

Frédéric Erard / Mathilde Heusghem / Clément Parisato

Recherche biomédicale et Open Data

Perspectives en droit suisse

Les politiques de libre accès à la recherche et aux données de recherche sont de plus en plus répandues dans notre société. Certaines visent notamment les données de santé utilisées dans le cadre de la recherche biomédicale. Cet article a pour objectif d'analyser les interactions entre politique de libre accès d'un côté et contraintes juridiques de l'autre, s'agissant de données de recherche biomédicale. Il propose également une liste d'éléments à considérer avant toute mise à disposition de données de recherche biomédicale.

Catégories d'articles : Articles scientifiques

Domaines juridiques : Droit de la santé ; Protection des données

Proposition de citation : Frédéric Erard / Mathilde Heusghem / Clément Parisato, Recherche biomédicale et Open Data, in : Jusletter 30 janvier 2023

Table des matières

1. Introduction
2. Contexte et définitions
 - 2.1. Open Science, Open Data et Open Research Data
 - 2.2. Contexte suisse
 - 2.3. Open Data et recherche biomédicale
3. Obstacles liés à la protection des données personnelles
 - 3.1. Cadre légal applicable et champ d'application
 - 3.2. Ouverture des données par l'anonymisation ?
 - 3.3. Réutilisation de données personnelles, en particulier à des fins de recherche sur l'être humain
 - 3.3.1. Régime de la LRH
 - 3.3.2. Régime hors de la LRH
 - 3.4. Nécessité d'adopter un cadre contractuel en cas de partage de données personnelles
 - 3.5. Recours aux modèles et technologies améliorant la protection de la vie privée (Privacy Enhancing Technologies, PET)
 - 3.6. Excursus : projet d'espace européen des données en Europe
 - 3.7. Constat intermédiaire et brèves considérations sur l'architecture juridique de la réutilisation des données à des fins de recherche
4. L'ouverture des données
 - 4.1. Les obstacles liés à l'ouverture
 - 4.1.1. Régimes juridiques applicables
 - 4.1.2. Exemples d'obstacles
 - 4.2. La mise en circulation des données : licences & waiver
5. Application pratique et check-list
6. Conclusion

1. Introduction¹

[1] Le paradigme de la science et des données ouvertes constitue indéniablement un facteur de développement pour les sociétés actuelles en ce sens qu'il vise à la fois à accroître la confiance dans les résultats de recherches générées et à permettre le développement de nouvelles connaissances. La pandémie de Covid-19 a une nouvelle fois démontré la nécessité de développer des solutions permettant la collecte et l'échange sécurisé d'informations liées aux données de santé. L'accès et la réutilisation de ces données scientifiques ont permis de faciliter le développement de vaccins, mais aussi d'adapter les mesures prises par les gouvernements en réaction à l'évolution de la pandémie.

[2] La philosophie du libre accès repose sur un principe relativement simple, à savoir « aussi ouvert que possible, aussi fermé que nécessaire ». En pratique toutefois, il n'est pas évident de déterminer dans quelle mesure les règles imposées par les différents régimes juridiques applicables aux données de recherche et leur éventuelle superposition peuvent permettre d'envisager leur ouverture en libre accès, conformément aux standards d'Open Data. Alors que la thématique générale des conditions-cadre et des incitatifs au partage des données de santé en Suisse fait l'objet

¹ Tous les liens internet mentionnés dans cet article ont été visités la dernière fois le 15 décembre 2022.

d'études régulières dans différentes disciplines, notamment en éthique médicale², elle a encore donné lieu à peu de contributions juridiques en Suisse.

[3] Après une brève présentation de la notion de libre accès (2) nous examinerons les obstacles que peuvent soulever l'ouverture des données biomédicales, en particulier du point de vue de la protection des données personnelles (3), afin de déterminer dans quelle mesure les données biomédicales peuvent être mises à disposition en conformité avec les principes prônés par l'Open Data (4). Nous terminerons en esquisant une brève proposition de check-list des aspects dont il faut tenir compte lorsqu'on envisage d'« ouvrir » des données à des fins de recherche. Certaines problématiques nous conduiront à examiner les approches développées et les solutions retenues notamment au niveau européen, mais les réflexions qui suivent seront menées essentiellement à l'aune du droit suisse.

2. Contexte et définitions

2.1. Open Science, Open Data et Open Research Data

[4] Le paradigme de libre accès aux données (ou Open Data) est une notion qui s'inscrit dans le mouvement plus global de l'Open Science, qui repose sur l'idée que le savoir académique doit être librement accessible au public sans restriction. Si la notion n'est pas strictement définie en droit positif, elle fait l'objet de définitions dans différentes déclarations et initiatives visant à la promouvoir.

[5] Ainsi, dans sa recommandation pour une science ouverte, l'UNESCO définit l'Open Science comme « [...] un concept inclusif qui englobe différents mouvements et pratiques visant à rendre les connaissances scientifiques multilingues, librement accessibles à tous et réutilisables par tous, à renforcer la collaboration scientifique et le partage des informations au profit de la science et de la société, ainsi qu'à ouvrir les processus de création, d'évaluation et de diffusion des connaissances scientifiques aux acteurs de la société au-delà de la communauté scientifique traditionnelle ».³

[6] L'Initiative de Budapest⁴ et la Déclaration de Berlin de 2003⁵ constituent deux initiatives majeures consacrant le mouvement de l'Open Science. Tandis que la première promeut le libre accès aux publications de recherche, la seconde élargit ce champ en proposant de garantir le libre accès à la littérature scientifique mondiale, c'est-à-dire aussi bien aux résultats de recherche qu'aux outils ayant permis de collecter les données. Elle indique que les auteurs et les titulaires des droits sur les contributions en libre accès accordent à tous les utilisateurs un droit gratuit, irrévocable et mondial d'accès à l'œuvre, ainsi qu'une licence. Cela couvre les résultats originaux

² Voir p. ex. ANDREA MARTANI *et al.*, *Evolution or Revolution? Recommendations to Improve the Swiss Health Data Framework*, *Front. Public Health*, 31 mai 2021, consultable ici : <https://doi.org/10.3389/fpubh.2021.668386> ; LESTER DARRYL GENEVIÈVE *et al.*, *Individual notions of fair data sharing from the perspectives of Swiss stakeholders*, *BMC Health Services Research*, 23 septembre 2021, consultable ici : <https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-021-06906-2>.

³ Recommandation de l'UNESCO sur une science ouverte, 23 novembre 2021, p. 7 (voir https://unesdoc.unesco.org/ark:/48223/pf0000379949_fre).

⁴ Budapest Open Access Initiative (BOAI) du 14 février 2002 (voir <https://www.budapestopenaccessinitiative.org/>).

⁵ Déclaration de Berlin sur le Libre Accès à la Connaissance en Sciences exactes, Sciences de la vie, Sciences humaines et sociales du 22 octobre 2003 (voir https://openaccess.mpg.de/68042/BerlinDeclaration_wsis_fr.pdf). Cette déclaration a été ratifiée par 768 institutions (<https://openaccess.mpg.de/319790/Signatories>).

de la recherche scientifique, les données brutes et les métadonnées, les documents sources, les représentations numériques de documents picturaux et graphiques et les documents scientifiques multimédia.⁶

[7] On peut ainsi retenir que l'Open Science suppose un accès et une réutilisation gratuite des résultats de recherche disponibles sur internet afin de favoriser leur dissémination à large échelle. Ce concept favorise à la fois le transfert des connaissances, mais aussi leur accès et leur visibilité aussi bien à l'égard de la communauté scientifique qu'à celui du grand public.

[8] L'Open Data constitue une sous-catégorie de l'Open Science (aux côtés par exemple de l'Open source software (logiciels libres)) et implique d'ouvrir, donc de rendre libre l'utilisation, la réutilisation, la conservation et la redistribution des données, dans un format exploitable et lisible par les humains et les machines. Les textes encourageant l'Open Data se réfèrent souvent, s'agissant de la mise en œuvre, aux principes FAIR (*Findable, Accessible, Interoperable, Reusable*), soit en traduction libre : trouvables, accessibles, interopérables et réutilisables. Élaborés par des chercheurs afin de guider la mise en place de stratégies d'Open Data,⁷ ils sont aujourd'hui largement reconnus et servent régulièrement de référence à des institutions, projets et initiatives de recherche d'envergure, tant au niveau suisse⁸ qu'international.⁹ La démarche d'ouverture des données de recherche, qui fait partie du paradigme de l'Open Data, est désignée par le terme Open Research Data.¹⁰

2.2. Contexte suisse

[9] Consciente de l'importance de ces enjeux, la Suisse a adopté en 2021 une Stratégie Nationale en matière d'Open Research Data (ORD)¹¹, élaborée conjointement par swissuniversities¹², le Domaine des EPF, le FNS et les Académies suisses des sciences.¹³

[10] La Stratégie Nationale ORD vise à faciliter l'accès aux données de recherche et leur réutilisation. Cette vision doit se matérialiser par le développement de pratiques basées sur le partage des données de recherche en Suisse, notamment par la régulation des services et infrastructures de soutien des chercheurs à cet effet.¹⁴ Les principes FAIR doivent être appliqués s'agissant de don-

⁶ Cf. Déclaration de Berlin.

⁷ MARK D. WILKINSON *et al.*, *The FAIR Guiding Principles for scientific data management and stewardship*, *Scientific Data* 3, 160018 (2016), consultable ici : <https://www.nature.com/articles/sdata201618>.

⁸ La Stratégie Nationale Suisse Open Research Data, dont il sera question dans la section suivante, s'y réfère explicitement (voir nbp 11).

⁹ C'est le cas notamment du European Open Science Cloud et des projets Horizon Europe.

¹⁰ Pour une définition (au niveau suisse), voir *infra* 2.2 et nbp 11.

¹¹ Stratégie Nationale Suisse Open Research Data du 23 avril 2021, consultable ici : https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/ORD/Swiss_National_ORD_Strategy_fr.pdf. La notion d'Open Research Data (ORD) y est expliquée ainsi (p. 4) : « En facilitant l'accès aux données de recherche et leur réutilisation, l'ORD promeut une recherche meilleure, plus efficace et plus percutante au profit de la société dans son ensemble. »

¹² swissuniversities est l'organisation faitière des hautes écoles suisses.

¹³ Voir : <https://www.swissuniversities.ch/fr/themes/digitalisation/open-research-data/strategie-nationale-et-plan-daction>.

¹⁴ Stratégie Nationale Suisse Open Research Data, p. 4.

nées de recherche financée par des fonds publics.¹⁵ La Stratégie Nationale ORD prend notamment acte du fait que la production, l'accès et l'utilisation de ces données présentent de nombreux défis juridiques, éthiques et sociaux. Certains sont liés à la structure fédéralisée de la Suisse, dont il résulte que des bases légales cantonales, fédérales, et parfois internationales coexistent. Ces dispositions légales régissent à la fois les processus impliquant des données (de l'acquisition à la réutilisation), ainsi que les différents niveaux de responsabilité incombant aux personnes et entités impliquées.¹⁶

[11] Les politiques d'Open Data sont souvent mises en œuvre par le biais d'instruments incitatifs. C'est le cas en Suisse notamment des règles de financement du Fonds National Suisse (FNS), qui prévoient que les bénéficiaires de subsides s'engagent à ce que les résultats de recherche soutenus par des ressources du FNS soient rendus accessibles au public de manière appropriée.¹⁷ Le FNS privilégie une approche *bottom-up* qui consiste à fournir des directives sur les bonnes pratiques et permet à chaque communauté scientifique de définir et d'appliquer ses propres standards avec une grande flexibilité.¹⁸ Cette approche a cependant ses limites, puisque le caractère contraignant de ces stratégies est restreint et peut être appliqué de façon très différente au sein de la communauté scientifique suisse.

[12] Dans le prolongement de la Stratégie Nationale ORD, en août 2022, la Commission de la science, de l'éducation et de la culture du Conseil des États a déposé une motion visant à charger le Conseil fédéral de créer, dans une loi-cadre, les bases nécessaires afin que des infrastructures spécifiques permettant de réutiliser des données dans des domaines stratégiques (incluant notamment la santé et la recherche) soient rapidement développées et mises en place.¹⁹ Le Conseil fédéral a relevé l'importance de la réutilisation des données, tout en rappelant que celle-ci n'était pas toujours évidente à concilier avec les conditions instaurées par la Loi fédérale sur la protection des données personnelles (LPD)²⁰ pour la collecte de données (finalités déterminées et reconnaissables pour la personne concernée). En effet, il est souvent difficile, voire impossible de prévoir l'utilité que des données pourraient présenter si elles étaient utilisées à d'autres fins, au moment de leur collecte. Le Conseil fédéral a donc proposé d'accepter la motion et précisé qu'il s'attacherait en particulier à déterminer les domaines dans lesquels une utilisation secondaire des données serait pertinente et proportionnée, ainsi que les infrastructures et autres exigences préalables qui seraient nécessaires pour exploiter des espaces de données fiables et interopérables. Le cadre juridique suisse devrait donc évoluer à l'avenir et se doter d'une réglementation en matière de réutilisation des données, notamment de santé.

[13] En parallèle des développements politiques, la communauté scientifique suisse fait progresser le partage de données à l'échelle nationale et met en œuvre des projets qui concrétisent les objectifs de la Stratégie Nationale ORD. Dans le domaine de la biologie par exemple, un des projets les plus ambitieux est certainement le *SwissBiodata ecosystem* (SBDe) qui implique à ce jour

¹⁵ Voir : <https://www.swissuniversities.ch/fr/themes/digitalisation/open-research-data/strategie-nationale-et-plan-daction>.

¹⁶ Stratégie Nationale Suisse Open Research Data, p. 10.

¹⁷ Règlement du Fonds national suisse relatif à l'octroi de subsides, art. 47.

¹⁸ Voir : <https://www.snf.ch/fr/FAiWVH4WvpKvohw9/dossier/points-de-vue-politique-de-recherche>.

¹⁹ Motion 22.3890 du 28 août 2022 de la Commission de la science, de l'éducation et de la culture du Conseil des États, Élaboration d'une loi-cadre sur la réutilisation des données. Consultable ici : <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaefft?AffairId=20223890>.

²⁰ Loi fédérale sur la protection des données personnelles du 19 juin 1992 (LPD ; RS 235.1).

un total de cinquante-trois plateformes, installations et groupes de recherche affiliés à dix-huit institutions suisses.²¹ Le projet SBDe se fonde sur le constat que les universités et instituts de recherche suisses ont largement développé leurs plateformes locales de génération et traitement de données et ont adopté les principes de l'ouverture des données de recherche (ORD). Cependant, l'élaboration d'une stratégie ORD à échelle nationale en Suisse nécessite un partage et une réutilisation des données plus efficaces. Il est donc essentiel d'adopter des normes de qualité et d'exploitation communes et d'instaurer une collaboration étroite entre les institutions suisses et leurs experts. Le SBDe envisage une infrastructure décentralisée qui ambitionne de relever ces défis, renforçant ainsi la capacité de la Suisse à convertir les données de recherche en savoir et en innovation, avec pour objectifs (i) d'accroître la qualité, la normalisation et l'efficacité de la chaîne de valeur des données (de la production de données à la génération de connaissances), par la fédération de plateformes ; (ii) de fournir un soutien de pointe à la communauté scientifique suisse pour rendre ses données, méthodes, et outils logiciels conformes aux principes FAIR ; et (iii) d'établir de nouvelles ressources qui renforceront la compétitivité internationale de la Suisse et sa position dans l'infrastructure de données pour les sciences du vivant.²²

2.3. Open Data et recherche biomédicale

[14] La recherche biomédicale, sur laquelle cet article se concentre, est directement confrontée aux incitatifs en matière d'Open Science et d'Open Data. Cette discipline regroupe l'étude des processus de la vie, de la prévention et du traitement des maladies, ainsi que des facteurs génétiques environnementaux liés à la maladie et à la santé.²³ C'est dans ce champ de recherche qu'ont notamment émergé l'informatique médicale²⁴ et la médecine personnalisée,²⁵ offrant de nombreuses perspectives de développement de traitements.

[15] En 2017, le Secrétariat d'État à la formation, à la recherche et à l'innovation (SEFRI) a lancé l'initiative nationale d'encouragement pour la médecine personnalisée sur une période de quatre ans, reconduite pour quatre ans supplémentaires en 2021.²⁶ Dans ce cadre, le SEFRI et l'Office fédéral de la santé publique (OFSP) ont confié à l'Association suisse des sciences médicales (ASSM)

²¹ Voir : https://www.sib.swiss/about/news/10943-the-swissbiodata-ecosystem-gains-further-traction?utm_source=LinkedIn&utm_medium=social&utm_campaign=organic&utm_content=SwissBioData-newpartners.

²² Voir : <https://www.sib.swiss/pages/10890-swissbiodata-ecosystem-boosting-switzerland-s-data-intensive-research-in-life-sciences>.

²³ Voir : <https://statesforbiomed.org/education/background-on-biomedical-research/biomedical-research-definitions/>.

²⁴ L'« informatique médicale » est la science du développement, de la gestion, du stockage, du traitement et de la mise à disposition systématiques de données, d'informations et de connaissances dans le domaine de la médecine et du système de santé, par le recours à des théories, méthodes, procédures et techniques informatiques. Bulletin ASSM 2/15, p. 1, consultable ici : https://www.samw.ch/dam/jcr:b169468a-9bcb-42c3-bf0d-853fda9eb67b/bulletin_assm_15_2.pdf.

²⁵ La médecine personnalisée vise à exploiter le potentiel de différents ensembles de données liées à la santé humaine (telles que les données « omics » (génomique, transcriptomique, protéomique, métabolomique, etc.)), les données cliniques, les données issues de biobanques ou les données de santé relevées par les individus eux-mêmes (self-tracking) pour optimiser la prise en charge médicale individuelle (voir : <https://sphn.ch/fr/organization/about-personalized-health/>). Voir également Swiss Academies Communications, Vol. 14, No 6, 2019, Ch. 1 (consultable ici : https://www.samw.ch/dam/jcr:b4aad9d0-c7d9-4c2f-b998-7f3c4727e539/assm_2019_medecine_personnalisee_chapitre_1.pdf).

²⁶ Voir : <https://www.sbf.admin.ch/sbfi/fr/home/recherche-et-innovation/la-recherche-et-linnovation-en-suisse/initiative-nationale-dencouragement-de-la-medecine-personnalisee.html>.

et au SIB (Institut Suisse de Bioinformatique) le mandat de mettre en œuvre un réseau suisse de santé personnalisée (*Swiss Personalized Health Network*, SPHN). L'initiative SPHN contribue notamment au développement, à l'implémentation et à la validation d'infrastructures coordonnées afin de rendre interopérables les données de santé pour la recherche en Suisse.²⁷ Elle a par ailleurs permis d'établir un environnement informatique sécurisé pour l'analyse et le partage de données sensibles garantissant la protection des données (infrastructure BioMedIT), mise à disposition des chercheurs dans toute la Suisse.²⁸ La deuxième période d'encouragement de l'initiative SPHN, initiée en 2021, se focalise sur la consolidation de cette infrastructure.²⁹

[16] Dans une autre perspective, le Conseil fédéral a publié en mai 2022 un rapport attendu et important intitulé : « Mieux utiliser les données médicales pour assurer l'efficacité et la qualité des soins » (Rapport Humbel).³⁰ Ce rapport s'articule autour de trois questions principales, à savoir (i) le potentiel de la réutilisation et de l'exploitation des données médicales, (ii) l'identification des conditions à remplir pour que les données médicales puissent être réutilisées par différents groupes d'utilisateurs et (iii) l'exploration des développements techniques, organisationnels et juridiques à initier pour permettre la réutilisation des données de santé. Le rapport propose l'instauration d'un « espace de données » dans le secteur de la santé, qui doit prendre en compte des aspects techniques, juridiques ou sémantiques, mais aussi le développement d'une culture commune de réutilisation des données. Il souligne notamment la nécessité de clarifier le régime de propriété des données, qui pourraient potentiellement être considérées comme un bien collectif si les personnes concernées y consentent.³¹ L'importance de respecter les principes FAIR est également mise en avant.³² Au terme du rapport, le Conseil fédéral souligne la nécessité de procéder à une adaptation du cadre légal pour établir un cadre de confiance et de sécurité juridique permettant la réutilisation des données à des fins diverses.³³ Le rapport reste néanmoins passablement vague sur les aspects juridiques.

[17] La recherche biomédicale exploite différents types de données, parmi lesquelles des données personnelles liées ou non à la santé. Le spectre de données couvert par la recherche biomédicale est très vaste et inclut notamment (en particulier dans le cadre de la médecine personnalisée) les données dites « -omiques », en référence au suffixe commun à certaines d'entre elles. Ce terme inclut par exemple les données génomiques (issues de techniques d'étude du génome entier), transcriptomiques (issues de techniques d'analyse d'ARNm³⁴ et du niveau d'expression des gènes

²⁷ Voir : <https://sphn.ch/fr/home/>.

²⁸ Voir : www.biomedit.ch.

²⁹ Voir : <https://www.sbf.admin.ch/sbfi/fr/home/services/publications/base-de-donnees-des-publications/s-n-2021-2/s-n-2021-2h.html>.

³⁰ Rapport du Conseil fédéral donnant suite au postulat 15.4225 Humbel du 18 décembre 2015, Mieux utiliser les données médicales pour assurer l'efficacité et la qualité des soins, 4 mai 2022, consultable ici : <https://www.admin.ch/gov/fr/accueil/documentation/communiqués.msg-id-88631.html> (Rapport Humbel du 4 mai 2022). Pour un résumé, voir : FRÉDÉRIC ERARD, Vers un nouveau modèle de réutilisation des données médicales en Suisse ?, 2 juillet 2022, in : www.swissprivacy.law/156.

³¹ Rapport Humbel du 4 mai 2022 (nbp 30), p. 26.

³² Rapport Humbel du 4 mai 2022 (nbp 30), p. 26 ss.

³³ Rapport Humbel du 4 mai 2022 (nbp 30), p. 43.

³⁴ Acronyme d'acide ribonucléique messenger. Pour une définition, voir : https://sciencesnaturelles.ch/covid19-vaccination-explained/mrna_vaccines/was_ist_eine_mrna_und_welche_funktion_hat_sie_.

dans un tissu), protéomiques (se rapportant à la composition des protéines dans un tissu) et métabolomiques (se rapportant aux produits métaboliques).³⁵

[18] De manière générale, il est important de souligner que l'encouragement de l'Open Science, en particulier l'Open Data, n'équivaut donc pas à imposer la mise en libre accès des données de recherche, de manière inconditionnelle. Au contraire, l'Open Research Data cohabite avec plusieurs régimes juridiques qui visent entre autres à protéger les intérêts des personnes dont les données sont exploitées. Plus largement encore, le Conseil fédéral observe, dans le cadre de la Stratégie en matière de libre accès aux données publiques en Suisse, que « [l]a publication des données en libre accès doit respecter les dispositions en matière de protection des données, de sécurité de l'information, de droit d'auteur et de secret professionnel ».³⁶ La rencontre entre la recherche biomédicale, le cadre juridique dans lequel elle s'inscrit et les encouragements à l'Open Data peuvent donc engendrer des tensions.

3. Obstacles liés à la protection des données personnelles

[19] Le droit fondamental à la protection de la vie privée dans le domaine de la santé est un droit largement reconnu aussi bien par les textes conventionnels³⁷ ou éthiques³⁸ internationaux que par le droit constitutionnel helvétique.³⁹ Dans le domaine de la santé et de la recherche, la protection de la vie privée des personnes concernées vise notamment à prémunir celles-ci contre les utilisations abusives de leurs données (sensibles de surcroît) qui pourraient conduire à des formes de discrimination par exemple. La nécessité de réguler les traitements de données personnelles relatives aux patients ou sujets de recherche sert également à préserver la confiance de la population à l'égard de la science.

[20] La protection de la vie privée des sujets de recherche peut entrer en conflit avec les intérêts de la recherche biomédicale, dont les avancées sont souvent tributaires de l'accès aux données personnelles des sujets de recherche. De manière plus générale, la recherche biomédicale a connu au cours des deux dernières décennies une évolution marquée par le partage croissant de données. Cela s'explique notamment par les avancées techniques, des investissements dans les infrastruc-

³⁵ Pour de plus amples informations, voir le site de l'Académie suisse des sciences naturelles : https://sciencesnaturelles.ch/personalized-health-explained/welche_methoden_werden_eingesetzt_/omcis.

³⁶ Stratégie en matière de libre accès aux données publiques en Suisse pour les années 2019 à 2023, adoptée par le Conseil fédéral le 30 novembre 2018, p. 858 (voir : <https://www.bfs.admin.ch/bfs/fr/home/services/ogd/strategie.html#:~:text=La%20%C2%ABStrat%C3%A9gie%20en%20mat%C3%A8re%20de,public%20sur%20le%20portail%20opendata>).

³⁷ Notamment : art. 8 Convention de sauvegarde des droits de l'homme et des libertés fondamentales du 4 novembre 1950 (CEDH; RS 0.101); art. 1 Convention pour la protection des Droits de l'Homme et de la dignité de l'être humain à l'égard des applications de la biologie et de la médecine du 4 avril 1997 (Convention sur les Droits de l'Homme et la Biomédecine; RS 0.810.2); Convention du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel du 28 janvier 1981 (Convention STE n° 108; RS 0.235.1).

³⁸ P. ex. Déclaration de l'Association Médicale Mondiale (AMM) sur les considérations éthiques concernant les bases de données de santé et les biobanques, octobre 2016 (communément désignée « Déclaration de Taipei »).

³⁹ Art. 13 al. 2 de la Constitution fédérale de la Confédération suisse du 18 avril 1999 (Cst.; RS 101), selon lequel toute personne a le droit d'être protégée contre l'emploi abusif des données qui la concernent.

tures permettant le partage des données, mais aussi par les exigences des agences de financement en matière d'Open Science.⁴⁰

[21] Or, l'augmentation des partages des données de recherche n'est pas sans conséquence du point de vue de la protection des données personnelles. Pour assurer un partage et une réutilisation efficaces des données, il est par exemple nécessaire d'établir des systèmes permettant l'appariement des données. L'appariement des données (*record linkage*) consiste à relier ou fusionner, à l'aide d'une ou plusieurs variables d'appariement, les données individuelles sur une même personne issues de deux jeux de données au moins.⁴¹ L'appariement génère alors la création de nouveaux jeux de données offrant davantage d'informations sur la personne concernée.

[22] La présente section offre un aperçu synthétique des règles susceptibles de s'appliquer en cas de mise à disposition de données personnelles à des fins de recherche biomédicale, dans une optique d'Open Science.

3.1. Cadre légal applicable et champ d'application

[23] Le traitement de « données personnelles » déclenche l'application des législations pertinentes en matière de protection des données personnelles. En Suisse, les traitements de données liées à la recherche sur l'être humain font l'objet d'un cadre législatif morcelé, composé de lois générales et de lois spéciales. Sous l'angle des lois générales, la Loi fédérale sur la protection des données, dont la révision totale entrera en vigueur le 1^{er} septembre 2023 (nLPD)⁴², régit les traitements de données personnelles effectués par les personnes privées (p. ex. fondations ou laboratoires privés) et les organes fédéraux (p. ex. EPFL et ETHZ).⁴³ À l'inverse, les traitements de données personnelles effectués par des organes publics cantonaux (p. ex. hôpitaux universitaires) sont régis à titre principal par les différentes lois cantonales sur la protection des données.

[24] Les lois générales sur la protection des données ont pour objectif principal de protéger les droits fondamentaux des personnes concernées, en imposant notamment des obligations aux responsables de traitements (p. ex. devoirs de respecter des principes généraux, devoirs d'informer, obligations de prendre des mesures de sécurité techniques ou organisationnelles) ou en reconnaissant des droits spécifiques aux personnes concernées (p. ex. droit d'accéder à ses données, droit de s'opposer à certains traitements de données).

[25] En parallèle des lois générales sur la protection des données, des législations dites « spéciales » peuvent imposer des règles spécifiques liées à certains types de traitements de données personnelles. En vertu de l'adage *lex specialis derogat legi generali*, les dispositions légales spéciales prennent en principe le pas sur les lois dites « générales ». En matière de protection des données personnelles, c'est du moins le cas si la réglementation spéciale vise à régler spécifiquement la protection des données elle-même ou si elle offre au moins une protection équivalente.⁴⁴ Dans le

⁴⁰ JANE KAYE, *The Tension Between Data Sharing and the Protection of Privacy in Genomics Research*, Annual Review of Genomics and Human Genetics 2012, p. 415 ss, p. 417.

⁴¹ À ce sujet et pour une description des différentes méthodes d'appariement, voir : Rapport Humbel du 4 mai 2022 (nbp 30), p. 12 ss.

⁴² FF 2020 7397 ss.

⁴³ Art. 2 al. 1 let. b LPD ; art. 2 al. 1 let. b nLPD.

⁴⁴ DAVID ROSENTHAL, *Die rechtlichen und gefühlten Grenzen der Zweitnutzung von Personendaten*, sic! 2021 p. 168 ss, p. 169.

contexte de la recherche biomédicale, la principale législation spéciale est la Loi fédérale relative à la recherche sur l'être humain (LRH).⁴⁵ Les dispositions de la LRH sont précisées par deux ordonnances d'exécution, l'Ordonnance fédérale relative à la recherche sur l'être humain (ORH)⁴⁶ et l'Ordonnance fédérale sur les essais cliniques (OClin).⁴⁷

[26] La LRH instaure entre autres des règles spéciales relatives aux traitements et à la réutilisation de données à des fins de recherche sur l'être humain (cf. *infra* 3.2). Les dispositions de la LRH ne s'appliquent cependant que dans la mesure où les activités visées tombent dans son champ d'application (art. 2 LRH). Le champ d'application de la LRH se limite en premier lieu à la recherche, c'est-à-dire la « recherche méthodologique visant à obtenir des connaissances généralisables »⁴⁸, sur les maladies humaines et sur la structure et le fonctionnement du corps humain.⁴⁹ Lorsque des données sont en jeu, le champ d'application de la LRH se limite de surcroît aux recherches pratiquées sur des « données personnelles liées à la santé », à l'exception des données liées à la santé qui ont été collectées anonymement ou qui sont anonymisées.⁵⁰ L'acte d'anonymiser des données constitue toutefois lui-même un traitement de données dont la réalisation peut faire l'objet de règles imposées par la LRH, à l'image de l'anonymisation de données génétiques à des fins de recherche.⁵¹

[27] Il est par ailleurs essentiel de s'assurer que le partage de données envisagé respecte les exigences imposées en termes d'éthique médicale ou de recherche. Les projets de recherche impliquant la réutilisation de données personnelles liées à la santé sont soumis à l'approbation d'une commission d'éthique (art. 45 al. 1 LRH). Cette dernière doit s'assurer, lorsqu'elle évalue un projet de recherche, que les exigences éthiques, juridiques et scientifiques prévues par la LRH sont remplies (art. 45 al. 2 LRH). En termes de mise à disposition et de partage de données, le texte de référence est certainement la Déclaration de l'Association Médicale Mondiale (AMM) sur les considérations éthiques concernant les bases de données de santé et les biobanques de 2016 (souvent désignée comme « Déclaration de Taipei »). Son objectif consiste à énoncer des principes éthiques relatifs à la collecte, au stockage et à l'utilisation de données identifiables au-delà des soins prodigués aux patients, notamment par leur utilisation dans des bases de données et des biobanques. La Déclaration de Taipei rappelle en particulier la nécessité de préserver le principe de confidentialité nécessaire pour préserver la confiance dans les bases de données (§ 10) ainsi que la nécessité de s'assurer que la collecte, le stockage et l'utilisation de données soient volontaires pour les personnes capables de consentir (§ 11). Elle dresse une liste des informations qui doivent être communiquées aux personnes qui consentent à la réutilisation de leurs données pour des utilisations multiples et indéfinies (§ 12). Sous l'angle des textes éthiques adoptés à l'éche-

⁴⁵ Loi fédérale relative à la recherche sur l'être humain du 30 septembre 2011 (LRH; RS 810.30).

⁴⁶ Ordonnance fédérale relative à la recherche sur l'être humain du 20 septembre 2013 (ORH; RS 810.301).

⁴⁷ Ordonnance fédérale sur les essais cliniques du 20 septembre 2013 (OClin; RS 810.305).

⁴⁸ Art. 3 let. a LRH.

⁴⁹ L'art. 3 let. b LRH définit la « recherche sur les maladies » comme « la recherche sur les causes, la prévention, le diagnostic, le traitement et l'épidémiologie des troubles physiques et psychiques de la santé ». L'art. 3 let. c LRH définit la « recherche sur la structure et le fonctionnement du corps humain » comme « la recherche fondamentale, en particulier dans les domaines de l'anatomie, de la physiologie et de la génétique du corps humain ainsi que la recherche non axée sur une maladie relative aux interventions sur le corps humain ».

⁵⁰ Art. 2 al. 1 let. e et art. 2 al. 2 let. c LRH.

⁵¹ Art. 32 al. 3 LRH, qui n'autorise l'anonymisation de données génétiques à des fins de recherche qu'à la condition que la personne concernée ou, le cas échéant, son représentant légal ou ses proches, ne s'y soient pas opposés, après avoir été informés.

lon helvétique, on peut notamment citer le *Ethical Framework for Responsible Data Processing in Personalized Health Research*, développé dans le cadre de l'initiative SPHN.⁵²

3.2. Ouverture des données par l'anonymisation ?

[28] La question de ce qui constitue ou non une donnée personnelle est centrale puisque les opérations de traitement de données *non personnelles* sortent des champs d'application de la LRH et des lois générales sur la protection des données. Dans une perspective d'ouverture des données, cela signifie que les données de recherche non personnelles (c'est-à-dire anonymes conformément au droit applicable) peuvent être mises à disposition d'autres chercheurs sans égard aux contraintes imposées par ces lois, étant cependant entendu que d'autres droits sont susceptibles de faire obstacle à la libre utilisation de telles données (droits exclusifs, secrets d'affaires, etc.). En pratique, la détermination du caractère personnel ou non d'une donnée de recherche donne souvent lieu à débat.

[29] Les données personnelles sur la santé sont définies par la LRH comme « les informations concernant une personne déterminée ou déterminable qui ont un lien avec son état de santé ou sa maladie, données génétiques comprises ».⁵³ À l'exception du fait que les données doivent nécessairement se rapporter à la santé⁵⁴ pour entrer dans le champ d'application de la LRH, la définition des données personnelles au sens de la LRH est similaire à celle de la LPD (même si la LPD utilise la terminologie « identifiée/identifiable » au lieu de « déterminée/déterminable »).⁵⁵ Une personne est déterminée lorsque son identité ressort directement des informations ; elle est déterminable lorsque les circonstances ou le contexte permettent d'identifier la personne par corrélation d'informations.⁵⁶

[30] À l'inverse des données personnelles, les données anonymes ou anonymisées sont celles qui ne peuvent pas être mises en relation avec une personne déterminée, ou alors seulement au prix d'efforts considérés comme démesurés.⁵⁷ L'art. 35 ORH précise que l'anonymisation implique de rendre définitivement méconnaissables ou de détruire toutes les informations qui, combinées, permettent de rétablir l'identité de la personne sans efforts disproportionnés. La même disposition indique qu'il faut au moins rendre méconnaissables ou détruire les noms, adresses, dates de naissance et numéros d'identification caractéristiques. La LRH et ses ordonnances ne définissent pas plus en détail les identificateurs indirects qui devraient être pris en compte pour réidentifier une personne et ne prévoient pas non plus de seuils chiffrés en dessous desquels une réidentification est considérée comme probable ou admise (p. ex. 20 patients partageant les mêmes caractéristiques).

⁵² Consultable ici : <https://sphn.ch/document/ethical-framework/>.

⁵³ Art. 3 let. f LRH.

⁵⁴ Selon le Conseil fédéral, la notion de données de santé est identique à celle de l'art. 3 let. c LPD, cf. Message du Conseil fédéral du 21 octobre 2019 sur la loi fédérale relative à la recherche sur l'être humain, FF 2009 7259 ss (Message LRH), p. 7310.

⁵⁵ Art. 3 let. a LPD ; art. 5 let. a nLPD. Message LRH (nbp 54), 7311 ; HFG-VAN-SPYK/RUDIN/SPRECHER/POLEDNA, art. 3 N 43, in : Bernhard Rütsche (édit.), *Humanforschungsgesetz (HFG)*, Berne 2015 (cit. SHK HFG-AUTEUR).

⁵⁶ Message du Conseil fédéral du 15 septembre 2017 concernant la loi fédérale sur la révision totale de la loi fédérale sur la protection des données et sur la modification d'autres lois fédérales, FF 2017 6565 ss (cit. Message nLPD), p. 6639. ATF 138 II 346 c. 6.1, JdT 2013 I 71. FRÉDÉRIC ERARD, *Les données codées dans le contexte de la recherche : personnelles ou anonymes ?*, AJP/PJA 2021 p. 606 ss, p. 608 ; ALEXANDRE JOTTERAND, *Personal Data or Anonymous Data : where to draw the lines (and why) ?*, in : Jusletter 15 août 2022, N 14 ss.

⁵⁷ Art. 3 let. i LRH.

téristiques)⁵⁸. Il convient donc d'axer l'analyse du caractère identifiable de données sur les efforts nécessaires pour réidentifier la personne, qui ne doivent pas être considérés comme démesurés. Il faut en principe prendre en compte l'investissement en coûts, en temps et en travail.⁵⁹ De manière générale, il est recommandé de faire preuve de prudence lorsque des données relatives à la santé sont en jeu⁶⁰, en particulier des données génétiques. En raison du lien particulièrement étroit entre ces données et la personne concernée, des sources externes toujours plus nombreuses avec lesquelles ces données peuvent être croisées et des moyens techniques et informatiques toujours plus performants, les risques de réidentification s'en trouvent renforcés, en particulier dans des États à la population restreinte comme la Suisse.

[31] Dans le contexte de la recherche biomédicale, les données personnelles sont par ailleurs fréquemment traitées sous forme « pseudonymisée » ou, pour reprendre la terminologie de la LRH, sous forme « codée ». Cette forme de traitement s'explique non seulement par le respect du principe de minimisation et de sécurité des données, mais aussi parce que la LRH facilite la réutilisation des données qui se trouvent sous forme codée (cf. *infra* 3.3.1) et que la pseudonymisation permet surtout de remonter au sujet de recherche concerné, ce qui peut souvent se révéler utile pour mener des investigations complémentaires. D'un point de vue juridique, la pseudonymisation et le codage au sens de la LRH sont des synonymes. Des données sont considérées comme codées si elles ne peuvent être mises en relation avec une personne déterminée qu'au moyen d'une clé.⁶¹ Pour être correctement codées, les données doivent apparaître comme anonymisées du point de vue d'une personne qui n'est pas en possession de la clé de correspondance avec les données originales.⁶²

[32] La question de savoir si une donnée personnelle codée ou pseudonymisée constitue encore une donnée personnelle à l'égard du destinataire qui ne détient pas la clé de correspondance suscite passablement de débats doctrinaux en Suisse et en Europe.⁶³ D'un point de vue dogmatique, les auteurs de doctrine ont généralement tendance à opposer les approches absolue et relative. Selon l'approche absolue, il suffit qu'un seul acteur de la communication (fournisseur ou destinataire) soit en mesure de réidentifier la personne concernée pour que les données soient considérées comme personnelles à l'égard de tous. À l'inverse, selon l'approche relative, des données pseudonymisées ou codées ne sont des données personnelles qu'à l'égard de ceux qui sont en mesure de réidentifier la personne, soit en principe les personnes qui sont en possession de la clé de correspondance si les données ont été correctement pseudonymisées.

⁵⁸ VALÉRIE JUNOD/BERNICE ELGER, Données codées, non-codées ou anonymes : des choix compliqués dans la recherche médicale rétrospective, in : Jusletter 10 décembre 2018, N 16.

⁵⁹ Message LRH (nbp 54), p. 7311.

⁶⁰ Sur les difficultés liées à l'anonymisation des données de santé, voir p. ex. ROLF H. WEBER/DOMINIC OERTLY, *Aus-höhlung des Datenschutzes durch De-Anonymisierung bei Big Data Analytics?*, in : Jusletter IT du 21 mai 2015.

⁶¹ Art. 3 let. h LRH.

⁶² Art. 26 al. 1 ORH.

⁶³ À ce sujet : JOTTERAND (nbp 56), N 17 ss. et références citées.

[33] Sous l'angle des lois générales relatives à la protection des données, la doctrine helvétique⁶⁴ et certains tribunaux cantonaux⁶⁵ ont aujourd'hui largement tendance à privilégier l'approche relative. C'est aussi la position retenue par le Conseil fédéral dans son message relatif à la future Loi fédérale sur la protection des données.⁶⁶ Si cette approche peut convaincre pour les traitements ordinaires de données personnelles, elle doit être rejetée pour les traitements de données tombant dans le champ d'application la LRH. Cette dernière établit en effet un régime juridique permettant la réutilisation facilitée de données personnelles à des fins de recherche (art. 32 à 34 LRH, cf. 3.3.1 *infra*). Or, ce régime règle directement les traitements de données codées, notamment en limitant le but pour lequel de telles données peuvent être traitées (un projet de recherche ou à des fins de recherche en général et non pour servir d'autres buts). Dans ce contexte, admettre que les données de recherche codées seraient anonymes du point de vue des destinataires aurait pour effet inadmissible de les exclure du champ d'application de la LRH (puisque la LRH ne s'applique pas aux traitements de données anonymes ou anonymisées). Les destinataires des données ne seraient alors plus liés par les exigences imposées par la LRH et pourraient par exemple s'affranchir d'adopter les mesures techniques et organisationnelles pour la conservation des données de recherche imposées par l'art. 43 LRH.⁶⁷ Il faut par conséquent considérer que les données codées restent des données personnelles à l'égard des destinataires si les activités concernées tombent dans le champ d'application de la LRH⁶⁸.

[34] De manière intéressante, certains auteurs proposent de sortir de l'approche purement binaire des théories absolue et relative. Selon JOTTERAND, l'élément central de l'analyse du caractère personnel ou non d'une donnée reposerait principalement sur l'environnement (dynamique) de celui qui traite les données.⁶⁹ Une donnée serait ainsi personnelle si le détenteur des données possède les données complémentaires pour réidentifier les données ou, s'il ne les possède pas, qu'elles existent et qu'il est probable qu'il puisse et veuille se les procurer pour réidentifier la personne concernée. À l'inverse, si les données complémentaires nécessaires pour une réidentification n'existent pas, ou que les données existent, mais qu'il apparaît peu probable que le détenteur des données tente de se les procurer pour réidentifier la personne, alors les données ne devraient pas être considérées comme personnelles du point de vue de ce dernier. Dans ce système, il appartient alors au responsable de traitement qui transfère des données pseudonymisées de prendre toutes les mesures nécessaires, notamment contractuelles, pour s'assurer que ce qui n'est pas autorisé pour le responsable de traitement ne devient pas permissible parce qu'il a

⁶⁴ P. ex. DAVID ROSENTHAL/SAMIRA STUDER/ALEXANDRE LOMBARD (pour la traduction), *La nouvelle loi sur la protection des données*, in : Jusletter 16 novembre 2020, N 20 ; DAVID ROSENTHAL/YVONNE JÖHRI, *Handkommentar zum Datenschutzgesetz*, Zurich 2008, Art. 3 let. a N 36 ; BEAT RUDIN, in : Bruno Baeriswyl/Kurt Pärli (édit.), *Datenschutzgesetz (DSG)*, Berne 2015, Art. 3 N 14 ; CÉLIAN HIRSCH/EMILIE JACOT-GUILLARMOD, *Les données bancaires pseudonymisées – Du secret bancaire à la protection des données*, RSDA 2020 p. 151 ss, p. 160–161.

⁶⁵ Handelsgericht zurichois, HGer ZH, HG150170, 30.03.2017, c. 5.3.5 (dans une affaire ayant pour objet la communication de données bancaires au fisc américain). Selon la décision rendue, il appartenait toutefois à la banque expéditrice des données de prouver qu'elle avait pris les mesures suffisantes contre la réidentification, ce qu'elle n'était pas parvenue à faire en l'occurrence. De plus, la décision soulignait que les technologies disponibles (notamment big data) laissaient peu de place pour une anonymisation irréversible.

⁶⁶ Message nLPD (nbp 56), p. 6640 : « La loi ne s'applique pas aux données qui ont été anonymisées si une réidentification par un tiers est impossible (les données ont été anonymisées complètement ou définitivement) ou ne paraît possible qu'au prix d'efforts tels qu'aucun intéressé ne s'y attellera. Cette dernière règle vaut aussi pour les données pseudonymisées. »

⁶⁷ Pour une argumentation plus étayée en ce sens : ERARD (nbp 56), p. 613 ss.

⁶⁸ Dans un sens similaire : ROSENTHAL (nbp 44), p. 169.

⁶⁹ JOTTERAND (nbp 56), N 46.

communiqué les données à un tiers qui ne dispose pas de la clé de réidentification, c'est-à-dire un tiers pour qui les données sont anonymes.⁷⁰ En ce qui concerne le contexte spécifique de la LRH, JOTTERAND relève que l'approche choisie ne devrait pas conduire à contourner les règles protectrices de la LRH. Il rejette néanmoins l'approche purement absolue et soutient que l'approche basée sur l'analyse de l'environnement de celui qui traite les données (en particulier le destinataire) est aussi transposable au régime de la LRH. Ainsi, selon cet auteur, une chercheuse qui reçoit des données codées est directement soumise à la LRH et peut uniquement mener sa recherche si l'ensemble des conditions posées par la LRH sont réunies. Toutefois, si la chercheuse peut raisonnablement considérer que les données reçues sont complètement anonymisées, elle peut mener sa recherche sans autre formalité.⁷¹ Cette approche est judicieuse en soi, mais son implication pratique semble limitée, du moins si les données émanent d'une institution de recherche helvétique. En effet, l'institution qui met à disposition les données de recherche (sous forme codée) ne devrait le faire que si l'ensemble des conditions légales pour la réutilisation sont réunies, en particulier l'approbation du projet par une commission d'éthique (art. 45 al. 1 LRH). Or, cette approbation ne peut être obtenue qu'en soumettant un protocole de recherche qui décrit précisément comment les données seront traitées, en particulier si elles feront l'objet d'un codage. Dans cette configuration, il paraît peu probable que la chercheuse entre en possession de données codées tout en considérant de bonne foi qu'elles sont complètement anonymisées.

[35] En synthèse, l'anonymisation des données personnelles peut certes constituer un moyen d'écartier les contraintes imposées par les législations sur la protection des données, mais elle doit être maniée avec la plus grande prudence, notamment en raison du lien étroit de ces données avec les personnes concernées. La recommandation de prudence vaut d'autant plus pour les données pseudonymisées/codées dans le contexte de la recherche sur l'être humain, qui doivent continuer à être considérées comme des données personnelles lorsque la LRH s'applique. Enfin, l'action d'anonymiser des données peut elle aussi faire l'objet de règles particulières (p. ex. information préalable et non-opposition de la personne concernée pour l'anonymisation de données génétiques à des fins de recherche selon l'art. 32 al. 3 LRH).

3.3. Réutilisation de données personnelles, en particulier à des fins de recherche sur l'être humain

3.3.1. Régime de la LRH

[36] Le concept de « réutilisation » de données est au cœur même du concept d'Open Data. Néanmoins, lorsque des données personnelles sont en jeu, une réutilisation de données peut porter atteinte au principe général de finalité. Selon ce principe, un traitement est conforme au principe de finalité lorsqu'on en informe la personne concernée, lorsque le traitement est prévu par la loi ou lorsqu'il ressort clairement des circonstances.⁷² Lors de chaque « réutilisation » de données personnelles,⁷³ il y a donc lieu de s'assurer que la réutilisation est conforme aux exigences légales

⁷⁰ JOTTERAND (nbp 56), N 61.

⁷¹ JOTTERAND (nbp 56), N 78.

⁷² Message nLPD (nbp 56), p. 6640.

⁷³ L'art. 24 ORH définit la réutilisation comme : « toute opération effectuée à des fins de recherche avec du matériel biologique déjà prélevé ou des données déjà collectées ». Sur la notion de réutilisation dans le contexte de la LRH : SHK HFG-RUDIN (nbp 55), Vorbemerkungen Art. 32–35, N 4 ss.

applicables, de telle sorte à respecter le principe général de licéité. Pour la réutilisation de données à des fins de recherche, il y a lieu de distinguer les utilisations qui tombent dans le champ d'application de la LRH (art. 2 LRH) des autres utilisations à des fins de recherche.

[37] Dans l'hypothèse où la réutilisation de données personnelles liées à la santé est envisagée dans un but de recherche sur l'être humain, les art. 32 à 34 LRH prévoient des conditions de réutilisation facilitées. Le système est toutefois établi selon des règles pour le moins complexes, qui dépendent du type de données (génétiques ou non génétiques) et de leur forme (non codées, codées ou anonymisées).⁷⁴ Ces règles de réutilisation, qui régissent également la réutilisation du matériel biologique, présupposent que les données aient déjà été collectées, par exemple dans un contexte de soins. Les données peuvent aussi avoir été collectées dans un projet de recherche préalable, dans le cas où la réutilisation envisagée excède le consentement donné initialement par la personne concernée.⁷⁵

[38] Sans entrer dans les détails, la systématique établie par la LRH exige qu'un consentement spécifique pour une recherche en particulier est nécessaire lorsque la réutilisation implique des données génétiques non codées (art. 32 al. 1 LRH).⁷⁶ Un consentement dit « général », à des fins de recherche – c'est-à-dire qui n'est pas limité à une recherche en particulier – est admis en droit suisse pour la réutilisation des données génétiques codées (art. 32 al. 2 LRH) et des données non génétiques non codées (art. 33 al. 1 LRH). En ce qui concerne la réutilisation des données non génétiques codées, la LRH admet qu'un droit d'opposition après information est suffisant (art. 33 al. 2 LRH). Néanmoins, pour des raisons éthiques, les hôpitaux qui proposent le consentement général renoncent à appliquer des règles distinctes pour les données non génétiques du seul fait qu'elles sont codées ou non codées.⁷⁷ En pratique, la réutilisation des données non génétiques codées est donc aussi sujette à l'obtention préalable du consentement général. Enfin, l'art. 34 LRH prévoit encore un régime d'exception dans les situations où les exigences requises par les art. 32 et 33 LRH ne sont pas satisfaites et que plusieurs conditions sont remplies, à savoir l'impossibilité ou les difficultés disproportionnées à recueillir le consentement de la personne concernée, l'absence de document attestant un refus de la personne concernée et la prépondérance de l'intérêt de la science sur celui de la personne concernée à décider de la réutilisation de ses données.⁷⁸

[39] Si les traitements de données anonymes tombent en principe hors du champ d'application de la LRH, l'art. 32 al. 3 LRH prévoit néanmoins un régime particulier pour l'opération d'anonymisation des données génétiques à des fins de recherche. Pour autant que l'anonymisation envisagée soit techniquement possible pour les données génétiques concernées, l'anonymisation à des fins

⁷⁴ Pour une critique du système : VALÉRIE JUNOD/BERNICE ELGER (nbp 58), N 16 ss.

⁷⁵ SHK HFG-RUDIN (nbp 55), Vorbemerkungen Art. 32–35, N 4.

⁷⁶ Le consentement peut aussi valablement porter sur plusieurs projets bien précis : SHK HFG-RUDIN (nbp 55), Art. 32, N 6.

⁷⁷ Voir p. ex. le modèle de formulaire de consentement général proposé par swissethics : <https://swissethics.ch/fr/templates/studieninformationen-und-einwilligungen>.

⁷⁸ Sur cette clause d'exception, voir en particulier : DOMINIQUE SPRUMONT/VLADISLAVA TALANOVA, *La recherche sans consentement : l'exceptionnelle exception*, in : Evelyne Clerc/Jean-Philippe Dunand/ Dominique Sprumont (édit.), *Alea jacta est : Santé! Mélanges en l'honneur d'Olivier Guillod*, Bâle 2021, p. 235 ss ; SUSANNE DRIESSEN/ANDRI CHRISTEN/PIETRO GERVAISONI, *Humanforschung, Weiterverwendung und informierte Einwilligung. Analyse zur Weiterverwendung von gesundheitsbezogenen Personendaten und biologischem Material sowie Anwendung von Artikel 34 HFG*, in : Jusletter 1^{er} février 2021.

de recherche est seulement autorisée si la personne concernée (ou son représentant) ne s'y est pas opposée après en avoir été informée.⁷⁹

[40] Le consentement général constitue un outil intéressant du point de vue du partage et de la réutilisation des données par des tiers. La loi permet en effet aux personnes concernées de consentir à l'utilisation de leurs données dans le strict cadre des recherches sur l'être humain, sans toutefois connaître à l'avance pour quelles recherches spécifiques leurs données seront utilisées. En autorisant le recours au consentement général, la LRH allège donc l'obligation de respecter strictement le principe de finalité. Dans la mesure où il s'agit d'un « consentement », celui-ci peut aussi être révoqué. Le cas échéant, l'art. 10 ORH prévoit que les données personnelles liées à la santé « doivent être anonymisées après avoir été analysées ». L'anonymisation n'est toutefois pas nécessaire si la personne concernée y renonce expressément ou s'il est évident depuis le début du projet de recherche que l'anonymisation n'est pas possible et que la personne en a été suffisamment informée au moment de participer au projet.

[41] Si le consentement général en tant que tel est aujourd'hui largement accepté en Suisse, il n'est pas exempt de critiques. D'abord, la formulation du consentement général peut varier d'hôpital en hôpital, ce qui est susceptible de créer des incertitudes dans le cadre de projets de recherche ayant recours à des données collectées auprès de différents hôpitaux. TALANOVA/SPRECHER, pour leur part, émettent plusieurs reproches à l'encontre des formulations actuellement utilisées, qui auraient entre autres pour conséquences d'affaiblir les droits des sujets de recherche.⁸⁰

[42] Tout en reconnaissant la nécessité d'obtenir le libre consentement des personnes concernées pour la réutilisation de leurs données à des fins de recherche, le Conseil fédéral a néanmoins récemment estimé que le consentement général devait faire l'objet d'une modernisation afin d'améliorer le potentiel de réutilisation et d'exploitation des données de santé.⁸¹ Selon une première option,⁸² la personne concernée pourrait consentir électroniquement à l'utilisation primaire de ses données et serait informée d'une possible réutilisation de ses données lors de la collecte initiale. Elle recevrait ensuite par voie électronique des informations sur les réutilisations envisagées et se verrait offrir la possibilité de s'opposer au cas par cas à de telles réutilisations (système d'*opt out*). Selon une seconde option,⁸³ présentée comme une alternative, la personne pourrait donner volontairement ses données nouvellement collectées pour une réutilisation libre non spécifiée, à la condition qu'aucun intérêt commercial ne soit poursuivi. L'utilisation des données serait autorisée à des fins qui excéderaient le cadre strictement limité de la recherche sur l'être humain. Notons qu'à l'aune du droit actuel, un consentement « dynamique », c'est-à-dire permettant aux personnes concernées de se voir proposer des projets de recherche et de décider au cas par cas si elles entendent mettre leurs données à disposition, est en principe déjà admissible.⁸⁴

[43] Une modernisation du consentement général peut certes être considérée dans une perspective d'amélioration de l'ouverture des données. Il convient néanmoins de faire preuve de pru-

⁷⁹ L'information qui doit être donnée dans ce cas est décrite à l'art. 30 ORH.

⁸⁰ VLADISLAVA TALANOVA/Franziska Sprecher, *Le consentement général : points à améliorer*, *Bulletin des médecins suisses*, 16 septembre 2020, consultable ici : <https://bullmed.ch/article/doi/bms.2020.19143>.

⁸¹ Rapport Humbel du 4 mai 2022 (nbp 30), p. 34.

⁸² Rapport Humbel du 4 mai 2022 (nbp 30), p. 39.

⁸³ Rapport Humbel du 4 mai 2022 (nbp 30), p. 40.

⁸⁴ SHK HFG-RUDIN (nbp 55), Art. 32, N 19. Pour une analyse détaillée de la question : JULIAN MAUSBACH, *Dynamische Einwilligung. Gedanken und Fragen zu einem neuen Einwilligungskonzept für die Forschung am Menschen*, in : Jusletter 27 janvier 2020.

dence. Suite à son introduction en 2014, le consentement général fait aujourd’hui l’objet d’un taux d’acceptation important chez les patients⁸⁵ et présente l’avantage de respecter la volonté des personnes qui refuseraient que leurs données soient utilisées à des fins de recherche, conformément au principe constitutionnel selon lequel un projet de recherche ne peut en principe être réalisé que si la personne y participant y a consenti (art. 118b al. 2 let. a Cst.).⁸⁶ En cas de révision ou de modernisation du consentement général, il sera primordial de veiller à ne pas vider le consentement de sa substance et de s’assurer que le consentement est compatible avec les standards éthiques internationaux, en particulier les règles imposées par la Déclaration de Taipei.

[44] Enfin, il faut souligner que les données personnelles qui ont été collectées à des fins de recherche ou qui ont fait l’objet d’une réutilisation au sens de la LRH ne peuvent ensuite être communiquées à d’autres fins qu’à des conditions particulières. Selon l’art. 41 LRH, une telle communication est seulement admissible si une base légale le prévoit ou si la personne concernée a donné son consentement éclairé « en l’espèce ». À la lumière de cette disposition, un consentement général ne peut donc pas couvrir d’autres finalités que celles de la recherche.

3.3.2. Régime hors de la LRH

[45] La réutilisation de données biomédicales peut viser d’autres buts que la recherche sur l’être humain au sens de la LRH, auquel cas cette dernière ne trouve pas application. Il en va par exemple ainsi de la réutilisation de données à des fins d’assurance qualité, à l’image de l’évaluation d’un processus appliqué dans le traitement d’hémorragies post-partum (introduit sur la base d’une étude) ou de l’évaluation de la sécurité d’une recommandation nationale fondée sur des études publiées.⁸⁷ Dans ce type de cas, il faut alors identifier les règles juridiques applicables, à commencer par les lois générales sur la protection des données personnelles.

[46] À la lumière de la nouvelle LPD par exemple,⁸⁸ une réutilisation de données est en principe admissible si elle est compatible avec les finalités pour lesquelles les données ont été collectées. Cette loi ne contient toutefois aucun critère à utiliser. Selon le Conseil fédéral, un traitement ultérieur est incompatible avec les finalités de collecte si la personne concernée peut légitimement le considérer comme « inattendu, inapproprié ou contestable ». ⁸⁹ Si l’on se réfère à l’art. 5 par. 4 let. b de la Convention internationale pour la protection des personnes à l’égard du traitement des données à caractère personnel (Convention 108+) et plus particulièrement à son rapport explicatif, il convient d’effectuer un test de compatibilité des finalités.⁹⁰ Ce dernier doit, entre autres, prendre en compte les liens entre la finalité initiale et la finalité ultérieure du traitement, le contexte de la collecte et les attentes raisonnables des personnes concernées, la nature des données, les conséquences du traitement ultérieur pour la personne concernée ou encore l’existence de garanties

⁸⁵ SPRUMONT/TALANOVA, (nbp 78), p. 249.

⁸⁶ À ce propos, avec une discussion prenant en compte les principes éthiques gouvernant la pratique de la recherche sur l’être humain : SPRUMONT/TALANOVA, (nbp 78).

⁸⁷ Sur la question de la distinction entre assurance de la qualité et recherche soumise à autorisation, voir les directives émises par swissethics, consultables ici : <https://swissethics.ch/fr/news/2020/02/04/qualitaetssicherung-oder-bewilligungspflichtige-forschung>.

⁸⁸ Art. 6 al. 3 nLPD.

⁸⁹ Message nLPD (nbp 56), p. 6645.

⁹⁰ Conseil de l’Europe, Convention 108 + Convention pour la protection des personnes à l’égard du traitement des données à caractère personnel, Rapport explicatif, 2018 N 49.

appropriées. Le test doit être mené à la lumière de l'ensemble des circonstances du cas d'espèce. On notera cependant que les données liées à la santé sont des données sensibles et que les traitements secondaires de telles données peuvent facilement se trouver en décalage avec les attentes raisonnables des personnes concernées. Dans ce contexte, il faut donc se montrer prudent lors de la réalisation du test de compatibilité.

[47] Une réutilisation de données personnelles incompatible avec le traitement initial n'est pas pour autant interdite, mais entraîne une atteinte à la personnalité des personnes concernées et doit alors reposer sur un motif justificatif.⁹¹ Ce dernier peut dans tous les cas reposer sur une base légale autorisant la réutilisation de données. En principe, la dérogation au principe général de finalité peut aussi être justifiée par le consentement de la personne ou, dans les cas où le traitement est effectué par une personne privée soumise à la LPD, par un intérêt prépondérant privé ou public.⁹²

[48] La LPD et la nLPD⁹³ contiennent des dispositions qui facilitent, sous certaines conditions, les traitements de données personnelles à des fins qui ne se rapportent pas à un ou plusieurs individus en tant que tels, mais à un groupe de personnes (traitements de données personnelles à des fins de recherche, de statistique ou de planification).⁹⁴ Dans ce type de situations, les données personnelles sont traitées sans rapport avec la personne concernée.⁹⁵ Les conditions légales posées pour ce type de réutilisations varient selon la loi applicable et la nature du responsable du traitement (personne privée ou organe public). Par exemple, l'art. 31 al. 2 let. e nLPD impose aux responsables de traitements privés qui souhaiteraient s'appuyer sur ce motif qu'ils s'assurent que les données soient anonymisées dès que possible, que ces dernières ne soient communiquées à des tiers que sous une forme ne permettant pas l'identification de la personne concernée (ou, si impossible, que des mesures soient prises pour que le destinataire ne traite les données qu'à des fins ne se rapportant pas aux personnes) et que les résultats soient publiés sous une forme ne permettant pas de réidentifier les personnes concernées. L'anonymat doit être préservé non seulement au moment de la communication ou de la publication, mais aussi tant que l'on peut raisonnablement supposer que des tiers auraient un intérêt à réidentifier les personnes concernées et qu'ils auraient ou pourraient se doter des moyens pour le faire.⁹⁶ Un organe public fédéral peut quant à lui traiter des données à des fins ne se rapportant pas à des personnes aux conditions posées par l'art. 39 nLPD, étant entendu que le traitement envisagé doit en principe reposer sur une base légale en vertu du principe de légalité (art. 34 nLPD). Toutefois, l'art. 39 al. 2 nLPD fait tomber l'exigence de la nature formelle de la base légale dans les situations visées par l'art. 34 al. 2 nLPD, notamment lorsqu'un traitement de données sensibles est en jeu. À noter que la plupart des lois cantonales sur la protection des données, applicables aux organes publics cantonaux,

⁹¹ Art. 12 al. 2 let. a *cum* art. 13 LPD ; art. 30 al. 2 let. a *cum* 31 nLPD. ROSENTHAL (nbp 44), p. 173.

⁹² Art. 13 LPD ; art. 31 nLPD.

⁹³ Art. 13 al. 2 let. e et art. 22 LPD ; art. 31 al. 2 let. e et 39 nLPD.

⁹⁴ Selon le Message du Conseil fédéral relatif à la nLPD, le champ d'application de ce motif justificatif ne semble pas se limiter aux seules activités de traitements liées à la recherche, à la planification ou à la statistique, mais s'étend à tout type de traitements de données personnelles à des fins qui ne se rapportent pas à des personnes. Message nLPD (nbp 56), p. 6692.

⁹⁵ DAVID ROSENTHAL/SAMIRA STUDER/ALEXANDRE LOMBARD (nbp 64), N 42 ; ROSENTHAL (nbp 44), p. 173.

⁹⁶ ROSENTHAL (nbp 44), p. 173.

contient aussi des dispositions visant des buts analogues, mais leur contenu et les conditions posées varient d'un canton à l'autre.⁹⁷

[49] Dans une optique d'Open Data, nous sommes d'avis que les dispositions légales permettant les traitements de données à des fins ne se rapportant pas à des personnes peuvent constituer une piste intéressante pour les recherches sans rapport avec les personnes concernées, par exemple à des fins d'assurance qualité. Pour que ce régime de réutilisation s'applique, il est néanmoins indispensable que la réutilisation ne tombe pas dans le champ d'application de la LRH (qui s'appliquerait alors en tant que loi spéciale). Par ailleurs, l'application de ce régime de réutilisation de données doit faire face à plusieurs obstacles lorsque des données sont en jeu :

- Les dispositions autorisant la réutilisation des données à des fins ne se rapportant pas à des personnes ne permettent pas de déroger aux obligations légales spéciales de confidentialité, à l'image du secret professionnel (art. 321 CP). En d'autres termes, un médecin doit se limiter à communiquer des données dans les situations prévues par l'art. 321 CP (consentement, disposition légale dérogeant spécifiquement au secret professionnel ou levée du secret par l'autorité cantonale compétente). Il ne peut donc pas s'appuyer sur les dispositions générales de réutilisation de données à des fins ne se rapportant pas à des personnes pour communiquer des données personnelles couvertes par le secret à des tiers. Une telle communication est néanmoins possible si les données sont anonymes (ou correctement pseudonymisées puisque la LRH ne trouverait pas application, cf. *supra* 3.2), de telle sorte que le destinataire n'est pas en mesure de réidentifier la personne concernée.⁹⁸
- Les données génétiques issues d'une analyse tombant dans le champ d'application de la Loi fédérale sur l'analyse génétique humaine (LAGH)⁹⁹ peuvent seulement être utilisées à une autre fin que le but de l'analyse initiale si la personne concernée a consenti librement et expressément à la réutilisation envisagée, même si les données se trouvent sous forme codée (art. 12 al. 1 LAGH). Une utilisation à une autre fin sous forme anonymisée n'est quant à elle possible que si la personne concernée en a été informée et qu'elle ne s'y est pas opposée (art. 12 al. 2 LAGH). En tant que disposition spéciale, l'art. 12 LAGH prime les dispositions autorisant les réutilisations de données à des fins ne se rapportant pas à des personnes.

[50] Enfin, comme la réutilisation ou la mise à disposition de données personnelles en vue de traitements ne se rapportant pas à des personnes constituent en principe des traitements de données, le responsable de traitement est généralement tenu d'offrir une information suffisante aux personnes concernées, en vertu de son obligation de respecter le principe de transparence.¹⁰⁰ Déterminer si une information est nécessaire ainsi que l'étendue de cette information sont des questions qui doivent être analysées au cas par cas.

⁹⁷ P. ex. art. 24 Loi vaudoise sur la protection des données personnelles, RS VD 172.65 ; art. 41 Loi genevoise sur l'information du public, l'accès aux documents et la protection des données personnelles, RS GE A 2 08.

⁹⁸ FRÉDÉRIC ERARD, *Le secret médical. Etude des obligations de confidentialité des soignants en droit suisse*, Zurich 2021, N 465.

⁹⁹ Loi fédérale sur l'analyse génétique humaine du 8 octobre 2004 (LAGH ; RS 810.12).

¹⁰⁰ Art. 14 LPD ; art. 19 nLPD.

3.4. Nécessité d'adopter un cadre contractuel en cas de partage de données personnelles

[51] Le responsable de traitement qui entend partager des données personnelles avec un tiers qui souhaiterait les utiliser à des fins de recherche biomédicale ne peut pas se limiter à vérifier que les conditions légales permettant un tel partage sont réunies. Comme le fournisseur est lui-même tenu au respect d'obligations imposées par la loi en matière de protection des données personnelles, notamment vis-à-vis des sujets de recherche, il doit prendre toutes les mesures nécessaires pour continuer à être en mesure d'assumer ses obligations après le partage des données avec un tiers. Par exemple, si le fournisseur des données personnelles n'est autorisé à traiter les données qu'à des fins particulières, il doit s'assurer que le destinataire des données ne les utilisera pas à d'autres fins, à moins que le droit applicable ne l'y autorise. Cela se fait usuellement par la conclusion d'un contrat entre le fournisseur et le destinataire des données. Les clauses contractuelles concernées peuvent alors figurer dans un contrat spécifique sur le partage de données (p. ex. *Data Transfer and Use Agreement*, *Data Transfer Agreement*, *Data Sharing Agreement*) ou être intégrées dans des contrats plus larges (p. ex. *Consortium Agreement*, *Research Agreement*). En Suisse, SPHN a développé des modèles de contrats pour faciliter le partage de données de recherche.¹⁰¹

[52] L'établissement des clauses nécessaires en vue d'un partage de données dans le contexte de la recherche sur l'être humain dépend de différents facteurs, à l'instar du statut des parties (responsable du traitement ou sous-traitant) ou de leur fonction (*data collector*, *data provider*, *data recipient*, *data owner*).¹⁰² Un examen détaillé des questions liées à l'établissement du cadre contractuel en vue du partage de données de recherche peut être consulté dans une contribution de JOTTERAND/ERARD,¹⁰³ à laquelle nous renvoyons. Nous nous limitons ici à énoncer brièvement et de manière non exhaustive quelques thématiques qui devraient faire l'objet d'un accord entre les parties :

- Cadre du traitement¹⁰⁴ : les parties devraient déterminer clairement les finalités du traitement par le destinataire, de telle sorte à s'assurer que les données ne seront pas utilisées pour des finalités illicites (p. ex. traitement à des fins d'assurance alors que les données doivent être réutilisées à des fins de recherche conformément aux art. 32 à 34 LRH).
- Garanties quant aux données : pour limiter ses risques sur le plan juridique, le destinataire des données doit s'assurer que les données qu'il va recevoir ont été collectées et sont partagées conformément à toutes les dispositions légales applicables. Cela se traduit par des garanties du fournisseur de données.
- Répartition des obligations imposées par le droit de la protection des données¹⁰⁵ : les responsables de traitements sont tenus de respecter toute une série d'obligations en vertu du droit de la protection des données, qu'il s'agisse de recueillir des consentements ou des au-

¹⁰¹ Disponibles ici : <https://sphn.ch/services/dtua>. Voir aussi : ALEXANDRE JOTTERAND/FRÉDÉRIC ERARD, Recherche sur l'être humain et données personnelles. Gestion des échanges et répartition des responsabilités, in : Jusletter 30 août 2021, N 90.

¹⁰² JOTTERAND/ERARD (nbp 101), N 92.

¹⁰³ JOTTERAND/ERARD (nbp 101).

¹⁰⁴ JOTTERAND/ERARD (nbp 101), N 94.

¹⁰⁵ JOTTERAND/ERARD (nbp 101), N 95.

torisations, de signaler à l'autorité une faille de sécurité ou d'assurer l'exercice des droits des personnes concernées (p. ex. droit d'accéder à ses données). Il est primordial que les parties à un partage de données s'entendent sur le partage de ces responsabilités.

- Mesures de sécurité¹⁰⁶ : comme le fournisseur des données est en principe directement responsable des traitements de données à l'égard des personnes concernées, il a tout intérêt à imposer contractuellement des mesures de sécurité ou de minimisation au destinataire des données, de telle sorte à limiter les risques d'un traitement négligent.

[53] La nécessité pour les parties (aussi bien le fournisseur que le destinataire) d'adopter un cadre contractuel pour le partage de données personnelles et de continuer à s'assurer du respect des droits des personnes concernées après le partage de données a pour effet de limiter le libre partage de données personnelles. Il est en effet nécessaire de garder un lien entre la personne concernée et l'utilisateur final des données personnelles, par exemple pour s'assurer que le retrait d'un consentement général conduit effectivement à écarter la mise à disposition des données concernées conformément aux exigences légales.¹⁰⁷ En d'autres termes, les exigences liées à la protection des données se combinent mal avec des réutilisations « en chaîne » de données personnelles, qui rendraient l'exercice des droits des personnes concernées particulièrement difficile. Par conséquent, si des données personnelles doivent être partagées dans une optique de science ouverte, elles devraient de préférence être partagées à partir d'un point d'accès unique garantissant un accès contrôlé.

3.5. Recours aux modèles et technologies améliorant la protection de la vie privée (Privacy Enhancing Technologies, PET)

[54] Sans se limiter au domaine de la recherche biomédicale, les contraintes posées aujourd'hui par les législations sur la protection des données ont conduit les milieux de recherche, en particulier ceux liés à l'informatique et à la cryptographie, à développer et proposer des modèles et des technologies qui doivent permettre de traiter et/ou partager des données personnelles de manière plus respectueuse de la vie privée. Les technologies concernées sont généralement désignées par la terminologie anglophone « *Privacy Enhancing Technologies* » (PETs).

[55] Comme la question du partage de données sensibles à grande échelle est une problématique centrale dans le secteur de la recherche biomédicale et que l'application des contraintes réglementaires peut avoir pour conséquences négatives de compliquer ou retarder la réalisation des projets de recherche, des efforts particuliers ont été fournis au cours des dernières années pour proposer des solutions de partage respectueuses de la vie privée dans ce secteur spécifique. Les méthodes suivantes, avec leurs avantages et leurs risques, sont par exemple décrites parmi d'autres par SCHEIBNER *et al.*¹⁰⁸ :

¹⁰⁶ JOTTERAND/ERARD (nbp 101), N 97.

¹⁰⁷ Art. 10 ORH.

¹⁰⁸ JAMES SCHEIBNER *et al.*, *Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies : Technical, Legal, and Ethical Synthesis*, *Journal of Medical Internet Research*, Vol. 23/2 (2021), 25 février 2021, consultable ici : <https://www.jmir.org/2021/2/e25120/>.

- *k-anonymity* : méthode d’anonymisation visant à réduire les possibilités de réidentification en établissant un modèle qui garantit que pour chaque combinaison d’identifiants, il existe au moins un certain nombre (k) d’individus qui partagent les mêmes attributs.¹⁰⁹ L’efficacité de cette méthode est toutefois limitée.
- Modèle d’analyses décentralisées : les données de recherche ne sortent pas des sites (p. ex. hôpitaux) et les analyses statistiques sont effectuées de manière décentralisée. Seuls les résultats sont envoyés vers l’institution centrale, qui les rassemble et procède à une méta-analyse.¹¹⁰
- Modèle d’analyse et d’apprentissage fédérés : il s’agit d’une évolution du modèle décentralisé, dans lequel les sites qui mettent les données à disposition entraînent sur leurs propres données un modèle d’apprentissage commun (global) mis à disposition par le site de recherche central. Les sites envoient ensuite les versions mises à jour de leur modèle au site de recherche central, qui met ensuite à jour le modèle global. Une fois reparamétré, le modèle global est renvoyé aux sites et le processus se répète jusqu’à la convergence du modèle global.¹¹¹
- Chiffrement homomorphique : ce type particulier de chiffrement permet le calcul directement sur des données chiffrées, sans impliquer leur déchiffrement.¹¹² Bien qu’il ne permette qu’un nombre limité de calculs, ce type de chiffrement peut aujourd’hui être utilisé dans des situations réelles.¹¹³

[56] Les développements de technologies en matière d’informatique et de cryptographie peuvent offrir des perspectives particulièrement attractives pour faciliter et permettre le partage de données (ou parfois de résultats d’analyses). Toutefois, le recours à de telles méthodes ou modèles n’exempte en rien les responsables de traitements de s’assurer qu’ils respectent l’ensemble des exigences légales qui leur sont applicables.

3.6. Excursus : projet d’espace européen des données en Europe

[57] Si la présente contribution se focalise essentiellement sur le droit suisse, la thématique traitée impose un petit écart vers l’Union européenne, qui compte parmi ses nombreux projets réglementaires en lien avec les données un projet qui doit précisément faciliter l’utilisation des données de santé : le projet d’Espace européen de données de santé (EHDS).¹¹⁴ Ce dernier vise principalement deux objectifs. Premièrement, il s’agira d’améliorer l’utilisation des données, soit l’accès numérique aux données de santé électroniques à caractère personnel ainsi que le contrôle sur celles-ci, tout en facilitant leur libre circulation (utilisation *primaire*). Deuxièmement, le projet a pour ambition de mettre en place un dispositif cohérent, fiable et efficace pour l’utilisation des

¹⁰⁹ SCHEIBNER *et al.* (nbp 108), p. 3.

¹¹⁰ SCHEIBNER *et al.* (nbp 108), p. 3–4.

¹¹¹ SCHEIBNER *et al.* (nbp 108), p. 4.

¹¹² SCHEIBNER *et al.* (nbp 108), p. 5.

¹¹³ Pour des exemples d’applications recourant à cette méthode de chiffrement : système MedCo (<https://medco.epfl.ch/>) et Tune Insight (<https://tuneinsight.com/>).

¹¹⁴ EHDS est l’acronyme anglais de l’Espace européen des données de santé (European Health Data Space).

données de santé à des fins de recherche, innovation, et élaboration des politiques et de réglementation (utilisation *secondaire*).¹¹⁵

[58] L’EHDS fait partie de la stratégie européenne pour les données, qui vise à créer un marché unique des données qui doit garantir la compétitivité mondiale et la souveraineté de l’Europe en matière de données.¹¹⁶ Il s’inscrit dans le prolongement du Règlement général sur la protection des données (RGPD)¹¹⁷, du Règlement sur la gouvernance des données¹¹⁸, du projet de règlement sur les données¹¹⁹ et de la directive sur la sécurité des réseaux et des systèmes d’information.¹²⁰ Une proposition de règlement européen sur les données de santé, ayant déjà fait l’objet d’une analyse d’impact et d’une consultation publique ouverte est actuellement en cours de discussion au Conseil de l’Union européenne.¹²¹

[59] Le chapitre IV de la proposition de règlement concerne la réutilisation des données de santé. L’art. 33 dresse une liste de données électroniques de santé que les détenteurs mettent à disposition à des fins d’utilisation secondaire. Il est précisé à l’art. 33 al. 4 que les données de santé électronique comportant des droits de propriété intellectuelle protégés et des secrets d’affaires d’entreprises privées protégés sont mises à disposition à des fins d’utilisation secondaire et que toutes les mesures nécessaires sont prises pour préserver ces droits.

[60] L’art. 34 de la proposition de règlement détermine les finalités pour lesquelles les données de santé électroniques peuvent être traitées à des fins d’utilisation secondaire. La liste inclut notamment des activités de développement et d’innovation pour les produits ou services contribuant à la santé publique ou à la sécurité sociale, ou à la garantie d’un niveau élevé de qualité et de sécurité des soins de santé, des médicaments ou des dispositifs médicaux (art. 34 al. 1 let. f), ainsi que la formation, le test et l’évaluation des algorithmes, entre autres dans les dispositifs médicaux, les systèmes d’intelligence artificielle et les applications de santé numériques, à la contribution à la santé publique ou à la sécurité sociale, ou à la garantie d’un niveau élevé de qualité et de sécurité des soins de santé, des médicaments ou des dispositifs médicaux (art. 34 al. 1 let. g).

[61] La proposition de règlement prévoit donc une obligation de mise à disposition de données de santé à des fins de réutilisation pour des buts autres que de la recherche. Ces deux dispositions ont fait l’objet de remarques, de la part du Comité européen de la protection des données et du Contrôleur européen de la protection des données dans leur avis conjoint sur la proposition de règlement. Ils ont exprimé leur préoccupation à ce sujet, estimant que les finalités n’étaient pas assez strictement délimitées.¹²² L’art. 35 de la proposition de règlement instaure une interdiction

¹¹⁵ https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_fr#pour-en-savoir-plus.

¹¹⁶ <https://digital-strategy.ec.europa.eu/fr/policies/strategy-data>.

¹¹⁷ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l’égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données, RGPD).

¹¹⁸ Règlement (UE) 2022/868 du Parlement Européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724.

¹¹⁹ Procédure 2022/0047/COD, Proposition de Règlement du Parlement européen et du Conseil fixant des règles harmonisées pour l’équité de l’accès aux données et de l’utilisation des données (règlement sur les données).

¹²⁰ Directive (UE) 2016/1148 du Parlement européen et du Conseil du 6 juillet 2016 concernant des mesures destinées à assurer un niveau élevé commun de sécurité des réseaux et des systèmes d’information dans l’Union.

¹²¹ Proposition de Règlement du Parlement européen et du Conseil relatif à l’espace européen des données de santé du 3 mai 2022, Procédure 2022/0140/COD.

¹²² EDPB-EDPS Joint Opinion 03/2022 on the Proposal for a Regulation on the European Health Data Space, du 12 juillet 2022, para. 85, consultable ici : https://edpb.europa.eu/system/files/2022-07/edpb_edps_jointopinion_202203_europeanhealthdataspace_en.pdf.

de demander l'accès et de traiter ces données en vue de certaines finalités. La liste des finalités visées inclut notamment la publicité, le marketing, le développement de produits ou services susceptibles de porter préjudice aux personnes, comme les drogues illicites, les boissons alcoolisées et les produits du tabac.

3.7. Constat intermédiaire et brèves considérations sur l'architecture juridique de la réutilisation des données à des fins de recherche

[62] Protection des données et ouverture des données ne sont pas incompatibles, mais les exigences liées à la protection des données personnelles soumettent l'accès et le partage de telles données au respect de conditions pour le moins contraignantes. En effet, même si le droit suisse offre des conditions de réutilisation facilitée aux données à des fins de recherche sur l'être humain (art. 32 à 34 LRH), le partage de telles données doit par exemple reposer sur l'autorisation d'une commission d'éthique de la recherche et faire l'objet d'un accord contractuel entre les parties concernées. L'anonymisation des données, si tant est qu'elle soit possible ou autorisée, permettrait certes d'exclure les traitements et les partages de telles données du champ d'application de la LRH et des lois sur la protection des données, mais une telle anonymisation n'est pas toujours adéquate pour poursuivre les objectifs de recherche sur l'être humain (nécessité de remonter au patient si nécessaire), implique parfois le respect d'exigences légales (information et absence de refus pour la réutilisation de données génétiques à des fins de recherche) et peut poser des difficultés pratiques de taille selon la nature des données (p. ex. anonymisation données génétiques). Comme les dispositions légales relatives à la recherche sur l'être humain (LRH) règlent également les traitements de données sous forme codée (ou pseudonymisée), elles ont donc un champ d'application particulièrement large, qui s'applique aussi aux traitements effectués par des personnes qui ne seraient pas en possession de la clé permettant de remonter aux personnes concernées. En conséquence, les données personnelles peuvent être partagées selon une logique d'Open Data, mais seulement dans un cadre qui assure un contrôle strict sur les données concernées.

[63] La nécessité de garder le contrôle sur les données personnelles peut engendrer des complications non négligeables pour l'ouverture des données dans le contexte de la recherche. À titre d'exemple, les projets les plus récents en génomique ou en médecine de précision requièrent des volumes importants de données et nécessitent de ce fait la collaboration de multiples institutions « fournisseuses de données ». Les jeux de données ainsi collectés ne tirent pas seulement leur valeur de leur volume, mais aussi du travail qui est effectué sur ces données (p. ex. opérations pour rendre les données interopérables, curation, interprétation) et il existe donc un intérêt important à les rendre accessibles à des fins ultérieures. Habituellement, ces projets présentent un caractère multicentrique et reposent sur une structure contractuelle de type « consortium ». Or, dans de telles configurations, la mise à disposition des jeux de données en vue de leur réutilisation par des tiers (qu'on appelle parfois « *third use* ») peut rapidement se révéler complexe si les données en question incluent des données personnelles. Le défi consiste entre autres à maintenir le lien entre les sujets de recherche concernés et les utilisateurs finaux de ces données, ce qui se traduit parfois par un enchaînement de contrats de partage de données. La situation se complexifie encore lorsqu'un des maillons de la recherche vient à disparaître, par exemple lorsque le consortium de recherche multicentrique ayant initialement collecté les données prend fin.

[64] L'exemple qui précède démontre qu'en pratique, une approche fondée sur la réutilisation en chaîne de données personnelles de recherche est peu viable, car elle engendre à terme des com-

plications importantes sous l'angle contractuel notamment. Pour faire face à ces défis, plusieurs options peuvent être envisagées, à commencer par le recours aux modèles ou méthodes respectueux de la vie privée (cf. *supra* 3.4). Néanmoins, si l'on souhaite offrir l'accès direct aux données, il semble alors opportun de favoriser une approche de type registre. En d'autres termes, il s'agit de permettre le dépôt de jeux de données dans une structure qui puisse mettre ces données à disposition de tiers.¹²³ Le registre peut alors servir de point de contact unique pour les tiers qui souhaiteraient accéder aux données concernées en vue de les réutiliser à des fins ultérieures de recherche. Cette approche permet aux fournisseurs de données de garder un contrôle direct sur les données et de s'assurer que les droits des sujets de recherche sont préservés, conformément au cadre légal en vigueur.

[65] Le développement d'un système de conservation des données figure d'ailleurs parmi les propositions émises dans le rapport Humbel.¹²⁴ Tout en écartant une solution où les données médicales disponibles en Suisse seraient stockées sur un support unique, le rapport soutient une solution de conservation des données décentralisée dans laquelle « les différentes collections de données de la même organisation seraient réunies et enregistrées dans un système local ».¹²⁵ À notre avis, une approche tendant vers un registre unique doit effectivement être rejetée, en particulier pour des raisons liées aux risques de sécurité trop importants qui en découleraient. Néanmoins, dans le cadre de ses réflexions à venir sur le sujet, le législateur devrait veiller à ne pas « imposer » une décentralisation trop contraignante, qui contribuerait à perpétuer la situation de silos de données qui existe aujourd'hui. Les projets de recherche dans le domaine biomédical, en particulier lorsqu'ils sont menés de manière multicentrique, doivent pouvoir bénéficier d'infrastructures fédérées dans lesquelles ils peuvent déposer les jeux de données générés en vue de les mettre à disposition de recherches futures, tout en respectant le droit applicable et les droits des sujets de recherche.

4. L'ouverture des données

[66] Comme l'a montré l'analyse menée dans la section précédente, les règles particulières applicables aux traitements de données personnelles limitent de manière importante l'ouverture de ces dernières. Les données traitées dans le contexte de la recherche biomédicale couvrent néanmoins un champ plus large que les données personnelles et peuvent aussi viser des données non personnelles, dont l'accès et l'utilisation peuvent être appréhendés par d'autres régimes juridiques.¹²⁶ Si ces derniers offrent généralement plus de flexibilité en matière de partage, ils peuvent aussi

¹²³ On notera que la conservation ou la saisie de données dans une banque de données constitue déjà une réutilisation au sens de la LRH, conformément à l'art. 24 let. c ORH. Voir aussi l'exemple des biobanques : SHK HFG-RUDIN (nbp 55), Vorbemerkungen Art. 32–35, N 7.

¹²⁴ Rapport Humbel du 4 mai 2022 (nbp 30), p. 41.

¹²⁵ Rapport Humbel du 4 mai 2022 (nbp 30), p. 41.

¹²⁶ Sur la question des droits d'accès et de réutilisation des données de recherche, voir YANIV BENHAMOU, *Intelligence artificielle : licence libre et gouvernance collective des données à travers l'altruisme des données et des data trusts*, in : RSDA 2021 p. 419 ss ; HÉLÈNE BRUDERER, *Les droits d'accès et de réutilisation des (bases de) données de recherche : de lege lata, de lege ferenda*, in : Jean-Philippe Dunand/Anne-Sylvie Dupont/Pascal Mahon (édit.), *Le droit face à la révolution 4.0*, Zurich 2019, p. 293 ss. Voir aussi EVA CELLINA, *La commercialisation des données personnelles – Aspects de droit contractuel et de protection des données*, Thèse Genève, Genève/Zurich/Bâle 2020 ; Rapport de l'Institut Fédéral de la Propriété Intellectuelle, Accès aux données non personnelles dans le secteur privé, 1^{er} mars 2021, consultable ici : <https://www.newsadmin.ch/newsadmin/message/attachments/66750.pdf>.

soulever certains obstacles pour l'Open Data. Après avoir examiné ces contraintes potentielles (4.1), nous analyserons certains outils contractuels favorisant la mise à disposition des données en libre accès, sans toutefois nous engager dans une démarche exhaustive (4.2).

[67] Afin d'éviter toute confusion, la terminologie de *données de recherche* sera employée dans la présente section pour désigner à la fois des données personnelles et non personnelles dont l'utilisation est envisagée à des fins de recherche biomédicale.

4.1. Les obstacles liés à l'ouverture

4.1.1. Régimes juridiques applicables

[68] Avant d'examiner les différents outils juridiques favorisant la réutilisation des données de recherche dans un contexte d'Open Access, il convient au préalable de procéder à un examen minutieux des potentiels obstacles susceptibles d'entraver leur ouverture en libre accès.¹²⁷ Outre les dispositions applicables en matière de protection des données personnelles, il faut prêter une attention particulière aux règles applicables qui peuvent découler d'autres régimes juridiques, en particulier celles prévues par le droit d'auteur¹²⁸ (LDA), la loi sur la concurrence déloyale¹²⁹ (LCD) et le droit des contrats. En raison de leur caractère protéiforme, les données de recherche font l'objet de multiples modes de présentation (imagerie médicale, tableaux, base de données, représentation et structure de biomolécules) grâce à la combinaison et à la mise en relation de données brutes (p. ex. mesures de tension artérielle) et d'informations plus complexes (annotation de variants multi-nucléotidiques). Celles-ci peuvent donc être appréhendées par différents régimes juridiques, dont les plus importants d'un point de vue pratique sont brièvement présentés ci-dessous.

[69] La protection offerte par le droit d'auteur s'ouvre dès lors que les conditions légales sont remplies, c'est-à-dire lorsque l'on est en présence d'une création de l'esprit, littéraire ou artistique, présentant un caractère individuel (art. 2 al. 1 LDA). Les données de recherche peuvent être exprimées sous différentes formes : tableaux, arbres phylogénétiques, représentations graphiques diverses ou encore sous forme de base de données. Elles peuvent être protégées au titre d'œuvre à contenu scientifique (art. 2 al. 2 let. d LDA), en tant qu'œuvre dérivée lorsqu'elles sont le fruit d'une reprise d'une œuvre préexistante (art. 3 al. 1 LDA) et enfin, au titre des recueils lorsqu'elles sont structurées sous forme de base de données (art. 4 al. 1 LDA). Aucune formalité n'est nécessaire pour que la protection du droit d'auteur s'ouvre : celle-ci s'applique dès la création de l'œuvre (art. 29 al. 1 LDA), même au simple stade de projet et aux parties d'œuvres dès lors que les conditions de protection sont remplies (art. 2 al. 4 LDA). Indépendamment de la qualification retenue (œuvre à contenu scientifique, recueil), la protection conférée par la LDA porte

¹²⁷ Le respect du principe « aussi ouvert que possible, aussi fermé que nécessaire » n'est pas intangible et peut souffrir de certaines exceptions. Voir p. ex. l'art. 39 al. 2 du Règlement (UE) 2021/695 du Parlement européen et du Conseil du 28 avril 2021 portant établissement du programme-cadre pour la recherche et l'innovation « Horizon Europe » et définissant ses règles de participation et de diffusion, et abrogeant les Règlements (UE) n°1290/2013 et (UE) n°1291/2013.

¹²⁸ Loi fédérale sur le droit d'auteur et les droits voisins du 9 octobre 1992 (LDA ; RS 231.1).

¹²⁹ Loi fédérale contre la concurrence déloyale du 19 décembre 1986 (LCD ; RS 241). Sur l'étude de la protection des bases de données par la LCD, PHILIPPE DUCOR, *Protection des bases de données et concurrence déloyale*, in : Alain Ragueneau (édit.), *Internet 2003 : travaux des journées d'étude organisées à l'Université de Lausanne les 21 mai et 26 novembre 2003*, CEDIDAC, Lausanne 2004, p. 153 ss.

uniquement sur la structure et la forme de la présentation des données et non sur les données en elles-mêmes.¹³⁰ Aucune protection ne peut être conférée aux idées et aux faits scientifiques, seule l'expression de l'idée – c'est-à-dire sa mise en forme – est protégeable au titre du droit d'auteur.¹³¹ Lorsque les contingences qui président la création d'une œuvre sont élevées, on admet en général que l'exigence d'individualité est remplie dès lors que les auteurs ont pu effectuer des choix personnels et qu'il ne s'agit pas d'un simple travail de routine. En présence d'une œuvre de l'esprit à caractère individuel, son auteur est alors investi d'un droit exclusif sur celle-ci lui conférant des prérogatives d'ordre moral et patrimonial.

[70] À toutes fins utiles, il convient de rappeler que la Suisse, contrairement à l'Union européenne, n'a pas introduit de droit *sui generis* pour les banques de données qui prohiberait toute extraction et/ou réutilisation du contenu d'une base de données ayant nécessité un investissement substantiel.¹³² En revanche, d'autres mécanismes de protection peuvent protéger une base de données, notamment le droit de la concurrence déloyale (art. 5 let. a et c et 6 LCD) ou le secret d'affaires (art. 4 let. c et 6 LCD).¹³³ Contrairement aux droits de propriété intellectuelle (p. ex. LDA), ces régimes juridiques ne confèrent aucun droit subjectif absolu : ils sanctionnent un comportement dénotant une certaine déloyauté ou l'accès indu au secret, exploité ou divulgué, par un tiers.

[71] Il convient encore de mentionner que les modalités d'accès et la réutilisation de données de recherche, indépendamment de la protection qui peut leur être conférée par le droit d'auteur, peuvent faire l'objet d'un contrat. Il est ainsi possible d'imposer des conditions d'utilisation spécifiques via une licence, voire d'autres obligations contractuelles telles qu'une obligation de confidentialité. C'est donc le principe de la liberté contractuelle qui prévaut dans ce cadre, sous réserve de l'observation d'éventuelles règles impératives susceptibles de trouver application (p. ex. la protection contre les engagements excessifs selon l'art. 27 al. 2 du Code civil). Contrairement aux droits de propriété intellectuelle, les contrats déploient toutefois uniquement leurs effets entre les parties et non à l'égard des tiers, en raison du principe de l'effet relatif des conventions.

[72] L'analyse succincte des différents régimes juridiques applicables aux données de recherche montre ainsi que ces dernières font l'objet d'une protection relativement complète bien qu'éparse permettant de garantir un réel contrôle, voire une « quasi-propriété ».¹³⁴

¹³⁰ Pour une étude de la protection des bases de données par le droit d'auteur voir, NATHALIE TISSOT, *Bases de données et droit d'auteur*, in : Alain Ragueneau (édit.) *Internet 2003 : travaux des journées d'étude organisées à l'Université de Lausanne les 21 mai et 26 novembre 2003*, CEDIDAC, Lausanne 2004, p. 187 ss. Voir aussi : Groupe Suisse de l'AIPPI, sic! 2020, p. 655 ss.

¹³¹ Pour un rappel de ce principe, ATF 133 II 306 c. 3, JdT 1988 I 304. Cf. BRUDERER (nbp 126), p. 298–299 et les références citées.

¹³² Directive 96/9/CE du Parlement européen et du Conseil du 11 mars 1996 concernant la protection juridique des bases de données.

¹³³ Pour une définition du secret d'affaires : arrêt du Tribunal fédéral du 8 juin 2010 4A_195/2010, c. 2.2. Voir aussi à ce sujet, RALPH SCHLOSSER, *Les secrets économiques dans les relations de travail, les collaborations et les procès civil*, in : Jacques de Werra (édit.), *La protection des secrets d'affaires*, p. 65 ss.

¹³⁴ Rapport IPI (nbp 126), section 5.1.3, p. 23.

4.1.2. Exemples d'obstacles

[73] Après ce bref survol des régimes juridiques applicables aux données de recherche, nous proposons d'analyser plus en détail trois types d'obstacles susceptibles de se dresser dans le contexte de l'ouverture des données à des fins de recherche. Il s'agit des obstacles potentiels suivants : (1) l'existence de contenu et/ou des connaissances préexistantes appartenant à des tiers dans les données que l'on souhaite diffuser de manière ouverte, (2) la titularité des données générées dans le cadre d'un projet de recherche et (3) l'application en concours des différents régimes juridiques applicables aux données.

[74] Le premier point d'analyse consiste à déterminer si les données de recherche nouvellement générées dans le cadre d'un projet de recherche, que l'on peut qualifier d'*Output Data* indépendamment de leur mode de présentation et du type de données en cause, intègrent du contenu et/ou des connaissances préexistantes appartenant à des tiers, c'est-à-dire des données initiales (*Input Data*).¹³⁵ Dans l'hypothèse où les *Output Data* intègrent des *Input Data*, il faut établir si leur réutilisation est soumise à des conditions contractuelles spécifiques telles qu'une obligation de garder le secret ou si elles font l'objet d'un droit subjectif absolu tel que le droit d'auteur par exemple. La mise en circulation d'*Output Data* intégrant sciemment du contenu appartenant à des tiers sans leur autorisation est susceptible de constituer une violation contractuelle et/ou d'obligations imposées par la loi. En l'absence de disposition claire spécifiant les conditions de réutilisation de contenu appartenant à des tiers et/ou de consentement explicite de ces derniers, la prudence commande de leur demander une autorisation écrite validant la réutilisation envisagée.¹³⁶ Selon les modalités d'exploitation applicables, les activités d'agrégation d'*Input Data* peuvent aussi impliquer d'observer des règles particulières, notamment en matière de citation. Il est donc primordial d'identifier les conditions contractuelles encadrant l'accès et la réutilisation de données initiales appartenant à des tiers aux fins d'apprécier l'étendue des obligations qui pèseront sur les titulaires des *Output Data* dans une perspective d'ouverture des données.

[75] Le second aspect à considérer concerne la titularité des *Output Data*. Au cours de la réalisation d'un projet de recherche, les *Output Data* sont généralement le fruit d'un travail collaboratif de plusieurs partenaires. Dans un tel contexte, les données de recherche peuvent par exemple être générées de la façon suivante : un premier partenaire met à disposition ses *Input Data* exprimées sous différentes formes (base de données, schémas et représentations scientifiques diverses), puis un deuxième partenaire se charge de les annoter et de les restructurer selon une méthodologie précise. Pour finir, un troisième partenaire réalise des opérations d'appariement des données permettant d'obtenir les données de recherche (*Output Data*). Dans une telle constellation de contributions successives, et en l'absence d'accord entre les parties impliquées, la maîtrise des

¹³⁵ Sur la distinction entre *Input Data* et *Output Data* et les enjeux soulevés en termes de propriété intellectuelle : YANIV BENHAMOU, *Big Data and the Law : a holistic analysis based on a three-step approach – Mapping property-like rights, their exceptions and licensing practices*, *Swiss Review of Business law* 2020 p. 393 ss, p. 404–407 ; RICHARD KEMP, *Legal aspects of managing Big Data*, *Computer Law & Security Review : The International Journal of Technology Law and Practice* 2014, Vol. 30 p. 482 ss, p. 489–490.

¹³⁶ FLORENT THOUVENIN, *Un droit de propriété sur les données en suisse ?*, in : Jacques de Werra (édit.), *Propriété intellectuelle à l'ère du Big Data et de la Blockchain*, Zurich 2017, p. 113–114, relève, à juste titre, que le processus de génération de données est davantage susceptible de faire l'objet de créations similaires, partant il est plus difficile de revendiquer la violation d'un droit exclusif préexistant. L'auteur en tire la conclusion que le titulaire d'un droit d'auteur ne peut interdire à un tiers d'utiliser une œuvre totalement ou largement identique à sa propre œuvre si ce dernier est en mesure de démontrer, qu'il a créé son œuvre de façon totalement indépendante.

données et leur exploitation peut susciter passablement de débats.¹³⁷ En outre, si la protection instituée par la LDA s'applique aux *Output Data*, il faut également déterminer si l'on se trouve en présence d'une œuvre dérivée (art. 3 al. 1 LDA) ou d'une œuvre commune distincte (art. 7 al. 1 LDA). Dans le premier cas, l'exploitation des *Output Data* nécessite l'accord du titulaire des données initiales réutilisées (art. 3 al. 4 LDA) tandis que dans le second cas, les titulaires sont libres de les utiliser dans les limites prévues par la LDA. La question de la titularité des *Output Data* et les conséquences qui en découlent méritent donc une attention toute particulière. L'absence d'accord en amont des parties impliquées dans le projet de recherche peut rapidement susciter des désaccords qui auront pour conséquence de nuire à la pleine exploitation des données, en particulier dans une perspective d'Open Access. Par ailleurs, la détermination de la titularité des *Output Data* présente un intérêt pratique important, car seuls les titulaires des *Output Data* sont habilités à délimiter l'étendue des droits d'usage conférés aux futurs utilisateurs et sont responsables du contenu publié en libre accès. Il est donc recommandé de porter une attention particulière à la rédaction du plan de recherche en veillant à définir clairement l'objectif et les résultats à générer (*Foreground IP*) ainsi que les différents contributeurs et/ou sous-traitants, notamment en identifiant précisément leurs contributions respectives et les connaissances préexistantes qu'ils mettront à disposition (*Background IP*). La flexibilité caractéristique du droit des contrats peut donc permettre de développer des solutions adaptées et originales pour régler les revendications des participants impliqués dans un projet de recherche à l'égard des *Output Data*, mais aussi de définir les modalités de mise en circulation de ces dernières.

[76] Le troisième point d'analyse consiste à évaluer les conséquences liées à l'application parfois simultanée de différents régimes légaux applicables¹³⁸ à l'égard des *Output Data*. Les implications seront sensiblement différentes selon qu'il est question de réutiliser des données sensibles de patients ou des données non personnelles, par exemple des données liées au séquençage génétique d'un virus tel que le SARS-CoV-2. Comme le relève à juste titre le rapport de l'IPI¹³⁹, il convient donc de procéder à un examen systématique du contenu des données préalablement à toute utilisation future, notamment pour déterminer si une réidentification des personnes est possible. En pratique toutefois, le caractère protéiforme des données peut rendre difficile l'examen de leur contenu et révéler la présence de données mixtes protégées par différents instruments juridiques. Dans tous les cas, les règles applicables en matière de protection des données personnelles (p. ex. LPD, LRH) prennent le pas sur les autres législations applicables aux données telles que le droit d'auteur.¹⁴⁰ Lorsque cela est possible et si cela est utile pour la diffusion des données, il convient d'isoler les données personnelles des données non personnelles.

[77] En résumé, la réutilisation d'*Input Data* et la création d'*Output Data* peuvent générer des obstacles susceptibles d'avoir une certaine incidence dans une perspective d'ouverture des résultats de recherche. Afin d'évaluer si une mise à disposition en libre accès est possible et de définir ses modalités, il est par conséquent primordial de bien évaluer le contenu des données en cause et d'identifier les différents régimes légaux applicables.

¹³⁷ Sur la question de la titularité des biens numériques en général et les règles applicables, YANIV BENHAMOU/LAURENT TRAN, *La circulation des biens numériques : de la commercialisation à la portabilité*, sic! 2006 p. 571 ss, p. 582-583.

¹³⁸ Sur ce sujet, BENHAMOU (nbp 135), p. 413-415; BRUDERER (nbp 126), p. 295-298.

¹³⁹ Rapport IPI (nbp 126), section 4.1, p. 20.

¹⁴⁰ En ce sens, voir BENHAMOU (nbp 138), p. 429 qui propose aussi cette approche lorsqu'il n'est pas possible d'opérer un tri des données.

4.2. La mise en circulation des données : licences & waiver

[78] En droit suisse, il n'existe pas de droit de propriété sur les données en tant que telles.¹⁴¹ Pour autant, cela n'exclut pas *per se* la possibilité de régler les modalités de leur accès et de leur utilisation contractuellement, en particulier à travers l'utilisation de licences ou de conditions générales. Étant entendu que les données de recherche peuvent s'exprimer sous différentes formes (bases de données, tableaux, images, schémas) en fonction des différentes disciplines (génomique, transcriptomique, etc.), la présente section se limite à présenter de manière générale les outils juridiques compatibles avec la démarche du libre accès et à mettre en évidence certaines problématiques pertinentes pour la pratique.

[79] Le mouvement du libre accès vise à ce que les titulaires des résultats de recherche conservent leurs droits sur les résultats de recherche et accordent des droits d'utilisation.¹⁴² Dans une perspective d'ouverture des données en Open Access, il est recommandé de recourir au mécanisme de licence. La pratique du *licensing* est un processus classique de valorisation des actifs immatériels qui s'opère via la figure du contrat de licence couvrant à la fois les biens immatériels protégés par des droits subjectifs absolus (p. ex. droit d'auteur, protection des données personnelles), mais aussi ceux qui ne sont pas protégés par un monopole juridique.¹⁴³ La licence est un contrat innommé *sui generis* par lequel le propriétaire d'un actif immatériel (donneur de licence) confère à un tiers (le preneur de licence) l'usage et la jouissance d'un bien immatériel moyennant l'éventuel paiement d'une redevance. La validité de ce type de contrat n'est pas subordonnée à l'observation d'une exigence de forme particulière. Il peut être écrit, oral, ou résulter d'actes concluants. L'étendue des droits d'exploitation est déterminée en fonction du type de licence¹⁴⁴ concédée par l'auteur : elle peut être simple ou exclusive, payante ou gratuite, totale ou partielle, pour une période déterminée ou indéterminée, limitée géographiquement, assortie ou non de la faculté d'accorder des sous-licences, ou encore être octroyée pour une finalité déterminée (p. ex. pour une utilisation à des fins académiques et/ou commerciales). Sous réserve du respect des dispositions impératives (p. ex. art. 27 al. 2 CC, art. 19 al. 1 LDA), c'est donc la liberté contractuelle qui prévaut dans ce domaine. En vertu de l'effet relatif des conventions, le contrat de licence n'a pas d'effet absolu (*erga omnes*), il déploie ses effets uniquement entre les parties concernées (*inter partes*), c'est-à-dire entre le donneur et le preneur de licence.

[80] L'ouverture des données de recherche en libre accès peut ainsi être formalisée en établissant un contrat de licence. Pour ce faire, il est possible de préciser les conditions d'accès et de réutilisation des données dans une section dédiée, que ce soit dans un contrat ou dans des conditions générales, ou de recourir à des modèles de licence gratuits et standardisés tels que les licences

¹⁴¹ Pour un examen approfondi de cette thématique : THOUVENIN (nbp 136), p. 61–62 et les références citées. La même approche est appliquée au niveau européen : Règlement (UE) 2018/1807 du Parlement européen et du Conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne, consultable ici : <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32018R1807&from=El>.

¹⁴² Cf. la Déclaration de Berlin, qui prévoit explicitement en ce sens que : « Les contributions en libre accès doivent satisfaire deux conditions : 1. Leurs auteurs et les propriétaires des droits afférents concèdent à tous les utilisateurs un droit gratuit, irrévocable et mondial d'accéder à l'œuvre en question, ainsi qu'une licence les autorisant à la copier, l'utiliser, la distribuer, la transmettre et la montrer en public, et de réaliser et de diffuser des œuvres dérivées, sur quelque support numérique que ce soit et dans quelque but responsable que ce soit, sous réserve de mentionner comme il se doit son auteur [...] ».

¹⁴³ THOMAS PROBST, *Der Lizenzvertrag : Grundlagen und Einzelfragen*, in : Jusletter 2 septembre 2013, N 10 ss ; CELLINA (nbp 126), p. 156 et les références citées.

¹⁴⁴ Pour une présentation générale des différents types de licence, PROBST (nbp 143), N 25 ss.

*Creative Commons*¹⁴⁵ ou les licences développées par l'*Open Knowledge Foundation*.¹⁴⁶ Ces deux approches ne sont pas exclusives et il est parfaitement envisageable de modifier des dispositions des licences *Creative Commons*.¹⁴⁷ Au moment de procéder au choix d'une forme de licence ou d'une autre, il convient néanmoins de vérifier l'existence d'éventuelles règles internes propres à l'institution de recherche concernée et, le cas échéant, de s'y conformer.

[81] Dans les cas de mise à disposition de données de recherche via un contrat de licence, il faut définir avec précision le champ d'application de la licence concédée. Afin d'être en accord avec la philosophie du libre accès, les modalités de la licence octroyée doivent selon nous comporter les caractéristiques suivantes : être non exclusive, gratuite et permettre à tous les utilisateurs d'utiliser, copier, modifier et de distribuer le contenu sous licence, y compris sous une forme dérivée. Il est par ailleurs utile et recommandé de préciser les éventuelles règles de citation applicables, de spécifier l'utilisation d'*Input Data* ou de contenu appartenant à des tiers qui tombent hors du champ de la licence concédée, de définir les règles de responsabilité applicables lors de la réutilisation du contenu, de mentionner un point de contact, mais aussi de préciser le droit applicable et le for compétent en cas de litige ainsi que les conditions de résiliation.¹⁴⁸

[82] En cas d'utilisation de licences publiques, les licences *Creative Commons* doivent être privilégiées. Elles constituent un instrument efficace permettant de déterminer aisément l'étendue des droits conférés aux utilisateurs à l'égard d'une œuvre protégée par le droit d'auteur.¹⁴⁹ Ces licences publiques jouissent d'une certaine popularité et leur utilisation est recommandée par plusieurs institutions publiques.¹⁵⁰ Les titulaires des droits peuvent définir les modalités d'exploitation de leur contenu en choisissant parmi les six catégories de licence *Creative Commons*. Sans entrer dans les détails de chaque licence, on peut néanmoins présenter brièvement leurs caractéristiques respectives. La licence la plus permissive est la licence d'attribution (BY). Elle offre à chaque utilisateur la liberté de réutiliser les données à condition toutefois de créditer les titulaires originaux et d'indiquer si des changements ont été apportés au contenu original. La licence du partage à l'identique (SA, *Share Alike*) oblige à citer la source et à partager l'œuvre nouvelle-

¹⁴⁵ Pour une présentation générale des différentes licences Creative Commons : MICHEL JACCARD/EVA CELLINA, *Les Creative Commons, avenir du droit d'auteur ?*, SJ 2017 II p. 229 ss, p. 235–239.

¹⁴⁶ Voir à ce sujet : <https://opendatacommons.org/licenses/>. À la différence des licences Creative Commons, les licences Open Data Commons couvrent seulement les bases de données mais pas leur contenu. Le préambule des trois licences Open Data Commons précise en ce sens que : « *Databases can contain a wide variety of types of content (images, audiovisual material, and sounds all in the same database, for example), and so this license only governs the rights over the Database, and not the contents of the Database individually.* » Partant, il est donc tout à fait envisageable d'appliquer à la fois une licence Open Data Commons à une base de données et une licence CC BY à son contenu si celui-ci est protégé au titre du droit d'auteur.

¹⁴⁷ Voir néanmoins les réserves mentionnées dans la FAQ des Creative Commons, <https://creativecommons.org/faq/#can-i-change-the-license-terms-or-conditions>.

¹⁴⁸ À titre d'exemple, l'accès et l'utilisation par des tiers des différentes données scientifiques disponibles sur le site internet GISAID font l'objet d'un *Data Access Agreement*. Chaque utilisateur se voit octroyer une licence non exclusive, mondiale, gratuite et non transférable pour accéder et utiliser la base de données EpiFlu. Voir plus en détail <https://gisaid.org/terms-of-use/> ; pour un autre exemple similaire : <https://www.surechembl.org/terms/>.

Voir aussi les recommandations proposées par BENHAMOU (nbp 138), p. 412–413.

¹⁴⁹ La section avertissement des licences Creative Commons insiste sur le fait qu'elles sont faites pour être utilisées « par les auteurs et les titulaires de droits dans la limite des lois et règlements en vigueur ». Cf. JACCARD/CELLINA (nbp 145), p. 240.

¹⁵⁰ Voir p. ex. les recommandations de l'infrastructure *Open Access for Infrastructure for Research in Europe* (OpenAIRE), consultables ici : <https://zenodo.org/record/2574619>, qui recommandent l'utilisation des licences Creative Commons CC BY 4.0 lorsque le matériel constitue une œuvre au sens de la LDA et CC0 pour les données et jeux de données qui ne sont pas structurés sous forme de base de données.

ment créée à partir d'œuvre préexistante sous la même licence CC BY-SA. La licence interdisant l'utilisation commerciale (CC BY-NC) prévoit les mêmes droits que la licence d'attribution, mais exclut toute utilisation à des fins commerciales. La licence *Non derivative* (CC BY-ND) permet un usage à titre commercial, mais interdit toute modification ou création d'œuvres dérivées. Pour finir, il convient d'évoquer les caractéristiques de la licence *Creative Commons Zéro* (CC0). Il ne s'agit pas à proprement parler d'une licence étant donné que son application a pour effet de placer le contenu dans le domaine public, leurs créateurs renonçant ainsi à l'exercice de leur droit d'auteur (*waiver*). Le contenu mis en circulation sous licence CC0 peut donc être modifié et réutilisé librement sans aucune restriction. Il s'agit d'un instrument efficace pour lever tout éventuel doute sur la question de savoir si la protection du droit d'auteur s'applique ou non étant donné que les auteurs renoncent à exercer leur droit exclusif à l'égard des biens immatériels visés par la licence dans les limites de ce qui est prévu par la loi.¹⁵¹ La licence CC0 présente aussi l'avantage de remédier à la problématique de l'*attribution stacking*, c'est-à-dire la nécessité de créditer un grand nombre de jeux de données réutilisées comprenant chacune des exigences propres en matière de citation. L'application d'une licence CC0 offre donc aux utilisateurs la possibilité d'agréger plusieurs jeux de données (*datasets*) sans devoir se préoccuper de créditer individuellement leur titulaire. Bien que les termes de la licence CC0 n'imposent aucune obligation de citation concernant le contenu réutilisé, il reste opportun de prévoir une règle de citation, car des considérations éthiques peuvent s'appliquer, notamment dans le milieu académique. Par ailleurs, les publications dans le milieu de la recherche constituent toujours un enjeu à haut potentiel pour les personnes impliquées dans la recherche biomédicale. On relèvera encore qu'il est important d'identifier soigneusement le contenu soumis aux licences Creative Commons et de mettre clairement en évidence d'éventuels éléments qui ne tomberaient pas dans le champ d'application de la licence choisie.¹⁵² Il reste à noter que les licences Creative Commons sont irrévocables et ne comportent aucune clause spécifique concernant l'intégration des données personnelles.¹⁵³

[83] Les licences Creative Commons ne doivent pas être utilisées pour la mise à disposition et le partage de données personnelles puisqu'elles conduiraient à une perte de contrôle contraire aux législations relatives aux traitements de données personnelles. Cela vaut de la même manière pour les traitements de données effectués dans le contexte de la recherche sur l'être humain et soumis à la LRH : la libre diffusion des données personnelles codées empêcherait non seulement les sujets de recherche à faire valoir leurs droits sur leurs données, mais aurait aussi pour conséquence qu'il deviendrait impossible de garantir que ces données ne seraient effectivement réutilisées qu'à des fins de recherche.¹⁵⁴ Comme énoncé précédemment (cf. *infra* section 3.4), il est nécessaire de recourir à des instruments contractuels spécifiques (p. ex. *Data Sharing Agreement*) et de spécifier le champ de la licence octroyée à l'égard de données de recherche intégrant des données personnelles sensibles. Autrement dit, les licences Creative Commons ne peuvent

¹⁵¹ Sur la question de la renonciation au droit d'auteur : DENIS BARRELET/WILLI EGLOFF, *Le nouveau droit d'auteur, Commentaire de la loi fédérale sur le droit d'auteur et les droits voisins*, 4^e éd., Berne 2021, Art. 16 LDA, N 18–19.

¹⁵² Sous licence CC-BY, on peut réutiliser librement le contenu protégé, mais les conditions de la licence imposent d'identifier le contenu réutilisé (y compris sous une forme modifiée) et d'identifier les auteurs de la contribution originale (art. 3(a)(1)).

¹⁵³ Cf. art. 2(b)(1) de la licence Creative Commons CC BY stipule que : « Les droits moraux, tels que le droit à l'intégrité de l'œuvre, ne sont pas accordés par la présente Licence publique, ni le droit à l'image, ni le droit au respect de la vie privée, ni aucun autre droit de la personnalité ou apparentée; [...] ».

¹⁵⁴ Voir aussi, JANE KAYE (nbp 40), p. 421–423.

être envisagées que si les données de recherche ont fait l'objet d'une anonymisation complète ou si elles ne contiennent aucune donnée personnelle.¹⁵⁵

[84] Compte tenu des éléments qui précèdent et sous réserve qu'aucune contrainte contractuelle et légale n'empêche une mise à disposition en libre accès, les licences publiques Creative Commons CC BY et CC0 restent les outils juridiques les plus adaptés pour mettre à disposition du contenu en Open Access. La licence CC BY-SA est aussi compatible avec le paradigme du libre accès, mais les utilisateurs qui ne souhaitent pas subir l'effet *copyleft*¹⁵⁶ préféreront renoncer à réutiliser le contenu en cause afin de ne pas être contraints d'appliquer les termes de la licence CC BY-SA. En revanche, les licences *Non Derivative* (CC BY-ND) et *Non Commercial* (CC BY-NC) apparaissent moins compatibles avec les exigences du libre accès. La condition *Non Derivative* interdit toute combinaison et agrégation avec de nouvelles données et la licence *Non Commercial* reste délicate à mettre en œuvre, car il n'est pas toujours évident de déterminer de ce qui relève ou non d'un usage commercial.¹⁵⁷

[85] En synthèse, les mécanismes du *waiver* et de la licence constituent des outils juridiques pertinents pour mettre en circulation des données de recherche en Open Access, à condition qu'elles ne contiennent pas de données personnelles. S'agissant du contrat de licence, il faut veiller à ce que ses modalités soient compatibles avec la philosophie du libre accès. Il sera toujours possible de s'inspirer des dispositions contractuelles mentionnées dans les licences *Creative Commons* CC0, CC BY et CC BY-SA pour arriver à une étendue d'exploitation similaire voire identique. Dans un souci de clarté et pour ne pas entraver inutilement la réutilisation du contenu en libre accès, il est plus prudent de ne pas modifier les dispositions d'une licence Creative Commons, car cela pourrait susciter de la confusion dans l'esprit du public sur l'étendue de l'exploitation offerte. Dans tous les cas, il est primordial de préciser le champ d'application de la licence octroyée, car en l'absence de mention spécifique à ce sujet, la prudence recommande d'appliquer les règles prévues par la LDA, c'est-à-dire de demander l'autorisation des titulaires du contenu avant toute réutilisation. À noter que dans un contexte ayant pour finalité principale la recherche scientifique, les utilisateurs peuvent par ailleurs s'appuyer sur l'exception légale de fouille de textes et de données pour reproduire un contenu disponible sans autorisation préalable de leurs auteurs.¹⁵⁸ Toutefois, s'ils souhaitent exploiter un tel contenu d'une manière qui excède la fouille de texte ou de données, ils doivent obtenir préalablement l'autorisation de ses auteurs.

5. Application pratique et check-list

[86] Les analyses menées jusqu'ici ont montré que la mise à disposition, le partage et la réutilisation des données à des fins de recherche biomédicale, dans une optique d'Open Data, pouvaient faire l'objet de diverses limitations de nature juridique et/ou impliquaient l'adoption de certaines

¹⁵⁵ Dans le même sens, Andreas Wiebe/Nils Dietrich (édit.), *Open Data Protection : Study on legal barriers to open data sharing – Data Protection and PSI*, Göttingen, Germany 2017, p. 199.

¹⁵⁶ L'effet *copyleft* est associé aux licences qui autorisent tout utilisateur à utiliser, copier, modifier et distribuer des versions modifiées d'une œuvre sous réserve que ladite œuvre, y compris les versions modifiées, soient distribuées sous les mêmes conditions.

¹⁵⁷ Sur ce point, voir JACCARD/CELLINA (nbp 145), p. 247–249.

¹⁵⁸ Art. 24d LDA. Voir aussi le Message du Conseil fédéral du 22 novembre 2017 relatif à la modification de la loi sur le droit d'auteur, à l'approbation de deux traités de l'Organisation Mondiale de la Propriété Intellectuelle et à leur mise en œuvre, FF 2018 559 ss, p. 594–595.

mesures, en particulier d'ordre contractuel. Dans une perspective pratique, nous proposons ici une check-list des questions principales à régler en amont de l'ouverture de données à des fins de recherche biomédicale :

[87] Phase 1 – Évaluation des obligations liées aux données concernées

- La mise à disposition des données est-elle imposée ou encouragée par des règles spécifiques, par exemple des règles imposées par un bailleur de fonds (p. ex. FNS) ou des règles internes de l'institution qui détient les données ? Dans l'affirmative, quelles sont les exigences imposées ?
- Les données concernées contiennent-elles des données personnelles ? Dans l'affirmative, il faut identifier l'ensemble des dispositions légales applicables (cf. *supra* 3.1) et déterminer si et à quelles conditions les données concernées peuvent être mises à disposition de tiers. Si la LRH est applicable et qu'un partage de données personnelles (codées ou non) est envisagé, on veillera entre autres à déterminer si les règles de réutilisation instituées par les art. 32 à 34 LRH sont réunies (notamment la validité des consentements) et à vérifier que le projet de recherche en vue duquel les données seront réutilisées a été validé par une commission d'éthique (art. 45 LRH). Dans tous les cas, des données personnelles ne peuvent pas être librement diffusées, elles doivent rester dans un environnement contrôlé.
- Les données concernées sont-elles couvertes par d'autres régimes juridiques ? Par exemple, les données font-elles l'objet d'une obligation spécifique de confidentialité, de licences conférées par des tiers, d'autres types d'obligations contractuelles qui empêcheraient une divulgation, un secret d'affaires ou de fabrication ?

[88] La phase 1 doit permettre d'acquérir une vision globale des droits et obligations liée au partage des données concernées et, ainsi, d'établir les conditions minimales auxquelles ces dernières peuvent être mises à disposition de tiers. La phase 1 doit aussi permettre de déterminer si des actions spécifiques doivent être entreprises.

[89] Phase 2 – Stratégie d'ouverture

- Déterminer le type de licence à appliquer pour les réutilisations futures des données. Le degré d'ouverture des données dépendra du contenu des données, en particulier si elles intègrent ou non des données personnelles.
- L'ouverture implique le partage de données personnelles :
 - Si le bailleur de fonds impose une obligation générale de rendre les données ou résultats accessibles au public, soumettre une demande de dispense si celle-ci est nécessaire en vue de se conformer aux obligations légales applicables.
 - En cas de partage et de réutilisation des données par des tiers, il convient de mettre en place des instruments contractuels spécifiques de type *Data Transfer Agreement* qui permettent de garder le contrôle sur les données, notamment en spécifiant le champ d'utilisation autorisé par le destinataire des données, mais aussi d'assurer le respect d'autres obligations spécifiques imposées par les règles applicables en matière de protection des données personnelles. Ces dernières priment sur les considérations liées à l'ouverture des données.

- L'ouverture n'implique pas le partage de données personnelles, notamment si les données de recherche ont été complètement (et correctement) anonymisées :
 - S'assurer qu'il n'existe aucune obligation contractuelle de garder le secret, par exemple à cause du dépôt d'une demande de brevet. Si des règles particulières sont imposées en matière d'Open Data par le bailleur de fonds ou par l'institution de recherche concernée, veiller à s'y conformer.
 - Le donneur de licence peut mettre à disposition les données de recherche en utilisant un contrat de licence propriétaire ou des licences publiques types Creative Commons. L'utilisation de la licence CC0 offre une capacité de réutilisation maximale.
 - Si cela est adapté, inclure des conditions contractuelles spécifiques concernant les points suivants : (a) description des éléments qui ne tombent pas dans le champ de la licence concédée et des règles applicables en cas de réutilisation de matériels appartenant à des tiers, (b) règle de citation, (c) droit et juridiction applicables (d) éventuelles règles de responsabilité applicables en cas de réutilisation du contenu sous licence, (e) conditions permettant de résilier l'accès et la réutilisation des données et (f) adresse de contact (p. ex. adresse courriel).
 - Évaluer l'opportunité de publier et diffuser les données de recherche dans un répertoire public (voir notamment le site www.re3data.org/ qui liste les différents répertoires publics selon les disciplines visées).

6. Conclusion

[90] L'Open Data, par la promotion de l'accès aux données et au savoir, est devenu un mouvement incontournable du secteur de la recherche. Ce phénomène s'observe aussi bien sur le plan international qu'à l'échelon helvétique, où il est directement pris en compte dans les stratégies les plus récentes du Conseil fédéral sur la réutilisation des données médicales (Rapport Humbel) ainsi que dans les conditions de financement du FNS par exemple. Néanmoins, les exigences posées en matière d'ouverture des données ne sont pas absolues et les textes en la matière réservent habituellement l'existence de motifs justificatifs contraires à une telle ouverture. La présente contribution avait pour objectif de dresser les contours des limitations que le droit suisse peut imposer à l'ouverture des données à des fins de recherche biomédicale.

[91] Au terme de nos réflexions, nous parvenons à la conclusion générale que l'Open Data et l'ouverture des données de recherche dans le contexte de la recherche biomédicale ne sont pas incompatibles avec le droit actuel, en particulier le droit de la protection des données, de la recherche sur l'être humain ou de la propriété intellectuelle. Selon le contexte et le type de données en jeu, la mise en œuvre de l'Open Data doit néanmoins faire l'objet d'aménagements importants.

[92] Sans grande surprise, la principale contrainte à prendre en compte se rattache à la protection d'éventuelles données personnelles. Si de telles données sont en jeu, le responsable du traitement doit procéder à une évaluation parfois complexe des règles juridiques applicables et déterminer à quelles conditions il peut rendre les données concernées accessibles. En guise de règle générale, il est nécessaire de s'assurer que les données personnelles restent constamment sous contrôle et, de ce fait, les réutilisations « en chaîne » de telles données doivent être évitées. Les technologies ou modèles respectueux de la vie privée peuvent contribuer à s'affranchir des contraintes imposées

pour le partage de données personnelles, mais une analyse circonstanciée doit être menée au cas par cas.

[93] Lorsque la question des données personnelles est évacuée et qu'une ouverture des données est envisageable, c'est-à-dire qu'aucun autre obstacle ne s'y oppose (p. ex. droits de tiers sur les données), il faut établir une stratégie réfléchie d'ouverture des données. Une telle stratégie passe généralement par l'établissement ou le choix de licences appropriées au regard des buts poursuivis.

[94] Récemment, le Conseil fédéral a donné au Département fédéral de l'Intérieur le mandat de développer une stratégie concrète de réutilisation de données de santé (Rapport Humbel) et un projet de loi-cadre fédérale sur la réutilisation de données (potentiellement de santé) est également en cours de réflexion.¹⁵⁹ Comme examiné, une telle démarche fait aujourd'hui l'objet de projets plus concrets en Europe, avec le projet d'Espace européen des données de santé. À l'échelon suisse, il s'agit de réflexions importantes qui pourraient engendrer des propositions de modifications législatives, à l'image du souhait affiché du Conseil fédéral de moderniser l'institution du consentement général. Les solutions proposées devront assurer un délicat équilibre entre la promotion d'une science ouverte et la protection des participants, dans le respect des garanties constitutionnelles et éthiques.

FRÉDÉRIC ERARD, Dr. iur., avocat, CIPP/E, Head, Legal and Technology Transfer Office au SIB Institut Suisse de Bioinformatique.

MATHILDE HEUSGHEM, LLM, avocate, Legal Officer au SIB Institut Suisse de Bioinformatique.

CLÉMENT PARISATO, MLaw, Senior Legal Officer au SIB Institut Suisse de Bioinformatique.

Le SIB Institut Suisse de Bioinformatique est chargé, en collaboration avec l'Académie suisse des sciences médicales (ASSM), de la mise en œuvre de l'initiative SPHN, qui comprend également l'établissement du réseau BioMedIT. Les salaires de Frédéric Erard et Mathilde Heusghem sont financés en partie par les fonds SPHN. Le projet SwissBiodata ecosystem (SBDe) est co-dirigé par l'Université de Berne et le SIB Institut Suisse de Bioinformatique.

Les analyses et réflexions menées dans cette contribution relèvent de l'opinion personnelle de leurs auteurs et n'engagent en rien celle de leur employeur, à savoir le SIB Institut Suisse de Bioinformatique, ou celle d'autres entités telles que SPHN.

Les auteurs remercient M. Marc Fillietaz pour sa relecture du texte et ses précieux commentaires.

¹⁵⁹ Cf. nbp 19.