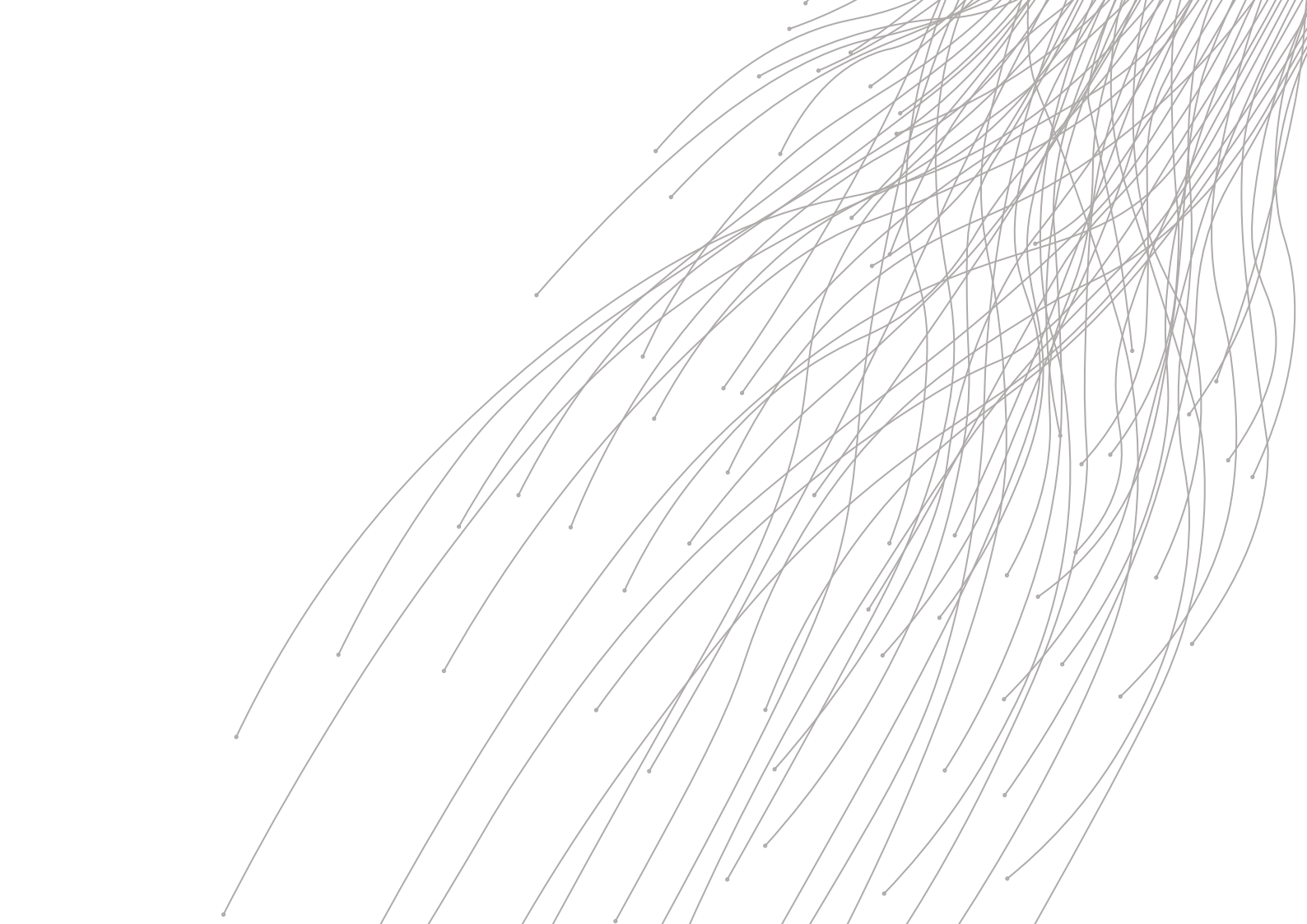


The logo consists of a stylized white icon of a person with arms raised, followed by the text "BY-COVID" in a bold, white, sans-serif font.

BY-COVID

Kit pédagogique.





D'où provient ce kit pédagogique?	3
Comment utiliser ce package éducatif?	4
Sujets	6
Introduction	9
INFORMATION DE BASE.....	9
Cas	12
CAS #1 / BANC DE POISSONS	13
CAS #2 / ACTION POLITIQUE	15
CAS #3 / EAUX USÉES.....	17
CAS #4 / LES EFFETS SECONDAIRES POSITIFS	19
CAS #5 / METADONNÉES.....	21
CAS #6 / FUTURE PANDÉMIES	25
CAS #7 / PARAPLUIE	28
CAS #8 / FAIR DATA	30
CAS #9 / SYSTÈMES FÉDÉRÉS	33
CAS #10 / COLLABORER POUR INNOVER.....	36

/ D'où provient ce kit pédagogique?



Ce kit pédagogique fait partie du projet BeYond-COVID (BY-COVID). Le projet est une collaboration scientifique entre 53 instituts partenaires dans 19 pays européens, qui se sont réunis pour améliorer la manière dont les données scientifiques et de santé sont gérées et partagées lors de pandémies. Les concepts du projet constituent un cadre à partir duquel nous pouvons soulever des questions éthiques plus larges sur les infrastructures de données de santé en général.

Le monde a généré d'énormes quantités de données en réponse à la pandémie de COVID-19, et continue d'en générer davantage. L'objectif du projet est de rendre les données COVID-19 accessibles à quiconque peut les utiliser, des scientifiques en laboratoire au personnel médical dans les hôpitaux et les responsables gouvernementaux. Ces données proviennent de sources diverses, et les identifier, les relier et les intégrer pour une analyse efficace reste un défi.

Le projet fournira également un cadre pour rendre les données d'autres maladies infectieuses ouvertes et accessibles à tous. Des données plus faciles d'accès, agrégées et analysées permettront aux scientifiques de réagir plus rapidement aux nouvelles souches de SARS-CoV-2 ou à de nouveaux virus. Cela aidera également les décideurs à évaluer l'impact de la maladie et à prendre les mesures les plus appropriées pour protéger les gens contre de nouvelles maladies infectieuses.

Cependant, les "données" ne sont pas simplement des chiffres abstraits dans une base de données abstraite. Elles concernent souvent les informations sur les citoyens individuels. La collecte de ces données personnelles permet des avancées scientifiques. De nombreuses questions éthiques se posent dans ce contexte. Où, comment et par qui les données doivent-elles être collectées, stockées ou partagées ? Combien d'informations voulons-nous recevoir sur ce qui arrive à nos données ? Et comment les différentes parties prenantes et les citoyens sont-ils impliqués dans cette utilisation des données ? Nous pensons que les salles de classe offrent un contexte idéal pour informer les jeunes citoyens et délibérer avec eux sur des questions sociétales importantes.

Comment utiliser ce package éducatif?

En tant qu'enseignant, vous trouverez ici des ressources pour aborder le sujet dans vos cours. Il s'agit d'une boîte à outils, ce qui signifie que vous êtes libre de sélectionner et d'adapter ce package éducatif en fonction de vos besoins. Le package se compose de suggestions d'accroches, de sujets proposés et d'informations de base introductives, le tout construit autour d'un certain nombre de visuels et de cas qui illustrent différents éléments de la manière dont les données sont utilisées, ou devraient être utilisées, pour prévenir ou gérer les pandémies. Le package est construit autour de questions plutôt que de réponses. Nous espérons que ce matériel inspirera les élèves à réfléchir à l'éthique des systèmes de données de santé et les encouragera à trouver leurs propres réponses. Nous les invitons à partager leur point de vue sur chaque cas individuel [ici](#).

Nous suggérons d'introduire les élèves au matériel en mettant l'accent sur ces deux points :

- **Cela s'est déjà produit lors de la pandémie :**

Pendant la pandémie, lorsque quelqu'un recevait un résultat positif au test PCR, il pourrait avoir eu besoin de s'isoler, il aurait pu recevoir un traitement et obtenir une preuve d'infection passée. Bien après le test, ces données continuent de servir à d'autres fins pour aider à comprendre davantage le virus. BY-COVID vise à faciliter et à améliorer l'utilisation de ces données pour prévenir et gérer efficacement les pandémies actuelles et futures. Cependant, il est important de savoir comment ces données sont gérées. Explorer les valeurs sous-tendant ce cadre nous aide à comprendre nos propres points de vue et ceux des autres sur le partage des données.

- **Cela a un impact sur vous et sur notre société :**

Depuis le début de la pandémie, [plus de 765 000 000 de cas ont été confirmés dans le monde](#). Pour comprendre chacun de ces cas, il faut savoir si la personne a été vaccinée (par exemple : quand, à quelle fréquence, quel vaccin), hospitalisée (par exemple : durée, type de traitement, en soins intensifs ou non), si elle est rentrée de l'étranger, a été en contact avec d'autres personnes, et quelle variante du virus elle a. Un chercheur pourrait également vouloir savoir si certains variants du virus affectent différemment les femmes enceintes de moins de 25 ans par rapport aux femmes enceintes de plus de 35 ans. Tout cela nécessite la liaison entre de nombreuses bases de données différentes.

Il est très important d'établir des liens entre différentes sources de données sur le COVID-19 et d'autres maladies infectieuses pour pouvoir informer les politiques, développer des traitements et prévenir la propagation des virus autant que possible. Les liens entre ces

différents types de données doivent être établis au niveau individuel. Les données vous concernent, et l'information, c'est le pouvoir. Qui peut avoir accès à ces informations ? Que peuvent-ils en faire ? Quels types d'informations aimeriez-vous voir protégé ?

Un ensemble d'outils, prêt à être adapté :

Ce package éducatif comprend différents cas, regroupés en différents sujets. Ces cas illustratifs peuvent servir de point de départ pour une discussion en classe et permettre aux élèves d'explorer leurs propres valeurs et principes concernant les écosystèmes de données de santé. Les sujets peuvent servir de guide aux élèves pour formuler des questions éthiques découlant du cas. Nous fournissons également un ensemble de questions liées aux sujets qui peuvent servir de point de départ pour une discussion plus approfondie sur un sujet spécifique. Les cours peuvent se concentrer sur un ou plusieurs sujets comme point de départ, mais les discussions toucheront probablement différents sujets, car ils sont tous interconnectés.

Cette courte vidéo peut être utilisée comme ressource supplémentaire pour illustrer l'importance de la réutilisation des données de santé. [Qu'est-ce que les données de santé ?](#) (2:33 min)

/ Sujets



Les chiffres influençant les politiques

Quelles considérations éthiques se posent lors de l'utilisation de données scientifiques et de santé pour informer les décisions politiques ? Quels sont les biais potentiels et les valeurs intégrées dans le processus de collecte et d'analyse des données ?

Comment aborder les implications éthiques de la sélection et de l'exclusion des données dans la formation des décisions politiques ? Comment naviguer dans les conséquences potentielles et les impacts des politiques basées sur les données sur les différents individus et les communautés ? Existe-t-il des conflits éthiques potentiels entre la quête de la connaissance scientifique et le bien-être ou les droits des individus ou des communautés ? Dans quelle mesure les données peuvent-elles justifier une réorganisation de la société ? Comment assurer la transparence et la responsabilité dans l'utilisation des données pour façonner les structures et les institutions de la société ? Quelles mesures peuvent être mises en œuvre pour atténuer les biais potentiels, la discrimination ou les inégalités susceptibles de découler de la prise de décision basée sur les données ? Comment incorporer des perspectives diverses et garantir l'inclusivité dans le processus d'utilisation de données scientifiques pour améliorer la santé publique ?



Cas liés: 2, 3, 6, 10

Confidentialité

Comment pouvons-nous équilibrer les avantages de la collecte et du traitement des données avec le risque de violations potentielles de la vie privée ? Quels sont les risques et les conséquences potentiels de la collecte de données ? Quelles mesures peuvent être mises en œuvre pour atténuer les risques de violations de la vie privée liés à la collecte et au traitement des données ? Quelles considérations doivent être prises en compte lors de la manipulation de données contenant des informations personnellement identifiables ? La manière dont les données sont utilisées influence-t-elle le niveau de protection requis ?



Cas liés: 1, 3, 4, 8

Données agrégées

Comment les pratiques d'agrégation et d'anonymisation des données peuvent-elles garantir la protection de la vie privée individuelle tout en permettant une analyse significative ? Comment le processus d'agrégation des données individuelles en un ensemble de données plus vaste influence-t-il les caractéristiques et les propriétés inhérentes des données d'origine ? Comment l'agrégation des données individuelles influence-t-elle l'interprétation et l'analyse de l'ensemble des données ?



Cas liés: 3, 9

Sécurité des données

Comment pouvons-nous nous assurer que les données combinées sont stockées en toute sécurité et protégées contre un accès non autorisé ou une utilisation abusive ? À quelle organisation accordons-nous la confiance pour stocker nos données ? Devons-nous stocker nos données en un seul endroit ou doivent-elles être séparées ? Quelles mesures techniques sont nécessaires pour garantir la sécurité des données ? Qui peut accéder à mes données ? Comment les données de santé peuvent-elles être réutilisées en toute sécurité ? Lorsque vous pensez aux conditions et aux garanties, à quoi pensez-vous spécifiquement ? Outils informatiques ? Lois et réglementations ? Experts et comités ? Autre chose ? Quels pourraient être les avantages et les inconvénients de ceux-ci ? Comment pourraient-ils être complémentaires ou conflictuels ? Dans quel aspect du processus de réutilisation des données de santé les garanties sont-elles les plus importantes ? Doivent-elles contrôler qui accède et réutilise ces données ? Devraient-elles plutôt se concentrer sur les objectifs et les moyens spécifiques de réutilisation ? Quels pourraient être les points vulnérables dans le processus de réutilisation des données de santé qui exigent une protection spécifique ? Comment devrions-nous équilibrer la protection de la vie privée et la promotion du progrès scientifique ? Comment la fédération des données peut-elle garantir la protection de la vie privée des individus et le maintien de la confidentialité des informations sensibles ?



Cas liés: 1, 2, 4, 8

Confiance

Quelles considérations éthiques se posent lors de la collecte et de l'analyse de données personnelles sans le consentement explicite des individus ? Comment pouvons-nous établir la confiance parmi les citoyens en ce qui concerne la gestion des données de santé grâce à des interventions ou des mesures efficaces ? Quelles sont les responsabilités éthiques des organisations et des institutions pour protéger les informations personnelles des individus contenues dans leurs données ?



Cas liés: 2, 4, 7

Qualité des données

Comment pouvons-nous nous assurer que les données combinées sont précises et fiables ? Quelles mesures peuvent être prises pour garantir que des données de haute qualité soient collectées et utilisées pour la recherche et la prise de décision ? Comment pouvons-nous garantir que la représentation des données est diversifiée et inclusive ?



Cas liés: 2, 5

Principes FAIR

Comment pouvons-nous équilibrer la promotion de la recherche, de l'accessibilité, de la réutilisation et de l'interopérabilité des données (principes FAIR) tout en respectant les droits à la vie privée des individus et en maintenant la sécurité des données ? Quelles sont les considérations éthiques lors de la prise de décisions concernant le niveau d'ouverture et d'accessibilité des données, en particulier lorsqu'il s'agit d'informations sensibles ou confidentielles, afin de garantir des pratiques responsables de partage des données ? Comment pouvons-nous garantir que les avantages de la combinaison des données de santé soient accessibles à tous et ne conduisent pas à davantage d'inégalités en matière de soins de santé ?



Cas liés: 4, 5, 8

IA dans les données de santé

Quelles considérations éthiques se posent lors de l'utilisation de l'intelligence artificielle (IA) pour la réutilisation des données de santé, en particulier en ce qui concerne la confidentialité, le consentement et la sécurité des individus dont les données sont utilisées ? Comment l'utilisation responsable et éthique de l'IA peut-elle être assurée, en tenant compte de problèmes tels que la représentation équitable et les risques de biais ou discriminations involontaires dans le processus de prise de décision ?



Cas liés: 4

Introduction

Information de base

La pandémie de COVID-19 a dominé nos vies pendant plus de deux ans. La lutte contre le coronavirus nous a contraints à modifier notre manière de travailler, de nous amuser, d'éduquer nos enfants et d'interagir les uns avec les autres.

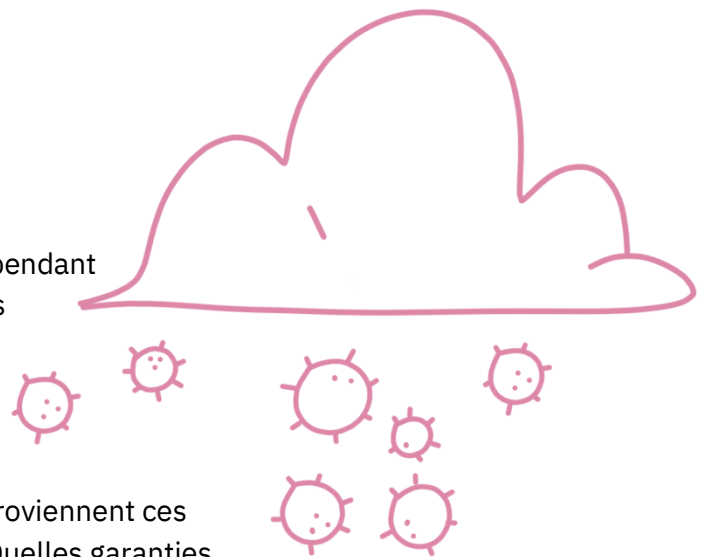
La justification de ces changements se

trouve souvent dans "les chiffres". Mais d'où proviennent ces données ? À quoi servent-elles exactement ? Quelles garanties

sont en place ? Et comment les différentes parties prenantes et les citoyens sont-ils impliqués dans cette réutilisation des données de santé ? Toutes ces questions sont liées à la toute première question, où et comment les données de santé sont-elles stockées ?

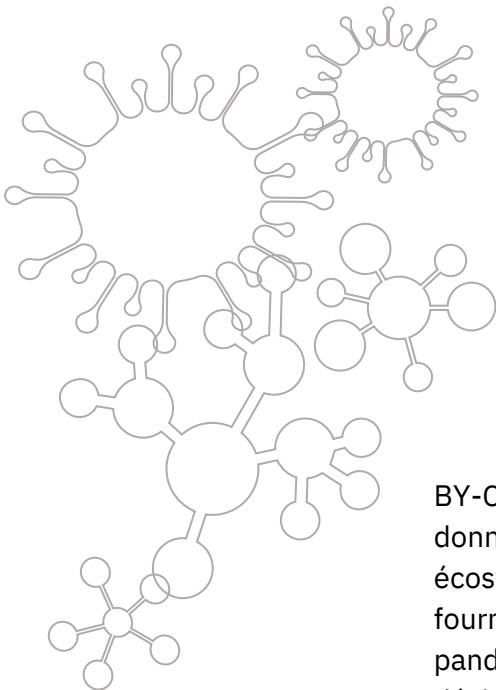
Actuellement, les données de santé sont stockées dans de nombreux silos séparés. Imaginez que vous ayez beaucoup d'informations importantes sur votre santé, comme votre historique médical, les résultats de tests et les médicaments. Maintenant, imaginez que ces informations soient stockées dans des endroits séparés qui ne communiquent pas entre eux. C'est comme avoir différents silos où chaque information est isolée des autres. C'est un problème car lorsque vos informations de santé sont dispersées dans ces silos, il devient difficile pour vos médecins et autres prestataires de soins de santé d'avoir une vue d'ensemble sur votre santé. Ils pourraient ne pas avoir toutes les informations nécessaires pour prendre les meilleures décisions concernant vos soins. Le problème des silos rend également difficile la collaboration entre différents médecins ou services et le partage d'informations. Un autre problème est que, lorsque les données de santé sont divisées en silos, il n'est pas facile d'utiliser ces informations pour la recherche ou pour trouver des schémas qui peuvent aider à améliorer les soins de santé pour tous. C'est comme avoir des pièces de puzzle dispersées à différents endroits, où vous ne pouvez pas assembler l'intégralité du puzzle pour voir le tableau d'ensemble.

Les données sont ce qui nous permet de décrire la réalité. Un manque de données conduit à une compréhension limitée du monde qui nous entoure. Nous pouvons illustrer l'importance d'un système pour lier efficacement les données en imaginant un extraterrestre nommé Xerox. Il visite la Terre pour l'explorer en tant que marché potentiel pour sa société de chaussures. Ne connaissant pas l'anatomie humaine, Xerox commence son enquête en entrant dans une salle de classe où il trouve un seul élève assis dans un coin. En se basant sur les pieds de l'élève, Xerox fait certaines hypothèses initiales sur les caractéristiques des



pieds humains. Il compte 10 orteils et mesure la longueur et la largeur. Cependant, sa compréhension est limitée car il n'a observé qu'une seule personne.

Peu de temps après, un groupe d'élèves supplémentaires entre dans la salle, fournissant à Xerox un échantillon plus large de pieds humains. En comparant les nouveaux individus, il réalise que ses hypothèses initiales sur la longueur et la largeur des pieds étaient incorrectes. Il découvre que les pieds humains se présentent sous différentes tailles et formes, et que la longueur de leurs pieds peut être plus courte ou plus longue que les 30 cm initiaux qu'il avait supposés. De plus, il remarque que certaines personnes ont des pieds plus étroits que d'autres, ce qui nécessite des ajustements pour s'adapter aux différentes largeurs de pied.



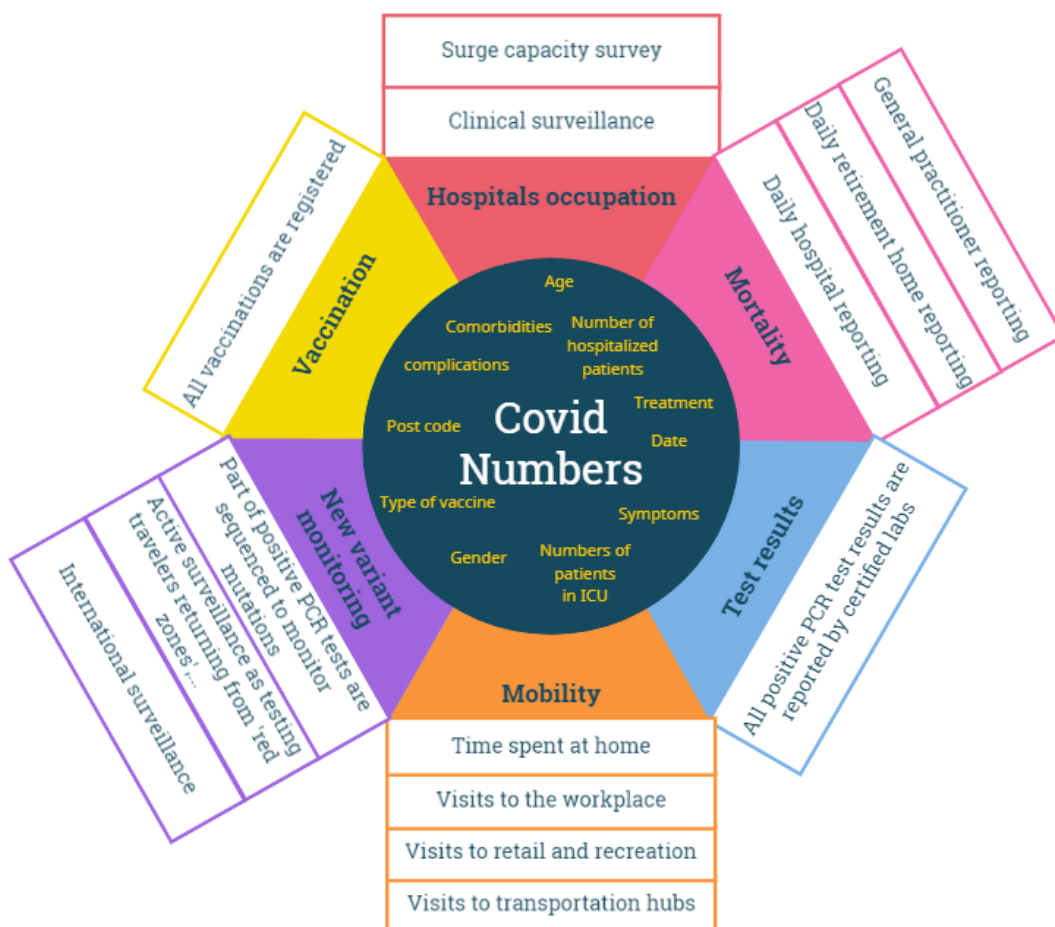
Si Xerox souhaite créer une gamme de chaussures adaptées à toute l'humanité, il est clair qu'il devra collecter des informations auprès du plus grand nombre de personnes possible. Plus il y en a, mieux c'est. Chaque paire de pieds supplémentaire qu'il peut inclure dans ses recherches ajoute un peu d'information et changera d'une certaine manière la collection finale de chaussures. Pour répondre à toute question de recherche concernant les données de santé, le même principe s'applique. Les données sont un pouvoir, d'autant plus lorsqu'elles sont liées et combinées.

BY-COVID soutient une structure fédérée pour connecter les données de ces silos séparés. Grâce à la mise en œuvre d'un écosystème de données fédéré, le projet BY-COVID vise à fournir une vue complète et consolidée de l'impact de la pandémie, permettant une meilleure analyse, suivi et prise de décision pour combattre efficacement la COVID-19 et les futures possibles pandémies.

La fédération désigne une approche de gestion des données dans laquelle les données restent à leur emplacement d'origine, mais peuvent être accessibles et analysées par des parties autorisées. Différentes organisations ou institutions collaborent et conservent le contrôle sur leur propre silo de données tout en permettant aux autres d'y accéder et de l'utiliser en toute sécurité. De cette manière, les silos contenant divers types d'informations médicales peuvent être liés sans avoir à stocker toutes les données en un seul endroit. Prenons l'exemple d'un chercheur en santé publique qui collecte des données pour comprendre comment le virus se propage dans différents pays. Au lieu d'exiger que toutes ces données soient envoyées à une base de données centrale, chaque instance individuelle représentant le pays fournit un accès aux données particulières pour répondre à la question de recherche. Chaque instance permet d'accéder aux pièces nécessaires d'un puzzle qu'elle possède et le chercheur peut assembler une image globale. Chaque question de recherche peut donner lieu à un puzzle différent.

Ils peuvent également décider s'ils veulent agréger les données qu'ils fournissent. Dans le contexte du projet BY-COVID, les données de santé seront souvent agrégées. Les données agrégées se réfèrent à la combinaison de données collectées auprès de diverses sources liées à la pandémie de COVID-19. Ces données incluent des informations sur les taux d'infection, les hospitalisations, les vaccinations et d'autres indicateurs pertinents. Lorsque les données sont agrégées, les lignes de données atomiques, généralement recueillies auprès de plusieurs sources, sont remplacées par des totaux ou des statistiques récapitulatives.

Cependant, rien de tout cela ne serait possible sans nos données personnelles. Dans l'illustration fournie, vous trouverez un aperçu de toutes les données différentes collectées, liées et utilisées pendant la crise du COVID-19. Cela commence, par exemple, par un individu se faisant tester pour le COVID-19. Dans un premier temps, ces informations seront utiles pour la personne en question et son médecin. Ce n'est que dans un second temps, lorsque les données font partie d'une structure plus vaste, qu'elles deviennent un outil puissant. L'infrastructure BY-COVID jette les bases pour rendre possible la liaison entre les statistiques des personnes infectées dans une région et le nombre de visites dans les commerces et les lieux de loisirs. Par conséquent, ce sont en partie mes données qui guideront les décideurs politiques pour fermer temporairement les écoles de musique ou assouplir les directives de distanciation sociale dans les clubs sportifs.



/ Cas

CAS #1 / BANC DE POISSONS	13
CAS #2 / ACTION POLITIQUE	15
CAS #3 / EAUX USÉES	17
CAS #4 / LES EFFETS SECONDAIRES POSITIFS	19
CAS #5 / METADONNÉES.....	21
CAS #6 / FUTURE PANDÉMIES.....	25
CAS #7 / PARAPLUIE	28
CAS #8 / FAIR DATA	30
CAS #9 / SYSTÈMES FÉDÉRÉS	33
CAS #10 / COLLABORER POUR INNOVER	36

Cas #1 / Banc de poissons

United, we stand



Divided, we fall



Together against
COVID-19



L'union fait la force !

Il peut s'avérer difficile pour un petit poisson de survivre seul, nageant dans un océan de prédateurs. C'est pour cela que de nombreuses espèces de poissons ont trouvé une technique redoutable face aux dangers des mers : le banc de poissons. Se déplacer en restant ainsi groupé présente de nombreux avantages : les poissons deviennent plus forts en s'unissant, ils sont donc beaucoup moins vulnérables face aux prédateurs, qui auront beaucoup plus de facilités à s'attaquer à une proie seule. Ils créent également du courant, ce qui leur permet d'économiser de l'énergie et d'être plus rapides. Le déplacement du groupe est pensé collectivement, chaque poisson s'adapte aux mouvements du voisin et leur espérance de vie augmente.



OK, intéressant, mais quel rapport avec le COVID-19?

Cette anecdote sert à faire un parallèle avec nos données de santé. Les données d'un seul individu sont peu utiles à la collectivité, ce n'est que lorsqu'elles sont combinées à un grand nombre de données qu'elles peuvent faire la différence. Par exemple, si nous voulons connaître l'impact du Covid-19 sur des personnes atteintes d'une maladie spécifique, comme le lupus, il est important d'analyser les données du plus grand nombre de patients possible. Plus les ensembles de données sont combinés, plus la recherche sera riche, augmentant ainsi les chances de trouver des traitements efficaces. Il en est de même pour les institutions, les pays et les chercheurs : la collaboration rend les données plus fortes et la recherche plus rapide. Combiner les données de différents domaines de recherche permet également d'approfondir la compréhension de la santé et des maladies.

Comment faire un « banc de données » ?

Pour rendre cela possible, il est nécessaire d'avoir un cadre pour collecter, gérer et permettre l'accès aux données au-delà des frontières internationales, en respectant les règles de protection des données de chaque pays. Le projet BY-COVID rend possible cette mise en commun des données, en réunissant les chercheurs et les institutions de nombreux pays européens, afin que les données de chaque pays puissent être combinées pour générer de la connaissance pour tout le monde. La combinaison de données accroît le pouvoir de la recherche et améliore les soins au patient. BY-COVID construit un système facilitant le stockage, le partage, la liaison et l'utilisation des données. Par exemple, pour s'assurer que l'ensemble des données combinées puisse être analysé, BY-COVID développe des métadonnées, qui étiquettent chaque élément sur les données afin qu'il puisse être trouvé et compris.

Si vous voulez en savoir plus...

- Pour en savoir plus sur les bancs de poissons : [Fish swimming in schools save energy regardless of their spatial position - PMC](#)
- Pour en savoir plus sur la valeur et l'importance du partage des données : [Coronavirus accelerates drive to share health data across borders | Research and Innovation](#)
- Pour en savoir plus sur [les objectifs de BY-COVID](#)
- Pourquoi les données sont nécessaires pour des maladies spécifiques : [Impact of COVID-19 on people with cystic fibrosis - PMC](#)
- Cet article montre l'importance de relier les ensembles de données dans le cas du cancer : [Linking data to improve health outcomes | The Medical Journal of Australia](#)

Cas #2 / Action politique



**With help,
it is easier
to make decisions**



Better data
Better knowledge

BY-COVID

De bonnes informations mènent à des résultats savoureux

Si vous voulez préparer un bon gâteau d'anniversaire, vous avez besoin d'ingrédients frais et d'une bonne recette pour les combiner correctement. Sans l'un ou l'autre, le gâteau est voué à l'échec.

Les décisions politiques concernant les pandémies sont à peu près les mêmes : il faut des informations fiables et un moyen de rassembler ces informations. Par exemple, si l'efficacité du vaccin ou le degré de sévérité du prochain variant sont inconnus, il s'avère impossible de formuler de bonnes politiques. Les données sont l'ingrédient-clé des politiques fondées sur des preuves, mais elles sont inutiles sans le cadre d'une bonne recette. Les données scientifiques doivent être collectées, stockées correctement et reliées avec d'autres données, le tout dans un environnement sécurisé et accessible.

Si les décideurs politiques sont les cuisiniers des mesures en santé publique pour gérer les pandémies, le projet BY-COVID fournit les meilleures recettes possibles pour combiner les données afin d'éclairer leurs décisions. L'objectif est de soutenir les décideurs politiques et leur fournir les outils nécessaires pour prendre des décisions ancrées dans la réalité et proportionnelles aux risques auxquels nous sommes confrontés..



Où aimeriez-vous manger?

Mangeriez-vous dans un restaurant utilisant des ingrédients périmés, des recettes farfelues jamais testées auparavant et ne respectant aucune mesure d'hygiène ? De la même manière, les politiques de santé publique ont besoin de bonnes données, de méthodes bien testées pour combiner les données et de normes pour garantir la protection des données. Avec les bonnes matières premières, les bons équipements et les bonnes ressources, la saveur des aliments ne dépendra donc que des compétences du cuisinier.

En qui faites-vous confiance pour élaborer des stratégies en période de pandémie ?

Sur quoi leurs décisions doivent-elles se fonder ?

Quelles sont les données les plus importantes à collecter ?

Si vous voulez en savoir plus...

- Sur la collaboration entre les scientifiques et les décideurs politiques : [Science advice in times of COVID-19 - OECD](#)
- Sur le lien entre la démocratie, l'action gouvernementale durant la pandémie et la confiance des citoyens: [Trust made the difference for democracies in COVID-19 - The Lancet](#)
- Sur les objectifs de BY-COVID : https://by-covid.org/pdf/BY-COVID_factsheet_dark.pdf
- Sur l'impact de l'incertitude scientifique sur la confiance publique dans la science et le soutien aux politiques fondées sur la science : [Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic](#)

Cas #3 / Eaux usées



**Did you know that
flushing
the toilet
can help
research?**

BY-COVID

Le saviez-vous ? Tirer la chasse peut aider la recherche !

Saviez-vous que la chasse d'eau de vos toilettes pouvait fournir des données pour la recherche ? Les eaux usées des toilettes ont été utilisées comme un outil précieux pour détecter les points chauds de la pandémie de COVID-19. En étudiant les eaux usées, nous pouvons anticiper les taux d'admission à l'hôpital liés à la maladie dans des zones spécifiques ou même prédire les futures épidémies d'autres maladies.

Nos excréments peuvent être considérés comme un enregistrement dynamique des événements qui se produisent dans notre corps. Par exemple, ils contiennent des fragments du virus COVID-19 lorsque nous sommes infectés par celui-ci. Cela signifie que nous pouvons mesurer la quantité de virus et les variants qui circulent dans la population ou dans des zones particulières, avec un système de gestion qui maximise les avantages, gère efficacement les risques et utilise ces informations pour concevoir une réponse de santé publique ciblée.



Comment les données sont-elles collectées et stockées?

BY-COVID travaille avec l'initiative italienne SARI, dans le cadre duquel les données sur les eaux usées sont collectées environ une ou deux fois par semaine dans chaque ville d'Italie. Les données sont ensuite stockées dans des bases de données et affichées sur des tableaux de bord dans lesquels les données sur la surveillance des eaux usées peuvent être saisies en temps réel pour avoir une image actualisée des vagues pandémiques. Les chercheurs du projet BY-COVID veillent ensuite à ce que les données extraites des échantillons d'eaux usées puissent être réutilisées dans d'autres études, par exemple pour établir des comparaisons entre différents pays, répondre à des questions de recherche sur la propagation du virus ou éclairer les décisions de santé publique.

Quels sont les avantages?

L'analyse des eaux usées est un outil efficace pour la détection et la surveillance des épidémies de maladies infectieuses. En connaissant les zones où l'agent pathogène est le plus concentré, les réponses politiques peuvent être mieux adaptées. Les ressources fournies par BY-COVID (par exemple les bases de données et les logiciels d'analyse de données) sont donc utiles pour lutter contre la pandémie actuelle mais aussi pour repérer les futures.

Quels sont les risques liés à l'utilisation de ces données?

Il est théoriquement possible de trouver des données personnelles dans les eaux usées si l'on analyse l'ADN humain présent dans les excréments. Cependant, dans ce domaine de recherche, seules des données sur les virus sont collectées, pas sur les humains.

L'infrastructure BY-COVID est mise en place pour contrer ces risques d'utilisation abusive des données : toutes les données utilisées dans l'analyse des eaux usées sont agrégées, ce qui signifie qu'il serait impossible d'identifier précisément toute information sur les individus.

Si vous voulez en savoir plus...

- Voici un exemple de l'utilisation de l'analyse des eaux usées dans un campus universitaire : <https://www.youtube.com/watch?v=ysZsx5wS2YM>
- Cet article présente un résumé des résultats de la première étude sur les variants du COVID-19 dans les eaux usées italiennes : [Flash survey on SARS-CoV-2 variants in urban wastewater in Italy 1st Report \(Investigation period: 04 – 12 July 2021\)](#)
- Pour en savoir plus sur l'utilisation des eaux usées : [Water and wastewater digital surveillance for monitoring and early detection of the COVID-19 hotspot: industry 4.0 | SpringerLink](#)
- Regardez une courte vidéo sur les données agrégées : <https://www.youtube.com/watch?v=PvohkMNV9Lo>

Cas #4 / Les effets secondaires positifs



Une pandémie peut-elle avoir des effets secondaires positifs?

Vous est-il déjà arrivé de chercher quelque chose de précis dans une pièce en désordre et de finalement tomber sur quelque chose d'inattendu ? C'est un phénomène courant dans le monde de la recherche, où certaines études menées dans un but précis se révèlent utiles pour autre chose. Si les données scientifiques sont suffisamment nombreuses, accessibles et bien exploitées, non seulement la recherche s'en trouve améliorée, mais cela ouvre également la possibilité de la réutilisation de ces données dans de nouvelles et différentes recherches.



Quelques exemples de (re)découvertes

La pandémie de Covid-19 a mobilisé des chercheurs, des institutions et des données du monde entier et a permis des avancées dans de nombreux domaines. Par exemple, des méthodes existantes, comme le test PCR ou les vaccins à ARN messager, ont été validées à grande échelle pendant la pandémie et sont désormais plus faciles à appliquer pour d'autres maladies (les [tests PCR pour la variole du singe](#) et les [vaccins à ARN messager pour le traitement de certains cancers](#)).

L'intelligence artificielle (IA) et l'Open Science (à savoir un mouvement qui vise à rendre la science plus largement accessible) existaient avant la pandémie, mais elles ont été déployées d'une manière encore jamais vue, entraînant de nouveaux développements. Les scientifiques ont économisé des mois d'expérimentation grâce aux prédictions générées par l'IA sur la [structure du virus](#), qui ont également aidé au séquençage du génome et à la prédiction de l'évolution de la pandémie.

En termes d'Open Science, la crise a accéléré l'[abandon de l'accès restreint et conditionnel à la connaissance](#). Des informations essentielles sur le virus ont été mises à la disposition de tous, permettant une meilleure compréhension de la maladie et la fabrication rapide de vaccins, de diagnostics et de traitements.

BY-COVID comme vecteur de progrès scientifique

Ces progrès ne sont possibles qu'avec une bonne infrastructure de données qui permet aux chercheurs et aux institutions de travailler ensemble pour fournir des données facilement accessibles afin de soutenir la découverte et la prise de décision.

L'une des façons dont le projet BY-COVID promeut l'Open Science est de contribuer à l'European Open Science Cloud (EOSC), qui "[vise à développer un Web de données et de services FAIR pour la science en Europe, sur lequel une large gamme de services à valeur ajoutée peut être construite](#)".

En réunissant un grand nombre de pays, de chercheurs et d'institutions et en rendant les données accessibles, BY-COVID crée une infrastructure qui permettra de tirer des enseignements positifs de la pandémie.

Si vous voulez en savoir plus...

Les vaccins à ARN messager: [New cancer treatments may be on the horizon—thanks to mRNA vaccines / L'ARN messager nous sauvera-t-il du cancer ? | 42, la réponse à presque tout | ARTE](#) (French)

L'Intelligence Artificielle : [AI and control of Covid-19 coronavirus](#)

<https://www.swissinfo.ch/eng/artificial-intelligence-helps-bring-about-record-fast-vaccines/46256752>

L'Open Science: [Open science saves lives: lessons from the COVID-19 pandemic | BMC Medical Research Methodology](#)

Comment le Covid-19 a eu un impact sur la recherche [How will COVID-19 reshape science, technology and innovation?](#) / [The impact of COVID-19 on research - PMC.](#)

Cas #5 / Metadonnées



Standards de qualité, nombreux accès

Lorsque nous voyageons à l'étranger, nous pouvons avoir des difficultés à charger notre téléphone. Sans le bon adaptateur, il n'est pas possible de connecter nos appareils et nous serons rapidement à court de batterie. Avec un adaptateur universel, tout le monde peut voyager partout dans le monde. Il en va de même pour les métadonnées : quel que soit le type de données ou l'endroit où elles sont stockées, les standards de métadonnées permettent à quiconque de trouver, d'accéder et d'utiliser des données pour ses études de recherche



C'est quoi une métadonnée ? Pourquoi est-il important d'avoir des standards ?

Les métadonnées sont des "données sur les données". Cependant, toutes les métadonnées ne sont pas utiles et les standards (ou normes) doivent être approuvés par la communauté des chercheurs et, idéalement, suivre un ensemble de lignes directrices appelées les [principes FAIR](#). Les normes relatives aux métadonnées contribuent à la création de catalogues de métadonnées pour rendre les données trouvables et accessibles, et peuvent également servir d'adaptateurs pour rendre les données interopérables et réutilisables.

Pour donner un exemple simple, lorsque vous recherchez un film sur une plateforme de streaming, vous trouverez des informations telles que l'année de sortie, le genre du film, le réalisateur, la durée. Ces informations décrivent le film et permettent d'en trouver plus facilement un qui vous intéresse et de décider si vous voulez le regarder.

Il en va de même pour les données de recherche : si les informations sur les données sont précises et détaillées, il sera beaucoup plus facile pour un chercheur de découvrir que les données existent et de les utiliser pour son analyse. Si un chercheur trouve deux ensembles de données potentiellement utiles, mais que l'un se réfère aux lieux par leur nom et l'autre par leur code postal, une norme de métadonnées est nécessaire pour combiner les données.

Nous utilisons tous des normes de métadonnées au quotidien : lorsque nous utilisons notre GPS pour nous rendre à la Tour Eiffel, nous devrions arriver au même endroit, quelle que soit l'application que nous utilisons. De bonnes normes de métadonnées permettent aux développeurs d'applications d'orienter les utilisateurs vers l'endroit où ils souhaitent se rendre, qu'ils aient saisi "Tour Eiffel", "Eiffel Tower" ou "75007".

Les données doivent être accompagnées de métadonnées dès la conception

Le meilleur livre de la bibliothèque ne sera pas trouvé s'il n'est pas indexé de manière standardisée. Les différents éléments de la bibliothèque ont différentes façons d'être trouvés. Par exemple, un magazine a un numéro, un livre peut être une édition spéciale, une bande dessinée peut faire partie d'une série plus importante. Il en va de même pour les données : les données sociales, les données médicales et les données biologiques nécessitent toutes des métadonnées différentes pour les décrire de manière appropriée.

Les normes relatives aux métadonnées sont nécessaires pour trouver, relier et utiliser ces données pour la recherche dans différents domaines. Les standards relatifs aux données et aux métadonnées contextuelles et expérimentales qui leur sont associées sont également connus sous le nom de standards de données, standards de métadonnées ou standards de contenu, et peuvent être classés en [quatre sous-types de standards](#) : lignes directrices ou listes de contrôle pour les rapports, modèles/formats ou syntaxe, artefacts terminologiques et schémas d'identification.

Il est essentiel que la saisie et la standardisation des métadonnées soient intégrées aux plans dès le début de tout projet de recherche, avant la collecte des données. Cela garantit que les données trouveront leur place dans un écosystème global d'informations. De bonnes métadonnées améliorent également la qualité et la fiabilité des données et la confiance dans les résultats de la recherche.

Standards de qualité, nombreux accès

Le projet BY-COVID comprend de nombreuses sources de données (par exemple, des bases de données, des dépôts et des bases de connaissances) provenant de différentes disciplines de recherche, notamment les biosciences, la recherche clinique et épidémiologique, ainsi que les sciences sociales et humaines. Ces sources de données sont décrites dans une collection [FAIRsharing](#) (en cours), ainsi que les standards de données et de métadonnées utilisés par chaque source de données. Un modèle commun de métadonnées a été développé pour représenter les métadonnées de chaque source et pour les rendre accessibles en un seul endroit : le [portail de données Covid-19](#).

Le développement d'un modèle de métadonnées commun est un défi majeur, car le projet implique un grand nombre de chercheurs de différents domaines scientifiques, et les sources de données de chaque partenaire utilisent des standards de métadonnées différents. L'approche consiste à cartographier les principales interrelations entre les métadonnées d'une manière logique et pratique à mettre en œuvre. Cela ouvre ensuite des possibilités passionnantes pour en savoir plus sur la façon dont les maladies infectieuses affectent les gens et pour éclairer l'élaboration de politiques fondées sur des preuves.

Si vous voulez en savoir plus...

- [FAIRsharing Educational](#) : découvrez les normes relatives aux données et aux métadonnées, comment le [registre FAIRsharing](#) vous aide si vous êtes un consommateur ou un producteur de données, les normes relatives aux métadonnées, les bases de données et les politiques en matière de données.
- Plus de détails sur les métadonnées : [Introduction à la gestion des métadonnées](#)
- Découvrez comment l'indexation est utilisée pour relier les données dans le projet BY-COVID : [Release of indexing system to link COVID-19 data across research disciplines](#)
- En savoir plus sur l'importance de disposer de standards en matière de métadonnées (en général) : [5 Minute Metadata - What is a standard?](#)
- En savoir plus sur l'importance des normes de métadonnées face à la pandémie COVID-19 : [COVID-19 pandemic reveals the peril of ignoring metadata standards | Scientific Data](#)
- Trouvez des outils et des directives pour vous aider à accéder aux données sur les maladies infectieuses, à les analyser et à les partager, et à réagir rapidement aux épidémies : [Infectious Diseases Toolkit](#)

- [BY-COVID - D3.1 - Metadata standards. Documentation on metadata standards for inclusion of resources in data portal | Zenodo](#)
- BY-COVID D2.1 : Harmonisation initiale des données et des métadonnées au niveau du domaine pour permettre des réponses rapides à COVID-19 : <https://doi.org/10.5281/zenodo.7017728>
- Apprenez-en davantage sur les recettes qui vous aident à rendre les données FAIR dans [FAIR Cookbook](#), une ressource en ligne de recettes pratiques pour les "faiseurs de FAIR" dans les sciences de la vie. [The FAIR Cookbook - pre-print: "The essential resource for and by FAIR doers"](#), vous fournit de plus amples informations sur sa création et son contenu.
- Partage de données FAIR, éthique et coordonné pour la réponse au COVID-19 : une revue des plateformes et registres de partage de données du COVID-19 : [FAIR, ethical, and coordinated data sharing for COVID-19 response: a review of COVID-19 data sharing platforms and registries | Zenodo](#)
- Packaging des artefacts de recherche avec RO-Crate : [Packaging research artefacts with RO-Crate - IOS Press](#)
- Modèle de preuve distribué allégé pour les environnements complexes du monde réel : [Lightweight Distributed Provenance Model for Complex Real-world Environments | Scientific Data](#)
- COVID-19 : Une exploration des obstacles systémiques consécutifs au partage des données relatives aux agents pathogènes pendant une pandémie : [\[2205.12098\] COVID-19: An exploration of consecutive systemic barriers to pathogen-related data sharing during a pandemic](#)

Cas #6 / Futures pandémies



Be innovative
in the face of
adversity...

...makes it
easier later



BY-COVID

Créer des routes pour le futur

La pandémie de Covid-19 a bouleversé nos vies, mais elle a aussi transformé le monde de la recherche. Des scientifiques du monde entier se sont réunis pour comprendre la maladie et mettre au point des traitements et des vaccins aussi rapidement que possible. Mais qu'advient-il de ces avancées une fois la crise passée ?



La science comme exploration

Les explorateurs partent en expédition pour découvrir des zones inconnues de notre monde. Ils peuvent être à la recherche d'une nouvelle espèce de champignon aux propriétés médicinales, ou tenter d'observer un nouveau comportement animal. La plupart d'entre eux trouveront leur chemin à l'aide d'outils testés et approuvés, comme le GPS, mais aussi grâce à leur capacité à innover face à des situations inconnues. Une fois que les pionniers ont trouvé un chemin dans la nature, leur expérience permet aux futurs voyageurs de suivre plus facilement leurs traces. Le risque de se perdre, de s'enliser dans des situations dangereuses ou de perdre du temps à chercher, est beaucoup plus faible lorsque le terrain a déjà été cartographié. Plus le chemin est utilisé, mieux il sera balisé et plus il sera facile d'accès. Il serait dommage de laisser ces chemins à l'abandon alors que nous savons qu'ils peuvent mener à des endroits intéressants.

Explorer les maladies infectieuses

Cette pandémie n'est ni la première, ni la dernière à laquelle nous serons confrontés. Cependant, nous serons mieux à même d'anticiper les futures pandémies grâce aux connaissances et à l'expérience acquises tout au long de la pandémie de la Covid-19. Tout comme l'exploration, les recherches et les expériences antérieures sont très utiles comme base pour la suite des travaux.

Maintenant que la pandémie est mieux maîtrisée, il est temps de revenir sur les [leçons apprises](#) et sur la manière dont elles peuvent être appliquées à l'avenir. Nous pouvons utiliser ces enseignements pour accélérer les efforts de réponse à la pandémie, par exemple en [optimisant la production de vaccins](#). Les recherches effectuées, les résultats trouvés, les procédures développées et les méthodes utilisées peuvent servir à prévenir ou à gérer les futures pandémies.

L'une des armes les plus efficaces contre le virus Covid-19 a été le [partage rapide de données provenant du monde entier](#). Ces données ont montré comment la maladie affectait le corps, quelles mesures étaient les plus efficaces pour stopper la transmission, comment le virus évoluait et quels étaient les meilleurs vaccins pour prévenir les infections.

Au-delà du COVID

Tels des explorateurs, les scientifiques ont ouvert la voie en collectant, reliant, stockant, utilisant et partageant les données pour aider à combattre la pandémie. Le Covid-19 a provoqué une crise mondiale, et les pionniers scientifiques ont dû travailler rapidement et de manière innovante, par exemple pour s'assurer que tous les cadres éthiques et juridiques étaient en place et trouver de nouvelles façons de connecter les données. Ce travail sous pression a montré l'avantage de mettre à la disposition de tous des données provenant de différents pays et domaines de recherche.

Certaines infrastructures de données ont été créées spécifiquement pour résoudre cette crise, mais se sont avérées si utiles qu'il est désormais important de veiller à leur maintien. Le projet BY-COVID suit les traces des scientifiques spécialisés dans les pandémies pour

établir des cartes, des guides, des ponts et des tremplins afin d'atténuer et de gérer les futures pandémies.

Vous souvenez-vous des défis que vous avez eu à relever ? Comme étudier pour vos examens ou vous entraîner pour un marathon ? Avez-vous tiré de ces réalisations des enseignements qui se sont avérés utiles pour d'autres défis ? À votre avis, quels ont été les principaux enseignements tirés des données pendant la pandémie ?

Si vous voulez en savoir plus...

- Un exemple d'une bonne infrastructure de données sur la santé (Danemark) : [Health data and registers - Sundhedsdatastyrelsen](#)
- Examen des déterminants moléculaires et cellulaires de la pneumonie critique à COVID-19 : [Human genetic and immunological determinants of critical COVID-19 pneumonia | Nature](#)
- Retour sur les leçons tirées de la pandémie : [Ten lessons from the first two years of COVID-19 | McKinsey](#)
- Un exemple de leçon tirée de la compréhension de Covid-19 (savoir comment utiliser toutes les données impliquées) : [Why the WHO took two years to say COVID is airborne](#)
- Ce que le COVID-19 nous a appris : [What COVID-19 has taught us: "Healthcare can no longer exist without technology" | Africa Renewal](#)
- Le cadre BY-COVID réutilisé pour la variole du singe : [Repurposing COVID-19 pipelines for the monkeypox virus](#)
- La santé numérique pendant la pandémie et comment l'utiliser après : [Data capture and sharing in the COVID-19 pandemic: a cause for concern - The Lancet Digital Health](#)
- Ce que covid nous a montré sur l'importance du partage des données et de la science ouverte : https://www.researchgate.net/profile/Kenneth-Zheng-2/publication/343892070_Data_sharing_during_COVID-19_pandemic_what_to_take_away/links/5f87cf9b458515b7cf81eda9/Data-sharing-during-COVID-19-pandemic-what-to-take-away.pdf

Cas #7 / Parapluie



**Our data
protect us.**

BY-COVID

Faisons-nous confiance aux données ?

Tout au long du développement et du déploiement d'un vaccin, un ensemble de données, allant des expériences en laboratoire aux essais cliniques, en passant par le suivi post-approbation, contribue à rendre les vaccins aussi efficaces et sûrs que possible. L'accès aux données est important pour les experts qui prennent des décisions en matière de santé publique, mais aussi pour les citoyens, car leur perception et leur attitude influencent fortement l'adoption de la vaccination.

Si l'on n'a pas confiance dans la sécurité et l'efficacité des vaccins Covid-19, il est difficile d'obtenir des taux de vaccination (volontaire) élevés dans une population donnée.



Les données influencent l'opinion publique

Les sondages nous apprennent que dans les pays de l'UE, l'hésitation à se faire vacciner est souvent fondée sur le manque d'accès à des données fiables. [Près d'une personne interrogée sur deux reconnaît qu'il est difficile de trouver des informations fiables sur le COVID-19 et les vaccins.](#)

La pandémie de Covid-19 a apporté avec elle des vagues de données (par exemple le nombre de citoyens testés, infectés et malades) et d'informations sur le fonctionnement des tests de diagnostic et des vaccins Covid-19. Pour de nombreuses personnes, non seulement les citoyens mais aussi les experts de la santé, il a été difficile d'interpréter toutes ces données. Mais il n'est pas possible de comprendre une pandémie sans données. Ce dont nous avons besoin, ce n'est pas de moins de données, mais de moyens et d'outils permettant de rendre facilement accessibles des données fiables sur le Covid-19, voire sur d'autres maladies infectieuses.

BY-COVID crée la confiance

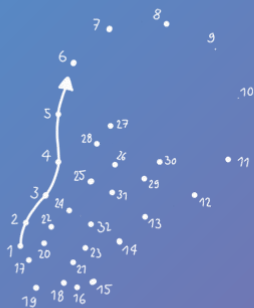
Le projet BY-COVID a créé un [portail de données sur le COVID-19](#) qui relie des données provenant de nombreux domaines de recherche différents, tels que la génomique et les sciences sociales. Il s'agit d'une ressource pour les scientifiques et les professionnels de la santé, mais elle peut contribuer à rassurer les citoyens sur l'existence de données fiables et leur accessibilité à tous.

Les données ne changent pas toujours facilement l'opinion publique. Selon le sondage EUROBAROMETER, la confiance dans la vaccination COVID-19 reflète la confiance générale dans les institutions telles que les gouvernements nationaux. Cette confiance reflète non seulement les politiques et actions actuelles des institutions, mais aussi leurs antécédents. La construction de la confiance est donc un processus lent, mais dans lequel des données fiables et accessibles jouent un rôle important.

Si vous voulez en savoir plus...

- Données Eurobarometer sur les attitudes des Européens à l'égard de la vaccination :
- [Eurobarometer](#) (données collectées en mai 2021)
- [Eurobarometer](#) (données collectées en février 2022)
- Catalogue de données CESSDA où l'on peut trouver quelques données intéressantes liées au COVID-19 et situées dans des archives européennes de données en sciences sociales : [CESSDA Data Catalogue](#)
- Une autre façon d'accéder aux données européennes sur les questions sociales liées au COVID-19 est de consulter la page de signalisation du CESSDA : [CESSDA and COVID-19](#)

Cas #8 / FAIR data



Data are
priceless,

when you give them
meaning.



 **BY-COVID**

Les données n'ont pas de prix, quand on leur donne un sens

Des données sont générées dans le monde entier pour aider à répondre aux pandémies et aux épidémies de maladies infectieuses. Cependant, chaque donnée isolée n'a qu'une valeur limitée si elle n'est pas reliée à des données connexes et mise à disposition de manière pertinente.

Une structure et une approche commune sont nécessaires pour relier différents types de données et pour garantir une gestion responsable et efficace des données. Les données sont ainsi plus faciles à partager et à réutiliser, et donc encore plus utiles et précieuses. Mais comment les données sont-elles gérées et rendues accessibles aux autres ? Il existe un ensemble de lignes directrices qui indiquent la voie à suivre : les principes FAIR.



Les principes FAIR, qu'est-ce que c'est?

Les principes FAIR soutiennent l'Open Science (Science ouverte), un mouvement mondial visant à rendre la science accessible à tous les chercheurs et à tous les niveaux de la société. Chacune des quatre lettres représente un élément distinct:

Findable. (Trouvable). Les scientifiques doivent trouver les données dont ils ont besoin. Plus les informations sur un ensemble de données (métadonnées, qui doivent également être standardisées) sont disponibles, plus il est facile de les trouver.

Accessible. L'accès aux données doit être aussi facile et ouvert que possible. Les données sont parfois mises à l'abri du regard du public. Il peut y avoir des raisons valables à cela, comme la protection des données personnelles d'une personne. Cependant, pour développer des médicaments ou des tests de diagnostic, les scientifiques doivent pouvoir accéder aux données de manière sûre et sécurisée.

Interoperable. Les scientifiques ont souvent besoin de combiner des données pour leur travail. Pour pouvoir le faire, elles doivent être disponibles dans des formats bien définis et compatibles, afin que les données puissent être liées, comparées et analysées efficacement.

Reusable. (Réutilisable). F, A et I sont mis en place pour rendre les données réutilisables. Les données, initialement collectées dans un but précis, peuvent améliorer la santé ou faire progresser les connaissances scientifiques dans d'autres domaines. Par exemple, des identifiants personnels sécurisés peuvent garantir l'utilité d'un ensemble de données à long terme.

Les données très sensibles et privées, par exemple le statut vaccinal, peuvent également être FAIR en plaçant des contrôles sur l'accessibilité. Cela permet de s'assurer que l'on peut accéder aux parties essentielles des données sans divulguer d'informations personnelles.

Les principes FAIR sont des lignes directrices et non une norme. Ils décrivent un continuum de réutilisation croissante, dont la pleine réalisation peut être entravée par la réglementation, le budget et les technologies disponibles. La collaboration entre toutes les parties prenantes (cliniciens, chercheurs, responsables de la santé publique ou décideurs politiques) est importante pour appliquer les principes autant que possible.

Principes FAIR pour donner du sens aux données et minimiser les erreurs

L'application des principes FAIR permet non seulement d'améliorer le partage et l'accessibilité des données, mais aussi de minimiser le risque de biais et d'erreurs dans la recherche. En effet, l'application de ce cadre permet d'obtenir des ensembles de données plus cohérents, pertinents pour la recherche et aussi fiables que possible. Avoir accès à

suffisamment de données FAIR permet d'évaluer ouvertement la sécurité et l'efficacité des interventions sur la base des preuves.

La contribution de BY-COVID aux principes FAIR

L'application des principes FAIR apporte des avantages au projet BY-COVID et, par extension, à la communauté des chercheurs. En appliquant ces principes, notamment en rendant les résultats de la recherche accessibles, le projet BY-COVID permet aux futurs chercheurs de réutiliser les données sur les maladies infectieuses afin de faire de nouvelles découvertes. L'accès aux données et aux résultats démontre également la transparence de la recherche, car il permet aux gens de vérifier comment celle-ci a été menée, quelles données ont été utilisées et comment.

Le projet BY-COVID contribue aux principes FAIR de plusieurs manières. Par exemple, en effectuant un [mappage inter-domaines des métadonnées](#) (un processus de mappage des métadonnées de différentes spécialités afin qu'elles puissent être utilisées ensemble), ou en améliorant la découverte, l'intégration et la citation des données, en utilisant des outils basés sur le [portail européen de données Covid-19](#) et [FAIRsharing](#). Ces outils permettent aux chercheurs de trouver facilement les données dont ils ont besoin, de les intégrer dans leurs travaux et de citer correctement leurs sources.

Pour mieux répondre aux futures pandémies, il est essentiel que des projets tels que BY-COVID appliquent les principes FAIR afin de passer d'un partage de données fragmenté et concurrentiel à un réseau de données interconnectées et accessibles

Si vous voulez en savoir plus...

- Sur les données FAIR : [FAIR data – Ghent University](#) / [How to make your data FAIR through Data Standardization – HERAX](#) / [Les principes FAIR de gestion et d'intendance des données de recherche](#) (en français)
- Partage de données FAIR, éthique et coordonné pour la réponse au COVID-19 : une revue des plateformes et registres de partage de données du COVID-19 : [FAIR, ethical, and coordinated data sharing for COVID-19 response: a review of COVID-19 data sharing platforms and registries | Zenodo](#)
- Les principes directeurs FAIR pour la gestion et l'intendance des données scientifiques : [The FAIR Guiding Principles for scientific data management and stewardship](#)
- L'European Open Science Cloud (EOSC) est un environnement permettant d'héberger et de traiter les données de recherche afin de soutenir la science européenne : [The European Open Science Cloud \(EOSC\) is an environment for hosting and processing research data to support EU science.](#)
- Si vous voulez en savoir plus, notamment sur l'interopérabilité : [Q&A: Interoperability and COVID-19, Part 1 - Watson Health Perspectives](#) and [Interoperability in Healthcare | IBM](#).
- Harmoniser les données cliniques pour faciliter la recherche en santé à grande échelle : [Harmonising clinical data to facilitate large-scale health research | News | CORDIS | European Commission](#)
- Comment COVID-19 a fondamentalement changé la recherche clinique en santé mondiale : [How COVID-19 has fundamentally changed clinical research in global health](#)

Cas #9 / Systèmes fédérés



Data visiting
is the new
data sharing.

BY-COVID

La visite des données est le nouveau partage des données

Pendant la pandémie de COVID-19, les réunions en ligne ont pris une place importante dans la vie professionnelle de beaucoup d'entre nous. Nous sommes restés chez nous en toute sécurité et, comme lors de n'importe quelle autre réunion, nous avons échangé des idées pour faire avancer nos projets. La différence : nous ne sommes pas sortis de chez nous. Cependant, les idées que nous avons partagées ont quitté notre maison pour rejoindre celles d'autres personnes. Nos pensées et nos opinions, mises ensemble, ont enrichi la discussion et permis de progresser.



Quel est le rapport entre les données et les réunions en ligne?

Les données de recherche sont stockées de différentes manières dans différents endroits et il n'est pas toujours facile d'y accéder. Avec un système fédéré, les données restent sur place, un peu comme les personnes qui travaillent à domicile, mais les chercheurs et les autres personnes qui bénéficient des données peuvent y avoir accès. Un système fédéré est composé de plusieurs fournisseurs de données, tels que des hôpitaux ou des universités, qui collaborent ensemble tout en conservant le contrôle de la gestion de leurs données.

Cette “fédération” peut être comprise dans un sens similaire aux réunions en ligne : chaque personne représente les données (qui ne bougent pas) tandis que les connaissances dérivées des données sont représentées par les idées de chaque personne (qui sont libres d'être partagées). L'alternative au système fédéré est connue sous le nom de système centralisé : les participants à la réunion (données) se rendent tous au même endroit pour se rencontrer en personne et partager leurs idées.

Quand le partage des données devient la visite des données

Un système fédéré facilite la réutilisation des données, conformément aux principes FAIR de trouvabilité (Findable), d'accessibilité (Accessible), d'interopérabilité (Interoperable) et de réutilisation (Reusable). De nombreux obstacles politiques, éthiques, administratifs, réglementaires et législatifs s'opposent à ce que les données quittent leur lieu d'origine ; le système fédéré permet aux données de rester là où elles se trouvent tout en étant réutilisées pour faire avancer la recherche.

BY-COVID en tant que réunion en ligne

Le projet BY-COVID utilise des systèmes centralisés et fédérés, en fonction du type de données. Dans les systèmes centralisés, les données sont partagées dans des bases de données communes qui résident dans un endroit différent de celui où se trouvent les données. Les systèmes centralisés sont utilisés pour les données qui ne sont pas sensibles, par exemple les données qui ne concernent pas les personnes.

Toutefois, pour répondre à des questions politiques intéressantes, l'accès à des données plus sensibles est également utile. Ces données, qui se trouvent à différents endroits, peuvent être consultées et analysées à l'aide de systèmes fédérés, dans le respect des règles de gouvernance nationales et des restrictions réglementaires. Les résultats peuvent être largement diffusés dans des articles scientifiques sans que les données ne soient déplacées, à l'instar de la diffusion des résultats d'une réunion en ligne alors que les participants restent chez eux.

Si vous voulez en savoir plus...

Sur les bases de données fédérées : [What is federation \(English\)](#) ([French](#))

Sur l'importance de la "visite" des données : [The way forward is not data sharing but data visiting](#)

Sur l'interopérabilité dans le développement d'une infrastructure de recherche fédérée : [Coping with interoperability in the development of a federated research infrastructure](#)

Sur le plan de gestion des données de BY-COVID : [BY-COVID data management plan](#)

Cas #10 / Collaborer pour innover



Lorsque nous faisons face à une urgence de santé publique, comme la pandémie de COVID-19, il est essentiel que les experts, qu'ils soient issus des universités ou des industries, s'unissent pour le bien commun, même si leurs intérêts divergent. Pour préserver la santé publique, il est essentiel qu'un maximum de données soient ouvertes et accessibles pour que chacun puisse apporter son expertise et bénéficier des connaissances générées. Maintenir la science ouverte favorise la transparence et la collaboration, ce qui stimule l'innovation et accélère le rythme de la recherche.

Compte tenu du volume et de la complexité des données produites au-delà des frontières nationales, elles doivent être gérées avec soin pour rester significatives et utiles. La gestion des données est un rôle endossé par les infrastructures de données financées par des fonds publics et travaillant pour le bénéfice de tous. Ces infrastructures développent des outils et mettent à disposition des données qui peuvent être largement utilisées non seulement par les universitaires, mais aussi par le secteur privé et les décideurs politiques. Cela permet une collaboration productive entre les secteurs public et privé, et permet au secteur privé de transformer les découvertes scientifiques en innovations bénéfiques pour la société.

Lors de la pandémie de COVID-19, grâce à la grande quantité de données disponibles et ouvertement partagées, les entreprises ont été en mesure de créer des solutions innovantes, telles que de nouveaux vaccins, diagnostics et traitements, à une vitesse jamais vue auparavant. Cela n'a été possible que grâce à la richesse des données disponibles dans les infrastructures financées par des fonds publics.

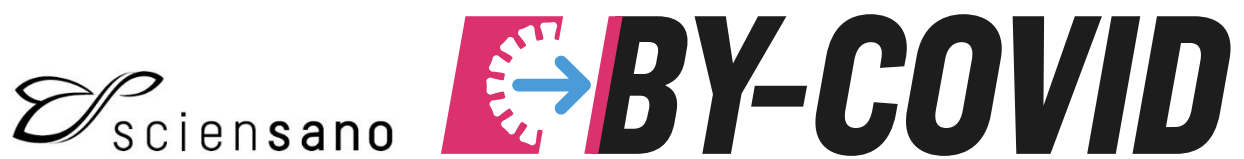


BY-COVID, un modèle d'innovation

Le projet BY-COVID vise à renforcer la capacité des infrastructures publiques à répondre aux besoins des futures pandémies. La rapidité de l'innovation au cours de la pandémie de COVID-19 a été rendue possible par des infrastructures de données antérieures à la pandémie. Bien qu'il s'agisse de ressources existantes, les infrastructures ont dû s'adapter rapidement pour faire face au déluge de données produites. Le projet BY-COVID s'appuie sur les outils et les ressources créés pendant la pandémie pour apporter un soutien face aux futures situations d'urgence et permettre la poursuite de la recherche sur les maladies infectieuses. Ce travail garantira que les ressources créées seront optimisées pour alimenter l'innovation future et que, quoi que nous ayons à affronter, nous y ferons face ensemble.

Si vous voulez en savoir plus...

- Sur le rôle de la collaboration dans le développement de ventilateurs alternatifs pendant la pandémie COVID-19 : [Accelerated innovation in crises: The role of collaboration in the development of alternative ventilators during the COVID-19 pandemic - ScienceDirect](#)
- Sur ce que la pandémie nous a appris sur la collaboration : [What COVID-19 taught us about collaboration – 7 lessons from the frontline](#)
- Sur le cas de la Malaisie pour la gestion de la pandémie : [Collaboration and innovation to overcome COVID-19 challenges – Dr Ir Siti Hamisah Tapsir – INGSA](#)
- Normaliser les partenariats et la collaboration dans la lutte contre le COVID-19 : [Partnership and collaboration: The new normal in the fight against COVID-19](#)



BY-COVID has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101046203.