Hours, Kleemola et de Leeuw 2019). C'est le cas du dépôt suisse FORSbase (DARIS 2018, p.29). Sur les dépôts disciplinaires, la conformité des formats à des standards spécialisés permet aussi aux utilisateurs et utilisatrices de les manipuler et de les interpréter selon leurs pratiques usuelles.

Citation des données : la formule de citation des données est générée automatiquement à partir des métadonnées descriptives (avec Datacite par exemple, si elles sont conformes au schéma DataCite Metadata). Elle est fournie aux utilisateurs et utilisatrices en même temps que les données récupérées. Dans les conditions et les termes sur la réutilisation des données de certains dépôts, ces utilisateurs s'engagent à citer les données de la façon qui leur est suggérée (voir par exemple FORS [sans date]). Sur FORSbase, ceux-ci s'engagent en outre à informer le dépôt de toute publication basée sur la réutilisation des données, permettant ainsi le traçage de cette réutilisation (FORS [sans date]).

Conditions de réutilisation et restrictions d'accès : le dépôt doit préciser les conditions de réutilisation des données, au moyen d'une licence sur les données ou des termes du copyright. La définition des données ouvertes de CASRAI(2) est la suivante (CASRAI [sans date]b):

« Structured data that are accessible, machine-readable, usable, intelligible, and freely shared. Open data can be freely used, re-used, built on, and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike. »

Selon CASRAI, les contraintes pour la réutilisation des données sont donc au maximum d'attribuer l'origine des données à leurs auteur-e-s et de les partager à l'identique. Les licences Creative Commons correspondantes sont les licences CC0, CC BY, et CC BY-SA, la licence CC0 étant celle dont l'ouverture est la plus élevée (Creative Commons [sans date]). Dans la suite, on considérera que les dépôts proposant une ou plusieurs de ces trois licences, pour au moins une partie de leurs données, hébergent des données ouvertes(3).

Sur certains dépôts, les dépositaires ont la possibilité de restreindre l'accès à leurs données : accès sur demande, accès privé (groupe de personnes identifiées) ou mise sous embargo. Aucune de ces configurations ne convient pour des données ouvertes, mais des données à l'accès restreint peuvent facilement devenir des données ouvertes, à la fin de la période d'embargo ou si les dépositaires suppriment les restrictions à la fin du projet. Restreindre l'accès pour des données personnelles ou sensibles dans le but de respecter leur confidentialité n'est pas approprié (elles resteront toujours consultables par les gestionnaires du dépôt). Par contre, une fois anonymisées, les données personnelles et les données sensibles peuvent être publiées (EPFL Library [sans date], Université de Lausanne [sans date]a).

Curation des données et des métadonnées : selon le Data Curation Network, la curation des données facilite leur découvrabilité et leur récupération, et contribue à leur réutilisabilité dans le temps (Johnston et al. 2017). Une définition plus complète des objectifs de la curation, incluant aussi le maintien de la qualité et l'ajout de valeur, est donnée par Cragin et al. (2007):

« *Data curation is the active and on-going management of data through its lifecycle of interest and usefulness to scholarship, science, and education; curation activities enable data discovery and retrieval, maintain quality, add value, and provide for re-use over time.* »

En pratique, elle consiste d'abord à préparer les données qui ont été sélectionnées afin qu'elles remplissent les conditions d'accès imposées par le dépôt : nommage des fichiers et des dossiers, structuration du jeu de données, changement de format éventuel et autres. La curation a aussi pour rôle de rendre et de maintenir les données FAIR, et de vérifier et de contrôler leur qualité. Les activités correspondantes incluent la création de métadonnées, la préparation de documentation, la vérification, la validation et l'enrichissement des données et des métadonnées. Elle peut aussi vérifier la conformité des données ou les rendre conformes aux règles légales et aux normes éthiques. Pour un détail des activités de curation, on peut consulter Johnston et al. (2016, 2017, 2018) et Johnston (2017).

Certaines activités de curation sont plus appliquées à la préservation. Elles concernent le maintien de la qualité des données sur le plus long terme, le soutien à leur préservation et leur transformation si nécessaire avec notamment la migration des formats mentionnée ci-dessus (Data Curation Network [sans date]). Des efforts supplémentaires sont aussi à fournir pour surveiller l'évolution possible des pratiques et des besoins au sein de la communauté cible. Si besoin, la curation met en œuvre les mesures d'adaptation nécessaires, telles que l'utilisation de nouveaux formats et standards pour le dépôt, l'utilisation de métadonnées plus riches, ou la mise à jour de la documentation.

Sur certains dépôts, les tâches de curation sont assurées par des « data steward » (OLOS 2020a) ou par des « data curator » (Dryad [sans date]).

Services complémentaires : outre les grandes fonctions assurées par les dépôts de données (l'ingestion, le stockage et la gestion des métadonnées et des données, et leur mise à disposition), ceux-ci proposent fréquemment des services complémentaires facilitant la visibilité et la découvrabilité des données.

Le résultat du moissonnage automatique de métadonnées conformes au schéma Dublin Core est utilisé par des services d'agrégation, des portails et des moteurs de recherche (tels que Google Dataset Search et Elsevier DataSearch) et renforce ainsi la visibilité des données à l'externe. Le protocole OAI-PMH est l'un des protocoles standards et/ou ouverts permettant ce moissonnage. Le dépôt doit disposer de l'interface OAI et s'enregistrer comme fournisseur de métadonnées auprès du service OAI-Data Provider de l'Open Archives Initiative (OAI) (DINI 2011, p.14, p.35). Thomson Reuters' Data Citation Index moissonne aussi le contenu des dépôts pour tracer les citations de données dans la littérature (Rice et Southall 2016, p.118).

Le web sémantique avec les Linked Open Data relie des entités décrites de façon structurée selon le cadre RDF, qui fournit un modèle pour la représentation, l'échange et l'interconnexion des métadonnées. Convertir les métadonnées en RDF et les exposer permet de les connecter à d'autres entités du web, et ce faisant, d'augmenter la découvrabilité des données auxquelles sont associées ces métadonnées (Arlitsch et al. 2016, Rice et Southall 2016).

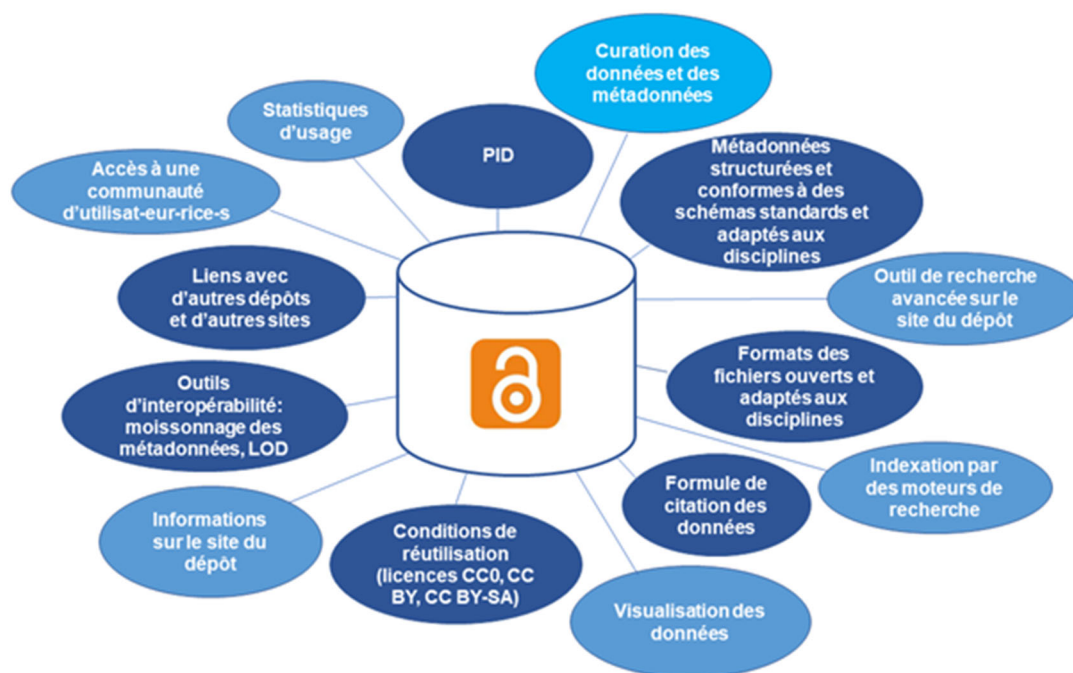
Avec certains dépôts, on peut faire le lien entre ses données déposées et sa page personnelle ORCID (par exemple avec Figshare [sans date]) ou avec les pages de ses projets (par exemple depuis Zenodo vers la page du projet sur le portail OpenAIRE ; zenodo [sans date]f). On peut parfois relier les publications et les données les sous-tendant (ETHZ Research Collection, ETHZ – ETH Bibliothek [sans date]), ou les données et le code qui a servi à les produire (zenodo [sans date]f), alors que ces éléments sont hébergés sur des dépôts distincts (par

exemple PLOS pour les publications, GitHub pour le code). Les dépôts permettant de faire ce type de liens reconnaissent les identifiants ORCID et GitHub. Ces fonctionnalités sont là encore un moyen de favoriser la visibilité des données.

Un outil de recherche avancée sur le site du dépôt est utile pour l'exploration et la découvrabilité des données. Des outils de visualisation donnent un premier aperçu du contenu des fichiers. L'information disponible sous forme de guide utilisateur et de rubrique FAQ renseigne sur le dépôt en général, sur les conditions d'hébergement ou sur les procédures de versement des données. Les réseaux sociaux, un forum ou un blog sont des moyens d'accéder à une communauté d'utilisateurs et d'utilisatrices du dépôt. Enfin, les métriques et les statistiques d'usage sur la fréquentation du dépôt rendent compte de son dynamisme et permettent d'évaluer si le dépôt est bien utilisé et reconnu par cette communauté. Quand ces métriques concernent les jeux de données eux-mêmes (nombre de vues et de téléchargements), elles donnent aux dépositaires une estimation de l'intérêt porté à leurs données.

Les outils et les services d'un dépôt de données de recherche hébergeant des données conformes aux principes FAIR et ouvertes sont schématisés sur la Figure 3.

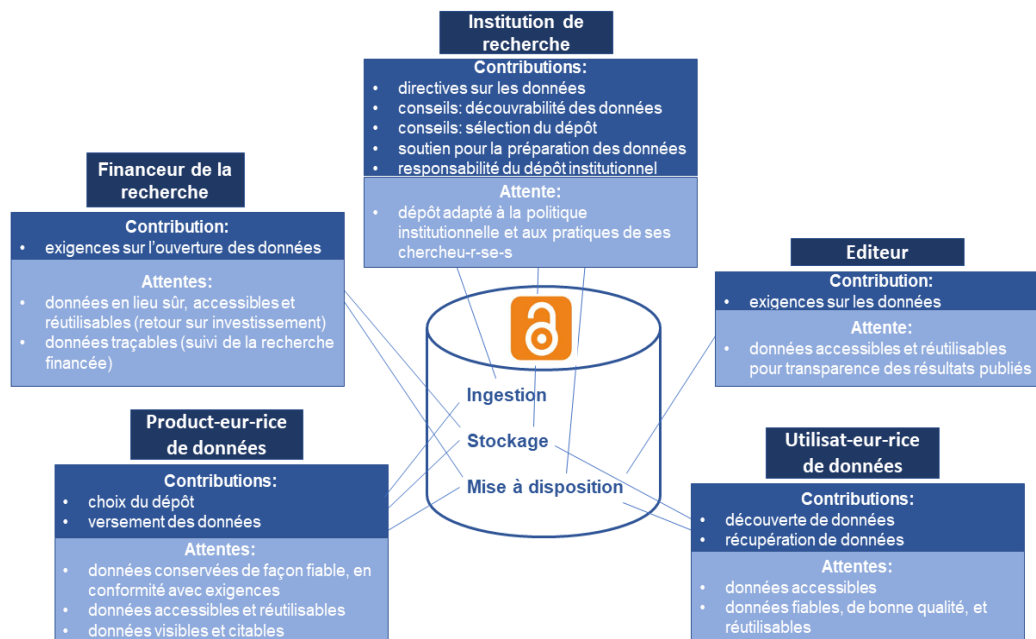
Figure 3 : Les outils et services d'un dépôt de données de recherche ouvertes



Critères de qualité

Les objectifs du dépôt de DR ouvertes sont donc de rendre et de maintenir les données accessibles et réutilisables à partir du moment du versement et dans le futur. Assurer que le dépôt est en capacité de remplir ces objectifs instaure la confiance des parties prenantes (présentées sur la Figure 4 avec leurs contributions et leurs attentes par rapport au dépôt de données). Cela leur garantit que les données sont en effet accessibles et réutilisables, mais aussi qu'elles sont conservées de façon sûre et qu'elles sont traçables (pour les agences de financement), qu'elles sont visibles et citables (pour les producteurs et productrices de données), et qu'elles sont fiables et de bonne qualité (pour les utilisateurs et utilisatrices).

Figure 4 : Parties prenantes du dépôt de DR ouvertes avec leurs contributions et leurs attentes



Depuis les années 1990, plusieurs démarches sur la base de normes et de standards ainsi que des certifications ont été élaborées pour évaluer de façon indépendante la qualité et la fiabilité d'un dépôt sur le long terme. Selon UK Data Archive, Standards of Trust (cité par l'Université d'Edimbourg, The University of Edinburgh 2019):

« The standards provide the basis of a framework by which different levels of trust of digital repositories can be demonstrated ».

Le cas échéant, le dépôt est reconnu « dépôt de confiance » (*trusted digital repository, trustworthy digital repository*).

L'approche d'évaluation de la qualité d'un dépôt la plus récente et la plus utilisée (Guirlet 2020, Tableau 14) est la Certification CoreTrustSeal mise au point en 2017 par un groupe de travail de la RDA, à partir de DSA et ICSU-WDS (Dillo et de Leeuw 2018 ; Corrado 2019 ; L'Hours, Kleemola et de Leeuw 2019). Les critères sont révisés tous les trois ans et la version la plus récente couvre maintenant la période 2020-2022 (CoreTrustSeal 2020a).

Plusieurs critères de CTS font référence aux standards du modèle OAIS (The Consultative Committee for Space Data Systems 2012) pour évaluer la fiabilité du dépôt à long terme (critères R9, R15 ; CoreTrustSeal 2020a). Le respect de cinq principes fondamentaux sur les données atteste que les données numériques sont archivées de façon durable. Selon ces principes, les données doivent être trouvables sur Internet, être accessibles en tenant compte de la législation en vigueur sur les informations personnelles et la propriété intellectuelle des données, être disponibles sous un format utilisable, être fiables et être référencables (Dillo et de Leeuw 2018). Enfin, un rôle important est donné à la qualité des métadonnées pour assurer la découvrabilité et l'accessibilité des données (ceci impliquant l'intervention de personnel qualifié ou la contribution d'expert-e-s externes), et à l'évaluation de la qualité de ces métadonnées. On pourrait d'ailleurs décrire de façon schématique la certification CTS comme une approche d'évaluation d'un dépôt englobant à la fois les exigences OAIS pour sa fiabilité

à long terme, les principes FAIR sur les données pour leur accessibilité et leur réutilisabilité, et des critères sur la qualité des données et des métadonnées.

Les critères de certification CTS sont organisés en trois catégories, selon le Tableau 2.

Tableau 2 : Catégories de critères CTS

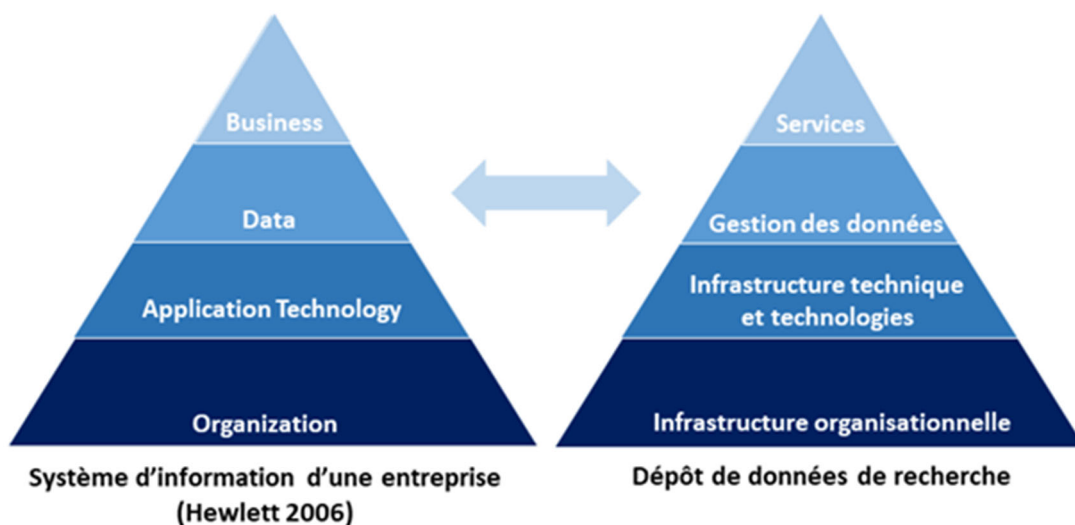
Catégorie	Nombre de critères	Détails
Infrastructure organisationnelle	6	Mission du dépôt, continuité d'accès, conformité avec les normes éthiques et de confidentialité, gouvernance, conseils d'expert-e-s, ...
Gestion de contenus numériques	8	Intégrité, authenticité, critères d'acceptation des données, documentation des procédures, plan de préservation, qualité des (méta)données, flux opérationnels, découvrabilité et référencement, réutilisation, ...
Technologie	2	Infrastructure technique et sécurité

(d'après CoreTrustSeal 2020b; CoreTrustSeal Standards and Certification Board 2020; Corrado 2019)

Elaboration du modèle

L'ajout à ces trois catégories de critères d'une nouvelle catégorie englobant les outils et les services d'un dépôt décrits précédemment permet de faire l'analogie avec la représentation par couches du système d'information d'une entreprise selon Hewlett (2006) (voir la Figure 5 ci-dessous). Dans cette représentation, l'infrastructure organisationnelle forme la base de la pyramide. Viennent ensuite l'infrastructure technique et les technologies, puis la gestion des données. Les services aux usager-e-s du dépôt forment la couche du sommet de la pyramide.

Figure 5 : Représentation par couches d'un dépôt de DR par analogie avec l'architecture du système d'information d'une entreprise



En pratique, dans le but de traduire cette conception du dépôt de DR ouvertes en une grille de description détaillée, plusieurs sources sont considérées pour identifier les éléments de cette grille :

- les critères d'évaluation de la qualité utilisés pour la certification CTS ;
- les outils et services d'un dépôt décrits plus haut et identifiés à partir de la revue de la littérature ; les éléments descriptifs correspondants ont été affinés grâce à une revue détaillée des dépôts généralistes, disciplinaires et institutionnels utilisés en Suisse (voir Guirlet 2020, Annexes 4 et 5).

On y ajoute des éléments supplémentaires permettant d'assurer la conformité aux exigences des financeurs de la recherche (e.g., dépôt non-commercial exigé par le FNS), au cadre légal, aux normes éthiques et disciplinaires en vigueur (gestion adéquate des données personnelles et sensibles), ainsi que des éléments relatifs à la qualité de l'expérience utilisateur ou utilisatrice (e.g., convivialité du site, facilité du versement, ...). La grille complète ainsi produite est donnée dans Guirlet (2020).

Cette grille de description peut être utilisée soit pour évaluer un dépôt existant (et aussi dans le but éventuel de l'améliorer), soit pour spécifier un dépôt à créer. La mise en œuvre satisfaisante de tous les aspects décrits par les éléments du modèle, que le dépôt existe déjà ou qu'il soit à l'état de projet, assure que ce dépôt est de qualité (car conforme aux critères CTS), qu'il héberge des données accessibles et réutilisables, conformes aux principes FAIR et ouvertes, et qu'il répond aux exigences des principaux financeurs de la recherche, ainsi qu'au cadre légal et aux normes sur la gestion des données personnelles et sensibles.

Panorama des dépôts utilisés et recommandés en Suisse

Dans le but d'établir un état des lieux des dépôts utilisés par les chercheurs et les chercheuses et recommandés par les institutions académiques en Suisse, la stratégie des plus engagées d'entre elles en termes d'ouverture des données de recherche (soit neuf institutions) a été passée en revue. Pour chaque institution en particulier, on a identifié, s'ils existent, le dépôt institutionnel de données ou de publications, le dépôt de données en projet, les dépôts disciplinaires développés dans le cadre d'un partenariat entre institutions, ainsi que les dépôts de données recommandés. Le détail de cette revue et les références bibliographiques associées sont donnés dans Guirlet (2020, Annexe 5). Elaborée sur la base de celle-ci mais en prenant aussi en compte l'ouverture récente d'OLOS (OLOS 2020b), la Figure 6 mentionne ces différents dépôts.

Figure 6 : Dépôts institutionnels et dépôts de DR recommandés pour neuf institutions académiques suisses ((p) : dépôt pour publications seulement; (p+d) : dépôt pour publications archivant aussi des DR ; disc. : disciplinaire)

<p>UNIGE</p> <p>Existants Archive ouverte (p) Yareta OLOS</p>	<p>Recommandations disc. avec Re3data Yareta Zenodo Dryad FORsbase OSF</p>	<p>UNIL</p> <p>Existants SERVAL (p) FORsbase la PlaTec</p> <p>En projet SWISSUBase</p>	<p>Recommandations Re3data avec checklist FNS Zenodo FORsbase la PlaTec</p>	<p>HES-SO</p> <p>Existants ArODES (p) OLOS</p>	<p>Recommandations top 10 par projets FNS Re3data, OpenAIRE, liste Nature FORsbase, DaSCH, GenBank OLOS, Zenodo, Dryad Figshare, autres FNS Yareta</p>
<p>EPFL</p> <p>Existant Infoscience (p)</p> <p>En projet ACOUA</p>	<p>Recommandations disc. ou com. avec Re3data Zenodo Dryad Figshare</p>	<p>ETHZ</p> <p>Existants Research Collection Collection (p+d) ETHZ Data Archive (préservation)</p>	<p>Recommandations selon directives Research Collection ETHZ Data Archive conseils de PLOS Re3data selon formats des données</p>	<p>UNIZ</p> <p>Existants Zora (p)</p> <p>En projet SWISSUBase</p>	<p>Recommandations liste FNS disc. disc. et FAIR avec Re3data liste dépôts disc. utilisés par les chercheurs et chercheuses de l'UNIZ</p>
<p>UNIBAS</p> <p>Existants edoc (p)</p> <p>DaSCH (Digital Humanities Lab)</p>	<p>Recommandations Zenodo Dryad disc. avec Re3data, OADirectory conseils de PLOS DaSCH FORsbase</p>	<p>UNIBE</p> <p>Existant BORIS (p+d) BerDA (préservation)</p> <p>En projet BORIS Research Data</p>	<p>Recommandations disc. avec Re3data, OADirectory conseils de PLOS Zenodo, Dryad, B2SHARE, Harvard Dataverse BORIS</p>	<p>UNIFR</p> <p>Existant edoc (p)</p>	<p>Recommandations FORsbase, DaSCH, GenBank Re3data, OpenAIRE, liste Nature Figshare, Zenodo, autres FNS</p>

(d'après les sources données dans Guirlet 2020, Annexe 5)

Cette revue met en lumière les différents cas de figure suivants selon les institutions (avec des recouvrements possibles).

- L'existence d'un dépôt de données institutionnel (Research Collection et Data Archive pour l'ETHZ ; BerDA pour l'Université de Berne) ;
- l'existence d'un dépôt à l'échelle cantonale (Yareta pour l'Université de Genève et pour la HES-SO), d'un dépôt à l'échelle nationale (OLOS, développé par l'Université de Genève et la HES-SO dans le cadre du projet DLCM) ou d'un dépôt disciplinaire auquel a contribué une entité de l'institution (DaSCH pour l'Université de Bâle ; FORsbase et la PlaTec pour l'Université de Lausanne) ;
- un projet de dépôt institutionnel (ACOUA pour l'EPFL) ou en partenariat avec d'autres institutions (SWISSUbase pour l'Université de Lausanne et l'Université de Zurich) ;
- un projet d'extension de l'archive institutionnelle de publications (BORIS Research Data pour l'Université de Berne) ;
- l'absence de dépôt institutionnel pour les données et pas de projet formulé pour en développer un (Université de Fribourg).

Ces institutions émettent toutes des recommandations pour le choix du dépôt de données de recherche (pour le détail, voir Guirlet 2020, Annexe 5). Celles-ci varient en fonction des disciplines de recherche au sein de l'institution, et de l'existence ou pas d'un dépôt institutionnel. Ces recommandations peuvent s'appuyer sur les exigences des agences de financement (FNS, H2020), et se référer aux conseils donnés par des éditeurs scientifiques (PLOS ONE [sans date]b, Springer Nature [sans date]b). Elles suivent aussi fréquemment les conseils du FNS, qui donne quatre exemples de dépôts généralistes répondant à ses exigences : Zenodo, Dryad, EUDAT et Harvard Dataverse (Swiss National Science Foundation 2017).

Du fait de la spécialisation et de l'expertise des dépôts disciplinaires, il est fréquent que les recommandations faites aux chercheurs et aux chercheuses orientent vers le dépôt disciplinaire adapté à leur spécialité avant un dépôt généraliste ou même avant le dépôt institutionnel. La position de l'Université de Berne sur ce point est bien marquée (Universität Bern, [sans date]a):

« Wherever possible, data should be deposited in disciplinary repositories. These are designed to meet the needs of the particular field, are aware of specific data formats and often also offer specific disciplinary metadata. »

L'utilisation de registres de dépôts tels que Re3data et ses filtres (Re3data.org [sans date]) pour la prise en compte d'autres critères est conseillée. Deux institutions renvoient aussi vers les pratiques d'une communauté : l'UNIZ fait référence aux dépôts utilisés par ses chercheurs et chercheuses (Universität Zürich, Hauptbibliothek 2019), et la HES-SO mentionne les pratiques des bénéficiaires de subsides FNS (HES-SO [sans date]).

Le Tableau 3 reprend les recommandations des neuf institutions pour le choix du dépôt.

Tableau 3 : Synthèse des recommandations des neuf institutions étudiées pour le choix du dépôt de données

Les recommandations d'une institution correspondent aux cellules marquées en bleu clair. Les recommandations des cellules marquées en bleu foncé sont celles d'une institution pour son propre dépôt institutionnel (marquage étendu à FORSbase et à la PlaTEC pour l'UNIL, à DaSCH pour UNIBAS et à Yareta et à OLOS pour UNIGE et la HES-SO).

	UNIGE	UNIL	HES-SO	EPFL	ETHZ	UNIZ	UNIBAS	UNIBE	UNIFR
Dépôts généralistes									
Zenodo	■	■	■	■		■	■	■	■
Dryad	■		■	■		■	■	■	■
OSF	■								
Figshare			■	■					■
B2SHARE EUDAT			■			■		■	■
Harvard Dataverse			■			■		■	■
Yareta (canton de Genève)	■		■						
OLOS (national)	■		■						
Dépôts disciplinaires									
Sans précision						■			
Disciplinaire avec Re3data	■				■		■	■	■
Disciplinaire et FAIR avec Re3data						■			
Disciplinaire et autres critères ⁴ avec Re3data				■					
FORSbase	■	■	■				■		■
La PlaTec		■							
DaSCH			■				■		■
GenBank			■						■
Dépôts institutionnels									
Research Collection					■				
ETHZ Data Archive ⁵					■				
BORIS ⁶								■	
Registres et bases de dépôts									
Re3data			■						■
OpenAIRE Explore			■						■
Open Access Directory							■	■	
Listes de dépôts conseillés									
Liste Springer Nature ⁷			■						■
Liste PLOS ONE ⁸					■		■	■	
Selon les pratiques de la communauté									
Dépôts utilisés par les collègues de l'institution						■			
Dépôts mentionnés dans les DMP pour le FNS			■						
Autres									
Critères de sélection de Science Europe ⁹			■						
Re3data selon la checklist FNS ¹⁰		■							

Bonnes pratiques observées

A partir de l'étude de la stratégie de ces institutions, on peut également dégager des exemples de bonnes pratiques sur plusieurs aspects.

Découvrabilité des données : parmi les ressources en ligne fournies, seules l'Université de Zurich et l'Université de Bâle donnent des conseils ou mentionnent des outils (tels que des registres et des moteurs de recherche) pour la recherche et la découvrabilité de données sur des dépôts (*Universität Basel [sans date]* ; Universität Zürich, Hauptbibliothek 2020). Les autres institutions, sans doute plus concentrées sur la conformité aux exigences des financeurs de la recherche, conseillent surtout sur la sélection d'un dépôt de données adapté. Etendre les conseils à la découvrabilité des données et à leur réutilisation est néanmoins essentiel pour donner sa pleine place au dépôt comme instrument de partage et de réutilisation des données.

Suivi des pratiques : on a mentionné précédemment que l'Université de Zurich, parmi ses recommandations pour le choix du dépôt, fournit une liste des dépôts fréquemment utilisés par ses chercheurs et chercheuses (Universität Zürich, Hauptbibliothek 2019). Pour toute institution, un suivi précis des dépôts utilisés et du nombre de jeux de données versés par dépôt peut lui fournir un état des lieux des pratiques de ses chercheurs et chercheuses. Ces informations sur les pratiques sont normalement fournies dans le DMP du projet. Rendre ce DMP obligatoire en interne permet ainsi à l'institution de suivre les pratiques institutionnelles. De même, tenir compte des données partagées par les chercheurs et les chercheuses pour évaluer leur carrière académique (voir plus bas), outre le fait d'inciter à partager ces données, permet aussi à leur institution d'affiliation de tracer ces données et les modalités de leur partage.

L'analyse par Milzow et al. (2020, fig.5) des DMP des projets financés en 2017-2018 par le FNS a identifié les dépôts qui y sont mentionnés en prévision du versement des données de recherche à la fin des projets. Comme mentionné plus haut, la HES-SO s'appuie sur la liste des dépôts les plus cités dans ces DMP pour ses recommandations sur le choix du dépôt (HES-SO [sans date]). La rédaction du DMP ayant été rendue obligatoire par le FNS en octobre 2017 (FNS 2017), et les projets correspondants arrivant prochainement à terme, il serait intéressant, en se basant sur la version finale des DMP, de faire un suivi des versements effectués (nombre de jeux de données, dépôts utilisés), éventuellement par discipline⁽¹¹⁾ et par institution. Ce suivi pourrait être utilisé par chaque institution pour éventuellement modifier sa stratégie, en ajustant ses recommandations pour le choix du dépôt ou en intensifiant ses efforts de sensibilisation et de communication, ou encore pour adapter ses mesures d'incitation.

Evaluation des chercheurs et des chercheuses : certaines données de stockage et d'utilisation des contenus hébergés par BORIS (dépôt pour publications) sont exploitées dans le processus d'évaluation de la recherche de l'Université de Berne (Universität Bern [sans date]b). Toutes les institutions étudiées ici sont signataires de la Déclaration de San Francisco sur l'évaluation de la recherche (DORA ou San Francisco Declaration on Research Assessment 2012, 2020), s'engageant ainsi à prendre en compte la valeur et l'impact des produits de la recherche autres que les publications. Un des moyens possibles de mettre en pratique cet engagement consiste à considérer également les métriques sur les dépôts de

données de recherche ouvertes (le nombre de jeux de données et de téléchargements, le suivi de la réutilisation) par chercheur ou chercheuse, par projet, par département, pour l'évaluation académique.

Mesures d'incitation et visibilité des pratiques par communauté : les mesures pour inciter à adopter de meilleures pratiques en gestion des données de recherche peuvent prendre diverses formes. La HES-SO a lancé un appel à projet pour l'obtention de fonds complémentaires soutenant le versement des données de recherche sur un dépôt FAIR (HES-SO 2020). L'EPFL, quant à elle, a mis en place une communauté de Data Champions, reconnus pour leur expertise et leurs bonnes pratiques en GDR, et offre visibilité et soutien à cette communauté (EPFL [sans date]). Sur ce modèle, toute institution pourrait aussi attribuer aux chercheurs et chercheuses qui partagent publiquement un grand nombre de données sur des dépôts des badges Open Science (Center for Open Science [sans date]), afin de reconnaître et de mettre en avant leurs bonnes pratiques. L'Université de Lausanne, de son côté, incite à ouvrir sur Zenodo des espaces communautaires par faculté (Université de Lausanne [sans date]b). Plusieurs autres institutions ou départements d'institutions y disposent déjà de leur espace communautaire(12). Sur des dépôts généralistes, cette organisation en communauté de pratiques et d'intérêt donne de la visibilité aux données de recherche par institution, et par faculté ou par département, et facilite le suivi des versements, selon la recommandation faite précédemment. Elle contribue aussi à l'identification et à la centralisation des données d'un domaine particulier, donnant ainsi une chance supplémentaire à ces données d'être découvertes et réutilisées.

Le Tableau 4 synthétise les recommandations que l'on peut formuler à l'adresse des institutions de recherche à partir de l'observation de ces bonnes pratiques.

Tableau 4 : Recommandations à l'adresse des institutions de recherche, basées sur les bonnes pratiques observées

Aspect	Exemple de pratique existante	Recommandation pour les institutions
Découvrabilité des données	Recommandations des Universités de Zurich et de Bâle pour la découverte de DR	Donner des conseils et mentionner des outils (registres, moteurs de recherche) pour la recherche et la découvrabilité des données sur les dépôts
Suivi des pratiques	Université de Zurich : publication de la liste des dépôts utilisés par ses chercheurs et chercheuses	Effectuer un suivi détaillé des versements par ses chercheurs et chercheuses ; selon les résultats de ce suivi, adapter éventuellement sa stratégie et ses recommandations pour le choix du dépôt
Evaluation des chercheurs et chercheuses	Université de Berne : prise en compte des données d'utilisation de BORIS (dépôt pour publications)	Prendre en compte le versement des données de recherche ouvertes sur des dépôts pour l'évaluation académique
Mesures d'incitation pour de bonnes pratiques en GDR et pour l'ouverture des DR	HES-SO: allocation de subsides pour le versement des données sur un dépôt FAIR; EPFL: communauté de Data Champions	Mettre en place des ressources (financières, humaines, de communication) pour inciter les chercheurs et chercheuses à partager publiquement leurs données
Visibilité des pratiques par communauté	EPFL, ETHZ, UNIBAS, UNIL-CHUV, UNIZ : communautés sur Zenodo	Créer des espaces communautaires sur les sites des dépôts de données de recherche ouvertes

Outils pour le choix de la stratégie

Au terme de cette étude, on a abouti à la production d'un vade-mecum pour le recueil d'informations et d'un guide décisionnel. Ces outils sont destinés aux personnes mandatées par une institution pour décider de la démarche à adopter concernant le soutien aux chercheurs et chercheuses pour le partage public de leurs données sur un dépôt. Dans l'idéal, ces personnes sont des représentant-e-s de services déjà engagés dans des activités de GDR et contribuant au dépôt existant ou potentiellement au futur dépôt : la bibliothèque de recherche, le service IT, le département légal, l'administration de la recherche, ainsi que des représentant-e-s des instances décisionnelles de haut niveau de l'institution, du département financier et des ressources humaines.

Le vade-mecum regroupe les questions auxquelles l'institution est invitée à répondre avant de commencer la démarche. Les informations ainsi collectées seront utiles à différentes étapes de la démarche. Le guide décisionnel se présente sous forme d'un logigramme. Un troisième document fournit des recommandations et des ressources complémentaires pour la mise en pratique de la stratégie, une fois celle-ci fixée à l'aide du guide décisionnel (Figure 7).

Figure 7 : Démarche pour le choix de la stratégie institutionnelle et outils correspondants

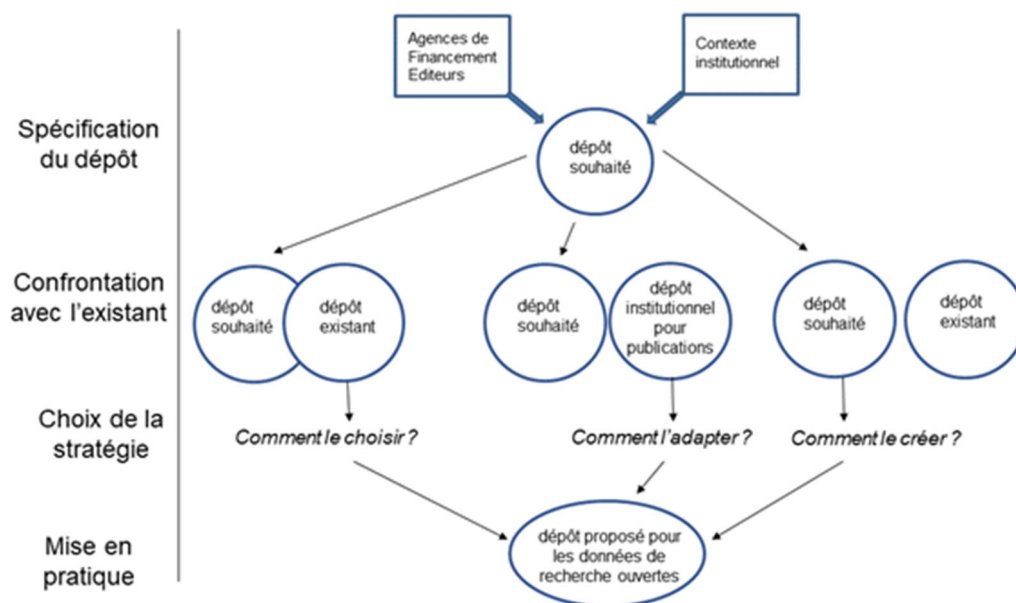


Présentation de la démarche décisionnelle

En fonction des exigences des agences de financement de la recherche et de celles des éditeurs de journaux scientifiques, l'institution caractérise le dépôt qui permettra à ses chercheurs et chercheuses de partager publiquement leurs données dans le respect de ces exigences. Ce dépôt souhaité doit aussi prendre en compte les spécificités du contexte institutionnel : la politique institutionnelle, les pratiques, la culture en matière de GDR, les ressources financières et humaines. La confrontation du dépôt souhaité avec les dépôts existants offre alors trois voies possibles (voir la Figure 8). La première consiste à orienter les chercheurs et chercheuses vers un dépôt existant (un dépôt disciplinaire, son dépôt institutionnel ou un dépôt généraliste) qui est suffisamment similaire à ce qu'on souhaite. Il faut pour cela identifier les critères permettant de sélectionner le dépôt adéquat. Dans le cas où aucun dépôt existant ne se rapproche suffisamment de ce qu'on souhaite, la deuxième voie possible est l'élargissement aux données de recherche du dépôt institutionnel pour publications, s'il existe. Et dans le cas où le dépôt pour publications n'existe pas, la troisième voie est celle de la création d'un nouveau dépôt de données de recherche.

Pour deux raisons principales, la priorité est donnée à l'utilisation d'un dépôt déjà existant plutôt qu'à la création d'un nouveau dépôt. D'une part, les ressources nécessaires pour développer et faire fonctionner un dépôt sont bien plus élevées que l'utilisation d'un dépôt existant (gratuite pour une grande partie des dépôts, dans une certaine limite de taille ou de nombre de fichiers versés) (Guirlet 2020, p.35 et suivantes). D'autre part, l'analyse ci-dessus des dépôts utilisés et recommandés par les institutions académiques suisses a montré une fragmentation poussée du paysage des dépôts de données de recherche. Cette fragmentation correspond à celle observée à une plus grande échelle par l'étude de von der Heyde (2019) qui recommande de ne pas financer de nouveaux dépôts, pour éviter d'accentuer ce phénomène, mais plutôt de consolider les dépôts déjà existants. Le FNS prévoit d'ailleurs d'étendre sur son site ses recommandations sur les dépôts conformes aux principes FAIR, afin de donner plus de visibilité aux dépôts spécialisés déjà existants (Milzow et al. 2020). Dans le guide décisionnel, pour le choix d'un dépôt existant, on s'oriente d'abord vers un dépôt disciplinaire, mieux adapté aux standards, à la culture et aux pratiques disciplinaires (Universität Bern [sans date]a), puis vers le dépôt institutionnel de données de recherche s'il existe, puis vers un dépôt généraliste.

Figure 8 : Les voies possibles pour le choix de sa stratégie par une institution



Cette démarche se traduit dans le vade-mecum et le guide décisionnel de la façon décrite ci-dessous.

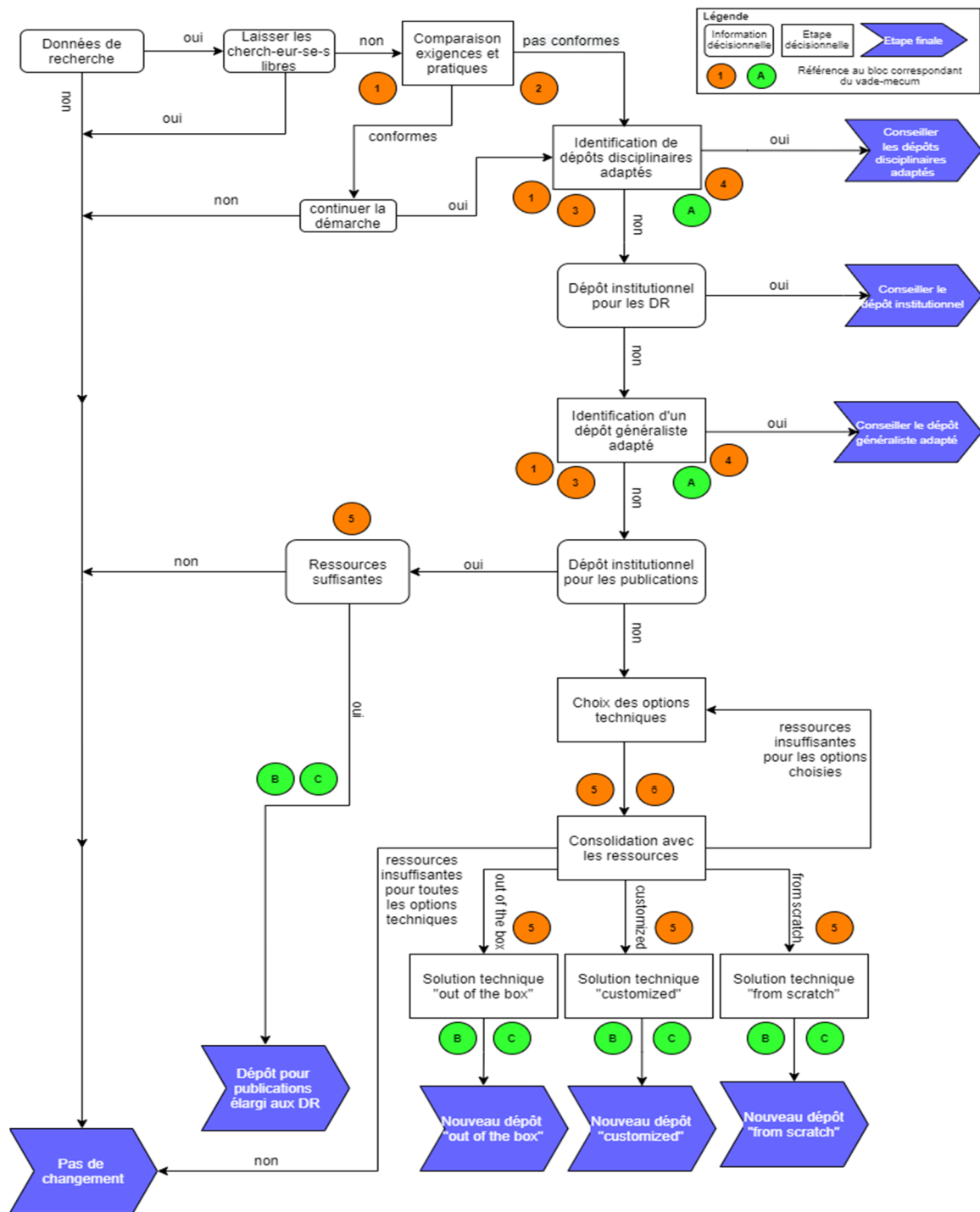
Le vade-mecum

Le vade-mecum permettant de collecter les informations nécessaires pour la démarche est organisé par rubrique. Chaque rubrique comprend une question principale et une liste de pistes suggérées pour explorer la question de manière approfondie (Tableau 5). Les informations récoltées correspondent soit à des paramètres décisionnels, décisifs pour l'orientation de la démarche (repérés avec des chiffres en orange), soit à des spécifications et des informations non décisionnelles, mais qui enrichissent le processus et contribuent au résultat final (par exemple les caractéristiques du dépôt que l'on va utiliser, adapter ou développer) (repérés avec des lettres en vert). Le chiffre ou la lettre de chaque rubrique est reporté sur le guide décisionnel à l'étape correspondante.

On commence par identifier les exigences et les directives pertinentes sur l'ouverture des données de recherche (rubrique 1). On s'intéresse alors aux pratiques des chercheurs et chercheuses de l'institution dans ce domaine (rubrique 2). On caractérise ensuite les données de recherche produites ou manipulées au sein de l'institution : les disciplines, la présence de données sensibles ou non, les volumes en jeu, les formats, ... (rubrique 3). On rassemble aussi les informations sur les ressources financières et humaines disponibles, qui seront notamment décisives pour le choix entre l'utilisation d'un dépôt existant (et lequel) et le développement d'un nouveau dépôt (rubriques 4 et 5), et pour les choix techniques associés (rubrique 6).

Les spécifications pour le choix d'un dépôt existant sont aussi précisées au moyen de ce vade-mecum (rubrique A), de même que les spécifications pour le nouveau dépôt éventuel (rubrique B). Elles incluent les éléments qui permettent de répondre aux exigences et aux directives dans la mesure des ressources internes, et d'autres souhaits possibles (la convivialité, l'adaptabilité, l'extensibilité, ...). Une rubrique spécifique concerne le cahier des charges à remplir pour l'élargissement ou la création d'un nouveau dépôt (rubrique C).

Tableau 5 : Rubriques du vade-mecum accompagnant le guide décisionnel



Le guide décisionnel

Le guide décisionnel se présente sous la forme du logigramme présenté sur la Figure 9.

Si l'institution choisit d'encadrer les chercheurs et les chercheuses pour le choix du dépôt, elle commence par évaluer leurs pratiques de partage public de leurs données (rubrique 2 du vade-mecum) en fonction des exigences sur les données de recherche (rubrique 1). Si ces pratiques ne sont pas satisfaisantes ou si l'institution décide de poursuivre quand même la démarche, elle est invitée à se tourner vers les dépôts déjà existants : des dépôts disciplinaires adaptés, le dépôt institutionnel de données de recherche s'il existe ou des dépôts généralistes. Pour les dépôts disciplinaires et les dépôts généralistes, on confronte les exigences, les caractéristiques

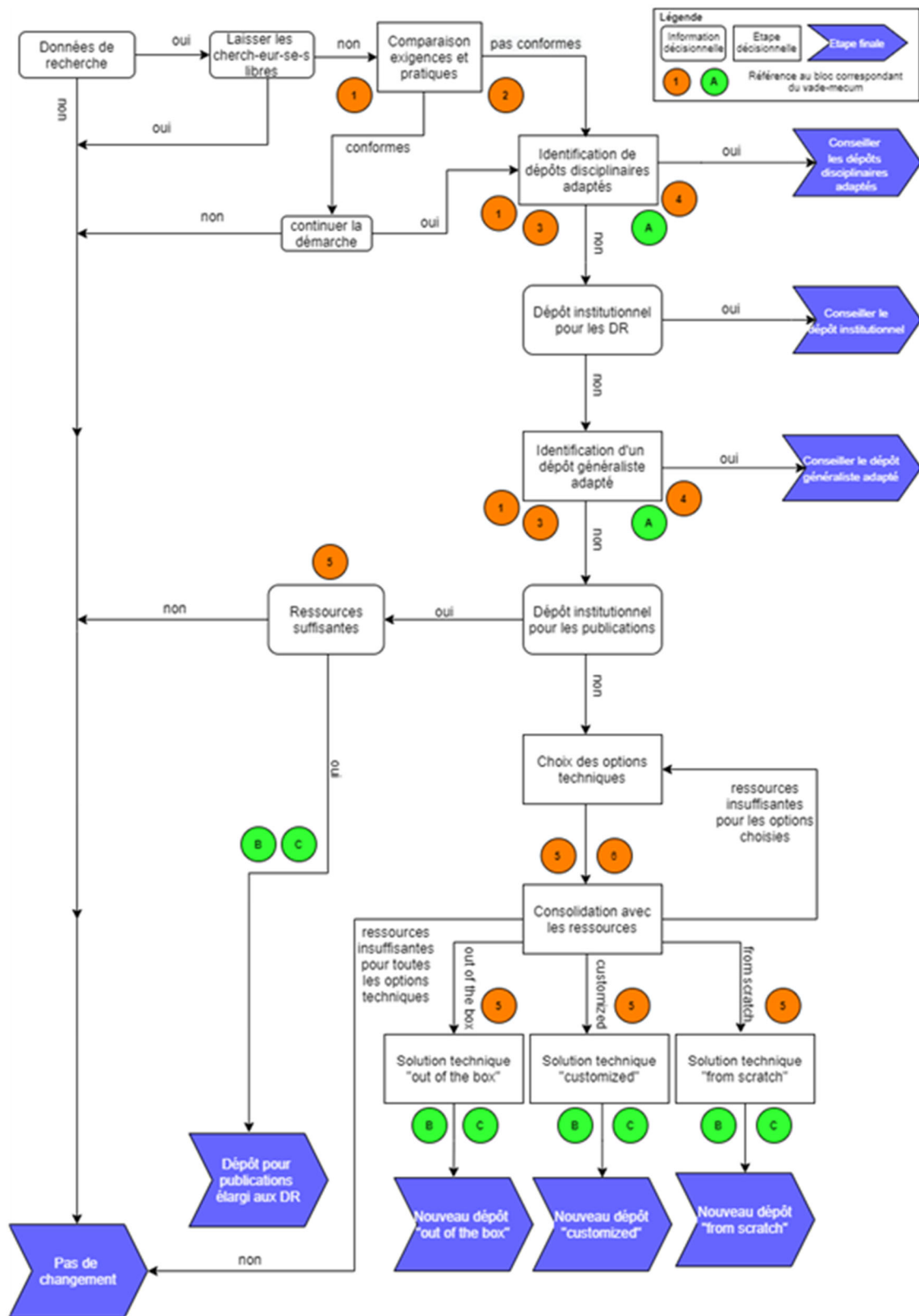
des données de recherche (rubrique 3) et les ressources disponibles pour l'utilisation d'un dépôt existant (rubrique 4) aux spécifications du dépôt existant qui serait le plus adapté (rubrique A).

Si aucun dépôt existant ne répond aux spécifications définies et si un dépôt institutionnel pour publications existe, on envisage d'élargir celui-ci aux données de recherche. La décision est prise en fonction des ressources et des compétences disponibles en interne (rubrique 5).

Enfin, s'il n'existe pas de dépôt pour publications, on envisage de créer un nouveau dépôt. Le choix de la solution technique (*out of the box/customized/from scratch*, propriétaire/Open Source, cloud ou externe/locale) se fait avec un outil fourni dans le vade-mecum (et détaillé dans Guirlet 2020) en fonction des préférences concernant l'échelle de temps et le niveau d'adaptabilité et de contrôle sur la solution et les données hébergées (rubrique 6). En fonction des ressources disponibles (rubrique 5), on confirme ces choix, on choisit d'autres options techniques ou on sort de la démarche.

L'élargissement du dépôt de publications ou la création d'un nouveau dépôt se fait en se basant sur les spécifications du nouveau dépôt (rubrique B) et sur le cahier des charges du projet correspondant (rubrique C).

Figure 9 : Guide décisionnel pour le choix de la stratégie institutionnelle sur les dépôts de données de recherche (la légende est en haut à droite)



Recommandations et ressources complémentaires

Des recommandations et des ressources utiles pour la mise en pratique de la stratégie fixée à l'étape précédente ont été regroupées dans un troisième outil (Tableau 6). Une partie de ces

recommandations peut aussi être utilisée pour améliorer les pratiques ou améliorer un dépôt existant.

Tableau 6 : Recommandations et ressources complémentaires pour la mise en pratique de la stratégie fixée à l'aide du guide décisionnel (les références et les détails sont donnés dans Guirlet 2020)

Rubrique	Suggestions
Utiliser des dépôts existants	<p>Former les chercheurs et les chercheuses au versement</p> <p>Faire un suivi des versements</p> <p>Demander les aides financières pour le versement</p> <p>Créer des communautés de pratiques et d'intérêt</p>
Créer ou améliorer un dépôt	<p>Préparer le projet de dépôt</p> <p>Rédiger la documentation</p> <p>Renforcer la découvrabilité des données et l'image du dépôt</p> <p>Assurer la promotion du dépôt (mentionné dans la politique et sur le site Internet institutionnels, les réseaux sociaux, ...)</p> <p>Faciliter et inciter à son utilisation (créer des communautés d'utilisateurs et d'utilisatrices, prendre en compte pour l'évaluation académique les données ouvertes, ...)</p> <p>Suivre l'utilisation du dépôt (données versées et données réutilisées)</p>
Prévoir pour le long terme	<p>Prévoir la stratégie en cas de dépassement des capacités du dépôt</p> <p>Évaluer régulièrement le dépôt (à l'aide des critères CTS)</p> <p>Veiller sur l'évolution du contexte OS et des exigences sur les DR</p> <p>Utiliser régulièrement le vade-mecum et le guide décisionnel pour évaluer la pertinence des recommandations institutionnelles et la conformité du dépôt institutionnel et/ou recommandé</p> <p>Faire les changements qui s'imposent (sur le dépôt, sur les recommandations)</p>
Ressources complémentaires fournies sur ...	<p>Les aides financières</p> <p>La préparation du projet : business model et gestion des risques</p> <p>La documentation à prévoir : exemples de conditions d'acceptation des données par le dépôt, de contrat de dépôt et de contrat d'utilisation</p> <p>La découvrabilité des données et l'image du dépôt : moissonnage des métadonnées (OAI-PMH) ; enregistrement du dépôt par Re3data ; certification du dépôt par CTS</p> <p>L'incitation à utiliser le dépôt : badges Open Science pour les chercheurs et chercheuses qui versent le plus de données ; prise en compte des données ouvertes dans l'évaluation académique (déclaration DORA¹³)</p>

(13)

Version pour navigateur Internet

Dans l'objectif de rendre la démarche plus souple et l'utilisation de ces outils interactive et collaborative, ces outils pourraient être rendus disponibles en version Internet. A titre de démonstration, quelques pages d'un prototype appelé InSTORED, pour *Institutional Strategy Tool for Open Research Data*, ont été développées (Figure 10). Avec cette version, l'accès aux outils décisionnels se fait à partir de l'un des quatre points d'entrée possibles (ou tâches) placés au même niveau sur la page d'accueil du site. Ces points d'entrée sont :

l'évaluation de l'alignement des pratiques actuelles avec les exigences sur les DR, l'identification d'un dépôt existant correspondant aux besoins, l'adaptation d'un dépôt pour publications aux DR, et la spécification et la création d'un dépôt de DR institutionnel (Figure 10, haut). Cette version inclut aussi toutes les rubriques du vade-mecum (la Figure 10, milieu, présente les rubriques sur la spécification du dépôt institutionnel à créer et sur le cahier des charges).

Dans le cas de la première tâche (l'évaluation de l'alignement des pratiques avec les exigences), les exigences sur les dépôts par le FNS et par H2020 sont rappelées. En cas de changement de ces exigences, ou de l'apparition de nouvelles exigences, il serait facile de mettre à jour cette rubrique. Il est également possible de remplir les exigences qui s'appliquent à l'ouverture des données de recherche dans un autre pays (Figure 10, bas), élargissant ainsi l'utilité des outils à d'autres institutions que les institutions suisses.

Ce format permet aussi de prévoir un espace collaboratif, où les institutions ayant déjà effectué cette démarche décisionnelle seraient invitées à déposer leurs retours d'expérience pour en faire bénéficier d'autres.

Figure 10 : Pages extraites du prototype [InSTORED](#) de version pour navigateur Internet des outils décisionnels : points d'entrée (ou tâches) proposés (haut); spécification du nouveau dépôt institutionnel pour données de recherche (milieu); transposabilité à d'autres pays de la comparaison des exigences et des pratiques (bas)



The screenshot shows the 'Prepare' step of the InStOReD tool. The main heading is '3. Specify and create your institutional data repository: b. Design the repository'. Below this, there are two sections: 'Specifications of the new data repository' and 'Project description'. The 'Specifications' section includes checkboxes for: Research disciplines, Data types and formats, Size limit on files, Metadata schema, Verification and validation, Handling of sensitive data, PID, Licences, and Citation. The 'Project description' section includes checkboxes for: Tasks and responsibilities, Resources, Timetable, Work packages, and Risk assessment and mitigation. To the right, a message states: 'You are ready to create your own repository according to the specifications and the technical solution you selected.' The interface also features a navigation bar with 'Home < New repository < Develop', a search bar, and language options (de, fr, it, en).

The screenshot shows the 'Assess needs' step of the InStOReD tool. The main heading is '1. Assess the alignment between practices and requirements'. Below this, there are two sections: 'Identify requirements and guidelines on research data' and 'Assess the institutional practices'. The 'Identify requirements' section includes checkboxes for: Research funders (with a dropdown menu showing 'In Switzerland' and 'Add my country'), Editors, and Institution. To the right, there are two boxes: 'Requirements from funders in Switzerland' (listing PNS: open and FAIR data, non-commercial repository and HQ020: open and FAIR data, certified repository) and 'Requirements from funders in (fill country name):' (with fields for Agency name 1 and Agency name 2, and a 'Go back to task 1' button). The interface also features a navigation bar with 'Home < Assess needs < Country', a search bar, and language options (de, fr, it, en).

Conclusion

Au terme de cette étude, on a abouti aux principaux résultats suivants.

On a identifié les directives pour l'ouverture des données de recherche en Suisse, ainsi que les moyens de les mettre en pratique. On a défini ce qu'est un dépôt pour le partage public des données et pour leur réutilisation, quelles formes il peut prendre, quels outils et services il peut fournir. On a identifié des critères de qualité pour ce dépôt et élaboré un modèle de description pouvant être utilisé soit pour l'évaluation soit pour la conception d'un dépôt de DR ouvertes. On a dressé un panorama des dépôts de données généralistes, disciplinaires et institutionnels utilisés et recommandés par neuf institutions académiques. On a identifié les informations et les paramètres importants pour le choix d'un dépôt ou la création d'un nouveau.

A partir des résultats précédents, on a produit des outils qui formalisent la démarche de choix de la meilleure stratégie possible par une institution de recherche en matière de dépôt de

données de recherche ouvertes. En suivant cette démarche, l'institution fait un choix éclairé qui lui permettra de répondre aux exigences en vigueur sur les données de recherche, tout en tenant compte des besoins et des pratiques des chercheurs et chercheuses et du contexte et des ressources à l'échelle locale.

Les institutions de recherche sont invitées à utiliser régulièrement ces outils. Le découpage modulaire de certaines étapes du guide décisionnel permet de les suivre indépendamment les unes des autres, en effectuant une partie de la démarche seulement. Avec les premières étapes de ce guide, on peut vérifier régulièrement la conformité entre les exigences et les pratiques, et adapter le cas échéant les conseils donnés aux chercheurs et chercheuses pour le choix du dépôt. Avec les étapes suivantes, on peut aussi réévaluer régulièrement, s'il existe, un ou des dépôts disciplinaires ou généralistes plus pertinents que ceux couramment utilisés par ses chercheurs et chercheuses, ou encore, adapter le dépôt institutionnel de données pour qu'il réponde aux besoins de façon plus satisfaisante.

Ces outils sont également adaptables en fonction des changements du paysage des dépôts et transposables à un autre contexte hors de Suisse. Avec la version prototype pour Internet développée au terme de cette étude, on facilite la mise à jour des outils, leur adaptabilité et leur transposabilité. Cette version offre en plus l'avantage d'inclure des aspects collaboratifs, pour le partage de retours d'expérience et l'échange de bonnes pratiques par des institutions ayant déjà effectué la démarche.

En aidant les institutions à proposer aux chercheurs et chercheuses le dépôt le plus adapté, ces outils les aident à mettre en pratique leur politique en matière d'ouverture des données de recherche, à répondre aux exigences en vigueur, et à fournir un instrument qui convient à leurs chercheurs et chercheuses et au contexte institutionnel. En proposant d'autres options que la création d'un nouveau dépôt, selon un argumentaire construit et adapté à chaque cas, dans un souci de rationalisation et de mutualisation des ressources, ces outils devraient contribuer aussi à limiter la fragmentation de l'offre et la multiplication inutile de dépôts de données de recherche en Suisse.

Notes

[\(1\)](#) On peut consulter la liste de Docuteam ([sans date]) des formats de fichiers reconnus comme adaptés à l'archivage des données. Cette liste est inspirée du catalogue des formats de données d'archivage du CECO ([sans date]).

[\(2\)](#) CASRAI est une organisation à but non lucratif travaillant sur la standardisation de formats pour la gestion et l'échange de l'information dans le domaine de la recherche (CASRAI [sans date]a).

[\(3\)](#) D'autres licences ouvertes telles que celles des Open Data Commons sont possibles (Open Data Commons [sans date] ; Ball 2011). Mais comme seules les licences CC ont été rencontrées dans le cadre de cette étude, on se limite à celles-ci.

[\(4\)](#) Ces autres critères sont : la facilité du versement, l'accessibilité, la découvrabilité, la curation, l'infrastructure de préservation, la pérennité de l'organisation, et le soutien pour les formats et les standards utilisés (EPFL Library, Research Data Library Team [sans date], p.31).

[\(5\)](#)(pour la préservation des données en accès restreint seulement)

[\(6\)](#)(pour les données liées à une publication aussi hébergée par BORIS)

[\(7\)](#)Springer Nature [sans date]b

[\(8\)](#)PLOS ONE [sans date]b

[\(9\)](#)Science Europe 2018

[\(10\)](#)FNS [sans date]b; voir aussi Perini 2019

[\(11\)](#)On peut déjà consulter la liste des dépôts par discipline établie par von der Heyde (2019, fig. 13 et fig. 18), à partir des réponses à ses questionnaires auprès des chercheurs et chercheuses de la communauté académique suisse en 2018.

[\(12\)](#)Voir les espaces communautaires de la FBM de l'Université de Lausanne et du CHUV (depuis 2016 ; zenodo [sans date]b), de l'Institute for Atmospheric and Climate Science ETH Zürich (depuis 2019 ; zenodo [sans date]c), l'espace Research Data University of Basel (depuis 2019 ; zenodo [sans date]d), et l'espace University of Zurich (depuis 2013 ; zenodo [sans date]e).

[\(13\)](#)Ou San Francisco Declaration on Research Assessment (2012)

Acronymes et abréviations

ACOUA	ACademic OUtput Archive
ARK	Archival Resource Key
ArODES	Archive Ouverte des Domaines de la HES-SO
BerDA	Bern Digital Archive
BORIS	Bern Open Repository and Information System
CASRAI	Consortia Advancing Standards in Research Administration Information
CC	Creative Commons
CCSDS	Consultative Committee for Space Data Systems
CECO électroniques	Centre de coordination pour l'archivage à long terme de documents électroniques
CHUV	Centre Hospitalier Universitaire Vaudois
CTS	CoreTrustSeal
DaSCH	Data and Service Center for the Humanities

DDI	Data Documentation Initiative
DINI	Deutsche Initiative für NetzwerkInformation
DLCM	Data Life Cycle Management
DMP	Data Management Plan
DOI	Digital Object Identifier
DORA	Declaration On Research Assessment
DR	Données de la Recherche
DSA	Data Seal of Approval
EUDAT	EUropean Data Infrastructure
EPFL	Ecole Polytechnique Fédérale de Lausanne
ETHZ Fédérale de Zurich	Eidgenössische Technische Hochschule Zürich ou Ecole Polytechnique
FAIR	Findable, Accessible, Interoperable, Reusable
FAQ	Frequently Asked Questions, Foire Aux Questions
FBM	Faculté de Biologie et de Médecine (UNIL-CHUV)
FNS	Fonds National Suisse de la Recherche Scientifique
GDR	Gestion des Données de la Recherche
H2020	Horizon 2020
HEG	Haute Ecole de Gestion de Genève
HES-SO	Haute Ecole Spécialisée de Suisse Occidentale
HEI	Haute Ecole Institutionnelle
ICPSR	Inter-university Consortium for Political and Social Research
ICSU-WDS	International Council for Science's World Data System
InSTOREd	Institutional Strategy Tool for Open Research Data
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting

OAIS	Open Archival Information System
ORCID	Open Researcher and Contributor ID
OS	Open Science
OSF	Open Science Framework
PID	Persistent Identifier
PURL	Persistent Uniform Resource Locator
RDA	Research Data Alliance
RDF	Resource Description Framework
Re3data	Registry of Research Data Repositories
RRID	Research Resource Identifier
Serval	Serveur académique lausannois
UNIBAS	Université de Bâle
UNIBE	Université de Berne
UNIFR	Université de Fribourg
UNIGE	Université de Genève
UNIL	Université de Lausanne
UNIZ	Université de Zurich
URN	Uniform Resource Name
ZORA	Zürich Open Repository Archive

Bibliographie

Amsterdam Call for Action on Open Science, 2016. [en ligne]. 04.04.2016.

[Consulté le 07.11.2020]. Disponible à l'adresse :

<https://www.government.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science>

ARLITSCH, Kenning, OBRIEN, Patrick, MIXTER, Jeffrey K., CLARK, Jason A. et STERMAN, Leila, 2016. Ensuring Discoverability of IR Content. In : CALLICOTT, Burton B., SCHERER, David et WESOLEK, Andrew. *Making Institutional Repositories Work* [en ligne]. Ed. Purdue University Press. [Consulté le 07.11.2020], pp. 31-50. ISBN 978-1-55753-902-1. Disponible à l'adresse : <http://www.jstor.org/stable/10.2307/j.ctt1wf4drg>

BALL, Alex, 2011. How to License Research Data. *dcc.ac.uk* [en ligne]. 09.02.2011. Version modifiée le 17.07.2014. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://dcc.ac.uk/guidance/how-guides/license-research-data>

BLUMER, Eliane et BURGI, Pierre-Yves, 2015. Data Life-Cycle Management Project: SUC P2 2015-2018. *Revue électronique suisse de science de l'information* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : http://www.ressi.ch/num16/article_110

BOSTON UNIVERSITY DATA SERVICES, [sans date]. Selecting a data repository. *bu.edu* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.bu.edu/data/share/selecting-a-data-repository/#openbu>

CASRAI [sans date]a. Welcome to CASRAI. *casrai.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://casrai.org/>

CASRAI [sans date]b. Research Data Management Glossary. *casrai.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://casrai.org/rdm-glossary/>

CECO, [sans date]. Catalogue des formats de données d'archivage. *kost-ceco.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://kost-ceco.ch/cms/formats-de-donnees.html>

CENTER FOR OPEN SCIENCE, [sans date]. Open Science badges enhance openness, a core value of scientific practice. *cos.io* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.cos.io/initiatives/badges>

CORETRUSTSEAL, 2020a. Core Certified Repositories. *CoreTrustSeal* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.coretrustseal.org/why-certification/certified-repositories/>

CORETRUSTSEAL, 2020b. *CoreTrustSeal Trustworthy Data Repositories Requirements 2020-2022, version 02.00* [en ligne]. S.l. : s.n. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.coretrustseal.org/why-certification/requirements/>

CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2020. *CoreTrustSeal Trustworthy Data Repositories Requirements: Extended Guidance 2020-2022, version 2.0* [en ligne]. S.l. : s.n. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.coretrustseal.org/why-certification/requirements/>, <https://zenodo.org/record/3632533>

CORRADO, Edward M., 2019. Repositories, Trust, and the CoreTrustSeal. *Technical Services Quarterly*. 02.01.2019. Vol. 36, n 1, p. 61-72. DOI [10.1080/07317131.2018.1532055](https://doi.org/10.1080/07317131.2018.1532055)

CRAGIN, Melissa H., HEIDORN, P. Bryan, PALMER, Carole L. et SMITH, Linda C., 2007. *An Educational Program on Data Curation* [en ligne]. 25.06.2007. [Consulté le 07.11.2020]. Disponible à l'adresse : <http://hdl.handle.net/2142/3493>

CREATIVE COMMONS [sans date]. About CC licenses. *creativecommons.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://creativecommons.org/about/ccllicenses/>

DARIS, 2018. Implementation of the CoreTrustSeal. *coretrustseal.org* [en ligne]. 20.03.2018. [Consulté le 02.12.2020]. Disponible à l'adresse: <https://www.coretrustseal.org/wp-content/uploads/2018/03/DARIS.pdf>

DASCH, [sans date]. Services. *dasch.swiss* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://dasch.swiss/services/#appendix>

DATA CURATION NETWORK, [sans date]. Mission. *datacurationnetwork.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://datacurationnetwork.org/about/our-mission/>

DATA CITE, [sans date]. Welcome to DataCite. *datacite.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://datacite.org/>

DATA CITE METADATA WORKING GROUP, 2019. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data v4.3. 2019. pp. 73 pages. Disponible à l'adresse: <https://doi.org/10.14454/7xq3-zf69>

DATAVERSE PROJECT, [sans date]. User Guide. *dataverse.org*. [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <http://guides.dataverse.org/en/4.20/user/>

DIGITAL PRESERVATION COALITION, [sans date]. Digital Preservation Handbook. *dpconline.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers>

DILLO, Ingrid et de LEEUW, Lisa, 2018. CoreTrustSeal. In : *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*. 19.07.2018. Vol. 71, n° 1, p. 162-170. DOI [10.31263/voebm.v71i1.1981](https://doi.org/10.31263/voebm.v71i1.1981)

DINI 2011. DINI-Zertifikat Dokumenten- und Publikationsservice 2010. Version 3.1. [en ligne]. 03.2011. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://edoc.hu-berlin.de/handle/18452/2145>

DOCUTEAM, [sans date]. Standard de versement. *docuteam.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://www.docuteam.ch/fr/prestations/archivage-electronique-docuteam-cosmos/standard-de-versement/#_ftn1

DRYAD, [sans date]. Submission process. *dryad.org* [en ligne]. [Consulté le 03.12.2020]. Disponible à l'adresse : https://datadryad.org/stash/submission_process#upload-methods

Dublin Core. 2020. *Wikipédia : l'encyclopédie libre* [en ligne]. Dernière modification de la page le 05.10.2020 à 8:14. [Consulté le 07.11.2020]. Disponible à l'adresse : https://fr.wikipedia.org/wiki/Dublin_Core#Autres_r%C3%A9f%C3%A9rentiels_de_m%C3%A9tadonn%C3%A9es

DUBLIN CORE METADATA INITIATIVE, 2020. *dublincore.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.dublincore.org/>

EPFL, [sans date]. EPFL Data Champions. *epfl.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.epfl.ch/campus/library/services/services-researchers/rdm-contacts-communities/epfl-data-champions/#more>

EPFL LIBRARY, [sans date]. Research Data Management Fast Guides. *epfl.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.epfl.ch/campus/library/services/services-researchers/rdm-guides-templates/>

EPFL LIBRARY, RESEARCH DATA LIBRARY TEAM, [sans date]. RDM Walkthrough Guide. *epfl.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.epfl.ch/campus/library/services/services-researchers/rdm-guides-templates/>

ETHZ - ETH-BIBLIOTHEK, [sans date]. FAQs de. *ethz.ch* [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse : <https://documentation.library.ethz.ch/display/RC/FAQs+de>

EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION, 2016. *H2020 Programme - Guidelines on FAIR Data Management in Horizon 2020* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION, 2017. *H2020 Programme - Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

FIGSHARE, [sans date]. How to upload and publish your data. *figshare.com* [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse : <https://help.figshare.com/article/how-to-upload-and-publish-your-data>

FNS, [sans date]a. Open Research Data. *snf.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open_research_data/Pages/default.aspx#D%E9claration%20de%20principe%20du%20FNS%20sur%20le%20libre%20acc%E8s%20aux%20donn%E9es%20de%20la%20recherche%20%28Open%20Research%20Data%29

FNS, [sans date]b. Data Management Plan (DMP) - Directives pour les chercheuses et chercheurs. *snf.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open_research_data/Pages/data-management-plan-dmp-directives-pour-les-chercheuses-et-chercheurs.aspx

FNS, 2017. Open Research Data : les requêtes devront inclure un plan de gestion des données. FNS. *snf.ch* [en ligne]. 06.03.2017. [Consulté le 07.11.2020]. Disponible à l'adresse : <http://www.snf.ch/fr/pointrecherche/newsroom/Pages/news-170306-open-research-data-bientot-une-realite.aspx>

FNS, 2020. Règlement d'exécution général relatif au règlement des subsides. *snf.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse:

http://www.snf.ch/SiteCollectionDocuments/fns-reglement_execution_general_relatif_au_reglement_subsidies_f.pdf#page=15

FOREIGN COMMONWEALTH OFFICE, 2013. G8 Science Ministers Statement. *gov.uk* [en ligne]. 12.06.2013. [Consulté le 07.11.2020]. Disponible à l'adresse :

<https://www.gov.uk/government/news/g8-science-ministers-statement>

FORS, [sans date]. Contrat utilisateur. *unil.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: https://forsbase.unil.ch/media/general_documentation/fr/User_contract_F.pdf

GUIRLET, Marielle, 2020. Guide décisionnel et vade-mecum pour la mise à disposition d'un dépôt de données de recherche ouvertes en Suisse. [en ligne]. Genève : Haute école de gestion de Genève. Travail de Master. Version révisée. 18.12.2020. DOI: 10.5281/zenodo.4357134. [Consulté le 18.12.2020]. Disponible à l'adresse: <https://zenodo.org/record/4357134>

HES-SO, [sans date]. Archiver ses données de recherche. *hes-so.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://openscience.hes-so.ch/fr/archiver-donnees-recherche-14819.html>

HES-SO, 2020. Appel à projets Open Data HES-SO. *hes-so.ch* [en ligne]. 02.06.2020. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.hes-so.ch/fr/appel-projets-open-data-hes-so-16822.html>

HEWLETT, Niles E, 2006. The USDA Enterprise Architecture Program. [en ligne]. 25.01.2006. [Consulté le 07.11.2020]. Disponible à l'adresse : https://web.archive.org/web/20070508175931/http://www.ocio.usda.gov/p_mgmt/doc/PM_Classes_EA_NEH_012506_Final.ppt

HODSON, Simon, JONES, Sarah et al., 2018. *Turning FAIR data into reality. Interim report of the European Commission Expert Group on FAIR data* [en ligne]. S.l. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://zenodo.org/record/1285272> ICPSR, [sans date]. History. *icpsr.umich.edu* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.icpsr.umich.edu/web/pages/about/history/>

JOHNSTON, Lisa R, CARLSON, Jake, HUDSON-VITALE, Cynthia, IMKER, Heidi, KOZLOWSKI, Wendy, OLENDORF, Robert et STEWART, Claire, 2016. *Data Curation Terms and Activities* [en ligne]. 23.10.2016. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://conservancy.umn.edu/bitstream/handle/11299/188638/DefinitionsofDataCurationActivities%20%281%29.pdf?sequence=1&isAllowed=y>

JOHNSTON, Lisa R., 2017. Data Curation Handbook Steps. In: JOHNSTON, Lisa R., 2017. *Curating Research Data Volume Two: A Handbook of Current Practice*. Ed: American Library Association, 2017. [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://conservancy.umn.edu/bitstream/handle/11299/183502/Data%20Curation%20Handbook%20Steps_v2.pdf?sequence=1&isAllowed=y

JOHNSTON, Lisa, CARLSON, Jake, HUDSON-VITALE, Cynthia, IMKER, Heidi, KOZLOWSKI, Wendy, OLENDORF, Robert et STEWART, Claire, 2017. *Data Curation*

Network: A Cross-Institutional Staffing Model for Curating Research Data [en ligne]. S.l. [Consulté le 07.11.2020]. Disponible à l'adresse : <http://hdl.handle.net/11299/188654>

JOHNSTON, Lisa R, CARLSON, Jacob, HUDSON-VITALE, Cynthia, IMKER, Heidi, KOZLOWSKI, Wendy, OLENDORF, Robert et STEWART, Claire, 2018. How Important is Data Curation? Gaps and Opportunities for Academic Libraries. *Journal of Librarianship and Scholarly Communication*. 26.04.2018. Vol. 6, n° 1, pp. 2198. DOI [10.7710/2162-3309.2198](https://doi.org/10.7710/2162-3309.2198)

JONG (de), Michiel, ZUIDERWIJK, Anneke, WILL, Nicole et JANSSEN, Marijn, 2020. Open Science: Sharing Your Research with the World [online course]. TU Delft. *edx.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.edx.org/course/open-science-sharing-your-research-with-the-world>

L'HOURS, Hervé, KLEEMOLA, Mari et de LEEUW, Lisa, 2019. CoreTrustSeal: From academic collaboration to sustainable services. In : *IASSIST Quarterly*. 10.05.2019. Vol. 43, n°1, p. 1-17. DOI [10.29173/iq936](https://doi.org/10.29173/iq936)

MILZOW, Katrin, VON ARX, Martin, SOMMER, Cornélia, CAHENZLI, Julia et PERINI, Lionel, 2020. *Open Research Data: SNSF monitoring report 2017-2018* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://zenodo.org/record/3618123>

NATURE, 2016. Data availability statements and data citations policy: Guidance for authors. *Nature* [en ligne]. 09.2016. [Consulté le 07.11.2020]. Disponible à l'adresse: <http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf>

OLOS, 2020a. *OLOS Specifications* [fichier texte Office Open]. Version 1. Dernière mise à jour le 19.05.2020. Document interne au projet.

OLOS, 2020b. Integrated data management solution for researchers and institutions. *olos.swiss*. [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse. <https://olos.swiss/>

OPEN DATA COMMONS, [sans date]. Open Data Commons, Legal tools for Open Data [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: <https://opendatacommons.org/licenses/odbl/>

Open Research Data (ORD) – the uptake in Horizon 2020, 2016. *EU Open Data Portal, europa.eu* [en ligne]. 10.05.2016. 19.04.2018. [Consulté le 03.12.2020]. Disponible à l'adresse : <https://data.europa.eu/euodp/en/data/dataset/open-research-data-the-uptake-of-the-pilot-in-the-first-calls-of-horizon-2020>

PERINI, Lionel, 2019. SNSF Open Research Data Policy. *Journée Open Science* [en ligne]. HES-SO, 18 March 2019. [Consulté le 07.11.2020]. Disponible à l'adresse: <https://www.hes-so.ch/data/documents/4-L.Perini-FNS-PolitiqueFNS-OpenData-10357.pdf>

PINFIELD, Stephen, COX, Andrew M. et SMITH, Jen, 2014. Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. In : LAUNOIS, Pascal (éd.), *PLoS ONE*. 08.12.2014. Vol. 9, n° 12, p. e114734. DOI: 10.1371/journal.pone.0114734 Pryor 2012

PLOS ONE, [sans date]a. Data availability. *plos.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: <http://journals.plos.org/plosone/s/data-availability>

PLOS ONE, [sans date]b. How to Store and Manage Your Data. *plos.org* [en ligne]. [Consulté le 06.12.2020]. Disponible à l'adresse : <https://plos.org/resource/how-to-store-and-manage-your-data/#choosing-repository>

PRYOR, Graham, 2012. Why manage research data? *Managing Research Data*. Londres: Facet Publishing, pp.1-16. ISBN 978-1-85604-756-2

RE3DATA.ORG, [sans date]. *re3data.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.re3data.org/>

RE3DATA.ORG, 2017. B2SHARE. *re3data.org* [en ligne]. 22.11.2017 [Consulté le 08.11.2020]. Disponible à l'adresse : <https://www.re3data.org/repository/r3d100011394>

RICE, Robin et SOUTHALL, John, 2016. *The data librarian's handbook*. London : Facet Publishing. ISBN 978-1-78330-047-1.

ROSENTHALER, Lukas, FORNARO, Peter et CLIVAZ, Claire, 2015. DaSCH: Data and Service Center for the Humanities. In : *Digital Scholarship in the Humanities*. 2015. Vol. 30, p. i43-i49. [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse : https://academic.oup.com/dsh/article/30/suppl_1/i43/365238

San Francisco Declaration on Research Assessment, 2012. DORA [en ligne]. 16.11.2012. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://sfdora.org/read/>

San Francisco Declaration on Research Assessment. 2020. DORA Signers. [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse: <https://sfdora.org/signers/>

Science Europe, 2018. Practical Guide to the International Alignment of Research Data Management. *scienceeurope.org* [en ligne]. 11.2018. [Consulté le 08.11.2020]. Disponible à l'adresse: https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

Sorbonne declaration on research data rights, 2020. [en ligne]. [Consulté le 07.12.2020]. Disponible à l'adresse: <https://sorbonedatadeclaration.eu/>

SPRINGER NATURE, [sans date]a. Research Data Policies FAQ. *Springer Nature* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: <https://www.springernature.com/gp/authors/research-data-policy/data-policy-faqs>

SPRINGER NATURE, [sans date]b. Research Data Policies. Recommended Repositories. *Springer Nature* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: <https://www.springernature.com/gp/authors/research-data-policy/recommended-repositories>

SWISS NATIONAL SCIENCE FOUNDATION, [sans date]. *Explanation of the FAIR data principles* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : http://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf

SWISS NATIONAL SCIENCE FOUNDATION, 2017. *Examples of data repositories* [en ligne]. 27.04.2017. [Consulté le 07.11.2020]. Disponible à l'adresse : http://www.snf.ch/SiteCollectionDocuments/FAIR_data_repositories_examples.pdf

THE CONCORDAT WORKING GROUP, 2016. *Concordat on Open Research Data* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.bolton.ac.uk/assets/ConcordatonOpenResearchData.pdf>

THE CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012. *Reference Model for an Open Archival Information System (OAIS) - Recommended Practice CCSDS 650.0-M-2* [en ligne]. Washington, DC, USA. CCSDS. [Consulté le 07.11.2020]. Recommendation for Space Data System Practices. Disponible à l'adresse : <https://public.ccsds.org/Pubs/650x0m2.pdf>

THE ROYAL SOCIETY, 2012. *Science as an Open Enterprise* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/>

THE UNIVERSITY OF EDINBURGH, 2019. Trustworthy Digital Repository. *ed.ac.uk* [en ligne]. 20.06.2019. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.ed.ac.uk/information-services/research-support/research-data-service/after/data-repository/trustworthy-digital-repository>

TRELOAR, Andrew, GROENEWEGEN, David et HARBOE-REE, Cathrine, 2007. The Data Curation Continuum: Managing Data Objects in Institutional Repositories. *D-Lib Magazine* [en ligne]. Septembre 2007. Vol. 13, n° 9/10. [Consulté le 07.11.2020]. DOI [10.1045/september2007-treloar](https://doi.org/10.1045/september2007-treloar). Disponible à l'adresse : <http://www.dlib.org/dlib/september07/treloar/09treloar.html>

TRELOAR, Andrew, 2012. Private Research, Shared Research, Publication, and the Boundary Transitions. Version 1.4.3 [en ligne]. 19.03.2012. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://andrew.treloar.net/research/diagrams/>

UNIVERSITÄT BASEL, [sans date]. Sharing data. *unibas.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://researchdata.unibas.ch/en/publish-and-share/>

UNIVERSITÄT BERN, [sans date]a. Universitätsbibliothek. Forschungsdatenmanagement. *unibe.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://www.unibe.ch/universitaet/dienstleistungen/universitaetsbibliothek/service/open_science/forschungsdatenmanagement/index_ger.html

UNIVERSITÄT BERN, [sans date]b. Universitätsbibliothek. BORIS Repository. *unibe.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : https://www.unibe.ch/universitaet/dienstleistungen/universitaetsbibliothek/service/elektronisch_publicieren/boris_repository/index_ger.html

UNIVERSITÄT ZÜRICH, HAUPTBIBLIOTHEK, 2019. Empfohlene Repositories. *uzh.ch* [en ligne]. 05.12.2019. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.hbz.uzh.ch/de/open-access-und-open-science/daten-repositories/empfohlene-repositories.html>

UNIVERSITÄT ZÜRICH, HAUPTBIBLIOTHEK, 2020. Nutzen von Daten in Repositories. *uzh.ch* [en ligne]. 06.02.2020. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.hbz.uzh.ch/en/open-access-und-open-science/daten-repositories/auffinden-von-daten-in-repositories.html>

UNIVERSITÉ DE GENÈVE, [sans date]. Données de recherche. Définitions. *unige.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.unige.ch/researchdata/fr/footer/definitions/>

UNIVERSITÉ DE LAUSANNE, [sans date]a. L'Open Science à l'UNIL. Données personnelles & sensibles. *unil.ch* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://www.unil.ch/openscience/home/menuinst/open-research-data/conformite--exigences/donnees-personnelles--sensibles.html>

UNIVERSITÉ DE LAUSANNE, [sans date]b. L'Open Science à l'UNIL. Archivage & partage. *unil.ch* [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse : <https://www.unil.ch/openscience/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/archivage--partage.html>

VON DER HEYDE, Markus, 2019. *Open Research Data: Landscape and cost analysis of data repositories currently used by the Swiss research community, and requirements for the future [Report to the SNSF]* [en ligne]. 22.05.2019 [Consulté le 07.11.2020]. Disponible à l'adresse : <https://zenodo.org/record/2643460>

WILKINSON, Mark D., DUMONTIER, Michel, AALBERSBERG, IJsbrand Jan, APPLETON, Gabrielle, AXTON, Myles, BAAK, Arie, BLOMBERG, Niklas, BOITEN, Jan-Willem, DA SILVA SANTOS, Luiz Bonino, BOURNE, Philip E., BOUWMAN, Jildau, BROOKES, Anthony J., CLARK, Tim, CROSAS, Mercè, DILLO, Ingrid, DUMON, Olivier, EDMUNDS, Scott, EVELO, Chris T., FINKERS, Richard, GONZALEZ-BELTRAN, Alejandra, GRAY, Alasdair J.G., GROTH, Paul, GOBLE, Carole, GRETHE, Jeffrey S., HERINGA, Jaap, 'T HOEN, Peter A.C, HOOFT, Rob, KUHN, Tobias, KOK, Ruben, KOK, Joost, LUSHER, Scott J., MARTONE, Maryann E., MONS, Albert, PACKER, Abel L., PERSSON, Bengt, ROCCA-SERRA, Philippe, ROOS, Marco, VAN SCHAIK, Rene, SANSONE, Susanna-Assunta, SCHULTES, Erik, SENGSTAG, Thierry, SLATER, Ted, STRAWN, George, SWERTZ, Morris A., THOMPSON, Mark, VAN DER LEI, Johan, VAN MULLIGEN, Erik, VELTEROP, Jan, WAAGMEESTER, Andra, WITTENBURG, Peter, WOLSTENCROFT, Katherine, ZHAO, Jun et MONS, Barend, 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*. décembre 2016. Vol. 3, n° 1, pp. 160018. DOI [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

ZENODO, [sans date]a. About Zenodo. *zenodo.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse : <https://about.zenodo.org/>

ZENODO, [sans date]b. Faculty of Biology and Medicine at University of Lausanne & Lausanne University Hospital. *zenodo.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: https://zenodo.org/communities/fbm_chuv/?page=1&size=20

ZENODO, [sans date]c. Atmospheric physics group, Institute for Atmospheric and Climate Science, ETH Zurich. *zenodo.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: https://zenodo.org/communities/eth_zurich_iac_atmospheric_physics/?page=1&size=20

ZENODO, [sans date]d. Research Data University of Basel. *zenodo.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: https://zenodo.org/communities/rdm_unibas/?page=1&size=20

ZENODO, [sans date]e. University of Zurich. *zenodo.org* [en ligne]. [Consulté le 07.11.2020]. Disponible à l'adresse: <https://zenodo.org/communities/uzh/?page=1&size=20>

ZENODO, [sans date]f. Frequently Asked Questions. *zenodo.org* [en ligne]. [Consulté le 08.11.2020]. Disponible à l'adresse : <https://help.zenodo.org/>