

h e g

Haute école de gestion
Genève

La gestion des données de recherche à l'Université de Lausanne : enjeux transdisciplinaires

The logo for UNIL (Université de Lausanne) is a stylized, cursive blue script that reads "Unil".

UNIL | Université de Lausanne

Ressources informationnelles
et archives

Travail de Bachelor réalisé en vue de l'obtention du Bachelor HES

par :

Carmen JAMBÉ

Conseiller au travail de Bachelor :

René Schneider, Prof. Dr. HEG-GE

Lausanne, le 29 mai 2015

Haute École de Gestion de Genève (HEG-GE)

Filière Information documentaire

Déclaration

Ce travail de Bachelor est réalisé dans le cadre de l'examen final de la Haute école de gestion de Genève, en vue de l'obtention du titre de *Bachelor of Science HES-SO en Information documentaire*.

L'étudiant atteste que son travail a été vérifié par un logiciel de détection de plagiat.

L'étudiant accepte, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans le travail de Bachelor, sans préjuger de leur valeur, n'engage ni la responsabilité de l'auteur, ni celle du conseiller au travail de Bachelor, du juré et de la HEG.

« J'atteste avoir réalisé seule le présent travail, sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Lausanne, le 29 mai 2015

Carmen Jambé

Note de l'auteure : Par commodité, le masculin est utilisé dans ce document mais s'applique bien évidemment aux deux genres.

Remerciements

Mes remerciements chaleureux s'adressent à toutes les personnes qui m'ont aidée à réaliser ce travail de Bachelor :

Mon conseiller pédagogique M. René Schneider, Prof. Dir. HEG-GE, que je remercie pour son soutien, son suivi régulier et ses conseils fort utiles ;

Mon mandant, M. Gérard Bagnoud, Chef du Service UNIRIS, mille mercis pour sa confiance, la qualité de son encadrement et sa très grande disponibilité ;

Ma jurée, Mme Eliane Blumer, Spécialiste en Information documentaire, UNIGE, merci d'avoir accepté cette responsabilité ;

Que soient également remerciés les collaboratrices et collaborateurs du Service UNIRIS pour leur accueil chaleureux. Mon intégration dans l'équipe a été facilitée par leur gentillesse et leur soutien. Je remercie particulièrement Eva, Nathalie, Olivier et Raphaël, leur expertise m'a indéniablement été précieuse durant ce projet.

Mon mandat m'a menée à rencontrer beaucoup d'actrices et acteurs au sein de l'Université de Lausanne – membres du Dicastère « Recherche & Relations internationales », du Secrétariat général et des Services centraux, chercheuses et chercheurs, personnel administratif et technique, entre autres. Je tiens à les remercier pour leur enthousiasme et leur disponibilité lors des entretiens.

Je remercie également toutes les personnes externes à l'UNIL, en particulier les collaboratrices et collaborateurs de l'EPFL, l'ETHZ, FORS et LIVES qui ont accepté de prendre le temps de me rencontrer et partager ainsi leur expérience enrichissante de la gestion des données de recherche.

Enfin je ne saurais assez remercier ma famille et mes proches, mes relectrices et relecteurs, pour leur indéfectible présence, leur patience et leurs constants encouragements.

Résumé

Le présent travail est effectué pour le Service des Ressources informationnelles et archives (UNIRIS) de l'Université de Lausanne (UNIL) dont la responsabilité couvre la gestion de l'ensemble du cycle de vie des informations et documents d'archives produits au sein de l'UNIL. À l'instar de n'importe quel lieu de savoir, l'UNIL héberge des chercheurs qui produisent des données de recherche hétérogènes et qui n'impliquent pas les mêmes problématiques d'accès, de diffusion, de partage ou de sécurisation par exemple. Certaines constituent le point départ d'une recherche là où d'autres en sont l'aboutissement. Néanmoins, toutes ces données de recherche ont en commun de constituer un « patrimoine » académique que l'UNIL, au travers son Service UNIRIS, entend conserver et valoriser. Ce travail de Bachelor propose ainsi de nourrir une réflexion sur les moyens qui peuvent être mis en œuvre, notamment par le Service UNIRIS, pour sensibiliser les chercheurs à la gestion des données de recherche et les accompagner dans ce processus qui nécessite l'aiguillage de compétences spécifiques transversales.

Il s'agit d'abord de présenter les concepts-clés et fondements théoriques puis fournir des éléments de contextes internes et externes pour établir ce qui a déjà été effectué et identifier qui en sont les acteurs (section 1). Sont ensuite présentés les résultats d'une large enquête et d'une série d'entretiens, effectués tout au long du mandat, destinés à établir un « état des lieux » UNIL et qui nous permettent de déterminer les besoins, les attentes ou les craintes des chercheurs face à un domaine dont l'importance est, certes reconnue, mais la maîtrise encore trop hasardeuse (section 2). De ces résultats, nous pouvons établir notre Guide des bonnes pratiques destiné à l'ensemble des chercheurs, indépendamment de la Faculté à laquelle ils sont rattachés (section 3). Enfin, nous proposons nos recommandations au Service UNIRIS pour développer des ressources propres à la gestion des données de recherche produites au sein de l'UNIL (section 4). Au terme de ce travail, nous espérons pouvoir démontrer l'importance d'envisager le problème de la gestion des données de recherche sous un angle résolument transdisciplinaire en faisant collaborer, autant que possible, l'ensemble de la communauté UNIL avec les spécialistes de l'Information documentaire.

Mots-clés : *Digital Curation* ; Données de recherche (*Research Data*) ; Gestion des données de recherche (*Research Data Management*) ; Plan de gestion des données (*Data Management Plan*) ; Université de Lausanne ; Service UNIRIS.

Table des matières

Déclaration.....	i
Remerciements	ii
Résumé	iii
Table des matières.....	iv
Liste des tableaux	vii
Liste des figures.....	vii
Abréviations et sigles	viii
Glossaire de termes spécifiques	xi
1. Introduction générale.....	1
1.1 Intérêt pour la thématique et problématique.....	1
1.2 Cadre général du travail : mandat, livrable et objectifs.....	4
1.3 Méthodologie générale	5
1.4 Limites et contraintes du travail	6
1.5 Structure du travail	7
SECTION 1 : La gestion des données de recherche (<i>Research Data Management</i>).....	8
2. La donnée de recherche (<i>Research Data</i>)	8
2.1 Définitions de la donnée de recherche.....	8
2.2 Nature, classification et typologie de la donnée de recherche.....	9
2.3 Le cycle de vie de la donnée de recherche	10
3. De la gestion des données de recherche (<i>Research Data Management</i>) à la <i>Digital Curation</i>	11
3.1 La <i>Digital Curation</i>	11
3.1.1 <i>DCC Curation Lifecycle Model</i>	12
3.2 La Politique de gestion des données	13
3.3 Le Plan de gestion des données (<i>Data Management Plan</i>).....	14
3.4 Parties prenantes : compétences, rôles et responsabilités	15
4. Contexte interne	15
4.1 L'Université de Lausanne (UNIL)	15
4.2 Le Service des ressources informationnelles et archives (UNIRIS).....	16
4.3 La recherche à l'UNIL.....	17
5. Contexte externe	17
5.1 Institutions nationales	17
5.1.1 Fonds national suisse de la recherche scientifique (FNS)	18
5.1.2 Swissuniversities – CUSP-2.....	18
5.1.3 Académies suisses des sciences (ASS).....	19

5.2 Projets développés sur le campus lausannois	19
5.2.1 FORS / DARIS	19
5.2.2 PRN LIVES	20
5.2.3 Initiatives au sein des Facultés.....	21
5.2.4 EPFL.....	22
5.3 Niveau international.....	22
5.3.1 Royaume-Uni.....	22
5.3.2 Europe	23
5.3.3 Une problématique mondiale	23
SECTION 2 : Enquête et entretiens.....	25
6. Enquête « Gestion des données de recherche à l'UNIL »	25
6.1 Méthodologie et objectifs de l'enquête.....	25
6.2 Présentation des résultats.....	28
7. Entretiens individuels	49
7.1 Méthodologie et objectifs des entretiens	49
7.2 Présentation des entretiens	50
7.2.1 Tableau croisé et synthèse des entretiens menés dans les Facultés	50
7.2.2 Analyse et synthèses des autres entretiens	53
SECTION 3 : Guide des bonnes pratiques pour la gestion des données de recherche.....	55
8. Méthodologie du guide	55
9. Guide des bonnes pratiques destiné aux chercheurs de l'UNIL	56
SECTION 4 : Recommandations, synthèse et conclusion.....	64
10. Recommandations pour le Service UNIRIS.....	64
10.1 Recommandations orientées vers la sensibilisation et la formation des chercheurs de l'UNIL.....	64
10.2 Recommandations orientées vers l'élaboration d'un Centre de soutien à la gestion des données de recherche à l'UNIL	64
10.2.1 Poursuite des relations engagées et élargissement du réseau	64
10.2.2 Constitution d'un groupe interne à l'UNIL.....	65
10.2.3 Création d'un Centre de soutien à la gestion des données de recherche à l'UNIL	66
11. Synthèse générale et conclusion.....	67
Bibliographie	69
Annexe 1 : Classification des données de recherche.....	75
Annexe 2 : Détail des étapes du cycle de vie de la donnée de recherche	76
Annexe 3 : Détail des actions du <i>DCC Curation Lifecycle Model</i>	77
Annexe 4 : <i>Research Lifecycle at University of Central Florida</i>.....	79

Annexe 5 : Politique de gestion des données de recherche de l'Université d'Oxford	80
Annexe 6 : Modèle de Plan de gestion des données	82
Annexe 7 : <i>Core skills for Data Management</i>.....	85
Annexe 8 : <i>Summary Table of Roles, Rights, Responsibilities and Relationships</i>.....	86
Annexe 9 : Organigramme de la Direction UNIL.....	87
Annexe 10 : <i>Manifesto on Data access and research transparency (DART) in Switzerland</i>.....	88
Annexe 11 : Existants à l'EPFL et l'ETHZ.....	90
Annexe 12 : Questionnaire « Gestion des données de recherche à l'UNIL »	91
Annexe 13 : Réponses et statistiques de l'enquête « Gestion des données de recherche à l'UNIL »	119
Annexe 14 : Guide d'entretien.....	129
Annexe 15 : Version abrégée du Guide et « checklist » de la GDR	131
Annexe 16 : Exemples de posters de notre enquête	132
Annexe 17 : « Brainstorming » Centre de soutien à la gestion des données de recherche.....	133

Liste des tableaux

Tableau 1 : Enquête UNIL / Détail des types de recherches par Faculté (N=531)	29
---	----

Liste des figures

Figure 1 : <i>Spectra of Data</i>	10
Figure 2 : <i>Research Data Lifecycle</i> (UK Data Archive)	11
Figure 3 : Le <i>DCC Curation Lifecycle Model</i>	13
Figure 4 : Enquête UNIL / Types de recherche par Faculté (N=525)	29
Figure 5 : Enquête UNIL / Recherche individuelle ou en équipe (N=310)	30
Figure 6 : Enquête UNIL / Modes de financement (N=511)	30
Figure 7 : Enquête UNIL / Types de données produites (N=1507)	31
Figure 8 : Enquête UNIL / Importance des données de recherche (N=311).....	32
Figure 9 : Enquête UNIL / Création du Data Management Plan (N=310).....	33
Figure 10 : Enquête UNIL / Stockage des données (N=1397)	35
Figure 11 : Enquête UNIL / Estimation du volume de stockage (N=310)	35
Figure 12 : Enquête UNIL / Pour ou contre le partage des données (N=310).....	37
Figure 13 : Enquête UNIL / Arguments pour le partage (N=937)	37
Figure 14 : Enquête UNIL / Arguments contre le partage (N=280)	38
Figure 15 : Enquête UNIL / Partage après publication (N=624)	39
Figure 16 : Enquête UNIL / Estimation de la durée de conservation (N=310)	40
Figure 17 : Enquête UNIL / Propriété des DR (N=540)	42
Figure 18 : Enquête UNIL / Participation à des formations (N=310).....	43
Figure 19 : Enquête UNIL / Besoins en formations (N=1221)	43
Figure 20 : Enquête UNIL / Besoins en information (N=310)	44
Figure 21 : Enquête UNIL / Politique spécifique institutionnelle UNIL (N=310)	44
Figure 22 : Enquête UNIL / Tranches d'âge par Faculté (N=310)	46
Figure 23 : Enquête UNIL / Répartition au sein des Facultés (N=310).....	46
Figure 24 : Enquête UNIL / Fonctions par Faculté (N=309)	47
Figure 25 : Enquête UNIL / Années d'expérience en recherche (N=310)	47

Abréviations et sigles

AAS	Académies suisses des sciences
ASSHS	Académie suisse des sciences humaines et sociales
BDD	Base(s) de données
CC	<i>Creative Commons</i>
CHUV	Centre hospitalier universitaire vaudois
CUS	Conférence Universitaire Suisse [aujourd'hui <i>Swissuniversities</i>]
CUS P-2	Programme CUS 2013-2016 P-2 « <i>Information scientifique : accès, traitement et sauvegarde</i> »
DARIAH	<i>Digital Research Infrastructure for the Arts and the Humanities</i>
DASCH	<i>Data and Service Center for the Humanities</i>
DC	<i>Digital Curation</i>
DCC	<i>Digital Curation Centre</i> [Grande-Bretagne]
CDP	Centre de données et de prestations de services de l'ASSHS
DHDC	<i>Digital Humanities Data Curation</i>
DMP	<i>Data Management Plan</i>
DOI	<i>Digital Object Identifier</i>
DR	Donnée(s) de recherche
ELN	<i>Electronic Laboratory Notebook</i>
Fac.	Faculté
FBM	Faculté de biologie et de médecine de l'UNIL
FDCA	Faculté de droit, des sciences criminelles et d'administration publique de l'UNIL
FORS	Centre de compétences national pour les sciences sociales
FNS	Fonds national suisse de la recherche scientifique
FTSR	Faculté de théologie et de sciences des religions de l'UNIL

GDR	Gestion des données de recherche
GSE	[Faculté des] Géosciences et environnement de l'UNIL
H2020	Horizon 2020 – Programme cadre de l'Union européenne pour la recherche et l'innovation
HEC	[Faculté des] Hautes Études Commerciales de l'UNIL
HEG-GE	Haute École de gestion de Genève
ID	Information documentaire
IDHEAP	Institut des Hautes études en administration publique intégré à l'UNIL
ISO	Organisation internationale de normalisation
IT	<i>Information technology</i> [ou technologies de l'information]
JISC	<i>Joint Information Systems Committee</i> [Grande-Bretagne]
LADHUL	Laboratoire de cultures et humanités digitales
Lettres	[Faculté des] Lettres de l'UNIL
LIMS	<i>Laboratory Information Management System</i>
NSF	<i>National Science Foundation</i> [États-Unis]
OA	<i>Open Access</i>
OAI-PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
OAIS	<i>Open Archival Information System</i> [ou Système ouvert d'archivage d'information]
OCDE	Organisation de coopération et de développement économiques
ORD@CH	<i>Open Research Data Platform Switzerland</i>
PACTT	Bureau de transfert de technologie de l'UNIL-CHUV
PGD	Plan de gestion des données
PI	<i>Principal Investigator</i>
PNR	Programmes nationaux de recherche du FNS
PRN	Pôle de recherche national du FNS

RDM	<i>Research Data Management</i>
SALSAH	<i>System for Annotation and Linkage of Sources in Arts and Humanities</i>
Sc. Crim.	Sciences criminelles
SEFRI	Secrétariat d'État à la formation, à la recherche et à l'innovation
SERVAL	Serveur académique lausannois
SIB	<i>Swiss Institute of Bioinformatics</i>
SSP	Faculté des sciences sociales et politiques de l'UNIL
TIC	Technologies de l'information et de la communication
UE	Union européenne
UNICOM	Service de communication et d'audiovisuel de l'UNIL
UNIGE	Université de Genève
UNIL	Université de Lausanne
UNIRIS	Service des ressources informationnelles et archives de l'UNIL
UNISIS	Service Système d'Information et Statistiques de l'UNIL
VRE	<i>Virtual Research Environment</i>

Glossaire de termes spécifiques

Creative Commons : licences développées comme alternative à la propriété intellectuelle et au *copyright*. L'auteur peut définir et combiner ensemble un certain nombre de conditions cadres qui régiront la future réutilisation de l'œuvre par des tiers : « paternité », « partage des conditions initiales à l'identique », « pas de modification » ou encore « pas d'utilisation commerciale ».

Dépôt ou **repository** : serveur hébergeant les données ou résultats de recherche et qui assure leur accessibilité, leur sécurité ou encore leur pérennité par l'utilisation de certains standards d'archivage (ex. OAIS) ou protocoles d'échange (ex. OAI-PMH pour les métadonnées).

DOI (*Digital object identifier*) : identifiant d'objet numérique qui permet de le localiser et maintenir le lien vers des objets digitaux de manière pérenne.

« **Golden road** » : modèle d'OA où les résultats de recherche, en plus d'être publiés dans une revue scientifique, sont mis à disposition dans un serveur institutionnel (SERVAL pour l'UNIL).

« **Green road** » : modèle d'OA où les résultats de recherche sont publiés dans une revue scientifique en OA. Certaines de ces revues imposent à l'auteur de payer pour être publié.

Impact Factor ou **facteur d'impact** : évaluation des revues scientifiques, des articles et publications, basée sur le calcul du nombre de citations.

Laboratory Information Management System (LIMS) ou **Systèmes de gestion de l'Information du Laboratoire** : systèmes utilisés dans les laboratoires de recherche composés de différentes applications, fonctionnalités et modules, et qui permettent par exemple la gestion des cahiers de laboratoires électroniques (*ELN*), la gestion de la communication, des coûts, des équipements, des relevés, entre autres.

Métadonnée : « donnée sur la donnée ». Les métadonnées peuvent être de différente nature : administrative (de gestion) et technique, de description, d'utilisation ou encore de préservation.

Modèle « hybride » : modèle de publication en OA à la fois en « *Green road* » et en « *Golden road* ».

Open Access (OA) : libre accès et mise à disposition des publications scientifiques en ligne. Ces dernières restent toutefois protégées par le droit d'auteur ou les licences libres (*Creative Commons* par exemple).

Output : résultat de recherche, généralement les articles et les publications.

Peer-review ou **évaluation par les pairs** : processus d'évaluation des résultats de recherche et articles, avant leur publication dans les revues scientifiques, par d'autres experts du domaine. Selon les revues, ces derniers peuvent être anonymes ou non.

Rankings : classements internationaux qui évaluent, selon plusieurs critères et indicateurs (parfois controversés), la qualité, la visibilité, le prestige, entre autres, des Universités et de leurs pôles de recherche.

Set de données ou **dataset** : collection ou corpus de données organisées et regroupées ensemble par le chercheur.

Virtual Research Environment (VRE) ou **environnements de recherche virtuels** : logiciels utilisés pour la recherche, le traitement et l'exploitation des données et dont les fonctionnalités peuvent être multiples : simulation, visualisation, data/text mining (« fouilles » de données/textes). Ils offrent également la possibilité aux chercheurs de travailler de manière collaborative.

1. Introduction générale

1.1 Intérêt pour la thématique et problématique

« *The new scientific revolution : Reproducibility at last* » voilà ce que titrait, fin janvier 2015, le Washington Post dans la rubrique Santé & Science de son édition en ligne. Véritable plaidoyer pour une recherche reproductible – il est bien trop souvent impossible de reproduire les expérimentations et de vérifier les résultats obtenus, ce qui ne signifie pas que ces derniers soient nécessairement faux ou frauduleux (Achenbach, 2015) –, cet article met en lumière un certain nombre de nouveaux phénomènes sociétaux et technologiques, qui bouleversent le monde scientifique et modifient les pratiques des chercheurs¹, tels que le développement des technologies de l'information et de la communication (TIC), la production de masse de données (*Big Data*), les mouvements citoyens en faveur d'un accès libre et ouvert (*Open Access*) – également pour la Science (*Open Science*) et pour les données (*Open Data*) – et qui soulèvent de nombreux intérêts économiques (des industriels et des éditeurs scientifiques en particulier). Mais il soulève justement les problèmes de l'opacité et du manque d'accès aux données sur lesquelles reposent les résultats de recherche.

Dans ce XXI^e siècle parfois qualifié de « siècle des données », la gestion des données de recherche (GDR)² – données produites par les chercheurs – apparaît indispensable et cruciale à de multiples égards : elle assure la conformité avec le cadre légal, réglementaire, éthique et déontologique ; elle répond aux exigences des bailleurs de fonds et éditeurs scientifiques qui encouragent de plus en plus l'*Open Access* (OA) et dont certains, parmi les premiers, exigent des Plans de gestion³ avant l'octroi de financements ; elle fait gagner du temps, facilite le partage des données et réduit les coûts de reproduction ; elle garantit l'authenticité, l'intégrité, la fiabilité, la transparence

¹ L'on parle désormais volontiers de *eScience* ou de *eResearch*, concepts qui ne sont pas encore totalement définis (Martin, 2014, p. 7) : « [...] *eResearch* is not simply an extension of *eScience*, even though it extends *eScience* practices to other disciplines. *eResearch* is a broader concept taking into consideration not only global computational science projects, but any research activity performed digitally at any scale. Furthermore, the concept of *eScience* is broadening to encompass nonscience disciplines such as the social sciences and humanities. These developments open up avenues for mutually beneficial collaboration between different research disciplines and, coupled with the use of advanced digital technology, will continue to « diffuse » disciplinary boundaries and foster breakthrough research. »

² Nous utiliserons également dans ce travail le terme équivalent anglais de *Research Data Management* (RDM).

³ Dans ce travail, nous alternerons les termes Plan de gestion des données (PDG) et *Data Management Plan* (DMP).

et la qualité des données⁴ ; enfin elle est un moyen de rendre davantage visible et de valoriser les travaux et résultats de recherche. Cette gestion ne relève pas uniquement de la responsabilité des chercheurs et des Universités, d'autres acteurs sont également concernés et doivent être associés à cette démarche transdisciplinaire : spécialistes en information documentaire (ID), informaticiens et experts IT, juristes ou encore éditeurs scientifiques.

Face à cet enjeu et à l'accumulation des données, le monde des Sciences de l'information voit également ses pratiques évoluer. Désormais, la gestion de la production des savoirs – assurée traditionnellement par les spécialistes en information documentaire dans des lieux spécifiques comme les bibliothèques ou les services d'archives – doit aussi se réaliser dans l'univers numérique qui n'est pas clos mais en mutation perpétuelle. Par conséquent, les activités de collecte, traitement, accès, mise à disposition, diffusion, valorisation, archivage ou encore conservation à long terme se voient modifiées et doivent s'adapter en permanence aux évolutions technologiques. De nouvelles infrastructures, collaborations et spécialisations émergent donc, tandis qu'un vocabulaire nouveau est créé : *Data Managers*, *Data Curators*, *Data Librarians*, *Data Specialists*, *Digital Curators*, *Metadata Librarians*, entre autres.

Dès lors, dans ce contexte, l'on parle désormais de la *Digital Curation (DC)* dont le but est de garantir la facilité d'accès aux masses de données de recherche (DR) réutilisables pour développer de nouvelles recherches, rendre possible la diffusion et la circulation du savoir et favoriser l'exploration des données associées dans le but d'enrichir les connaissances tout en assurant le suivi et l'usage des données durant leur cycle de vie par les professionnels et le public. On constate donc que l'enjeu est de taille : cependant la longue expérience du savoir-faire déjà existant en matière de gestion et préservation des écrits constitue un solide socle d'appui pour conduire ces projets nouveaux.

En Suisse, ce nouveau paradigme de gestion a récemment pris son essor notamment avec le lancement en 2013 du programme national *CUS 2013-2016 P-2 « Information scientifique : accès, traitement et sauvegarde »* (CUS P-2) qui réunit Universités et Hautes écoles autour de cette question au moyen de différents projets destinés à développer et adapter les infrastructures, méthodes et outils au monde de la recherche de demain. Sur le plan mondial, la situation est diversifiée. Dans certains pays, ce sujet est débattu et théorisé depuis une dizaine d'années environ. Des initiatives, des

⁴ Principes directeurs de la norme ISO 15489 « Information et documentation – *Records Management* ».

démarches et des projets de gestion ont déjà été réalisés et ont abouti à des résultats concrets.

Dans le cadre du travail de Bachelor, ces diverses expériences serviront de point de départ et d'éventuels modèles d'inspiration pour nourrir notre réflexion et apporter des solutions spécifiques à notre champ d'action : l'état de la gestion des données de recherche à l'Université de Lausanne (UNIL). Se définissant comme lieu d'enseignement et de recherche académique fondée sur l'exploitation et la remise en question d'un certain nombre de données *du passé*, pour produire les données *du présent et du futur*, rendant ainsi le « savoir vivant⁵ », l'UNIL ne peut donc ignorer ces récents et rapides changements et échapper aux nombreuses questions qu'ils soulèvent. Sa *Politique de records management et d'archivage pour une gouvernance informationnelle* évoque ainsi cette thématique de la gestion des données de recherche :

« *La gestion des données issues de l'activité académique (recherche et contenu des enseignements⁶) n'est pas prise en compte dans ce document, mais devra l'être à terme.* »
(UNIL, UNIRIS, 2014, p. 5)

Ce constat, qui relève explicitement l'absence de la prise en compte de la gestion des données à ce jour mais indique qu'elle **devra** l'être à terme, se trouve à l'origine de ce travail de Bachelor. La problématique générale, et partant l'intérêt de ce projet, découle d'une part de la nécessité de réfléchir à des solutions et des réponses pratiques permettant de sensibiliser les chercheurs⁷ de cette institution à la gestion de leurs propres données et d'autre part de constituer et mettre en œuvre un Guide des bonnes pratiques pouvant servir de référence en la matière mais aussi servir de base à un futur projet institutionnel de gestion des données de recherche. En effet, faute d'une politique spécifique en la matière – de directives générales, règlements particuliers, procédures claires – la responsabilité de la gestion des données de recherche est laissée actuellement au soin des chercheurs de l'UNIL, qui reçoivent ou non des consignes de leur Faculté. Or pour que cette gestion soit efficace et que l'accès aux données soit sécurisé tant au plan du chercheur que de l'institution, il est important de pouvoir accompagner ces derniers, les former et les soutenir afin de pouvoir accéder à leurs données, les explorer et rendre ainsi possible leur échange et leur reproductivité au sens de l'article cité.

⁵ Concept général, fil conducteur de l'Université de Lausanne, le « savoir vivant » exprime l'intention de faire évoluer perpétuellement la connaissance et de ne pas la figer dans le temps.

⁶ La gestion des contenus des enseignements n'est pas traitée dans ce travail, néanmoins cette thématique mériterait de faire l'objet d'une étude approfondie.

⁷ Dans le cadre de ce travail, les chercheurs de l'UNIL sont les doctorants et tous les autres chercheurs, indépendamment de leur titre ou fonction, actifs dans la recherche à l'UNIL.

1.2 Cadre général du travail : mandat, livrable et objectifs

Le mandat du travail de Bachelor est réalisé pour le Service des Ressources informationnelles et archives (UNIRIS) de l'UNIL. Il est placé sous la supervision de M. Gérard Bagnoud, Chef du service et responsable du *Records Management*. Le projet consiste à fournir un travail de Bachelor dont le mandat repose sur deux axes principaux.

Le premier axe est constitué d'une exploration de la thématique. Il introduit le sujet de la gestion des données de recherche, de manière théorique et d'un point de vue professionnel de l'information documentaire, sans toutefois faire l'économie des autres parties prenantes impliquées. De cette réflexion générale sur l'état de la situation actuelle concernant les enjeux majeurs et les bonnes pratiques en la matière, des réponses conceptuelles et organisationnelles seront présentées sous forme de recommandations plus spécifiques destinées au service UNIRIS.

Le second axe repose sur l'identification des attentes et des besoins concrets – en matière de gestion des données de recherche – propres à l'ensemble de la communauté des chercheurs de l'UNIL. Pour ce faire, une enquête en ligne et des entretiens individuels seront réalisés. Par la suite, l'analyse des résultats de cette enquête permettra l'élaboration du livrable, un Guide des bonnes pratiques, outil pratique qui se veut transdisciplinaire, donc adapté à l'ensemble des chercheurs de l'UNIL indépendamment de leur Faculté et de leur discipline de recherche. En annexe seront également fournis : une version allégée du guide (de type *checklist*), un modèle de Plan de gestion des données (*Data Management Plan*) ainsi que quelques exemples de posters (infographies) présentant certains résultats de l'enquête.

Objectifs généraux et spécifiques pour UNIRIS

1. Fournir une première réponse pragmatique à la Politique de records management et d'archivage pour une gouvernance informationnelle d'UNIRIS et de l'UNIL, en intégrant la dimension de la gestion des données de recherche ;
 - 1.1. Élaborer une réflexion sur la thématique avec une approche conceptuelle et organisationnelle ;
 - 1.2. Expliciter les enjeux majeurs et les principaux acteurs impliqués dans le domaine ;
 - 1.3. Comparer et analyser les pratiques actuelles externes, afin de mettre en lumière les bonnes pratiques (best practice) et fournir des recommandations adaptées au service UNIRIS ;

2. Identifier les attentes et les besoins des chercheurs de l'UNIL en matière de gestion des données de recherche ;
 - 2.1. Mener une approche transdisciplinaire ;
 - 2.2. Réaliser une enquête auprès des chercheurs ainsi que des entretiens individuels ciblés, afin de collecter des données à la fois quantitatives et qualitatives ;
 - 2.3. Analyser les réponses et les résultats obtenus, en établissant les points de convergence et/ou de divergence au sein de la communauté des chercheurs ;
3. Rédiger un Guide des bonnes pratiques pour la gestion des données de recherche, destiné aux chercheurs de l'UNIL ;
 - 3.1. Adapter et rédiger le guide selon les attentes et les besoins des chercheurs ;
 - 3.2. Proposer des réponses concrètes aux diverses problématiques rencontrées par les chercheurs ;
 - 3.3. Élaborer une version abrégée du guide (de type prospectus ou autre) ;
 - 3.4. (Facultatif, selon le temps restant à disposition : Proposer un modèle de Plan de gestion des données – *Data Management Plan*).

Objectifs généraux et spécifiques pour la HEG

1. Appliquer, dans un cadre pratique, les enseignements théoriques et les compétences acquises à la HEG ;
 - 1.1. Intégrer à la réflexion les fondements des domaines de l'archivistique, du Records Management, de l'informatique et de la gestion de projet ;
2. Produire un travail dont la thématique présente un intérêt pour le monde de la recherche au sein des Hautes écoles spécialisées (HES) dont fait partie la HEG.

1.3 Méthodologie générale

Les moyens mis en œuvre pour réaliser le travail de Bachelor sont à la fois théoriques – par la consultation, la lecture et l'analyse de ressources informationnelles existantes – et également de nature pratique – par des entretiens avec les différents acteurs internes ou externes à l'UNIL. Ainsi, les principaux outils à disposition sont les suivants :

- Sources pour mener la réflexion : ouvrages de référence et littérature professionnelle (monographies, articles et revues) ; études, mémoires et travaux de recherche ; documentation officielle produite par le milieu universitaire, de la recherche, des associations, institutions et organismes publics, nationaux et internationaux ; actes de colloques, conférences, comptes rendus d'événements ;
- Veille documentaire : sites Internet spécialisés, blogs, réseaux sociaux (Twitter), flux RSS, alertes Google ;
- Articles de presse généraliste (journaux grand public), émissions de radio et de télévision ;
- Techniques d'enquête et méthodes de sondage pour déterminer les attentes et besoins des chercheurs et fournir une analyse quantitative des résultats ;

- Contacts, rencontres et entretiens semi-directifs pour préciser ces attentes et besoins et fournir une analyse qualitative des résultats ;
- Dans la mesure du possible, visites d'institutions similaires et/ou partenaires pour pouvoir bénéficier de leur expérience et de comparer les pratiques entre institutions.

1.4 Limites et contraintes du travail

La principale difficulté rencontrée est liée à l'étendue et à la complexité du sujet qui est en perpétuelle évolution. Cette thématique de la gestion des données de recherche soulève de nombreux questionnements tant éthiques, politiques que scientifiques et génèrent une multitude de publications. Il a donc fallu opérer un tri parmi les sources à disposition pour ne retenir que celles pertinentes pour ce travail d'étude. De ce fait, les exemples d'initiatives externes (présentés dans la partie 5. « Contexte externe » du travail) ne contiennent volontairement que les informations et références jugées les plus essentielles et utiles pour la compréhension générale du sujet. Nous partons en effet du principe que le lecteur désireux d'en savoir plus ou de suivre un projet précis ira naturellement consulter la source primaire et officielle tenue à jour.

Autre complexité, celle du monde académique et universitaire. Comme dans n'importe quelle Université, l'UNIL abrite en son sein une multitude d'acteurs aux professions différentes et impliqués à différents niveaux hiérarchiques. Connaître le détail des activités des chercheurs, des Facultés, de la Direction, des Services, des associations ou encore des collaborations externes est impossible, il est donc nécessaire de dégager une vue d'ensemble de la situation interne. De même, maintenir cette vue d'ensemble est indispensable pour pouvoir sensibiliser, faire adhérer et réunir cette communauté autour d'un projet de gestion des données de recherche puisque les attentes et besoins spécifiques en la matière ne sont pas nécessairement les mêmes pour tous.

S'ajoute à ces deux observations un autre obstacle d'ordre politique et décisionnel. A l'heure actuelle, la responsabilité de la gestion des données de recherche n'est ni définie à échelle de la Confédération, ni à échelle de l'ensemble des Universités et Hautes écoles suisses. Il en résulte que l'attribution de cette responsabilité et les moyens pour la mettre en œuvre dépendront très certainement des politiques institutionnelles soutenues par les Directions des Universités. Cet état de fait, qui s'ajoute à l'absence de pratiques uniformisées et définies au niveau national, peut entraîner une difficulté supplémentaire à prendre en compte.

1.5 Structure du travail

Le travail de Bachelor est articulé en quatre grandes sections. La première section a pour but d'introduire le sujet dans les grandes lignes, elle présente donc les termes et les concepts clés – la donnée de recherche, la *Digital Curation*, la Politique de gestion des données de recherche, le Plan de gestion des données (*Data Management Plan*), ou encore les modèles de cycle de vie – en s'appuyant sur certains cas réels et situations existantes. Le contexte interne (UNIL) et le contexte externe (national, européen et international) figurent également dans celle-ci. La deuxième section se divise en deux parties, à savoir l'enquête « Gestion des données de recherche à l'Université de Lausanne » et les entretiens individuels. Chacune des deux parties contient une méthodologie, une présentation des résultats ainsi que leur analyse et synthèse. La section 3 « Guide des bonnes pratiques » contient la méthodologie et ledit guide. Enfin, la dernière section présente les recommandations pour le Service UNIRIS, la synthèse générale et la conclusion du travail.

SECTION 1 : La gestion des données de recherche (*Research Data Management*)

2. La donnée de recherche (*Research Data*)

2.1 Définitions de la donnée de recherche

La donnée (*data*), « *représentation formalisée de l'information, adaptée à la communication, l'interprétation ou le traitement* »⁸ peut être collectée et produite dans un environnement numérique ou non. Selon Bellinger et al. (2004), cette donnée devient information par la compréhension des *relations* ; savoir par la compréhension des *modèles* ; enfin sagesse par la compréhension des *principes*. A la base de la pyramide des connaissances⁹, comment qualifier cette donnée lorsqu'elle est issue du monde scientifique et qu'elle est créée par un chercheur au cours d'une activité et d'un processus de recherche ? Faut-il employer l'expression « donnée scientifique », « donnée de recherche », « donnée de la recherche », « donnée du chercheur¹⁰ », « donnée du travail savant » ou encore « donnée du savoir scientifique » ?¹¹

A l'heure actuelle, il n'existe pas une seule et même définition communément admise par l'ensemble des acteurs et institutions concernés. Néanmoins, nous constatons que bon nombre de définitions se basent sur celle de l'Organisation de coopération et de développement économiques (OCDE) où les données de recherche sont étroitement associées à la validité et aux résultats de recherche. Ainsi, dans ses *Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics*, l'OCDE définit les données de recherche comme des :

« *Enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche.* »
(OCDE, 2007, p. 18)

Essentiellement adaptée à l'environnement numérique, cette définition de l'OCDE exclut certains « *produits ou résultats de recherche* » – carnets de laboratoire, analyses préliminaires, examens par les pairs (*peer reviews*), correspondance ou encore objets matériels – et explicite ce que sont les « sets de données », à savoir les données

⁸ Définition du Comité Consultatif pour les Systèmes de Données Spatiales dans son *Modèle de référence pour un Système ouvert d'archivage d'information (OAIS, 2003)*.

⁹ Modèle hiérarchique fréquemment utilisé en *Knowledge Management* (gestion des connaissances).

¹⁰ Voir les travaux de Bert (2014) sur les « archives du chercheur » (références en bibliographie).

¹¹ Le monde anglo-saxon a su trancher rapidement cette question sémantique en utilisant le terme *Research Data* dont la traduction française la plus proche, à savoir « donnée(s) de recherche », sera employée dans ce travail.

regroupées et organisées en collections et qui constituent une « *représentation systématique, partielle du sujet qui fait l'objet d'étude* » (OCDE, op. cit.).

Du côté des Universités, les définitions ne se limitent toutefois pas uniquement à la relation données-résultats de recherche. Elles tiennent aussi compte de l'ensemble des documents produits durant le processus de recherche et de l'activité du chercheur, précisant parfois l'origine, la nature, le type ou le format des données de recherche. Généralement, elles reprennent les principes directeurs énoncés par les fonds nationaux de recherche qui mettent l'accent sur le partage et l'accès libre aux données (*Open Access*). Dans le cadre de ce travail, nous élaborons notre propre définition [qui est adaptée au contexte de l'UNIL] :

« *Les données de recherche sont tous les enregistrements factuels – de toute nature, de tout type et sur tout support – collectés et produits [dans le cadre de recherches menées à l'Université de Lausanne] et qui permettent de produire des idées originales et des résultats de recherche validés par la communauté scientifique* ».

2.2 Nature, classification et typologie de la donnée de recherche

La nature, la structure et les types de données de recherche sont variables. Selon leur contexte de création (capture ou production), leur exploitation, leur analyse et les traitements qu'elles subissent, les données de recherche peuvent donc être brutes (*raw*), dérivées (*derived*), formatées (*formatted*) nettoyées (*cleaned*), primaires (*primary*), secondaires (*secondary*) ou encore traitées (*processed*). Par conséquent, elles sont de tous types (p. ex. textes, chiffres, statistiques, quantitatives, qualitatives, audio, vidéo,...) et contenues dans divers supports (p. ex. papier, carnets de laboratoire, documents électroniques,...).

Le modèle de classification du *Research Information Network*, qui reprend en grande partie la classification du *National Science Board* aux États-Unis, distingue les données en cinq catégories – données **d'observation** ; données **expérimentales** ; données **computationnelles**, de **modèles** ou de **simulations** ; données **dérivées** ou **compilées** ; données de **référence** ou **canoniques** – selon la manière dont elles ont été produites, c'est-à-dire selon les méthodes et processus utilisés dont elles sont issues (voir notre tableau détaillé en annexe 1).

Lorsqu'elles sont créées dans un environnement numérique particulièrement, l'on peut différencier les données « très bien structurées » des données « hétérogènes » comme le fait l'organisation *Research Data Alliance Europe* (RDA Europe) dans son rapport issu du premier *Science Workshop* organisé en avril 2014 :

Figure 1 : *Spectra of Data*

Well-structured data	Heterogeneous data sets
Data with automatically generated metadata	Data with complex metadata issues
Static data	Dynamically changing data
Data acquired under controlled conditions	Crowd-sourced data
Centrally managed databases	Widely distributed data, no clear curation
Data that are computationally simple to handle	Data needing massive computing
Data that are used "raw"	Data that are understandable only after processing
Numerical data	Text data
Communities knowledgeable about data processing	Communities scared of data
Communities with trust	Communities with no tradition of sharing, even with distrust
Open data	Proprietary/embargoed data, data with copyright issues
Impersonal data	Data with privacy issues
Privately generated data	Data with publicly funded stakeholders

(Source : RDA Europe, 2014, p.11)

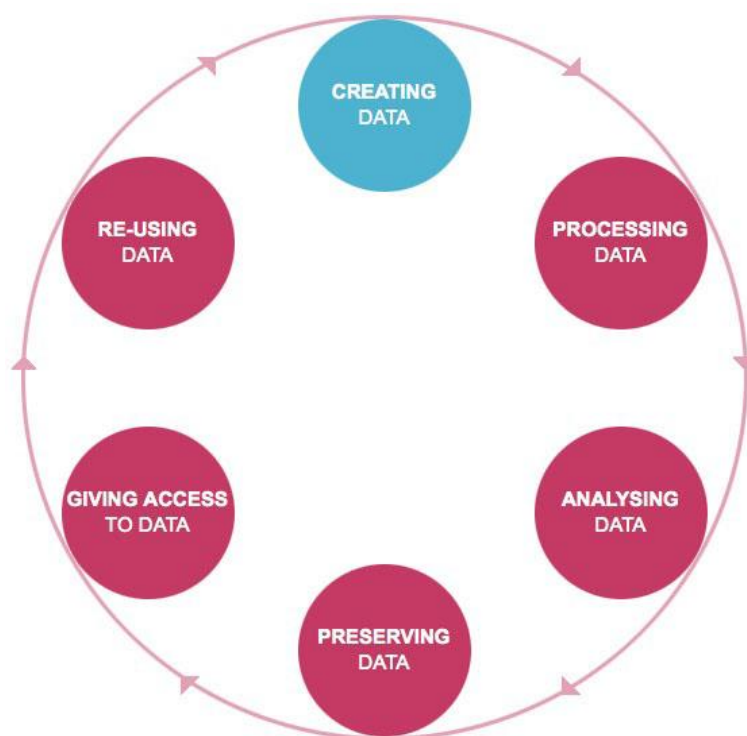
2.3 Le cycle de vie de la donnée de recherche

Le modèle de référence *Research Data Lifecycle* élaboré par *UK Data Archive*¹² définit six principales étapes (voir les détails en annexe 2) par lesquelles les données passent, ou repassent si la rotation du cycle est perpétuelle :

1. **Création** des données
2. **Traitement** des données
3. **Analyse** des données
4. **Préservation** des données
5. **Accès** aux données
6. **Réutilisation** des données

¹² Au Royaume-Uni, UK Data Archive est le centre national chargé de la gestion et de la curation – acquisition, mise à disposition, archivage et préservation notamment – des données issues des sciences humaines et sociales. Voir son site Internet : <http://www.data-archive.ac.uk/>

Figure 2 : *Research Data Lifecycle* (UK Data Archive)



(Source : UK Data Archive, 2002-2015)

3. De la gestion des données de recherche (*Research Data Management*) à la *Digital Curation*

3.1 La *Digital Curation*

La *Digital Curation*¹³ est un concept créé, développé et théorisé par le *Digital Curation Centre (DCC)*¹⁴ : « *Digital curation is maintaining and adding value to a trusted body of digital research data for current and future use; it encompasses the active management of data throughout the research lifecycle.* » (DCC, 2010)

La *Digital Curation* ne consiste pas uniquement à archiver ou préserver sur le long terme les données numériques, elle regroupe en effet l'ensemble des stratégies, processus et activités de « gestion active » des données durant l'entier de leur cycle de vie et avant même leur création. Elle est caractérisée par une série de processus appliqués aux objets digitaux durant l'entier de leur cycle de vie, depuis leur création jusqu'à leur

¹³ Le terme anglais sera utilisé dans ce travail puisque son équivalent en français « curation digitale » n'est pas communément adopté à ce jour.

¹⁴ Créé le 1^{er} mars 2004 en Grande-Bretagne à l'initiative du JISC (*Joint Information Systems Committee*), le *Digital Curation Centre* est le centre national chargé de développer les concepts, élaborer les bonnes pratiques et résoudre les grands défis en matière de *Digital Curation*, pour l'ensemble des institutions et disciplines de recherche. Voir son site Internet : <http://www.dcc.ac.uk/>

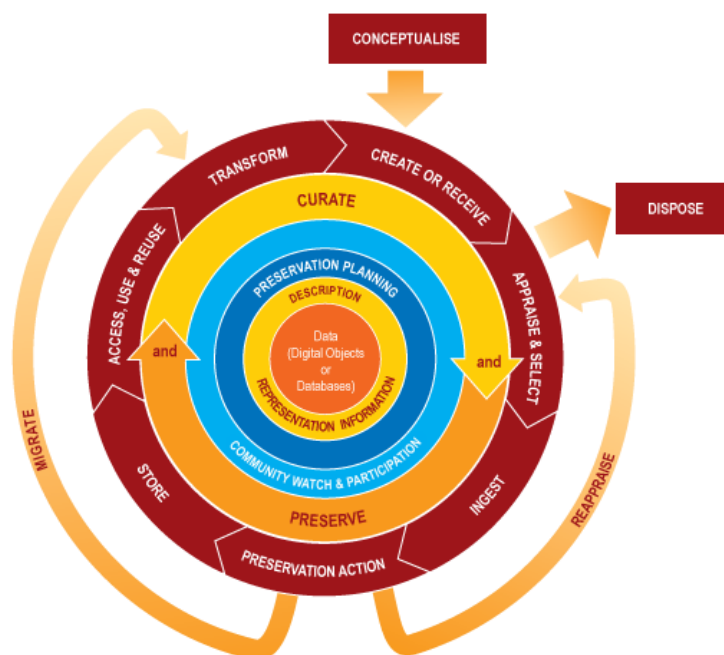
élimination finale ; une préoccupation constante des problématiques de reproductibilité, de responsabilité et de conservation à long terme ; le fait d'ajouter de la valeur aux objets numériques pour qu'ils puissent être mis à disposition et réutilisés ; le fait d'impliquer un large éventail de parties prenantes et de transcender les frontières disciplinaires ; un intérêt fort pour les solutions *Open Source* ; enfin des liens forts entre la recherche et la pratique (Harvey, 2010, pp. 15-16).

3.1.1 DCC Curation Lifecycle Model

Depuis longtemps, les Sciences de l'information – en particulier les milieux de l'archivistique et du *Records Management* – ont privilégié l'approche par cycle de vie (*Lifecycle*) pour gérer et préserver l'information (Higgins, 2008). Dans le contexte de la *Digital Curation*, l'approche par cycle de vie est nécessaire. Elle constitue un scénario « idéal » et est justifiée par trois raisons principales (Pennock, 2007) : la fragilité des *digital materials* et leurs possibles altérations dues aux évolutions technologiques ; leur bonne ou (surtout) mauvaise gestion, à une certaine étape, répercutée sur les étapes suivantes ; leur réutilisation, rendue possible par la *Digital Curation* qui assure l'authenticité, la fiabilité et l'intégrité. Destiné à la fois aux créateurs de données, aux curateurs et aux utilisateurs, le *DCC Curation Lifecycle Model* – LE modèle de référence en matière de *Digital Curation* – a été conçu pour être applicable dans n'importe quel contexte et discipline de recherche (Harvey, 2010, p. 34). Il est défini comme tel :

« The DCC Curation Lifecycle Model provides a graphical high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. The model can be used to plan activities within an organisation or consortium to ensure that all necessary stages are undertaken, each in the correct sequence. The model enables granular functionality to be mapped against it; to define roles and responsibilities; and build a framework of standards and technologies to implement. It can help with the process of identifying additional steps which may be required, or actions which are not required by certain situations or disciplines, and ensuring that processes and policies are adequately documented. »
(Higgins, 2008)

Figure 3 : Le *DCC Curation Lifecycle Model*



(Source : DCC, 2015)

Avec au centre la donnée, l'objet numérique ou la base de donnée – qui ont leur propre définition – ce *DCC Curation Lifecycle Model* identifie trois principaux types d'actions (Harvey, 2010, p. 34) : *Full Lifecycle Actions*¹⁵ (couches jaunes, bleues et flèches jaunes-oranges) ; *Sequential Actions*¹⁶ (en rouge) ; enfin les *Occasional Actions*¹⁷ (flèches extérieures oranges). Le détail des définitions et des actions se trouve en annexe 3. Nous pouvons toutefois mentionner l'existence d'autres modèles de cycles adaptés à notre thématique générale, en particulier les modèles plus généraux de cycle du processus de recherche ou *Research Lifecycle* (voir un exemple de ce type de modèle réalisé par la bibliothèque de *University of Central Florida* en annexe 4).

3.2 La Politique de gestion des données

Appuis nécessaires pour délimiter et instaurer un cadre de gouvernance des données de recherche, les Politiques de gestion des données de recherche (*Research Data Policies*) sont des outils qui garantissent une gestion efficace et transparente des données de recherche. Les Universités sont donc fortement encouragées à élaborer de

¹⁵ Description et représentation de l'information, planification de la préservation, participation et suivi de la communauté, curation et préservation.

¹⁶ Conceptualisation, création et collecte, évaluation et sélection, *ingest*, actions de préservation, stockage, accès, utilisation et réutilisation, transformation.

¹⁷ Élimination, réévaluation, migration.

tels documents (voir la Politique d'Oxford en annexe 5) – généralement composés d'une dizaine de points qui :

- Rappellent les grands principes directeurs qui encadrent la recherche (p. ex. éthique, intégrité, transparence, accessibilité, reproductibilité) ainsi que le cadre légal et réglementaire ;
- Précisent les rôles et responsabilités de chacun dans la gestion des données de recherche (en particulier les chercheurs, l'Université et ses services institutionnels) ;
- Spécifient parfois certains aspects plus précis de la gestion active des données (p. ex. stockage, sécurité¹⁸, accès, partage, préservation, conservation à long terme, destruction) ;
- Définissent les termes et les notions essentielles (définitions, glossaire) ;
- Mentionnent l'existence de politiques (directives, règlements internes) préexistantes ;
- Anticipent l'actualisation et la future mise-à-jour de ladite politique de gestion des données de recherche.

3.3 Le Plan de gestion des données (*Data Management Plan*)

Le Plan de gestion des données (voir notre modèle de plan en annexe 6) n'est utile et efficace que si un cadre de gestion des données de recherche est instauré et soutenu par l'existence de la Politique de gestion que nous venons d'aborder plus haut. De plus en plus exigé par les bailleurs de fonds – notamment sous l'impulsion du *National Science Foundation (NSF)* aux États-Unis qui les impose depuis 2011 – le *Data Management Plan* est un document à réaliser avant même le début d'un projet de recherche, il n'est pas figé dans le temps – dans le sens où il est évolutif et peut/doit être complété tout au long de la recherche. Les principales informations qu'il contient précisent (Jones, 2011) :

- **La nature et le contexte du projet de recherche** : question principale de recherche, chercheur principal (*Principal Investigator*), bailleurs de fonds ;
- **Les données de recherche** : leur type, formats, standards, méthodes de capture et production ;
- **Questions éthiques, juridiques et déontologiques** : droit d'auteur, propriété intellectuelle, confidentialité et protection des données sensibles ;
- **Accès, partage et réutilisation des données** : droits, permissions, infrastructure, collaborations ;
- **Stockage à moyen terme et gestion des données** : assurance qualité, backups, procédures, responsabilités ;

¹⁸ Nous mentionnons ici la *Politique de sécurité des données de recherche* de l'Université d'Harvard qui a fait le choix de définir cinq niveaux de sécurité – donc cinq types de données – selon le potentiel risque et préjudice que la divulgation de ces dernières pourrait provoquer.

- **Dépôt et préservation à long terme** : sélection (évaluation), sécurité, accessibilité.

3.4 Parties prenantes : compétences, rôles et responsabilités

Pour assurer la préservation à long terme et le partage des données, il est essentiel de mobiliser tous ces acteurs et de mettre en commun leur expertise (Ray, 2014, p. 10). La gestion des données de recherche et la *Data Curation* implique donc un certain nombre de parties prenantes aux compétences (*skills*), rôles et responsabilités (détaillés aux annexes 7 et 8) :

- **Chercheurs (ou producteurs)**
- **Universités**
- **Spécialistes et professionnels de l'Information documentaire (ID)**
- **Informaticiens / IT**
- **Juristes**
- **Bailleurs de fonds – Fonds nationaux de recherche**
- **Éditeurs scientifiques**
- **Spécialistes de la communication**

Impliquant ces différents acteurs, la gestion des données de recherche ne peut donc être efficace dans sa réalisation autant que dans son coût économique qu'à la seule condition d'être « transversale ». Nous voulons insister sur le fait que les données de recherche s'inscrivent, au propre comme au figuré, dans un réseau si bien que la collaboration entre les différents pôles s'avère nécessaire. Toutefois, cette gestion transdisciplinaire soulève aussi son lot de questions, souvent d'ordre éthique, telles que : les données de recherche appartiennent-elles au chercheur, au bailleur de fonds ou encore à l'éditeur ? Dans un souci d'intégrité, les données de recherches doivent-elles être systématiquement publiées « en annexe » de la recherche ?

4. Contexte interne

4.1 L'Université de Lausanne (UNIL)

Les missions¹⁹ de l'Université de Lausanne sont définies dans la Loi cantonale vaudoise sur l'Université de Lausanne (LUL) du 6 juillet 2004 :

« L'Université a pour missions : a. de transmettre les connaissances et développer la science par l'enseignement et la recherche ; b. de favoriser le développement de la vie intellectuelle et la diffusion de la culture ; c. d'assurer la relève académique et scientifique ; d. de favoriser la valorisation des résultats de la

¹⁹ Il nous semble que la gestion des données de recherche, et plus généralement la *Data Curation*, s'accordent pleinement avec cette mission générale, notamment le point d. sur la valorisation des résultats – donc des données – de recherche.

recherche ; e. de préparer aux professions nécessitant une formation académique ; f. d'organiser une formation continue dans les domaines qui relèvent de sa compétence ; g. d'exercer une fonction de service en faveur de la collectivité et de stimuler le débat de société. »

(Grand Conseil du Canton de Vaud, 2005)

L'UNIL compte sept Facultés : Faculté de théologie et de sciences des religions (FTSR) ; Faculté de droit, des sciences criminelles et d'administration publique (FDCA) ; Faculté des lettres ; Faculté des sciences sociales et politiques (SSP) ; Faculté des hautes études commerciales (HEC) ; Faculté des géosciences et de l'environnement (GSE) ; Faculté de biologie et de médecine (FBM). En termes d'effectifs, ces Facultés hébergent 14'165²⁰ étudiants – dont 2'072 doctorants – inscrits au semestre d'automne 2014.

4.2 Le Service des ressources informationnelles et archives (UNIRIS)

Créé en 2014²¹ et destiné à la fois aux utilisateurs internes – la communauté UNIL – et aux usagers externes, le Service UNIRIS²² a pour mission « *de permettre à la Direction, aux unités de l'Université de disposer, en toute confiance, des informations nécessaires à leur activité au bon moment et sous une forme exploitable* » (UNIRIS, 2014, p. 5). Dans cette perspective, UNIRIS est donc responsable de la gestion de l'entier du cycle de vie des informations et documents d'archives produits au sein de l'Université. Ainsi, les collaborateurs du service apportent leur expertise dans les secteurs d'activité suivants : gestion des archives historiques ; information documentaire et gestion électronique des documents ; numérisation ; technologies de l'information et de la communication ; *Records Management* ; veille documentaire stratégique. Pour répondre à sa mission de « *gouvernance informationnelle* »²³, le Service UNIRIS met donc en œuvre divers projets principalement avec les unités de l'UNIL et instaure un cadre de bonnes pratiques de gestion, en délivrant des formations et en produisant une riche documentation d'aide à ses usagers (p. ex. guides, manuels, procédures, règles,...).

Gérer, archiver et préserver les données de recherche suscite naturellement l'intérêt d'un Service tel UNIRIS. Organisateur des *Journées des archivistes des universités et*

²⁰ Données fournies par le Service Système d'Information et Statistiques (UNISIS) de l'UNIL.

²¹ Rattaché au Secrétariat général de l'UNIL (voir l'organigramme de la Direction en annexe 9), le Service UNIRIS succède ainsi au Service des archives de l'Université de Lausanne (SAUL) fondé en 2004.

²² Voir le site Internet du Service : <http://unil.ch/uniris/home.html>

²³ Définie dans la Directive de la Direction 0.12 « Ressources informationnelles et archives de l'Université de Lausanne » comme étant le « *dispositif stratégique composé de normes, de processus, de rôles qui responsabilisent les collaborateurs de l'UNIL dans la production, la gestion, la sécurité, la conservation, l'utilisation et la destruction de l'information de telle façon qu'ils respectent les objectifs de l'institution et y contribuent.* » (UNIL, 2015, p.2)

*hautes écoles*²⁴ lors desquelles cette thématique a été débattue, UNIRIS prend également part au sous-groupe « Gestion des données de recherche » du groupe Mnémo-Pôle²⁵

4.3 La recherche à l'UNIL

« Inséparable »²⁶ de l'enseignement, la recherche est au cœur des activités de l'UNIL. Composée de 140 unités d'enseignement et de recherche – ce qui représente environ 2'700 chercheurs dont 500 professeurs (UNIL, 2015) – cette communauté contribue, au travers de projets de recherche nationaux et internationaux, à la fois au progrès scientifique, au rayonnement et à la renommée de l'institution. Cette recherche « étendue à tous les niveaux »²⁷ est encadrée à la fois légalement (ex. Loi sur la recherche sur l'être humain) ainsi qu'institutionnellement, par un certain nombre de documents internes qui placent les conditions cadres – *Plan d'intentions de l'Université de Lausanne 2012-2016* ; Directives de la Direction (particulièrement celles 4.1. à 4.4.) ou encore la *Charte du doctorat*. A l'UNIL, les chercheurs ont la possibilité de déposer leurs travaux (thèses, articles, rapports, publications, etc.) sur le dépôt institutionnel SERVAL – Serveur académique lausannois – qui comptabilise à ce jour près de 3'380²⁸ documents en *Open Access*.

5. Contexte externe

5.1 Institutions nationales

Malgré l'absence d'une politique nationale pilotée par la Confédération²⁹ elle-même, l'on assiste, dans le milieu académique, à l'émergence de projets mis en place par différentes institutions qui sont impliquées, à différentes échelles, dans la gestion des données de recherche ou qui, conscientes de la problématique, encouragent les chercheurs à investir dans des solutions pour gérer le *Big Data*. Nous souhaitons

²⁴ Journées qui se sont déroulées le 16 octobre 2014 et dont est issu un dossier thématique « *Archives des savoirs : de la gestion des données de recherche vers une gestion des données pour la recherche* » (UNIL, 2014)

²⁵ Groupement des « *archives et lieux de conservation au service de la connaissance sur le campus lausannois* ». Voir son site Internet : <http://wp.unil.ch/mnemopole/>

²⁶ Terme repris de la Charte de l'Université de Lausanne.

²⁷ Expression issue du *Plan d'intentions de l'Université de Lausanne 2012-2016* (UNIL, 2011) dont la mesure 2.1.1 propose « *d'intégrer, toutes les fois que cela s'avère pertinent, le personnel administratif et technique dans l'élaboration des projets de recherche, au vu du rôle important qu'il est amené à jouer dans leur réalisation.* » (UNIL, 2011, p. 33)

²⁸ Données fournies par la Bibliothèque cantonale et universitaire de Lausanne (BCUL) mi-mai 2015.

²⁹ La Confédération soutient par ailleurs la recherche par le biais de son Secrétariat d'Etat à la formation, à la recherche et à l'innovation (SEFRI), à hauteur d'environ 24 milliards de francs CHF (SEFRI, 2015).

présenter ici quelques-unes des principales institutions nationales ainsi que des projets (parfois nationaux) développés sur le campus lausannois et que nous avons eu la chance de pouvoir rencontrer les acteurs, sur le terrain.

5.1.1 Fonds national suisse de la recherche scientifique (FNS)

Le Fonds national suisse de la recherche scientifique (FNS) est l'un des plus importants bailleurs de fonds et de nombreux chercheurs y sont rattachés soit en touchant des subsides, soit en participant aux programmes ou aux pôles de recherche nationaux. Il accorde une place particulière à l'*Open Access* en « *sout[enant] et encourage[ant] sur le plan national et international le principe du libre accès électronique (open access) pour les publications scientifiques. Il engage les bénéficiaires de subsides à rendre accessibles au public les résultats de la recherche qu'il a soutenue* » (FNS, 2015), via les voies « Green road » et « Golden road ». Précisons encore qu'à ce jour, il n'existe pas de directive particulière concernant la gestion des données de recherche ou l'élaboration d'un éventuel Plan de gestion.

5.1.2 Swissuniversities – CUSP-2

Sous l'égide de Swissuniversities (anciennement CUS³⁰), le projet national *CUS 2013-2016 P-2 « Information scientifique : accès, traitement et sauvegarde »*³¹ représente une avancée majeure en matière de gestion des données de recherche. Ce programme pilote et soutient financièrement des projets d'envergure orientés vers les technologies de demain et a pour objectif central de :

« Encourage[r] le regroupement et le développement des efforts que les hautes écoles déploient actuellement de manière dispersée pour mettre à disposition et traiter des informations scientifiques. Pour renforcer la compétitivité de la place scientifique suisse au niveau international, une réorganisation globale doit permettre de fournir aux chercheurs, aux enseignants et aux apprenants une offre complète de contenus numériques à caractère scientifique ainsi que des instruments optimisés pour leur traitement. Par le biais d'un encouragement ciblé, CUS P-2 initie et pilote le développement de cette offre et garantit un fonctionnement durable. »
(Swissuniversities, 2015)

Parmi la quinzaine de projets actuellement en cours de développement, nous retenons particulièrement certains directement en lien avec notre thématique :

- **DLCM** (*Data Life- Cycle Management*) : avant-projet orienté vers la gestion du cycle de vie de la donnée
- **Train 2 Dacar** (*Train the Trainer for Data Curation in Advanced Research*) : projet de formation de formateurs à la *Data Curation*

³⁰ CUS : Conférence Universitaire Suisse.

³¹ Voir le site Internet du programme : <http://www.swissuniversities.ch/fr/organisation/projets-et-programmes/programme-cus-2013-2016-p-2-information-scientifique-acces-traitement-et-sauvegarde/>

- **Pilot-ORD@CH** (*Open Research Data*) : plateforme³² suisse de méta-catalogue de données de recherche ouvertes
- **DICE+** (*Digital Copyright for E-learning*) : site³³ de ressources et formations à la question du droit d'auteur dans l'enseignement supérieur
- **Projets SWITCH** (infrastructure informatique) : *Swiss edu-ID* (identifiants persistants pour les membres du secteur académique) & *SCALE* (*Swiss Cloud for Academic and Learning Experts*)

Il est à noter que l'Université de Lausanne participe activement à certains de ces projets (voir ci-après), notamment dans le projet DLCM.

5.1.3 Académies suisses des sciences (ASS)

Les Académies suisses des sciences (ASS), regroupant plusieurs disciplines de recherches – Sciences humaines et sociales, sciences médicales, sciences techniques et sciences naturelles – sont un organe fédéral qui veille particulièrement à « l'intégrité dans la recherche scientifique ». Via ses *Principes de bases et procédures* et notamment dans son *Mémoire sur l'intégrité scientifique*, les Académies se positionnent essentiellement sur les problématiques de l'éthique du chercheur et de la recherche mais également sur celles du droit d'auteur³⁴. Elles posent donc un cadre général dans lequel s'intègre la gestion « *intègre et responsable* »³⁵ des données de recherche. Nous relevons que l'Académie suisse des sciences humaines et sociales (ASSHS) mène actuellement un projet-pilote, conjointement dirigé par les Universités de Bâle, Berne et Lausanne, destiné à la création d'un Centre de données et de prestation de services pour les données de recherche en Sciences humaines (CDP/DaSCH) et dont le rapport final du projet est paru en mai 2015³⁶.

5.2 Projets développés sur le campus lausannois

Dans le cadre du mandat de ce travail de Bachelor, nous sommes allée à la rencontre de personnes actives dans certains des projets soutenus par les institutions précitées.

5.2.1 FORS / DARIS

L'Université de Lausanne accueille en son sein le Centre de compétences national pour les sciences sociales FORS³⁷, spécialisé dans l'élaboration et l'analyse d'enquêtes « de

³² La plateforme est disponible à l'adresse : <http://openresearchdata.ch/>

³³ Voir le site Internet du projet : <http://www.diceproject.ch/>

³⁴ Intégrité scientifique et citation des auteurs font l'objet d'analyse et de recommandations – « Qualité d'auteur des publications scientifiques » – disponibles à l'adresse : <http://www.akademien-schweiz.ch/fr/index/Schwerpunktthemen/Wissenschaftliche-Integritaet.html>

³⁵ Expression reprise du Comité éthique du CNRS (Centre national pour la recherche scientifique – France) et de son guide *Promouvoir une recherche intègre et responsable*.

³⁶ Ce rapport est disponible à l'adresse : <http://www.sagw.ch/fr/sagw/laufende-projekte/ddz.html>

³⁷ Voir le site Internet de FORS : <http://forscenter.ch/fr/>

masse » et dont la grande partie des données récoltées sont de nature quantitative. Chargé de la gestion de ces données issues des Sciences sociales, le Service des données et d'information sur la recherche (DARIS) a développé une vision³⁸ et une série de mesures se déclinant en plusieurs axes et dont nous pouvons lister les points forts. Conformément à la ligne défendue par le FNS ou des programmes tel CUSP-2, desquels FORS dépend en partie, les chercheurs en Sciences sociale sont encouragés dans la voie de l'*Open Access* et l'*Open Data*. A cet effet, ils sont incités à déposer leurs données dans le dépôt FORSbase³⁹ librement accessible. Conscients que les chercheurs ne sont pas nécessairement habitués à ces nouvelles dispositions, le service DARIS a élaboré des procédures et des guides destinés à les aider lors des différentes étapes de la gestion de leurs données, qu'elles soient d'ordre technique – lors du dépôt des données – ou d'ordre administratif – lors de contrats à élaborer avec d'éventuels partenaires. La publication fin 2014 d'un *Manifesto to advance Data Access and Research Transparency (DART) in Switzerland* (le Manifeste se trouve en annexe 10) couvrant aussi bien des dispositions générales que des considérations éthiques est aussi à mettre à son actif. A l'international, FORS collabore activement à l'élaboration de standards internationaux en Sciences sociales tel que le standard DDI⁴⁰. On notera finalement que FORS sensibilise et forme des chercheurs en organisant des workshops thématiques, des conférences ou encore des congrès nationaux. A l'échelle de l'Université de Lausanne, FORS, son infrastructure et ses services annexes représentent sans doute le projet le plus abouti en matière de gestion des données de recherche.

5.2.2 PRN LIVES

En plus de FORS, l'Université de Lausanne héberge également le Pôle de Recherche National (PRN) du FNS en Sciences sociales « LIVES – Surmonter la vulnérabilité : perspective du parcours de vie »⁴¹. Il est prévu qu'à terme, les données récoltées par les chercheurs du projet LIVES – données à la fois quantitatives mais également qualitatives – soient déposées chez FORS et rendues accessibles via FORSbase. Cette procédure spécifique a incité LIVES à se doter des compétences d'un *Data Manager* également chargé d'établir des procédures pour la description des données et d'assister les chercheurs dans diverses démarches. Cette personne a notamment récemment mis à

³⁸ La « Vision 2020 » de DARIS est disponible à l'adresse : <http://forscenter.ch/wp-content/uploads/2013/11/DARIS-vision-20201.pdf>

³⁹ FORSbase est disponible à l'adresse : <https://forsbase.unil.ch/>

⁴⁰ Elaboré par la *Data Documentation Initiative*, le DDI est un standard international de description et de conservation des métadonnées utilisées en Sciences humaines sociales. Il est principalement utilisé pour les données statistiques ou issues d'enquêtes. Voir le site Internet du DDI : <http://www.ddialliance.org/>

⁴¹ Voir le site Internet de LIVES : <https://www.lives-nccr.ch/>

jour de nouvelles procédures « *Documenter et archiver les enquêtes de LIVES* » (document interne).

5.2.3 Initiatives au sein des Facultés

Moins institutionnalisés mais pas moins inintéressants, d'autres projets voient le jour sous l'impulsion de chercheurs de l'UNIL qui ont parfois des intérêts convergents, notamment lorsqu'il s'agit d'exploiter, de réutiliser, de mettre à disposition ou encore de préserver les données de recherche. C'est, par exemple, le cas du Laboratoire des Cultures et Humanités Digitales (LADHUL)⁴² – rattaché à SSP mais commun à la Faculté des Lettres et la FTSR – qui soulève à lui seul l'enjeu des *Digital Humanities*⁴³. Impliqué et actif dans des projets nationaux (p. ex. SALSAH⁴⁴) et européens (p. ex. Erasmus+⁴⁵ et DARIAH⁴⁶), LADHUL organise d'ailleurs des journées de formation – nous avons pu dans le cadre de ce travail participer à l'une d'entre elles intitulée « Domestiquez vos données ! »⁴⁷.

Nous pouvons également mentionner le projet RechAlp⁴⁸ de la Faculté GSE qui, dans l'idée de récolter et de décrire tous types de données sur les Alpes vaudoises, a développé sa propre typologie de données ainsi que des formulaires pour leur mise à disposition. Dans l'optique d'une recherche éthique, ce projet s'est également doté d'une charte déontologique⁴⁹ reprise du Centre de données sur les espèces de Suisse (*InfoSpecies*⁵⁰), autre projet scientifique d'envergure nationale et soutenu par la Confédération.

Tous ces projets mettent en lumière que la problématique de la gestion des données de recherche est devenue incontournable pour un grand nombre de chercheurs et que

⁴² Voir le site Internet de LADHUL : <http://www.unil.ch/ladhul/fr/home.html>

⁴³ Les humanités digitales se définissent comme : « *Digital Humanities is an important multidisciplinary field, undertaking research at the intersection of digital technologies and humanities. It aims to produce applications and models that make possible new kinds of research, both in the humanities disciplines and in computer science and its allied technologies. It also studies the impact of these techniques on cultural heritage, memory institutions, libraries, archives and digital culture.* » (UCL, 2015)

⁴⁴ SALSAH (System for Annotation and Linkage of Sources in Arts and Humanities) est un projet dirigé par l'Université de Bâle qui a pour objectif de développer des outils informatiques de type *Virtual Research Environment* pour les chercheurs en Arts et Humanités. Voir le site Internet du projet : <http://www.salsah.org/>

⁴⁵ Programme européen pour l'éducation, la formation, la jeunesse et le sport.

⁴⁶ *Digital Research Infrastructure for the Arts and Humanities*. Voir le site Internet de DARIAH Europe : <https://www.dariah.eu/>

⁴⁷ Le programme de la journée est disponible à l'adresse : <http://www.unil.ch/ladhul/files/live/sites/ladhul/files/shared/actualites/Domestiquez%20vos%20donn%C3%A9es.pdf>

⁴⁸ Voir le site Internet du projet : <http://rechalpvd.unil.ch/>

⁴⁹ Ces « Lignes directrices sur la propriété, la diffusion et l'utilisation des données » sont disponibles à l'adresse : <http://rechalpvd.unil.ch/static/files/deontologie-rechalp.pdf>

⁵⁰ Voir son site Internet : <http://www.infospecies.ch/fr/>

ceux-ci, même sans l'appui d'une politique de gestion harmonisée à un niveau national, prennent volontiers part au développement de cette thématique. Nous soulignons encore que cette thématique n'est pas exclusivement réservée aux Sciences dites « dures » – à l'UNIL, l'on pense surtout aux sciences biomédicales qui ont une plus longue expérience de la gestion des données de recherche – mais que bon nombre de projets, parfois interdisciplinaires, sont portés par des chercheurs en Sciences humaines et sociales.

5.2.4 EPFL

Voisine de l'Université de Lausanne, l'École polytechnique fédérale de Lausanne (EPFL) n'est pas non plus en reste, en témoigne le projet pilote *Open Research Data at EPFL*⁵¹ – débuté le 1^{er} février 2015 – a pour finalité la création d'un *DMP support service*, un service d'aide à la réalisation du *Data Management Plan* d'une part, mais également de soutien auprès des chercheurs. La direction de ce projet (Comité de pilotage) intègre à la fois les spécialistes ID (via Bibliothèque de l'EPFL), les chercheurs (via l'*Office Research*) et les Systèmes d'information (*Vice Presidency for Information Systems* VPSI). Ces trois pôles de compétences ont donc la mission de définir le cadre institutionnel de gestion des données de recherche, d'accompagner et de former les chercheurs et enfin de développer et mettre en place une infrastructure informatique et technique (voir le détail des outils proposés par l'EPFL en annexe 11) qui soutienne les chercheurs avant la création des données avec les *DMP* ; pendant la recherche avec les *LIMS* et les espaces de stockage ; à la fin du projet avec les dépôts de données (*repositories*).

5.3 Niveau international

5.3.1 Royaume-Uni

À l'échelle internationale, il n'y a guère que le Royaume-Uni qui s'est doté de moyens spectaculaires pour mettre en œuvre une politique nationale de gestion de données de recherche, via le *Joint Information Systems Committee* (JISC)⁵². Ce comité public, en charge de soutenir la recherche et l'enseignement supérieur – en livrant aux institutions concernées des conditions cadres en matière de technologie de l'information et de la communication – est notamment à l'origine de la création du DCC. A eux deux, ils détiennent des compétences et des moyens financiers non négligeables pour élaborer

⁵¹ Voir la conférence organisée par la Bibliothèque de l'EPFL quelques mois plus tôt BIBLIOTHEQUE DE L'EPFL, 2014. *Open Research Data : The Future of Science*, Lausanne, EPFL, Forum Rolex Learning Center, October 28, 2014. *Site web de la Bibliothèque de l'EPFL* [en ligne]. 13.11.2014. [Consulté le 1^{er} mai 2015]. Disponible à l'adresse : <http://library2.epfl.ch/conf/opendata>

⁵² Voir le site Internet du JISC : <https://www.jisc.ac.uk/>

des guides de bonnes pratiques, fournir de l'aide aux chercheurs et institutions ou encore définir des normes et des standards pour la gestion des données de recherche et la *Digital Curation*. Ces deux initiatives nationales, conjuguées à celles de fonds nationaux (publics ou privés) de recherche – notamment le *Research Councils UK (RCUK)* – et à celles des institutions académiques elles-mêmes ont mené les Universités britanniques à élaborer des *Data Management Plans* et à les décliner pour chacune des disciplines de recherche. Le modèle britannique inspire largement les spécialistes du monde entier.

5.3.2 Europe

La Commission européenne a émis en 2013 des *Lignes directrices pour le libre accès aux publications scientifiques et aux données de recherche dans Horizon 2020*⁵³ ainsi que des *Lignes directrices pour la gestion des données dans Horizon 2020*. Ces deux textes s'insèrent dans la vision d'une recherche transparente et reproductible, rendue notamment possible par la création de Plans de gestion des données qui désormais devront être réalisés pour tout projet financé. Suite à la votation fédérale du 9 février 2014, les autorités helvètes sont actuellement en pleine renégociation des accords avec le programme Horizon 2020. En dépit de cette situation, les Universités suisses sont donc *techniquement* soumises à ces *Data Management Plans* pour tous les projets qu'elles portent avec leurs partenaires européens.

5.3.3 Une problématique mondiale

Toujours est-il que l'on observe, à l'échelle mondiale une nette tendance au développement de moyens et de ressources mises à disposition dans le cadre de programmes de gestion des données de recherche, confirmant ainsi que les problématiques engendrées (financement, sécurité, intégrité, entre autres) sont devenues, pour les États et organismes publics de recherche, une priorité – preuve en est par la récente annonce d'un plan d'action national, aux États-Unis, pour un accès public aux résultats de recherche⁵⁴. Les Universités ont également bien compris cet enjeu global puisque bon nombre d'entre-elles⁵⁵ ont créé des services dédiés à la gestion des données de recherche (*Research Data Management Services*), services chargés de fixer le cadre institutionnel de bonnes pratiques et d'accompagner les

⁵³ H2020 est le Programme-cadre de l'Union européenne pour la recherche et l'innovation. Dans ces lignes directrices, les métadonnées associées constituent aussi des données de recherche.

⁵⁴ La déclaration officielle du *National Science Foundation* est disponible à l'adresse : https://www.nsf.gov/news/news_summ.jsp?cntn_id=134478

⁵⁵ Principalement des Universités du monde anglo-saxon où ce sont généralement les bibliothèques (*Libraries*) qui ont la responsabilité de ces *Research Data Management Services*. Rappelons également que ces bibliothèques sont directement rattachées aux Universités, contrairement à la Suisse où les bibliothèques cantonales disposent d'un statut autonome.

chercheurs dans cette voie-là. De plus, certaines d'entre elles se regroupent à un niveau européen – par exemple la Ligue européenne des universités de recherche (LERU) – pour élaborer des recommandations et directives communes pour une gestion transparente et de qualité des données de recherche⁵⁶.

Les éditeurs scientifiques sont également conscients des multiples bouleversements suscités par cette thématique. L'évolution du mode de citation – désormais, l'on cite également les sets de données en plus des résultats de recherche (la visibilité et l'impact du chercheur et de la revue sont donc augmentés) – induit de nouvelles pratiques éditoriales pour améliorer la « *discoverability* » des données de recherche déposées dans les *repositories*, comme l'explique la revue *Science* :

« In 2015, we want to work with authors and readers to identify which of those repositories Science should promote because they are well managed, have long-term support, and are responsive to community needs. [...] We also will evaluate different ways to tag data sets and integrate such tagging into our peer-review process. For example, one might associate a digital identifier for a data set with a figure in a paper. [...] »
(McNutt, 2015)

Ces trois acteurs peuvent encore compter sur le soutien d'organismes internationaux⁵⁷ favorables à l'ouverture, au libre accès et au partage des données de recherche et partisans d'une Science ouverte. Ainsi, de l'Australie – avec son *Australian National Data Service (ANDS)* – en Europe – avec les directives de l'UE, la Grande-Bretagne qui joue le rôle de moteur, ou encore la France⁵⁸ – en passant par les États-Unis – dont les bailleurs de fonds publics imposent les *DMP* et se positionnent clairement en faveur de l'*Open Research Data* –, cette thématique de gestion des données de recherche soulève donc de multiples enjeux au-delà des frontières.

⁵⁶ Voir la feuille de route « *LERU roadmap for Research Data* » à l'adresse :

http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf

⁵⁷ *Center Open Science, Open Data, OpenLearning, Open Science Framework, Open Science Working Group, Open Knowledge, Research Data Alliance, ROpenScience*, parmi de nombreux autres.

⁵⁸ La France développe cette thématique par le biais de son Institut de l'information scientifique et technique (INIST-CNRS) notamment.

SECTION 2 : Enquête et entretiens

6. Enquête « Gestion des données de recherche à l'UNIL »

6.1 Méthodologie et objectifs de l'enquête

Interroger les principales personnes concernées, à savoir les chercheurs de l'UNIL qui gèrent leurs données de recherche au quotidien, est nécessaire pour mieux comprendre un certain nombre d'éléments : quelles données de recherche sont collectées et produites au sein des sept Facultés ; la manière dont elles sont gérées avant, pendant et après un projet de recherche ; les pratiques en matière de stockage, sauvegarde, sécurité, accès, publication, archivage, préservation et réutilisation des données de recherche ; les éventuels problèmes rencontrés ; enfin les attentes et besoins⁵⁹ des chercheurs. L'enquête a donc pour objectif principal de fournir une image précise et réelle de la situation à l'UNIL à un temps t et de nous fournir des résultats pertinents sur lesquels se baser pour, dans un second temps, élaborer et rédiger le Guide des bonnes pratiques⁶⁰ qui tienne compte des réponses des chercheurs.

Pour savoir de manière générale comment et par où débiter l'enquête, nous avons utilisé, dans un premier temps, la méthodologie conseillée par la HEG de Muet (2007) puis le poster *Conducting a Data Interview*⁶¹ développé par Witt & Carlson (2007, Bibliothèque de l'Université de Purdue, USA) décrit comme un « *point de départ [...] non pas une stratégie complète mais un outil pratique destiné à mettre en évidence les informations à considérer pour évaluer l'adéquation d'un set de données avec la collection ainsi que les besoins en infrastructure et services nécessaires pour la data curation* » et qui contient dix questions⁶² centrales à poser. Une fois ce cadre général clarifié, nous avons consulté des enquêtes similaires préexistantes – *Research Data [Management] Surveys* – qui posent toutes à peu près les mêmes questions sur des aspects concrets et pratiques du cycle de vie de la donnée de recherche ainsi que sur la gestion des données de recherche (ex. types de données produites, stockage,

⁵⁹ Deuxième objectif général du travail de Bachelor et ses sous-objectifs (voir partie « objectifs généraux et spécifiques »).

⁶⁰ Troisième objectif général du travail de Bachelor et ses sous-objectifs 3.1 et 3.2 (voir partie « objectifs généraux et spécifiques »).

⁶¹ Ce modèle de *Data Interview* est notamment repris par d'autres (Digital Commons, 2014).

⁶² « #1 *What is the story of the data ?*, #2 *What form and format are the data in ?*, #3 *What is the expected lifespan of the dataset ?*, #4 *How could the data be used, reused, and repurposed ?*, #5 *How large is the dataset, and what is its rate of growth ?*, #6 *Who are the potential audiences for the data ?*, #7 *Who owns the data ?*, #8 *Does the dataset include any sensitive information ?*, #9 *What publications or discoveries have resulted from the data ?*, #10 *How should the data be made accessible ?* » (Witt & Carlson, 2007).

sauvegarde, sécurité, partage, archivage, etc.). Ces enquêtes réalisées par d'autres sont listées ci-dessous par ordre chronologique décroissant :

- *Digital Curation Centre* [RDM Strategy to Action Survey] (Whyte, Sisu, 2014)
- *SIM4RDM* [*Support Infrastructure Models for Research Data Management*] (Mossink and Bijsterbosch, 2013)
- *Humboldt-Universität zu Berlin* (Simukovic, Kindling and Schirmbacher, 2013)
- *University of Nottingham* (Parsons, Grimshaw and Williamson, 2013)
- *University of Leeds* [en ligne] (2012)
- *University of Oxford* (Wilson et al., 2012)
- *ETH-Bibliothek* (Töwe and Scheid, 2011)
- *University of Southampton* (Gibbs, 2009)
- *University of Edinburgh* (Ekmekcioglu and Rice, 2009)
- *UK Research Data Service (UKRDS)* (Beagrie, Beagrie and Rowlands, 2009)

Comptabiliser le total exact des chercheurs actifs à l'UNIL, au sein des sept Facultés, ainsi que la répartition des différentes fonctions associées à la recherche (professeurs, maîtres d'enseignement, assistants doctorants, etc.) s'avère relativement compliqué, notamment de par l'absence de statistiques officielles – indicateurs institutionnels et données centralisées – sur la recherche, p. ex. le nombre de projets de recherche actuellement menés globalement, par Faculté et par institut ; le pourcentage de projets de recherche nationaux ou internationaux ; les données bibliométriques générales sur les publications scientifiques des chercheurs, etc.

Du Centre informatique, nous avons retenu trois listes de distribution – doctorants (~2'100 pers.), enseignants (~2'930 pers.) et chercheurs FNS (~350 pers.) – pour s'assurer de toucher tout le monde. Or il serait inexact d'utiliser comme valeur de référence la somme de ces trois nombres – l'on arriverait à un total de ~5'400 pers. – puisque certaines personnes appartiennent d'une part à plusieurs listes (p. ex. les doctorants FNS) et que ces listes contiennent des personnes qui ne sont pas toutes actives dans la recherche d'autre part. Ce nombre de ~5'400 est donc trop élevé. Par conséquent, c'est par le biais des données issues du Service des ressources humaines qu'une estimation du nombre total de personnes actives ainsi que les fonctions concernées par la recherche a pu être faite, avec l'aide du Service UNISIS. En se basant sur le nombre de contrats⁶³ principaux⁶⁴ actifs au 31 décembre 2014 et en appliquant

⁶³ Nous nous basons sur les contrats plutôt que sur l'équivalent temps plein (ETP) puisque ce sont les individus et leurs pratiques qui nous intéressent et non pas leur charge ou leur capacité de travail.

⁶⁴ Pour éviter de comptabiliser plusieurs fois une seule même personne si elle a plusieurs contrats, seuls les contrats principaux en matière de temps de travail sont pris en compte

une granulométrie fine la plus adéquate – à savoir celle du Bureau de l'égalité de l'UNIL –, l'on obtient un total de **3'150** personnes réparties en quatre grandes catégories⁶⁵ de fonctions :

- *Enseignants habilités à diriger une UO*⁶⁶ (559 pers.) : Professeur assistant (en pré-titularisation conditionnelle ou non) / Prof. associé / Prof. ordinaire ;
- *Autres enseignants* (758 pers.) : Chargé de cours / Maître d'enseignement et de recherche (MER suppléant, de type 1 ou de type 2) / Privat-docent / Prof. invité / Prof. remplaçant / Prof. titulaire ;
- *Assistants et collaborateurs scientifiques* (1408 pers.) : Assistant boursier / Ass. diplômé / Ass. FNS / Ass. Sur fonds externe / Maître assistant / Personnel de recherche / 1^{er} assistant (UNIL ou FNS) ;
- *Personnel de direction, administratif et technique* (425 pers.) : Personnel technique et de laboratoire.

Vu l'importance du nombre de personnes (3'150), le choix s'est rapidement porté vers l'enquête en ligne réalisée avec le logiciel *Limesurvey* via le compte du Service UNIRIS pour conserver le thème (la ligne graphique) des questionnaires UNIL. D'entente avec le mandant, il a été décidé d'élaborer un questionnaire (en annexe 12) relativement complet afin de couvrir l'ensemble des thématiques liées à la gestion des données de recherche. Pour ne pas décourager d'emblée les répondants et assurer un taux de réponse jugé correct, le questionnaire nécessitera au maximum un temps de réponse de 15-20 minutes et sera également disponible en langue anglaise.

Élaboration

Largement inspiré des exemples des questionnaires cités précédemment, notre questionnaire est donc composé de 18 grandes catégories – 1) Projet de recherche ; 2) Données de recherche ; 3) Plan de gestion des données (*Data Management Plan*) ; 4) Description et documentation ; 5) Métadonnées ; 6) Stockage ; 7) Sauvegarde ; 8) Sécurité ; 9) Partage et accès ; 10) Publication ; 11) Archivage et préservation ; 12) Réutilisation ; 13) Juridique, éthique et déontologie ; 14) Aide, soutien et formation ; 15) Information sur le sujet ; 16) Service UNIRIS ; 17) Profil personnel ; 18) Commentaires, suggestions et propositions d'amélioration – et comporte 62 questions au total.

(non pas tous les contrats). Avec cette méthode, une personne qui a 3 contrats par exemple ne sera donc pas comptée 3 fois. Dans le cas où une personne aurait deux contrats à 50%, le principal est défini soit par la fonction la plus élevée soit lors de sa signature.

⁶⁵ Ces grandes catégories correspondent à la granulométrie SIUS (Système d'information universitaire suisse) définie et employée par l'Office fédéral de la statistique (OFS).

⁶⁶ UO : Unité d'organisation.

Diffusion

Par la suite, une invitation à participer à l'enquête a été envoyée par la Direction de l'UNIL par courriel aux trois listes de distribution mentionnées précédemment le 10 avril 2015 et une piqûre de rappel a été faite deux semaines et demi après la date du lancement. Les réponses des participants ont été récoltées durant un peu plus de trois semaines, l'enquête a donc été clôturée le 04 mai 2015.

6.2 Présentation des résultats

Au total, 429 réponses ont été enregistrées⁶⁷ : 310 complètes⁶⁸ (72%) et 199 incomplètes⁶⁹ (28%) et sur ces 429 réponses, 69 (soit 16%⁷⁰) sont en anglais (49 réponses complètes et 20 réponses incomplètes). Seules les 310 réponses complètes – qui représentent un taux de participation de presque 10% – ont été prises en compte.

Traitement des données

Nous avons traité les réponses de manière anonyme. Les pourcentages présentés sont arrondis. Il est important de garder en mémoire que lorsque nous utilisons les termes génériques « participants, répondants ou chercheurs », il s'agit bien des 310 personnes qui nous ont répondu. Nous avons sélectionné quelques questions nous paraissant les plus intéressantes pour notre analyse des besoins des chercheurs. Toutes ne sont pas reformulées mais l'ensemble du questionnaire est reporté en annexe 12 et le détail des réponses et des statistiques en annexe 13.

Résultats Catégorie 1 – Projet de recherche

Remarque préliminaire : Les quatre figures et tableaux suivants nous situent le cadre, le contexte et le mode de financement du projet de recherche mené par les participants à l'enquête.

⁶⁷ Concrètement, cela signifie qu'à chaque fois qu'une personne clique sur le lien URL et ouvre la page web du questionnaire en ligne, elle est comptabilisée par *Limesurvey*, peu importe le fait qu'ensuite elle réponde ou non aux questions posées.

⁶⁸ Les personnes ont donc répondu à toutes les questions posées.

⁶⁹ Ces questionnaires sont inexploitable dans la mesure où ils comportent trop de lacunes, pour environ la moitié d'entre eux la première question n'a même pas été complétée et pour l'autre moitié, ils ont été interrompus relativement rapidement.

⁷⁰ Ce taux de 16% est valable à la fois lorsque l'on calcule le pourcentage, taux de participation générale (réponses complètes et incomplètes) et lorsque l'on calcule le pourcentage des réponses en anglais par rapport aux réponses complètes uniquement.

Q1 : Quel type de recherche menez-vous actuellement ?

Figure 4 : Enquête UNIL / Types de recherche par Faculté (N=525)

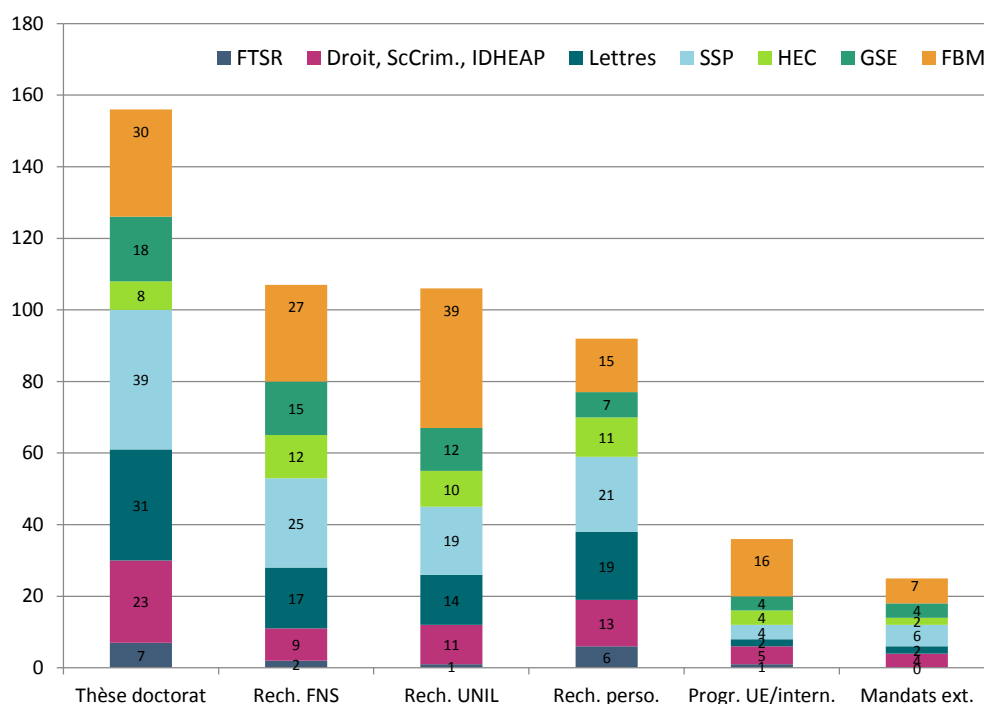


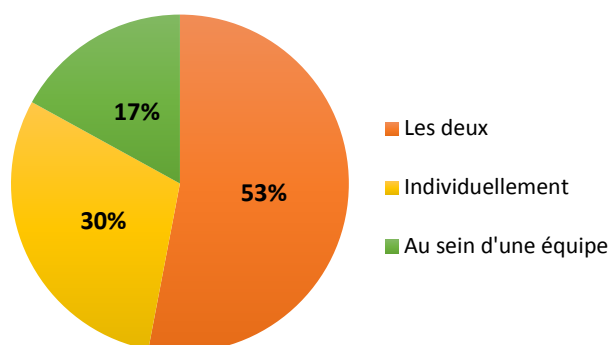
Tableau 1 : Enquête UNIL / Détail des types de recherches par Faculté (N=531)

	FTSR	FDCA	Lettres	SSP	HEC	GSE	FBM	Autres instit.	TOTAL
Thèse doctorat	7	23	31	39	8	18	30	0	156
Rech. FNS	2	9	17	25	12	15	27	1	108
Rech. UNIL	1	11	14	19	10	12	39	1	107
Rech. perso	6	13	19	21	11	7	15	1	93
Mandat ext.	0	4	2	6	2	4	7	1	26
Progr. europ./UE	0	3	1	0	3	1	11	1	20
Progr. intern.	1	2	1	4	1	3	5	1	18
Autre rech.	0	1	0	1	0	0	1	0	3
TOTAL	17	66	85	115	47	60	135	6	531

Commentaire : Nous constatons que : 50% des participants réalise actuellement une thèse de doctorat ; 35% réalise une recherche UNIL ; 35% réalise une recherche FNS ; 30% réalise une recherche personnelle ; 12% fait partie des programmes européens/UE et internationaux ; enfin 8% réalise une recherche sur mandat externe.

Q2 – Menez-vous votre projet de recherche individuellement ou au sein d'une équipe ?

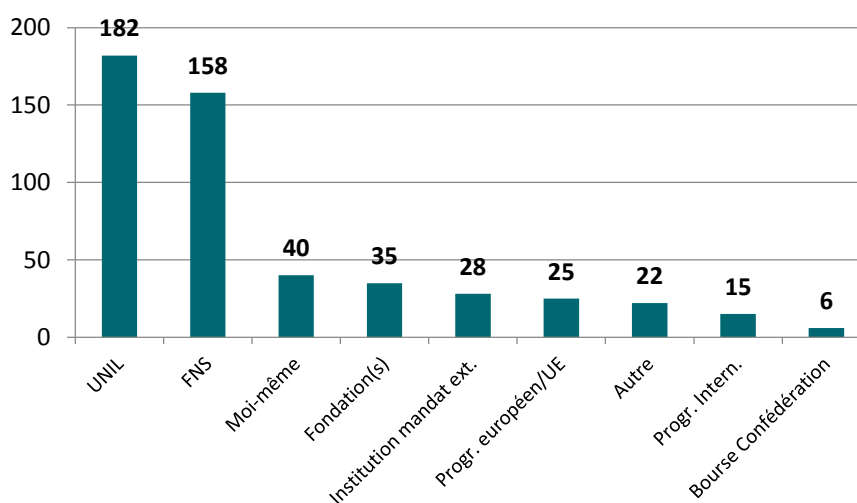
Figure 5 : Enquête UNIL / Recherche individuelle ou en équipe (N=310)



Commentaire : La majorité des chercheurs effectuent leur projet de recherche à la fois individuellement et au sein d'une équipe (53%), un tiers de manière individuelle et enfin 17% uniquement au sein d'une équipe.

Q3 – Quel est le mode de financement de votre recherche ?

Figure 6 : Enquête UNIL / Modes de financement (N=511)



Commentaire : L'UNIL arrive en tête des bailleurs de fonds, elle finance 60% des personnes interrogées. Le FNS se place en deuxième position, 51% des chercheurs touchent des subventions – $\frac{3}{4}$ sont des subsides, le reste est réparti entre les PNR et PRN. A relever que 13% des répondants s'auto-financent et 10% en moyenne touchent des fonds tiers (11% les fondations et 9% les institutions mandantes externes). Enfin, sur le plan des collaborations internationales, 8% des chercheurs bénéficient de financements des programmes européens/UE et 5% des programmes internationaux.

Q4 – Recevez-vous des financements particuliers pour gérer vos DR ?

Quasiment tous les répondants (93%) ne reçoivent aucun financement dédié à la gestion des données de recherche. Les 7% restants sont majoritairement financés par le FNS (77%), par l'UNIL (45%), par les programmes internationaux à égalité avec les fondations (23%) et enfin par les institutions externes (14%).

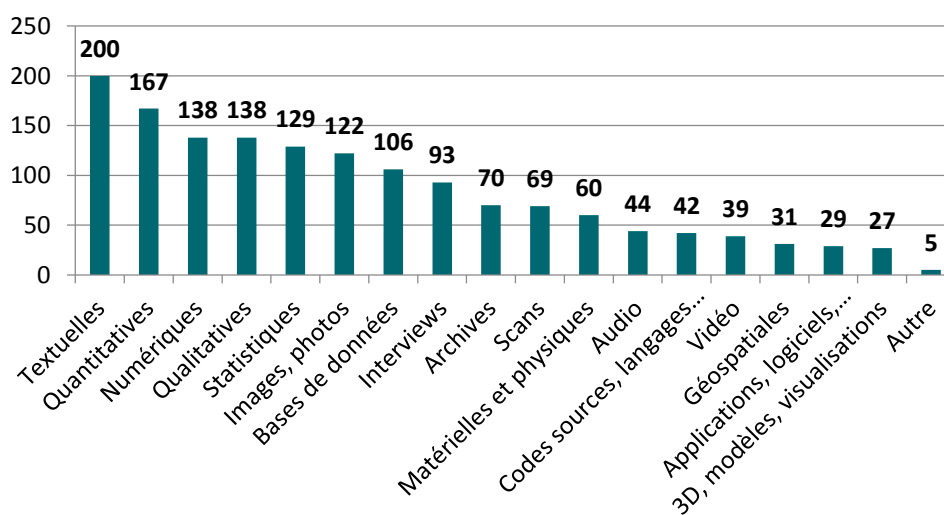
Q5 – Votre bailleur de fonds a-t-il des exigences particulières concernant vos DR ?

51% des chercheurs interrogés répondent que leur bailleur de fonds n'a pas d'exigence particulière concernant les données de recherche. Dans le cas contraire, les exigences les plus souvent mentionnées sont la confidentialité et la sécurité des données (ex. anonymisation, accessibilité restreinte voire interdite, destruction des données après utilisation, etc.) ou encore l'obligation de diffuser et de publier en *Open Access*. Notons un certain manque de clarté de la part des bailleurs de fonds puisqu'environ un tiers des chercheurs répond ne pas savoir si de telles exigences existent ou non.

Résultats Catégorie 2 – Les données de recherche (DR)

Q7 – Quels types de données collectez-vous et produisez-vous ?

Figure 7 : Enquête UNIL / Types de données produites (N=1507)



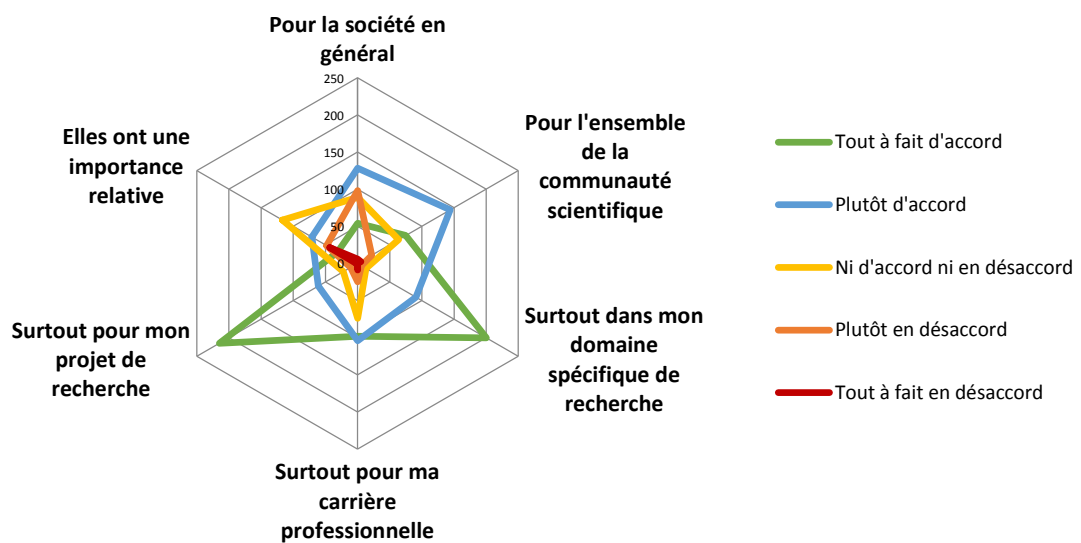
Commentaire : Les principaux types de données collectées et produites par les chercheurs sont : 1) données textuelles (pour 65% des pers.) ; 2) données quantitatives (pour 54% des pers.) ; 3) données numériques, à égalité avec les données qualitatives (pour 45% des pers.) ; 4) données statistiques (pour 42% des pers.) et enfin 5) images et photographies (pour 39% des pers.).

Q8 – D'une manière générale, qui est responsable de gérer les données collectées et produites durant votre recherche ?

La quasi-totalité des répondants (94%) dit gérer soi-même ses données de recherche – la responsabilité relève donc du chercheur – bien qu'aucune aide financière ne soit allouée pour cela (seulement 7% bénéficient d'une aide financière spécifique).

Q9 – Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos DR ?

Figure 8 : Enquête UNIL / Importance des données de recherche (N=311)



Commentaire : La perception qu'ont les chercheurs de l'importance de leurs données de recherche peut être indirectement observée par le degré d'adhésion à chacune des affirmations soumises aux répondants, phrases qui commencent toutes par « *Mes données de recherche sont importantes pour...* ». Différents contextes sont proposés pour compléter ces phrases et, en cumulant les avis favorables (tout à fait d'accord et plutôt d'accord), l'on observe que les chercheurs hiérarchisent l'importance qu'ont leurs données de recherche pour : 1) le domaine spécifique de recherche (oui pour 94% des répondants) ; 2) le projet de recherche (oui pour 89% des répondants) ; 3) l'ensemble de la communauté scientifique (oui pour 71% des répondants) ; 4) la carrière professionnelle (oui pour 65% des répondants) et enfin 5) la société en général (oui pour 59% des répondants).

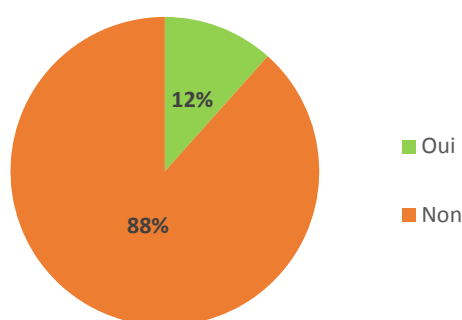
Résultats Catégorie 3 – Plan de gestion des données (*Data Management Plan*)

Q10 – Selon vous, planifier la gestion de ses DR est....

Pour 24% des répondants, la planification est *cruciale* ; *très importante* pour 37% ; *plutôt importante* pour 28% ; *plutôt pas importante* pour 5% ; *pas du tout importante* pour 1% ; et enfin 5% des répondants *ne savent pas*. En additionnant les valeurs *cruciale*, *très importante* et *plutôt importante*, l'on constate donc que planifier la gestion des données de recherche est, pour 90% des personnes interrogées, une tâche importante.

Q11 – Avez-vous déjà réalisé un DMP ?

Figure 9 : Enquête UNIL / Création du Data Management Plan (N=310)



Commentaire : La grande majorité des chercheurs (88%) n'a jamais réalisé de Plan de gestion des données.

Q12 – Si vous avez réalisé un DMP était-ce une exigence de votre bailleur de fonds ?

Lorsqu'un DMP a été réalisé, dans la plupart des cas (64%) il n'était pas une exigence du bailleur de fonds. C'est donc sur propre initiative du chercheur ou de l'équipe de recherche qu'il a été réalisé – sous forme libre – pour un certain nombre de raisons : organiser le projet de recherche et assurer son bon déroulement en général, définir un cadre et formaliser certaines procédures spécifiques à la saisie et au stockage des données, uniformiser les pratiques au sein de l'équipe de recherche, faciliter le partage des données de recherche avec d'autres chercheurs et anticiper leur diffusion ultérieure ou encore pour des raisons éthiques (recherche intègre et responsable). Sur les 36% de répondants pour lesquels le Plan de gestion des données était exigé par le bailleur de fonds (lors de la demande de subvention et/ou après avoir obtenu le financement), 67% touchent des financements du FNS, 47% de l'UNIL et 15% en moyenne des programmes européens/UE ou internationaux, des mandants externes ou des fondations.

Résultats Catégorie 4 – Description et documentation

Q14-Q15 – Création et contenu de la documentation

82% des répondants ont réalisé de la documentation (p. ex. carnets, guides, journaux, rapports) pour décrire leur projet de recherche. Lorsque de la documentation est créée, les principales informations qui la constituent sont : 1) la méthodologie et les processus de recherche (pour 71% des répondants) ; 2) le contexte de la recherche (pour 64% des répondants) ; 3) la récolte des données (pour 63% des répondants) ; 4) les variables utilisées (pour 35% des répondants) ; 5) les instruments et manipulations (pour 31% des répondants).

Résultats Catégorie 5 – Utilisation des métadonnées

Q16-Q17 – Utilisation des métadonnées, normes et standards

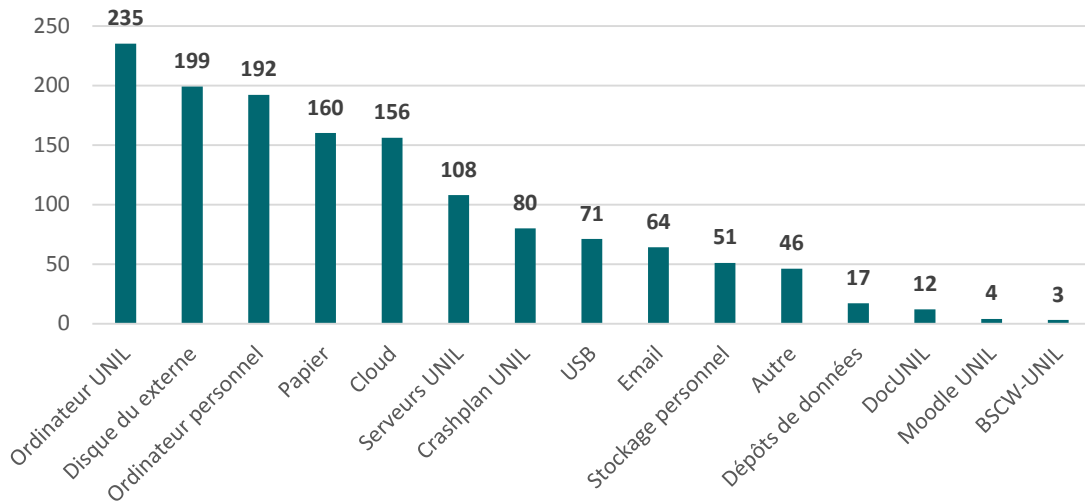
La moitié (50%) des personnes interrogées répond ne pas utiliser de métadonnées, tandis qu'un peu plus d'un tiers (32%) en utilise. Enfin 17% répond ne pas savoir. Lorsque des métadonnées sont utilisées, dans presque la moitié des cas (48%) elles ne sont pas normalisées, standardisées ou structurées. Au contraire, dans 23% des cas, sont principalement utilisés : les standards DDI et TEI, des métadonnées normalisées contenues dans des fichiers générés par des applications et programmes spécifiques (p. ex. fichiers de références bibliographiques, systèmes d'information géographique,...) ou encore des structures et arborescences « faites maison ». Enfin presque un tiers (29%) répond ne pas savoir si les métadonnées utilisées sont standardisées ou non.

Résultats Catégorie 6 – Stockage des DR

Remarque préliminaire : Les pratiques de stockage des données sont similaires à celles pour la sauvegarde des données – ce dernier processus est généralement automatisé. A titre informatif, les sept Facultés utilisent près de 65% du volume de stockage total de l'UNIL et parmi elles, FBM est celle arrive largement en tête du volume de stockage utilisé.

Q18 – Où stockez-vous vos DR ?

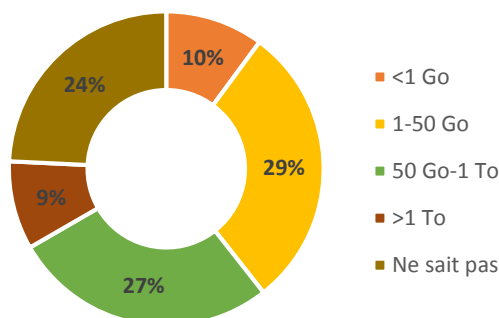
Figure 10 : Enquête UNIL / Stockage des données (N=1397)



Commentaire : Pour 76% des répondants, le stockage des données de recherche est fait sur ordinateur de l'UNIL ; pour 64% : sur disque dur externe ; pour 62% : sur ordinateur personnel ; pour 52% : sur papier ; pour 50% : sur le *Cloud* (majoritairement Dropbox) ; pour 35% : sur les serveurs UNIL ; pour 26% : via Crashplan UNIL ; pour 23% : sur clé USB/Flash ; pour 21% : via l'email ; pour 16% : stockage personnel ; pour 5% : dans des dépôts (*repositories*) de données nationaux ou internationaux (ex. Fors, Genebank) ; pour 4% : via DocUNIL ; enfin pour 7% : via d'autres solutions (par ordre de grandeur : CD/DVD, serveurs CHUV, SIB (Vital-IT) et autres Hautes écoles et Universités, prestataires externes privés, VHS/K7/Disquette).

Q19 – A combien estimez-vous le volume de stockage de vos DR ?

Figure 11 : Enquête UNIL / Estimation du volume de stockage (N=310)



Commentaire : 10% des personnes interrogées estiment le volume de stockage de leurs données de recherche à moins de 1Go ; environ un tiers (29%) entre 1-50Go ; 27% entre 50Go-1To ; 9% 1To et plus ; enfin 24% répond ne pas savoir.

Résultats Catégorie 7 – Sauvegarde des DR

Q20-22 – Existence, fréquence et lieux de sauvegarde

De manière générale, 97% des chercheurs sauvegardent (ou sauvegarderont⁷¹) soit toutes leurs données de recherche (72% des pers.) soit seulement une sélection (23% des pers.). Les 3% restants ne savent pas s'il existe des sauvegardes de leurs données de recherche.

Lorsque les données de recherche sont sauvegardées, dans 35% des cas elles le sont à une fréquence journalière ; 25% à une fréquence hebdomadaire⁷² ; 17% à une fréquence mensuelle⁷³ ; 18% à une fréquence irrégulière⁷⁴ et enfin 5% répond ne pas savoir.

Toujours lorsque les données de recherche sont sauvegardées, le recours à des médias amovibles (disque dur externe, clé USB) demeure le principal moyen pour sauvegarder les données de recherche : 65% des chercheurs utilisent le disque dur externe et 21% des clés USB. Un peu moins d'un tiers des chercheurs en moyenne⁷⁵ délègue la responsabilité au Centre informatique qui assure la sauvegarde automatique des données stockées sur les serveurs UNIL et sur Crashplan UNIL. Enfin environ un tiers également (28%) sauvegarde ses données sur le *Cloud* (toujours majoritairement sur Dropbox).

Résultats Catégorie 8 – Sécurité des DR

Q23-25 – Existence, niveau et moyens pour la sécurité

43% des répondants disent que leurs données de recherche ne sont pas sécurisées et qu'ils n'ont pas l'intention de le faire. A contrario, 39% sécurisent soit toutes leurs données de recherche (14%) soit seulement une partie (25%). Enfin, 9% ont l'intention de le faire et 10% répond ne pas savoir si les données sont sécurisées.

Lorsque les données de recherche sont sécurisées, elles le sont à un niveau standard pour un peu plus de la moitié des chercheurs (51%) ; 21% répondent à un haut niveau ; 15% à un niveau minimal ; 7% à un très haut niveau et enfin 6% à un niveau maximal.

⁷¹ 2% des personnes interrogées ont l'intention de le faire.

⁷² La fréquence hebdomadaire regroupe les réponses « 1x/semaine » et « plusieurs fois par semaine ».

⁷³ La fréquence mensuelle regroupe les réponses « 1x/mois » et « plusieurs fois par mois ».

⁷⁴ La fréquence irrégulière regroupe les réponses « fréquence irrégulière », « 1x/semestre », « 1x/an » et « autre ».

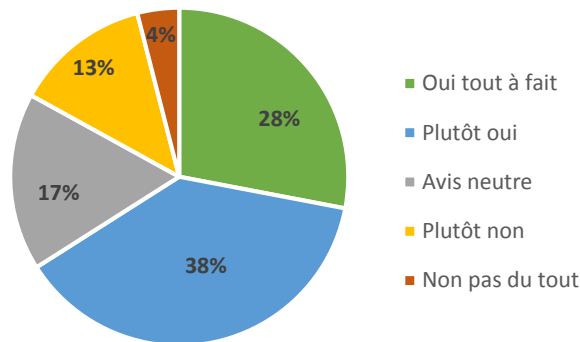
⁷⁵ Moyenne égale à 26.5% (respectivement 28% pour les serveurs UNIL et 25% pour Crashplan UNIL).

Toujours lorsque les données de recherche sont sécurisées, les moyens suivants sont principalement utilisés : 1) login / mots de passes (33% des pers.) ; 2) serveurs sécurisé (22% des pers.) ; 3) pare-feu, antivirus, logiciels anti-hacking, etc. (17% des pers.) et 4) contrôle des accès aux laboratoires, locaux et salles (15% des pers.).

Résultats Catégorie 9 – Partage et accès aux DR

Q26 – De manière générale, êtes-vous favorable au partage de vos DR ?

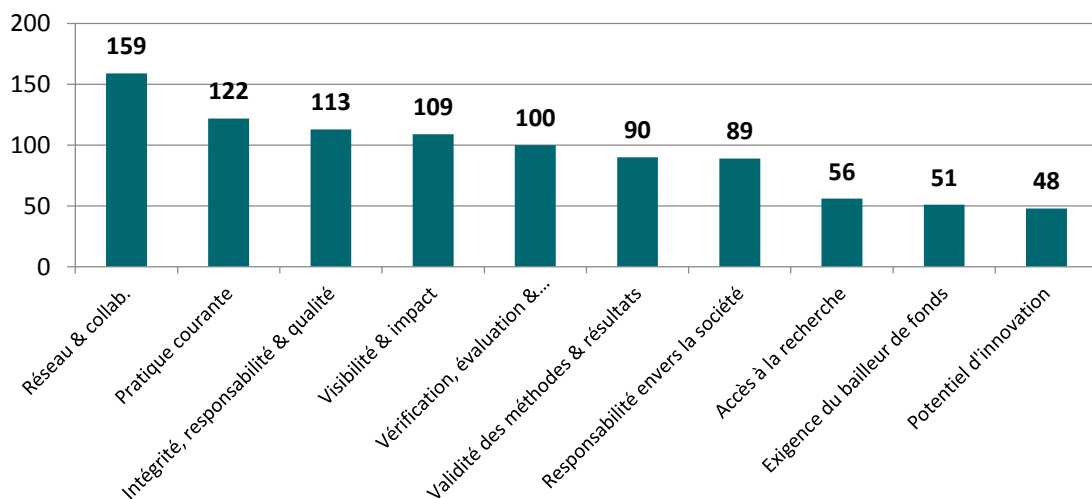
Figure 12 : Enquête UNIL / Pour ou contre le partage des données (N=310)



Commentaire : De manière générale, 66% des personnes interrogées se disent favorables⁷⁶ au partage des données de recherche, contre 18% défavorables⁷⁷. 17% ne se prononce pas (« avis neutre »).

Q27 – Si vous partagez vos DR, quelles raisons motivent votre choix ?

Figure 13 : Enquête UNIL / Arguments pour le partage (N=937)



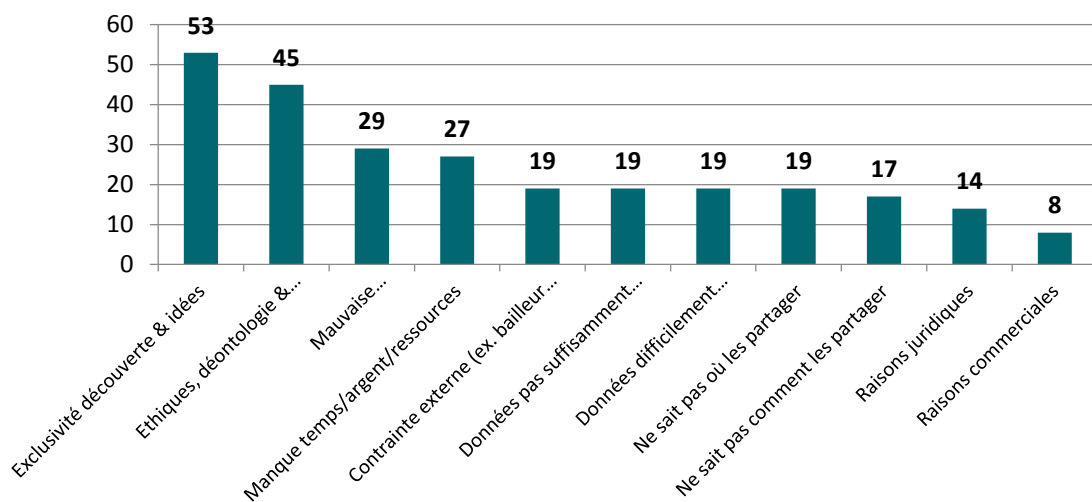
⁷⁶ Cumul des réponses « oui tout à fait » (28%) et « plutôt oui » (38%).

⁷⁷ Cumul des réponses « plutôt non » (13%) et « non pas du tout » (5%).

Commentaire : Lorsque les données de recherche sont partagées, les principaux arguments en faveur du partage sont : 1) réseau, relations et collaborations professionnelles (51% des pers.) ; 2) pratique courante au sein de la communauté scientifique (39% des pers.) ; 3) prouve que la recherche est intègre, responsable et de qualité (36% des pers.) ; augmente la visibilité et l'impact du chercheur / de l'UNIL / de la recherche (35% des pers.) ; 4) facilite la vérification, l'évaluation (*peer review*) et la reproductibilité de la recherche (32% des pers.) et 5) renforce la validité des méthodes et des résultats (29% des pers.) à égalité avec la responsabilité du chercheur envers la société.

Q28 – Si vous ne partagez pas vos DR, quelles raisons motivent votre choix ?

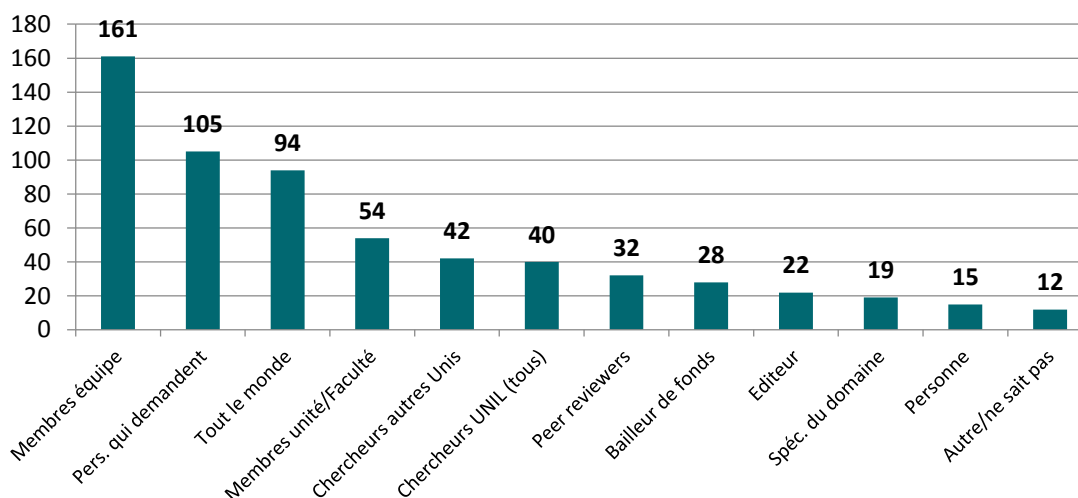
Figure 14 : Enquête UNIL / Arguments contre le partage (N=280)



Commentaire : Lorsque les données de recherche ne sont pas partagées, les principaux arguments en défaveur du partage sont : 1) protection de l'exclusivité de la découverte et des idées (17% des pers.) ; 2) raisons éthiques, de déontologie et de confidentialité (15% des pers.) et 3) risque de fausses interprétations ou d'utilisation fallacieuse des données (9% des pers.) à égalité avec le manque de temps / argent / ressources pour le faire.

Q29 – Après publication, qui peut accéder à vos DR ?

Figure 15 : Enquête UNIL / Partage après publication (N=624)



Commentaire : Après publication des résultats de la recherche, les données de recherche sont partagées avec, de préférence : 1) les membres de l'équipe (52% des pers.) ; 2) uniquement les personnes qui en font la demande (33% des pers.) et 3) tout le monde (30% des pers.).

Résultats Catégorie 10 – Publication des DR

Q30-33 – *Open Access* et publication des DR

De manière générale, 87% des personnes interrogées se disent favorables⁷⁸ au principe de libre accès des publications scientifiques (*Open Access*), contre 6% défavorables⁷⁹. 7% ne se prononce pas (« avis neutre »).

46% des chercheurs ont déjà publié dans des revues⁸⁰ en *Open Access* (OA), selon divers modèles⁸¹. Plus d'un tiers (36%) répond avoir l'intention de publier en OA à l'avenir – pour l'instant la plupart n'ont pas encore publié. Les personnes qui ne souhaitent pas publier en OA représentent 8%. Enfin, 11% des chercheurs ne connaissaient pas cette possibilité.

⁷⁸ Cumul des réponses « oui tout à fait » (53%) et « plutôt oui » (34%).

⁷⁹ Cumul des réponses « plutôt non » (5%) et « non pas du tout » (1%).

⁸⁰ P. ex. *Europe's Journal of Psychology*, *Frontiers in Microbiology*, *Journal of Chemical Ecology*, *Judgment and Decision Making*, *PLOS ONE* ou encore *Springer ebook*.

⁸¹ 15% en « golden road », 9% en « green road », 7% en modèle « hybride » et 15% ne savent pas selon quel modèle.

Un tiers (31%) des répondants dit ne pas connaître la possibilité de fournir les données de recherche en annexe (*supplementary material*) à la publication. 43% sont enclines⁸² à cette démarche, contre 25% qui disent ne pas avoir l'intention de le faire – sauf si les revues scientifiques l'exigent – principalement pour des raisons éthiques, de confidentialité des données, ou encore par obligation juridique et contractuelle.

Dans environ trois quarts des cas (74%), la publication des données de recherche, en annexe, n'était pas une exigence de l'éditeur scientifique. Dans le cas contraire (26%), certaines revues sont mentionnées par les chercheurs : *Bioinformatics*, *eLife*, *Molecular Systems Biology*, *Nucleic Acids Research* ou encore *Plant physiology*.

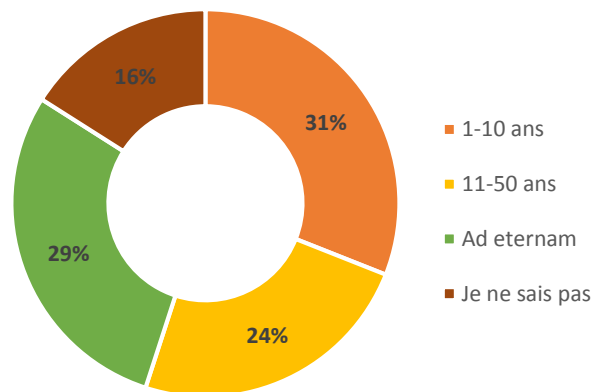
Résultats Catégorie 11 – Archivage et préservation des DR

Q34 – Conservez-vous vos DR à la fin d'un projet ?

Seul 1% des personnes interrogées dit ne pas vouloir conserver ses données de recherche à la fin d'un projet, les 99% autres conservent (ou conserveront⁸³) soit toutes leurs données (76% des pers.) soit une partie⁸⁴ d'entre elles (21% des pers.).

Q35 – Combien de temps pensez-vous que vos DR devraient être conservées ?

Figure 16 : Enquête UNIL / Estimation de la durée de conservation (N=310)



Commentaire : Un tiers environ (31%) des chercheurs pense que la durée de conservation de leurs données de recherche devrait être de ≤ 10 ans⁸⁵, presque à égalité avec le tiers environ (29%) qui considère au contraire que ses données devraient être

⁸² Cumul des personnes qui ont déjà publié leurs données dans des revues en OA (15%) – p. ex. *American Economic Review*, *Clio*, *eLife*, *Environmental Science & Technology*, *Journal of Proteome Research* ou encore *ProteomeXchange* – et des personnes qui ne l'ont pas encore fait mais qui ont l'intention de le faire (25%).

⁸³ 3% répond avoir l'intention de le faire.

⁸⁴ Cumul des réponses « oui env. 75% [des données] » (14%), « oui env. 50% [des données] » (5%) et « oui env. 25% [des données] » (2%).

⁸⁵ Cumul des réponses « 1-4 ans » (5%) et « 5-10 ans » (26%).

conservées « ad eternam ». 24% répondent entre 11-50 ans⁸⁶ ; enfin 16% dit ne pas savoir.

Q36 – De manière générale, êtes-vous favorable au dépôt de vos DR dans un dépôt / répertoire (*repository*) de données ou centre d'archives ?

69% des personnes interrogées se disent favorables⁸⁷ au principe de déposer ses données de recherche dans un dépôt (*repository*), contre 6% défavorables⁸⁸. Enfin, 22% ne se prononcent pas (« avis neutre »).

Q37 – Avez-vous déjà déposé vos DR dans un dépôt (*repository*) ?

De manière générale, les répondants sont plutôt favorables au dépôt des données : 15% des répondants disent avoir déjà déposé leurs données⁸⁹ et 19% ont l'intention de le faire dans le futur. Au contraire, 21% n'ont pas l'intention de le faire. Enfin, 45% des personnes répondent qu'elles ne connaissaient pas cette possibilité.

Q38 – Si vous déposez (ou avez l'intention de déposer) vos DR dans un dépôt (*repository*), à quel endroit souhaiteriez-vous les déposer ?

Les lieux privilégiés pour un éventuel dépôt de données sont majoritairement des solutions institutionnelles UNIL, viennent ensuite les solutions disciplinaires nationales ou internationales. Les principaux lieux plébiscités pour le dépôt des données sont donc : 1) dépôt institutionnel UNIL (25%) ; 2) dépôt de l'unité de recherche (24%) ; 3) dépôt de la Faculté (15%) et 4) dépôt disciplinaire national ou international⁹⁰ (17%).

Q39 – Si vous déposez (ou avez l'intention de déposer) vos DR dans un dépôt (*repository*), quelles raisons motivent votre choix ?

Les principales raisons qui motivent les chercheurs à déposer leurs données dans des dépôts sont les suivantes : 1) préservation à long terme et archivage (35%) ; 2) sécurité (26%) ; 3) gratuité (25%) ; 4) capacité de volume de stockage (21%) et 5) gestion des accès et des droits (20%).

⁸⁶ Cumul des réponses « 11-20 ans » (14%) et « 21-50 ans » (10%).

⁸⁷ Cumul des réponses « oui tout à fait » (53%) et « plutôt oui » (34%).

⁸⁸ Cumul des réponses « plutôt non » (5%) et « non pas du tout » (1%).

⁸⁹ Pour la grande majorité seulement une partie d'entre elles, dans des dépôts de l'unité de recherche ou des dépôts disciplinaires comme celui de FORS par exemple.

⁹⁰ Cumul des réponses « disciplinaire national » (10%) et « disciplinaire international » (7%)

Résultats Catégorie 12 – Réutilisation des DR

Q40 De manière générale, acceptez-vous que d'autres chercheurs réutilisent vos données de recherche ?

70% des personnes interrogées se disent favorables⁹¹ et acceptent que d'autres chercheurs réutilisent leurs données, tandis que 15% se disent opposées⁹², à égalité avec celles qui ne se prononcent par (« avis neutre »).

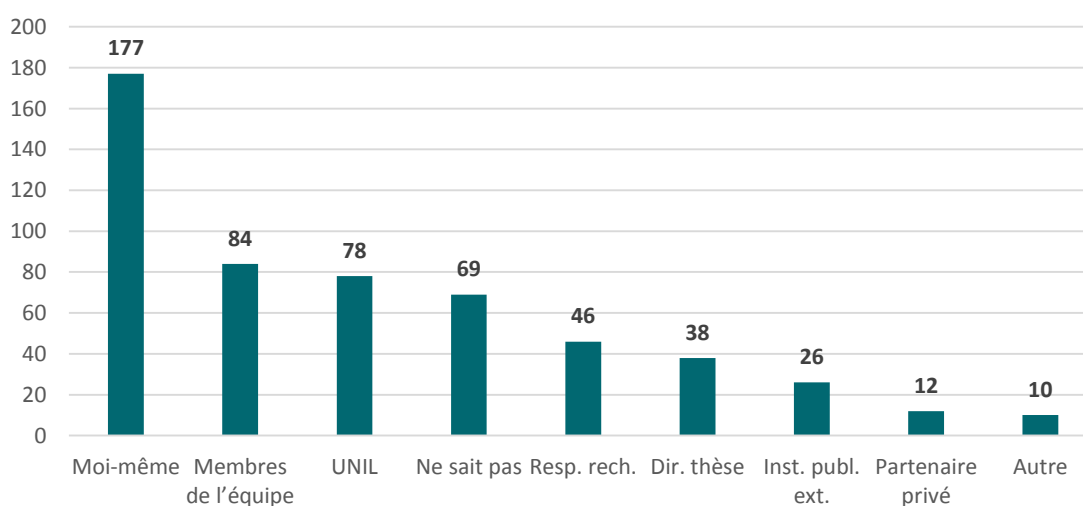
Q41-42 Avez-vous déjà réutilisé les données de recherche d'autres chercheurs dans le cadre de votre propre recherche ? Si oui, pour quelles raisons ?

Environ deux tiers (59%) des répondants ont déjà réutilisé⁹³ les données de recherche d'autres chercheurs – ou réutiliseront (ont l'intention de le faire) pour 17%. Environ un quart (24%) répond ni l'avoir fait ni avoir l'intention de le faire. Les trois principales raisons qui motivent les chercheurs à réutiliser les données d'autres chercheurs sont : 1) pour les citer, à égalité avec comparer les données (39%) ; 2) pour développer une nouvelle recherche (31%) ; 3) pour éviter de refaire l'expérience à double (14%).

Résultats Catégorie 13 – Questions juridiques, éthiques et déontologiques

Q43 A qui appartiennent les données que vous créez et collectez lors de votre activité de recherche ?

Figure 17 : Enquête UNIL / Propriété des DR (N=540)



⁹¹ Cumul des réponses « oui tout à fait » (30%) et « plutôt oui » (40%).

⁹² Cumul des réponses « plutôt non » (11%) et « non pas du tout » (4%).

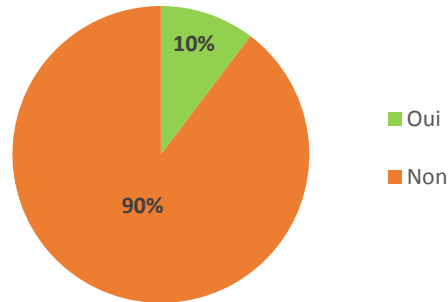
⁹³ Principalement « quelques fois ».

Commentaire : Selon les personnes interrogées, les données de recherche appartiennent d'abord à : 1) Le chercheur (57%) ; 2) Tous les membres de l'équipe de recherche (27%) ; 3) l'UNIL (25%). 22% disent ne pas savoir.

Résultats Catégorie 14 – Aide, soutien et formations

Q47 Avez-vous déjà suivi des formations sur la GDR ?

Figure 18 : Enquête UNIL / Participation à des formations (N=310)



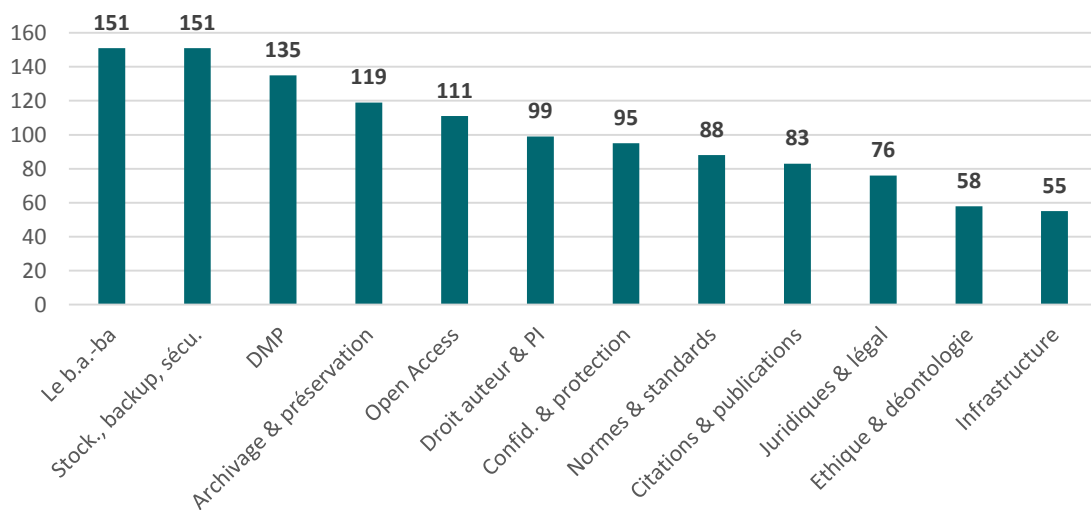
Commentaire : 90% des répondants n'ont jamais suivi de formation particulière en la matière.

Q48 Parmi les différents types de formation, lesquels préférez-vous ?

Les types de formations principalement plébiscités sont : 1) Ateliers pratiques et workshops pour plus de la moitié des personnes interrogées (55%) ; 2) Documentation et ressources en ligne (46%) ; 3) Cours, séminaires intégrés au cursus doctoral ; 4) Site web dédié (33%) ; 5) Aide et conseils personnalisés (31%).

Q49 Si vous deviez participer à une formation, quels sujets vous intéresseraient le plus ?

Figure 19 : Enquête UNIL / Besoins en formations (N=1221)

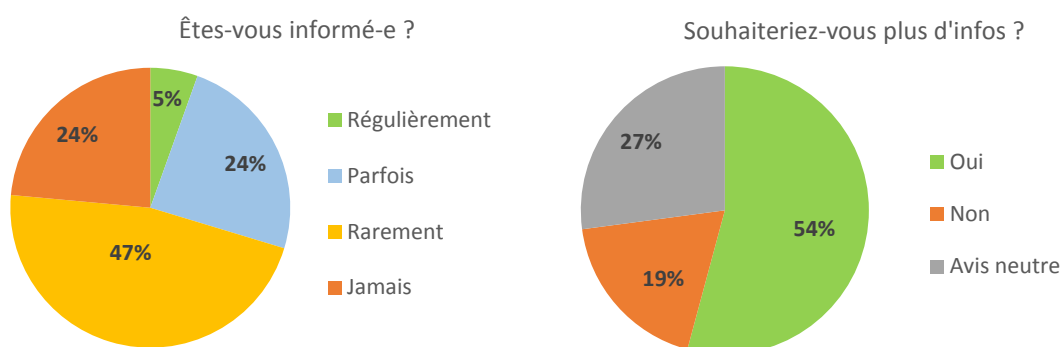


Commentaire : Presque la moitié (49%) des répondants souhaite que ce soient les fondamentaux (b.a.-ba, la base) qui leur soient transmis, à égalité avec les problématiques concrètes de stockage, sauvegarde, sécurité. 44% souhaite recevoir des formations sur les DMP en particulier, enfin les thématiques d'archivage et de préservation à long terme (38%) ou d'*Open Access / Open Data* (36%) sont également privilégiées.

Résultats Catégorie 15 – Informations sur la GDR

Q50-51 Informations sur la GDR

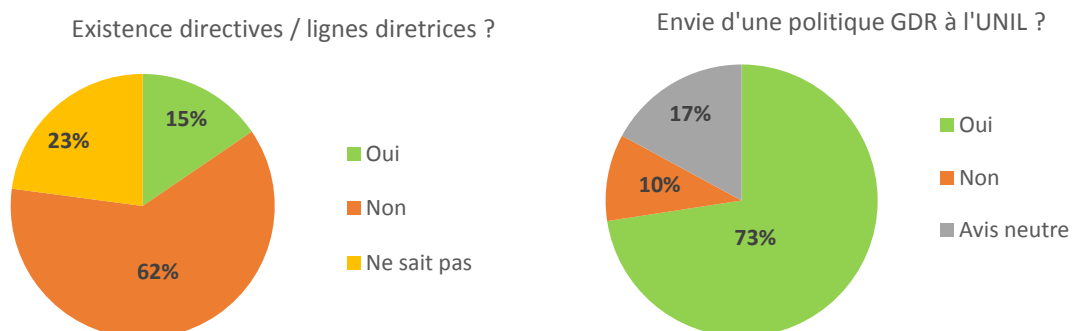
Figure 20 : Enquête UNIL / Besoins en information (N=310)



Commentaire : Près de trois quarts (71%) des personnes interrogées disent être informées rarement (47%) ou jamais (24%). A contrario, un tiers environ (29%) est informé de temps en temps (24%) ou régulièrement (5%). Plus de la moitié (54%) des chercheurs exprime son besoin de d'avantage d'informations sur la gestion des données de recherche. Un tiers environ (27%) ne se prononce pas (« avis neutre »), enfin 19% estime ne souhaite pas recevoir plus d'information.

Q52-53 Directives, lignes directrices et politique de GDR à l'UNIL

Figure 21 : Enquête UNIL / Politique spécifique institutionnelle UNIL (N=310)



Commentaire : Environ deux tiers (62%) des chercheurs répondent qu'il n'existe pas de directives spécifiques pour la gestion des données de recherche au sein de la l'unité de recherche ou de la Faculté. 23% ne sait pas, enfin 15% mentionne l'existence de directives (ex. stockage, accès, diffusion OA, destruction, directives LIVES).

73%⁹⁴ expriment le souhait que l'UNIL développe une politique de gestion des données de recherche. 17% ne se prononcent pas (« avis neutre ») enfin 10%⁹⁵ sont plutôt opposés.

Résultats Catégorie 16 – Service UNIRIS

Q54-55 Connaissance du Service UNIRIS

Un peu moins de trois quarts (74%) des personnes interrogées dit ne pas du tout connaître le Service UNIRIS. 23% ne le connaît uniquement de nom. Enfin, parmi les 4% qui le connaissent, 1% a déjà sollicité le Service UNIRIS (pour des numérisations et recherches de fonds d'archives).

Résultats Catégorie 17 – Profil personnel

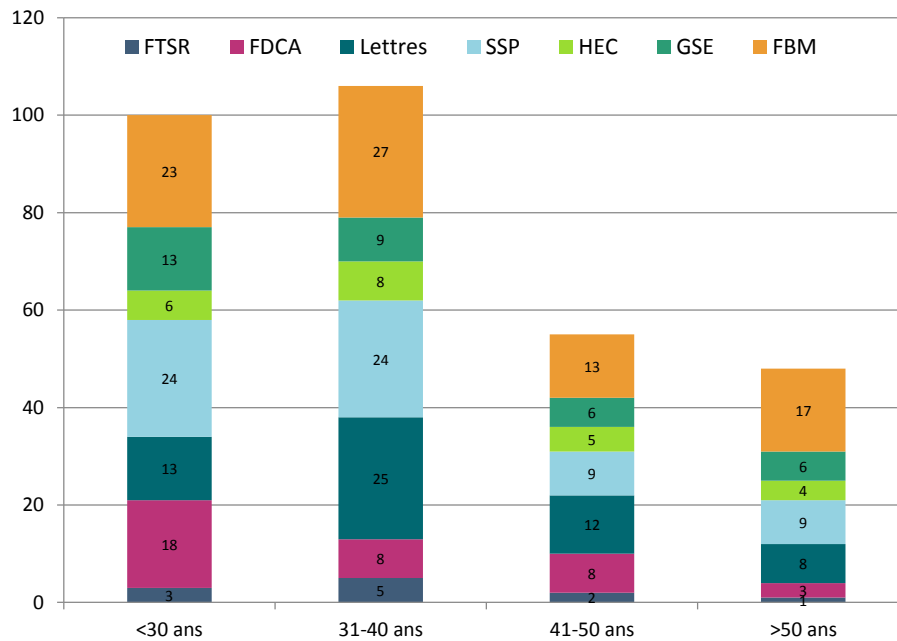
Remarque préliminaire : Les trois questions suivantes démontrent toutes sous un angle différent que ce sont plutôt des jeunes chercheurs relativement novices (doctorants, assistants) qui ont répondu à notre questionnaire.

⁹⁴ Cumul des réponses « oui tout à fait » et « plutôt oui ».

⁹⁵ Cumul des réponses « plutôt non » et « non pas du tout ».

Q57 Quel âge avez-vous ?

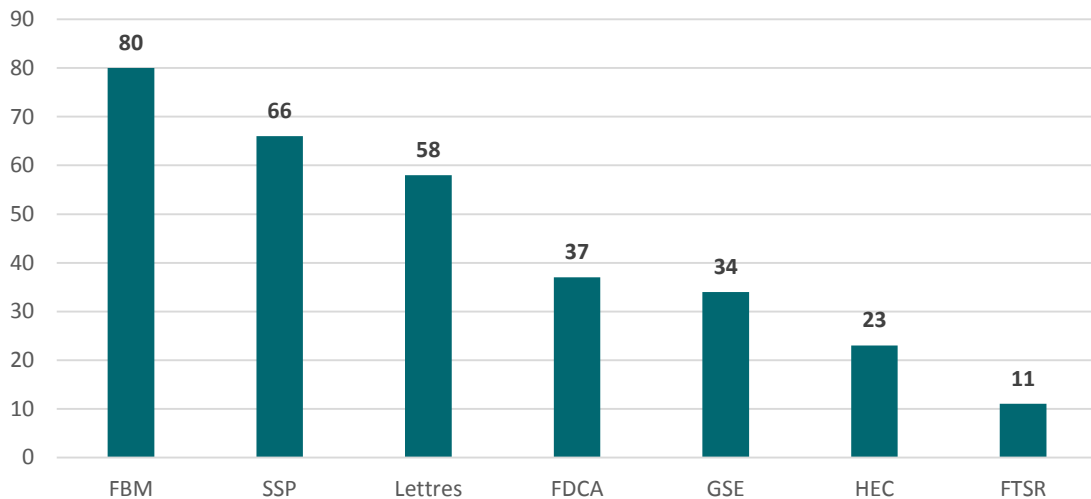
Figure 22 : Enquête UNIL / Tranches d'âge par Faculté (N=310)



Commentaire : 66%⁹⁶ des participants à notre enquête ont moins de 40 ans.

Q58 A quelle Faculté êtes-vous rattaché-e ?

Figure 23 : Enquête UNIL / Répartition au sein des Facultés (N=310)

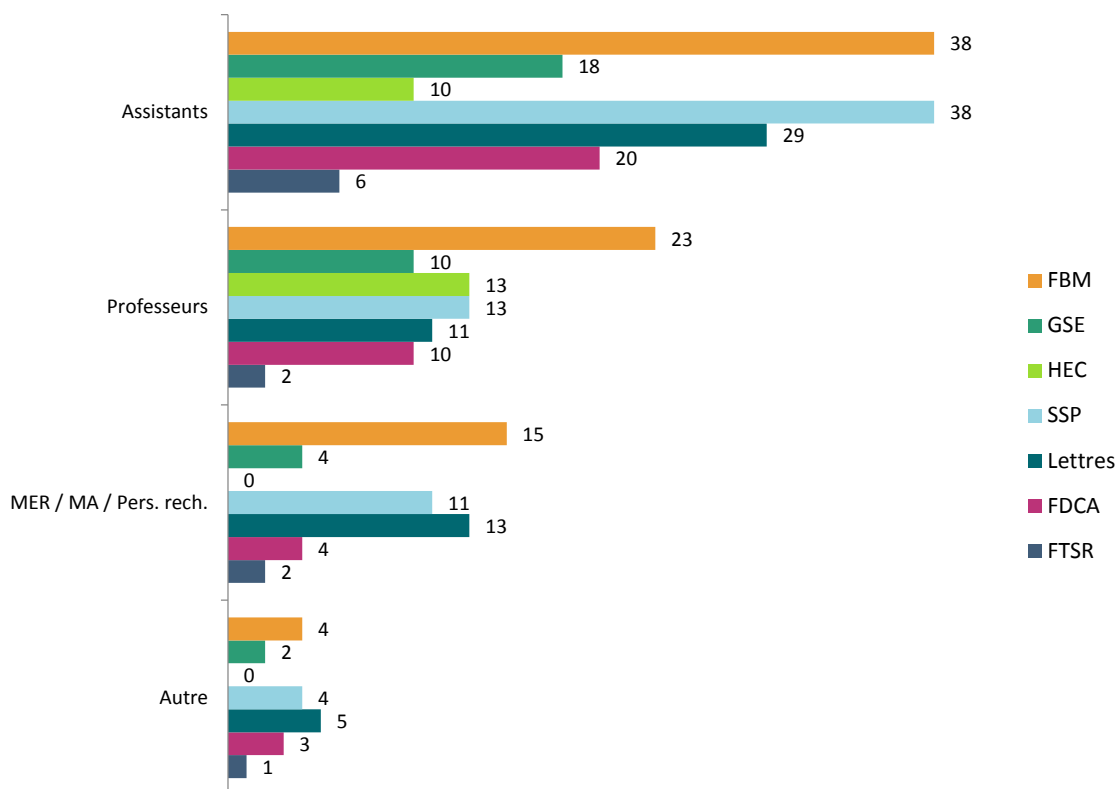


Commentaire : Environ un quart (26%) des répondants sont rattachés à la FBM ; 21% à SSP ; 19% Lettres ; 12% FDCA ; 11% GSE ; 7% HEC ; 4% FTSR.

⁹⁶ Cumul des réponses « - de 30 ans » et « 31-40 ans ».

Q60 Quelle est votre fonction (ou la fonction la plus élevée si vous en avez plusieurs) ?

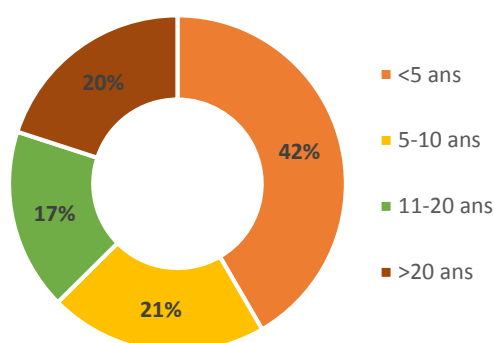
Figure 24 : Enquête UNIL / Fonctions par Faculté (N=309)



Commentaire : Les assistants⁹⁷ représentent 51% des personnes interrogées, l'ensemble des professeurs 26% et le corps intermédiaire (MER, MA et personnel de recherche) 13%.

Q61 Combien d'années d'expérience dans la recherche avez-vous ?

Figure 25 : Enquête UNIL / Années d'expérience en recherche (N=310)



⁹⁷ Assistants boursiers, diplômés, FNS, sur fonds externe et premiers assistants.

Commentaire : 42% ont moins de cinq ans d'expérience dans la recherche. De manière générale donc, 63%⁹⁸ des personnes interrogées ont 10 – ou moins – années d'expérience dans la recherche. ; 42% ont moins de cinq ans d'expérience dans la recherche.

Résultats Catégorie 18 – Commentaires libres

Q62 Ici vous pouvez nous faire part de vos attentes et besoins spécifiques en matière de gestion des données de recherche, de vos commentaires, suggestions et propositions d'amélioration.

13% de répondants ont pris le temps de faire des remarques et des suggestions dans l'espace à disposition dans l'enquête. L'analyse des réponses n'a rien de scientifique mais donne un aperçu plus personnalisé, parce que formulé, des interventions.

Dans les grandes lignes, nous pouvons dire que les attentes sont diverses, parfois contradictoires, et relever le regard critique des répondants sur le statut actuel qui ne les laisse pas indifférents. Certains souhaitent des solutions de gestion à un niveau « intra-universitaire », l'optimisation des systèmes déjà en vigueur, la mise à disposition d'un service transversal, l'usage d'un DMP ou des solutions spécifiques au type de recherche. Ceux qui semblent s'accommoder du statu quo suggèrent d'améliorer les outils déjà existants. Au contraire, d'autres manifestent une certaine forme d'attachement à la liberté de gérer, à la responsabilité individuelle et aux solutions de gestion de proximité. Les enjeux de l'interdisciplinarité, de l'accès, du dépôt, de la visibilité et de la reproductivité sont évoqués, y compris pour des partenaires hors de l'UNIL. Un souci de bien faire en terme d'éthique, de déontologie et même d'écologie est aussi soulevé. Certains suggèrent de développer l'aide sous forme de cours même très basiques, voire de les intégrer au cursus d'étude, de créer des d'espaces ressources et de mieux développer l'information.

⁹⁸ Cumul des réponses « <5 ans » et « 5-10 ans ».

7. Entretiens individuels

7.1 Méthodologie et objectifs des entretiens

Pour mettre en œuvre nos entretiens individuels, les techniques existantes et particulièrement la méthodologie de Muet (2007) a été suivie pour dans un premier temps élaborer notre guide d'entretien (en annexe 14), puis pour mener les entretiens. Nous avons fait le choix de conduire des entretiens semi-directifs afin de recueillir des données qualitatives qui compléteront les données quantitatives issues de notre enquête en ligne. Les entretiens devraient nous permettre de mieux comprendre à la fois le contexte interne à l'UNIL – monde académique et autres Services – et également le contexte externe – pratiques d'autres institutions partenaires.

Après avoir identifié les personnes clés à interroger, nous avons pu rencontrer et interviewé les personnes suivantes sur le campus de l'UNIL :

- Personnes impliquées au sein du Dicastère « *Recherche & Relations internationale* »
- Directions et Décanats de toutes les Facultés
- Collaborateurs d'autres Services de l'UNIL :
 - Centre informatique
 - Service des statistiques UNISIS
 - Bureau des relations UNIL-FNS
 - Bureau des transferts technologiques PACTT
- Chercheurs en lien avec les données de recherche (LaDHUL et RechAlp)
- Associations de doctorants (ACIDUL⁹⁹ et PhDnet¹⁰⁰)

Nous avons également rencontré des personnes externes à l'UNIL, dans des institutions partenaires :

- FORS – DARIS et les personnes impliquées dans le projet CUS P-2 ORD@CH
- ETHZ : présentation et visite du Service
- EPFL et les personnes impliquées dans le projet CUS P-2 DLCM
- SIB et les personnes impliquées dans le projet CUS P-2 DLCM

Tout comme pour notre enquête, nous présenterons ci-après les résultats des entretiens de manière à garantir l'anonymat des personnes interviewées, c'est-à-dire sans information de type : nom, fonction ou titre de la personne.

⁹⁹ Association du corps intermédiaire et des doctorants de l'UNIL.

¹⁰⁰ Association des étudiants doctorants de HEC.

7.2 Présentation des entretiens

7.2.1 Tableau croisé et synthèse des entretiens menés dans les Facultés

Faculté	FTSR & Lettres	Droit	Sc. Crim.	IDHEAP	SSP	HEC	GSE	FBM
Directives, règlements GDR ?	Non	Non	Sécurité oui, d'autres sont en élaboration	Non	Non sauf pour les gros projets de recherche (ex. LIVES)	Non	Non	Éthique oui Sécurité DSI-CHUV oui
Types de DR	Publications Majorité des sources primaires non créées par le chercheur Utilisation des BDD.	Publications Majorité des sources primaires non créées par le chercheur	Hautement sensibles	Hétérogènes Études longitudinales et analyses comparatives	Hétérogènes Qualitatives Quantitatives Sensibles Nouvelles DR LADHUL	Hétérogènes	Hétérogènes	Hautement sensibles Fondamentales et cliniques
Problèmes, attentes & besoins spécifiques GDR	<u>BDD</u> : - Vol. stockage sur serveurs UNIL - Accès sécurisé - Mise à disposition pour la communauté	Non signalés	Besoin infrastructure informatique hautement sécurisée Confidentialité vs. accès à l'information	Accès, partage et diffusion Besoin infrastructure simple d'utilisation	Vol. stockage Sécurité Confidentialité anonymisation OA : favoriser le dépôt dans SERVAL	Connaissances limitées des infrastructures et solutions existantes Besoin de partage facile, rapide synchronisation	OA : favoriser le dépôt dans SERVAL	Vol. stockage Sécurité Confidentialité Puissance de calcul de masse Collaboration Partage

Favorable à un projet institutionnel GRD-UNIL	Très favorable Nécessité d'impliquer tout le monde y.c. les Facs.	Plutôt favorable Difficulté à adopter des pratiques communes	Très favorable Nécessité de collaborer avec les spéc. ID et le Ci	Favorable Nécessité d'impliquer tout le monde y.c. les Facs	Favorable Nécessité d'impliquer tout le monde y.c. les Facs	Favorable Nécessité d'impliquer tout le monde (Facs & fonctions)	Favorable	Favorable « Longue » expérience de la GDR
Favorable à des formations en GDR	Très favorable Séminaires formations intégrées au cursus	Plutôt favorable	Très favorable Cible : étudiants et doctorants	Plutôt favorable Difficulté à adopter des pratiques communes	Favorable Incitation à participer aux formations existantes (ex. FORs)	Plutôt favorable	Favorable Cible : étudiants et doctorants	Favorable Cible : étudiants et doctorants
Enjeux & perspectives futures	Création, accès, gestion des droits, pérennisation Interdisciplinarité Valorisation & vulgarisation Éthique et plagiat GDR = efficacité et gain temps	Peu d'enjeux directement liés à la GDR Collaborations et projets internationaux	Sécurité Stockage Protection des données Accès, partage et diffusion Conservation, élimination Mais aussi interdisciplinarité	Accès, partage et diffusion	Humanités digitales (LADHUL) Big Data Interdisciplinarité Valorisation, visibilité & vulgarisation	Collaborations et projets internationaux OA vs. avantage compétitif et facteur d'impact	Valorisation & vulgarisation Interdisciplinarité	Big Data Banques de données Sécurité et protection des données Harmonisation des pratiques au niveau suisse Interdisciplinarité Potentiel d'innovation et de découvertes

Divers	Projet de création d'une BDD UNIL pour les Lettres	Rech. individuelles Peu de publications dans SERVAL Réticence des éditeurs face à l'OA	La forte augmentation d'étudiants modifie les pratiques des chercheurs		LADHUL actif dans divers projets GDR Incitation à déposer les données chez FORS Réutilisation fréquente des DR pour des études secondaires	Utilisation très fréquente du <i>Cloud</i> Réticence des éditeurs face à l'OA	Projet RechAlp (DR sur les Alpes vaudoises)	Formations sur l'éthique et la déontologie existantes Partenariat avec <i>SIB</i>
---------------	--	--	--	--	--	--	---	--

7.2.2 Analyse et synthèses des autres entretiens

UNIL, FORS et LIVES

Riches d'enseignement, ces entretiens nous ont permis de mieux comprendre qui sont et que font les personnes dans les Services et Facultés et dans quelle mesure elles sont concernées par la gestion des données de recherche. Ces rencontres ont mis en évidence une possible collaboration future et intégration relativement simple de ces personnes – puisqu'elles sont internes à l'UNIL – dans la perspective de la création d'un groupe de soutien. Institution partenaire, FORS, a une expérience de la conduite de projets de gestion des données de recherche, des infrastructures qui existent déjà et des retours de fonctionnement plutôt positifs. Nous avons eu accès à certains documents internes confidentiels, notamment des supports de formations pour la gestion des données de recherche en Sciences sociales, et des démonstrations pratiques sur les outils utilisés. C'est en partie grâce aux contacts de FORS que nous avons pu rencontrer le *Data Manager* du PRN LIVES qui nous a également fourni de nombreuses informations et qui, lui aussi, nous a montré l'infrastructure informatique utilisée. Nous pourrions donc nous appuyer sur toutes ces compétences, ces modèles et infrastructures existantes dans la perspective de la création d'un Centre de soutien à la gestion des données de recherche à l'UNIL (voir point 10.2.3 plus loin).

ETHZ et EPFL

Nous avons eu la chance de pouvoir rencontrer des membres du *Digital Curation Office* de l'ETHZ qui fournit, depuis 2014, la possibilité aux chercheurs de déposer leurs données de recherche dans l'archive institutionnelle *ETH Data Archive*¹⁰¹. En plus de l'accès et du signalement des données – par l'utilisation des *Digital object identifiers* (DOIs) développés par le consortium *DataCite*¹⁰² –, la préservation à long terme est garantie pour une durée minimale de dix ans. Lors du dépôt de données, les chercheurs sont guidés (via des procédures et lignes directrices), notamment pour le choix des formats, ajouts de métadonnées, permissions accordées aux utilisateurs ou encore la définition des droits d'accès (libres, restreints, après période d'embargo). Actuellement une centaine de données de recherche et sets sont librement accessibles sur le catalogue en ligne de la bibliothèque. Dans les échanges, nous avons pu prendre

¹⁰¹ Les technologies utilisées sont *Ex Libris Rosetta* et *DocuTeam Packer*.

¹⁰² Consortium international créé en 2009 et qui réunit bibliothèques et autres acteurs du monde des Sciences de l'information. Il est spécialisé dans les questions d'archivage numérique et particulièrement du signalement et de l'accessibilité des ressources en ligne (articles et publications scientifiques essentiellement). Il est à l'origine du DOI, qui ressemble aux URIs (*Uniform Resource Identifiers*) mais qui sont bien plus stables et pérennes à long terme. Voir le site Internet de DataCite : <https://www.datacite.org/>

conscience des obstacles parfois rencontrés lors du déroulement d'un projet ainsi que du temps de mise en œuvre de telles structures et processus (voir le workflow en annexe 11). Nous avons également eu accès aux résultats d'une enquête réalisée par la Bibliothèque de l'ETH dont nous avons pu nous inspirer pour élaborer la nôtre.

Du côté de l'EPFL, nous sommes dans l'actualité du projet qui est en chantier puisqu'il faudra impérativement répondre aux exigences de l'Union européenne (obligation de réaliser un *DMP* pour les projets de recherche du programme H2020). Par ailleurs, tout comme l'UNIL, l'EPFL participe également au projet DLCM du programme national CUS P-2, centré exactement sur nos préoccupations du cycle de vie des données. Bien conscients des enjeux de la gestion des données de recherche, une équipe a été formée pour venir en aide aux chercheurs et leur indiquer les bonnes pistes. Une collaboration paraît souhaitable et facilitée par notre proximité géographique. Dans l'immédiat, nous pourrions profiter de leur expérience de l'assistance qu'ils offrent aux chercheurs.

Enfin, grâce aux rencontres à l'EPFL et l'ETHZ, nous avons pu aborder la question des nouveaux outils informatiques utilisés par les chercheurs, tels que les environnements de recherche virtuels (*Virtual Research Environment* ou *VRE*) et les systèmes de gestion de l'Information du Laboratoire (*Laboratory Information Management System* ou *LIMS*), puisque ces deux institutions développent notamment de tels systèmes de gestion et les implémentent de plus en plus dans leurs équipes de recherche.

SECTION 3 : Guide des bonnes pratiques pour la gestion des données de recherche

8. Méthodologie du guide

Suivant le mandat qui nous a été confié, nous avons élaboré un Guide des bonnes pratiques – également une version abrégée et des exemples de posters – le plus conforme possible aux attentes et besoins des chercheurs exprimés au travers de notre large enquête et des entretiens conduits.

Constatant qu'à la question « *Si vous deviez participer à une formation, quels sujets vous intéresseraient le plus ?* » environ la moitié des personnes interrogées ont répondu « le b.a.-ba et les fondamentaux » et que de surcroît les ressources en lignes – plutôt que de la documentation imprimée – étaient plébiscitées, nous avons fait le choix de réaliser un guide qui soit aussi fondamental que possible, en y intégrant des liens vers des *Data Management Plans* déjà existants destinés à ceux qui voudraient en savoir davantage.

Par conséquent, nous avons observé quelles étaient les actuels préexistants, à la fois en matière de formations générales sur la gestion des données d'une part, mais également en matière de ressources et documentations à destination des chercheurs – guides, manuels, modes d'emplois, flyers, posters, etc. Nous en avons recensé¹⁰³ plusieurs et avons dû faire une sélection pour ne retenir que les principales sur lesquelles nous nous sommes basée (uniquement cette sélection figure dans la bibliographie) :

- Guides de l'*Australian National Data Service* ;
- Guides du DCC « *Checklist for a Data Management Plan* » et « *How to Develop a Data Management and Sharing Plan* » ;
- Guide de UK DATA ARCHIVE « *Managing and sharing data : best practice for researchers* » ;
- Outils en ligne « *DMPonline* » et « *DMPTool* ».

¹⁰³ Voici pour information une liste non-exhaustive : guides de l'*Australian National Data Service* ; guides « *How to do...* » du Digital Curation Centre ; guides « *Data Curation in Digital Humanities* » du DHDC ; formation [en ligne] *MANTRA – Research Data Management Training* de l'Université d'Édimbourg (et autres ressources de l'Université) ; formation [en ligne] « *What counts as research data ?* » de l'Université de Bristol ; module [en ligne] Une introduction à la gestion et au partage des données de la recherche de l'Institut de l'information scientifique et technique (Inist) du CNRS (Centre national de la recherche scientifique) ; enfin plus largement les pages et sites [web] des Services de gestion des données de recherche d'Universités étrangères (ex. Édimbourg, Minnesota Monash, Nottingham, Oxford, Sheffield, Queensland).

9. Guide des bonnes pratiques destiné aux chercheurs de l'UNIL

Remarque préliminaire : les propres annexes du guide – version abrégée et exemples de posters – se trouvent aux annexes 15 et 16 du travail.



GUIDE DES BONNES PRATIQUES DE GESTION DES DONNÉES DE RECHERCHE DESTINÉ AUX CHERCHEUSES ET CHERCHEURS DE L'UNIL

Avant de débuter votre projet de recherche, pensez à...



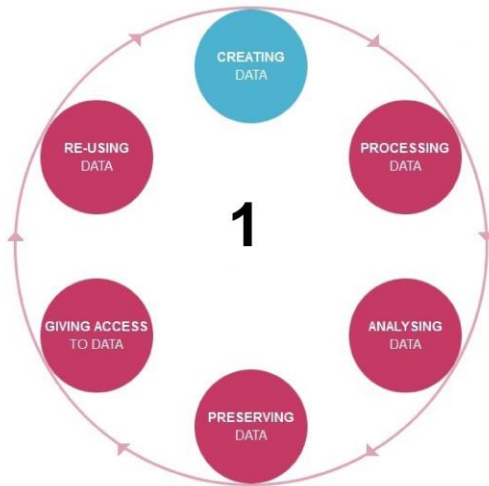
- ▶ Visualiser les grandes étapes que vous allez traverser (cf. schéma ci-contre)
- ▶ Planifier votre recherche en vous basant sur les « Plans de gestion des données » (*Data Management Plans*)
- ▶ Identifier quelles données seront collectées et produites et comment elles seront documentées
- ▶ Anticiper les questions éthiques et juridiques (contrats et formulaires de consentement, brevets, licences,...)

Le *Data Management Plan (DMP)* contient les informations principales suivantes :

- ▶ Contexte de la recherche
- ▶ Rôles et responsabilités
- ▶ Données de recherche (types, formats, modes de création,...)
- ▶ Questions éthiques, juridiques et déontologiques
- ▶ Conditions/droits d'accès/partage/réutilisation des données
- ▶ Stockage, sauvegardes (*backups*) et sécurité des données

Quelques outils en ligne pour créer son *DMP* :

- ▶ **DMPonline** (*Digital Curation Centre*)
<https://dmponline.dcc.ac.uk/>
- ▶ **DMP Tool** (*Uni. of California*)
<https://dmp.cdlib.org/>
- ▶ **Modèles (templates) de *DMP* existants**
<http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>



Lorsque vous créez vos données, pensez à...

- ▶ Privilégier, dans la mesure du possible, les outils et formats ouverts (*Open Source*).
- ▶ Sécuriser les accès à vos données
- ▶ Stocker vos données dans un endroit sécurisé (UNIL de préférence, voir plus loin)
- ▶ Vérifier que vos données sont régulièrement sauvegardées

Nommage des fichiers et organisation des dossiers

- ▶ Pour nommer les fichiers, utilisez une logique que vous comprenez et retenez facilement pour retrouver plus facilement vos fichiers !

ex. *TBCJ2015_GDRUNIL_Into_2015.05.13*

- ▶ Utilisez des abréviations claires pour les différentes versions de vos fichiers

ex. *VP2 (version provisoire) / VF (version finale) / VcorrCJ (version corrigée par ...)*

=> Voir la Liste des abréviations faite par le Service UNIRIS : http://www.unil.ch/uniris/files/live/sites/uniris/files/documents/unites/UNIL_POL_RM_Regles_nommage_annexe_abreviations_VF.pdf

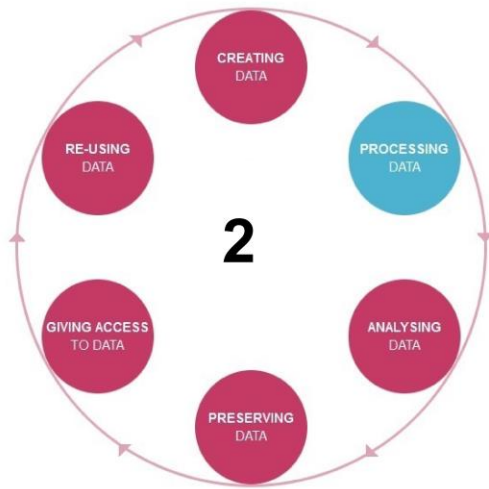
- ▶ Pensez à créer une arborescence pour classer vos fichiers et reprendre la structure de base pour d'autres projets de recherche ultérieurement.

Formats de fichiers

- ▶ Privilégiez les formats standards d'**archivage** et de **conservation**

ex. *formats OpenDocument, PDF/A, CSV, SIARD, TIFF, WAVE, MPEG-4, MPEG2000,...*

=> Voir les formats recommandés par les Archives fédérales suisses : [BAR_Formats_de_fichiers_adapt%C3%A9s_%C3%A0+l_archivage_2014-01463.pdf](http://www.bar.admin.ch/fr/system/attachment/data/document/1463/1463.pdf)

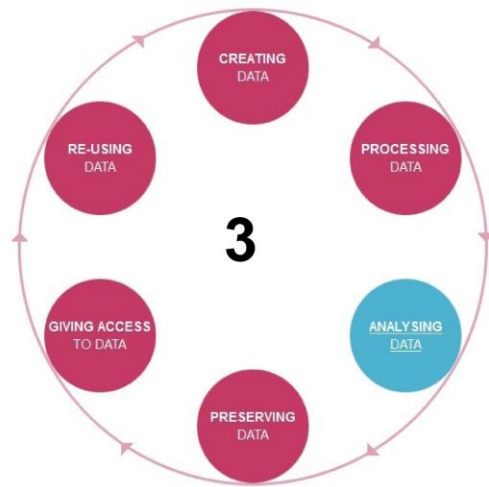


Lorsque vous traitez vos données, pensez à...

- ▶ Saisir, transcrire vos données de manière à ce qu'elles soient valides
- ▶ Contrôler vos données (nettoyage, réécriture et resaisie)
- ▶ Assurer la confidentialité des données (anonymisation)

=> Voir les procédures de FORS pour la confidentialité, l'anonymisation, le consentement et l'accès aux données :

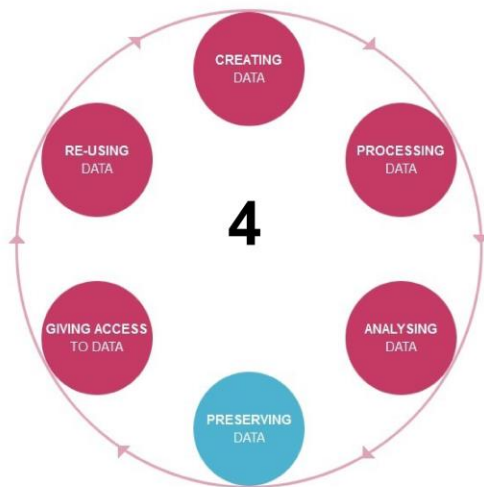
http://forscenter.ch/wp-content/uploads/2013/10/Donnees_qualitatives_archivees_chez_FORS_-_Politique_et_Procedures.pdf



Puis, lorsque vous analysez, interprétez et dérivez vos données

- ▶ Anticiper la publication des futurs résultats de recherche
- ▶ Anticiper la préservation des données
- ▶ Archiver les données, les déposer dans des répertoires (*repositories*)

III



Lorsque vous préservez vos données, pensez à...

- ▶ Migrer vos données vers un format stable et un support approprié
- ▶ Sauvegarder les données (*backups*)
- ▶ Archiver les données, les déposer dans des répertoires (*repositories*)

Les dépôts de données (*repositories*)

- ▶ Assurent la conservation et la préservation à long terme des données
- ▶ Ont une grande capacité de stockage et sont hautement sécurisés
- ▶ Ils utilisent des normes et standards internationaux
- ▶ Visibilisent les données en leur attribuant des DOIs (citation)
- ▶ Favorisent l'*Open Data* tout en respectant la confidentialité des données
- ▶ Les chercheuses et chercheurs sont libres de définir les conditions d'accès ou de restriction
- ▶ Le dépôt est accompagné et fait l'objet de contrats spécifiques

Sur le campus, il existe FORS :

- ▶ Forsbase est le dépôt des données en Sciences sociales : <https://forsbase.unil.ch/>
- ▶ Il existe d'autres dépôts institutionnels utilisés en Suisse (ex. *ETH Data Archive* ou *Zenodo*)

- ▶ Pour rechercher un dépôt de données en qui sont en Open Access : **OpenDOAR** (Directory of Open Access Repositories) <http://www.opendoar.org/>
- ▶ Pour rechercher un dépôt de données de recherche : **re3data** (Registry of Research Data Repositories) : www.re3data.org/

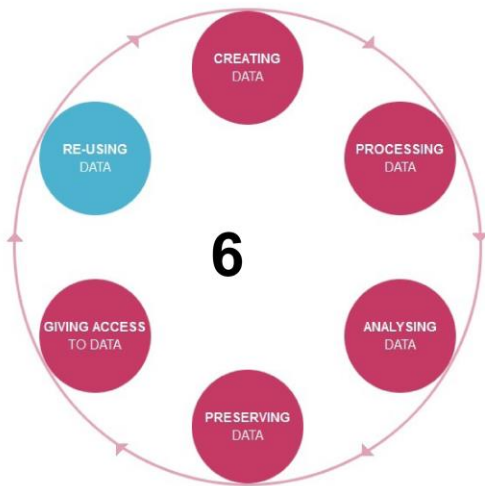
IV



Lorsque vous donnez accès à vos données, pensez à...

- ▶ Définir les droits d'auteur, de propriété intellectuelle et les licences (voir plus loin)
- ▶ Définir les conditions d'accès et de partage des données
- ▶ Valoriser les données en les distribuant et en les partageant

=> voir page suivante



Lorsque vous réutilisez des données, pensez à...

- ▶ Situer les données par rapport à votre nouvelle recherche
- ▶ Contrôler la validité des résultats de recherche
- ▶ Encourager la réutilisation des données pour de nouvelles recherches

v

Finale^{ment}, pensez à **valoriser** vos données de recherche, via :

- ▶ Les dépôts (*repositories*) de données de recherche ouvertes
- ▶ Les revues scientifiques, en annexant vos données aux résultats de recherche
 - => il est souhaité d'attribuer un identifiant *Digital object identifier (DOI)* aux données pour garantir leur identification à long terme
 - => Les DOIs augmentent le nombre de citations
 - => Voir le Guide du *DCC* pour la citation des sets de données :
<http://www.dcc.ac.uk/resources/how-guides/cite-datasets>

Open (Research) Data en Suisse

- ▶ **Opendata.ch**
fr.opendata.ch/
- ▶ **Open Research Data Platform Switzerland**
<http://openresearchdata.ch/>
- ▶ **Open Social Sciences Data**
<http://opendata.forscenter.ch/>
- ▶ **Open Life Sciences Data**
www.lifesciencedata.ch

Pour approfondir le sujet de la gestion des données de recherche...

- ▶ Australian National Data Service – Guides disponibles à l'adresse :
<http://www.ands.org.au/guides/index.html>
- ▶ Guides du Digital Curation Centre – Guides disponibles à l'adresse :
<http://www.dcc.ac.uk/resources/how-guides>

Pour se former en ligne...

- ▶ MANTRA (Research Data Management Traininig) – University of Edinburgh
<http://datalib.edina.ac.uk/mantra/>
- ▶ Module en ligne « Une introduction à la gestion et au partage des données de la recherche » – Inist-CNRS
www.inist.fr/donnees/co/Donnees_recherche_web.html

VI

Pour finir, quelques informations supplémentaires....

Stockage et sauvegardes des données à l'UNIL

Stockage des données à l'UNIL (et sauvegardes automatiques) :

- ▶ Serveurs UNIL plutôt que les disques durs des ordinateurs
- ▶ Crashplan
- ▶ SWITCHdrive (disponible à la fin de l'année)

=> Voir le magazine du Centre informatique (CiNN - printemps 2015) :

<http://wp.unil.ch/cinn/2015/04/ou-sont-stockees-mes-donnees-visite-dun-monde-tres-organise/>

=> Quelques bonnes pratiques

- ▶ Multiple stockage (ex. serveurs UNIL et disque dur externe)
- ▶ Formats de fichiers ouverts et non-propriétaires
- ▶ Générez des PDF/A (format PDF pour la préservation à long terme)
- ▶ Vérifiez régulièrement si les données sont toujours lisibles et intègres (non altérées)

!!! SAUVEGARDE ≠ CONSERVATION A LONG TERME !!!

Cloud Computing

▶ En utilisant des solutions *Cloud* (ex. Dropbox, Google Drive, iCloud, etc.), vous n'êtes plus maître à 100% de vos données !

▶ Exemple - Nouvelles conditions générales de Dropbox (mai 2015):

« [This may required] our systems to access, store and scan Your Stuff*. You give us permission to do those things, and this permission extends to our affiliates and trusted third parties we work with. »

« Copyright Policy » basée sur la législation américaine

« Any juridical proceeding to resolve claims [...] will be brought in the federal or state courts of San Francisco, California. »

« By continuing to use or access the Services after the revisions come into effect, you agree to be bound by the revised Terms. »

*Your Stuff = vos données

VII

Droit d'auteur, propriété intellectuelle et éthique

Cadre législatif

- ▶ Loi fédérale du 9 octobre 1992 sur le droit d'auteur et les droits voisins (Loi sur le droit d'auteur, LDA)
=> *Pour aller plus loin avec le droit d'auteur dans le contexte de l'enseignement supérieur* : **DICE** (Digital Copyright in Education) : <http://www.diceproject.ch/>
« [Our goal is] to provide you with information, resources and a learning tool on how to deal with your personal copyright problems related to higher education. »
- ▶ Autres lois et règlements : <http://www.unil.ch/central/home/menuinst/organisation/les-documents-officiels.html>

Licences

- ▶ Licence **Creative Commons** la «plus restrictive» qui protège le mieux vos données :
CC BY-NC-ND (Attribution - Pas d'Utilisation Commerciale - Pas de Modification) :
<http://creativecommons.org/licenses/>
- ▶ Guide du *DCC* pour attribuer des licences à vos données : <http://www.dcc.ac.uk/resources/how-guides/license-research-data>

Ethique et déontologie

- ▶ *Charte du doctorat UNIL* : http://www.unil.ch/researcher/files/live/sites/researcher/files/shared/Charte_doctorat.pdf
- ▶ *Principes de base et procédures des Académies suisses des sciences* : <http://www.usi.ch/directive-ass-229374.pdf>

N'oubliez pas de contacter les **Services spécialisés à l'UNIL** :

- ▶ **Service des ressources informationnelles et archives (UNIRIS)** : <http://www.unil.ch/uniris/fr/home.html>
- ▶ **Centre informatique (Ci)** : www.unil.ch/ci/fr/home.html
- ▶ **Service juridique** : <http://www.unil.ch/service-juridique/fr/home.html>
- ▶ **Bureau PACTT UNIL-CHUV** : <http://www.pactt.ch/>
- ▶ **Grant office UNIL, Relations FNS, Euresearch, et autres contacts au Dicastère «Recherche & Relations internationales»** : <http://www.unil.ch/central/home/menuinst/organisation/les-autorites/details-direction-unil.html#rech>

Les schémas cycliques qui figurent dans ce guide sont tirés de :
UK Data Archive – ©Copyright 2002-2015 University of Essex. All rights reserved.
JISC (Brown, Bruce and Kernohan, 2015) – © Jisc Published under the CC BY 4.0 licence

VIII

SECTION 4 : Recommandations, synthèse et conclusion

10. Recommandations pour le Service UNIRIS

Au terme de ce travail, des considérations, propositions et recommandations générales peuvent être adressées au Service UNIRIS. Ces recommandations peuvent être une base sur laquelle s'appuyer pour développer et mettre en œuvre un éventuel futur projet de gestion des données de recherche à l'Université de Lausanne.

10.1 Recommandations orientées vers la sensibilisation et la formation des chercheurs de l'UNIL

Notre enquête nous a démontré que les chercheurs de l'UNIL – les répondants du moins – manifestent un intérêt certain pour la gestion des données mais qu'ils se trouvent démunis face à sa mise en œuvre. Par conséquent, il convient de leur procurer les outils adaptés au besoin et au niveau de chacun, puisque leur rôle est avant tout de faire de la recherche, non pas d'endosser celui de spécialistes en GDR. Une sensibilisation à la thématique, un accompagnement pour les bonnes pratiques, l'explicitation des enjeux scientifiques, juridiques et éthiques inhérents devrait leur être offerts. Pour ce faire, nous pouvons envisager d'agir, sous différents angles et avec différentes approches, en proposant l'élaboration future de :

- Site / plateforme web qui mette à disposition des ressources téléchargeables, qui offre la possibilité de réaliser en ligne un *Data Management Plan* (suivant les modèles d'outils tels les *DMP Tools online*) voire même un module de formation à la gestion des données de recherche ;
- Guides adaptés à chacune des sept Faculté qui précisent les actions spécifiques à réaliser et qui tiennent compte de la discipline de recherche. Ces guides pourraient pourquoi pas mentionner les personnes à contacter, au sein des Facultés, en cas d'urgence (SOS) ;
- Guides qui détaillent certaines actions / thématiques particulières (ex. normes et standards, sécurité, stockage, dépôt des données, conservation à long terme, *Open Access*,...).

Notre guide des bonnes pratiques n'a pas la prétention d'y répondre mais il peut ouvrir la voie.

10.2 Recommandations orientées vers l'élaboration d'un Centre de soutien à la gestion des données de recherche à l'UNIL

10.2.1 Poursuite des relations engagées et élargissement du réseau

Grâce aux entretiens menés dans le cadre de ce travail, nous avons pu identifier, rencontrer, et donc finalement constituer un réseau de personnes intéressées par la

thématique de la gestion des données de recherche et impliquées, à divers degrés et niveaux hiérarchiques – dans des projets en lien avec le nôtre. Nous préconisons vivement au Service UNIRIS de maintenir ces contacts, de les intensifier, poursuivre cette démarche et d'ouvrir ce réseau à d'autres acteurs concernés, à échelle locale et nationale. Idéalement et dans la mesure du possible, ce réseau devrait également inclure des spécialistes internationaux. La transmission de l'expérience par des relations directes nous semble efficace dans le sens qu'elle permet d'identifier plus rapidement les démarches qui ont fait leur preuve pour en tirer profit de manière à avancer plus vite.

10.2.2 Constitution d'un groupe interne à l'UNIL

Dans un premier temps, il nous semble plus pertinent et utile de s'appuyer sur les forces vives du campus lausannois pour réunir et constituer, à court terme un groupe chargé des questions de gestion des données de recherche. Selon nous, ce groupe devrait être piloté par un comité (Copil), articulé en cinq pôles d'expertise, et constitué de représentants de :

- Service UNIRIS (pôle Sciences de l'information)
- Le Dicastère « Recherche & Relations internationales » et ses Services¹⁰⁴, les sept Facultés et la Commission de la recherche scientifique (pôle recherche)
- Centre informatique et experts IT (pôle informatique)
- Service juridique et PACTT (pôle juridique)
- Service UNICOM (pôle communication et information au public)

En plus du Copil, ce groupe pourrait réunir une série d'autres parties prenantes internes concernées (voir notre *brainstorming* en annexe 17). En outre, des membres d'institutions partenaires externes pourraient être regroupés en « comité consultatif d'experts » – nous pensons à la BCUL, à l'EPFL, FORS, et au SIB notamment. L'objectif principal de ce groupe serait de soumettre à la Direction de l'UNIL un projet de création d'un « Centre de soutien à la gestion des données de recherche » destiné à l'ensemble de la communauté UNIL. Pour réaliser cet objectif et espérer pouvoir obtenir l'aval et le soutien de la Direction, nous recommandons que le groupe définisse préalablement un cadre de gestion des données de recherche, en produisant notamment les documents suivants : la Politique et la Stratégie de gestion des données de recherche ainsi qu'un *Business plan* et un plan de viabilité du projet. A ces documents pourraient s'ajouter d'autres « arguments » en faveur du Centre de soutien, tels qu'un éventuel rapport détaillé des résultats de notre enquête qui démontre un réel besoin d'accompagnement, la valorisation des projets actuellement menées dans les Facultés (cf. point 5.2.3) ou

¹⁰⁴ Services au nombre de trois : Service Relations internationales, Euresearch et PACTT.

encore l'appui de personnes impliquées pour l'UNIL dans les projets nationaux (cf. point 5.1). Ces observations conduisent à notre prochaine recommandation :

10.2.3 Création d'un Centre de soutien à la gestion des données de recherche à l'UNIL

A l'instar d'autres Universités qui disposent déjà de Services dédiés à la gestion des données de recherche, nous recommandons vivement la création d'un « Centre de soutien¹⁰⁵ » à la gestion des données de recherche. Ce Centre serait chargé d'assurer que la politique et la stratégie soient respectées et mises en œuvre, il aurait également pour mission de soutenir les chercheurs de l'UNIL en leur offrant un cadre favorisant la bonne gestion des données de recherche : infrastructure informatique et technique d'un côté, mais également des formations (intégrées au cursus, workshops, formations en groupe,...), des conseils personnalisés, des ressources à disposition (guides méthodologiques, procédures, marches-à-suivre,...) ou toute autre forme d'aide, à la fois par le biais d'intervention de spécialistes (ID, IT, juristes) et à la fois en ligne (Internet). Dans cette optique, il nous semble nécessaire de s'appuyer sur l'expertise et le savoir-faire du Service UNIRIS en matière de « gouvernance informationnelle » pour créer un Centre qui, lui, serait orienté vers la « gouvernance des données »¹⁰⁶. De même, nous encourageons l'utilisation du nombre de récentes publications – ouvrages de référence (en bibliographie) en littérature professionnelle ID sur cette thématique et que nous avons pu voir fleurir durant les cinq mois consacrés à notre projet – dont certaines sont dédiées à la mise en œuvre et le développement de tels services au sein des Universités. Enfin, toujours dans l'optique de la création du Centre, nous ne pouvons qu'encourager à s'inspirer des expériences vécues par d'autres, soit au travers des *case studies* (études de cas) de la littérature, soit, encore mieux, au travers de collaborations avec des institutions partenaires et plus proches géographiquement, comme FORS par exemple.

¹⁰⁵ Terme inspiré du *Centre de soutien à l'enseignement* (CSE) de l'UNIL.

¹⁰⁶ La gouvernance des données (*Data Governance*), est définie par Smith (2014) comme :
« [...] the system of decision rights and responsibilities covering who can take what actions with data, when, under what circumstances, and using what methods ».

11. Synthèse générale et conclusion

Nous pouvons d'abord établir qu'actuellement, la gestion de masses de données de recherche est un enjeu scientifique majeur mais aussi un enjeu transdisciplinaire au sens où il implique des acteurs issus de pôles d'intérêts variés qu'il s'agit de mettre en réseau pour collaborer. Dans ce cadre, les spécialistes ID ont un rôle important à jouer en raison de leurs compétences propres : garants de la ligne ou du « continuum » de *Curation* (Treloar et al., 2007), des bonnes pratiques en matière d'efficacité de gestion, d'accessibilité, de partage, de publication, de mise en valeur, d'archivage et de conservation pérenne de l'information, ils doivent contribuer à veiller au respect des standards en vigueur mais aussi participer à l'élaboration de nouvelles normes de travail et développer les outils de gestion adéquats.

Au travers du mandat d'UNIRIS, nous avons observé que les chercheurs souhaitent des réponses concrètes pour leur pratique. Ainsi, ils ont des besoins et des demandes d'information et de soutien tant au sujet des fondamentaux de la gestion de leurs données que de l'usage des infrastructures existantes que certains ignorent. Ils souhaitent également solliciter de l'aide pour des solutions techniques, juridiques qui sont à leur disposition. Nous avons encore constaté qu'une majorité des chercheurs est favorable à l'idée de partager librement ses données de recherches à condition que des garde-fous soient respectés : garantir la paternité de la découverte scientifique et de la recherche ou encore prévenir les utilisations fallacieuses ou abusives par des tiers sont des conditions non négociables à l'accès libre des données de recherches.

Nous avons relevé que l'archivage et la préservation des documents à long terme revêtent une importance certaine aux yeux des chercheurs mais qu'ils ont une connaissance hasardeuse ou maladroite des dépôts de données (*repositories*) chargés d'assurer l'authenticité, l'intégrité, la fiabilité, l'exploitabilité et la conservation des données, raison pour laquelle ils en font trop peu usage. Nous avons enfin pris note qu'ils sollicitent clairement des solutions « institutionnelles » encadrées par des infrastructures professionnelles aussi bien pour la gestion que la mise en valeur et la conservation de leurs données. Malgré tout, les chercheurs sont déjà mis à contribution dans d'autres tâches que la recherche et sont soumis à des exigences, parfois chronophages, comme les rapports ou la publication de projet pour lever des fonds. Ils ne souhaitent donc pas que la gestion des données représente une charge de travail supplémentaire.

Fort du constat qu'il existe une demande, la question est de savoir comment y répondre compte tenu de la diversité des Facultés, des intervenants et des types de recherches. Dans l'immédiat, notre guide est une première réponse pratique à l'intention des chercheurs. À long terme, comme nous l'avons recommandé précédemment, l'idéal serait de créer un centre de soutien dédié à la gestion des données de recherche. Pour ce faire, notre enquête et notre travail en général pourrait constituer une base de réflexion solide. Il est clair que, dans cette perspective, l'analyse de nos résultats mériterait d'être affinée car les éléments de réponse à notre disposition nous ont permis de dégager un profil de la situation actuelle et des attentes principales des chercheurs à un niveau qui demeure général. Selon nous, il pourrait être utile de poursuivre notre démarche pour être en mesure de cibler les besoins pour chacune des Facultés. C'est la raison pour laquelle, constatant que les entretiens individuels sont d'une aide précieuse et nous fournissent des informations *de nature qualitatives*, nous en préconisons la poursuite. Au terme de ce travail, nous sommes convaincue de l'importance de la mise en œuvre d'un outil de gestion des données de recherche qui simplifierait le travail des chercheurs, ceci dans un cadre institutionnel souple pour permettre leur adhésion mais suffisamment contraignant pour faciliter l'accès aux données.

Ce travail de Bachelor m'a permis, dans le cadre bien structuré d'UNIRIS d'explorer une thématique dans le domaine de la gestion des données de recherche. J'ai pu approfondir mes connaissances au moyen de sources variées et mettre en œuvre les outils professionnels acquis à la HEG dans le but de bâtir un projet concret dans une terre encore en friche. J'ai pu mesurer concrètement l'investissement en termes de temps et d'engagement que nécessitent la préparation, la création et la diffusion d'un questionnaire ainsi que la compilation et l'analyse des résultats sous forme de tableaux lisibles. J'ai pu me rendre compte de l'importance des entretiens qui reflètent la réalité vécue, les exigences de terrain et les contradictions qui peuvent surgir entre différents intérêts et point de vue selon les besoins des chercheurs. J'ai pu observer ce qui se fait déjà et réfléchir à élaborer des réponses à ces besoins en espérant que mon travail et mes échanges contribuent à améliorer la situation actuelle. Confrontée à l'étendue de ce champ d'action, il a fallu faire des choix afin de bien le limiter. Consciente que mon travail n'investigue qu'une toute petite partie du problème soulevé dans l'introduction du thème, j'ai volontairement mis de côté d'autres sujets comme la propriété, la protection des données, le droit d'auteur, le choix de l'outil de stockage qui sont intimement liés à la gestion.

Bibliographie

Ouvrages de référence

HARVEY, Ross, 2010. Digital Curation : a how-to-do-it manual. New York : Neal-Schuman Publishers, 2010. How-to-do-it manuals, no 170. ISBN 978-1-55570-694-4

KRIER, Laura and STRASSER, Carly A., 2014. Data Management for Libraries : a LITA Guide. Chicago : ALA TechSource, an imprint of the American Library Association, 2014. LITA guides. ISBN 978-1-55570-969-3 (paper)

MARTIN, Victoria, 2014. Demystifying eResearch : a primer for librarians. Santa Barbara, CA : Libraries Unlimited, 2014. ISBN 978-1-61069-520-6

PRYOR, Graham [ed.], 2012. Managing research data. London : Facet Publishing, 2012. ISBN 978-1-85604-756-2

PRYOR, Graham [ed.], JONES, Sarah [ed.] and WHYTE, Angus [ed.], 2014. Delivering research data management services : fundamentals of good practice. London : Facet Publishing, 2014. ISBN 978-1-85604-756-2

RAY, Joyce M. [ed.], 2014. Research Data Management : Practical Strategies for Information Professionals. West Lafayette : Purdue University Press, 2014. ISBN 978-1-55753-664-8

Définitions de la donnée de recherche

BELLINGER, CASTRO AND MILLS, 2004. Data, Information, Knowledge, and Wisdom. *Site Internet de Systems Thinking* [en ligne]. 2004. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.systems-thinking.org/dikw/dikw.htm>

BERT, Jean-François, 2014. *Qu'est-ce qu'une archive de chercheur ?* [en ligne]. Marseille : OpenEdition Press, 11 juin 2014. [Consulté le 28 mai 2015]. Encyclopédie numérique, 4. ISBN numérique 9782821834620. Disponible à l'adresse : <http://books.openedition.org/oep/438?lang=fr>

COMITE CONSULTATIF POUR LES SYSTEMES DE DONNEES SPATIALES, 2005. Modèle de référence pour un Système ouvert d'archivage d'information (OAIS) [Livre Bleu]. *Site Internet du groupe PIN de l'association Aristote* [en ligne]. Mars 2005. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://pin.association-aristote.fr/lib/exe/fetch.php/public/documents/norme_oais_version_francaise.pdf

GAILLARD, Rémi, 2014. *De l'open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?* [en ligne]. Lyon : Enssib. Mémoire d'étude. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf>

ORGANISATION DE COOPERATION ET DE DEVELOPPEMENT ECONOMIQUE, 2007. Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics. *Oecd.org* [en ligne]. Avril 2007. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.oecd.org/fr/sti/sci-tech/38500823.pdf>

UK DATA ARCHIVE, 2002. Create and manage data : research data lifecycle. *Site du UK DATA ARCHIVE* [en ligne]. 2002-2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.data-archive.ac.uk/create-manage/life-cycle>

RESEARCH DATA ALLIANCE EUROPE, 2015. Report on the RDA-MPG Science Workshop on Data. *Site Internet du RDA Europe* [en ligne]. Avril 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : <https://europe.rd->

alliance.org/sites/default/files/report/RDA-Europe-Science-Workshop-Report_final_April2014.pdf

Digital Curation et Research Data Management

DIGITAL CURATION CENTRE, 2004. DCC Charter and Statement of Principles. *Site Internet du DCC* [en ligne]. 2004-2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.dcc.ac.uk/about-us/dcc-charter/dcc-charter-and-statement-principles>

DIGITAL CURATION CENTRE, 2004. DCC Curation Lifecycle Model. *Site Internet du DCC* [en ligne]. 2004-2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

DONNELLY, Martin, 2008. RDMF2 [Research Data Management Forum] : Core Skills Diagram. *Site Internet du Research Data Management Forum* [en ligne]. 17 décembre 2008. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://data-forum.blogspot.ch/2008/12/rdmf2-core-skills-diagram.html>

HARVARD UNIVERSITY, 2014. Harvard Research Data Security Policy. *Vpr.harvard.edu* [en ligne]. 2012. Révision en automne 2014. [Consulté le 20 mars 2015]. Disponible à l'adresse : http://files.vpr.harvard.edu/files/vpr-documents/files/hrdsp_10_14_14_final_edits.pdf

HIGGINS, Sarah, 2008. The DCC Curation Lifecycle Model. *Site Internet du International Journal of Digital Curation* [en ligne]. 2008. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.ijdc.net/index.php/ijdc/article/viewFile/69/48>

JONES, Sarah, 2012. Developments in Research Funder Data Policy. *Site Internet du International Journal of Digital Curation* [en ligne]. 2012. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://ijdc.net/index.php/ijdc/article/view/209/278>

LYON, Liz, 2007. Dealing with Data: Roles, Rights, Responsibilities and Relationships : Consultancy Report. 19 juin 2007. *Site Internet d'UKOLN* [en ligne]. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf

PENNOCK, Maureen, 2007. Digital Curation : A Life-Cycle Approach to Managing and Preserving Usable Digital Information. [Article for Libraries & Archives Journal, Issue 1, 2007]. *Site Internet d'UKOLN* [en ligne]. Janvier 2007. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf

TRELOAR, Andrew, GROENEWEGEN, David and HARBOE-REE, Cathrine, 2007. The Data Curation Continuum. *D-Lib Magazine* [en ligne]. Volume 13, Number9/10. Septembre/octobre 2007. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.dlib.org/dlib/september07/treloar/09treloar.html>

UNIVERSITY OF CENTRAL FLORIDA. Libraries Research Lifecycle Committee. (2012). The research lifecycle at UCF [Online Graphic]. *Site Internet de l'UCF* [en ligne]. 2012. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://library.ucf.edu/ScholarlyCommunication/UCFResearchLifecycle.pdf>

UNIVERSITY OF OXFORD, 2015. University of Oxford Policy on the Management of Research Data and Records. *Site Internet de l'Université d'Oxford* [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://researchdata.ox.ac.uk/university-of-oxford-policy-on-the-management-of-research-data-and-records/>

Guides et formations (notre sélection)

AUSTRALIAN NATIONAL DATA SERVICE, 2011. ANDS Guides and Publications. *Ands.org* [en ligne]. 2011. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://ands.org.au/guides/index.html>

DIGITAL CURATION CENTRE, 2014. DCC Checklist for a Data Management Plan (v.4.0, 2014). *Site Internet du DCC* [en ligne]. 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.dcc.ac.uk/resources/data-management-plans/checklist>

DIGITAL CURATION CENTRE, 2015. *DMPonline* [en ligne]. 2004-2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <https://dmponline.dcc.ac.uk/>

DIGITAL CURATION CENTRE, 2011. How to Develop a Data Management and Sharing Plan. *Site Internet du DCC* [en ligne]. 8 septembre 2011. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.dcc.ac.uk/resources/data-management-plans/checklist>

UK DATA ARCHIVE, 2011. Managing and sharing data : best practice for researchers. *Site du UK DATA ARCHIVE* [en ligne]. Mai 2011. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>

UNIVERSITY OF CALIFORNIA, 2015. *DMPTool [Build your Data Management Plan]* [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <https://dmp.cdlib.org/>

Enquête et entretiens

BEAGRIE, Neil, BEAGRIE, Robert and ROWLANDS, Ian, 2009. Research Data Preservation and Access: The Views of Researchers. [Findings from a UKRDS survey]. *Site Internet de Ariadne [Web Magazine for Information Professionals]* [en ligne]. 30 juillet 2009. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.ariadne.ac.uk/issue60/beagrie-et-al>

BEPRESS, 2014. Data Interview. Data. Paper 2. *Site Internet de Digital Commons* [en ligne]. 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://digitalcommons.bepress.com/cgi/viewcontent.cgi?article=1004&context=data>

GIBBS, Harry, 2009. [University of] Southampton Data Survey: Our Experience and Lessons Learned. DISC-UK DataShare Project. *Disc-uk.org* [en ligne]. Mars 2007-2009. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.disc-uk.org/docs/SouthamptonDAF.pdf>

EKMEKCIOGLU, Çuna and RICE, Robin, 2009. Edinburgh Data Audit Implementation Project: Final Report. [Programme/Project deposit]. *Jisc.ac.uk* [en ligne]. 30 janvier 2009. Dernière modification le 1^{er} juillet 2013. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://repository.jisc.ac.uk/id/eprint/283>

MOSSINK, Wilma and BIJSTERBOSCH, Magchiel, 2013. European Landscape Study of Research Data Management. *Site Internet de SIM4RDM* [en ligne]. Janvier 2013. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20final%2025.01.12.pdf>

MUET, Florence, 2003/2007. *Concevoir et exploiter un guide d'entretien. Notes techniques* [document PDF]. Support de cours : Cours « Analyse des besoins », Haute école de gestion de Genève, filière information documentaire, années académiques 2003/2007.

MUET, Florence, 2003/2007. *Conduire une enquête auprès de son public. Notes techniques* [document PDF]. Support de cours : Cours « Analyse des besoins », Haute école de gestion de Genève, filière information documentaire, années académiques 2003/2007.

MUET, Florence, 2003/2007. *Conduire un entretien semi-directif. Notes techniques* [document PDF]. Support de cours : Cours « Analyse des besoins », Haute école de gestion de Genève, filière information documentaire, années académiques 2003/2007.

MUET, Florence, 2003/2007. *Rédiger un questionnaire. Notes techniques* [document PDF]. Support de cours : Cours « Analyse des besoins », Haute école de gestion de Genève, filière information documentaire, années académiques 2003/2007.

PARSONS, Thomas, GRIMSHAW, Shirley and WILLIAMSON, Laurian, 2013. Research data management survey : report. Project Report. *Site Internet de University of Nottingham* [en ligne]. 9 février 2013. Dernière modification le 6 mars 2013. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://eprints.nottingham.ac.uk/id/eprint/1893>

SIMUKOVIC, Elena, KINDLING Maxi and SCHIRMBACHER, Peter, 2013. Humboldt-Universität zu Berlin Research Data Management Survey Results. *Zenodo.org* [en ligne]. 21 octobre 2013. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://zenodo.org/record/7468>

SIMUKOVIC, Elena, KINDLING Maxi and SCHIRMBACHER, Peter, 2013. Humboldt-Universität zu Berlin Research Data Management Survey Results. Comparing respondent groups « Professor » and « Research associate ». *Zenodo.org* [en ligne]. 21 octobre 2013. [Consulté le 28 mai 2015]. Disponible à l'adresse : <https://zenodo.org/record/7469>

TÖWE, Matthias, SCHEID, Susanne, 2011. Töwe, Matthias and Scheid, Susanne. User expectations in archived research data [in Eidgenössische Technische Hochschule Zürich]. *Site Internet de ETH-Bibliothek* [en ligne]. 19 octobre 2011. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://dx.doi.org/10.3929/ethz-a-006691697>

UNIVERSITY OF LEEDS, 2012. Research data survey. *Site Internet de Leeds University Library* [en ligne]. Juillet-octobre 2012. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://library.leeds.ac.uk/info/377/roadmap/122/roadmap_project_outputs/7

UNIVERSITY OF OXFORD, WILSON, James et al., 2012. University of Oxford Research Data Management Survey 2012 : The Results. *Blog DaMaRO de University of Oxford* [en ligne]. 3 janvier 2013. [Consulté le 28 mai 2015]. Disponible aux adresses : <http://blogs.it.ox.ac.uk/damaro/2013/01/03/university-of-oxford-research-data-management-survey-2012-the-results/> et http://damaro.oucs.ox.ac.uk/docs/OxfordRDMsurvey2012_public.xlsx

WHYTE, Angus, and SISU, Diana, 2014. Digital Curation Centre's 2014 RDM Strategy to Action Survey. *Zenodo.org* [en ligne]. 30 juin 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://zenodo.org/collection/user-dcc-rdm-2014>

WITT, Michael and CARLSON, Jake, 2007. Conducting a Data Interview. *Site Internet de Purdue University* [en ligne]. 12 décembre 2007. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1092&context=lib_research

Contexte national suisse

ACADEMIES SUISSES DES SCIENCES, 2008. L'intégrité dans la recherche scientifique : principes de base et procédures. *Site Internet des ASS* [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.akademien-schweiz.ch/fr/index/Schwerpunktthemen/Wissenschaftliche-Integritaet.html>

ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE, 2015. Les données de la recherche : une mine d'or à domestiquer et valoriser. *Flash journal interne* [en ligne]. 21 mai 2015. No 4, pp. 16-17. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://medias.epfl.ch/files/content/sites/medias/files/Flash/Flash%2004-2015.pdf>

FONDS NATIONAL SUISSE DE LA REHERCHE SCIENTIFIQUE, 2015. Open Access. Site Internet du FNS [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open-access/Pages/default.aspx>

FORS CENTER, 2014. Manifesto on Data access and research transparency (DART) in Switzerland. *Site Internet de FORS* [en ligne]. 21 novembre 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://forscenter.ch/wp-content/uploads/2014/11/Manifesto-DART.pdf>

Loi du 6 juillet 2004 sur l'Université de Lausanne (LOI 414.11). *Site officiel du Canton de Vaud* [en ligne]. Adopté le 06.07.2004, entrée en vigueur le 01.01.2005 - Etat au 01.02.2011 (en vigueur). [Consulté le 20 mars 2015]. Disponible à l'adresse : http://www.rsv.vd.ch/rsvsite/rsv_site/doc.pdf?docId=685872&Pvigueur=&Padoption=&Pcurrent_version=99&PetatDoc=abroge&Pversion=&docType=&page_format=A4_3&isSV=true&isSjL=true&outformat=pdf&isModifiante=false

TÖWE, Matthias, 2014. Open Data and Digital Curation: Why each is only worth half without the other. In : *EPFL. Open research data: the future of science, Lausanne, Forum Rolex Learning Center (EPFL Campus), 28 octobre 2014* [présentation ppt en ligne]. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://library2.epfl.ch/conf/opendata>

UNIVERSITÉ DE LAUSANNE. La Charte [de l'Université de Lausanne]. *Unil.ch* [en ligne]. [Consulté le 19.01.2015]. Disponible à l'adresse : http://www.unil.ch/central/files/live/sites/central/files/docs/charte_UNIL.pdf

UNIVERSITÉ DE LAUSANNE, 2004. Directive de la Direction 0.12. Ressources informationnelles et archives de l'Université de Lausanne. *Unil.ch* [en ligne]. Modifications adoptées par la Direction dans sa séance du 30 mars 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.unil.ch/interne/files/live/sites/interne/files/textes_leg/0_aff_gen/dir0_12_archives3.pdf

UNIVERSITÉ DE LAUSANNE, 2011. Plan d'intentions de l'Université de Lausanne 2012-2016 [Version remise au Département de la Formation et de la Jeunesse du Canton de Vaud]. *Unil.ch* [en ligne]. 15 décembre 2011. [Consulté le 19.01.2015]. Disponible à l'adresse : http://www.unil.ch/central/files/live/sites/central/files/docs/plan_intentions_12_16.pdf

UNIVERSITÉ DE LAUSANNE. Service des ressources informationnelles et archives, 2014. Politique de records management et d'archivage pour une gouvernance informationnelle. *Unil.ch* [en ligne]. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.unil.ch/uniris/files/live/sites/uniris/files/documents/references/UNIL_POL_Records_management_archivage_VF.pdf

UNIVERSITÉ DE LAUSANNE. Service des ressources informationnelles et archives, 2014. Mission, objectifs et prestations d'UNIRIS. *Unil.ch* [en ligne]. État au 18.06.2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.unil.ch/uniris/files/live/sites/uniris/files/documents/references/UNIRIS_Mission_objectifs_prestations_2014.pdf

UNIVERSITÉ DE LAUSANNE, 2015. Recherche. A la pointe de la recherche scientifique. *Unil.ch* [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.unil.ch/central/home/menuinst/recherche.html>

UNIVERSITÉ DE LAUSANNE, 2015. La Direction de l'UNIL : organigramme. *Unil.ch* [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.unil.ch/central/home/menuinst/organisation/la-direction.html>

SECRETARIAT D'ETAT A LA FORMATION, A LA RECHERCHE ET A L'INNOVATION, 2015. Encouragement de la formation, de la recherche et de l'innovation 2013-2016. *Site Internet du SEFRI* [en ligne], 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.sbf.admin.ch/org/01645/index.html?lang=fr>

SWISSUNIVERSITIES, 2015. PROGRAMME CUS 2013-2016 P-2 INFORMATION SCIENTIFIQUE: ACCES, TRAITEMENT ET SAUVEGARDE. SITE INTERNET DE SWISSUNIVERSITIES [EN LIGNE]. 2015. [CONSULTE LE 28 MAI 2015]. DISPONIBLE A L'ADRESSE : <http://www.swissuniversities.ch/fr/organisation/projets-et-programmes/programme-cus-2013-2016-p-2-information-scientifique-acces-traitements-et-sauvegarde/>

Autres

ACHENBACH, Joel, 2015. The new scientific revolution: Reproducibility at last. *The Washington Post* [en ligne]. 27 janvier 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.washingtonpost.com/national/health-science/the-new-scientific-revolution-reproducibility-at-last/2015/01/27/ed5f2076-9546-11e4-927a-4fa2638cd1b0_story.html

CNRS. Comité d'éthique du CNRS, 2014. Promouvoir une recherche intègre et responsable : un guide. *Cnrs.fr* [en ligne]. Juillet 2014. [Consulté le 28 mai 2015]. Disponible à l'adresse : http://www.cnrs.fr/comets/IMG/pdf/guide_promouvoir_une_recherche_integre_et_responsable_8septembre2014.pdf

MCNUTT, Marcia, 2015. Data, eternal. [Editorial]. *Science Magazine* [en ligne]. 2 janvier 2015. Vol. 347 no. 6217 p. 7. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.sciencemag.org/content/347/6217/7.full>

UNIVERSITY COLLEGE LONDON, 2015. MA/MSc in Digital Humanities. What we do. *Site internet de l'UCL* [en ligne]. 1999–2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://www.ucl.ac.uk/dis/taught/pg/tmadhmsing01>

PIKTOCHART INFOGRAPHICS, 2015. Piktochart.com [en ligne]. 2015. [Consulté le 28 mai 2015]. Disponible à l'adresse : <http://piktochart.com/>

Annexe 1 : Classification des données de recherche

Données	Méthodes et processus	Finalité et valeur	Exemples
<i>D'observation</i>	Capturées en temps réel ; généralement uniques et irremplaçables	Conservation généralement pérenne	Imagerie cérébrale, relevés climatiques et météorologiques, images satellite
<i>Expérimentales</i>	Capturées grâce à des équipements de labos ; souvent reproductibles ; peuvent être très coûteuses	Conservation à long terme justifiée, en fonction des moyens investis et du potentiel de réutilisation des données	Séquençage ADN, chromatographie, spectrométrie de masse
<i>Computationnelles, de modèles ou de simulations</i>	Générées depuis des modèles d'essais ; reproductibles (s'il existe une documentation détaillée et accessible)	Conservation pérenne non nécessaire s'il existe de la documentation ; par contre, conservation indispensable des modèles et jeux de métadonnées	Modèles climatiques, économiques, mathématiques
<i>Dérivées ou compilées</i>	Issues du traitement et de l'analyse des données brutes ; reproductibles ; très coûteuses	Conservation pérenne des différentes versions justifiée dans certaines circonstances	Bases de données compilées, text/data mining, modèles 3D
<i>De référence ou canoniques</i>	Préalablement organisées, gérées voire publiées	Conservation à long terme généralement assurée par d'autres	Banques de données de séquences ADN (ex. GenBank), de structures chimiques ou portails de données spatiales

(Source : Gaillard, 2014)

Annexe 2 : Détail des étapes du cycle de vie de la donnée de recherche

- **Création** des données
 - Concevoir et conceptualiser la recherche
 - Plan de gestion des données (*Data Management Plan*)
 - Plan de consentement pour le partage (contrats, formulaires,...)
 - Localiser et repérer les données existantes
 - Collecter les données (expérimenter, observer, mesurer, simuler)
 - Capturer et créer les métadonnées
- **Traitement** des données
 - Saisir, numériser, transcrire, traduire les données
 - Vérifier, valider, nettoyer les données
 - Anonymiser les données, si nécessaire
 - Décrire les données
 - Gérer et stocker les données
- **Analyse** des données
 - Interpréter les données
 - Dériver les données
 - Produire des résultats de recherche (output)
 - Publications d'auteur
 - Préparer les données pour la préservation
- **Préservation** des données
 - Migrer les données vers le meilleur format
 - Migrer les données vers un support approprié
 - Sauvegarder (back up) et stocker les données
 - Créer les métadonnées et la documentation
 - Archiver les données
- **Accès** aux données
 - Distribuer les données
 - Partager les données
 - Contrôler les accès
 - Établir les droits d'auteur (copyright)
 - Promouvoir les données
- **Réutilisation** des données
 - Suivi de la recherche
 - Nouvelle recherche
 - Entreprendre les évaluations de la recherche (reviews)
 - Contrôler les résultats de recherche
 - Enseigner et apprendre

(Source : *UK Data Archive*, 2002-2015)

Annexe 3 : Détail des actions du *DCC Curation Lifecycle Model*

The Curation Lifecycle	
<p>The DCC Curation Lifecycle Model provides a graphical high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. The model can be used to plan activities within an organisation or consortium to ensure that all necessary stages are undertaken, each in the correct sequence. The model enables granular functionality to be mapped against it; to define roles and responsibilities; and build a framework of standards and technologies to implement. It can help with the process of identifying additional steps which may be required, or actions which are not required by certain situations or disciplines, and ensuring that processes and policies are adequately documented.</p>	

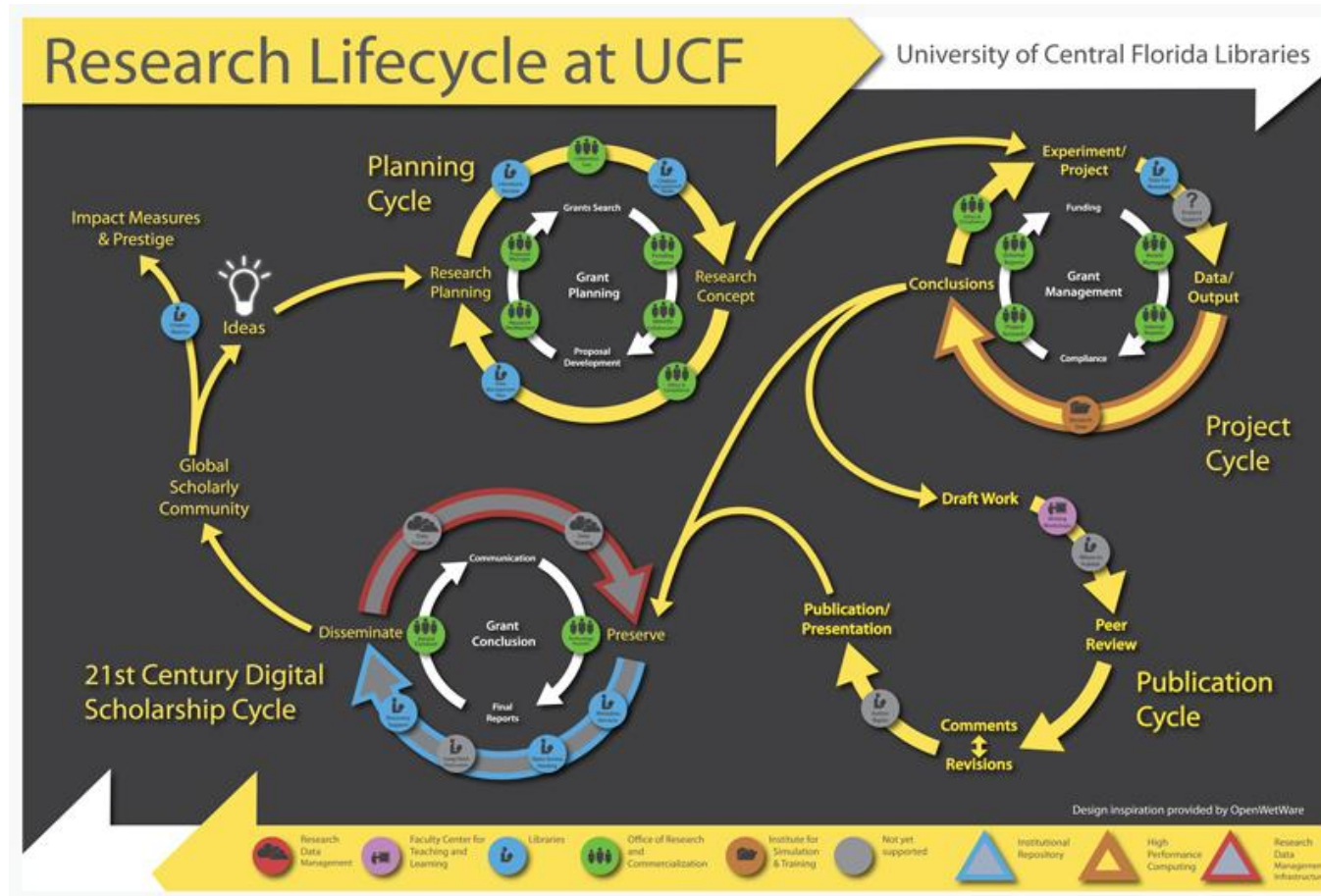
Data (Digital Objects, or Databases)	
	Data, any information in binary digital form, is at the centre of the Curation Lifecycle. This includes:
Digital Objects	<ul style="list-style-type: none"> - Simple Digital Objects are discrete digital items; such as textual files, images or sound files, along with their related identifiers and metadata. - Complex Digital Objects are discrete digital objects, made by combining a number of other digital objects, such as websites.
Databases	Structured collections of records or data stored in a computer system.

Full Lifecycle Actions	
Description and Representation Information	Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.
Preservation Planning	Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.
Community Watch and Participation	Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.
Curate and Preserve	Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.

Sequential Actions	
Conceptualise	Conceive and plan the creation of data, including capture method and storage options.
Create and Receive	<ul style="list-style-type: none"> - Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation. - Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.
Appraise and Select	Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.
Ingest	Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.
Preservation Action	Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.
Store	Store the data in a secure manner adhering to relevant standards.
Access, Use and Reuse	Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.
Transform	<p>Create new data from the original, for example</p> <ul style="list-style-type: none"> - By migration into a different format. - By creating a subset, by selection or query, to create newly derived results, perhaps for publication.

Occasional Actions	
Dispose	Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data are destroyed. The data's nature may, for legal reasons, necessitate secure destruction.
Reappraise	Return data which fails validation procedures for further appraisal and reselection.
Migrate	Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.

Annexe 4 : Research Lifecycle at University of Central Florida



(Source : UCF, 2012)

Annexe 5 : Politique de gestion des données de recherche de l'Université d'Oxford



Policy on the Management of Research Data and Records

1. The University of Oxford seeks to promote the highest standards in the management of research data and records¹ as fundamental to both high quality research² and academic integrity.
2. The University recognises that accurate and retrievable research data are an essential component of any research project and necessary to verify and defend, when required, the process and outcomes of research. Research data are valuable to researchers for the duration of their research, and may well have long-term value for research, teaching and for wider exploitation for the public good, by individuals, government, business and other organisations, as a project develops and after research results have been published.
3. The University acknowledges its obligations under research funders' data-related policy statements³ and codes of practice to ensure that sound systems are in place to promote best practice, including through clear policy, guidance, supervision, training and support.
4. Researchers⁴, departments/faculties, divisions, central administrative units and service providers and, where appropriate, research sponsors and external collaborators, need to work in partnership to implement good practice and meet relevant legislative, research funder and regulatory requirements.
5. Research data and records should be:
 - a. Accurate, complete, authentic and reliable;
 - b. Identifiable, retrievable, and available when needed;
 - c. Secure and safe;
 - d. Kept in a manner that is compliant with legal obligations and, where applicable, the requirements of funding bodies and project-specific protocols approved under the University Policy on the Ethical Conduct of Research Involving Human Participants and Personal Data⁵
 - e. Able to be made available to others in line with appropriate ethical, data sharing and open access principles.
6. Research data and records should be retained for as long as they are of continuing value to the researcher and the wider research community, and as long as specified by research funder, patent law, legislative and other regulatory requirements. The minimum retention period for research data and records is three (3) years after publication or public release of the work of the research. In many instances, researchers will resolve to retain research data and records for a longer period than the minimum requirement.

¹ Research data and records are defined as the recorded information (regardless of the form or the media in which they may exist) necessary to support or validate a research project's observations, findings or outputs.

² Research is defined as per the Frascati manual, i.e. creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications.

³ An overview of the major research funders' data policies is available at <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>. See also the RCUK Common Principles on Data Policy (2011) at <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>

⁴ Researchers are defined as members of the University including staff and doctoral students, and those who are not members of the University but who are conducting research on University premises or using University facilities.

⁵ <http://www.admin.ox.ac.uk/curec/policystatement/>

7. Where research is supported by a contract with or a grant to the University that includes specific provisions regarding ownership, retention of and access to data, the provisions of that agreement will take precedence.
8. If research data and records are to be deleted or destroyed, either because the agreed period of retention has expired or for legal or ethical reasons, this should be done so in accordance with all legal, ethical, research funder and collaborator requirements and with particular concern for confidentiality and security.
9. Researchers are responsible for:
 - Managing research data and records in accordance with the principles and requirements in 5-8 above;
 - Developing and documenting clear procedures for the collection, storage, use, re-use, access and retention or destruction of the research data and records associated with their research. This shall include, where appropriate, defining protocols and responsibilities in a joint or multi-institution collaborative research project. This information should be incorporated, where appropriate, in a research data management plan;
 - Planning for the ongoing custodianship (at the University or using third-party services) of their data after the completion of the research or, in the event of their departure or retirement from the University, reaching agreement with the head of department/faculty (or his/her nominee) as to where such data will be located and how this will be stored;
 - Ensuring that any requirements in relation to research data and records management placed on their research by funding bodies or regulatory agencies or under the terms of a research contract with the University are also met.
10. The University is responsible for:
 - Providing access to services and facilities for the storage, backup, deposit and retention of research data and records that allow researchers to meet their requirements under this policy and those of the funders of their research;
 - Providing researchers with access to training, support and advice in research data and records management;
 - Providing the necessary resources to those operational units charged with the provision of these services, facilities and training.
11. The University's Research and Information Sub-Committee, a sub-committee of the University Research Committee, is responsible for guiding the development and updating of this policy.

Relationship with existing policies

12. This policy will operate in conjunction with other University policies such as:
 - Academic Integrity in Research (<http://www.admin.ox.ac.uk/personnel/cops/researchintegrity/>)
 - Policy on the ethical conduct of research involving human participants and personal data (<http://www.admin.ox.ac.uk/curec/policystatement/>)
 - Intellectual property policy (http://www.admin.ox.ac.uk/statutes/790-121.shtml#_Toc28143157)
 - Data protection policy (<http://www.admin.ox.ac.uk/dataprotection/policy/>)
 - Freedom of Information (<http://www.admin.ox.ac.uk/foi/>)
 - Information Security Policy (<http://www.it.ox.ac.uk/infosec/ispolicy/>)

(Source : University of Oxford, 2015)

Annexe 6 : Modèle de Plan de gestion des données¹⁰⁷



UNIL | Université de Lausanne

Ressources informationnelles
et archives

Université de Lausanne – Plan de gestion des données (Data Management Plan)

Document créé le : 02.05.2015

Version actualisée le : 21.05.2015

Nom(s)	<i>Jambé</i>
Prénom(s)	<i>Carmen</i>
Adresse mail	Carmen.Jambe@unil.ch
ID (ex. ORCID / ResearcherID / Scopus Author ID)	[Si aucun, veuillez laisser vide] ex. orcid.org/0000-0001-5399-3948 (http://orcid.org/)
Institution(s)	ex. Haute école de gestion de Genève
Institution(s) partenaire(s)	ex. Université de Lausanne

<u>Projet de recherche</u>	
Date début projet : <i>01.01.2015</i>	Date fin projet : <i>16.06.2015</i>
Titre du projet de recherche <i>ex. Travail de Bachelor HEG-ID 2015 / Gestion des données de recherche à l'UNIL</i>	
Description du projet (résumé) <i>[Nature du projet, questions de recherche, objectifs principaux, contexte et déroulement du projet, ...]</i>	
Mots-clés <i>ex. Sciences de l'information ; Research Data Management, Digital Curation</i>	
Responsable / chercheur principal	<i>Carmen Jambé</i>
Co-chercheur(s)	[Si personne, veuillez laisser vide]
Mode(s) de financement – bailleur(s) de fonds	[ex. UNIL, subsides FNS, financement H2020, ...]

¹⁰⁷ La structure de ce modèle reprend celle du « DDC Checklist for a Data Management Plan, v4.0 » (2013)

Cadre normatif (loi, politiques, directives existantes)	<i>[Lois, politiques et procédures formelles qui encadrent le projet de recherche, règlements spécifiques (ex. partage, sécurité) de l'équipe / unité / département / institution ; éventuelles exigences du bailleur de fonds,...]</i>
--	---

<u>Données de recherche</u>	
Données collectées (sources primaires) et produites	<i>[Références et description des sources primaires utilisées et autres données existantes réutilisées] [ex. nature, types de données (ex. statistiques, qualitatives, texte,...), formats (ex. formats OpenSource, formats d'archivage), contenu, structure, volume (ex. ~50 Go) des données produites]</i>
Modes de capture ou de collecte des données ;	<i>[Méthodologie, processus de recherche ; instruments et manipulations, standards utilisés (ex. DDI), étalonnages, échantillons, mesures,...]</i>
Gestion des dossiers et fichiers ;	<i>[Arborescence classificatoire utilisés (ex. structure des dossiers et fichiers) ; gestion des différentes versions des fichiers]</i>
Traitement et analyse des données	<i>[Anonymisation, contrôle qualité des données, nettoyage, retranscription,...]</i>

<u>Documentation et métadonnées</u>	
Matériel d'accompagnement (documentation) et métadonnées utilisés	<i>[Documentation qui accompagne les données en vue leur réutilisation et réinterprétation future lors d'analyses secondaires par d'autres chercheurs; la documentation contient des informations de base pour comprendre le projet de recherche et retrouver les données, p. ex. : - Contexte du projet de recherche, méthodologie et processus, récolte des données, instruments et manipulations, formats et types de données, standards et variables utilisés (noms, questions, descriptions, algorithmes, syntaxes,...), unités de mesure, hypothèses, procédures d'analyse,...] Dans la mesure du possible, utiliser des métadonnées descriptives normalisées, structurées et standardisées (ex. DDI)]</i>

<u>Éthique, déontologie et respect du cadre légal</u>	
Éthique et déontologie	<i>[Formulaires de consentement, anonymisation des données, conservation / destruction / partage / protection / sécurité des données sensibles,...] [Règles éthiques (ex. Commission éthique) et codes déontologiques utilisés]</i>
Droit d'auteur et propriété intellectuelle	<i>[Propriété des données, conditions d'accès et de réutilisation des données (définition des droits / restrictions), diffusion, partage et publication des données (ex. licences, Open Access,...), brevets, inventions,...]</i>

<u>Stockage, sauvegardes et sécurité</u>	
Stockage	<i>[Capacités / volume / lieux / sécurité / responsabilités / ... du stockage]</i>

	<i>[Le stockage se fait de préférence sur les serveurs de l'UNIL]</i>
Sauvegardes	<i>[Modes / lieux / fréquence / automatisation / sécurité / mesures en cas d'incident (perte des sauvegardes) / responsabilités / ...des sauvegardes]</i>
Sécurité	<i>[Politique de sécurité, évaluation du risque, sécurité des accès, du partage, du stockage, des transferts, mesures spécifiques pour la protection des données sensibles,...]</i>

<u>Sélection et préservation à long terme</u>	
Sélection des données qui seront conservées / partagées et/ou préservées à la fin du projet	<i>[Sélection des données à conserver, partager, préserver ; obligations légales ou contractuelles pour l'élimination et la destruction des données ; futures réutilisations prévisibles des données (ex. peer-review, nouvelles recherches,...) ; durée de conservation à long terme (ex. 5-10 ans, illimitée,...)]</i>
Mesures et plans pour l'archivage, la conservation, le dépôt des données et la préservation à long terme	<i>[Choix du dépôt (repository) de données (ex. dépôt institutionnel, FORS, dépôt international,...) et contrats de dépôt ; coûts éventuels, charges, investissement en temps pour l'archivage dans les dépôts, préparation des données avant leur archivage et leur préservation (ex. migration des formats, descriptions des sets de données,...)]</i>

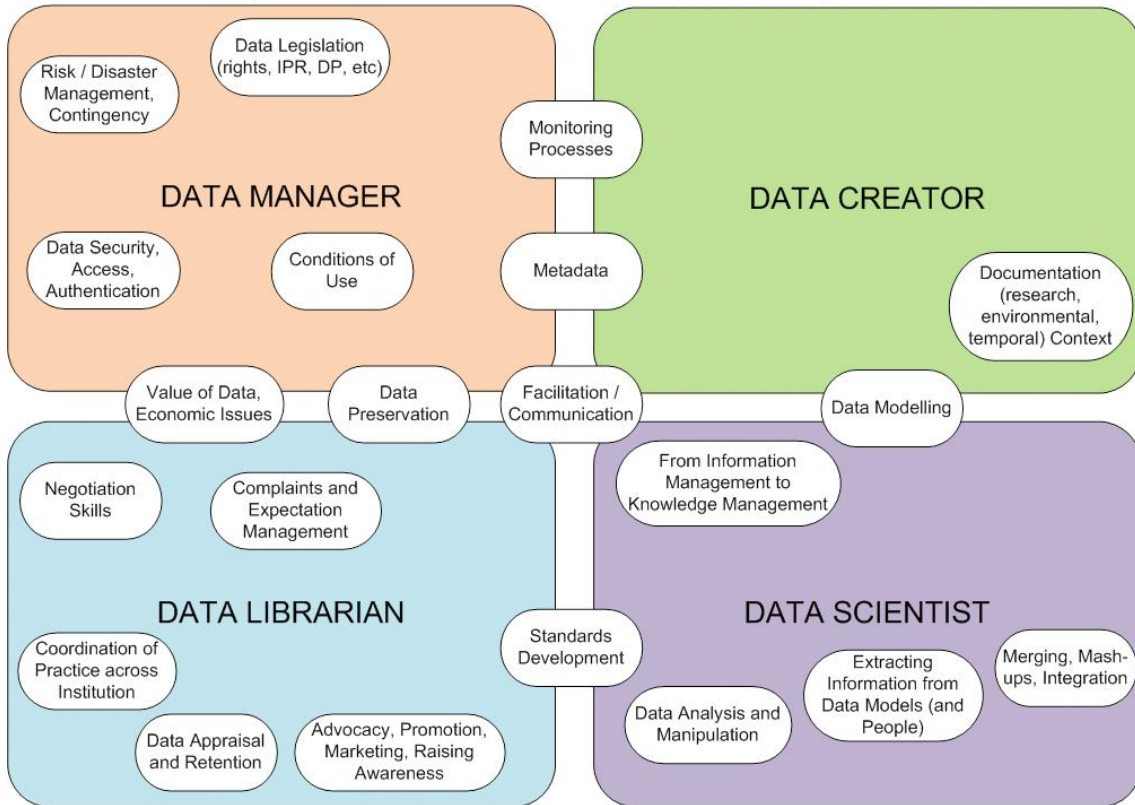
<u>Partage des données</u>	
Moyens utilisés pour partager les données	<i>[Avec qui et sous quelles conditions ? Accessibilité, mise à disposition, dépôts (repositories), visibilité, périodes de rétention (avant publication), publication (Open Access ou non), localisation des données et identification (ex. identifiants uniques et persistants, DOIs,...), valorisation des données,...]</i>
Éventuelles restrictions pour le partage des données	<i>[Restrictions dues aux obligations légales ou contractuelles, durée de restrictions, contrats pour le partage ou la non-divulgateion,...]</i>

<u>Responsabilités et ressources</u>	
Responsabilités pour la gestion des données	<i>[ex. : élaboration / relecture / révision du DMP, responsabilités à chacune des étapes (ex. collecte, création, analyse, traitement, stockage, sauvegarde, sécurité, accès, mise à disposition, partage, archivage, dépôt, préservation, réutilisation...), responsabilités au sein de l'équipe, lors de collaborations, des autres chercheurs, contrats avec des partenaires, propriété des données,...]</i>
Ressources à disposition pour la gestion des données	<i>[Moyens humains (membres de l'équipe), interventions de spécialistes externes (ex. archivistes, spécialistes de l'information, informaticiens / IT, juristes, ...), coûts financiers, ressources matérielles (infrastructure technique et informatique),...] <i>[Financements particuliers pour la gestion des données de recherche ? Prise en charge institutionnelle ou non ?]</i></i>

(Source : DCC, 2013)

Annexe 7 : Core skills for Data Management

A follow-up from the second DCC Research Data Management Forum (November 2008)



(Source : Donnelly, 2008)

Annexe 8 : Summary Table of Roles, Rights, Responsibilities and Relationships

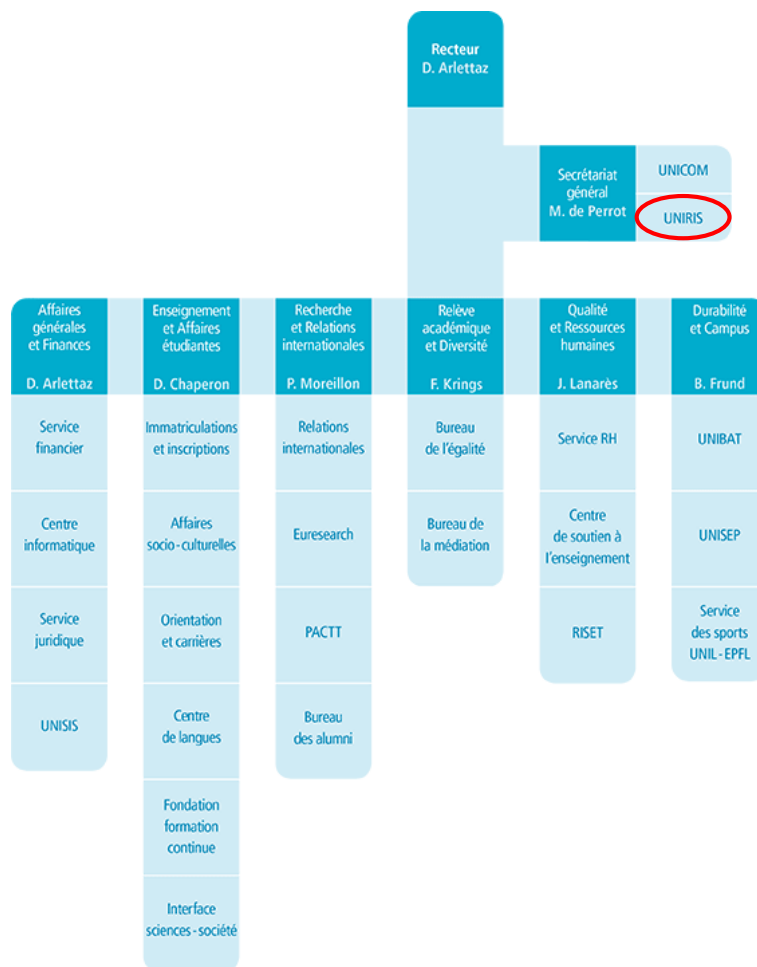
DEALING WITH DATA

Role	Rights	Responsibilities	Relationships
<i>Scientist</i> : creation and use of data	Of first use. To be acknowledged. To expect IPR to be honoured. To receive data training and advice.	Manage data for life of project. Meet standards for good practice. Comply with funder / institutional data policies and respect IPR of others. Work up data for use by others.	With institution as employee. With subject community With data centre. With funder of work.
<i>Institution</i> : curation of and access to data	To be offered a copy of data.	Set internal data management policy. Manage data in the short term. Meet standards for good practice. Provide training and advice to support scientists. Promote the repository service.	With scientist as employer. With data centre through expert staff.
<i>Data centre</i> : curation of and access to data	To be offered a copy of data. To select data of long-term value.	Manage data for the long-term. Meet standards for good practice. Provide training for deposit. Promote the repository service. Protect rights of data contributors. Provide tools for re-use of data.	With scientist as "client" With user communities. With institution through expert staff. With funder of service.
<i>User</i> : use of 3 rd party data	To re-use data (non-exclusive licence). To access quality metadata to inform usability.	Abide by licence conditions. Acknowledge data creators / curators. Manage derived data effectively.	With data centre as supplier. With institution as supplier.
<i>Funder</i> : set/react to public policy drivers	To implement data policies. To require those they fund to meet policy obligations.	Consider wider public-policy perspective & stakeholder needs. Participate in strategy co-ordination. Develop policies with stakeholders. Participate in policy co-ordination, joint planning & fund service delivery. Monitor and enforce data policies. Resource post-project long-term data management. Act as advocate for data curation & fund expert advisory service(s). Support workforce capacity development of data curators.	With scientist as funder. With institution. With data centre as funder. With other funders. With other stakeholders as policy-maker and funder of services.
<i>Publisher</i> : maintain integrity of the scientific record	To expect data are available to support publication. To request pre-publication data deposit in long-term repository.	Engage stakeholders in development of publication standards. Link to data to support publication standards. Monitor & enforce public. standards.	With scientist as creator, author and reader. With data centres and institutions as suppliers.

56

(Source : Lyon, 2007)

Annexe 9 : Organigramme de la Direction UNIL



(Source : UNIL, 2015)

Annexe 10 : *Manifesto on Data access and research transparency (DART) in Switzerland*

Workshop "Improving Data Access and Research Transparency (DART) in Switzerland" - 7 November 2014

Manifesto to advance Data Access and Research Transparency (DART) in Switzerland

Manifesto presented at the workshop November 7 2014 in Bern, Switzerland.

Acknowledging that,

- information and data sharing is essential to advance knowledge for the benefit of society;
- academic research is based on the premise that all procedures relevant for evidence-based conclusions are transparent and can potentially be verified and replicated;
- research transparency and data access strengthen the credibility of published research results;
- data collections that are publicly funded should also be publicly available;
- making data accessible is beneficial for the visibility of the projects and individuals involved in data collection;
- the confidentiality and anonymity of information provided by individuals who participate in research projects must be ensured;

we encourage the various relevant stakeholders and actors to take action to improve data access and research transparency in Switzerland, with respect to the following principles.

Individual researchers,

- who collect or generate data have the right to use those data first;
- have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency and analytic transparency, so that their work and conclusions can be tested or replicated;
- should provide access to the data and other replication material used in their research in a way that ensures accessibility in the long run;
- offer full account of the procedures used to collect or generate the data without any unjustified delays or embargos in compliance with relevant and applicable laws, including copyright laws, with the exception of privacy and confidentiality concerns specified in relevant human research regulations;

Universities, departments, and faculties,

- should promote and reward data access and research transparency;
- should make sure that students and researchers understand the importance of data access and research transparency and train them in the procedures and standards how to ensure transparency of the entire research process and how to make data and replication material available;
- should make it clear to both faculty and students that misconduct in research will be punished.

Research funding agencies, in particular the Swiss National Science Foundation,

- should ensure that data collected with public funds are made available without unjustified delays and embargos to other researchers for secondary analysis;
- will look into possibilities to make the necessary resources available so that researchers can properly document and archive data according to best practices and international standards, including resources for data management training.

Data archives,

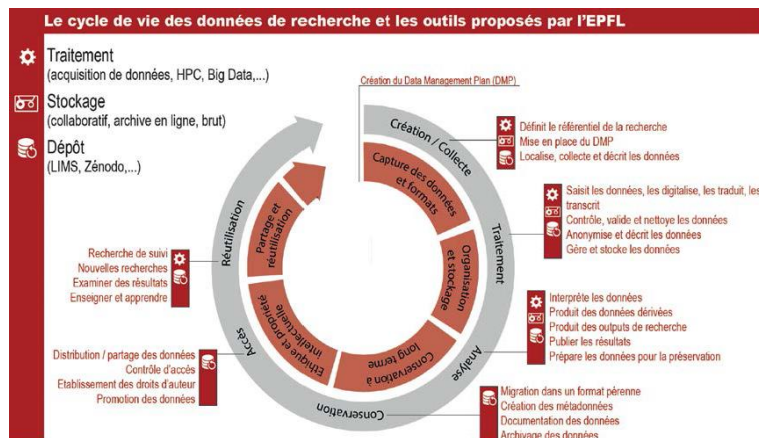
- should ensure that the necessary procedures and tools are in place to deposit, store, distribute and cite data;
- should take all necessary measures to protect the confidentiality of study respondents, including sufficient anonymisation of data and the imposition of contractual obligations for data users;
- should promote data discovery and visibility by employing lasting references to stored files (persistent identifiers such as DOIs)" and by providing users with standardised data citations for each data set distributed.

Editors of scientific journals and publishing houses,

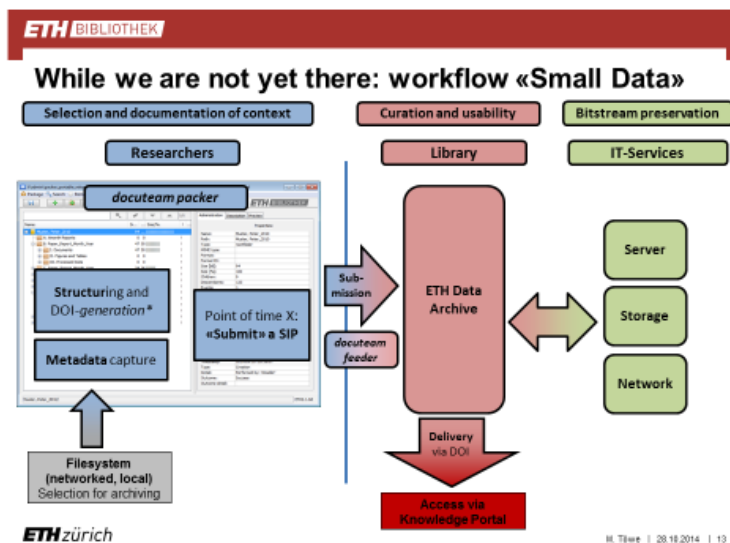
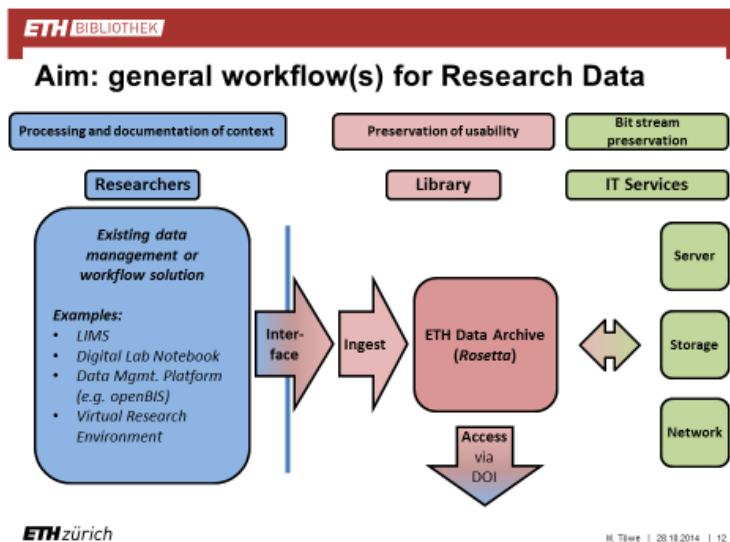
- should require that data used in publications are properly cited and included in lists of references to ensure that data collectors are properly acknowledged;
- should require that data used in publications in their journals, as well as corresponding replication material, are available for secondary analysis, either through their own websites, through data archives or the institutional or personal websites of their authors.

(Source : FORS, DART, 2014)

Annexe 11 : Existants à l'EPFL et l'ETHZ



(Source : EPFL, 2015)



(Source : Töwe, 2014)

Annexe 12 : Questionnaire « Gestion des données de recherche à l'UNIL »



UNIL | Université de Lausanne

| le savoir vivant |

UNIRIS - Service des ressources informationnelles et archives

Gestion des données de recherche à l'Université de Lausanne

Préambule

Nous vous remercions de participer à cette enquête du Service des ressources informationnelles et archives (UNIRIS) de l'UNIL. D'une durée de 15-20 minutes, cette enquête s'adresse à l'ensemble de la communauté **active** dans la recherche à l'UNIL.

Notre thématique est celle de la « gestion des données de recherche », à savoir toutes les stratégies et les actions réalisées à chacune des grandes étapes de : création, traitement, analyse, conservation, accès, réutilisation des données issues de la recherche. Nous cherchons à comprendre :

Quelles données de recherche vous collectez et produisez

La manière dont vous gérez ces données, avant-pendant-après un projet de recherche

Vos pratiques en matière de : stockage, sauvegarde, sécurité, accès, publication, archivage, préservation et réutilisation des données de recherche

Vos attentes, besoins et éventuels problèmes rencontrés

L'objectif de cette démarche est d'établir un guide d'aide à la gestion des données de recherche à l'UNIL.

Notion préalable : la donnée de recherche

Dans le cadre de cette enquête, la « donnée de recherche » est définie comme l'« Ensemble des informations et matériaux - quels que soient leur nature et leur support - observés, collectés et créés par les chercheurs-euses à des fins d'analyse, dans le but de produire des résultats de recherche originaux ».

Aide pour répondre aux questions :

Les questions obligatoires sont signalées par *. Lorsque vous n'êtes pas concerné-e par une question qui débute par "Si vous...", **ne** répondez **pas** et passez directement aux questions suivantes.

Confidentialité des données :

Les cookies et données conservées par votre navigateur web ne sont pas utilisés dans cette enquête. Les réponses seront traitées de manière anonyme. Elles ne pourront en aucun cas permettre l'identification des participants-es.

Le Service UNIRIS est responsable de la récolte, de l'exploitation et de la préservation des données issues de l'enquête en conformité

avec les règles, normes et principes déontologiques accessibles aux liens suivants :

<http://www.unil.ch/uniris/home/menuinst/references/lois-reglements-et-normes.html>

<http://www.unil.ch/uniris/home/menuinst/consulter-les-archives/autorisation-dutilisation.html>

<http://www.unil.ch/uniris/home/menuinst/notre-service/vision-et-deontologie.html>

Les résultats de l'enquête seront présentés par son auteure, Mme Carmen Jambé, dans son travail de Bachelor à la Haute école de gestion Genève, réalisé sur mandat et sous la direction de M. Gérard Bagnoud, responsable du Service UNIRIS.

Il y a 62 questions dans ce questionnaire

Votre projet de recherche

1 Quel type de recherche menez-vous actuellement ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Recherche personnelle
- Recherche UNIL
- Thèse de doctorat
- Recherche Fonds national suisse (FNS)
- Programme européen/UE
- Programme international
- Recherche sur mandat externe
- Autre:

2 Menez-vous votre projet de recherche individuellement ou au sein d'une équipe ? *

Veuillez sélectionner une seule des propositions suivantes :

- Individuellement
- Les deux : à la fois individuellement et au sein d'une équipe
- Au sein d'une équipe

3 Quel est le mode de financement de votre recherche ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Recherche financée par l'UNIL
- Subsidés FNS
- Programme national de recherche (PNR) du FNS
- Pôle de recherche national (PRN) du FNS
- Bourse d'excellence de la Confédération

- Programme européen/UE
- Programme international
- Institution mandante externe
- Fondation(s)
- Moi-même
- Autre:

4 Recevez-vous des financements particuliers pour gérer vos données de recherche (ex. les sauvegarder, publier, conserver,...) ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

5 Votre bailleur de fonds a-t-il des exigences particulières concernant vos données de recherche (ex. leur sécurité, utilisation, diffusion, conservation,...) ? Si oui, veuillez préciser de quelles exigences il s'agit. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Je ne sais pas

Faites le commentaire de votre choix ici :

Vos données de recherche

6 A quelle "grande famille" appartient vos données de recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Données d'observation
- Données d'expérimentation
- Données computationnelles, de modèles ou de simulation
- Données dérivées ou compilées
- Données canoniques ou de référence
- Je ne sais pas

7 Quels types de données collectez-vous et produisez-vous ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Données matérielles et physiques
- Données textuelles (ex. documents, textes)
- Données numériques (ex. tableurs)
- Données géospatiales
- Données statistiques
- Données qualitatives
- Données quantitatives
- Bases de données
- Entretiens, interviews
- Audio
- Vidéo
- Images, photographies
- Numérisations (scans)
- Archives, sources historiques
- 3D, modèles, visualisations
- Codes sources, langages de programmation
- Applications, logiciels, programmes informatiques
- Autre:

8 D'une manière générale, qui est responsable de gérer les données collectées et produites durant votre recherche ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Moi-même
- Assistant-e
- Personne désignée au sein de l'équipe
- Responsable de l'équipe/du projet
- Directeur-trice de thèse
- Centre informatique/informaticiens
- Prestataire externe
- Personne
- Je ne sais pas
- Autre:

9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes... *

Choisissez la réponse appropriée pour chaque élément :

	Tout à fait d'accord	Plutôt d'accord	Ni d'accord ni en désaccord	Plutôt en désaccord	Tout à fait en désaccord
Pour la société en général	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pour l'ensemble de la communauté scientifique	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Surtout dans mon domaine spécifique de recherche	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Surtout pour ma carrière professionnelle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Surtout pour mon projet de recherche	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Elles ont une importance relative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Plan de gestion des données (Data Management Plan)

Un "Plan de gestion des données" (*Data Management Plan*) est un document qui décrit quelles actions seront réalisées pour gérer les données de recherche, p. ex. :

Quelles données seront produites et comment

Comment elles seront décrites, documentées et organisées

Comment elles seront stockées et sécurisées

A quelles conditions et comment elles seront accessibles et partagées

Comment elles seront archivées et préservées sur le long terme

10 Selon vous, planifier la gestion de ses données de recherche est... *

Veuillez sélectionner une seule des propositions suivantes :

- Crucial
- Très important
- Plutôt important
- Plutôt pas important
- Pas du tout important
- Je ne sais pas

11 Avez-vous déjà réalisé un « Plan de gestion des données » (Data Management Plan) ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

12 Si vous avez réalisé un Data Management Plan, était-ce une exigence de votre bailleur de fonds ? Si non, à qui d'autre était-il destiné et pour quelles autres raisons l'avez-vous réalisé ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

Faites le commentaire de votre choix ici :

13 Si vous avez réalisé un DMP, vous a-t-il été utile (ou non) et pour quelles raisons ? Veuillez préciser les raisons. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
 Plutôt oui
 Plutôt non
 Non pas du tout

Faites le commentaire de votre choix ici :

Décrire et documenter votre projet et vos données de recherche

14 Avez-vous créé de la documentation (ex. carnets, guides, journaux, rapports,...) pour décrire votre projet de recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

15 Si vous avez créé de la documentation, quelles principales informations contient-elle ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Contexte de la recherche
- Méthodologie et processus de recherche
- Récolte des données (ex. leur type, nature, contenu, structure,...)
- Instruments et manipulations
- Variables (ex. noms, questions, descriptions, algorithmes, syntaxes,...)
- Autre:

Utiliser les métadonnées

Les métadonnées sont des "*données sur les données*" qui permettent de définir et de décrire certaines caractéristiques des données, p. ex. :

Aspects administratifs des données (ex. droits d'accès et d'utilisation)

Aspects techniques des données (ex. date)

Contenant des données (ex. format de fichier)

Contenu des données (ex. titre, sujet)

Contexte de création des données (ex. auteur, institution)

Structure logique et physique des données (ex. liens entre les fichiers)

16 Utilisez-vous des métadonnées pour documenter et décrire vos données de recherche ? *

Veuillez sélectionner une seule des propositions suivantes :

- Oui
- Non

- Je ne sais pas

17 Si vous utilisez des métadonnées, sont-elles normalisées, structurées ou standardisées ? Si oui, veuillez préciser quels standards vous utilisez. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Je ne sais pas

Faites le commentaire de votre choix ici :

Stocker vos données de recherche

18 Où stockez-vous vos données de recherche ? *Si autre, veuillez préciser.* *

Veillez choisir toutes les réponses qui conviennent :

- Sur papier
- Disque dur - ordinateur/laptop/netbook de l'UNIL
- Disque dur - ordinateur/laptop/netbook personnel
- Disque dur externe
- Stockage en réseau (NAS) personnel
- Serveurs UNIL
- Crashplan UNIL
- DocUNIL
- Espace de travail collaboratif BSCW-UNIL
- Moodle UNIL
- Dépôt/répertoire institutionnel/disciplinaire national (ex. FORSbase) => veuillez préciser dans "autre"
- Dépôt/répertoire institutionnel/disciplinaire international (ex. GenBank) => veuillez préciser dans "autre"
- Dépôt/répertoire général national/international (ex. Figshare, Dryad, Zenodo,...) => veuillez préciser dans "autre"
- USB/Flash
- CD/DVD
- VHS/K7/Disquette
- Email/messagerie
- Prestataire externe privé
- Cloud (ex. Amazon S3, Dropbox, Google Docs, iCloud,...) => veuillez préciser dans "autre"
- Solution GED (gestion électronique des documents) => veuillez préciser dans "autre"
- Je ne sais pas
- Autre:

19 A combien estimez-vous le volume de données que vous collectez et produisez ? *

Veillez sélectionner une seule des propositions suivantes :

- < 1 Go
- 1-50 Go
- 50-100 Go
- 100-500 Go
- 500 Go-1 To
- 1-50 To
- 50-100 To
- > 100 To
- Je ne sais pas

Sauvegarder vos données de recherche

20 Existe-t-il des sauvegardes (back-ups, copies de secours) de vos données de recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui pour toutes mes données
- Oui pour une partie seulement
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire
- Je ne sais pas

21 Si vos données sont sauvegardées, à quelle fréquence le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veillez sélectionner une seule des propositions suivantes :

- Tous les jours
- 1x/semaine
- Plusieurs fois par semaine
- 1x/mois
- Plusieurs fois par mois
- 1x/semestre
- 1x/an
- Fréquence irrégulière
- Je ne sais pas
- Autre

Faites le commentaire de votre choix ici :

22 Si vos données sont sauvegardées, où le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Sur papier
- Disque dur externe
- USB/Flash
- CD/DVD
- Serveur UNIL
- Crashplan UNIL
- Serveur externe
- Cloud (ex. Dropbox, OpenDrive, CrashPlan,...) => veuillez préciser dans "autre"
- Dépôt/répertoire institutionnel/disciplinaire national (ex. FORSbase) => veuillez préciser dans "autre"
- Dépôt/répertoire institutionnel/disciplinaire international (ex. GenBank) => veuillez préciser dans "autre"
- Dépôt/répertoire général national/international (ex. Figshare, Dryad, Zenodo,...) => veuillez préciser dans "autre"
- Je ne sais pas
- Autre:

Sécuriser vos données de recherche

23 Vos données de recherche sont-elles sécurisées (ex. contrôle des accès, cryptage, destruction des données,...) ? *

Veuillez sélectionner une seule des propositions suivantes :

- Oui pour toutes mes données
- Oui pour une partie seulement
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire
- Je ne sais pas

24 Si vos données sont sécurisées, à quel niveau de sécurité le sont-elles ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez sélectionner une seule des propositions suivantes :

- Niveau maximal
- Très haut niveau
- Haut niveau
- Niveau standard
- Niveau minimal

25 Si vos données sont sécurisées, comment le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Contrôle des accès aux labos/locaux/salles
- Utilisation d'un coffre-fort
- Sécurisation du transport des données
- Pare-feu, logiciels de sécurité antivirus et contre le piratage informatique
- Serveurs sécurisés
- Cryptage des données
- Protocoles de transfert sécurisés
- Login/mots de passe pour l'accès
- Destruction des données
- Autre:

Partager et donner accès à vos données de recherche

26 De manière générale, êtes-vous favorable au partage de vos données de recherche ? *

Veuillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
- Plutôt oui
- Avis neutre
- Plutôt non
- Non pas du tout

27 Si vous partagez vos données de recherche, quelles raisons motivent votre choix ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Exigence de mon bailleur de fonds/éditeur/institution/partenaire externe
- Pratique courante au sein de la communauté scientifique
- Réseau, relations et collaborations professionnelles
- Permet d'accéder et de retrouver plus facilement la recherche
- Augmente la visibilité et l'impact du chercheur/UNIL/recherche
- Prouve que la recherche est intègre, responsable et de qualité
- Facilite la vérification, l'évaluation (peer review) et la reproductibilité de la recherche
- Renforce la validité des méthodes et des résultats
- Augmente le potentiel d'innovation
- Responsabilité du chercheur envers la société
- Autre:

28 Si vous NE partagez PAS vos données de recherche, quelles raisons motivent votre choix ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Contrainte du bailleur de fonds/éditeur/institution/partenaire externe
- Mes données ne sont pas suffisamment documentées
- Mes données ne sont pas suffisamment pertinentes
- Mes données sont difficilement compréhensibles ou interprétables

- Pas de temps/argent/ressources pour le faire
- Je ne sais pas comment les partager
- Je ne sais pas où les partager
- Pour protéger l'exclusivité de ma découverte et de mes idées
- Risque de fausse interprétation ou d'utilisation fallacieuse de mes données
- Raisons commerciales et stratégiques (ex. brevets)
- Raisons éthiques, de déontologie, de confidentialité
- Raisons juridiques et contractuelles
- Raisons de sécurité nationale
- Raisons personnelles
- Autre:

29 APRES PUBLICATION de votre recherche, qui peut accéder à vos données de recherche ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Les membres de mon équipe
- Les membres de mon unité/de ma Faculté
- Tous les chercheur-euse-s de l'UNIL
- Les chercheurs d'autres Universités
- Uniquement les spécialistes de la discipline et du domaine de recherche
- Mon bailleur de fonds
- Mon éditeur
- Peer reviewers
- Tout le monde
- Uniquement les personnes qui font la demande
- Personne
- Autre:

Publier vos données de recherche

30 De manière générale, êtes-vous favorable au principe de libre accès des publications scientifiques (Open Access) ? *

Veuillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
- Plutôt oui
- Avis neutre
- Plutôt non
- Non pas du tout

31 Avez-vous déjà publié en Open Access ? Si oui, veuillez préciser où. *

Veuillez sélectionner une seule des propositions suivantes :

- Oui selon le modèle « golden road » (= publication dans une revue OA)
- Oui selon le modèle « green road » (= publication commerciale + dépôt sur serveur OA, ex. SERVAL)
- Oui selon le modèle « hybride » (= le chercheur paie pour publier)
- Oui mais je ne sais pas selon quel modèle
- Non je ne connaissais pas cette possibilité
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire

Faites le commentaire de votre choix ici :

32 Avez-vous déjà annexé à votre article/publication, vos données de recherche dans une revue scientifique ? Si oui, veuillez préciser le nom de la revue. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non je ne connaissais pas cette possibilité
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire

Faites le commentaire de votre choix ici :

33 Si vous avez publié vos données en annexe à votre article/publication, était-ce une exigence de votre éditeur scientifique ? Si oui, veuillez préciser le nom de la revue. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Faites le commentaire de votre choix ici :

Archiver et préserver vos données de recherche

34 Conservez-vous vos données de recherche à la fin d'un projet ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui toutes
- Oui env. 75%
- Oui env. 50%
- Oui env. 25%
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire

35 Combien de temps pensez-vous que vos données de recherche devraient être conservées ? *

Veillez sélectionner une seule des propositions suivantes :

- 1-4 ans
- 5-10 ans
- 11-20 ans
- 21-50 ans
- Ad eternam
- Je ne souhaite pas que mes données soient conservées
- Je ne sais pas

36 De manière générale, êtes-vous favorable au dépôt de vos données de recherche dans un dépôt/répertoire de données ou service d'archives ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
- Plutôt oui
- Avis neutre
- Plutôt non
- Non pas du tout

37 Avez-vous déjà déposé vos données de recherche dans un dépôt/répertoire de données ou service d'archives ? Si oui, veuillez préciser où. *

Veillez sélectionner une seule des propositions suivantes :

- Oui pour toutes mes données
- Oui pour une partie de mes données
- Non je ne connaissais pas cette possibilité
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire

Faites le commentaire de votre choix ici :

38 Si vous avez l'intention de déposer vos données de recherche, à quel endroit souhaiteriez-vous les déposer ? Archive/dépôt/répertoire... Si autre, veuillez préciser le type de dépôt. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- De mon unité de recherche
- De ma Faculté
- Institutionnel UNIL
- Institutionnel/disciplinaire national (ex. FORSbase) => veuillez

préciser dans "autre"

- Institutionnel/disciplinaire international (ex. GenBank) => veuillez préciser dans "autre"
- Général national/international (ex. Dryad, Figshare, Zenodo,...) => veuillez préciser dans "autre"
- Commercial, privé
- Autre:

39 Si vous déposez (ou avez l'intention de déposer) vos données de recherche dans un répertoire/dépôt de données ou service d'archives, quelles raisons motivent votre choix ? Si autre, veuillez préciser les raisons. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veuillez choisir toutes les réponses qui conviennent :

- Recommandé par mes collègues/spécialistes du domaine
- Renommée du répertoire/dépôt
- Nombre de données déposées sur le dépôt/répertoire
- Conditions générales d'utilisation
- Gestion des accès et des droits
- Gratuité du service
- Capacité du volume de stockage
- Sécurité des données
- Préservation et archivage à long terme
- Impact, citation, visibilité des données déposées
- Réseau, relations et collaborations professionnelles
- Avis/choix personnel
- Autre:

Réutiliser les données de recherche

40 De manière générale, acceptez-vous que d'autres chercheurs réutilisent vos données de recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
- Plutôt oui
- Avis neutre
- Plutôt non
- Non pas du tout

41 Avez-vous déjà réutilisé les données de recherche d'autres chercheurs dans le cadre de votre propre recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui très régulièrement
- Oui quelques fois
- Non mais j'ai l'intention de le faire
- Non et je n'ai pas l'intention de le faire

42 Si vous avez réutilisé les données de recherche d'autres chercheurs, pour quelles raisons les avez-vous réutilisées ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Veillez choisir toutes les réponses qui conviennent :

- Pour les citer dans mon travail
- Pour les comparer à mes propres données
- Pour développer une nouvelle recherche
- Pour évaluer la recherche d'autrui (peer review)
- Pour éviter de refaire à double l'expérience
- Pour gagner du temps
- Pour économiser de l'argent
- Autre:

Questions juridiques, éthiques et déontologiques

43 A qui appartiennent les données que vous créez lors de votre activité de recherche ? Si autre, veuillez préciser. *

Veillez choisir toutes les réponses qui conviennent :

- Moi-même
- Tous les membres de l'équipe de recherche
- Responsable de l'équipe/du projet
- Directeur-trice de thèse
- L'Université de Lausanne
- Institution publique externe
- Partenaire privé externe
- Je ne sais pas
- Autre:

44 Des contrats (ex. de secret, de confidentialité ou de non-divulgence, de transfert de matériel biologique, formulaire de consentement,...) régissent-ils votre projet de recherche ? Si oui, veuillez préciser lesquels. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Faites le commentaire de votre choix ici :

45 Collectez-vous et gérez-vous des données à caractère confidentiel (ex. données personnelles, sensibles, profil de la personnalité, d'identité, médicales,...) ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

46 Dans la gestion de vos données de recherche, vous référez-vous à un code déontologique ou éthique ? Si oui, veuillez préciser de quel code il s'agit. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

Faites le commentaire de votre choix ici :

Aide, soutien et formations

47 Avez-vous déjà suivi des formations sur la gestion des données de recherche ? Si oui, veuillez préciser quel type de formation, où et par qui elle vous a été donnée. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
 Non

Faites le commentaire de votre choix ici :

48 Parmi les différents types de formation ci-dessous, lesquels préférez-vous ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Cours, séminaires intégrés au cursus doctoral
- Ateliers pratiques, workshops, formations en groupe
- Site web dédié à la gestion des données de recherche
- Cours en ligne (e-learning, moocs,...)
- Documentation en ligne (fiches techniques, guides, modes d'emploi, check-lists, tutoriels...)
- Documentation imprimée (fiches techniques, guides, modes d'emploi, check-lists, tutoriels...)
- Aide et conseils personnalisés fournis par un-e spécialiste (ex. archiviste, informaticien, juriste,...)
- Autre:

49 Si vous deviez participer à une formation (ex. cours, séminaires, workshops,...), quels sujets vous intéresseraient le plus ? Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Le b.a.-ba de la gestion des données de recherche
- Data Management Plan
- Normes et standards (formats de fichiers, métadonnées)
- Stockage, sauvegardes (back-up), sécurité

- Infrastructure technique/informatique
- Anonymisation, confidentialité, protection des données
- Citations, publications, partage
- Open Access, Open (Linked) Data
- Archivage, préservation à long terme
- Ethique et déontologie
- Questions juridiques, légales et contractuelles
- Droit d'auteur et propriété intellectuelle
- Autre:

Informations sur la gestion des données de recherche

50 De manière générale, êtes-vous informé-e (via avis officiels, communiqués internes, séances,...) sur la question des données de recherche ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui régulièrement
- De temps en temps
- Rarement
- Jamais

51 Souhaiteriez-vous recevoir plus d'informations sur cette question ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Cela m'est égal

52 Existe-t-il au sein de votre Faculté/unité de recherche/laboratoire des directives, lignes directrices, recommandations internes pour la gestion des données de recherche ? Si oui, veuillez préciser lesquelles. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Je ne sais pas

Faites le commentaire de votre choix ici :

53 Souhaiteriez-vous que l'UNIL développe une politique de gestion des données de recherche pour soutenir les chercheurs ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui tout à fait
- Plutôt oui
- Plutôt non
- Non pas du tout
- Cela m'est égal

Service UNIRIS

54 Connaissez-vous le Service des ressources informationnelles et archives (UNIRIS) de l'UNIL ? *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Seulement le nom
- Non pas du tout

55 Avez-vous déjà sollicité ce service ? Si oui, veuillez préciser pour quelles raisons. *

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Faites le commentaire de votre choix ici :

Votre profil personnel

56 Vous êtes... *

Veillez sélectionner une seule des propositions suivantes :

- Féminin
- Masculin

57 Quel âge avez-vous ? *

Veillez sélectionner une seule des propositions suivantes :

- de 30 ans
- 31-40 ans
- 41-50 ans
- + de 50 ans

58 A quelle Faculté de l'UNIL (ou autre institution) êtes-vous rattaché ? Si "autre institution", veuillez préciser laquelle. *

Veuillez sélectionner une seule des propositions suivantes :

- Faculté de biologie et de médecine
- Faculté de droit, des sciences criminelles et d'administration publique
- Faculté des géosciences et de l'environnement
- Faculté des Hautes Etudes Commerciales (HEC Lausanne)
- Faculté des lettres
- Faculté des sciences sociales et politiques
- Faculté de théologie et de sciences des religions
- Autre institution

Faites le commentaire de votre choix ici :

59 Quel est votre principal domaine de recherche ?
Si autre, veuillez préciser. *

Veuillez choisir toutes les réponses qui conviennent :

- Administration publique
- Arts
- Biologie
- Droit

- Etudes du tourisme
- Géosciences
- Histoire
- Langues
- Médecine
- Philosophie
- Psychologie
- Sciences criminelles
- Sciences de l'environnement
- Sciences des religions
- Sciences du sport et de l'éducation physique
- Sciences économiques
- Sciences humaines
- Science politique
- Sciences sociales
- Théologie
- Autre:

60 Quelle est votre fonction (ou la fonction la plus élevée si vous en avez plusieurs) ? Si autre, veuillez préciser. *

Veillez sélectionner une seule des propositions suivantes :

- Assistant-e boursier-ère
- Assistant-e diplômé-e
- Assistant-e FNS
- Assistant-e fonds externe
- Chargé-e de cours
- Maître d'enseignement et de recherche (suppléant-e ou MER1 ou MER2)
- Maître assistant-e
- Personnel de recherche
- Personnel technique et de laboratoire
- Premier-ère assistant-e
- Premier-ère assistant-e FNS
- Privat-docent
- Prof. assistant-e (PAST ou PAST-PTC ou boursier FNS)
- Prof. associé-e
- Prof. honoraire

- Prof. invité-e
- Prof. ordinaire
- Prof. remplaçant-e
- Prof. titulaire
- Autre

61 Combien d'années d'expérience dans la recherche avez-vous ? *

Veillez sélectionner une seule des propositions suivantes :

- de 5 ans
- 5-10 ans
- 11-20 ans
- + de 20 ans

Vos commentaires, suggestions et propositions d'amélioration

62 Ici vous pouvez nous faire part de vos attentes et besoins spécifiques en matière de gestion de données de recherche, de vos commentaires, suggestions et propositions d'améliorations.

Veuillez écrire votre réponse ici :

Nous vous remercions d'avoir pris le temps de participer à notre enquête !

Pour plus d'informations, vous pouvez contacter le Service UNIRIS, du lundi au vendredi de 9h à 17h, à l'adresse :

*Université de Lausanne
Service des ressources informationnelles et archives
Unicentre - Bureau 214
CH-1015 Lausanne
+41 21 692 20 28
uniris@unil.ch*

Envoyer votre questionnaire.
Merci d'avoir complété ce questionnaire.



Unicentre - CH-1015 Lausanne - Suisse - Tél. +41 21 692 20
28 - Fax +41 21 692 20 35 **swissuniversities**

Annexe 13 : Réponses et statistiques de l'enquête « Gestion des données de recherche à l'UNIL »

Nombre total d'enregistrements pour ce questionnaire : **310**
 Pourcentage du total : **100.00%**
 Réponse **Décompte** **Pourcentage**

Q1 Quel type de recherche menez-vous actuellement ? Si autre, veuillez préciser.

Recherche personnelle (1)	93	30.00%
Recherche UNIL (2)	107	34.52%
Thèse de doctorat (3)	156	50.32%
Recherche Fonds national suisse (FNS) (4)	108	34.84%
Programme européen/UE (5)	20	6.45%
Programme international (6)	18	5.81%
Recherche sur mandat externe (7)	26	8.39%
Autre	3	0.97%

Q2 Menez-vous votre projet de recherche individuellement ou au sein d'une équipe ?

Individuellement (1)	94	30.32%
Les deux : à la fois individuellement et au sein d'une équipe (2)	163	52.58%
Au sein d'une équipe (3)	53	17.10%

Q3 Quel est le mode de financement de votre recherche ? Si autre, veuillez préciser.

Recherche financée par l'UNIL (1)	182	58.71%
Subsides FNS (2)	118	38.06%
Programme national de recherche (PNR) du FNS (7)	29	9.35%
Pôle de recherche national (PRN) du FNS (8)	11	3.55%
Bourse d'excellence de la Confédération (3)	6	1.94%
Programme européen/UE (4)	25	8.06%
Programme international (5)	15	4.84%
Institution mandante externe (6)	28	9.03%
Fondation(s) (9)	35	11.29%
Moi-même (10)	40	12.90%
Autre	22	7.10%

Q4 Recevez-vous des financements particuliers pour gérer vos données de recherche (ex. les sauvegarder, publier, conserver,...) ?

Oui (Y)	22	7.10%
Non (N)	288	92.90%

Q5 Votre bailleur de fonds a-t-il des exigences particulières concernant vos données de recherche (ex. leur sécurité, utilisation, diffusion, conservation,...) ? Si oui, veuillez préciser de quelles exigences il s'agit.

Oui (1)	55	17.74%
Non (2)	158	50.97%
Je ne sais pas (3)	97	31.29%
Commentaires	44	14.19%

Q6 A quelle "grande famille" appartiennent vos données de recherche ?

Données d'observation (1)	125	40.32%
Données d'expérimentation (2)	73	23.55%
Données computationnelles, de modèles ou de simulation (3)	23	7.42%
Données dérivées ou compilées (4)	21	6.77%
Données canoniques ou de référence (5)	20	6.45%
Je ne sais pas (6)	48	15.48%
Sans réponse	0	0.00%

Q7 Quels types de données collectez-vous et produisez-vous ? Si autre, veuillez préciser.

Données matérielles et physiques (1)	60	19.35%
Données textuelles (ex. documents, textes) (2)	200	64.52%
Données numériques (ex. tableurs) (3)	138	44.52%
Données géospatiales (4)	31	10.00%
Données statistiques (5)	129	41.61%
Données qualitatives (6)	138	44.52%
Données quantitatives (7)	167	53.87%
Bases de données (8)	106	34.19%
Entretiens, interviews (9)	93	30.00%
Audio (10)	44	14.19%
Vidéo (11)	39	12.58%

Images, photographies (12)	122	39.35%
Numérisations (scans) (13)	69	22.26%
Archives, sources historiques (14)	70	22.58%
3D, modèles, visualisations (15)	27	8.71%
Codes sources, langages de programmation (16)	42	13.55%
Applications, logiciels, programmes informatiques (17)	29	9.35%
Autre	5	1.61%

Q8 D'une manière générale, qui est responsable de gérer les données collectées et produites durant votre recherche ? Si autre, veuillez préciser.

Moi-même (1)	292	94.19%
Assistant-e (2)	49	15.81%
Personne désignée au sein de l'équipe (3)	36	11.61%
Responsable de l'équipe/du projet (4)	44	14.19%
Directeur-trice de thèse (5)	18	5.81%
Centre informatique/informaticiens (6)	6	1.94%
Prestataire externe (7)	7	2.26%
Personne (8)	4	1.29%
Je ne sais pas (9)	1	0.32%
Autre	2	0.65%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

a) [Pour la société en général]

Tout à fait d'accord (1)	54	17.42%
Plutôt d'accord (2)	128	41.29%
Ni d'accord ni en désaccord (3)	89	28.71%
Plutôt en désaccord (4)	33	10.65%
Tout à fait en désaccord (5)	6	1.94%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

b) [Pour l'ensemble de la communauté scientifique]

Tout à fait d'accord (1)	75	24.19%
Plutôt d'accord (2)	144	46.45%
Ni d'accord ni en désaccord (3)	64	20.65%
Plutôt en désaccord (4)	22	7.10%
Tout à fait en désaccord (5)	5	1.61%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

c) [Surtout dans mon domaine spécifique de recherche]

Tout à fait d'accord (1)	200	64.52%
Plutôt d'accord (2)	91	29.35%
Ni d'accord ni en désaccord (3)	13	4.19%
Plutôt en désaccord (4)	6	1.94%
Tout à fait en désaccord (5)	0	0.00%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

d) [Surtout pour ma carrière professionnelle]

Tout à fait d'accord (1)	98	31.61%
Plutôt d'accord (2)	104	33.55%
Ni d'accord ni en désaccord (3)	74	23.87%
Plutôt en désaccord (4)	25	8.06%
Tout à fait en désaccord (5)	9	2.90%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

e) [Surtout pour mon projet de recherche]

Tout à fait d'accord (1)	215	69.35%
Plutôt d'accord (2)	61	19.68%
Ni d'accord ni en désaccord (3)	22	7.10%
Plutôt en désaccord (4)	11	3.55%
Tout à fait en désaccord (5)	1	0.32%

Q9 Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche ? Mes données de recherche sont importantes...

f) [Elles ont une importance relative]

Tout à fait d'accord (1)	31	10.00%
Plutôt d'accord (2)	71	22.90%

Ni d'accord ni en désaccord (3)	117	37.74%
Plutôt en désaccord (4)	48	15.48%
Tout à fait en désaccord (5)	43	13.87%

Q10 Selon vous, planifier la gestion de ses données de recherche est...

Crucial (1)	74	23.87%
Très important (2)	116	37.42%
Plutôt important (3)	88	28.39%
Plutôt pas important (4)	15	4.84%
Pas du tout important (5)	2	0.65%
Je ne sais pas (6)	15	4.84%

Q11 Avez-vous déjà réalisé un « Plan de gestion des données » (Data Management Plan) ?

Oui (Y)	36	11.61%
Non (N)	274	88.39%

Q12 Si vous avez réalisé un Data Management Plan, était-ce une exigence de votre bailleur de fonds ? Si non, à qui d'autre était-il destiné et pour quelles autres raisons l'avez-vous réalisé ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Oui (1)	13	4.19%
Non (2)	43	13.87%
Commentaires	18	5.81%
Sans réponse	254	81.94%

Q13 Si vous avez réalisé un DMP, vous a-t-il été utile (ou non) et pour quelles raisons ? Veuillez préciser les raisons. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Oui tout à fait (1)	16	5.16%
Plutôt oui (2)	19	6.13%
Plutôt non (3)	2	0.65%
Non pas du tout (4)	2	0.65%
Commentaires	16	5.16%
Sans réponse	271	87.42%

Q14 Avez-vous créé de la documentation (ex. carnets, guides, journaux, rapports,...) pour décrire votre projet de recherche ?

Oui (Y)	253	81.61%
Non (N)	57	18.39%

Q15 Si vous avez créé de la documentation, quelles principales informations contient-elle ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Contexte de la recherche (1)	199	64.19%
Méthodologie et processus de recherche (2)	220	70.97%
Récolte des données (ex. leur type, nature, contenu, structure,...) (3)	194	62.58%
Instruments et manipulations (4)	96	30.97%
Variables (ex. noms, questions, descriptions, algorithmes, syntaxes,...) (5)	109	35.16%
Autre	10	3.23%

Q16 Utilisez-vous des métadonnées pour documenter et décrire vos données de recherche ?

Oui (1)	100	32.26%
Non (2)	156	50.32%
Je ne sais pas (3)	54	17.42%

Q17 Si vous utilisez des métadonnées, sont-elles normalisées, structurées ou standardisées ? Si oui, veuillez préciser quels standards vous utilisez. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Oui (1)	26	8.39%
Non (2)	54	17.42%
Je ne sais pas (3)	32	10.32%
Commentaires	17	5.48%
Sans réponse	198	63.87%

Q18 Où stockez-vous vos données de recherche ? Si autre, veuillez préciser.

Sur papier (21)	160	51.61%
Disque dur - ordinateur/laptop/netbook de l'UNIL (1)	235	75.81%
Disque dur - ordinateur/laptop/netbook personnel (2)	192	61.94%
Dique dur externe (5)	199	64.19%
Stockage en réseau (NAS) personnel (6)	46	14.84%
Serveurs UNIL (7)	108	34.84%
Crashplan UNIL (23)	79	25.48%
DocUNIL (8)	12	3.87%
Espace de travail collaboratif BSCW-UNIL (9)	3	0.97%

Moodle UNIL (10)	4	1.29%
Dépôt/répertoire institutionnel/disciplinaire national (ex. FORSbase) (11)	5	1.61%
Dépôt/répertoire institutionnel/disciplinaire international (ex. GenBank) (12)	7	2.26%
Dépôt/répertoire général national/international (ex. Figshare, Dryad, USB/Flash (14)	5	1.61%
CD/DVD (15)	71	22.90%
VHS/K7/Disquette (16)	20	6.45%
Email/messagerie (17)	1	0.32%
Prestataire externe privé (18)	64	20.65%
Cloud (ex. Amazon S3, Dropbox, Google Docs, iCloud,...) (19)	6	1.94%
Solution GED (gestion électronique des documents) (20)	106	34.19%
Je ne sais pas (22)	0	0.00%
Autre	1	0.32%
	73	23.55%

Q19 A combien estimez-vous le volume de données que vous collectez et produisez ?

< 1 Go (1)	32	10.32%
1-50 Go (2)	91	29.35%
50-100 Go (3)	31	10.00%
100-500 Go (4)	35	11.29%
500 Go-1 To (5)	18	5.81%
1-50 To (6)	27	8.71%
50-100 To (7)	1	0.32%
< 100 To (8)	0	0.00%
Je ne sais pas (9)	75	24.19%

Q20 Existe-t-il des sauvegardes (back-ups, copies de secours) de vos données de recherche ?

Oui pour toutes mes données (1)	222	71.61%
Oui pour une partie seulement (2)	73	23.55%
Non mais j'ai l'intention de le faire (3)	6	1.94%
Non et je n'ai pas l'intention de le faire (4)	1	0.32%
Je ne sais pas (5)	8	2.58%

Q21 Si vos données sont sauvegardées, à quelle fréquence le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Tous les jours (1)	105	33.87%
1x/semaine (2)	52	16.77%
Plusieurs fois par semaine (3)	21	6.77%
1x/mois (4)	34	10.97%
Plusieurs fois par mois (5)	16	5.16%
1x/semestre (6)	7	2.26%
1x/an (8)	1	0.32%
Fréquence irrégulière (9)	40	12.90%
Je ne sais pas (10)	16	5.16%
Autre (11)	5	1.61%
Commentaires	23	7.42%
Sans réponse	13	4.19%

Q22 Si vos données sont sauvegardées, où le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Sur papier (10)	62	20.00%
Disque dur externe (1)	202	65.16%
USB/Flash (2)	64	20.65%
CD/DVD (3)	11	3.55%
Serveur UNIL (4)	86	27.74%
Crashplan UNIL (11)	78	25.16%
Serveur externe (5)	20	6.45%
Cloud (ex. Dropbox, OpenDrive, CrashPlan,...) (6)	86	27.74%
Dépôt/répertoire institutionnel/disciplinaire national (ex. FORSbase) (7)	2	0.65%
Dépôt/répertoire institutionnel/disciplinaire international (ex. GenBank) (8)	1	0.32%
Dépôt/répertoire général national/international (ex. Figshare, Dryad, Je ne sais pas (12)	2	0.65%
Autre	4	1.29%
	53	17.10%

Q23 Vos données de recherche sont-elles sécurisées (ex. contrôle des accès, cryptage, destruction des données,...) ?

Oui pour toutes mes données (1)	42	13.55%
Oui pour une partie seulement (2)	77	24.84%
Non mais j'ai l'intention de le faire (3)	27	8.71%
Non et je n'ai pas l'intention de le faire (4)	132	42.58%
Je ne sais pas (5)	32	10.32%

Q24 Si vos données sont sécurisées, à quel niveau de sécurité le sont-elles ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Niveau maximal (1)	7	2.26%
Très haut niveau (2)	8	2.58%
Haut niveau (3)	25	8.06%
Niveau standard (4)	61	19.68%
Niveau minimal (5)	18	5.81%
Sans réponse	191	61.61%

Q25 Si vos données sont sécurisées, comment le sont-elles ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Contrôle des accès aux labos/locaux/salles (1)	39	12.58%
Utilisation d'un coffre-fort (2)	1	0.32%
Sécurisation du transport des données (3)	4	1.29%
Pare-feu, logiciels de sécurité antivirus et contre le piratage informatique (4)	44	14.19%
Serveurs sécurisés (5)	55	17.74%
Cryptage des données (6)	13	4.19%
Protocoles de transfert sécurisés (7)	8	2.58%
Login/mots de passe pour l'accès (8)	83	26.77%
Destruction des données (9)	3	0.97%
Autre	4	1.29%

Q26 De manière générale, êtes-vous favorable au partage de vos données de recherche ?

Oui tout à fait (1)	88	28.39%
Plutôt oui (2)	117	37.74%
Avis neutre (3)	52	16.77%
Plutôt non (4)	39	12.58%
Non pas du tout (5)	14	4.52%

Q27 Si vous partagez vos données de recherche, quelles raisons motivent votre choix ? Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Exigence de mon bailleur de fonds/éditeur/institution/partenaire externe (1)	51	16.45%
Pratique courante au sein de la communauté scientifique (2)	122	39.35%
Réseau, relations et collaborations professionnelles (3)	159	51.29%
Permet d'accéder et de retrouver plus facilement la recherche (4)	56	18.06%
Augmente la visibilité et l'impact du chercheur/UNIL/recherche (5)	109	35.16%
Prouve que la recherche est intègre, responsable et de qualité (6)	113	36.45%
Facilite la vérification, l'évaluation (peer review) et la reproductibilité de la	100	32.26%
Renforce la validité des méthodes et des résultats (8)	90	29.03%
Augmente le potentiel d'innovation (9)	48	15.48%
Responsabilité du chercheur envers la société (10)	89	28.71%
Autre	8	2.58%

Q28 Si vous NE partagez PAS vos données de recherche, quelles raisons motivent votre choix ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Contrainte du bailleur de fonds/éditeur/institution/partenaire externe (1)	19	6.13%
Mes données ne sont pas suffisamment documentées (2)	19	6.13%
Mes données ne sont pas suffisamment pertinentes (3)	6	1.94%
Mes données sont difficilement compréhensibles ou interprétables (4)	19	6.13%
Pas de temps/argent/ressources pour le faire (5)	27	8.71%
Je ne sais pas comment les partager (6)	17	5.48%
Je ne sais pas où les partager (7)	19	6.13%
Pour protéger l'exclusivité de ma découverte et de mes idées (8)	53	17.10%
Risque de fausse interprétation ou d'utilisation fallacieuse de mes données	29	9.35%
Raisons commerciales et stratégiques (ex. brevets) (10)	8	2.58%
Raisons éthiques, de déontologie, de confidentialité (11)	45	14.52%
Raisons juridiques et contractuelles (12)	14	4.52%
Raisons de sécurité nationale (13)	1	0.32%
Raisons personnelles (14)	4	1.29%
Autre	7	2.26%

Q29 APRES PUBLICATION de votre recherche, qui peut accéder à vos données de recherche ? Si autre, veuillez préciser.

Les membres de mon équipe (1)	160	51.61%
Les membres de mon unité/de ma Faculté (2)	54	17.42%
Tous les chercheur-euse-s de l'UNIL (3)	40	12.90%
Les chercheurs d'autres Universités (4)	42	13.55%
Uniquement les spécialistes de la discipline et du domaine de recherche (5)	19	6.13%
Mon bailleur de fonds (6)	28	9.03%
Mon éditeur (7)	22	7.10%

Peer reviewers (8)	32	10.32%
Tout le monde (9)	92	29.68%
Uniquement les personnes qui font la demande (10)	102	32.90%
Personne (11)	15	4.84%
Autre	18	5.81%

Q30 De manière générale, êtes-vous favorable au principe de libre accès des publications scientifiques (Open Access) ?

Oui tout à fait (1)	164	52.90%
Plutôt oui (2)	104	33.55%
Avis neutre (3)	22	7.10%
Plutôt non (4)	17	5.48%
Non pas du tout (5)	3	0.97%

Q31 Avez-vous déjà publié en Open Access ? Si oui, veuillez préciser où.

Oui selon le modèle « golden road » (= publication dans une revue OA) (1)	46	14.84%
Oui selon le modèle « green road » (= publication commerciale + dépôt sur	27	8.71%
Oui selon le modèle « hybride » (= le chercheur paie pour publier) (3)	22	7.10%
Oui mais je ne sais pas selon quel modèle (4)	47	15.16%
Non je ne connaissais pas cette possibilité (5)	33	10.65%
Non mais j'ai l'intention de le faire (6)	111	35.81%
Non et je n'ai pas l'intention de le faire (7)	24	7.74%
Commentaires	33	10.65%

Q32 Avez-vous déjà annexé à votre article/publication, vos données de recherche dans une revue scientifique ? Si oui, veuillez préciser le nom de la revue.

Oui (1)	48	15.48%
Non je ne connaissais pas cette possibilité (2)	96	30.97%
Non mais j'ai l'intention de le faire (3)	88	28.39%
Non et je n'ai pas l'intention de le faire (4)	78	25.16%
Commentaires	32	10.32%

Q33 Si vous avez publié vos données en annexe à votre article/publication, était-ce une exigence de votre éditeur scientifique ? Si oui, veuillez préciser le nom de la revue. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Oui (1)	15	4.84%
Non (2)	43	13.87%
Commentaires	8	2.58%
Sans réponse	252	81.29%

Q34 Conservez-vous vos données de recherche à la fin d'un projet ?

Oui toutes (1)	237	76.45%
Oui env. 75% (2)	42	13.55%
Oui env. 50% (3)	14	4.52%
Oui env. 25% (4)	6	1.94%
Non mais j'ai l'intention de le faire (5)	8	2.58%
Non et je n'ai pas l'intention de le faire (6)	3	0.97%

Q35 Combien de temps pensez-vous que vos données de recherche devraient être conservées ?

1-4 ans (1)	16	5.16%
5-10 ans (2)	80	25.81%
11-20 ans (3)	44	14.19%
21-50 ans (4)	30	9.68%
Ad eternam (5)	91	29.35%
Je ne souhaite pas que mes données soient conservées (6)	2	0.65%
Je ne sais pas (7)	47	15.16%

Q36 De manière générale, êtes-vous favorable au dépôt de vos données de recherche dans un dépôt/répertoire de données ou service d'archives ?

Oui tout à fait (1)	114	36.77%
Plutôt oui (2)	98	31.61%
Avis neutre (3)	68	21.94%
Plutôt non (4)	25	8.06%
Non pas du tout (5)	5	1.61%

Q37 Avez-vous déjà déposé vos données de recherche dans un dépôt/répertoire de données ou service d'archives ? Si oui, veuillez préciser où.

Oui pour toutes mes données (1)	3	0.97%
Oui pour une partie de mes données (2)	43	13.87%
Non je ne connaissais pas cette possibilité (3)	141	45.48%

Non mais j'ai l'intention de le faire (4)	58	18.71%
Non et je n'ai pas l'intention de le faire (5)	65	20.97%
Commentaires	29	9.35%

Q38 Si vous avez l'intention de déposer vos données de recherche, à quel endroit souhaiteriez-vous les déposer ? Archive/dépôt/répertoire... Si autre, veuillez préciser le type de dépôt. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

De mon unité de recherche (1)	73	23.55%
De ma Faculté (2)	47	15.16%
Institutionnel UNIL (3)	77	24.84%
Institutionnel/disciplinaire national (ex. FORSbase) => veuillez préciser	32	10.32%
Institutionnel/disciplinaire international (ex. GenBank) => veuillez préciser	23	7.42%
Général national/international (ex. Dryad, Figshare, Zenodo,...) => veuillez	17	5.48%
Commercial, privé (7)	4	1.29%
Autre	24	7.74%

Q39 Si vous déposez (ou avez l'intention de déposer) vos données de recherche dans un répertoire/dépôt de données ou service d'archives, quelles raisons motivent votre choix ? Si autre, veuillez préciser les raisons. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Recommandé par mes collègues/spécialistes du domaine (1)	41	13.23%
Renommée du répertoire/dépôt (2)	29	9.35%
Nombre de données déposées sur le dépôt/répertoire (3)	15	4.84%
Conditions générales d'utilisation (4)	42	13.55%
Gestion des accès et des droits (5)	62	20.00%
Gratuité du service (6)	79	25.48%
Capacité du volume de stockage (7)	65	20.97%
Sécurité des données (8)	80	25.81%
Préservation et archivage à long terme (9)	110	35.48%
Impact, citation, visibilité des données déposées (10)	58	18.71%
Réseau, relations et collaborations professionnelles (11)	60	19.35%
Avis/choix personnel (12)	15	4.84%
Autre	5	1.61%

Q40 De manière générale, acceptez-vous que d'autres chercheurs réutilisent vos données de recherche ?

Oui tout à fait (1)	93	30.00%
Plutôt oui (2)	123	39.68%
Avis neutre (3)	48	15.48%
Plutôt non (4)	34	10.97%
Non pas du tout (5)	12	3.87%

Q41 Avez-vous déjà réutilisé les données de recherche d'autres chercheurs dans le cadre de votre propre recherche ?

Oui très régulièrement (1)	39	12.58%
Oui quelques fois (2)	142	45.81%
Non mais j'ai l'intention de le faire (3)	54	17.42%
Non et je n'ai pas l'intention de le faire (4)	75	24.19%

Q42 Si vous avez réutilisé les données de recherche d'autres chercheurs, pour quelles raisons les avez-vous réutilisées ? Si autre, veuillez préciser. Si vous n'êtes pas concerné-e, veuillez ne pas répondre et passer à la question suivante.

Pour les citer dans mon travail (1)	120	38.71%
Pour les comparer à mes propres données (2)	121	39.03%
Pour développer une nouvelle recherche (3)	97	31.29%
Pour évaluer la recherche d'autrui (peer review) (4)	24	7.74%
Pour éviter de refaire à double l'expérience (5)	44	14.19%
Pour gagner du temps (6)	32	10.32%
Pour économiser de l'argent (7)	11	3.55%
Autre	11	3.55%

Q43 A qui appartiennent les données que vous créez lors de votre activité de recherche ? Si autre, veuillez

Moi-même (1)	177	57.10%
Tous les membres de l'équipe de recherche (2)	84	27.10%
Responsable de l'équipe/du projet (3)	46	14.84%
Directeur-trice de thèse (4)	38	12.26%
L'Université de Lausanne (5)	78	25.16%
Institution publique externe (6)	26	8.39%
Partenaire privé externe (7)	12	3.87%
Je ne sais pas (8)	69	22.26%
Autre	10	3.23%

Q44 Des contrats (ex. de secret, de confidentialité ou de non-divulgence, de transfert de matériel biologique, formulaire de consentement,...) régissent-ils votre projet de recherche ? Si oui, veuillez préciser lesquels.

Oui (1)	105	33.87%
Non (2)	205	66.13%
Commentaires	64	20.65%

Q45 Collectez-vous et gérez-vous des données à caractère confidentiel (ex. données personnelles, sensibles, profil de la personnalité, d'identité, médicales,...) ?

Oui (Y)	122	39.35%
Non (N)	188	60.65%

Q46 Dans la gestion de vos données de recherche, vous référez-vous à un code déontologique ou éthique ? Si oui, veuillez préciser de quel code il s'agit.

Oui (1)	106	34.19%
Non (2)	204	65.81%
Commentaires	69	22.26%

Q47 Avez-vous déjà suivi des formations sur la gestion des données de recherche ? Si oui, veuillez préciser quel type de formation, où et par qui elle vous a été donnée.

Oui (1)	32	10.32%
Non (2)	278	89.68%
Commentaires	19	6.13%

Q48 Parmi les différents types de formation ci-dessous, lesquels préférez-vous ? Si autre, veuillez préciser.

Cours, séminaires intégrés au cursus doctoral (1)	128	41.29%
Ateliers pratiques, workshops, formations en groupe (2)	169	54.52%
Site web dédié à la gestion des données de recherche (3)	102	32.90%
Cours en ligne (e-learning, moocs,...) (4)	65	20.97%
Documentation en ligne (fiches techniques, guides, modes d'emploi, check-	143	46.13%
Documentation imprimée (fiches techniques, guides, modes d'emploi, check-	35	11.29%
Aide et conseils personnalisés fournis par un-e spécialiste (ex. archiviste,	96	30.97%
Autre	7	2.26%

Q49 Si vous deviez participer à une formation (ex. cours, séminaires, workshops,...), quels sujets vous intéresseraient le plus ? Si autre, veuillez préciser.

Le b.a.-ba de la gestion des données de recherche (1)	151	48.71%
Data Management Plan (2)	135	43.55%
Normes et standards (formats de fichiers, métadonnées) (3)	88	28.39%
Stockage, sauvegardes (back-up), sécurité (4)	151	48.71%
Infrastructure technique/informatique (5)	55	17.74%
Anonymisation, confidentialité, protection des données (6)	95	30.65%
Citations, publications, partage (7)	83	26.77%
Open Access, Open (Linked) Data (8)	111	35.81%
Archivage, préservation à long terme (9)	119	38.39%
Ethique et déontologie (10)	58	18.71%
Questions juridiques, légales et contractuelles (11)	76	24.52%
Droit d'auteur et propriété intellectuelle (12)	99	31.94%
Autre	11	3.55%

Q50 De manière générale, êtes-vous informé-e (via avis officiels, communiqués internes, séances,...) sur la question des données de recherche ?

Oui régulièrement (1)	17	5.48%
De temps en temps (2)	75	24.19%
Rarement (3)	145	46.77%
Jamais (4)	73	23.55%

Q51 Souhaiteriez-vous recevoir plus d'informations sur cette question ?

Oui (1)	168	54.19%
Non (2)	58	18.71%
Cela m'est égal (3)	84	27.10%

Q52 Existe-t-il au sein de votre Faculté/unité de recherche/laboratoire des directives, lignes directrices, recommandations internes pour la gestion des données de recherche ? Si oui, veuillez préciser lesquelles.

Oui (1)	48	15.48%
Non (2)	206	66.45%
Je ne sais pas (3)	56	18.06%
Commentaires	33	10.65%

Q53 Souhaiteriez-vous que l'UNIL développe une politique de gestion des données de recherche pour

Oui tout à fait (1)	72	23.23%
---------------------	----	--------

Plutôt oui (2)	153	49.35%
Plutôt non (3)	24	7.74%
Non pas du tout (4)	8	2.58%
Cela m'est égal (5)	53	17.10%

Q54 Connaissez-vous le Service des ressources informationnelles et archives (UNIRIS) de l'UNIL ?

Oui (1)	11	3.55%
Seulement le nom (2)	71	22.90%
Non pas du tout (3)	228	73.55%

Q55 Avez-vous déjà sollicité ce service ? Si oui, veuillez préciser pour quelles raisons.

Oui (1)	4	1.29%
Non (2)	306	98.71%
Commentaires	6	1.94%

Q56 Vous êtes...

Féminin (F)	140	45.16%
Masculin (M)	170	54.84%

Q57 Quel âge avez-vous ?

- de 30 ans (1)	100	32.26%
31-40 ans (2)	106	34.19%
41-50 ans (3)	56	18.06%
+ de 50 ans (4)	48	15.48%

Q58 A quelle Faculté de l'UNIL (ou autre institution) êtes-vous rattaché ? Si "autre institution", veuillez

Faculté de biologie et de médecine (1)	80	25.81%
Faculté de droit, des sciences criminelles et d'administration publique (2)	37	11.94%
Faculté des géosciences et de l'environnement (3)	34	10.97%
Faculté des Hautes Etudes Commerciales (HEC Lausanne) (4)	23	7.42%
Faculté des lettres (5)	58	18.71%
Faculté des sciences sociales et politiques (6)	66	21.29%
Faculté de théologie et de sciences des religions (7)	11	3.55%
Autre institution (8)	1	0.32%
Commentaires	3	0.97%

Q59 Quel est votre principal domaine de recherche ? Si autre, veuillez préciser.

Administration publique (1)	11	3.55%
Arts (2)	13	4.19%
Biologie (20)	60	19.35%
Droit (3)	15	4.84%
Etudes du tourisme (4)	1	0.32%
Géosciences (5)	22	7.10%
Histoire (6)	34	10.97%
Langues (7)	24	7.74%
Médecine (8)	28	9.03%
Philosophie (9)	8	2.58%
Psychologie (10)	28	9.03%
Sciences criminelles (11)	14	4.52%
Sciences de l'environnement (12)	13	4.19%
Sciences des religions (13)	9	2.90%
Sciences du sport et de l'éducation physique (14)	8	2.58%
Sciences économiques (15)	19	6.13%
Sciences humaines (16)	19	6.13%
Science politique (17)	23	7.42%
Sciences sociales (18)	49	15.81%
Théologie (19)	5	1.61%
Autre	14	4.52%

Q60 Quelle est votre fonction (ou la fonction la plus élevée si vous en avez plusieurs) ? Si autre, veuillez

Assistant-e boursier-ère (1)	5	1.61%
Assistant-e diplômé-e (2)	90	29.03%
Assistant-e FNS (3)	40	12.90%
Assistant-e fonds externe (4)	6	1.94%
Chargé-e de cours (5)	5	1.61%
Maître d'enseignement et de recherche (suppléant-e ou MER1 ou MER2) (6)	25	8.06%
Maître assistant-e (7)	9	2.90%
Personnel de recherche (8)	7	2.26%
Personnel technique et de laboratoire (9)	1	0.32%
Premier-ère assistant-e (10)	13	4.19%

Premier-ère assistant-e FNS (11)	5	1.61%
Privat-docent (12)	2	0.65%
Prof. assistant-e (PAST ou PAST-PTC ou boursier FNS) (13)	13	4.19%
Prof. associé-e (14)	20	6.45%
Prof. honoraire (15)	1	0.32%
Prof. invité-e (16)	0	0.00%
Prof. ordinaire (17)	45	14.52%
Prof. remplaçant-e (18)	2	0.65%
Prof. titulaire (19)	1	0.32%
Autre	20	6.45%

Q61 Combien d'années d'expérience dans la recherche avez-vous ?

- de 5 ans (1)	129	41.61%
5-10 ans (2)	65	20.97%
11-20 ans (3)	54	17.42%
+ de 20 ans (4)	62	20.00%

Q62 Ici vous pouvez nous faire part de vos attentes et besoins spécifiques en matière de gestion de données de recherche, de vos commentaires, suggestions et propositions d'améliorations.

Sans réponse	269	86.77%
--------------	-----	--------

Annexe 14 : Guide d'entretien

Objectifs de l'entretien

L'entretien semi-directif a pour objectif principal de recueillir des informations et des données qualitatives qui compléteront l'enquête *Limesurvey* menée auprès des chercheurs. De plus, un certain nombre de thématiques et problématiques concernant la gestion des données de recherche (GDR) seront discutées avec les personnes interviewées afin de pouvoir notamment :

- Cerner le profil personnel de la personne (fonction, activités et tâches) ;
- Comprendre sa réalité et son expérience vécue au quotidien ;
- Comprendre ses liens avec le monde de la recherche et les chercheurs ;
- Evaluer son degré de connaissance de la thématique GDR ;
- Comprendre le contexte général dans lequel elle évolue (Faculté, Département, Unité, Service,...) ;
- Identifier les attentes et les besoins à la fois généraux (Faculté ou autre) et spécifiques (niveau individuel) ;
- Comprendre quels sont les enjeux futurs pour la Faculté (ou autre) en matière de GDR ;
- Evaluer le niveau/degré d'adhésion de la Faculté (ou autre) à un éventuel futur projet de GDR ;
- Evaluer le niveau de visibilité du Service UNIRIS.

Profil personnel

- ✓ Faculté(s)
- ✓ Unité(s)
- ✓ Fonction(s)
- ✓ Activités et tâches en relation avec les chercheurs
- ✓ Quels types de demandes vous adressent les chercheurs ?
- ✓ Parmi ces demandes, lesquelles sont les plus récurrentes ?
- ✓ Quels problèmes particuliers rencontrent-ils ?

Connaissance du sujet

- ✓ La notion de gestion des données de recherche vous parle-t-elle ?
- ✓ Le terme anglais de Research Data Management est-il plus significatif ?
- ✓ Que signifie pour vous la gestion des données de recherche ?
- ✓ Que signifie pour vous une donnée de recherche ?

Situation réelle au sein de la Faculté

- ✓ Quels types de données sont produits au sein de la Faculté ?
- ✓ La thématique de la gestion des données de recherche suscite-t-elle de l'intérêt ?
- ✓ Existe-t-il des procédures et/ou des directives internes mises en place pour gérer les données de recherche ?
- ✓ Savez-vous ce qu'est un Data Management Plan ?
- ✓ Avez-vous déjà eu des demandes spécifiques liées au DMP ? (Si oui, lesquelles ?)

Attentes et besoins au sein de la Faculté

- ✓ Quels sont selon vous les principaux problèmes rencontrés par les chercheurs lorsqu'ils gèrent leurs données ?
- ✓ Ex. Accessibilité, archivage, classement, communication, conservation, description (métadonnées), droit d'auteur, organisation, partage, pérennisation, propriété intellectuelle, réutilisation, sécurité, stockage des données de recherche,...
- ✓ Les chercheurs expriment-ils leurs attentes et besoins en matière de GDR ?
- ✓ Si oui, lesquels ?

Projet de gestion des données de recherche

- ✓ Trouveriez-vous utile qu'un projet de GDR soit mené à l'UNIL ? (vs niveau national)
- ✓ Par qui ce projet devrait-il être mené, qui devrait être impliqué dans le projet ?
- ✓ Qui, au niveau de votre Faculté, devrait participer activement à un tel projet ?

Formation à la GDR

- ✓ Pensez-vous qu'un « programme de formation » à la GDR serait utile et pourquoi ?
- ✓ Qui, selon vous, devrait proposer cette formation ?
- ✓ Parmi les différentes offres de formation, lesquelles jugez-vous indispensables, lesquelles vous seraient vraiment utiles ?
- ✓ Ex. Site web dédié, formation en ligne (autoformation, e-learning), workshops, formations en groupe ou individuelles, séminaires intégrés au cursus doctoral, documentation (ex. guides pratiques, prospectus, autre)

Perspectives futures

- ✓ Quels sont, selon vous, les principaux enjeux futurs pour le monde de la recherche et des chercheurs ?
- ✓ Même question, mais pour la gestion des données de recherche ?

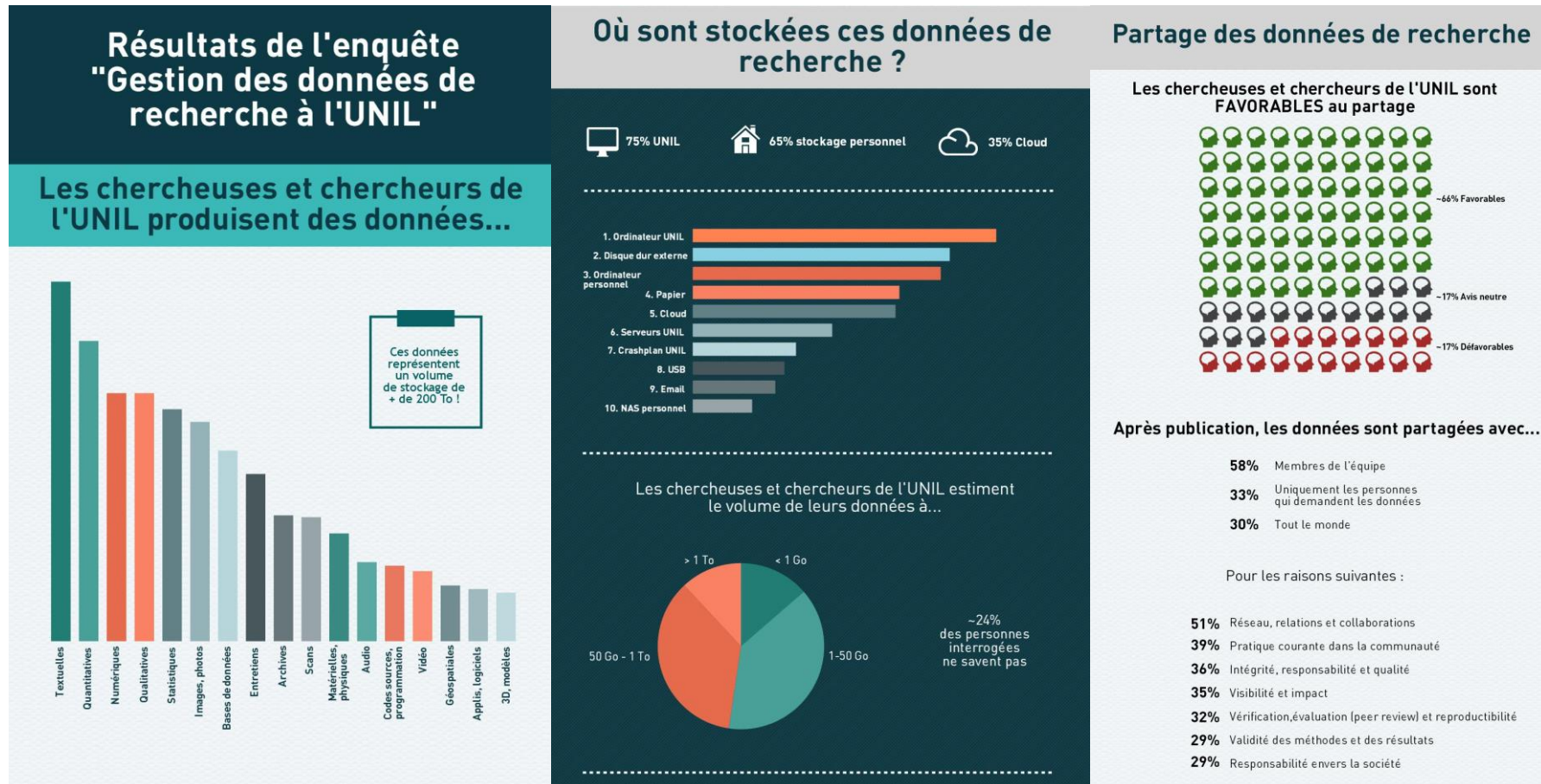
Annexe 15 : Version abrégée du Guide et « checklist » de la GDR



Checklist de la gestion des données de recherche

- ✓ **Politiques, procédures, normes et standards** utilisés ?
- ✓ **Exigences** des bailleurs de fonds et exigences légales
- ✓ **Cycle de vie** : création, traitement, analyse, préservation, accès et réutilisation
- ✓ **Description et documentation** du projet de recherche (utilisation de métadonnées) : contexte, méthodologie, processus, récolte, instruments, manipulations, mesures, standards, variables,...
- ✓ Organisation des **dossiers et fichiers** ; choix des **formats de fichiers** (ouverts et pour l'archivage) ; gestion des versions
- ✓ **Stockage, sauvegardes, migrations et sécurité**
- ✓ **Données sensibles** : anonymisation, confidentialité, formulaires de consentement (accès, diffusion, partage, destruction), protection, sécurité des données ; **éthique et déontologie**
- ✓ **Droit d'auteur et propriété intellectuelle** : licences, brevets, *Open Access*, **publication**, périodes de rétention
- ✓ **Archivage, conservation** à long terme et préservation : dépôts (*repositories*) de données
- ✓ **Partage et réutilisation** des données : conditions (droits et restrictions), visibilité (DOIs, citations), valorisation et nouvelles recherches ; *à qui ? quand ? comment ? pourquoi ?*
- ✓ **Rôles et responsabilités** : propriété des données, responsabilités, ressources à disposition (humaines, financières, temporelles, matérielles,...)

Annexe 16 : Exemples de posters¹⁰⁸ de notre enquête



¹⁰⁸ Toutes nos infographies ont été réalisées avec l'outil en ligne Piktochart.

Annexe 17 : « Brainstorming » Centre de soutien à la gestion des données de recherche

